TRANSCRIPTIONAL REGULATION OF THE *VIBRIO HARVEYI* *LUX* OPERON

BY

Elana Esther Swartzman
Department of Biochemistry
McGill University
Montreal, Canada

October, 1992

A thesis submitted to the Faculty of Graduate Studies
and Research in partial fulfillment of the requirements
of the degree of Doctor of Philosophy.

© Elana E. Swartzman, 1992

Abstract

The molecular mechanisms involved in transcriptional regulation of the *Vibrio harveyi lux* operon were examined. The 5' and 3' ends were mapped and two new *lux* genes, *luxG* and *luxH*, were discovered upstream of the termination site. A transcription terminator that was able to function bidirectionally was located downstream of only one new *lux* gene, *luxG*, in the *lux* operon of *Vibrio fischeri*.

Differences in transcriptional start sites were detected when the *V. harveyi lux* mRNA was mapped *in vivo* and *in vitro*. Studies using recombinant *Escherichia coli* containing the *V. harveyi lux* operon demonstrated that the promoter in *E. coli* mapped to the same site as found *in vitro*. The *in vivo* start site mapped to 26 base pairs (bp) and the *in vitro* start site mapped to 123 bp upstream of the initiation codon of the *luxC* gene. The expression of *luxR* in recombinant *E. coli* shifted the start site from position -123 to position -26, demonstrating that LuxR is required for transcriptional activation of the -26 *lux* promoter.

A DNA-binding protein, detected exclusively in cell extracts of *V. harveyi*, bound specifically to a fragment of DNA upstream of the *lux* promoter in mobility shift assays. Footprint analysis showed that the protein bound to two A + T-rich regions of DNA upstream form the *V. harveyi lux* promoter. The protein was purified to homogeneity and was determined to be LuxR. The *V. harveyi lux* operon is therefore regulated by LuxR, a novel poly(dA-dT) DNA-binding protein that is required for accurate transcription of the *V. harveyi lux* promoter.

Résumé

Les mécanismes moléculaires impliqués dans la régulation transcriptionelle de l'opéron *lux* de la bactérie *Vibrio harveyi* furent étudié. Les extémités 5' et 3' furent cartographiées de plus deux nouveaux gènes, *luxG* et *luxH*, furent découverts en amont du site de terminaison. Un site de terminaison bidirectionel fut découvert dans l'opéron de *Vibrio fischeri*, qui contenait seulment un nouveau gène, *luxG*.

Des différences dans les sites d'initiation furent détectées durant la cartographie de l'opéron *lux* de *V. harveyi* *in vivo* et *in vitro*. Des études avec *Escherichia coli* contenant l'opéron *lux* de *V. harveyi* ont démontré que le promoteur dans *E. coli* se trouve à des sites équivalents *in vivo* et *in vitro*. Le site d'initiation *in vivo* a été cartographié à 26 paires de bases et le site d'initiation *in vitro* à 123 paires de bases en amont du premier codon ATG du gène *luxC*. Le facteur manquant dans *E. coli* recombinant pour l'activation de transcription du promoteur situé à la position -26 fut déterminé comme étant le produit du gène *luxR*.

Une protéine se liant à l'ADN, détectée exclusivement dans des extraits cellulaires de *V. harveyi*, était liée spéciquement au fragment d'ADN en amont du promoteur de *lux* dans des essais de changement de mobilité. Les analyses d'empreinte démontrent que la protéine se lie à deux régions rich en A + T, situées en amont du promoteur *lux* de *V. harveyi*. Cette protéine fut purifiée jusqu'à homogénéité et identifiée comme étant LuxR. L'opéron *lux* de *V. harveyi* est par conséquent régulé par LuxR, une nouvelle protéine se liant à l'ADN qui est requise pour une transcription précise du promoteur *lux* de *V. harveyi*.

TABLE OF CONTENTS

FORWARD

The present thesis includes the text of original papers submitted for publication. In compliance with the guidelines for thesis prepations provided by the Faculty of Graduate Studies and Research, the text of section 2 is cited below:

"Manuscript and Authorship

The candidate has the option, subject to approval of their department, of including as part of the thesis the text, or duplicated published text, of an original paper or papers.

- Manuscript-style thesis must still conform to all other requirements explained in the Guidelines Concerning Thesis Preparation.
- Additional material (procedural and design data as well as descriptions of equipment) must be provided in sufficient detail (eg. in appendices) to allow clear and precise judgement to be made of the importance and originality of the research reported.
- The thesis should be more than a mere collection of manuscripts published or to be published. It must include a general abstract, a full introduction and literature review and a final overall conclusion. Connecting texts which provide logical bridges between different manuscripts are usually desirable in the interest of cohesion.
- It is acceptable to include, as chapters, authentic copies of papers already published, provided these are duplicated clearly and bound as an integral part of the thesis. In such instances, connecting texts are mandatory and supplementary explanatory material is always necessary.
- Photographs or other materials which do not duplicate well must be included in their original form.
- While the inclusion of manuscripts co-authored by the candidate and others is acceptable, the candidate is required to make an explicit statement in the thesis of who contributed to such work and to what extent, and supervisors must attest to the accuracy of the claims at the Ph.D. Oral Defence. Since the task of the Examiners is made more difficult in these cases, it is in the candidate's interest to make the responsibilities of authors perfectly clear."

ACKNOWLEDGEMENTS

I am indebted to my supervisor, Dr. Edward Meighen, for his unfailing support and encouragement throughout my graduate training. He has given me excellent advice and guidance on my research project as well as on oral and written presentations. His trust in my judgement and capabilities has given me the confidence and courage to pursue a career in independent research.

I am grateful to past and present members of my lab for their friendship, advice and help: Stefano Ferri, Wei Qun Sun, Chan Yong Lee, Zhi Lee, Rose Szittner, Robert Soly, Jack Cao, Mike Boylan, and Xiaoning Zhao.

A special thanks to Carol Miyamoto, who has taught me everything there is to know about molecular biology (and then some!). She has been a mentor and a great friend throughout the years. I will miss her helpful insights and our daily chats and coffee break, but she will forever be in my thoughts and heart.

I would like to thank past and present members and friends of the department for making grad life bearable, even enjoyable: Helena, Dean, Carole, Jerry, Rob, Isabel, Mireille, Maria, Jinhi, Julie, Joelle, Laura Lee, Mitch, Serge, and many, many others. Thanks to Dr. Branton's lab for the beer club (it sure beats the coffee club we had in the same lab many years ago!) and thanks to everyone who helped endure the coldest but cosiest ski trip to Mt. St. Anne.

I would also like to thank Dr. John Silvius and Dr. Robert Mackenzie for enabling me to strive for new heights in the Adirondacks. I am sure that memories of the panoramic views and camaraderie will eventually outlive memories of the pain.

I am especially grateful to my parents, Roslyn and Monte, for their constant support and motivation. They encouraged me to pursue my own interests and, by example, inspired me to realize my goals. I would also like to thank my sisters, Mia and Leora, for showing me how it's done, and my grandmother, Baba Freda.

PUBLICATIONS

Swartzman, E., Silverman, M., and Meighen, E. A. (1992) The *luxR* Gene Product is a Transcriptional Activator of the *Vibrio harveyi lux* Promoter. *J. Bacteriol.* **174**:7490-7493.

Swartzman, E., and Meighen, E. A. (1992) *Vibrio harveyi* RNA Polymerase: Purification and Resolution from Gyrase A. *Biochem. Cell. Biol.* **70**:698-702.

Swartzman, E., Cao, J. G., Miyamoto, C., Graham, A., and Meighen, E. A. (1991) Differences in the Regulatory Signals Controlling Expression of the *Vibrio lux* Systems. In *Bioluminescence and Chemiluminescence: A Current Status* (P. E. Stanley and L. J. Kricka, eds) pp. 27-30. John Whiley and Sons.

Swartzman, E., Kapoor, S., Graham, A. F., and Meighen, E. A. (1990) A New *Vibrio fischeri lux* Gene Precedes a Bidirectional Termination Site for the *lux* Operon. (1990) *J. Bacteriol.* **172**:6797-6802.

Swartzman, E., Miyamoto, C., Graham, A. F., and Meighen, E. A. (1990) Delineation of the Transcriptional Boundaries of the *lux* Operon of *Vibrio harveyi* Demonstrates the Presence of Two New *lux* genes. *J. Biol. Chem.* **256**:3513-3517.

Cao, J. G., **Swartzman, E.,** Miyamoto, C., and Meighen, E. A. (1989) Regulatory Mutants of the *Vibrio harveyi lux* System. *J. Biolum. Chemilum.* **3**:207-212.

ABSTRACT

Swartzman, E., and Meighen, E. A. (1988) Identification of Promoters in the Luminescent Operon of the Marine Bacterium *Vibrio harveyi* using S1 Nuclease Mapping and Run-off Transcription. *Proc. Can. Fed. Biol. Soc.* **31**:240.

ORIGINAL CONTRIBUTIONS TO KNOWLEDGE

I. The *V. harveyi lux* operon consists of two new *lux* genes, *luxG* and *luxH*, followed by a rho-independent transcription terminator. The *V. fischeri lux* operon contains one new *lux* gene, *luxG*, that precedes a rho-independent bidirectional transcription terminator.

II. The *V. harveyi* RNA polymerase was purified and shown to have a similar subunit composition to that of the *E. coli* enzyme. *In vitro* transcription studies demonstrated that the *V. harveyi* and *E. coli* RNA polymerases share the same promoter specificity and initiate transcription 123 bp upstream from the initiation codon of *luxC*.

III. A 100 kDa protein copurified with the *V. harveyi* RNA polymerase and constituted over 30% of the protein eluting from a dsDNA-cellulose column. N-terminal sequence analysis suggested that it is the A subunit of gyrase, which has been linked to light production in bioluminescent bacteria.

IV. The transcriptional start site of the *V. harveyi lux* operon was mapped to 26 bp upstream from the initiation codon of *luxC* (-26). This is in contrast to the start site for *lux* mRNA isolated from recombinant *E. coli* containing the *V. harveyi lux* operon and for RNA synthesized *in vitro* from a *lux* DNA template, which mapped to position -123. Expression of *luxR* in recombinant *E. coli* shifted the start site from position -123 to -26, demonstrating that LuxR is required for accurate transcription of the *V. harveyi lux* promoter.

V. A protein found exclusively in cell extracts of *V. harveyi* that bound to a region of DNA upstream from the *V. harveyi lux* promoter was purified and identified as LuxR. LuxR bound to two A + T rich regions: -290 to -253 and -116 to -170, and poly(dA-dT) successfully competed with the *lux* DNA for binding by LuxR. LuxR of *V. harveyi* is the first DNA-binding protein to be detected, purified and identified from a luminescent bacterium. *In vitro* transcription studies showed that LuxR blocked transcription initiation at the -123 promoter.

LIST OF FIGURES

CHAPTER 4

CHAPTER 5

CHAPTER 6

LIST OF TABLES

LIST OF ABBREVIATIONS

ATP	adenosine triphosphate
bp	base pairs
CoA	coenzyme A
cpm	counts per minute
DTT	dithiothreitol
HEPES	N-2-hydroxyethylpiperazine-N-2-ethanesulphonic acid
kbp	kilobase pairs
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
kDa	kilodaltons
NADP	ß-ninotinamide adenine dinucleotide phosphate
PAGE	poly-acrylamide gel electrophoresis
PMSF	phenylmethanesulfonyl fluoride
SDS	sodium dodecyl sulfate
TRIS	tris(hydroxymethyl)amino methane

CHAPTER 1

GENERAL INTRODUCTION*

References for chapters 1 through 7 can be found under "GENERAL REFERENCES".

HISTORICAL PERSPECTIVES

The phenomenon of bioluminescence is characterized as the emission of visible light by a living organism without the generation of heat, a feature that was recognized by Aristotle in the fourth century BC. Bioluminescent organisms are found across the evolutionary spectrum and inhabit terrestrial, marine and freshwater niches (Harvey, 1952). The bioluminescent reactions, including the enzymes (luciferases) and substrates (luciferins), are varied and evolutionary distinct, save for the absolute requirement for and extreme sensitivity to O_2 by all the systems studied to date. It has been proposed that bioluminescence first developed in anaerobes as a means of metabolizing O_2 and has since evolved alongside aerobic metabolism, conferring selective advantage to some organisms (McElroy and Seliger, 1962). It is generally believed, for example, that the firefly emits light as a mating signal, while most marine bioluminescent bacteria exist in a symbiotic relationship with certain fish and squid. The requirement for O_2 was first observed by Robert Boyle, who noted that air was required for light emission from bacteria or fungi residing on rotten wood (Boyle, 1668). Later, the French physiologist Raphaël Dubois used hot- and cold-water extracts of the luminous mollusc *Pholus dactylus* to demonstrate the presence of a material he called luciferin (from Lucifer, or light-bearer) in the hot-water extract that was essential for light emission by a substance in the-cold water extract he called luciferase (Dubois, 1887). He recognized that luciferase, being heat-sensitive, had enzymatic properties, while luciferin was active in both cold and hot water.

Our current understanding of bioluminescent systems has been significantly advanced with the onset of modern biochemistry and molecular biology. Several luciferases have been purified and their corresponding genes have been cloned and expressed from bacteria, fireflies, clickbeetles, jellyfish, and crustaceans (Meighen, 1991). The ease and simplicity of light detection has led to the widespread application of luciferase as a sensor for specific metabolites within the cell, as well as a reporter for gene expression and regulation. There is also now an expanding field exploring the complex regulation of bioluminescence in marine bacteria. The

following thesis examines the molecular mechanisms involved in transcriptional regulation of the luminescent (*lux*) operon of the marine bacterium *Vibrio harveyi*.

BACTERIAL GENE REGULATION

The bacterial operon consists of genes related by function coordinately expressed in one transcriptional unit. Thus genes involved in, for example, tryptophan synthesis or light production are regulated as a functional unit and are translated from a single mRNA species. Features of the operon include promoters, operators and terminators, each serving to regulate transcription initiation and termination.

Promoters

Transcription initiation occurs at a region of DNA, or promoter, that directs the RNA polymerase binding and subsequent activation of transcription initiation (von Hippel et al., 1984; McClure, 1985). A compilation of known promoter regions using the σ^{70} RNA polymerase holoenzyme in *E. coli* has yielded a consensus sequence for two conserved regions of DNA centred at -10 and -35 base pairs from the transcriptional start start (Rosenberg and Court, 1979; Siebenlist et al., 1980). The consensus sequence for the -35 region is TTGaca and TAtaaT for the -10 region, with a spacer region of approximately 17 base pairs. The upper case letters represent the most highly conserved bases within the hexamer. The general rule follows that promoters with the most highly conserved sequences have maximal promoter strength and mutations that deviate from the consensus result in weaker promoters. Synthetic promoters containing the -10 and -35 consensus sequences initiate transcription at very high frequencies both *in vivo* and *in vitro* (McClure, 1985).

Transcription Initiation

RNA polymerase first recognizes and binds to the promoter forming an inactive, transient closed promoter complex (Chamberlin, 1974). This complex then isomerizes to the open complex form where the RNA polymerase forms a tight association with the DNA from approximately position -55 to +20 from the

transcriptional start site with major contacts in the -35 and -10 consensus sequences. This step also involves "melting" or unwinding of the double helix from position -9 to +3 (Siebenlist et al., 1980). Transcription initiation occurs when the first complementary nucleotide, usually ATP or GTP, binds to the enzyme and forms a phosphodiester bond with the second complemetry nucleotide (von Hippel et al., 1984). This transforms the open complex to the more stable ternary complex (polymerase-DNA-RNA). Abortive initiation has been observed to occur *in vitro*, where short polynucleotides of 2 to 9 residues are synthesized even in the presence of all four nucleotides. It is believed that RNA polymerase remains bound to the DNA and promoter region during abortive initiation. The ternary complex shifts to the elongation mode when the σ factor (see section on RNA polymerase) dissociates from the holoenzyme to yield core polymerase, which continues incorporating nucleotides into the nascent RNA chain. The σ factor is then recycled by binding to free core polymerase and reinitiating at a free promoter site.

Regulation at the Promoter

Aside from such considerations as promoter strength and RNA polymerase promoter specificity (see below), transcription initiation can be positively and negatively regulated by a number of factors. In general, repressor proteins block initiation by competing with RNA polymerase for binding within the promoter region (Reznikoff et al., 1985). As expected, there are numerous exceptions, including AraC, a repressor for the arabinose operon which binds to two distinct regions of DNA 211 bp apart. Repression is achieved only when AraC occupies both sites and forms a loop of DNA (Lee and Schleif, 1989). Activators serve to stimulate transcription by several different means. The classical model for transcriptional activation held that DNA-bound regulators would interact with an adjacent RNA polymerase molecule to facilitate access to the promoter site. It is clear now that they may bind near or far from the promoter (Raibaud and Schwartz, 1984) or at enhancer-like elements (Su et al., 1990). Transcriptional activators may indirectly affect initiation by altering the DNA conformation of the promoter and thus favour

open complex formation (Reznikoff et al., 1985). However, there is mounting evidence indicating that direct activator-polymerase interactions are responsible for transcriptional activation (see Catabolite Repression section). As the location of binding sites relative to the start site of transcription can vary, it has been hypothesized that DNA-looping is an important aspect of transcriptional regulation (Schleif, 1988). Direct visualization of a looped structure has been observed for DNA bound by RNA polymerase and NtrC, the general nitrogen regulatory protein (Su et al., 1990). Other factors may be required for regulator-polymerase interactions. It has been shown that IHF (integration host factor) binds between the *nifH* promoter and an upstream activator binding site (NifA) and bends the DNA to facilitate NifA-RNA polymerase contacts (Hoover et al., 1990). It has been also postulated that regulator-induced DNA bending results in a conformational change in the promoter region, creating a favourable RNA polymerase binding site (Wu and Crothers, 1984).

Transcription Termination

Elongation continues until the RNA polymerase encounters a termination signal that causes cessation of transcription, release of transcript from the ternary complex and dissociation of the enzyme from the template (Platt, 1986). This signal can be one of two forms: rho-independent or rho-dependent. At a rho-independent signal, termination *in vitro* can be accomplished by purified RNA polymerase alone, without ancillary proteins. It consists of a GC-rich stem-loop structure with a free energy of -10 to -30 kcal/mol centred 16-20 residues upstream from the 3'-terminus, followed by a stretch of 6 to 8 uridine residues. Following formation of the stem-loop structure in the nascent RNA chain, the RNA polymerase pauses within the stretch of poly dA-dT residues and causes disruption of the 5'-portion of the RNA-DNA hybrid helix. As the remaining rU-dA hybrid helix is considerably unstable, the remaining transcript dissociates from the template allowing the unwound DNA helix to reassociate and the RNA polymerase to dissociate from the template.

There exists a class of terminators that requires an additional factor, rho, for

transcription termination (Platt, 1986). The termination endpoints are heterogeneous and a comparison of several different rho-dependant termination sites has shown that these sequences lack strong homology with one another. The model proposes that rho, a RNA-dependent ATPase, interacts with a region of untranslated, unstructured RNA upstream from the termination site and catalyses the release of the transcript at an RNA polymerase pause site (Bear and Peabody, 1988).

RNA Polymerase and Sigma Factors

Core RNA polymerase is a multienzyme complex consisting of four subunits with the stoichiometry $\beta\beta'\alpha_2$. This complex contains the catalytic polymerase function and requires σ factor for promoter recognition. Although the holoenzyme can bind specifically to DNA, σ factor does not bind DNA in the absence of core enzyme (Burgess et al., 1969) The majority of promoters requires the predominant σ^{70} factor in *E. coli* or σ^{43} in *Bacillus subtilis*. There are, however, several other alternate σ factors in both organisms that are required for specific promoter recognition. In *B. subtilis*, many of the alternate σ factors are essential in coordinating gene expression during endospore formation. Sporulation is seen as a result of a cascade of transcriptional activation, mainly coordinated by the sequential expression of the alternate σ factors (Losick and Pero, 1981). In *E. coli*, the two best characterized alternate transcription factors are involved in nitrogen regulation (σ^{54}), and the heat-shock response (σ^{32}). Bacteriophages of both organisms also encode one or several alternate σ factors that are able to specifically recognize phage promoters. The σ factors have been organized into three groups of σ^{70}-like factors (Lonetto et al., 1992). Group 1 consists of the primary σ factors from various prokaryotes and share at least 51% amino acid sequence identity. Most of the divergence is restricted to poorly conserved regions and can be explained by genetic drift. Group 2 σ factors, including HrdA, HrdC, and HrdD from *Streptomyces coelicolor* (Buttner et al., 1990) and SigS (Lange and Hengge-Aronis, 1990) from *E. coli* are not required for exponential cell growth. Although they share minimum sequence identity with the

group 1 sigmas, they are most similar in the DNA-binding region, indicating that they recognize similar promoter sequences. Group 3 contains subgroups of the alternate σ factors that are responsible for various physiological functions including heat-shock, sporulation and flagellar synthesis. These sigmas share at maximum 27% identity with the primary σ factors. The bacteriophage sigmas are only weakly related in sequence to the σ^{70} family, while no similarity has been found between members of the σ^{70} family and σ^{54} , the factor responsible for transcribing nitrogen- regulated genes.

Based on the alignment of sequences obtained from σ^{70} family members, the subunit can be divided into four main regions, each containing two or more subregions (Helmann and Chamberlin, 1988; Lonetto et al., 1992). Region 1 consists of the highly conserved amino terminal in the primary sigmas of *E. coli* and *B. subtilis*. Its two subregions are divided by a stretch of comparatively less conserved amino acids. While subregion 1.1 is found exclusively in the primary sigmas (and SigS), several residues in subregion 1.2 are conserved in many primary and alternate σ factors. Subregion 1.2 may be involved in core binding; however its specific function remains unknown. Region 2, containing some of the most highly conserved amino acids among the σ factors, is subdivided into four regions. Based on the similarity to eukaryotic RNA-binding proteins and the presence of highly conserved aromatic residues, subregions 2.1 and 2.3 are thought to be involved in unwinding of the DNA duplex during open promoter complex formation. The aromatic residues are believed to intercalate with the double helix and facilitate DNA melting. Mutations altering recognition of the -10 consensus sequence have been found within subregion 2.4. Thus this segment has been postulated to be involved in direct contact with -10 promoter sequences. There is no known function for subregion 2.2, the most conserved region of σ factors. Region 3 is found in all primary but in only some alternate σ factors. Subregion 3.1 contains sequences that may form a helix-turn-helix DNA binding domain, while conserved residues within subregion 3.1 are mainly acidic. Region 4 is thought to be involved in -35 promoter specificity, as mutations within this segment affect recognition at the -35 hexamer. Whereas

subregion 4.1 may form an amphipathic α-helix, subregion 4.2 contains a highly conserved helix-turn-helix DNA-binding domain, consistent with its role in promoter recognition.

Catabolite Repression

There exists a class of operons whose expression is depressed when glucose or other rapidly metabolized sugars are used as carbon sources in the growth media. This occurrence, termed catabolite repression, is mediated by CRP or cyclic AMP Receptor Protein (also known as CAP - Catabolite gene Activator Protein) (de Crombrugghe et al., 1984). Intracellular cAMP levels rise in response to diminished carbohydrate concentrations activating various catabolic operons, including among others *lac*, *ara*, *mal*, and *gal*. In the presence of cAMP, CRP binds to a consensus sequence upstream from the promoter and activates transcription. The promoter usually lacks a recognizable -35 consensus sequence and the CRP binding site may vary from position -106 (for *malE*) to -41 (for *gal*). It has been ascertained that, in the case of *lac* activation, RNA polymerase and CRP bind cooperatively to their respective sites indicating that there are direct CRP-RNA polymerase contacts (Reznikoff, 1992). Further evidence of protein-protein contacts has been provided by the generation of CRP and RNA polymerase α subunit mutations (Eschenlauer and Reznikoff, 1991; Igarashi and Ishihama, 1991). Mutations within distinct domains of each polypeptide interfered with transcriptional regulation. CRP is known to induce a 90°C bend in the DNA target sequence (Shultz et al., 1991) and curved DNA can stimulate transcription in the absence of CRP (Bracco et al., 1989). However no direct role for CRP-induced DNA bending has been found for its role in transcriptional activation. Because catabolite sensitive promoters generally show poor homology with the -35 consensus sequence, it is thought that CRP may be able to assist the RNA polymerase in recognizing the correct promoter sequences (de Crombrugghe et al., 1984). CRP may also function as a repressor by binding at or near a promoter element and blocking access to RNA polymerase (Malan and McClure, 1984).

BIOLUMINESCENT BACTERIA

Luminescent bacteria are classified into four genera: *Vibrio, Photobacterium, Alteromonas,* which are marine in origin, and *Xenorhabdus* (Baumann et al., 1983; Hastings et al., 1985). The bacteria of the latter genus are found in a symbiotic relationship with nematodes in a parasitic infection of caterpillars (Thomas and Poinar, 1979), and have recently been isolated from human wounds (Farmer et al., 1989). The distribution of luminous bacteria in the marine environment is wide spread, existing as free-living, symbiotes, parasites, and saprophytes. The most common form appears to be as symbiotes in the light organ or digestive tract of certain fish. Light emission from the light organ appears to facilitate feeding, flight from prey and communication (Nealson and Hastings, 1979). The benefit to the symbiotic bacteria appears to be in the provision of an exclusive habitat to sustain their growth. There are common species of light organ-luminous bacteria: *V. fischeri, P. leiognathi, and P. phosphoreum.* Although *lux* genes have been cloned from all three species, the most thoroughly studied is *V. fischeri* with respect to *lux* gene regulation. In contrast, *V. harveyi* has been isolated only from the ocean water, where it exists as a free-living species (Nealson and Hastings, 1979; Baumann and Baumann, 1981). There exists several key differences in the *lux* gene organization and regulation of *V. harveyi* and *V. fischeri*, though it is not clear whether these are related to the differcences in their habitats.

Bioluminescence Reaction

The light emitting reaction in bioluminescent bacteria is catalyzed by luciferase, a heterodimer consisting of an α and a β subunit. A long-chain aldehyde and $FMNH_2$ are oxidized to form the corresponding fatty acid and FMN:

$$FMNH_2 + O_2 + R\text{-}CO\text{-}H \longrightarrow FMN + R\text{-}COOH + H_2O + light$$

The mechanism involves formation of an enzyme-bound 4a-peroxyflavin that forms a stable intermediate with aldehyde. Photons with a wavelength maximum of 490nm are emitted as the substrates are oxidized (Hastings et al., 1985).

The aldehyde substrate is provided by the fatty acid reductase complex, a multi-enzyme complex consisting of a reductase, a transferase and a synthetase subunit. The fatty acid reductase complex can be purified from *P. phosphoreum* as a weakly interacting multienzyme complex of 5 x 10^5 Da (Wall and Meighen, 1986). The complex consists of a central core of four reductase subunits, each interacting with a synthetase subunit which in turn is weakly associated with a transferase subunit. Each of the subunits can be specifically acylated during the reaction (Wall et al., 1984), and can thus be identified in bacterial extracts. Although fatty acid reductase activity has been detected only in extracts of *Photobacterium* spp., acylated polypeptides corresponding to the transferase, synthetase, and reductase have been identified in *V. harveyi* (Wall et al., 1984) and *V. fischeri* (Boylan et al., 1985).

The transferase diverts fatty acids from the fatty acid biosynthetic pathway into the luminescence system. It catalyses the transfer of activated fatty acyl groups to water or other oxygen and thiol acceptors, with the acyl group forming a covalent bond with the enzyme during the reaction:

$$RCOX + H_2O \longrightarrow COOH + XH$$

The acyl donner can be acyl-ACP (acyl carrier protein), acyl-coenzyme A, or acyl-*p*-nitrophenol, and the enzyme shows maximal activity for acyl groups with chain lengths of 14 carbons (Ferri and Meighen, 1991). This is consistent with the finding that tetradecanal has been found in lipid extracts of luminescent bacteria (Ulitzer and Hastings, 1979) and is the preferred substrate for luciferase (Meighen et al., 1982). The transferase was found to be related to eukaryotic serine esterases, which diverts short-chain fatty acids from the fatty acid biosynthetic pathway. A serine residue was identified as the site of acylation, and a charge relay system involving a conserved histidine residue has been proposed to be involved with the serine active site (Ferri and Meighen, 1991).

The synthetase subunit activates the fatty acid with ATP, and forms an acyl-AMP bound intermediate (Wall et al., 1986). The site of acylation on the synthetase was determined to be at a cysteine residue near the carboxyl terminal, where the enzyme

bound acyl group can interact with the reductase subunit to effect acyl group transfer between these two proteins (Soly and Meighen, 1991). The reductase subunit then catalyses the NADPH-dependent reduction of the activated fatty acid:

$$RCOOH + ATP + NADPH \longrightarrow NADP + AMP + PP_i + RCHO$$

The aldehyde product is used as a substrate for luciferase and is susequently converted back to the fatty acid form.

Lux Structural Genes

Genes corresponding to the polypeptides required for light generation and aldehyde biosynthesis have been cloned from several species of bioluminescent bacteria, including *V. harveyi, V. fischeri, P. phosphoreum, P. leiognathi,* and *X. luminescens* (Meighen, 1991) and are located within one operon. All contain the *luxA* and *luxB* genes, which code for the α and ß subunits of luciferase, as well as the *luxC*, *luxD*, and *luxE* genes, which code for the reductase, transferase, and synthetase subunits, respectively. With one exception, the gene order *luxCDABE* is the same for all species examined to date. In most strains of the *Photobacterium* species, an additional gene, *luxF*, is located between the *luxB* and *luxE* genes. *luxF* is homologous in amino acid sequence to *luxA* and *luxB* and is thought to have arisen through gene duplication (Soly et al., 1988). Although its function in relation to bioluminescence is unknown, it has been identified as a flavoprotein that had been previously purified from *P. leiognathi* (O'Kane et al., 1987).

Regulation of Bioluminescence

The most distinctive feature of the development of bioluminescence in marine bacteria is a phenomenon termed autoinduction. Luminescence at first lags behind cellular growth and then increases up to several thousand-fold at high cell densities (Nealson et al., 1970). While this lag has been attributed in part to the removal of inhibitors in the growth media, it is primarily due to the accumulation in the growth media of a small molecule called autoinducer (Eberhard, 1972; Rosson and Nealson,

1981). It is for this reason that luminescence occurs only when the cells are confined in a closed environment, such as in the light organ of symbiotic fish. There are several other factors that have been linked to the regulation of bioluminescence including: catabolite repression, the SOS response, as well as arginine, iron and oxygen levels (Nealson et al., 1972; Ulitzer, 1989; Nealson et al., 1970; Haygood and Nealson, 1985; Nealson and Hastings, 1977). Of all of these, however, only catabolite repression has been shown to be intimately involved in the regulatory process.

Autoinducer

The autoinducer of *V. fischeri* has been purified and identified as *N*-ß-ketocaproyl homoserine lactone (Eberhard et al., 1981). Because it is freely diffusible across biological membranes (Kaplan and Greenberg, 1985), the autoinducer is able to accumulate in the cells and medium during growth and induces the luminescence system at threshold concentrations. This autoinducer is species specific; it is able to stimulate luminescence in *V. fischeri* and *V. logei*, but not in other luminous species (Eberhard, 1972; Nealson, 1977; Engebrecht et al., 1983). The precursors of the *V. fischeri* autoinducer, which is composed of an amino acid and a fatty acid metabolite, are thought to be *S*-adenosylmethionine and either 3-oxo-hexanoyl coenzyme A or 3-oxo-hexanoyl acyl carrier protein (Eberhard et al., 1991). A factor that stimulates light production in *V. harveyi* was detected in the growth medium (Eberhard, 1972), and was purified and identified as *N*-ß-hydroxybutryl homoserine lactone (Cao and Meighen, 1989). Although it is similar in structure to the *V. fischeri* autoinducer, there are some significant chemical and functional differences. The *V. harveyi* autoinducer does not stimulate luminescence in *V. fischeri* and, unlike the *V. fischeri* autoinducer, is heat labile (Eberhard, 1972) and more polar (Cao and Meighen, 1989). A study of several analogues of the *V. fischeri* autoinducer suggest that the bioluminescence system can tolerate only minor changes in autoinducer structure (Eberhard et al., 1986). As both *V. harveyi* and *V. fischeri* autoinducers are composed of an intermediate from fatty acid metabolism, and the luminescence

system diverts fatty acids from the normal lipid metabolic pathways, it is possible that the autoinducer is an indicator of the level of metabolites required as substrates for light production (Meighen, 1988). Both autoinducers are similar in structure to A factor (isocapryl-δ-butryl lactone), a regulatory factor produced by some species of *Streptomyces* that is responsible for self-induction of sporulation and streptomycin synthesis. These molecules can be referred to as bacterial pheromones or hormones that act as environmental sensory molecules (Silverman et al., 1989).

The *V. fischeri lux* Operons

Through transposon mutagenesis and complementation studies, the genes required for structural and regulatory functions of the *V. fischeri* luminescence system were assigned to two operons, the left (L) and right (R) *lux* operons (Engebrecht et al., 1983; Engebrecht and Silverman, 1984). *E. coli* recombinants containing both operons were able to generate light in the same quantities and in the same cell density-dependent manner as *V. fischeri*. The genes encoding luciferase (*luxA* and *luxB*), and aldehyde biosynthesis (*luxC*, *luxD*, and *luxE*) were located in operon R along with *luxI* in the order of *luxICDABE*. Transposon insertions within the *luxI* gene rendered the bacteria dim and autoinducer deficient. While the mutants could respond to exogenously added autoinducer, they were unable to produce autoinducer activity. The *luxI* gene was therefore determined to be necessary for autoinducer synthesis. *LuxR*, located immediately upstream and transcribed in the opposite direction of *luxICDABE*, constitutes operon L. As with the *luxI* gene, mutations within *luxR* generated bacteria that were dark and unable to produce autoinducer. However, because these mutants were also unable to respond to exogenously added autoinducer, it was deduced that *luxR* must encode a receptor for this sensory molecule. Through *lacZ* transcriptional fusion studies it was determined that expression of operon R, containing *luxI*, is dependent on autoinducer activity and, as with development of luminescence, occurs at high cell densities. Thus the gene responsible for autoinducer synthesis is autoregulated by its gene product. In

minicell programming experiments (Engebrecht and Silverman, 1984) and transposon mutagenesis (Engebrecht and Silverman, 1986) it was shown that the autoinducer and its receptor negatively regulate the expression of *luxR* possibly through posttranscriptional control of operon L. The following model was then proposed to explain how light production is induced with cellular growth. At low cell densities autoinducer and *lux* proteins are produced at constitutive levels. As cell density increases, the autoinducer accumulates within the cell and growth medium and at a critical threshold level, interacts with the product of the *luxR* gene or receptor protein. A positive feedback loop is established when the complex activates transcription of the operon R, generating more light and increasing autoinducer levels. The exponential rise in light production reaches a plateau when LuxR and autoinducer negatively control LuxR levels within the cell. This model was used as the basis for further investigation into the regulation of bioluminescence, in particular with regard to catabolite repression.

Catabolite Repression in Bioluminescence

Catabolite repression of bioluminescence was observed in both *V. harveyi* (Nealson et al., 1972) and *V. fischeri* (Ruby and Nealson, 1976), although the effect in *V. fischeri* is transient. While addition of glucose to batch cultures of *V. fischeri* resulted in a temporary decrease in light emission, this repression was not reversed by cAMP and prior growth in the presence of glucose eliminated its effect. In contrast, when grown in phosphate-limited chemostat cultures, light production in *V. fischeri* was permanently repressed by glucose and was restored with the addition of cAMP or autoinducer (Friedrich and Greenberg, 1983). Unlike *V. fischeri*, repression by glucose in batch cultures of *V. harveyi* is permanent and can be overcome by addition of cAMP to the growth medium (Nealson et al., 1972).

LuxR-CRP Regulatory Circuit

Direct evidence for the involvement of CRP in the regulation of bioluminescence

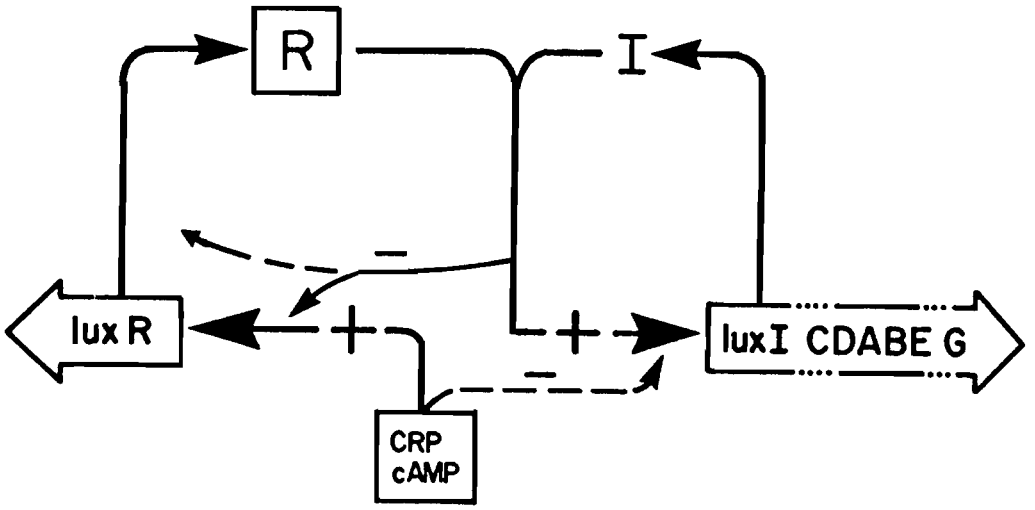
in *V. fischeri* was obtained by examining *lux* gene regulation in *E. coli* cAMP and CRP mutants (Dunlap and Greenberg, 1985; Dunlap and Greenberg, 1988; Dunlap and Ray, 1989). It was shown that transcription from the *luxR* promoter (operon L) was stimulated by cAMP-CRP and was negatively regulated by high levels of LuxR and autoinducer. In contrast, low levels of the autoinducer-LuxR complex caused an increase in *luxR* expression (Shadel and Baldwin, 1991). It was also apparent that transcription of *luxICDABE* (operon R) was activated by cAMP and CRP in the presence of and depressed in the absence of autoinducer and LuxR. The activation of *luxICDABE* by cAMP-CRP is indirect, as overexpression of *luxR* eliminated the requirement of cAMP-CRP. The model (see Figure 1) proposes that at low cell densities, cAMP-CRP activates transcription of *luxR* while simultaneously decreasing expression of *luxICDABE*. As LuxR begins to accumulate, it interacts with basal levels of autoinducer and stimulates expression of both the right and left *lux* operons, fuelling a double positive feedback loop. At high concentrations of LuxR and autoinducer transcription of *luxR* is inhibited, thereby counterbalancing the effect of cAMP-CRP and restraining the runaway feedback loop.

Because this model was based on studies in *E. coli* CRP and cAMP mutants, it failed to explain the transient nature of glucose repression in batch cultures of *V. fischeri* and wild type *E. coli* recombinants. Although the generation of *cya*-like and *crp*-like mutants of *V. fischeri* confirmed the involvement of cAMP-CRP in *lux* gene regulation (Dunlap 1989) the contrast seen in chemostat- versus batch-grown *V. fischeri* remains unresolved.

V. fischeri lux Promoters and Operator

The DNA sequence encompassing *luxI*, *luxR*, and the intervening region was determined and the promoter elements were identified (Engebrecht and Silverman, 1987; Devine et al., 1988). The two genes are divergently transcribed with a span of 155 residues separating the two transcriptional start sites. While the promoter region for operon L (P_1) contains the characteristic -10 and -35 consensus sequences, the

Figure 1: Model of *lux* gene regulation in *V. fischeri*. The regulatory circuit of LuxR-autoinducer and CRP-cAMP are depicted. The positive regulation of *luxR* by LuxR-autoinducer at low cell densities is not shown (From Meighen, 1991).


R
I
lux R
luxI CDABE G
CRP
cAMP

promoter region of operon R does not contain a recognizable -35 sequence. This is an indication that a positive regulator may be required for RNA polymerase binding. A CRP binding site (TGTCACA) was located at 59 bp upstream of the transcriptional start site for *luxR* as well as a 20 bp palindrome (ACCTGTAGGA/ TCGTACAGGT) centred 40 bp upstream of the transcriptional start site for operon R. The ability of CRP to bind to its site was confirmed with DNA footprinting studies (Shadel et al., 1990). Mutational analysis of the palindrome demonstrated that transcription from P_r required intact palindrome, LuxR and autoinducer, indicating that this sequence was the site of LuxR-autoinducer binding (Devine et al., 1989) and was required for both *luxR* positive and negative autoregulation (Shadel and Baldwin, 1991).

In addition to LuxR, autoinducer and the *lux* operator, repression of *luxR* was shown to require sequences within operon R (*luxICDABE*) (Shadel and Baldwin 1992a). Deletion analysis of the operon identified a negative acting element within *luxD* located 2.0 kb from the *lux* operator. It is similar in sequence to the *lux* operator (11/20 identity) and can simulate LuxR-mediated transcriptional regulation, albeit weakly, when it replaces the *lux* operator in the control region. A model for both positive and negative regulation of *luxR* by LuxR was formulated to include this intragenic LuxR-binding site. At low LuxR and autoinducer levels early in cellular growth, autoinducer-LuxR binds to the high affinity *lux* operator within the control region and activates both *luxR* and *luxICDABE* transcription. At high cell densities and increased autoinducer and LuxR concentrations, autoinducer-LuxR binds to the low affinity intragenic operator as well as to the high affinity operator, resulting in negative transcriptional autoregulation of *luxR*. Although the mechanism of repression has yet to be elucidated, it has been proposed that this might involve DNA looping between the LuxR-bound operator sites (Shadel and Baldwin, 1992a).

The role of CRP in transcriptional regulation of P_l was elucidated by examining the effects of a 4 bp change within the CRP binding site (Shadel and Baldwin, 1992b). The mutation caused a complete loss of *luxR* positive autoregulation, which

was restored when *luxR* was provided *in trans*. Therefore LuxR-autoinducer is able to stimulate expression of *luxR* in a cAMP-CRP dependent and independent manner. Primer extension analysis demonstrated three sites of initiation of *luxR* transcription, two of the three requiring cAMP-CRP for activation.

V. fischeri luxR Gene Product

Attempts to examine LuxR activity *in vitro* have failed, in spite of its overproduction and purification from *E. coli* (Kaplan and Greenberg, 1987). Only non-specific DNA-binding activity was observed with the purified protein, which was was autoinducer-independent. Furthermore, autoinducer-LuxR binding could not be demonstrated *in vitro*. Overproduction of LuxR in *E. coli* caused granular inclusion bodies to form which could only be solubilized in 6 M guanidine hydrochloride. It was suggested that the inability of purified LuxR to bind autoinducer and *lux*-specific DNA was an indication that the protein was not renatured in its functional state.

Much more information regarding LuxR activity was obtained by combining mutational analysis and *in vivo* activity to define functional domains of LuxR. Randomly generated point mutations within LuxR identified two regions required for transcriptional activation of the P_1 promoter (Slock et al., 1990; Shadel et al., 1990). Several of the amino acid substitutions within the amino-terminal of LuxR (residues 79-127) resulted in mutations that could be suppressed in the presence of an excess of exogenously added autoinducer. It was concluded that this region was the site of autoinducer binding. A second cluster of point mutations that affected LuxR activity was found in the carboxyl-terminal, spanning residues 184-230. Because this region shows sequence similarity to the DNA-binding domains of a class of transcriptional regulators, including members of the two-component environmental-sensing systems (Henikoff et al., 1990), it was extrapolated that this region is responsible for *lux*-operator binding. This family of transcriptional activators can be subgrouped based on similarities within the amino- and carboxyl-terminals. LuxR is most similar to LasR and 28K-UvrC: LasR is a transcriptional activator of *lasB*, the *Pseudomonas*

elastase structural gene (Gambello and Iglewski, 1991), while 28K-UvrC is thought to be a positive regulator of *uvrC* (Kunst et al., 1988). Elastase, which contributes to the pathogenicity of the microorganism, is a metalloprotease that is produced only in late-logarithmic and stationary growth phases. This has lead to the hypothesis that, like the bioluminescence system, elastase is expressed in a cell-density dependent manner by responding to environmental factors (Gambello and Iglewski, 1991).

The other members of this group of transcriptional activators, the two component regulatory proteins, share similar amino-terminal domains (Ronson et al., 1987). These are systems that respond to environmental stimuli, including among others, osmolarity (*envZ/ompR*), nitrogen limitation (*ntrB/ntrC*), and phosphate limitation (*phoR/phoB*), and are composed of a sensor kinase and a transcriptional regulator. The sensor, usually a transmembrane protein with a periplasmic and cytoplasmic domain, detects an environmental stimulus and perhaps through conformational change transmits the signal to the cytoplasmic domain. This domain then phosphorylates the amino-terminal of the regulator, thereby activating transcription of its target gene(s). The DNA-binding domain of the regulator, located within the carboxyl-terminal, contains a predicted helix-turn-helix motif that shares significant homology with region 4 of RNA polymerase primary σ factors (Kahn and Ditta, 1991; Stout at al., 1991), the region thought to be involved in recognizing the -35 consensus promoter sequence. It is interesting to note that like σ factor, the regulatory proteins of the two-component systems (including LuxR) do not bind specific DNA sequences *in vitro*. The exception is MalT, which contains an amino-terminal five times as large as that of LuxR (Henikoff et al., 1990), and requires both maltotriose and ATP for activation (Raibaud and Richet, 1987; Vidal-Ingigliardi et al., 1991). As LuxR contains an amino-terminal that differs from that of regulators of the two-component system, it would be consistant with autoinducer binding rather than phosphorylation causing transcriptional activation (Slock et al., 1990).

Further detailed analyses of functional regions within LuxR was obtained by the generation of a series of amino-terminal (Choi and Greenberg, 1991) and carboxyl-

terminal (Choi and Greenberg, 1992) deletion mutants. Complete removal of the autoinducer-binding amino-terminal domain allowed for autoinducer-independent transcriptional activation. It was proposed that in the absence of autoinducer, the amino-terminal of full length LuxR (residues 20-162) blocks the DNA-binding activity of an independently folded carboxyl-terminal (residues 163-250), whereas autoinducer-bound amino-terminal unmasks this function. It was also determined that residues 10-20 were required for *luxR* autoinducer-dependent autorepression but not required for *luxICDABE* activation (Choi and Greenberg, 1991). The ability of 3'-deletion mutations of LuxR to both activate transcription of *luxICDABE* and autorepress *luxR* was tested in order to separate the DNA-binding domain from the transcriptional activating domain in the carboxyl-terminal. Although deletions up to but not including the helix-turn-helix motif rendered the protein unable to activate *luxICDABE* expression, they were able to autoregulate *luxR* transcription, indicating that residues 211-250 were essential for transcriptional activation. Truncated proteins missing some or all of the helix-turn-helix region (residues 196-210) were unable to perform both tasks, confirming that this predicted DNA-binding region does interact with *lux* DNA (Choi and Greenberg, 1992).

Other Regulatory Controls

It has become apparent that other factors, besides autoinducer-dependant autoregulation, are involved in regulation of bioluminescence in *V. fischeri*. Low levels of iron or oxygen in the culture medium results in slower growth but allows for induction of luciferase synthesis to occur at lower cell densities (Nealson and Hastings, 1977; Haygood and Nealson, 1985b). It appears, though, that the effect of iron on *lux* gene transcription is indirect and does not involve the LuxR regulatory circuit (Dunlap and Greenberg, 1991). It has been proposed that oxygen (Nealson, 1979) and iron (Haygood and Nealson, 1985a) limitation may function as a means for the host organism to limit growth of *V. fischeri* in the light organ while maintaining a high level of bacterial light production.

The *lux* operator is similar in sequence to the *E. coli* LexA binding sequence (Ulitzer, 1989) and binding of LexA to the operator has been demonstrated *in vitro* with DNA footprint analysis (Shadel et al., 1990a). However, there is no data demonstrating the involvement of the SOS response or LexA in the control of bioluminescence in *V. fischeri*. The alternate sigma factor mediating the heat shock response, σ^{32}, was thought to be involved in *lux* transcriptional regulation (Ulitzer and Kuhn, 1988). However, it was determined that σ^{32} acts indirectly since luminescence was dependent on the product of a σ^{32}-controlled gene, *groE* (Dolan and Greenberg, 1992). GroEL and GroER are chaperonins with a wide range of functions including: protein folding, oligomer assembly, and stabilization of unfolded proteins (Gething and Sambrook, 1992). It was concluded that GroEL and GroES are required for proper folding of LuxR into an active transcriptional regulator (Dolan and Greenberg, 1992).

Finally, it appears that the *V. fischeri* luminescence system can be regulated by cell density in the absence of autoinducer and LuxR (Dunlap and Kuo, 1992). Transcriptional *lux*::*lacZ* fusion mutants of *V. fischeri* were constructed to examine the effects of regulatory and environmental factors on *lux* gene transcription within *V. fischeri*. A mutant with a deletion of *luxR*, *luxICD*, and the intervening control region exhibited a 8- to 10-fold decrease in ß-galactosidase activity and rise back during the growth period, a pattern similar to that of autoinducer-LuxR controlled gene expression. Repression of bioluminescence by glucose in *V. fischeri* has not been linked to cellular cAMP levels. In these experiments it was shown that repression by both glucose and iron was mediated by a cAMP-independent and autoinducer-LuxR independent control mechanism. This mechanism may involve gene dosage modulation due to DNA replication, or changes in DNA superhelicity induced by cell density.

V. harveyi lux Gene Organization

The *V. harveyi lux* structural genes are organized in a similar manner to that of

operon R of *V. fischeri* in the order of *luxCDABE* (Miyamoto et al., 1988a). However, there is no analogous *luxI* or *luxR* gene located immediately upstream of *luxC*. Instead, the upstream region is rich in A + T residues and contains an open reading frame (ORF) located 600 bp from *luxC* (Miyamoto et al., 1988b). Although it is transcribed in the opposite direction and is situated in the same relative position as the *luxR* gene of *V. fischeri*, the ORF shares no sequence homology with *luxR* or any other known gene in the database.

V. harveyi lux Gene Regulation

Unlike the situation for *V. fischeri*, recombinant *E. coli* containing the *V. harveyi luxCDABE* genes and large regions of flanking DNA failed to produce cell-density dependant autoregulation of bioluminescence (Miyamoto et al, 1987). It was concluded that genes coding for regulatory proteins were not all present within this 18 kbp DNA fragment and must be located elsewhere on the *V. harveyi* genome (Miyamoto et al. 1988b). High levels of light production in *E. coli* can be obtained if the *luxCDABE* operon is placed under the strong T7 phage promoter and transcribed by T7 RNA polymerase (Miyamoto et al., 1988a), or when transformed into mutant *E. coli* (Miyamoto et al., 1987). These mutants, termed 43R and 43H, were isolated by their ability to produce high levels of light when transformed with the *V. harveyi lux* DNA. However, because the bioluminescence did not exhibit the same growth dependant regulation as seen for the parental *V. harveyi* strain, it was thought that the *E. coli* mutants bypassed the requirement for cell-density dependant *lux* gene regulation (Miyamoto et al., 1987). Attempts to reproduce cell-density dependant light production in recombinant *E. coli* containing the *V. harveyi lux* operon complemented with the *luxI* and *luxR* genes of *V. fischeri* failed, and Southern hybridization failed to identify any *luxI* or *luxR* homologs in *V. harveyi* (L. Cragg, unpublished results). This clearly demonstrates that the *V. harveyi lux* regulatory gene organization and function has diverged from that of *V. fischeri*.

A set of polycistronic mRNAs transcribed from the *V. harveyi lux* DNA was

identified by Northern blotting (Miyamoto et al., 1985; 1988a). The most abundant species originate just upstream of *luxD* and *luxA* and terminate downstream of *luxB*, reflecting the high levels of luciferase (LuxA and LuxB) and transferase (LuxD) in the cell. It is unclear whether these polycistronic mRNAs arise from internal initiation and termination sites or from specific mRNA processing. A nucleotide sequence capable of forming a stem loop structure has been located downstream of *luxB* (Baldwin et al., 1989) and may help stabilize the upstream mRNA from exonucleotide degradation.

As *E. coli* proved to be an unsuitable host to study *lux* gene regulation, a conjugative transfer system was developed to transfer cloned *lux* DNA with a reporter gene back into *V. harveyi* (Miyamoto et al., 1990). A region of DNA located approximately 500 bp upstream of *luxC* was able to promote transcription of the reporter gene (CAT) in a cell-density dependant manner and exhibited characteristics of being catabolite repressed. It was unclear, however, if this promoter region was under direct control of cAMP-CRP, or indirect via another control mechanism. A consensus CRP binding site has been located 455 bp upstream of *luxC* and 192 bp upstream of the ORF transcribed in the opposite direction (Miyamoto et al., 1988b). Transcription from the ORF promoter was not growth regulated nor glucose repressed, questioning the involvement of the putative CRP site and the ORF in *lux* gene regulation.

V. harveyi luxR Gene

Transposon mutagenesis was used to identify two unlinked regions of the *V. harveyi* genome required for bioluminescence (Martin et al., 1989). The first loci contained, as expected, the *luxCDABE* genes (region I), while the second loci (region II) was required for *luxCDABE* mRNA expression. Because these mutants were not defective in cAMP/CRP function or autoinducer synthesis, it was concluded that region II codes for a regulatory protein required for transcriptional activation of *luxCDABE*.

The regulatory gene located within region II was cloned and was found to encode a 23 kDa protein containing a Cro-like helix-turn-helix DNA-binding domain (Showalter et al., 1990). Although the gene product shares no identity with the LuxR protein of *V. fischeri* or any other sequence in the Genbank and EMBL data bases, this gene was named *luxR*. *In vivo* complementation studies in *E. coli* demonstrated that *luxR* is required for high expression of *luxCDABE*. However, because the light production did not exhibit the same growth dependant regulation as seen for *V. harveyi* and did not respond to exogenously added autoinducer, it was concluded that other regulatory factors were still required for autoinducer-controlled bioluminescence.

CHAPTER 2

DELINEATION OF THE TRANSCRIPTIONAL BOUNDARIES OF THE *LUX* OPERON OF *VIBRIO HARVEYI* DEMONSTRATES THE PRESENCE OF TWO NEW *LUX* GENES

©
A.F. Graham

PREFACE

The following chapter has been published as presented: Swartzman, E., C. Miyamoto, A. F. Graham, and E. A. Meighen (1990) *J. Biol. Chem.* **256**:3513-3517.

Carol Miyamoto determined the nucleotide sequence of the *luxG* and *luxH* genes, and Robert Soly provided the N-terminal amino acid sequence of *P. phosphoreum* LuxG.

ABSTRACT

The 5' and 3' ends of the *lux* mRNA of *Vibrio harveyi*, which extends over 8 kilobases, have been mapped, and two new genes *luxG* and *luxH*, were identified at the 3' end of the *lux* operon. Both S1 nuclease and primer extension mapping demonstrated that the start site for the *lux* mRNA was 28 bases before the initiation codon of the first gene, *luxC*. The promoter region contained a typical -10 but not a recognizable -35 consensus sequence. By using S1 nuclease mapping the mRNA was found to be induced in a cell density- and arginine-dependent manner. The DNA downstream of the five known *V. harveyi lux* genes, *luxCDABE*, was sequenced and found to contain coding regions for two new genes, designated as *luxG* and *luxH*, followed by a classical rho-independant termination signal for RNA polymerase. *luxG* codes for a protein of 233 amino acids with a molecular weight of 26,108, and *luxH* codes for a protein of 230 amino acids with a molecular weight of 25,326. The termination signal is active *in vivo* as demonstrated by 3' S1 nuclease mapping, confirming that the two genes are part of the *V. harveyi lux* operon. Comparison of the *luxG* amino acid sequence with coding regions immediately downstream from *luxE* in other luminescent bacteria has demonstrated that this gene may be a common component of the luminescent systems in different marine bacteria.

INTRODUCTION

The regulation of luminescence in marine bacteria has been the target of intense investigation over the last few years. Structural genes responsible for light production have been isolated from several strains of luminescent bacteria, including *Vibrio harveyi*, *Vibrio fischeri*, and *Photobacterium phosphoreum* (Meighen, 1988). There are five common *lux* structural genes; *luxC*, *luxD*, and *luxE* code for the reductase, transferase, and synthetase components, respectively, of a fatty acid reductase complex and *luxA* and *luxB* code for the α and β subunits of luciferase (Baldwin et al., 1984; Belas et al., 1982; Engebrecht et al., 1983; Engebrecht and Silverman, 1984; Boylan et al., 1989). The fatty acid reductase complex is responsible for producing an aldehyde substrate which, along with O_2 and $FMNH_2$, are necessary for the light-emitting reaction catalyzed by luciferase. An additional gene, *luxF*, has been found in the *P. phosphoreum lux* operon (Soly et al., 1988). Although its specific function is unknown, the protein exhibits 30% homology with the β subunit of luciferase.

The mode of regulation of only one bacterial *lux* system, that of *V. fischeri*, has been well documented (Engebrecht et al., 1983; Engebrecht and Silverman, 1986). There are two (left and right) operons involved in the *V. fischeri lux* system which are transcribed in opposite directions. The right operon contains the *luxI* gene, which is responsible for producing a small molecule (autoinducer) that causes induction of the luminescence system. This regulatory gene is followed by the five structural *lux* genes, *luxA-E* in the order *luxCDABE*. The left operon contains the *luxR* gene which encodes a protein that has been proposed to function as a receptor for the autoinducer. This complex then stimulates transcription of the right operon. A positive feedback loop is therefore established and autoinduction of the luminescent system is achieved in a cell-density dependent manner. The autoinducers of *V. fischeri* and *V. harveyi* have been purified and identified and have similar chemical structures (Eberhard et al., 1981; Cao and Meighen, 1989).

Analysis and expression of *V. harveyi lux* DNA has shown that the *luxA-E* genes

are arranged in the same order, *luxCDABE*, as in the *V. fischeri lux* system. There is no gene corresponding to the *luxI* gene of *V. fischeri* immediately upstream from the *V. harveyi* structural genes. Instead, the first open reading frame of greater than 40 codons is located more than 630 bases upstream from *luxC*. It has the same relative position and orientation as the *luxR* gene of *V. fischeri* (Miyamoto et al., 1988), but this upstream *V. harveyi* gene does not correspond in sequence to either regulatory *lux* gene of *V. fischeri* (Engebrecht and Silverman, 1987; Devine et al., 1988). The role, if any, of this gene in the regulation of the *lux* operon of *V. harveyi* has yet to be determined. Although the structural genes of the luminescent systems as well as the structures of the autoinducers of both bacteria are comparable, the mechanism of regulation and/or the organization of the *lux* regulatory genes are different. Transposon mutagenesis of *V. harveyi* resulting in Lux$^-$ phenotypes has demonstrated that two unlinked regions of the genome are essential for luminescence (Martin et al., 1989). Region I contains the *luxCDAB* and *E* genes while region II appears to have a regulatory function, suggesting that the regulatory genes of the luminescent operon of *V. harveyi* are not linked to the structural genes.

The present work defines the boundaries of the transcriptional unit of the *lux* operon of *V. harveyi* that extends over 8 kb and demonstrates that the *lux* mRNA is induced during development of luminescence. The 5' and 3' ends of the mRNA have been mapped and two new *lux* genes, *luxG* and *luxH*, have been located at the 3' end of the operon. Moreover, evidence has been obtained that *luxG* is found not only in *V. harveyi*, but also in *V. fischeri* and *P. phosphoreum*.

MATERIALS AND METHODS

Materials

Restriction enzymes were from Boehringer Mannheim or Pharmacia. S1 nuclease, avian myeloblastosis virus (AMV) reverse transcriptase, Klenow, T4 polynucleotide kinase and T4 DNA polymerase were purchased from Pharmacia. ^{32}P-Labelled nucleotide and deoxynucleotide triphosphates (>3000 Ci/mmol) were from ICN Biomedicals Canada Ltd. and [^{35}S]AMPτS (1400 Ci/mmol) was obtained from Dupont-New England Nuclear.

Growth Conditions and RNA Extraction

V. harveyi was grown in 1% NaCl complex or minimal medium (Reindeau and Meighen, 1979) with or without added arginine (1.0 mg/ml). RNA was extracted as previously described (Miyamoto et al., 1985) from uninduced cultures of *V. harveyi* at OD_{660} = 0.3, and from induced cultures at OD_{660} = 1.5. RNA was extracted from *V. harveyi* grown in minimal medium (± arginine) at the point of maximum light emission at OD_{660} = 0.3. In minimal medium the cells achieve maximum light emission at a lower OD_{660} value than that reached for cells grown in complex medium.

S1 Nuclease Mapping

A modified procedure based on the one developed by Berk and Sharp (1978) was used to map the 5' end of the *lux* mRNA. A *Sac*I-*Bam*HI 0.9 kbp dsDNA fragment extending into *luxC*, and 5' labelled at the *Bam*HI site with T4 polynucleotide kinase and [τ-^{32}P]ATP (0.02 pmol, 4.4x10^6 cpm/pmol), was coprecipitated with 50 ug of RNA. The pellet was resuspended in hybridization buffer (80% formamide, 40 mM PIPES, pH 6.6, 0.4 M NaCl, 1 mM EDTA), heated to 85° for 10 min, then slow-cooled to 50° for 12-16 h. The hybrid mixture was then diluted to a final volume of 0.3 ml with 30 ul 10x S1 nuclease buffer (0.3 M NaOAc, pH 4.6, 1 M NaCl, 10 mM $Zn(OAc)_2$, 50% glycerol) and H_20. S1 nuclease reactions were carried out for 30 min

at 37° with the indicated amounts of enzyme. Reaction products were phenol-extracted, precipitated with ethanol, and resuspended in an 80% formamide, 10 mM NaOH-dye solution. Samples were then heat-denatured and loaded onto a 6% acrylamide-7 M urea sequencing gel. Samples were ran at a sufficiently high temperature (50-55°) so that all secondary structure was eliminated in the urea gel. Under these conditions, the electrophoretic mobility of ssDNA is independent of base composition in a denaturing gel (Maniatis and Efstratiadis, 1980). Consequently, a sequence ladder of M13mp18 served as size standards.

To map the 3' end of the *lux* mRNA the protocol just described was followed except that the probe used was a *Bam*HI-*Sac*I 0.7 kbp dsDNA fragment located about 1.5 kbp downstream from *luxE* and 3' labelled at the *Bam*HI site with the Klenow fragment of *E. coli* DNA polymerase and [α-^{32}P]dCTP (0.04 pmol, 8.8x10^6 cpm/pmol). The reaction products were electrophoresed on an 8% acrylamide-7 M urea sequencing gel. The 3'-end of the mRNA transcribed from the opposite strand and terminating near the 3'-end of the *lux* mRNA was investigated using the same *Bam*HI-*Sac*I dsDNA fragment, except that it was 3' labelled at the *Sac*I site with T4 DNA polymerase.

Primer Extension Mapping

The 5' end of the mRNA was confirmed using primer extension analysis under conditions described by Lagacé et al. (1987). The probe, a 100 base *Cla*I-*Ssp*I ssDNA fragment 5' labelled at the *Ssp*I site with T4 polynucleotide kinase and [τ-^{32}P]ATP (0.1 pmol, 7.8x10^5 cpm/pmol), was sealed in a glass microcapillary tube with RNA in a total volume of 10 ul containing 10 mM PIPES, pH 6.6, and 0.4 M NaCl. Hybridization was carried out at 55° for 12-16 h. The primer was extended with 15 units of AMV reverse transcriptase at 42° for 60 min in a total volume of 100 ul containing 50 mM Tris, pH 8.2, 10 mM DTT, 6 mM $MgCl_2$, and 0.5 mM of the four dNTPs. The reaction products were phenol-extracted, ethanol-precipitated, resuspended in an 80% formamide, 10 mM NaOH-dye solution, heat denatured and

resolved on an 8% acrylamide-7 M urea sequencing gel.

DNA Sequencing

The sequence of the DNA found downstream from *luxE* was obtained using the dideoxy chain termination method of Sanger (Sanger et al., 1977) according to procedures previously described (Miyamoto et al., 1988a). The cloning strategy for the downstream DNA is given in the text. Analysis of data was performed using the DNASIS and PROSIS programs of Hitachi Software Engineering Co., Ltd.

RESULTS

A set of polycistronic mRNAs that codes for the *lux* proteins has been identified in *V. harveyi* (Miyamoto et al., 1985; 1988a), and it has been shown that the mRNAs detected by Northern blot analysis extending across *luxD*, *luxA*, *luxB* and *luxE* are induced during development of luminescence. However, the mRNAs starting at *luxC* are not readily detected and it could not be determined whether they are induced. Moreover, the exact upstream and downstream termini of the mRNA have not yet been determined.

Localization of the 5' end of the lux mRNA

To elucidate the 5' terminus of the *lux* mRNA, total mRNA was isolated from *V. harveyi*, hybridized to a DNA probe encompassing the 5' region of the *luxC* gene and treated with S1 nuclease. The probe used was a *Sac*I-*Bam*HI restriction fragment (see Fig. 1), 5' ^{32}P-labelled at the *Bam*HI site. As shown in Fig. 2, only mRNA isolated from *V. harveyi* cells after induction of luminescence partially protected this fragment, demonstrating that the 5' terminus of the induced *lux* mRNA occurs between the *Sac*I and *Bam*HI sites. Since arginine is known to augment luminescence in minimal medium (Nealson et al., 1970), the S1 nuclease patterns of mRNA obtained from *V. harveyi* grown in the presence and absence of arginine in minimal medium were also investigated. Only mRNA from cells grown in minimal medium containing arginine protected the DNA probe, and the same sized fragment was obtained as during induction in complex medium. The partially protected DNA fragment was sized on a sequencing gel to accurately determine the 5' end of the mRNA. The S1 nuclease reactions were carried out with 20, 50, 100 and 300 units of enzyme (Fig. 3a). The sharpest band can be seen when 50 units of S1 nuclease is used (lane 2, Fig. 3a), placing the start of the message 16 nucleotides before the initiation codon of the *luxC* gene. However, start sites ranging from 11 to 28 nucleotides in front of *luxC* can be measured depending on the amounts of S1 nuclease used. With 20 units of S1 nuclease, a minor band (arrow, Fig. 3a) can be

Figure 1: Restriction map of *V. harveyi lux* DNA. The placement of the *lux* genes is depicted below the restriction map. The probes used to determine the 5' and 3' ends of the mRNA are as follows: (a) *Sac*I-*Bam*HI 998 bp dsDNA fragment, 5' ^{32}P-labelled at the *Bam*HI site, used for 5' S1 nuclease mapping. (b) *Cla*I-*Ssp*I 100 base ssDNA fragment, starting at the first *Cla*I site, 5' ^{32}P-labelled at the *Ssp*I site for primer extension mapping. (c) *Bam*HI-*Sac*I 709 bp dsDNA fragment, 3' ^{32}P-labelled at the *Bam*HI or *Sac*I site for 3' S1 nuclease mapping. B, *Bam*HI; S, *Sac*I; C, *Cla*I; Ss, *Ssp*I.

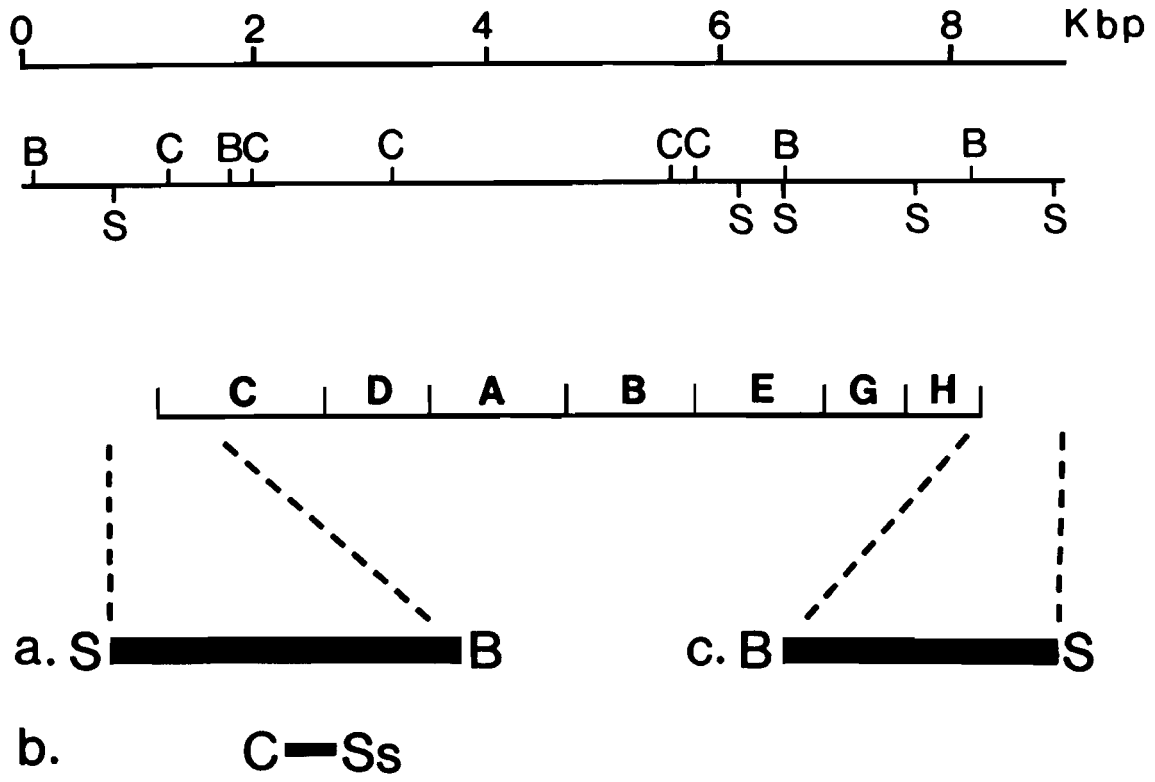

0
2
4
6
8
Kbp
B
C
BC
C
CC
B
B
S
S
S
S
S
C
D
A
B
E
G
H
a. S
B
c. B
S
b.
C
Ss

Figure 2: Induction of *V. harveyi lux* mRNA. 5' S1 nuclease mapping was performed with RNA extracted from *V. harveyi* grown under various conditions using a *Sac*I-*Bam*H1 dsDNA fragment 5' ^{32}P-labelled at the *Bam*HI site as the probe and 50 units of S1 nuclease. Lane 1, *V. harveyi* grown in complex medium and harvested before induction of luminescence; lane 2, *V. harveyi* grown in complex medium and harvested after induction of luminescence; lane 3, *V. harveyi* grown in minimal medium; lane 4, *V. harveyi* grown in minimal medium +1.0 mg/ml arginine. No S1 nuclease products could be detected when RNA isolated from *E. coli* was hybridized to the same probe, as shown in lane 5. The DNA sequence of M13mp18 in lanes marked A, C, G, and T is given as a size standard.

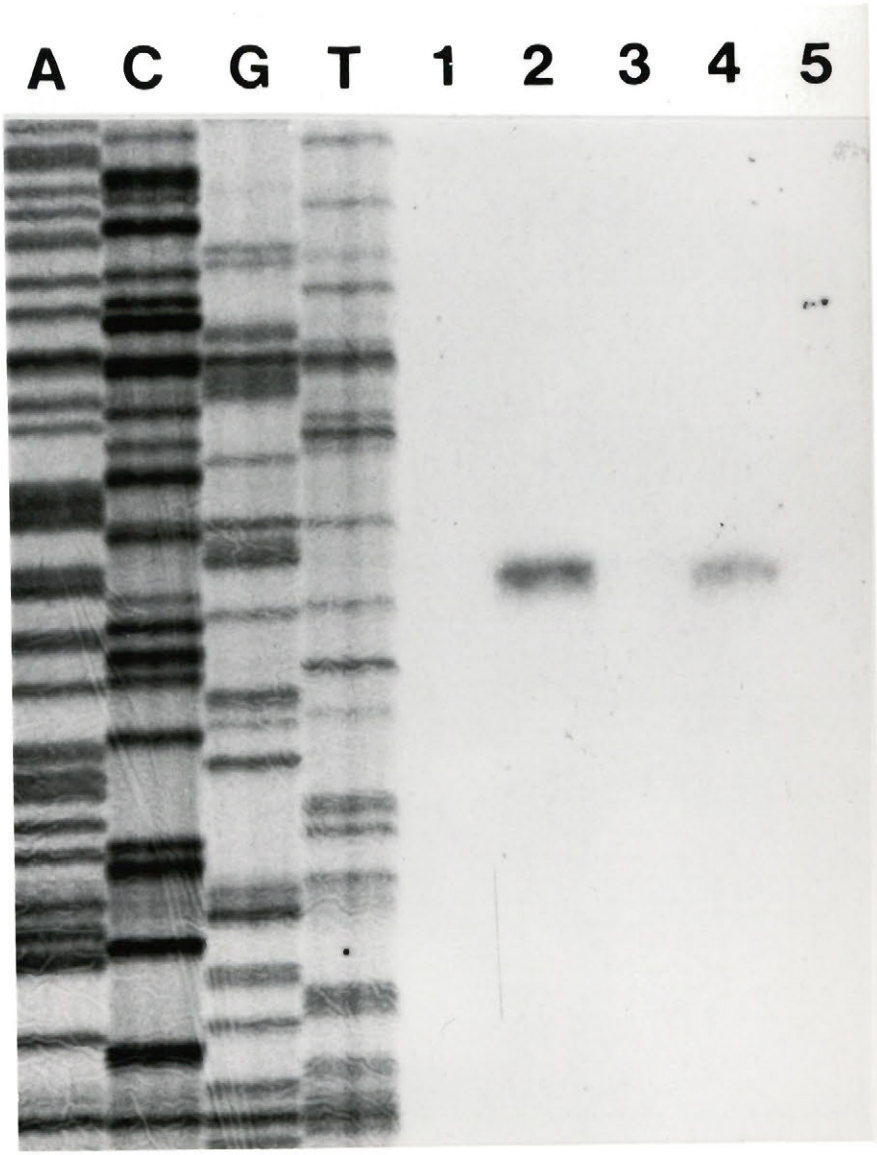
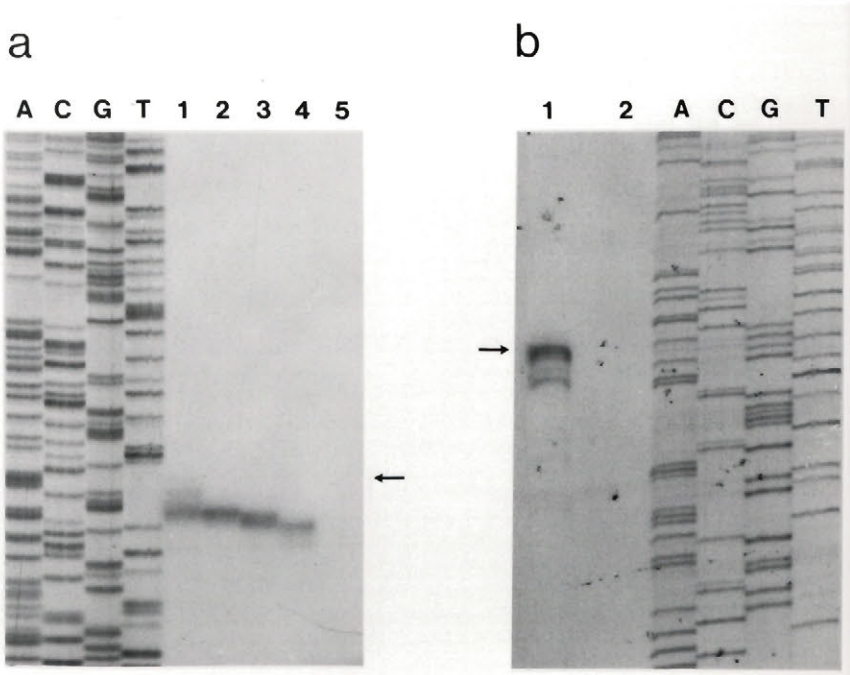

A C G T 1 2 3 4 5

Figure 3: Mapping the 5' end of the *lux* mRNA. (a) 5' S1 nuclease mapping of *V. harveyi* RNA using a *Sac*I-*Bam*HI dsDNA fragment 5' ^{32}P-labelled at the *Bam*HI site as the probe. The amounts of S1 nuclease used are as follows: lane 1, 20 units; lane 2, 50 units; lane 3, 100 units; lane 4, 300 units. Lane 5 is the control with *E. coli* RNA, using 50 units of S1 nuclease. (b) Primer extension mapping of *V. harveyi* RNA using a *Cla*I-*Ssp*I ssDNA fragment 5' ^{32}P-labelled at the *Ssp*I site as the primer. Lane 1 contains *V. harveyi* RNA; lane 2 is the control with *E. coli* RNA. The DNA sequence reactions (A, C, G and T) of M13mp18 served as size standards. The arrows indicate where the longest reaction products are found. (c) Nucleotide sequence of the DNA found upstream of *lux*C as reported by Miyamoto et al. (1988b). The first nucleotide of the mRNA is located at position +1 as indicated by the arrow. The proposed -10 promoter and Shine-Dalgarno sequences are underlined.



c

```
          -60                    -40                    -20
TTGAACTGTTTAATGTATTTGGTTAAAAGTTTTTAATTAACTTTAAAAAAATGATCC

                +1                     +20
AAGGAATTAATGTTTTCCAAAATTTAAAAGAGAAGCTTGAT ATG GAA AAA CAC
      -10                                 Met Glu Lys His
                                          ---------->
                                            luxC
```

seen corresponding to a start site 28 nucleotides in front of *luxC* (position +1, Fig. 3c). Because of the somewhat ambiguous nature of the S1 nuclease mapping results, the 5' end of the mRNA was also determined by primer extension mapping (Fig. 3b). A 100 base ssDNA fragment (*Cla*I-*Ssp*I) ^{32}P-labelled at the 5' end was hybridized to RNA, and was extended with reverse transcriptase. The 100 base primer migrated as predicted with respect to the M13mp18 sequence ladder (data not shown), while the primer extension products migrated as a doublet with sizes of 184 and 185 bases. From the size of the larger fragment (Fig. 3b), the 5' end of the message could be assigned to nucleotide +1 (Fig. 3c) in agreement with the results obtained by S1 nuclease mapping using the lowest amount of enzyme. It is interesting to note that an increase in S1 nuclease concentration results in the removal of the AT-rich 5' end by as much as 13 nucleotides. Upon examining the DNA sequence, a putative -10 promoter sequence can be recognized just upstream from the message start site (Fig.3c). There does not, however, appear to be any recognizable -35 promoter consensus sequence.

Nucleotide sequence of two new genes found within the lux operon

Since the mRNAs extend 2-3 kb downstream from the *luxE* gene, the last known gene of the *lux* operon, the downstream region was sequenced (using the strategy outlined in Fig. 4) in order to determine the specific 3' terminator site and whether or not other *lux* genes are encoded in this area. The DNA sequence downstream from *luxE* was found to contain two previously unrecognized genes, *luxG* and *luxH*, transcribed in the same direction as *luxE* (Fig. 5). The first gene, *luxG*, starts just one nucleotide after the stop codon of the *luxE* gene and consequently the Shine-Dalgarno (Shine and Dalgarno, 1974) sequence of *luxG* resides in the 3' end of the coding region of the *luxE* gene. This gene codes for a protein of 233 amino acids with a molecular weight of 26,108. The second gene, *luxH*, starts 22 nucleotides after *luxG* and codes for a protein of 230 amino acids with a molecular weight of 25,326. The first 150 codons of *luxG* were compared to an open reading frame found just

Figure 4: Cloning strategy for the sequence of DNA found downstream of *luxE*. The open reading frames of *luxG* and *luxH* are located after *luxE* as indicated. The arrows above and below the restriction map represent the direction and extent of the DNA sequences read from each clone. The DNA encompassing the 3' end of *luxE* and the *luxG* and *luxH* genes was sequenced in both directions from overlapping clones. B, *Bam*HI; S, *Sac*I; H, *Hin*dIII; G, *Bgl*III.

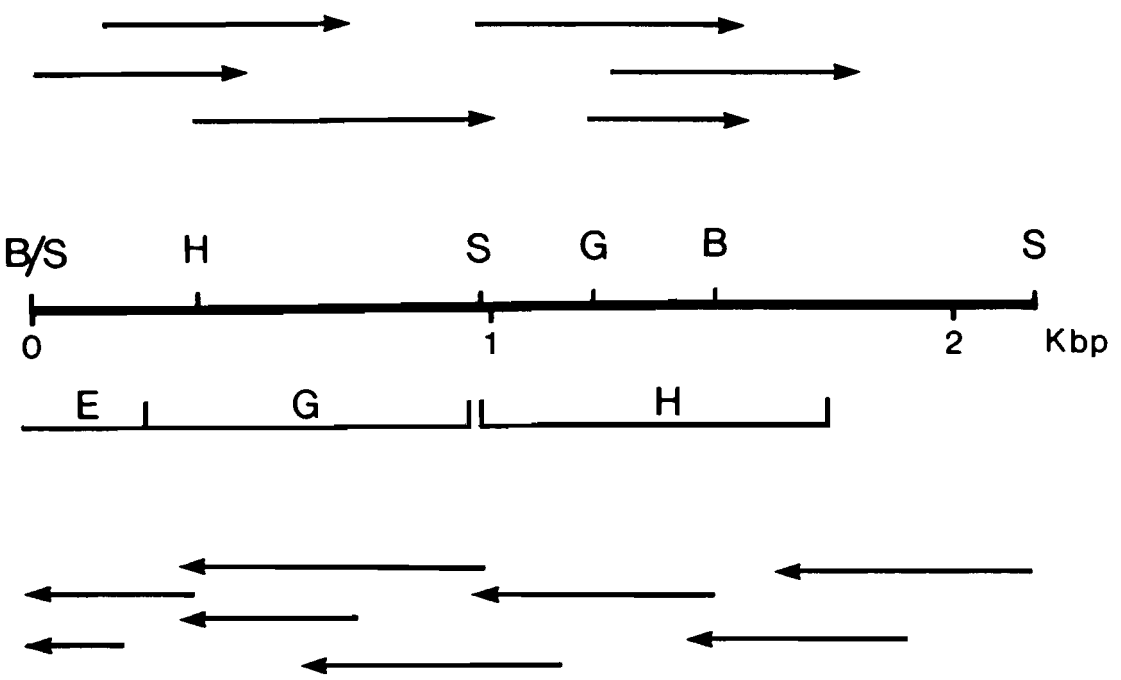
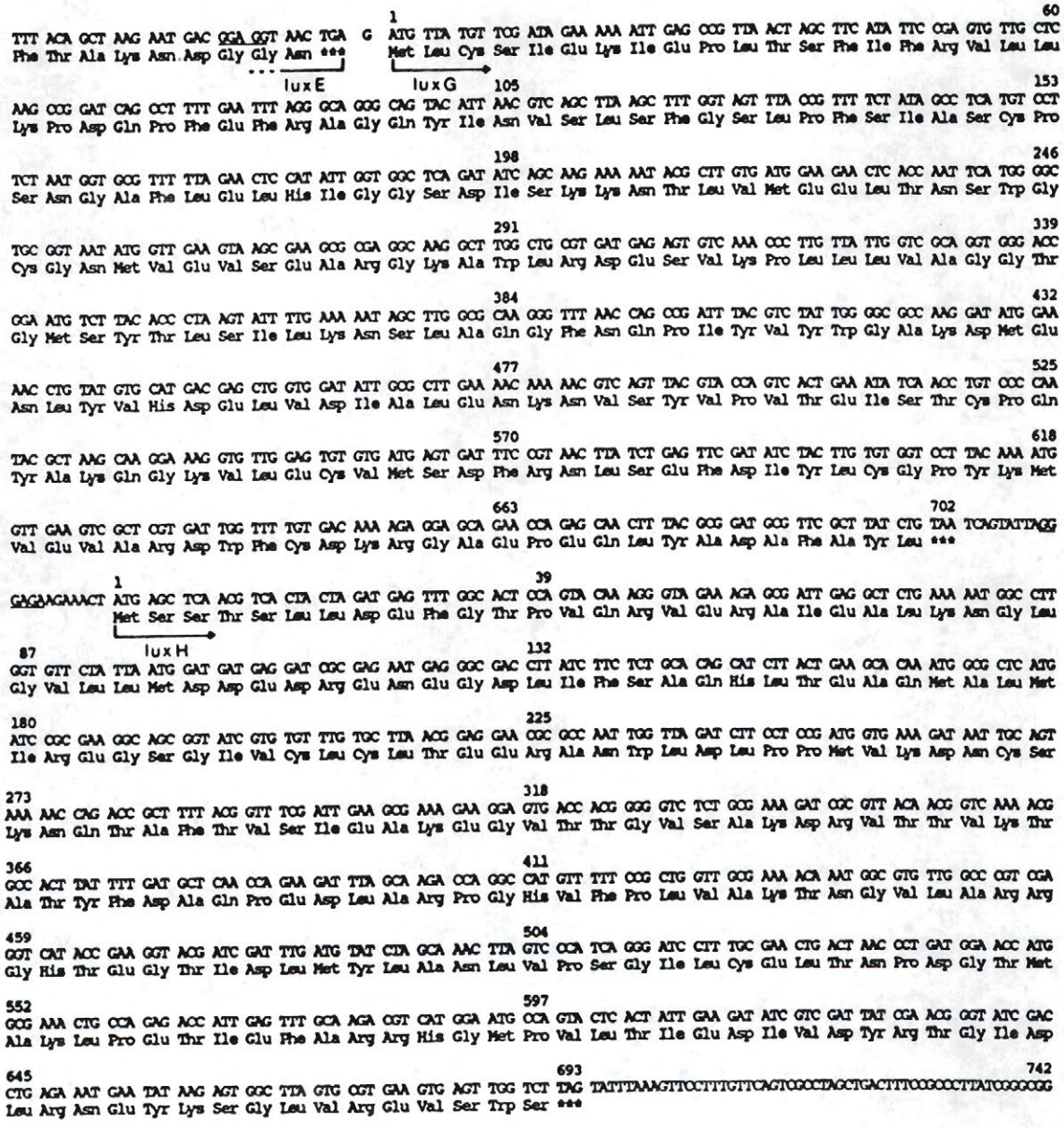

B/S
H
S
G
B
S
0
1
2
Kbp
E
G
H

Figure 5: Nucleotide and accompanying amino acid sequence of the *luxG* and *luxH* genes of *V. harveyi*. The first nucleotide of each gene is designated as base 1. *luxG* starts 1 nucleotide after the stop codon of *luxE*, and *luxH* starts 22 nucleotides after *luxG*. The proposed Shine-Dalgarno sequences of *luxG* and *luxH* are underlined, with the Shine-Dalgarno sequence of *luxG* residing within the coding region of *luxE*.


luxE
luxG
luxH

downstream from the *luxE* gene of a different genera of luminescent bacteria, *P. phosphoreum* (Soly, R., unpublished results) (fig. 6). A 39% homology found between the amino acid sequences demonstrates that *luxG* is common to both *V. harveyi* and *P. phosphoreum*.

3' End of the lux operon as determined by nucleotide sequencing and S1 nuclease mapping

Analysis of this DNA resulted in the recognition of a putative rho-independent termination signal for RNA polymerase (Platt, 1986) located about 40 nucleotides downstream from the stop codon of *luxH* (Fig. 7c). It contains a classical GC rich hairpin loop, followed by a string of T's. In order to determine if this signal is active *in vivo*, the 3' end of the *lux* mRNA was mapped using S1 nuclease. A DNA probe ^{32}P-labelled at the 3' end was hybridized to total RNA isolated from *V. harveyi* cells. The S1 nuclease reactions and subsequent treatment of the protected fragments were performed as before for the 5' end of the mRNA. It is of interest to note that in order to accurately determine the size of the S1 nuclease reaction products, it was necessary to run the sequencing gel at a very high temperature. The DNA ran anomalously at lower temperatures possibly due to the secondary structure of the hairpin loop in the probe. The same size product was obtained in S1 nuclease reactions containing 20, 50, and 100 units of enzyme (Fig. 7a) indicating that S1 nuclease was not removing residues from the duplex at the 3' end as was observed for the 5' end. According to S1 nuclease mapping, the mRNA terminates after the hairpin structure, at one of two nucleotides within the run of T's (Fig. 7c). When the probe was hybridized to mRNA isolated from *V. harveyi* cells before and after induction of luminescence (Fig. 7b) more *lux* mRNA could be seen in the cultures after induction of luminescence, consistent with the 5' S1 nuclease mapping results.

Just downstream from the termination signal for the *lux* operon is found an analogous, rho-independent termination signal in the complementary strand (Fig. 7c). The 3' termination of an open reading frame that extends at least 400 bp occurs 19

Figure 6: Comparison of *V. harveyi* and *P. phosphoreum* **LuxG amino acid sequences.** The first 150 and 151 amino acids of LuxG from *V. harveyi* and *P. phosphoreum* , respectively, were compared and found to contain 39% homology. The numbered positions refers to the *V. harveyi* LuxG sequence. Maximum homology was obtained by inserting a space (-) within the LuxG sequence of *P. phosphoreum*.

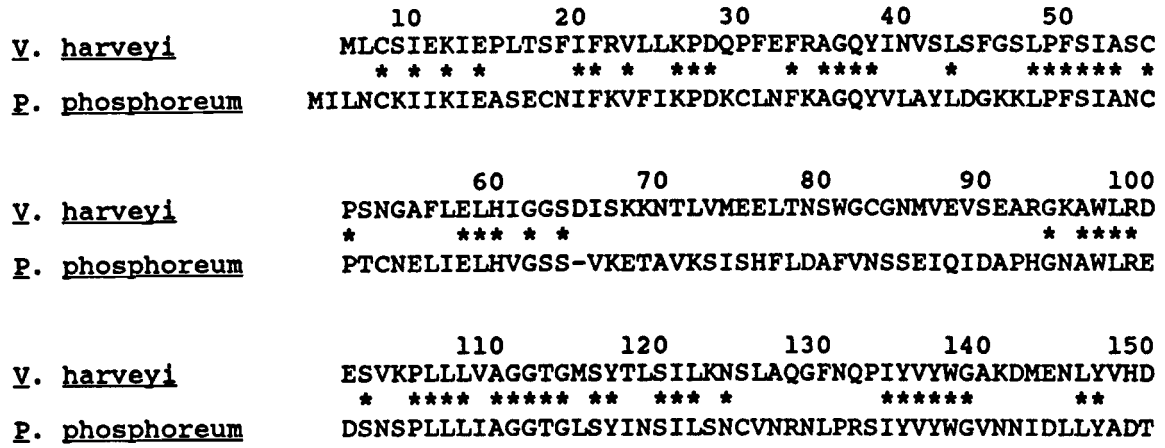
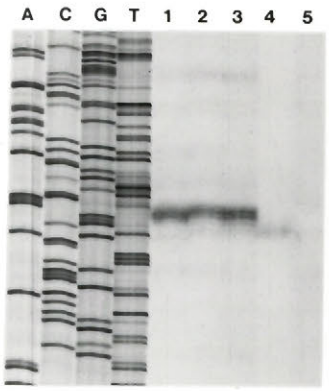

10 20 30 40 50
V. harveyi MLCSIEKIEPLTSFIFRVLLKPDQPFEFRAGQYINVSLSFGSLPFSIASC
P. phosphoreum MILNCKIIKIEASECNIFKVFIKPDKCLNFKAGQYVLAYLDGKKLPFSIANC
60 70 80 90 100
V. harveyi PSNGAFLELHIGGSDISKKNTLVMEELTNSWGCGNMVEVSEARGKAWLRD
P. phosphoreum PTCNELIELHVGSS-VKETAVKSISHFLDAFVNSSEIQIDAPHGNAWLRE
110 120 130 140 150
V. harveyi ESVKPLLLVAGGTGMSYTLSILKNSLAQGFNQPIYVYWGAKDMENLYVHD
P. phosphoreum DSNSPLLLIAGGTGLSYINSILSNCVNRNLPRSIYVYWGVNNIDLLYADT

Figure 7: Mapping the 3' end of the *lux* mRNA. (a) 3' S1 nuclease mapping of *V. harveyi* RNA using a *Bam*HI-*Sac*I dsDNA fragment 3' ^{32}P-labelled at the *Bam*HI site as the probe. The amounts of S1 nuclease used are as follows: lane 1, 20 units; lane 2, 50 units; lane 3, 100 units; lane 4, 300 units. Lane 5 is the control with *E. coli* RNA using 50 units of S1 nuclease. The DNA sequence reactions (A, C, G, and T) of M13mp18 served as size standards. (b) 3' S1 nuclease mapping of *V. harveyi* RNA isolated from cells before (lane 1) and after (lane 2) induction of luminescence was performed using the same *Bam*HI-*Sac*I dsDNA probe and 50 units of S1 nuclease. Lane 3 is the control using *E. coli* RNA. The arrow indicates the protected fragment obtained after S1 nuclease digestion. (c) Nucleotide sequence found downstream of *luxH*. Two rho-independent termination signals for RNA polymerase are found on opposite strands and converge end-on-end. The first one, starting 38 nucleotides after the coding region of *luxH*, has an energy of -17.6 kcal. The 3' end of the *lux* mRNA terminates at one of two nucleotides within the run of T's found after the hairpin loop, as indicated by the arrows. The second termination signal found in the opposite strand has an energy of -11 kcal and is located 19 nucleotides after the end of an open reading frame that extends at least 400 bp to the right. Nucleotides are numbered starting from the initiation codon of *luxH* as given in fig. 5.

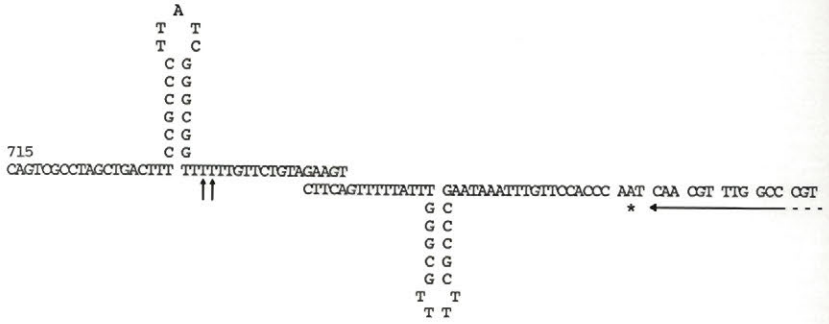
a


A C G T 1 2 3 4 5

b


1 2 3

c


715
CAGTCGCCTAGCTGACTTT TTTTTTGTTCTGTAGAAGT
CTTCAGTTTTTATTT GAATAAATTTGTTCCACCC AAT CAA CGT TTG GCC CGT

nucleotides upstream from this signal. When a probe (*Bam*HI-*Sac*I, 3'-labelled at the *Sac*I site) was hybridized to RNA isolated from uninduced and induced cultures of *V. harveyi*, no message could be detected (data not shown). It appears then that the mRNA corresponding to this gene is in low abundance and is not coordinately induced with light production, indicating that this open reading frame is not likely to code for a *lux* gene involved in the luminescent system. The presence of the two termination signals in opposite strands with converging coding regions at the end of *luxH*, along with *in vivo* verification of the 3' end confirms that the end of *luxH* is the 3' terminus of the *lux* operon.

DISCUSSION

In this paper, the transcriptional endpoints of the mRNA from the *lux* operon of *V. harveyi* have been defined and two new genes encoded by the *lux* mRNA have been identified. S1 nuclease and primer extension mapping were the two techniques used to map the 5' end of the mRNA, and gave identical results providing that the amount of S1 nuclease was carefully controlled. Just upstream from the startpoint a -10 but no corresponding -35 recognition sequence for RNA polymerase could be found. This may suggest that a regulatory protein is required for proper transcription initiation by the RNA polymerase. The promoter region for the right operon of the *V. fischeri* luminescent system also lacks a -35 consensus sequence and is believed to require a positive regulator for transcription (Engebrecht and Silverman, 1987).

Previous studies using Northern blots have shown that polycistronic messages of varying lengths exist for the *V. harveyi lux* operon (Miyamoto et al., 1985; 1988a). While it is clear that the mRNAs extending across the *luxCDABE* genes and downstream DNA were induced, those starting at the *luxC* gene were not readily detected. Nor could it be determined whether or not they were induced during the development of luminescence. By application of S1 nuclease mapping in these experiments, it has been possible to show that the mRNA originating at the *luxC* gene is indeed induced, consistent with the synthesis of all the proteins within this operon being coregulated during induction of light emission. Similarly, S1 nuclease mapping has shown that arginine causes an increase in the *lux* mRNA level in *V. harveyi* grown in minimal medium, indicating that arginine acts to stimulate luminescence at the transcriptional level.

The DNA located downstream of the *luxE* gene was sequenced in an effort to understand why the mRNA extends beyond the last known gene of the *lux* operon. Two new genes were found, designated as *luxG* and *luxH*. Downstream from *luxH*, two classical rho-independent termination signals for RNA polymerase on opposite strands and separated by less than 30 bp could easily be identified. The termination signal for the *lux* mRNA, which has an energy of -17.6 kcal, was confirmed *in vivo*

using 3' S1 nuclease mapping. The other termination signal, with an energy of -11 kcal, is located just after the end of a convergent open reading frame coding for a protein of unknown function. The next best candidate that could exhibit a hairpin loop structure in the downstream region after *luxE* has an energy of only -6.9 kcal and lies within the *luxH* gene. The presence of the termination signal immediately after *luxH* along with the induction of the corresponding mRNA with light production provide strong evidence that the *luxG* and *luxH* genes are part of the *lux* operon.

Elucidation of the functions of the *luxG* and *luxH* genes may provide a key to understanding the role of luminescence in bacteria. Homologies between the proteins coded by *luxG* and *luxH* and sequences of proteins of known functions have not yet been detected. The presence of these genes in free living bacteria is not essential for light production since clones containing only the *luxCDABE* genes are able to emit light (Miyamoto et al., 1985; 1987) and transposon mutagenesis has failed to produce any Lux$^-$ phenotypes with insertions in these genes (Martin et al., 1989). It is possible that *luxG* and *luxH* are regulatory proteins, but this is difficult to test because *E. coli* is unable to support regulated light generated by *V. harveyi* DNA (Miyamoto et al., 1985). Preliminary sequence data (unpublished results) have shown that *V. fischeri* contains a homologous gene to *luxG* at the same relative position in the operon. This result suggests that the *luxG* gene product is not essential for regulation by autoinduction, since clones containing only the *luxCDABE* and regulatory genes of *V. fischeri* are able to produce regulated light in *E. coli* (Engebrecht et al., 1983). Transposon mutagenesis of *V. harveyi* (Martin et al., 1989) has provided further evidence that *luxG* and *luxH* are not required for regulation since disruption of the transcription of downstream genes does not affect induction of the reporter gene, ß-galactosidase. Comparison of the first 150 amino acids of *luxG* to an amino acid sequence found just downstream from *luxE* in *P. phosphoreum* demonstrated 39% identity. *luxG* is therefore common to and located in the same relative position in the *lux* operons of *V. harveyi*, *V. fischeri* and *P. phosphoreum*. It is possible that *luxG* and *luxH* may produce proteins that fine-tune the expression or properties of the light-

emitting reaction without affecting induction of light production. Alternatively, it is possible that the downstream genes are required for an essential function relating the *lux* system to the survival and/or symbiosis of luminescent bacteria in the marine environment.

CHAPTER 3

A NEW *VIBRIO FISCHERI LUX* GENE PRECEDES A BIDIRECTIONAL TERMINATION SITE FOR THE *LUX* OPERON

©
S. Kapoor

PREFACE

The following chapter has been published as presented: Swartzman, E., S. Kapoor, A. F. Graham, and E. A. Meighen. (1990) *J. Bacteriol.* **172**:6797-6802.

Shalini Kapoor, Carol Miyamoto and Rose Szittner determined the sequence of the *luxG* gene.

ABSTRACT

The DNA downstream of the *lux* structural genes in the *Vibrio fischeri lux* operon has been sequenced and a new *lux* gene (*luxG*) identified. A hairpin loop that begins with a poly A region and ends with a poly T region and thus can function as a bidirectional termination site for *luxG* and a convergent gene is located immediately downstream of *luxG*. 3' S1 nuclease mapping has demonstrated that the *luxG* mRNA was induced in a cell density dependent fashion consistent with it being part of the *lux* system and that the *lux* mRNA terminated immediately after the hairpin loop. The mRNA coded by an open reading frame convergent to *luxG* on the complementary strand was also shown by S1 nuclease mapping to overlap the *lux* mRNA for at least 20 nucleotides before termination. Expression of DNA containing the hairpin loop, placed between a strong promoter and a reporter gene and transferred by conjugation into luminescent bacteria, demonstrated the very high efficiency of termination by this hairpin loop orientated in either direction. These results also demonstrate that the organization of the genes at the 3' ends of the *lux* operons of *V. fischeri* and *V. harveyi* has clearly diverged.

INTRODUCTION

The *V. fischeri* luminescence system is currently the best characterized of the bacterial *lux* systems. Two operons encompassing seven *lux* genes (*A-E, I, R*) are sufficient for expression and regulation of luminescence on transformation into *E. coli* (Engebrecht et al., 1983; Engebrecht and Silverman, 1984). The right operon contains the *lux* structural genes in the order *luxCDABE* preceded by the regulatory gene, *luxI*. The left operon, immediately upstream of *luxI*, contains a second regulatory gene, *luxR*, transcribed in the opposite direction. It has been proposed that *luxI* specifies a product responsible for synthesis of an autoinducer that on binding with a receptor produced by *luxR* is responsible for the growth-dependent induction of luminescence (Engebrecht and Silverman, 1986). The *V. fischeri lux* structural genes are composed of two genes(*luxA* and *B*) coding for the α and β subunits of luciferase and three genes (*luxC, D* and *E*) that code for polypeptides that catalyze the formation of the aldehyde substrate for the luminescent reaction (Boylan et al., 1985).

Transposon mutagenesis of the *V. fischeri lux* system has shown that all transposon insertions that block luminescence were located within the two regulatory and five structural *lux* genes (Engebrecht et al., 1983, Engebrecht and Silverman, 1984). Based on this genetic analyses of the *V. fischeri* luminescence system, it would appear that the complete *lux* system had been defined.

The organization of the *lux* systems from *V. harveyi* and *P. phosphoreum* have at least partially been determined (Miyamoto et al., 1988a, Mancini et al., 1988). In contrast to the *V. fischeri lux* system, regulatory genes are not located immediately upstream of the structural genes in the *V. harveyi lux* system. Instead, the DNA upstream is very rich in AT residues and contains numerous stop codons in all reading frames extending for over 600 nucleotides (Miyamoto et al, 1988b). The order of the *lux* structural genes is the same in these two systems as in the *V. fischeri lux* system except the *P. phosphoreum lux* system contains an extra gene, *luxF*, that is located between *luxB* and *luxE*. The *luxF* gene, which has some homology to the

luxA and *luxB* genes, codes for a flavoprotein of unknown function (Soly et al., 1988).

Recently, two new genes, *luxG* an *luxH*, have been located immediately downstream of *luxE* in the *V. harveyi lux* system and shown to be within the *lux* transcriptional unit (Swartzman et al., 1990b). The 5' terminus of a gene homologous to *luxG* was also identified immediately downstream of *luxE* in the *P. phosphoreum lux* system. No structural or regulatory role has yet been found for these two new *lux* genes.

As the *lux* system of *V. fischeri* has been well-defined by genetic and biochemical analyses in terms of both expression and regulation, it was of particular interest to determine whether *luxG*, *luxH*, or other *lux* genes may be downstream of *luxE* in the *V. fischeri lux* system or whether the downstream *lux* genes in the other systems may be unique to particular luminescent bacteria. The results have demonstrated that the *V. fischeri lux* system contains one new downstream *lux* gene, *luxG*, and that a strong bidirectional termination site is located immediately after this gene at the end of the *lux* operon.

METHODS AND MATERIALS

Materials

Restriction enzymes, S1 nuclease, T4 DNA polymerase and the large fragment of DNA polymerase (Klenow) were purchased from Pharmacia. ^{32}P-Labelled deoxynucleotide triphosphates (>3000 Ci/mmol) were from ICN Biomedicals Canada, [^{35}S]dATP (1400 Ci/mmol)was from Du Pont - New England Nuclear and ^{14}C-chloramphenicol (57 mCi/mmol) was obtained from Amersham.

Growth Conditions and RNA Extraction

V. fischeri ATCC 7744 was grown in 3% NaCl complex medium (Boylan et al., 1985) at 20°C and *E. coli* RR1 transformed with pVfB1 (Miyamoto et al., 1990) was grown at 20°C in LB medium supplemented with 100 ug of ampicillin per ml. RNA was extracted as previously described (Swartzman et al., 1990).

3' S1 Nuclease Mapping

S1 nuclease mapping was performed according to procedures previously described (Swartzman et al., 1990b) except that 100 ug of RNA was used in the hybridizations. A *Sau*96I-*Hin*dIII 0.65 kbp dsDNA fragment located approximately 0.6 kbp downstream from the *luxE* gene was used as the probe to map the end of the *V. fischeri* right *lux* operon. It was 3'-labeled at the *Sau*96I site with the Klenow fragment of *E. coli* DNA polymerase and [α-^{32}P]dCTP (0.02 pmol, 1x10^6 cpm/pmol). To determine the 3' end of the convergent gene a 0.4 kbp *Apa*I-*Ssp*I dsDNA fragment 3' labeled at the *Ssp*I site with T4 DNA polymerase and [α-^{32}P]dATP was used as the probe(0.03 pmol, 2.4x10^5 cpm/pmol).

DNA Sequencing

The sequence of the DNA found downstream of *luxE* was obtained using the dideoxy-chain termination method according to procedures described previously (Miyamoto et al., 1988b). The DNASIS and PROSIS programs of Hitachi Software Engineering Co. Ltd. were used to analyze sequence data.

In vivo Termination Assay

The strength of the bidirectional terminator was determined by using an *in vivo* gene expression assay developed by Miyamoto et al. (1990). A 700 bp *Sau*96I-*Hin*dIII DNA fragment containing the inverted repeat was blunt ended with the Klenow fragment of *E. coli* and inserted in both directions into the filled-in *Hin*dIII site of pMGM110 (Miyamoto et al., 1990) between the *V. harveyi lux* promoter and the CAT gene and transferred by conjugation into *V. harveyi*. As a control, a 1.3 kbp *Hin*dIII fragment from within the right *lux* operon of *V. fischeri* was tested as well. Cloning techniques, bacterial manipulations and CAT assays have been described in detail elsewhere (Miyamoto et al., 1990).

RESULTS

Nucleotide Sequence of luxG

A restriction map of *V. fischeri lux* DNA showing the gene organization is given in Figure 1A. In order to analyze the DNA downstream of *luxE*, DNA extending from the *Eco*RV at 7.1 kbp extending to the *Stu*I site was sequenced. The strategy for sequencing the DNA across this region is outlined in Figure 1B. An open reading frame extending for 711 nucleotides and transcribed in the same direction as the genes in the *V. fischeri lux* operon starts only 3 nucleotides after the termination codon for *luxE* (Figure 2). A Shine-Dalgarno (Shine and Dalgarno, 1974) sequence can be recognized just upstream of the ATG initiation codon within the coding region of *luxE*. The deduced amino acid sequence was found to share 41% identity with LuxG of *V. harveyi* (Figure 3) and is located in the same relative position in the *lux* operon. For these reasons, the new *V. fischeri lux* gene was designated as *luxG*. Analysis of potential open reading frames downstream of *luxG* revealed that the only possible open reading frame was for a convergent gene transcribed in the opposite direction and extending for some 500 nucleotides. This result was somewhat surprising since the *V. harveyi lux* system has an additional closely coupled *lux* gene, *luxH*, starting only 18 nucleotides after the termination codon for *luxG*.

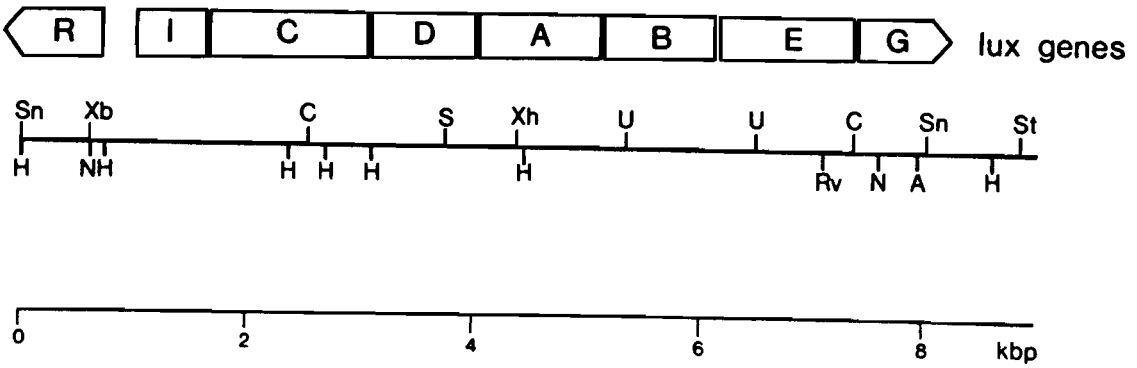
Analysis of the Bidirectional Terminator

Examination of the DNA sequence downstream of *luxG* revealed that a hairpin loop (base pair free-energy of -13 kcal/mol) with one mismatched base pair begins approximately 12 nucleotides after *luxG* and ends 2 nucleotides before the termination codon for the convergent gene (Figure 4C). As this hairpin loop starts with a poly(A) stretch and terminates with a poly(T) stretch, it appears that it could function as a rho-independent bidirectional termination site. It is also apparent that the poly(A) and poly(T) stretches can base pair so that the stem of the hairpin loop may vary in length. In order to determine if this site is active *in vivo*, S1 nuclease mapping was performed with mRNA isolated from the native *V. fischeri* strain as well

Figure 1: A) Restriction map of the *V. fischeri lux* regulon. The right *lux* operon contains the *luxICDABE* genes and the left *lux* operon contains the *luxR* gene. The restriction sites until the end of the *luxE* gene were obtained from the sequence data of Baldwin et al. (1989). Sn, *Sna*BI; H, *Hin*dIII; Xb, *Xba*I; N, *Nco*I; C, *Cla*I; S, *Sac*I; Xh, *Xho*I; U, *Pvu*II; Rv, *Eco*RV; A, *Apa*I; St, *Stu*I

B) Cloning strategy used to sequence the DNA downstream of *luxE*. The arrows represent the direction and extent of the sequence obtained from each clone. The open reading frames of *luxG* and the convergent gene were obtained from overlapping clones sequenced in both directions. Rv, *Eco*RV; C, *Cla*I; N, *Nco*I; Sn, *Sna*BI; H, *Hin*dIII; St, *Stu*I.

A


R
I
C
D
A
B
E
G
lux genes
Sn
Xb
C
S
Xh
U
U
C
Sn
St
H
NH
H
H
H
H
Rv
N
A
H
0
2
4
6
8
kbp

B

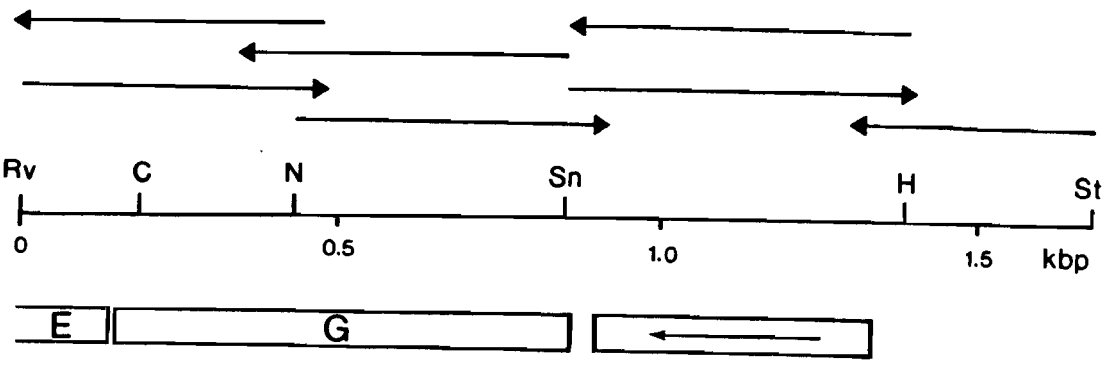

Rv
C
N
Sn
H
St
0
0.5
1.0
1.5
kbp
E
G

Figure 2: Nucleotide sequence of *luxG* and predicted amino acid sequence of the *luxG* product. *luxG* contains 711 nucleotides and codes for a protein of 236 amino acids with a molecular weight of 26,501. It starts 3 nucleotides after the stop codon of *luxE* and its Shine-Dalgarno sequence (underlined) is located within the 3' end of the *luxE* gene.

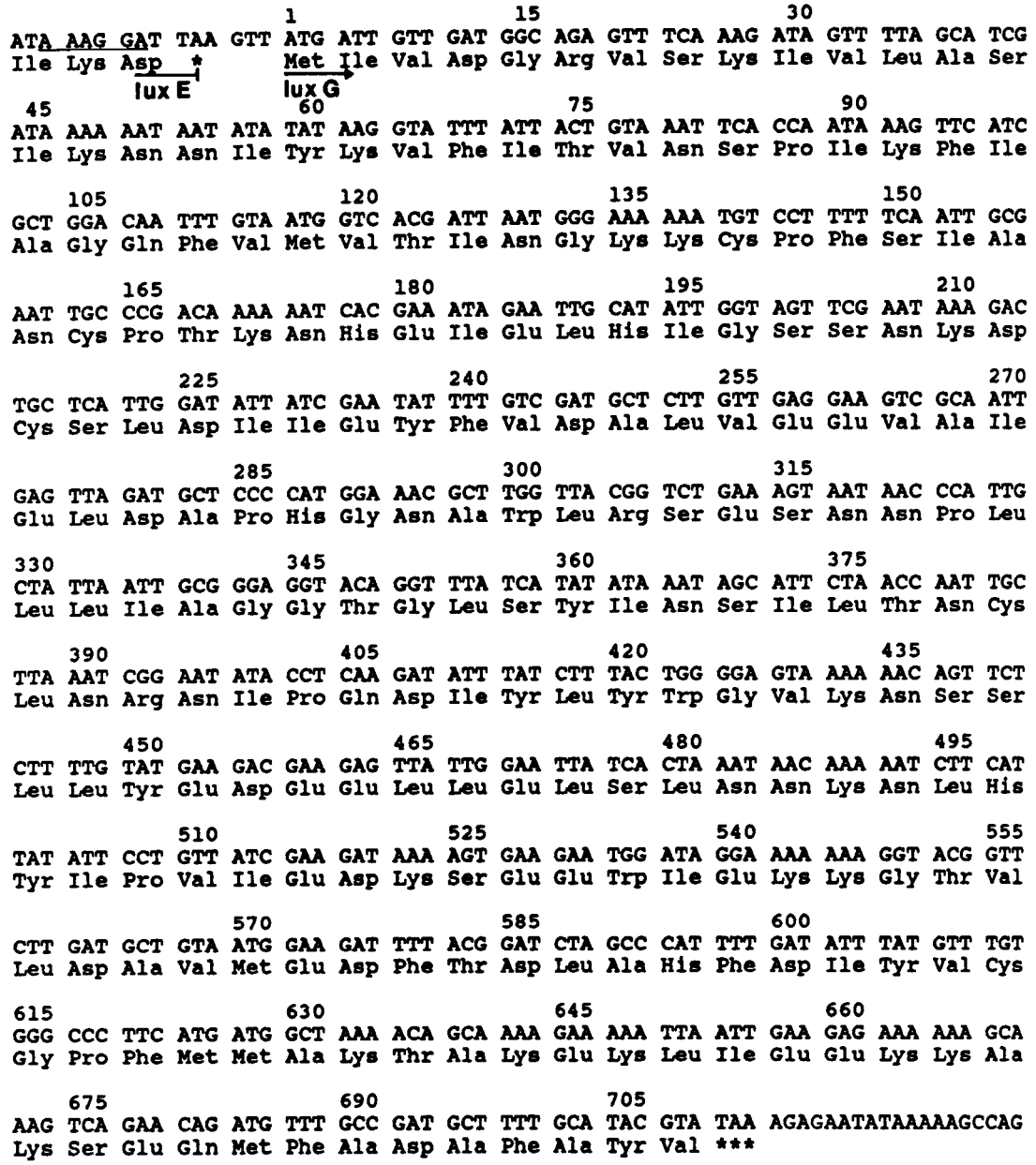

1
15
30
ATA AAG GAT TAA GTT ATG ATT GTT GAT GGC AGA GTT TCA AAG ATA GTT TTA GCA TCG
Ile Lys Asp * Met Ile Val Asp Gly Arg Val Ser Lys Ile Val Leu Ala Ser
luxE
luxG
45
60
75
90
ATA AAA AAT AAT ATA TAT AAG GTA TTT ATT ACT GTA AAT TCA CCA ATA AAG TTC ATC
Ile Lys Asn Asn Ile Tyr Lys Val Phe Ile Thr Val Asn Ser Pro Ile Lys Phe Ile
105
120
135
150
GCT GGA CAA TTT GTA ATG GTC ACG ATT AAT GGG AAA AAA TGT CCT TTT TCA ATT GCG
Ala Gly Gln Phe Val Met Val Thr Ile Asn Gly Lys Lys Cys Pro Phe Ser Ile Ala
165
180
195
210
AAT TGC CCG ACA AAA AAT CAC GAA ATA GAA TTG CAT ATT GGT AGT TCG AAT AAA GAC
Asn Cys Pro Thr Lys Asn His Glu Ile Glu Leu His Ile Gly Ser Ser Asn Lys Asp
225
240
255
270
TGC TCA TTG GAT ATT ATC GAA TAT TTT GTC GAT GCT CTT GTT GAG GAA GTC GCA ATT
Cys Ser Leu Asp Ile Ile Glu Tyr Phe Val Asp Ala Leu Val Glu Glu Val Ala Ile
285
300
315
GAG TTA GAT GCT CCC CAT GGA AAC GCT TGG TTA CGG TCT GAA AGT AAT AAC CCA TTG
Glu Leu Asp Ala Pro His Gly Asn Ala Trp Leu Arg Ser Glu Ser Asn Asn Pro Leu
330
345
360
375
CTA TTA ATT GCG GGA GGT ACA GGT TTA TCA TAT ATA AAT AGC ATT CTA ACC AAT TGC
Leu Leu Ile Ala Gly Gly Thr Gly Leu Ser Tyr Ile Asn Ser Ile Leu Thr Asn Cys
390
405
420
435
TTA AAT CGG AAT ATA CCT CAA GAT ATT TAT CTT TAC TGG GGA GTA AAA AAC AGT TCT
Leu Asn Arg Asn Ile Pro Gln Asp Ile Tyr Leu Tyr Trp Gly Val Lys Asn Ser Ser
450
465
480
495
CTT TTG TAT GAA GAC GAA GAG TTA TTG GAA TTA TCA CTA AAT AAC AAA AAT CTT CAT
Leu Leu Tyr Glu Asp Glu Glu Leu Leu Glu Leu Ser Leu Asn Asn Lys Asn Leu His
510
525
540
555
TAT ATT CCT GTT ATC GAA GAT AAA AGT GAA GAA TGG ATA GGA AAA AAA GGT ACG GTT
Tyr Ile Pro Val Ile Glu Asp Lys Ser Glu Glu Trp Ile Glu Lys Lys Gly Thr Val
570
585
600
CTT GAT GCT GTA ATG GAA GAT TTT ACG GAT CTA GCC CAT TTT GAT ATT TAT GTT TGT
Leu Asp Ala Val Met Glu Asp Phe Thr Asp Leu Ala His Phe Asp Ile Tyr Val Cys
615
630
645
660
GGG CCC TTC ATG ATG GCT AAA ACA GCA AAA GAA AAA TTA ATT GAA GAG AAA AAA GCA
Gly Pro Phe Met Met Ala Lys Thr Ala Lys Glu Lys Leu Ile Glu Glu Lys Lys Ala
675
690
705
AAG TCA GAA CAG ATG TTT GCC GAT GCT TTT GCA TAC GTA TAA AGAGAATATAAAAAGCCAG
Lys Ser Glu Gln Met Phe Ala Asp Ala Phe Ala Tyr Val ***

Figure 3: Comparison of the *V. harveyi* and *V. fischeri* LuxG amino acid sequences. The deduced amino acid sequence of *V. harveyi* LuxG (Swartzman et al., 1990) was compared to that of *V. fischeri* and the aligned sequences showed 41% identity. Maximum homology was obtained when a space (-) was placed within the *V. fischeri* sequence.

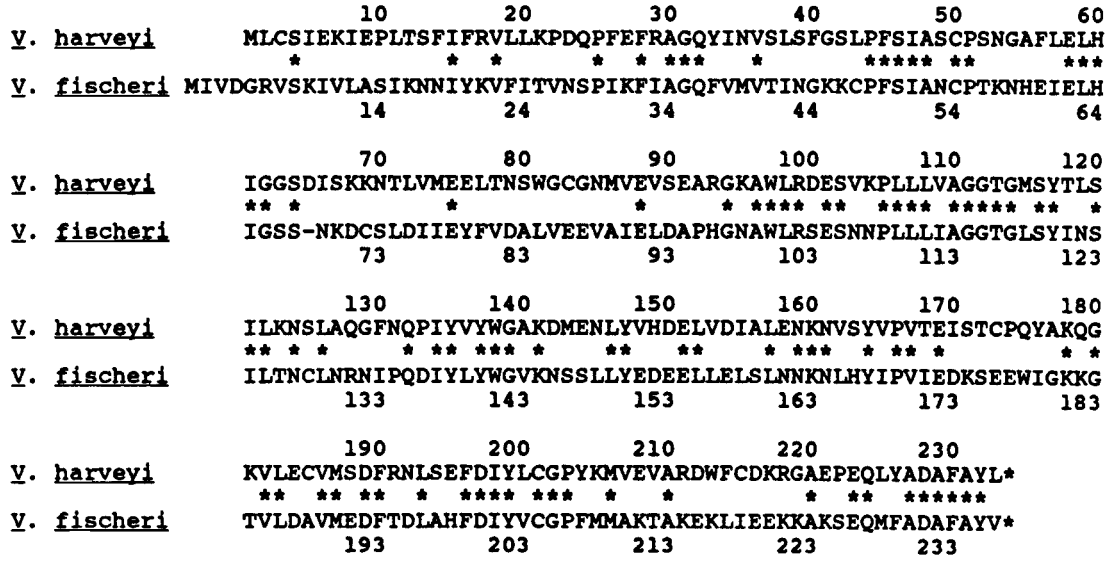
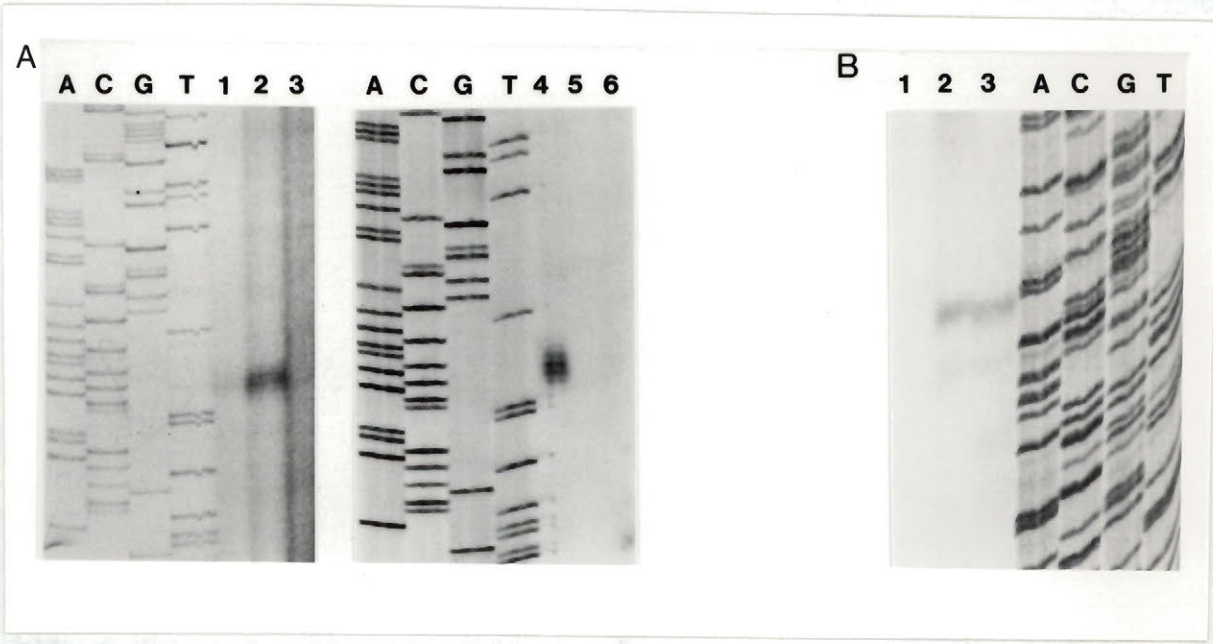
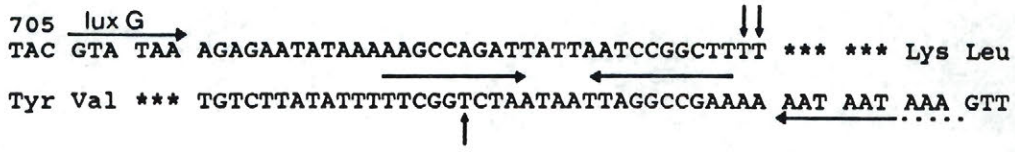

10 20 30 40 50 60
V. harveyi MLCSIEKIEPLTSFIFRVLLKPDQPFEFRAGQYINVSLSFGSLPFSIASCPSNGAFLELH
V. fischeri MIVDGRVSKIVLASIKNNIYKVFITVNSPIKFIAGQFVMVTINGKKCPFSIANCPTKNHEIELH
14 24 34 44 54 64
70 80 90 100 110 120
V. harveyi IGGSDISKKNTLVMEELTNSWGCGNMVEVSEARGKAWLRDESVKPLLLVAGGTGMSYTLS
V. fischeri IGSS-NKDCSLDIIEYFVDALVEEVAIELDAPHGNAWLRSESNNPLLLIAGGTGLSYINS
73 83 93 103 113 123
130 140 150 160 170 180
V. harveyi ILKNSLAQGFNQPIYVYWGAKDMENLYVHDELVDIALENKNVSYVPVTEISTCPQYAKQG
V. fischeri ILTNCLNRNIPQDIYLYWGVKNSSLLYEDEELLELSLNNKNLHYIPVIEDKSEEWIGKKG
133 143 153 163 173 183
190 200 210 220 230
V. harveyi KVLECVMSDFRNLSEFDIYLCGPYKMVEVARDWFCDKRGAEPEQLYADAFAYL*
V. fischeri TVLDAVMEDFTDLAHFDIYVCGPFMMAKTAKEKLIEEKKAKSEQMFADAFAYV*
193 203 213 223 233

Figure 4: 3' S1 nuclease mapping. A) 3' S1 nuclease mapping of the end of the *luxG* gene. Two independent experiments compare the mRNA from *V. fischeri* and *E. coli* transformed with pVfB1 and demonstrate the induction of expression. The type of RNA used for each hybridization is as follows: *V. fischeri*, lane 1; *E. coli* (pVfB1) harvested before (lane 5) and after (lanes 2 and 4) induction of luminescence at OD_{660}=0.5 and 1.6, respectively; *E. coli*, lanes 3 and 6. The autoradiogram containing lane 3 was overexposed in order to detect any bands over background. B) 3' S1 nuclease mapping of the end of the convergent gene. *E. coli*, lane 1, *E. coli*(pVfB1) harvested before (lane 2) and after (lane 3) induction of luminescence. The sizes of the mRNAs (133 and 134 bases for A and 297 bases for B) were determined by comparison to a sequence ladder of M13mp18 extending from the 17 base universal primer (5'-GTAAAACGACGGCCAGT) primer, as mobility is independent of base composition in denaturing polyacrylamide gels (Maniatis and Efstratiadis, 1980). C) Nucleotide sequence of the bidirectional terminator. The converging arrows represent the stem of the hairpin loop. Because of the poly(A) and poly(T) regions flanking this structure, the size of the stem may vary. The vertical arrows indicate where the mRNAs for the *luxG* and convergent gene terminate according to 3' S1 nuclease mapping. The ends of the upper and lower arrows indicate the positions of the termination codons for *luxG* and the convergent gene, respectively.


A
A C G T 1 2 3
A C G T 4 5 6
B
1 2 3 A C G T

C


705
lux G
TAC GTA TAA AGAGAATATAAAAAGCCAGATTATTAATCCGGCTTTT *** *** Lys Leu
Tyr Val *** TGTCTTATATTTTTCGGTCTAATAATTAGGCCGAAAA AAT AAT AAA GTT

as from E. coli transformed with the complete *V. fischeri lux* system. Hybridization of the mRNA with a 650 bp *Sau*96I-*Hin*dIII restriction fragment 3' ^{32}P-labeled at the Sau96I site (Figure 5), followed by S1 nuclease digestion of unhybridized mRNA and DNA and high resolution electrophoresis revealed a DNA fragment partially protected against S1 nuclease digestion. This fragment was the same size whether the probe was hybridized to mRNA from *V. fischeri* or mRNA from *E. coli* transformed with the *V. fischeri lux* genes (Figure 4A). By comparison to markers in this denaturing gel, the position of termination could be localized to the poly(T) region immediately after the hairpin loop (nucleotides 747 and 748). As mRNA could be more readily isolated from *E. coli* than *V. fischeri*, presumably reflecting differences in nuclease levels, most experiments were conducted with transformed *E. coli*. Protection of this DNA fragment from S1 nuclease digestion could only be accomplished using mRNA isolated from the cells after induction of luminescence.

To determine if this terminator can function bidirectionally, mRNA was hybridized to a 400 bp *Apa*I-*Ssp*I restriction fragment 3' ^{32}P-labelled at the *Ssp*I site (Figure 5). S1 nuclease mapping shows that message corresponding to the convergent gene is present in cells before and after induction of luminescence (Figure 4B). The termination site was located within the stem of the hairpin loop (nucleotide 729) at the site of the mismatched base pair (Figure 4C) instead of within the run of T residues as expected, possibly due to limited nuclease digestion. No signal could be detected when RNA from *E. coli* was hybridized to either probe (Figure 4A and 4B) indicating that only specific mRNA-DNA hybridizations were observed. Although not shown , the only other visible bands were the full length, undigested dsDNA probes.

In Vivo Termination Assay Using CAT as the Reporter Gene

The strength of the bidirectional terminator was determined using an *in vivo* termination assay. A DNA fragment containing the hairpin loop was inserted in both directions in a plasmid between the promoter for the *V. harveyi lux* system and a reporter gene (CAT) as shown in Figure 6A. The recombinant plasmids were then

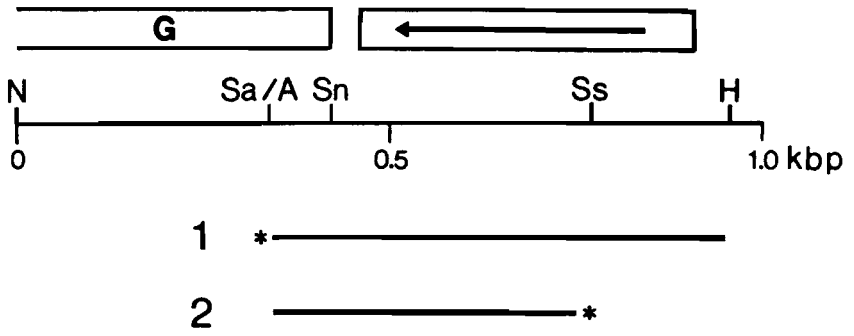
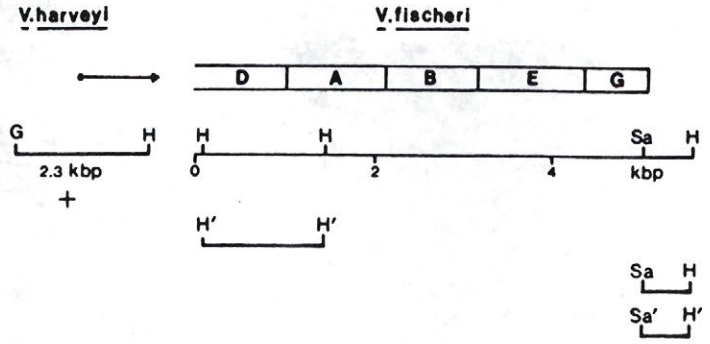

G
N
Sa/A
Sn
Ss
H
0
0.5
1.0 kbp
1
2

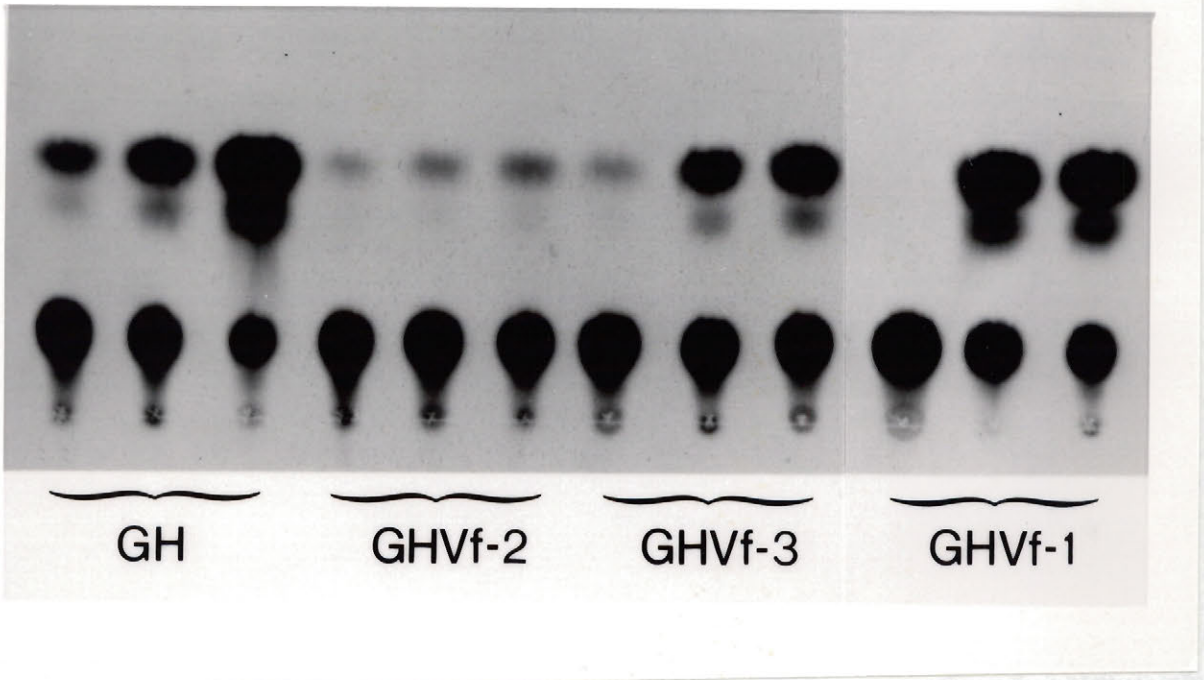
Figure 6: *In vivo* **termination assay.** A) Inserts of *V. fischeri* DNA were placed between the inducible promoter of the *V. harveyi lux* operon (GH) and the CAT reporter gene and transferred by conjugation into *V. harveyi*. GHVf-1 and GHVf-3 were inserted backwards with respect to direction of transcription off the *V. harveyi* promotor. The specific activity of CAT was calculated from lysates of cells harvested between OD_{660} = 1.5 and 2.0. G, *Bgl*II; H, *Hin*dIII; Sa, *Sau*96I. B) CAT activity was assayed from extracts of *V. harveyi* containing GH (at OD_{660} = 0.58, O.96 and 1.5), GHVF-2 (at OD_{660} = 0.52, 1.0 and 2.0), GHVf-3 (at OD_{660} = 0.57, 0.95 and 1.7), and GHVf-1 (at OD_{660}=0.44, 0.99 and 1.5). Twenty times more protein was used for CAT assays of cells containing GHVf-2 and GHVf-3.

A



INSERT	ACTIVITY (units/mg)
GH	925
GHVf-1	756
GHVf-2	3
GHVf-3	17

B



transferred by conjugation into the luminescent bacterium *V. harveyi*. CAT activity was assayed during different stages of growth (Figure 6B) and the specific activity of CAT was calculated. The level of expression of the CAT gene is decreased by over 54-fold on insertion of this DNA in the backward direction (GHVf-3), and over 300-fold when inserted in the forward direction (GHVf-2). A control involving insertion of a much larger fragment of *V. fischeri lux* DNA (GHVf-1) between the promoter and the CAT gene had little effect on the level of expression when compared to the control with no *V. fischeri* DNA.

Discussion

Sequencing the distal end of the right *lux* operon of *V. fischeri* has revealed the existence of an open reading frame located immediately downstream of *luxE*, coding for a protein of 236 amino acids. S1 nuclease mapping has shown that the mRNA transcribed from this gene is induced in a cell density-dependant manner indicating that this gene is coregulated with the *luxICDABE* genes. Because the amino acid sequence for this open reading frame was found to share 41% identity with *luxG* of *V. harveyi* and is located in the same relative position in the *lux* operon (Swartzman et al., 1990), the new gene is designated as *luxG*. A recent study (Baldwin et al., 1989) presenting the entire sequence of the left and right *lux* operons of *V. fischeri* included the start of an open reading frame after *luxE* which corresponds in sequence to the amino terminal of LuxG presented in this work.

A hairpin loop that functions as a bidirectional transcription terminator was located immediately after *luxG*. Although relatively rare, there have been reports of bidirectional terminators in *E. coli* (Hudson and Davidson, 1984; Jerlstrom et al., 1989; Lloubes et al., 1988; Postle and Good, 1985; Schollmeier et al., 1985), other bacteria (Carlomagna et al., 1985; Chen et al., 1987; Neal and Chater, 1987; Platt, 1986), and even mitochondria (Christianson and Clayton, 1986) and chloroplasts (Westhoff and Gitterman, 1988). It has been suggested that a bidirectional terminator may play a regulatory role in gene expression (Platt, 1986; Postle and Good, 1985). Antitermination through a weak termination site would produce anti-sense RNA blocking translation of the convergent gene (Mizuno et al., 1984). It is also possible that the complementary 3' ends of the mRNAs hybridize to help stabilize the messages against nuclease digestion. mRNA corresponding to *luxG* was only detectable in cells producing high levels of light while mRNA corresponding to the convergent gene was present in cells before and after induction of luminescence. This indicates that the convergent gene does not respond to the same regulatory signals as the right *lux* operon and is not likely to be involved with the *lux* system. The 3' end of the *V. fischeri lux* mRNA was identical using RNA isolated from *V. fischeri* and from an *E. coli* clone containing the *lux* operons. It was also found that the

hairpin loop was highly effective in terminating transcription in both directions in *V. harveyi*, although the terminator is 6-fold more efficient in the forwards than the backwards direction. The *V. fischeri* terminator is therefore recognized by the RNA polymerases from three different bacteria: *V. fischeri*, *E. coli*, and *V. harveyi*, demonstrating a common mechanism of transcription termination.

The function of *luxG* is still unknown and homologies with proteins of known function have not yet been found. *luxG* is apparently not required for light production or *lux* gene regulation in *V. fischeri* since transposon mutagenesis has failed to detect any Lux- phenotypes for disruptions in any other genes except *luxRICDAB* and *E* (Engebrecht et al., 1983). Since *luxG* has so far been found in the *lux* operons of three different species of luminescent bacteria (*V. harveyi*, *P. phosphoreum* [Swartzman et al., 1990] and *V. fischeri*), it would appear that *luxG* is a common gene in the luminescent systems of marine bacteria. However, the *V. fischeri* gene organization contrasts to that of the *V. harveyi lux* system where an additional gene, *luxH*, was found after *luxG* before the transcription terminator. Consequently, these results show that the 3' ends as well as the 5' ends (Miyamoto et al., 1988b) of the *lux* regulons of *V. harveyi* and *V. fischeri* have diverged.

CHAPTER 4

VIBRIO HARVEYI RNA POLYMERASE: PURIFICATION AND RESOLUTION FROM GYRASE A

PREFACE

The following chapter has been published as presented: Swartzman, E., and E. A. Meighen. (in press) *Biochem. Cell. Biol.* vol: **70**

ABSTRACT

RNA polymerase was purified from *Vibrio harveyi* and was found to contain polypeptides (ß, ß', α, and σ) closely corresponding to those of the *Escherichia coli* enzyme. *In vitro* transcription studies using *V. harveyi* and *E. coli* RNA polymerase demonstrated that the purified *V. harveyi* RNA polymerase is functional and that the two enzymes have the same promoter specificity. Chromatography through a monoQ column was required to remove a 100 kDa protein that was present in large amounts and copurified with the RNA polymerase. N-terminal amino acid sequencing showed that the first 18 amino acids of the 100 kDa protein shares 78% sequence identity with the A subunit of gyrase or topoisomerase II. The abundance of the gyrase A protein is unprecedented and may be linked to bioluminescence.

INTRODUCTION

Genes responsible for light production have been identified and cloned from several species of marine bacteria (for review see Meighen 1991), although regulation of luminescence has been studied in only two species, *Vibrio harveyi* and *Vibrio fischeri*. The two bacteria share common structural *lux* genes, *luxCDABE* (Engebrecht and Silverman 1984, Miyamoto et al., 1988a) but differ with respect to *lux* gene organization and regulation. The light generating reaction is catalyzed by luciferase, encoded by *luxA* and *B*, utilizing O_2, reduced flavin, and aldehyde as substrates. The aldehyde is provided by the gene products of *luxCD* and *E*, polypeptides of the fatty acid reductase complex (Boylan et al., 1985). Two new genes of unknown function have been recently identified as part of the *lux* operon. *luxG* and *H* are located downstream of *luxE* in *V. harveyi*, whereas *V. fischeri* contains only the *luxG* gene (Swartzman et al., 1990b, Swartzman et al., 1990a).

The study of *lux* gene regulation had been initially limited to *V. fischeri* since *Escherichia coli* cells transformed with the *V. fischeri lux* genes are able to produce wild-type levels of regulated light emission (Engebrecht et al., 1983). The cells produce an autoinducer, ß-ketocaproyl homoserine lactone (Eberhard et al., 1981), that diffuses freely across the cell wall into the growth medium (Kaplan and Greenberg 1985). *luxI*, the gene thought to be responsible for autoinducer production, is located immediately upstream of and cotranscribed with *luxCDABEG* as part of the right *lux* operon. It has been proposed that the receptor protein, a product of the *luxR* gene, binds to autoinducer and positively regulates transcription of the right *lux* operon. *luxR* is located upstream of *luxICDABEG* but is transcribed in the opposite direction, comprising the left *lux* operon. When autoinducer concentration reaches a critical threshold, transcription of the right *lux* operon is induced and a positive feedback loop is established (Engebrecht et al., 1983).

Efforts to elucidate the mechanism of light induction in *V. harveyi* have been hampered by our inability to obtain clones that reproduce wild-type levels of regulated light emission in *E. coli* (Miyamoto et al., 1987). Luminescence in *V.*

harveyi is cell-density controlled, as in *V. fischeri*, and is dependent on a structurally similar autoinducer, ß-hydroxybutryl homoserine lactone (Cao and Meighen 1989). However, we have been unable to locate a corresponding luxI or *luxR* gene near the structural *lux* genes of *V. harveyi* (Miyamoto et al., 1988b) or produce regulated luminescence in *E. coli* with complementing clones. The cloning of a *V. harveyi* gene that stimulates expression of *lux* genes *in trans* has been reported and although it has no sequence homology to the *V. fischeri luxR* gene, it appears to have a DNA-binding motif (Martin et al., 1989, Showalter et al., 1990). It was concluded, though, that *lux* gene regulation in *E. coli* requires an additional regulatory element that has yet to be identified.

The *V. harveyi lux* promoter has already been mapped (Swartzman et al., 1990b) and studied *in vivo* using conjugative transfer of cloned DNA into *V. harveyi* (Miyamoto et al., 1990). Further insight may be achieved through *in vitro* studies of the *V. harveyi lux* promoter. To that end, we describe here the purification of *V. harveyi* RNA polymerase and its use in an *in vitro* transcription assay. Large amounts of a 100 kDa protein copurified with the RNA polymerase throughout the purification. N-terminal sequence analysis demonstrates that this protein is the A subunit of gyrase or topoisomerase II which controls the superhelical tension of the bacterial chromosome. Bioluminescence in marine bacteria appears to be osmoregulated and linked to gyrase activity (Watanabe et al., 1991). The presence of large amounts of the A subunit of gyrase in *V. harveyi* may be related to its role in bioluminescence.

MATERIALS AND METHODS

Purification of V. harveyi *RNA Polymerase.*

V. harveyi was grown in 8 L of 1% complex media (Miyamoto et al., 1985) at 27° and was harvested and stored at -70° before use. The cells (27 g) were thawed at 37° before addition of 100 mL of lysis buffer (50 mM HEPES pH 7.9, 2 mM EDTA, 1 mM DTT, 1 mM PMSF, and 5% glycerol). After stirring for 1 h at 4°, 100 μg of lysozyme was added, followed by an additional h of stirring. Twelve g of solid NaCl was added (1 M final) and stirred for a final 30 min. The lysed cells were centrifuged at 44,000 x g for 1 h to obtain a clear, but viscous lysate (260 mL). One tenth volume (26 mL) of saturated ammonium sulfate (3.5 M) and 94 g of solid ammonium sulfate was added to the high speed supernatant and stirred for 1 h. After spinning at 25,000 x g for 30 min, the pellet was resuspended in 69 mL of buffer A (10 mM Tris-Cl pH 7.9, 0.1 mM EDTA, 1 mM DTT, 5% glycerol) containing 100 mM KCl and dialysed overnight against 4 L of the same buffer.

After centrifuging to remove debris, the sample (94 mL) was loaded onto a 40 mL dsDNA-cellulose (Sigma) column. The column was washed with 90 ml of buffer A containing 100 mM KCl and protein was eluted with buffer A containing 75 mM KCl. Two volumes of 3.5 M ammonium sulfate were added to the pooled fractions and stored overnight at 4°. The precipitate was centrifuged at 25,000 x g, and the pellet resuspended in 1 mL of buffer A containing 0.5 M KCl.

The concentrated dsDNA-cellulose pool was loaded onto a 1.5 x 48 cm Biogel A 1.5 (Bio-Rad) column that had been equilibrated with buffer A containing 0.5 M KCl and was eluted with the same buffer. One mL fractions were collected and two active fractions were pooled separately and diluted with buffer A until the KCl concentration reached 100 mM.

The diluted Biogel A 1.5 fractions (A and B) were applied separately onto a monoQ column using fast performance liquid chromatography (Pharmacia) and a gradient of buffer A from 100 mM to 1 M KCl was used to elute the protein. The active fractions were pooled and dialysed against buffer A containing 100 mM KCl

and 50% glycerol, and then stored at -20°.

In vitro *Transcription*

A *Bam*HI-*Cla*I 1.0 kbp fragment spanning the 5'-non coding and coding regions of the *luxC* gene was used for all transcription reactions (see Miyamoto et al., 1988b for detailed restriction map). DNA (100 ng) was preincubated with RNA polymerase (1 μl of the *V. harveyi* enzyme [~0.25 units] or 0.5 units of *E. coli* RNA polymerase purchased from Pharmacia) and 1.5 μl 10 x transcription buffer in a total volume of 12.5 μl for 5 min at 37°. One μl of 2 mg/mL heparin was added to inactivate free RNA polymerase and the incubation continued for a further 5 min. The transcription reaction was started by the addition of 1.5 μl of 10 x nucleotide mixture and continued for 15 min. The reaction was stopped by phenol/chloroform extraction, and the RNA was ethanol precipitated with 50 μg of *E. coli* tRNA added as carrier.

The 10 x transcription buffer consists of 0.2 mg/mL BSA, 0.2 M Tris-acetate pH 8.0, 40 mM $MgCl_2$, 1.0 M KCl, 1.0 mM EDTA, 1.0 mM DTT, and 40% glycerol. The 10 x nucleotide mixture contained 1.5 mM each of UTP, ATP, and GTP, 0.2 mM of CTP and 0.33 μM [α-^{32}P]CTP (3000 Ci/mmol). The ^{32}P-labelled transcripts were resuspended in a formamide-dye mix and heated at 65° for 2 min before loading on a 6.5% acrylamide-7 M urea sequencing gel. A sequencing ladder of M13mp18 served as size markers.

RNA Polymerase Assay

The assay contained 5 μL 10 x transcription buffer, 5 μL 10 x nucleotide mix with 0.06 μM [α-^{32}P] CTP (3000 Ci/mmol), and 5 μL of 1.5 mg/mL sonicated calf thymus DNA in a total volume of 50 μL. The reactions were incubated at 37° for 15 min and the transcription products were precipitated with 5 μL of 100% trichloroacetic acid. After 15 min on ice, the precipitate was collected on glass fibre filters (#30 Schleicher and Schuell) and quantified by Cerenkov counting. One unit of RNA

polymerase corresponds to the incorporation of 1 nmol of CMP.

Other Methods

N-terminal sequence analysis was done at the Sheldon Biotechnology Centre, McGill University. SDS-Page gels (4% stacking, 10% resolving) were by the method of Laemmli (1970) and stained with Kodak Coomassie blue R-250.

RESULTS

Purification of V. harveyi *RNA Polymerase.*

The strategy (Table 1) used to purify the *V. harveyi* RNA polymerase was based on one developed by Burgess and Jendrisak (1975), for the purification of the *E. coli* enzyme, with some modifications. Instead of polymin P precipitation, total protein was precipitated away from the clarified cell extract with ammonium sulfate. The dsDNA-cellulose column provided a rapid enrichment of DNA-binding proteins and was followed by a gel filtration step. As can be seen by SDS-PAGE analysis of the fractions eluted from the Biogel A 1.5 column (Fig. 1), polypeptides corresponding to RNA polymerase are eluted initially, followed closely by large amounts of a 100 kDa protein. Although activity eluted in one peak, fractions 48 to 52 (A) and 53 to 58 (B) were pooled separately in order to maximize recovery without compromising purity. The separate pools were then subjected to MonoQ fast performance liquid chromatography, resulting in a great increase of purity with minimal loss of activity. Analysis of the final fractions on SDS-PAGE (Fig. 2) reveals that the monoQ A fraction is over 90% pure, while the 100 kDa protein compromises over 45% of the monoQ B fraction.

Table 1 summarizes the recovery of purified RNA polymerase from *V. harveyi*. The specific activity of the *V. harveyi* RNA polymerase monoQ A fraction (102 units/mg) is less than but comparable in value to that for most preparations of *E. coli* RNA polymerase. The lower activity may be attributed to less than saturating amounts of the σ subunit. Approximately 2.3-2.5 μg of *V. harveyi* RNA polymerase produced the same amount of transcript as 1 μg of the *E. coli* enzyme using the *lux* promoter (see *in vitro* transcription below). The subunit composition of the purified RNA polymerase is similar to that of the *E. coli* enzyme, with molecular weights for the component polypeptides of 160,000(β'), 150,000(β), 88,000(σ), and 68,000(α) as estimated by migration on SDS-PAGE.

Table 1

Summary of RNA polymerase purification from *V. harveyi*

Fraction	Volume (ml)	mg/ml	Total protein (mg)	Total Units	Specific Activity (units/mg)	Yield (%)	Fold Purification
Clarified lysate	260	11	2860	1095	0.38	-	-
Ammonium Sulfate	94	16	1504	203	0.13	18.5	0.34
DNA-cellulose	1.0	11.6	11.6	35	3.0	3.2	7.9
Biogel A 1.5 A	5.0	0.38	1.9	61	32	5.6	84
Biogel A 1.5 B	6.0	0.62	3.7	66	18	6.0	47
MonoQ A	0.25	2.5	0.63	64	102	5.8	268
MonoQ B	0.28	2.3	0.62	48	78	4.4	205

Figure 1: SDS-polyacrylamide gel electrophoresis of fractions from Biogel A 1.5 column with RNA polymerase activity. Five μl of the indicated fractions were analyzed on a 10% polyacrylamide gel. Fractions 48 to 52 and 53 to 58 were pooled separately.

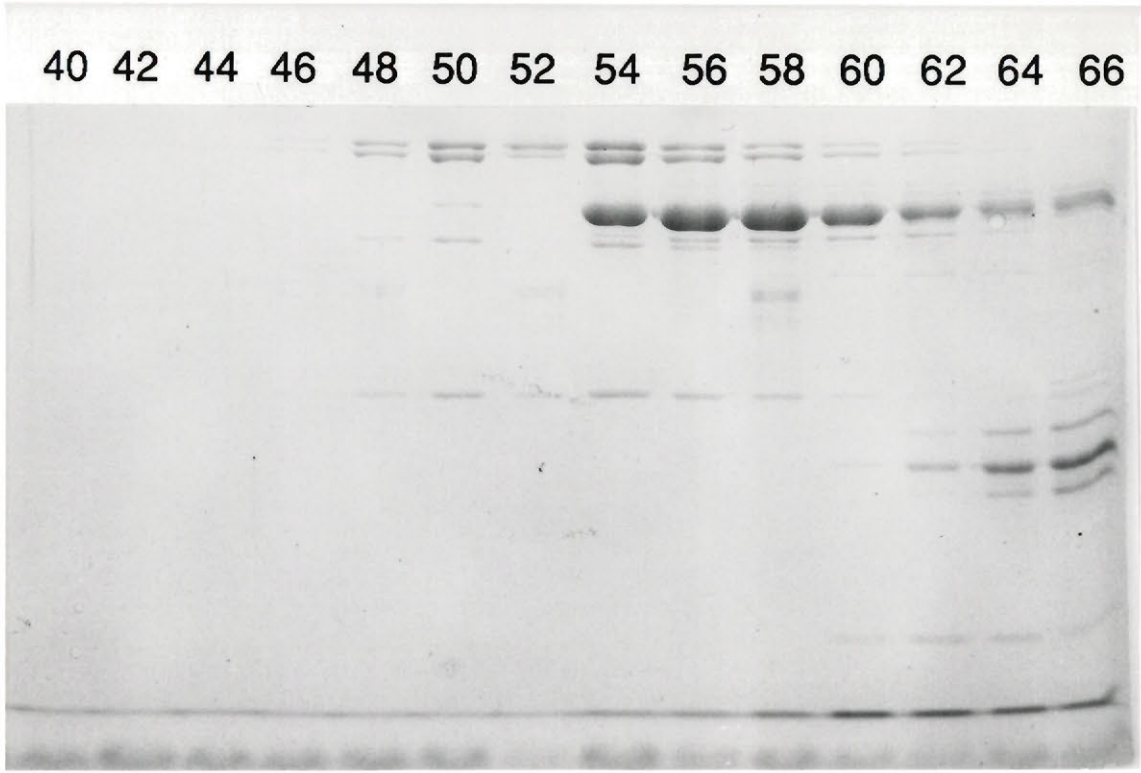
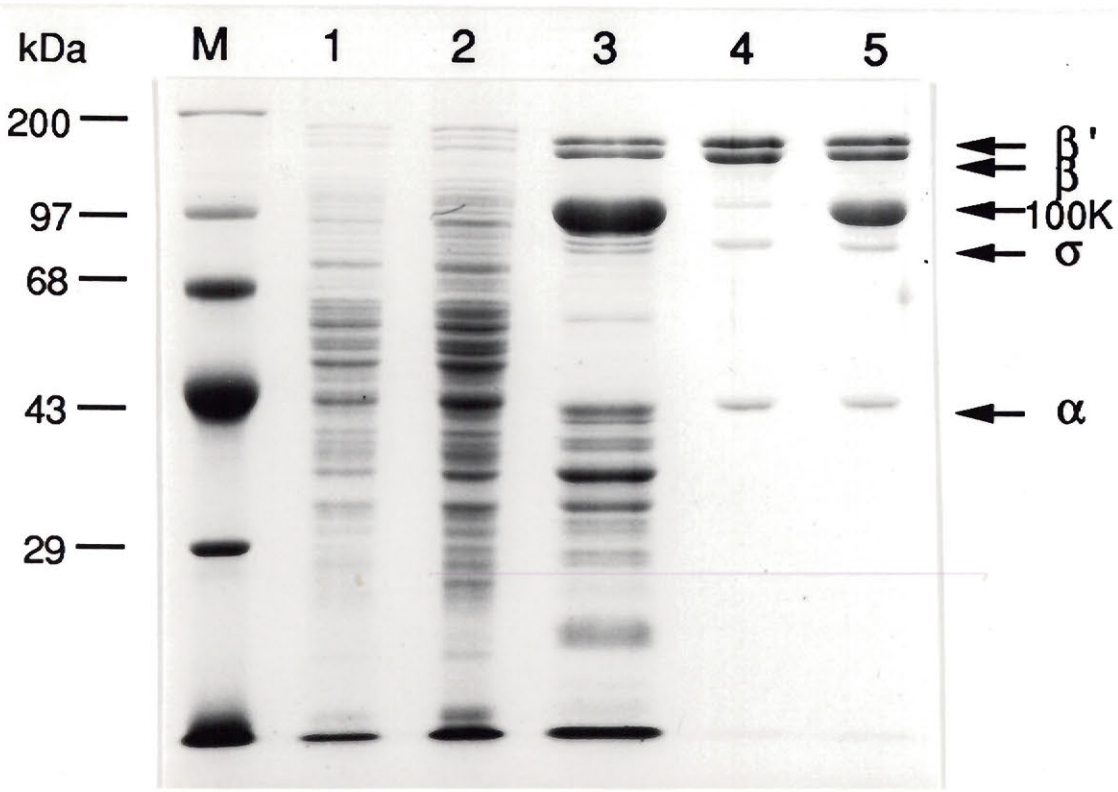

40 42 44 46 48 50 52 54 56 58 60 62 64 66

Figure 2: SDS-polyacrylamide gel electrophoresis of fractions with RNA polymerase activity at different stages of purification: 1) 11 μg of high speed supernatant, 2) 16 μg of resuspended ammonium sulfate pellet, 3) 12 μg of dsDNA-cellulose fraction, 4) 2.5 μg of dialysed monoQ A fraction, 5) 2.3 μg of dialysed monoQ B fraction. The molecular weight of protein markers are given in lane M. The β, β', α, and σ subunits of the *V. harveyi* RNA polymerase as well as the 100 kDa protein are indicated.


kDa
M
1
2
3
4
5
200
97
68
43
29
β'
β
100K
σ
α

In vitro *Transcription*

The purified *V. harveyi* RNA polymerase was tested in an *in vitro* transcription assay using a DNA fragment containing the *lux* promoter. Both fractions, monoQ A and monoQ B, were used in the assay and compared with the *E. coli* RNA polymerase product. As can be seen in Figure 3, the monoQ A (lane 2) and monoQ B (lane 3) fractions produced identical transcripts in size and amount, indicating that the 100 kDa protein does not interfere with transcription. *E. coli* RNA polymerase produced a transcript (lane 1) that is identical in size to that of the *V. harveyi* enzyme, establishing that the two enzymes have the same specificity at this particular promoter. The transcripts are 186 nucleotides in length, which corresponds to 130 nucleotides upstream of the first ATG codon of the *luxC* gene. Because a DNA marker was used, these measurements may not reflect the exact size of the transcripts. A previous study (Swartzman et al., 1990b) had shown that the 5'-end of *lux* mRNA isolated *in vivo* mapped to 28 bp upstream of the *lux* gene.

The 100 kDa Protein is Gyrase A

The presence of the 100 kDa protein was of interest since it comprised over 30% of the proteins eluting from the dsDNA-cellulose column, as estimated by densitometry scanning. Pure 100 kDa protein was obtained from a fraction off the monoQ B column and the first 18 amino acids were elucidated via automated N-terminal Edman degradation. A search through the Genebank database revealed that the sequence shares 78% identity with the first 19 amino acids of the A subunit of *E. coli* gyrase, also a 100 kDa protein (see Fig. 4). The purified protein was tested for gyrase activity, but because the A subunit requires the B subunit for superhelical and/or relaxation activity (Higgins et al., 1978, Gellert 1981), we were unable to detect any activity (data not shown).

Figure 3: ***In vitro*** **transcription of the** ***lux*** **promoter.** A 1.0 kbp *Bam*HI-*Cla*I DNA fragment spanning the 5'-coding and noncoding region of the *luxC* gene was transcribed using different preparations of RNA polymerase: 1) *E. coli* RNA polymerase, 2) *V. harveyi* RNA polymerase (monoQ A fraction), 3) *V. harveyi* RNA polymerase (monoQ B fraction). The arrow indicates the position of the most prominent full-length products. A sequence ladder of M13mp18 was used as a size reference.

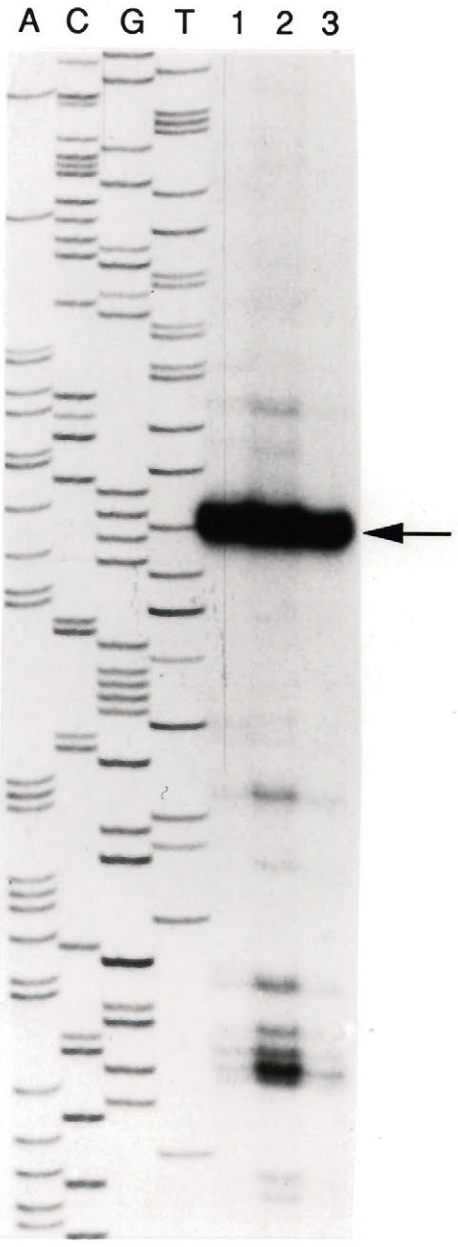

A C G T 1 2 3

Figure 4: The first 18 amino acids of the 100 kDa protein determined by automated N-terminal Edman degradation are aligned with the amino terminal sequence of *E. coli* gyrase A (Yoshida et al., 1988). The N-terminal methionine is apparently missing from the 100 kDa protein. The asterisks denote exact sequence identity.


100 kDa S D L A K E I T P V N I E D E L R G
* * * * * * * * * * * * * *
gyrase A M S D L A R E I T P V N I E E E L K S

DISCUSSION

Analysis of *lux* gene regulation has demonstrated some interesting and unique features. In both *V. harveyi* and *V. fischeri*, luminescence is cell-density dependent and catabolite repressed (Dunlap and Greenberg 1985, Nealson et al., 1972). *Lux* gene regulation in *V. fischeri* involves a complicated positive and negative feedback loop consisting of the receptor-inducer complex and CRP-cAMP regulating expression of the left and right *lux* operons (Dunlap and Greenberg 1988). There is also evidence that *lexA* (Ulitzer 1989) may play a role in *V. fischeri lux* gene regulation.

Much less is known about the mechanism of *lux* gene regulation in *V. harveyi*, primarily because we have been unable to reconstitute regulated light emission in *E. coli*. This problem can be overcome by establishing an *in vitro* transcription system where purified proteins are added and tested for their ability to affect transcription of the *lux* promoter. It was necessary to first purify the *V. harveyi* RNA polymerase and examine its activity in the *in vitro* transcription assay. The results demonstrated that the *V. harveyi* and *E. coli* RNA polymerase initiate transcription at the same site *in vitro*, indicating that the two enzymes share the same specificity at the *lux* promoter. Consequently it should be possible to use the more readily available *E. coli* RNA polymerase as well as the *V. harveyi* RNA polymerase for future *in vitro* transcription studies. The size of the *in vitro* synthesized transcripts indicate that the 5'-ends of the *in vivo* and *in vitro* transcripts are different. This result suggests that another sigma factor or regulatory protein is required *in vivo* for proper transcription at the -28 promoter. *In vitro* transcription can be used to assay for proteins responsible for this difference.

Large amounts of a 100 kDa protein copurified with the *V. harveyi* RNA polymerase up until the last stages of the purification, and constituted over 30% of the DNA-binding proteins eluting from the dsDNA-cellulose column. N-terminal sequence analysis suggests that this protein is the A subunit of gyrase or topoisomerase II. To our knowledge, gyrase A has not previously been found in high

quantities during purification of RNA polymerase from other bacteria. Bioluminescence in marine bacteria has recently been found to be dependant on osmolarity and indirectly linked to gyrase activity (Watanabe et al., 1991). Because gyrase is required for proper transcription and replication of the bacterial chromosome (Gellert 1981), the abundance of gyrase A in *V. harveyi* may be essential.

CHAPTER 5

THE *LUXR* GENE PRODUCT IS A TRANSCRIPTIONAL ACTIVATOR OF THE *VIBRIO HARVEYI LUX* PROMOTER

PREFACE

A modified version of this chapter has been published: Swartzman, E., M. Silverman, E. A. Meighen. (in press) *J. Bacteriol.*

Dr. Micheal Silverman provided *V. harveyi* BB7 wild type and mutant strains, as well as the *E. coli* recombinant containing the *V. harveyi luxR* gene.

ABSTRACT

Expression of the *lux* operon from the marine bacterium *Vibrio harveyi* is dependant on cell density and requires an unlinked regulatory gene, *luxR*, and other cofactors for autoregulation. *Escherichia coli* transformed with the *lux* operon emits very low levels of light, and this deficiency can be partially alleviated by co-expression of *luxR* in *trans*. The *V. harveyi lux* promoter was examined *in vivo* and *in vitro* with primer extension mapping to determine the mechanism of activation by *luxR*. RNA isolated from *Escherichia coli* transformed with the *Vibrio harveyi lux* operon was shown to have a start site at 123 bp upstream of the first ATG codon of *luxC*, as was RNA synthesized *in vitro* using a DNA fragment containing the *lux* promoter and *E. coli* or *V. harveyi* RNA polymerase. This is in sharp contrast to the start site found for *lux* RNA isolated from *V. harveyi*, at 26 bp upstream of the *luxC* initiation codon. However, when *E. coli* was cotransformed with both the *lux* operon and *luxR*, the start site of the *lux* mRNA shifted from -123 to -26. Furthermore, expression of the *luxR* gene caused a 350-fold increase in *lux* mRNA levels. The results suggest that LuxR is a transcriptional activator required for accurate transcription of the *V. harveyi lux* promoter.

INTRODUCTION

The generation of light from bioluminescent marine bacteria is influenced by a variety of factors (Dunlap and Greenberg, 1991), and is regulated by cell-density dependant autoinduction. Although genes responsible for light production have been cloned from several different species of bacteria (Meighen, 1991), the best characterized luminescent system is from *Vibrio fischeri* since all the genes required for producing a regulated luminescence phenotype have been isolated and expressed in *E. coli* (Engebrecht et al., 1983).

The light generating reaction is catalyzed by the α and β subunits of luciferase (products of the *luxA* and *luxB* genes, respectively) and requires O_2, $FMNH_2$ and aldehyde as substrates. The aldehyde is provided by polypeptides of the fatty acid reductase complex, encoded by the *luxCD* and *E* genes (Boylan et al., 1985). In *V. fischeri*, the genes are located within one operon in the order of *luxICDABEG*, where *luxG* has no known function and *luxI* is responsible for autoinducer production (Swartzman et al., 1990; Engebrecht and Silverman, 1983). The autoinducer, identified as N-ß-ketocaproyl homoserine lactone, is a freely diffusible molecule that accumulates in the growth medium as cell density increases (Eberhard et al., 1981, Kaplan and Greenberg, 1985). It has been proposed that the autoinducer establishes a positive feed back loop by binding to a receptor protein and stimulating transcription of the *luxICDABEG* operon (Engebrecht and Silverman, 1986). The receptor protein is encoded by the *luxR* gene which is located immediately upstream of the *lux* operon but is transcribed in the opposite direction (Engebrecht and Silverman, 1984). The operator for LuxR binding has been identified (Devine et al., 1989), although the mechanism of transcriptional activation has not yet been determined.

The *lux* operon of *V. harveyi* has also been isolated and consists of *luxCDABEGH* (Miyamoto et al., 1988b; Swartzman et al., 1990b). There is no known function for *luxH*, although it shares significant homology with *htrP* of *E. coli*, an essential gene required for bacterial growth at high temperatures (Raina et al., 1991). Unlike *V.*

fischeri, *E. coli* expressing the *luxCDABEGH* genes of *V. harveyi* does not emit high levels of light in a cell density dependant manner (Miyamoto et al., 1987). It has been established that the regulatory genes analogous to *luxI* and *luxR* of *V. fischeri* are not linked to the operon (Miyamoto et al., 1988b). The autoinducer of *V. harveyi* has been isolated and identified as N-ß-hydroxybutryl homoserine lactone (Cao and Meighen, 1989) and, in view of its similar structure to the autoinducer of *V. fischeri*, may function in the same manner.

A gene that is able to stimulate light production in *E. coli* containing the *V. harveyi lux* operon has recently been cloned and is designated as *luxR*, although there is no sequence homology to the *luxR* gene of *V. fischeri* (Martin et al., 1989; Showalter et al., 1990). This stimulation is not cell-density dependant and does not require autoinducer, yet *luxR* is required for light production in *V. harveyi*. In this paper we address the problem of why *E. coli* recombinants lacking *luxR* cannot produce high levels of light, and how *luxR* is able to complement this deficiency. Primer extension mapping of transcripts produced *in vivo* and *in vitro* demonstrated that LuxR is able to activate transcription at the *lux* promoter and causes a shift in the transcriptional start site by almost 100 bp.

MATERIALS AND METHODS

Bacterial strains and media

V. harveyi strain B392 was grown in 1% complex media (2 g of Na_2HPO_4, 1 g of KH_2PO_4, 0.5 g of $(NH_4)_2HPO_4$, 0.2 g of $MgSO_4 \cdot 7H_2O$, 5.0 g of tryptone [Difco Laboratories], 0.5 g of yeast extract [Difco], 10 g of NaCl, and 2 ml of glycerol per liter) at 27°. *E. coli* RR1 or 43R (Miyamoto et al., 1987) recombinants were grown in LB medium (10 g of tryptone [Difco], 5g of yeast extract [Difco], and 10 g of NaCl per liter) at 30°. *E. coli* transformed with pSau1U, pSau2U, or pSH was propagated in LB containing 100 μg of ampicillin per ml and *E. coli* cotransformed with pSH and pMR1403 was grown in LB containing ampicillin and 10 μg of tetracycline per ml.

Primer extension mapping

RNA purification and primer extension mapping was done as described before (Swartzman et al., 199b), except that a synthetic oligomer, a 31-mer complementary to nucleotides +4 to +33 of the *luxC* gene, was used as the primer and hybridization was carried out at 45°. A sequence ladder generated from the same primer served as size marker. RNA for the primer extension mapping was either synthesized *in vitro* (described below) or isolated from *V. harveyi* strain B392 or *E. coli* recombinants containing *lux* DNA. The same results were obtained when either *E. coli* RR1 or 43R (Miyamoto et al., 1987), a mutant strain that is able to support luminescence, was used. The cultures were harvested for RNA purification between $OD_{660}=1.5$ and $OD_{660}=2.0$. At these cell densities, *V. harveyi* is fully induced for luminescence. The plasmids pSau1U and pSau2U are subclones of pVhSau1 and pVhSau2 (Miyamoto et al., 1988a), respectively, cut and religated at the *Pvu*II site. pSH (Miyamoto et al., 1988a) and pMR1403 (Showalter et al., 1990), a cosmid containing the *luxR* gene, have been previously described. The plasmids pSau1U, pSau2U and pSH contain *V. harveyi lux* DNA, but only pSH contains all the structural genes required for light production. *E. coli* strain RR1 was cotransformed with pSH and pMR1403, as the 43R strain could not be transformed with both vectors.

In vitro transcription

A *Bam*HI-*Cla*I 1.0 kbp fragment spanning the 5'-non coding and coding regions of the *luxC* gene was used for all transcription reactions. DNA (100 ng) was preincubated with RNA polymerase (0.25 units of *V. harveyi* RNA polymerase purified in this laboratory or 0.5 units of *E. coli* RNA polymerase purchased from Pharmacia) and 1.5 μl 10 x transcription buffer in a total volume of 12.5 μl for 5 min at 37°. One μl of 2 mg/ml heparin was added to inactivate free RNA polymerase and the incubation continued for a further 5 min. The transcription reaction was started by the addition of 1.5 μl of 10 x nucleotide mixture and continued for 15 min. The reaction was stopped by phenol/chloroform extraction, and the RNA was ethanol precipitated with 50 μg of *E. coli* tRNA added as carrier. The transcripts were then processed as described above for primer extension.

The 10 x transcription buffer consists of 0.2 mg/ml BSA, 0.2 M Tris-acetate pH 8.0, 40 mM $MgCl_2$, 1.0 M KCl, 1.0 mM EDTA, 1.0 mM DTT, and 40% glycerol. The 10 x mix consisted of 1.5 mM each of ATP, CTP, GTP and UTP.

Other

A comparison of *lux* mRNA levels was done by densitometry scanning of the autoradiographs using Bio Image from Millipore.

RESULTS

Comparison of lux mRNA from V. harveyi and recombinant E. coli

In order to better understand the regulation of the *V. harveyi lux* promoter in *E. coli*, primer extension mapping was performed on mRNA isolated from *E. coli* cells transformed with three different *V. harveyi lux* clones (Fig. 1). Despite the variation in length of the 5' noncoding regions, the transcriptional start sites map to the same location, at nucleotide -123 with respect to the initiation codon of the *luxC* gene (Fig. 2B). In contrast, the 5' end of the *lux* mRNA isolated from *V. harveyi* maps to position -26 (Fig. 2A), a difference of 97 nucleotides. Although the transcriptional start site of the *V. harveyi lux* mRNA has been previously determined to be at position -28 (Swartzman and Meighen, 1990b), the new site is likely more accurate due to the better resolution of the primer extention product. Several smaller transcripts are noticeable for the reactions seen in figures 2A and 2B, however *lux* mRNA from *E. coli* recombinants was not observed to have a 5' end at -26 (see Fig. 4, lane 3). The multiple transcripts could be due to partial RNA degradation, incomplete extension of the primer, or internal transcription initiation. Consequently, the largest primer-extended product of each reaction was used to map the 5' end of the transcripts. There is some inconsistency in *lux* mRNA levels from the 3 different *E. coli* recombinants as seen in Figure 2B. These differences might be due to a variation of plasmid copy number or mRNA stability. Regardless, 10-fold more RNA and a 5-fold longer exposure time was required to visualize the primer-extended products of RNA from the *E. coli* recombinants in Figure 2B as compared to the RNA from *V. harveyi* in Figure 2A. It has already been shown that the amount of *lux* mRNA increases with induction of luminescence (Miyamoto et al., 1985; Showalter et al., 1990; Swartzman et al., 1990b). Accordingly, the low level of light produced by the *E. coli* recombinants can be attributed to the low abundance of *lux* mRNA.

Sequences upstream of the start sites were examined for consensus -10 (TATAAT) and -35 (TTGACA) *E. coli* promoter sequences (Hawley and McClure,

1983) (Fig 2C). The sequence TATAAT is located 7 bp upstream from the -123 start site, in complete agreement with the -10 hexamer with respect to position and sequence. The sequence TTACGA is located 34 bp upstream from the -123 start site, showing some homology to the -35 hexamer. The low level of transcription at this promoter *in vivo* may be a reflection of the imperfect -35 sequence. The -26 start site has the sequence ATTAAT 6 bp upstream from the start site, in good agreement with the -10 sequence, although no comparable -35 sequence could be located within this promoter.

Comparison of in vitro and in vivo transcripts

There is a possibility that the difference in transcriptional start sites is due to a difference in promoter specificity of the *E. coli* versus *V. harveyi* RNA polymerase. *V. harveyi* RNA polymerase has been purified and can replace *E. coli* RNA polymerase in an *in vitro* transcription assay (Swartzman and Meighen, in press). To map the exact 5' end of *lux* RNA synthesized *in vitro*, primer extension mapping was performed on transcripts produced *in vitro* and compared to RNA isolated from *E. coli* transformed with *lux* DNA (Fig. 3). *In vitro* transcripts produced by *V. harveyi* RNA polymerase (lane 1) or *E. coli* RNA polymerase (lane 2) mapped to the same location (at -123) as the *in vivo* transcript from *E. coli* (lanes 3 and 4). The difference seen in transcriptional start sites is therefore not due a difference in promoter specificity of the *E. coli* versus *V. harveyi* RNA polymerase. The lack of a recognizable -35 consensus sequence upstream of the -26 start site is an indication that this promoter will not be recognized by RNA polymerase alone. There must be a factor (or factors) missing in the *in vitro* transcription assay as well as in *E. coli* that is essential for transcription initiation at the -26 *lux* promoter.

LuxR is a transcriptional activator

Because *luxR* has been shown to stimulate luminescence in recombinant *E. coli* containing the *V. harveyi lux* operon (Showalter et al., 1990), the effect of *luxR* on

Figure 1: Restriction map of *V. harveyi lux* operon. Primer extension mapping was performed on RNA isolated from *E. coli* transformed with vectors 1) pVhSau1U, 2) pVhSau2U, and 3) pSH, containing the indicated inserts of *V. harveyi lux* DNA. The *Bam*HI-*cla*I 1.0 kbp DNA fragment used in the *in vitro* transcription assay is indicated by a thick line within the restriction map. C, *Cla*I; H, *Hin*dIII; B, *Bam*HI; S, *Sac*I; U, *Pvu*II.

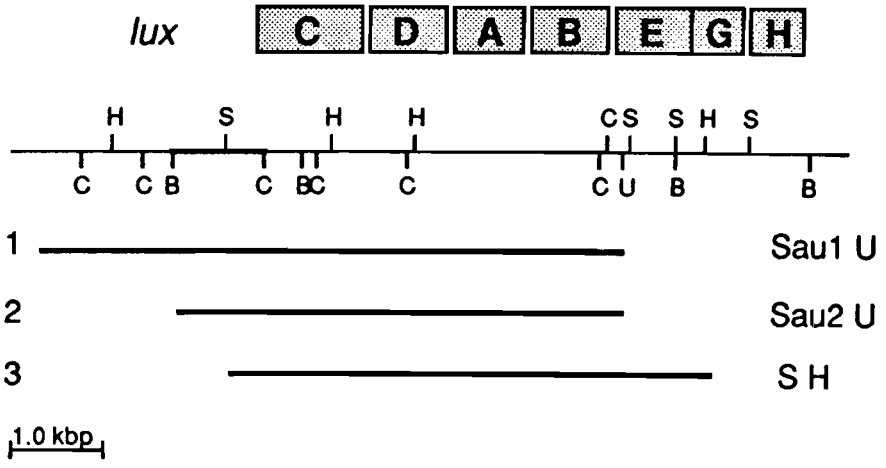
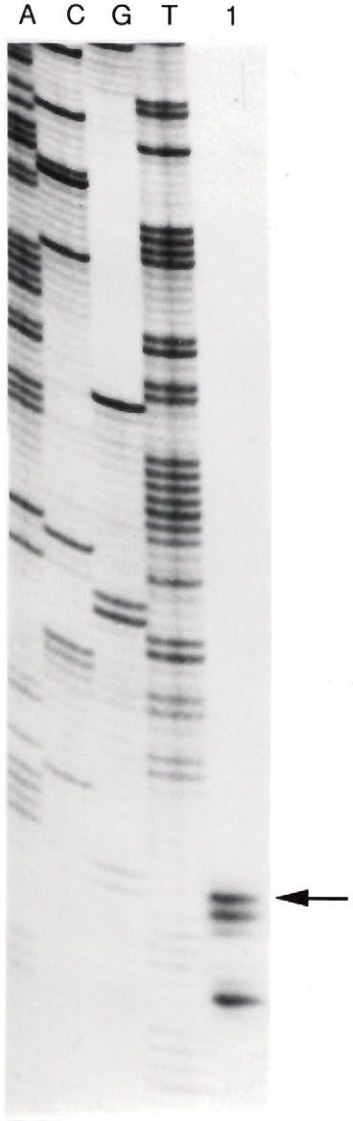
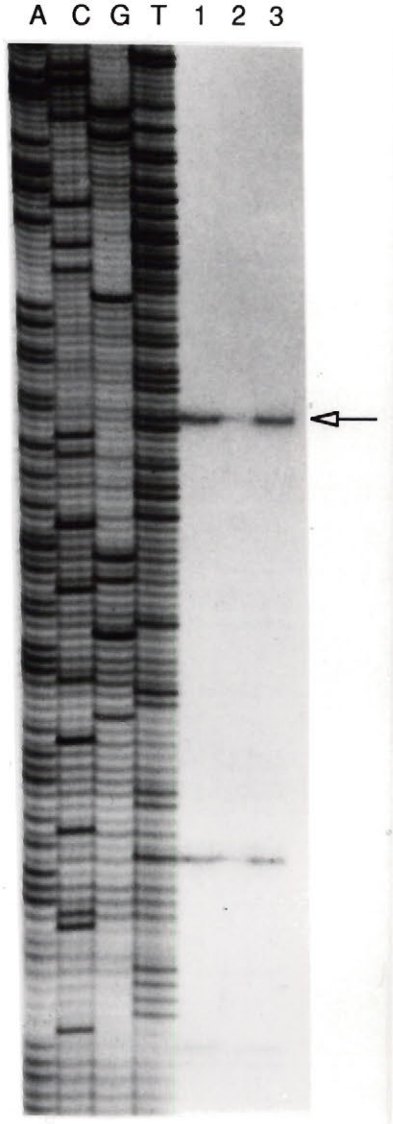

lux
C
D
A
B
E
G
H
H
S
H
H
C S
S
H
S
C
C B
C
B C
C
C U
B
B
1
Sau1 U
2
Sau2 U
3
S H
1.0 kbp

Figure 2) Primer extension mapping of *lux* mRNA. A) Lane 1 is the primer extension product using 10 μg of RNA isolated from *V. harveyi*. B) Primer extension products using 100 μg of RNA isolated from *E. coli* transformed with: 1) pVhSau1U, 2) pVhSau2U, and 3) pSH. The primer used in both A and B was used to generate the sequence ladders in order to accurately locate the 5' end of the primer extension products. C) Nucleotide sequence of 5'-end of *luxC* gene and upstream region. The first nucleotide of *luxC* is designated as +1. The filled arrow indicates the 5' end of *lux* mRNA isolated from *V. harveyi* and the hollow arrow indicates the 5' end of *lux* mRNA isolated from *E. coli* transformed with *lux* DNA. The underlined sequences correspond to the -10 and -35 promoter consensus sequences. The DNA sequence was obtained from Miyamoto *et al.* (Miyamoto et al., 1988b).

A



B



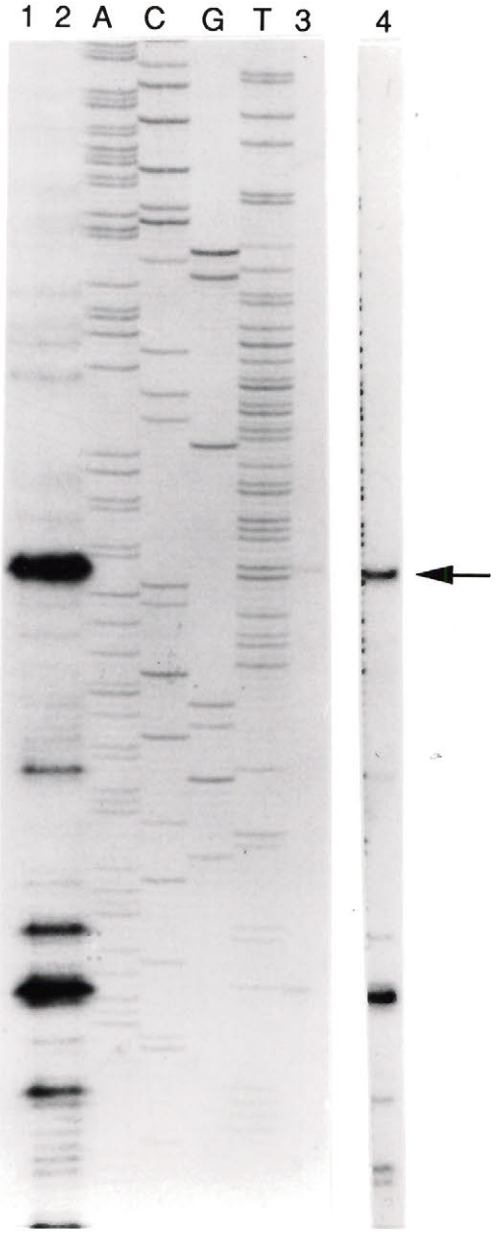
C

-170 GTTAAGTGTT ACGACTAATT ATAGATAAGA AGAACTATAA TTAAATTAAG TGATAATAGT TCTCGTTACT
-35 -10 -123

-100 TTGAACTGTT TAATGTATTT GGTTAAAAGT TTTTAATTAA CTTTAAAAAA ATGATCCAAG GAATTAATGT
-10

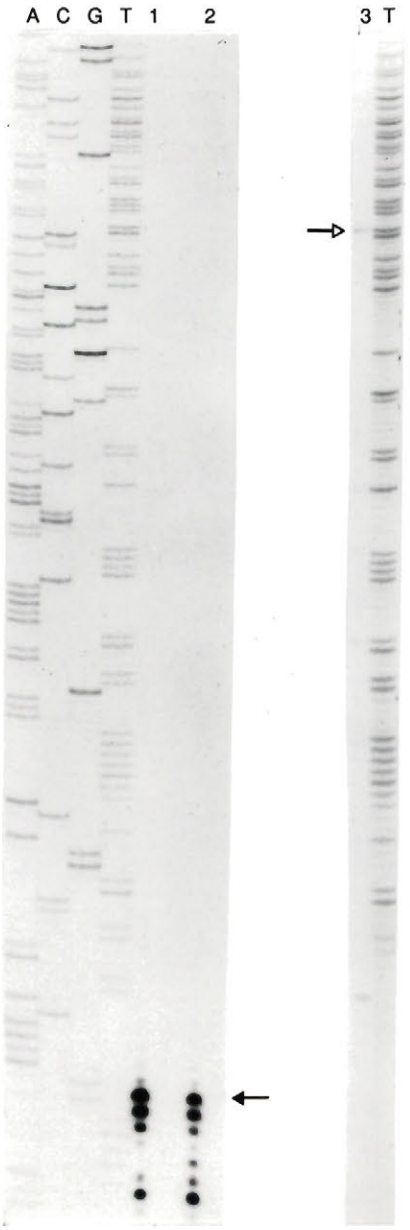
+1
-30 TTTCCAAAAT TTAAAAGAGA AGCTCTTGAT ATG GAA AAA CAC TTA CCT TTA ATA GTA AAT GGA
-26

Figure 3: Comparison of *in vitro* and *in vivo* transcription. Primer extension mapping was performed on RNA synthesized *in vitro* with 1) *E. coli* RNA polymerase and 2) *V. harveyi* RNA polymerase as well as with 3) 100 μg of RNA isolated from *E. coli* transformed with pSH. Lane 4 is an overexposure of lane 3. The same primer used in the extension reactions was used to generate the sequence ladder. The arrow indicates the position of the full length primer extension products.


1 2 A C G T 3
4

transcription of the *lux* promoter was tested *in vivo*. RNA isolated from *E. coli* transformed with both the *lux* operon and *luxR* in *trans* was analyzed by primer extension and was compared with RNA from *V. harveyi* and from *E. coli* containing just the *lux* operon. In Figure 4, lane 1, the 5' end of the *lux* RNA from *E. coli* recombinants containing both the *lux* operon and *luxR* maps to -26, the same start site as for *V. harveyi* RNA (lane 2) and different from that seen for RNA isolated from *E. coli* containing just the *lux* operon (-123, lane 3). Expression of *luxR* in *E. coli* is therefore responsible for the shift in transcriptional start site from -123 to -26. Moreover, it was determined that there was 350-fold more *lux* mRNA in lane 1 than in lane 3, indicating that the shift of start site by LuxR is responsible for increasing levels of *lux* mRNA in *E. coli*.

Figure 4: Primer extension mapping of *lux* RNA with and without expression of *luxR*. Primer extension mapping was performed on 1) 10 μg of RNA isolated from *E. coli* cotransformed with pSH and pMR1403, 2) 10 μg of RNA isolated from *V. harveyi*, and on 3) 150 μg of RNA isolated from *E. coli* transformed with pSH. The hollow arrow indicates the -123 start site and the filled arrow indicates the -26 start site. Lanes A, C, G, and T contain the sequence ladder obtained using using the same primer as in the primer extension reactions, while only the T tract of the same sequence reaction is shown next to lane 3.


A C G T 1 2
3 T

DISCUSSION

Autoinducer-dependant regulation of the *V. harveyi lux* operon has never been reconstituted in recombinant *E. coli* and, as a result, the mechanism of transcriptional regulation of the *V. harveyi lux* promoter has remained elusive. Transposon mutagenesis of *V. harveyi* led to the isolation and cloning of *luxR*, a gene that is unlinked to the *lux* operon and yet is required for bioluminescence (Martin et al., 1989; Showalter et al., 1990). Because expression of *luxR* in *E. coli* recombinants significantly enhances light production (Showalter et al., 1990), luxR is thought to be an activator of bioluminescence.

It was necessary to first understand why *E. coli* recombinants containing just the *lux* operon could not produce high levels of light in order to determine how *luxR* functions as an activator. A comparison of 5' ends of transcripts produced *in vivo* and *in vitro* showed that the transcriptional start site used in the *E. coli* recombinant containing just the *lux* operon was at -123, almost 100 bp upstream form the start site used in *V. harveyi*. This difference is independent of the length of the 5' untranscribed region, eliminating the possibility of missing *cis*-activating elements upstream of the *luxC* gene. This difference is also not due to a difference in σ subunit specificity of *V. harveyi* and *E. coli* RNA polymerase as both enzymes initiate transcription at the same site *in vitro*.

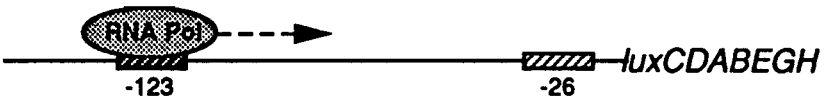
Because the transcripts produced *in vitro* had the same 5' end as the transcripts produced by the *E. coli* recombinants, it was concluded that there was a factor missing in *E. coli* as well as in the *in vitro* assay that was required for proper initiation at the *V. harveyi lux* promoter. Analysis of RNA isolated from *E. coli* recombinants containing both the *lux* operon and the *luxR* gene revealed that *luxR* was essential for accurate transcription of the *V. harveyi lux* promoter. In addition to increasing levels of *lux* mRNA, the expression of *luxR* in *E. coli* shifted the transcriptional start site by almost 100 bp to the -26 promoter site seen in *V. harveyi*. Without *luxR*, transcription initiation only occurs at the weak -123 promoter both *in vivo* and *in vitro*. LuxR is a transcriptional regulator that is required for RNA

polymerase to activate the otherwise non-functional -26 promoter (Fig. 5).

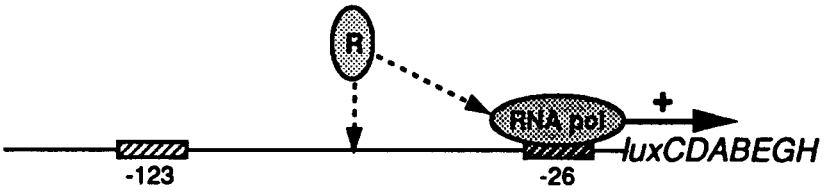
The possibility that the contrast seen in the 5' ends of the transcripts produced *in vivo* is caused by specific RNA processing seems unlikely, particularly as the deduced amino acid sequence of *luxR* was found to contain a Cro-like DNA-binding domain that is common to many prokaryotic transcription regulators (Showalter et al., 1990), and no sequence homology to any RNAses or RNA-binding proteins was found. There is also evidence that LuxR binds specifically to DNA sequences upstream of the *luxC* gene (unpublished data). Although the mechanism of activation remains elusive, it has now been established that LuxR of *V. harveyi* is an activator of the *V. harveyi lux* promoter.

Figure 5: Model of transcription initiation at the *V. harveyi lux* promoter. (A) In the absence of LuxR, RNA polymerase is unable to bind to the -26 promoter and instead initiates transcription at the weak -123 promoter. (B) LuxR stimulates the binding of RNA polymerase to the -26 promoter (or stimulates open compex formation) and activates transcription of the *lux* operon, while the -123 promoter in *V. harveyi* is ineffective and most likely has no functional significance. LuxR has a putative DNA-binding domain and may mediate transcription activation through LuxR-DNA and/or LuxR-RNA polymerase interactions.

A



B



CHAPTER 6

PURIFICATION OF A POLY(dA-dT) *LUX*-SPECIFIC DNA-BINDING PROTEIN FROM *VIBRIO HARVEYI* AND IDENTIFICATION AS LUXR

ABSTRACT

A *lux*-specific DNA-binding protein was purified from *Vibrio harveyi* and identified as LuxR, a transcriptional regulator of the *V. harveyi lux* operon. LuxR bound to two regions of DNA upstream of the *V. harveyi lux* promoter, region A: -290 to -253, and region B: -116 to -170, relative to the initiation codon of *luxC*. Both regions contain the octamer AGTGTTAC and both are A + T rich. Synthetic poly(dA-dT) but not poly(dA)-poly(dT) competed with the *lux* DNA for binding to LuxR, suggesting the LuxR is a novel prokaryotic poly(dA-dT) binding protein. As *in vitro* transcription assays with LuxR did not reconstitute *lux* promoter activation, it appears that another factor or factors is/are required for transcription initiation from the *V. harveyi lux* promoter. A model of transcriptional regulation of the *V. harveyi lux* operon is presented.

INTRODUCTION

The *Vibrio harveyi luxCDABEGH* operon encodes all the enzymatic functions required for light production, yet lacks the regulatory controls necessary for growth-dependent luminescence in *Escherichia coli* (Miyamoto et al., 1987). The different strategies employed in the past to search for the missing regulatory elements have met with varying degrees of success. Attempts to complement *E. coli* already containing the *luxCDABEGH* genes with a library of *V. harveyi* DNA failed to produce any recombinants exhibiting cell density-dependent luminescence (Miyamoto et al., 1990). Mutagenesis of the *V. harveyi* genome proved to be more effective in identifying loci required for light production. A gene designated as *luxR* was shown to be required for expression of *luxCDABEGH* mRNA, yet is unlinked to the structural *lux* genes. When expressed in recombinant *E. coli in trans* with the *luxCDABEGH* operon, *luxR* stimulated luminescence by at least 10,000-fold (Martin et al., 1989; Showalter et al., 1990). However, this stimulation did not exhibit the characteristic cell density-dependent modulation, nor did it respond to exogenously added autoinducer. Although the sequence of *V. harveyi* LuxR did not show any significant identity with LuxR of *V. fischeri*, it does contain a Cro-like DNA-binding domain consistent with it being a transcriptional regulator. Another approach to investigate the differences of *lux* gene regulation in *E. coli* and *V. harveyi* was to examine the transcriptional start sites *in vivo* and *in vitro* (Swartzman and Meighen, in press; Swartzman et al., in press). The transcriptional start site of *lux* mRNA in recombinant *E. coli* and in an *in vitro* transcription assay using *V. harveyi* or *E. coli* RNA polymerase mapped to the same site, at -123 bp upstream of *luxC*. This is in contrast to the start site found in *V. harveyi*, located at position -26. It was concluded that the *in vitro* assay as well as *E. coli* lacked a crucial factor required for accurate transcription initiation. The missing element was identified as LuxR, as expression of *luxR in trans* with *luxCDABEGH* in *E. coli* shifted the transcriptional start site from -123 to -26.

Mobility shift assays were utilized in an effort to search for other regulatory

proteins that may be involved in *lux* gene regulation and resulted in the identification of a *lux*-specific DNA-binding protein found exclusively in extracts of *V. harveyi* (Swartzman et al, 1991). In this study, we examine the properties of the DNA-binding activity and purify the protein to homogeneity. It has a high affinity for alternating A + T residues and binds to two distinct regions of *lux* DNA. Amino-terminal sequencing and functional assays reveal that the DNA-binding protein is the transcriptional activator LuxR.

MATERIALS AND METHODS

Bacterial strains and media

V. harveyi strain B392 was grown in 1% complex media (2 g of Na_2HPO_4, 1 g of KH_2PO_4, 0.5 g of $(NH_4)_2HPO_4$, 0.2 g of $MgSO_4 \cdot 7H_2O$, 5.0 g tryptone [Difco Laboratories], 0.5 g of yeast extract [Difco], 10 g of NaCl, and 2 ml of glycerol per litre) at 27°C. *V. harveyi* strain BB7 (wild type and mutants) were grown in beef heart fusion medium (25 g heart infusion broth [Difco] and 20 g NaCl per litre) (Martin et al., 1989). *E. coli* DH5α recombinant containing pMR1403 (Showalter et al., 1990) was grown in LB medium (10 g of tryptone, 5g of yeast extract, and 10 g of NaCl per litre) containing 10 μg of tetracycline per ml at 30°C.

The transposon-generated mutants strains of *V. harveyi* are desribed elsewhere (Martin et al., 1989) and the cosmid pMR1403 contains the *V. harveyi luxR* gene (Showalter et al., 1990).

Gel retardation assay

The DNA-binding assay was performed according to established procedures (Fried and Crothers, 1981). The 20 μl reaction mixture, containing approximately 1 ng of a 5' ^{32}P-end labelled DNA fragment (1000 dpm) and the indicated amount of protein extract in 10 mM Hepes pH 7.9, 100 mM KCl, 2 mM DTT, 0.4 mM PMSF, 0.2 mM EDTA, 10% glycerol, and 3 μg of poly(dI-dC) (Pharmacia), was incubated for 30 min at room temperature. After adding bromophenol blue and xylene cyanol FF to 0.025%, the sample was loaded onto a 5% polyacrylamide gel and electrophoresed in 0.5 % TBE at 15 milliamps.

DNase I Footprinting

The protein-DNA mixture was incubated at room temperature as described above, except 10 ng (10,000 dpm) of a 5' ^{32}P-labelled DNA fragment was used. After 30 min, 1 μl of 0.1 M $MgCl_2$ was added and the mixtures were placed on ice. The reaction was started by the addition of 1 μl of 40 μg of DNase I (BRL) per ml

and stopped after 30 s with 80 μl of stop buffer (20 mM Tris-Cl pH 7.9, 20 mM EDTA, 250 mM NaCl, 0.5% SDS, and 1 μg of tRNA). After phenol/chloroform extraction and ethanol precipitation, each sample was dissolved in loading buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF), boiled for 2 min, and loaded onto a 6.5% sequencing gel. A sequence ladder of M13mp18 served as size markers.

Preparation of DNA fragments

Fragment 1, extending from nucleotides -212 to -465 relative to the initiation codon of *luxC* (+1) was obtained from a subclone of *V. harveyi* upstream *lux* DNA in M13mp18 used for DNA sequencing that had been subjected to exonuclease III digestion (Miyamoto et al., 1988b). A 270 bp fragment was excised with *Bam*HI and *Eco*RI, each cutting within the polylinker flanking the *lux* DNA. For gel retardation assays, both sites were ^{32}P-labelled, and for the DNase I protection assay, the *Eco*RI site was ^{32}P-labelled using the large fragment of DNA polymerase I. Fragment 2 was obtained by isolating a 525 bp *Sac*I-*Cla*I insert extending from -465 to +58 and was ^{32}P-labelled at the *Cla*I site for footprint analysis. The DNA fragments were isolated and purified from low melting agarose (Seakem).

Preparation of cell extracts

Extracts were prepared from cells harvested at OD_{660} = 1.5 for *E. coli* and cultures of *V. harveyi* after luminescence induction and from cells harvested at OD_{660} = 0.3 for cultures of *V. harveyi* before luminescence induction. The cell pellets were resuspended in 25 ml of lysis buffer (50 mM Hepes pH 7.9, 2 mM EDTA, 1 mM DTT, 1 mM PMSF, 5% glycerol) per litre of cell culture and frozen in a dry ice/methanol bath followed by thawing at 37°C twice before addition of 12.5 mg of lysozyme per litre of cell culture. After gentle mixing at 4°C for 30 min, solid NaCl was added to a 1 M final concentration and gentle mixing continued for an additional 30 min. The lysed cells were centrifuged at 44, 000 *g* for 1 h (for large volumes) or

in an Eppendorf centrifuge for 30 min (for small volumes) to obtain clear lysates. One tenth volume of saturated ammonium sulfate (3.5 M, adjusted to pH 7.9 with NaOH) and solid ammonium sulfate (0.33 g per ml of lysate) was added and mixed for 1 h at 4°C to precipitate protein. After centrifugation as above, the pellets were resuspended in approximately 10 ml of dialysis buffer (20 mM Hepes pH 7.9, 100 mM KCl, 1 mm DTT, 0.2 mM PMSF, 0.2 mM EDTA, 20% glycerol) per litre of cell culture, and dialysed overnight against the same buffer. The cell extracts could then be stored indefinitely at -70°C.

Protein purification

The DNA-binding activity was purified from 8 litres of *V. harveyi* cells harvested at OD_{660}=1.5. The cells were lysed essentially as described above, except protein that precipitated between 25% to 80% ammonium sulfate saturation was collected and dialysed against buffer A (10 mM Tris-Cl pH 7.9, 100 mM KCl, 1 mM DTT, 0.1 mM EDTA, 5% glycerol).

The lysate (150 ml) was loaded onto a 100 ml DEAE-Sepharose column (Pharmacia) and the column washed with buffer A until the OD_{280} was under 0.05 (approximately 300 ml). Protein was eluted using a 250 ml gradient of buffer A from 0.1 M KCl to 1.0 M KCl. Fractions with DNA-binding activity were pooled and dialysed overnight against 2 litres of buffer A.

The dialysate was loaded onto a 40 ml dsDNA-cellulose column (Sigma) and the column was washed with buffer A until the OD_{280} was under 0.5 (approximately 120 ml). Protein was eluted using a 100 ml gradient of buffer A from 0.1 M KCl to 1.0 M KCl. Fractions with DNA-binding activity were pooled and dialysed overnight against 2 litres of saturated ammonium sulfate (>4M, adjusted to pH 7.9 with NaOH).

The precipitate was collected by centrifugation and resuspended in buffer A containing 0.5 M KCl to a final volume of 200 μl. The concentrated dsDNA-cellulose pool was then applied onto and eluted from an Superose 12 gel filtration column (Pharmacia) using fast performance liquid chromatography (FPLC;

Pharmacia) using buffer A containing 0.5 M KCl. The active fractions were pooled, diluted to 5 ml with buffer A lacking KCl, and applied onto a mono Q column (Pharmacia) using FPLC. A shallow gradient from 0.34 M KCl to 0.48 M KCl in buffer A was used to elute the DNA-binding activity. Fractions were pooled, dialysed against buffer A containing 50% glycerol and stored at -20°C.

Other

The *in vitro* transcription assay was performed as described (Swartzman and Meighen, in press). Protein concentration was determined by the modified Lowry assay (Markwell et al., 1981). SDS-PAGE gels (4% stacking, 12% resolving) were prepared by the method of Laemmli (1970) and stained with Kodak Coomassie blue R-250. N-terminal sequence analysis was done at the Sheldon Biotechnology Centre, McGill University.

RESULTS

Gel Retardation Assays

The ability of a protein in cell extracts of *V. harveyi* to bind to *lux* DNA containing putative regulatory sites upstream of the *lux* promoter was characterized with respect to luminescence induction and the extent of cell-growth. Extracts were prepared from cells of *V. harveyi* before and after induction of luminescence and different amounts were tested for their ability to bind to the *lux* DNA fragment located between -465 and -212 nucleotides from the *luxC* initiation codon. Although no difference can be observed between the cell extracts at high protein concentration, a small difference can be seen when as little as 0.5 ug of protein is used (Figure 1 A), which is sufficient for binding by protein extacted from *V. harveyi* after induction of luminescence but not for binding by protein extacted from *V. harveyi* before induction of luminescence. This indicates that the synthesis of the DNA-binding factor in *V. harveyi* may be induced to some degree with cell-growth.

The DNA upstream of the *luxC* gene, including the fragment used in the gel retardation assay, is very A + T rich (Miyamoto et al., 1988b). It was of interest to determine whether the binding could be abolished with different A + T polymer DNAs, as has been shown for other some other DNA-binding proteins (Garreau and Williams, 1983; Levinger and Varshavsky, 1982; Winter and Varshavsky, 1989; Solomon et al., 1986). Poly(dA-dT) but not poly(dA)-poly(dT) was able to successfully compete out *lux* DNA binding (Figure 1B), demonstrating that alternating A + T polymers are important for specific DNA-protein interactions.

DNase I Footprinting

The region of DNA bound by the protein in the *V. harveyi* extract was located by DNA footprinting analysis. Using the same DNA probe as in the mobility shift assays, the protein protects a region of DNA spanning from -253 to -290 from DNase I digestion (Figure 2, lane 2). The protection is competed out with poly(dA-dT) (Figure 2, lane 4), and does not occur when a cell extract prepared from *V. harveyi*

Figure 1: Gel retardation assays. (A) The indicated amounts of protein extacted from *V. harveyi* before and after luminescence induction were combined with ^{32}P-labelled fragment 1 (see figure 4) and electrophoresed as described in Materials and Methods. Lanes 1 and 7 contain no added protein. (B) Increasing amount of poly(dA-dT) and poly(dA)-poly(dT) were added as competitor to reaction mixtures containing 2 μg of protein extracted from *V. harveyi* after luminescence induction and ^{32}P-labelled fragment 1. Lanes 1 and 7 contain no added protein or competitor DNA, while lanes 2 and 8 contain no competitor DNA.

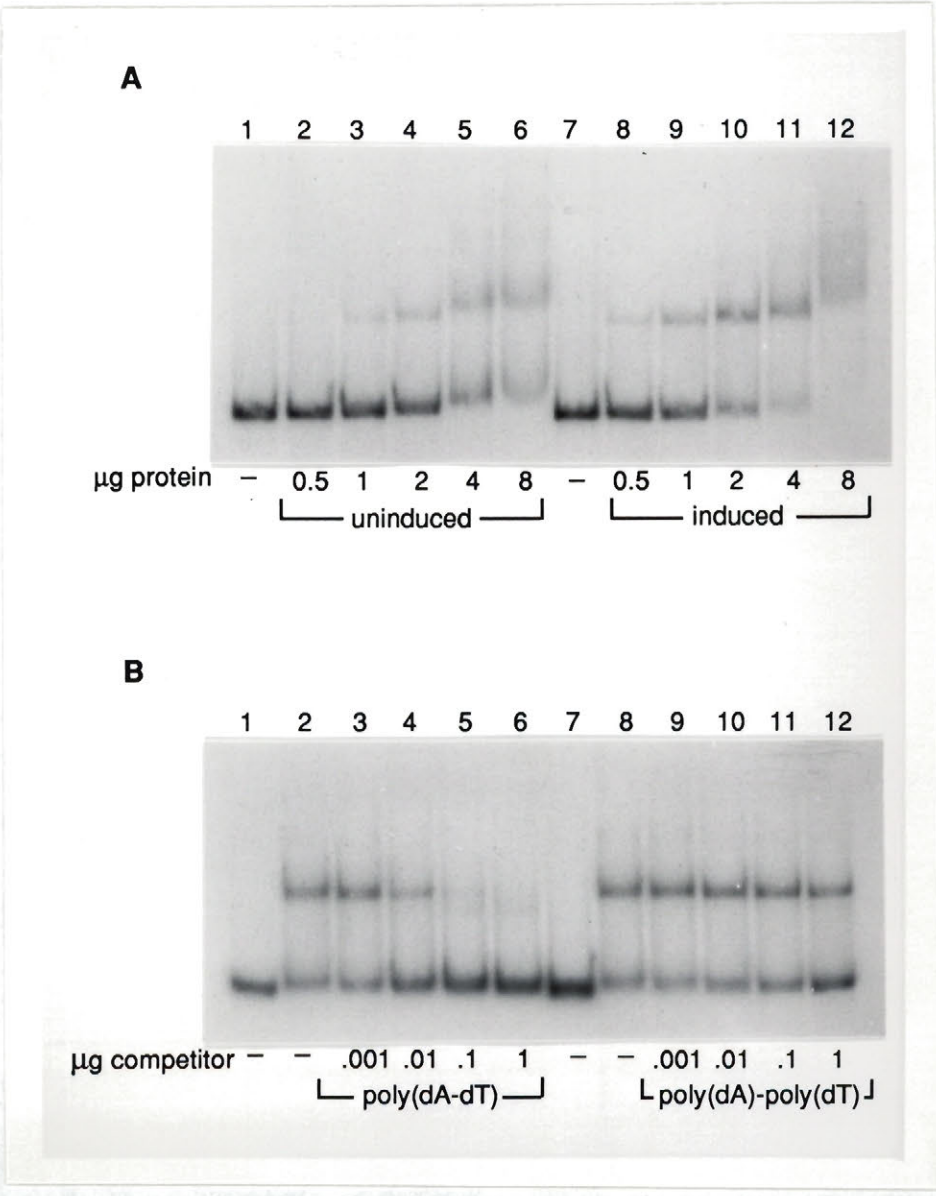
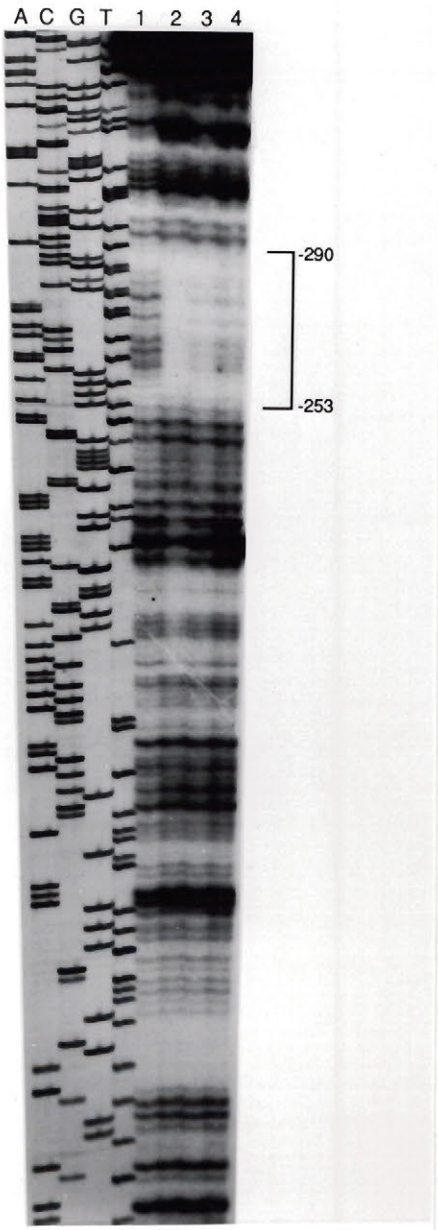

A
1 2 3 4 5 6 7 8 9 10 11 12
µg protein – 0.5 1 2 4 8 – 0.5 1 2 4 8
uninduced
induced
B
1 2 3 4 5 6 7 8 9 10 11 12
µg competitor – – .001 .01 .1 1 – – .001 .01 .1 1
poly(dA-dT)
poly(dA)-poly(dT)

Figure 2: DNase I protection analysis. DNase I digest of 5' end-labelled fragment 1 (see figure 4) with 10 μg of BSA (lane 1), 10 μg of cell extract from *V. harveyi* after luminescence induction (lane 2), 10 μg of cell extract from *V. harveyi* before luminescence induction (lane 3), and 10 μg of cell extract from *V. harveyi* after induction of luminescene plus 1 μg of poly(dA-dT) (lane 4). The protected region is indicated by the bracket, from -253 to -290 upstream from the initiation codon of *luxC*.


A C G T 1 2 3 4
-290
-253

before luminescence induction is used in the assay (Figure 2, lane 3).

A larger fragment of DNA extending from -465 to +58 was then used in the footprinting assay to locate any other protein-binding sites. With the second DNA probe, two regions of DNA were protected from DNase I digestion (Figure 3). Region A (-253 to -290) is the same site located using the smaller DNA probe, while region B spans from -116 to -170. The characteristics of protein-binding for both regions are the same: binding is competed out with poly(dA-dT) and protein extracted from *V. harveyi* before induction of luminescence fails to bind. It is therefore likely that a single protein is responsible for binding to both regions of DNA.

The nucleotide sequence of both binding sites was examined for any similarities (Figure 4). Although the sizes of the regions differ (37 for region A versus 55 for region B), both are A + T rich, and both contain the octamer AGTGTTAC. It is interesting to note that the *in vitro* and recombinant *E. coli* transcription start site at -123 is located within region B.

Purification of the DNA-Binding Activity

The gel retardation assay was used to monitor the DNA-binding activity during the purification. Figures 5, 6 and 7 illustrate the protein elution and activity profiles during DEAE-Sepharose, DNA-cellulose, and gel filtration chromatography, respectively. A final column using anion exchange chromatography on mono Q was employed to concentrate the protein and remove any residual contaminants (data not shown). The recovery of the DNA-binding protein is presented in Table 1 and displayed in Figure 8. The protein migrates in SDS-Page gels with a molecular mass of 23 kDa. Migration of the purified protein through a gel filtration column compared with those of known standards (Figure 9) revealed that the native protein has a molecular mass of 48 kDa and most likely exists as a dimer in solution. The purified protein bound to both regions A and B in the footprint assay (data not shown), confirming that this is the DNA-binding protein first identified in *V. harveyi* extracts.

Figure 3: DNase I protection analysis. DNase I digest of 5' end labelled fragment 2 (see figure 4) with 10 μg of BSA (lane 1), 10 μg of cell extract from *V. harveyi* after luminescence induction (lane 2), 10 μg of cell extract from *V. harveyi* before luminescence induction (lane 3), and 10 μg of cell extract from *V. harveyi* after luminescence induction plus 1 μg of poly(dA-dT) (lane 4). The protected regions are indicated by the brackets, region A extending from -253 to -290, and region B extending from -116 to -170.

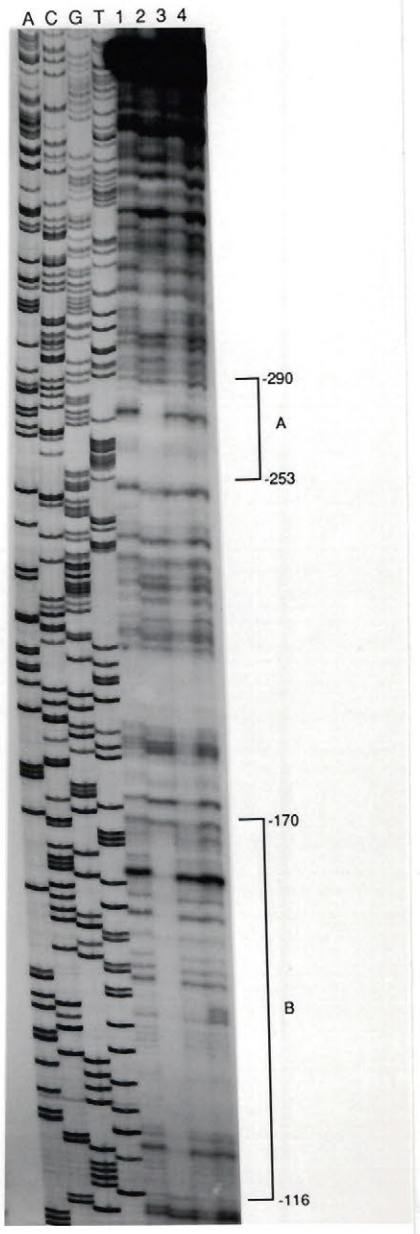
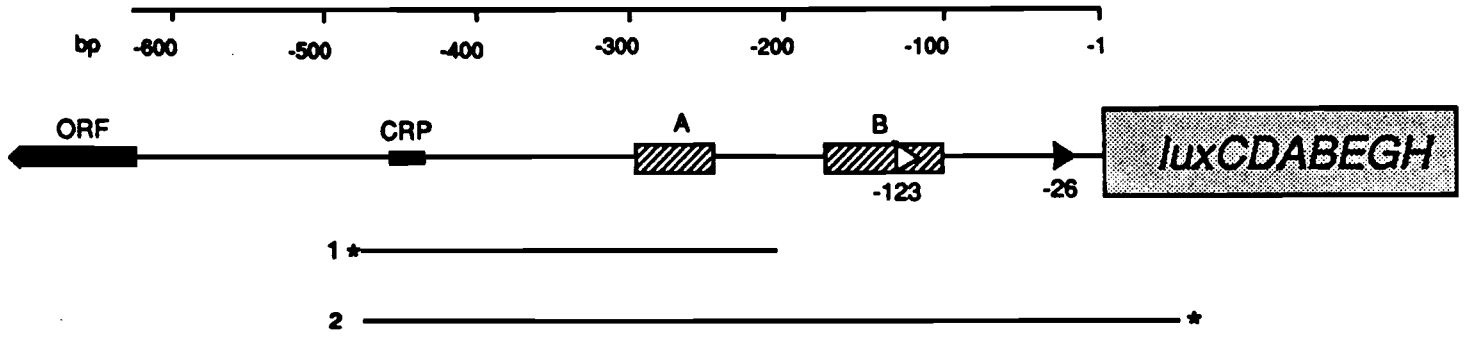

A C G T 1 2 3 4
-290
A
-253
-170
B
-116

Figure 4: (A) Schematic diagram of DNA-binding regions. The regions of DNA bound by protein in the *V. harveyi* cell extract are indicated by boxes A and B. The putative CRP recognition sequence is indicated as well as the -26 (filled arrow) and -123 (empty arrow) transcription initiation sites. The fragments used in the gel retardation and footprinting assays are depicted, with the star illustrating the location of the 5' ^{32}P-label. (B) Nucleotide sequence of *V. harveyi* upstream *lux* DNA. The regions of DNA bound by protein are boxed and labelled A and B. The repeated octamer within each region is underlined and the -26 (filled arrow) and -123 (hollow arrow) transcription initiation sites are indicated. The first nucleotide of the *luxC* gene is designated as +1.

A



B

```
                                                                A
-320 TCTAGTAGTA ACAACTACTT AAACATCATT ACTTATTTAT AGTGTTACAA ATAACATTAA TAATTTATGC AATATTTATG AGTATGATTC
                                                                                                B
-230 CGCTAGTGTT TAATAGCGCT GATAAAATCA AAAACTAAAG TGTTTGAAGT TTAGTTATTT GTTAAGTGTT ACGACTAATT ATAGATAAGA
-140 AGAACTATAA TTAAATTAAG TGATAATAGT TCTCGTTACT TTGAACTGTT TAATGTATTT GGTTAAAAGT TTTTAATTAA CTTTAAAAAA
                         △                                                  +1
 -50 ATGATCCAAG GAATTAATGT TTTCCAAAAT TTAAAAGAGA AGCTCTTGAT ATG GAA AAA CAC TTA CCT TTA ATA GTA AAT GGA
                         ▲
```

Figure 5: DEAE-Sepharose chromatography. (A) Elution profile of the DEAE- Sepharose column. Fractions of 2.3 ml were collected and protein concentrations were monitored by absorption at OD_{280}. (B) Gel retardation assay using fragment 1 and 1 μl of a 1:10 dilution of the indicated fractions. Fractions 51 to 65 were pooled.

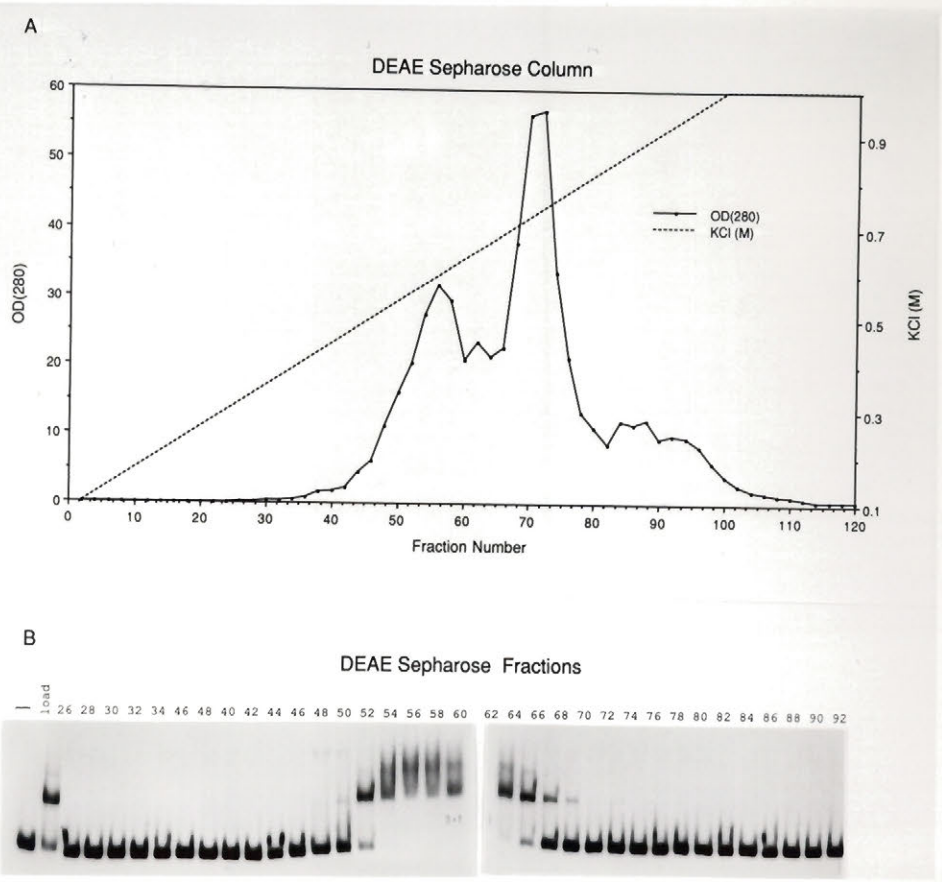
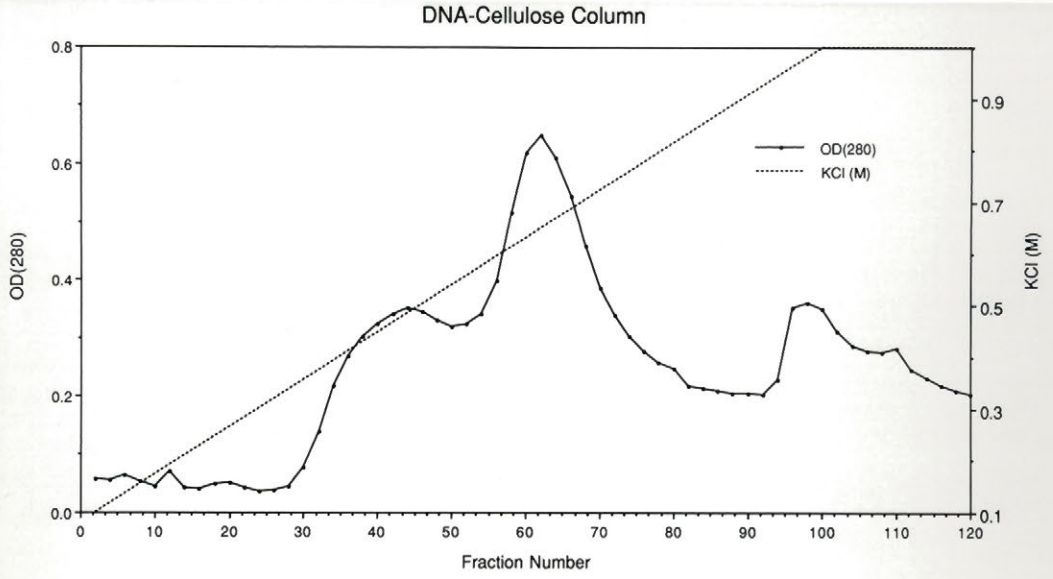

A
DEAE Sepharose Column
OD(280)
KCl (M)
Fraction Number
B
DEAE Sepharose Fractions
load
26 28 30 32 34 46 48 40 42 44 46 48 50 52 54 56 58 60 62 64 66 68 70 72 74 76 78 80 82 84 86 88 90 92

Figure 6: dsDNA-cellulose chromatography. (A). Elution profile of the dsDNA-cellulose column. Fractions of 1.0 ml were collected and protein concentrations were monitored by absorption at OD_{280}. (B) Gel retardation assay using fragment 1 and 1 μl of a 1:20 dilution of the indicated fractions. Fractions 58 to 68 were pooled.


DNA-Cellulose Column
OD(280)
KCl (M)
Fraction Number
0.0
0.2
0.4
0.6
0.8
0.1
0.3
0.5
0.7
0.9
0
10
20
30
40
50
60
70
80
90
100
110
120

B

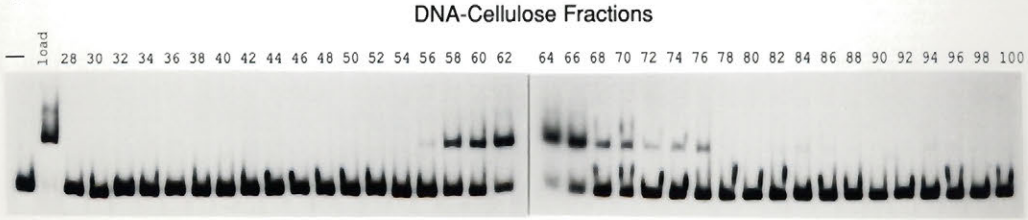
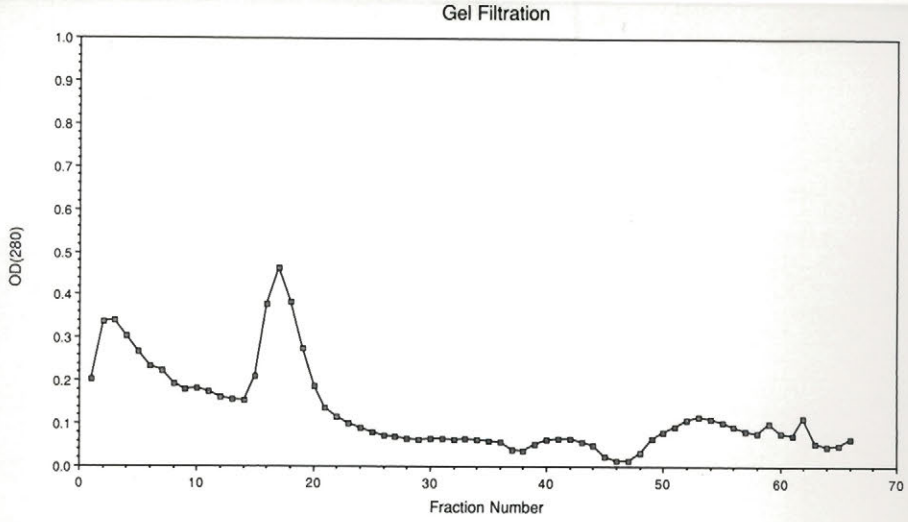

DNA-Cellulose Fractions
load
28 30 32 34 36 38 40 42 44 46 48 50 52 54 56 58 60 62
64 66 68 70 72 74 76 78 80 82 84 86 88 90 92 94 96 98 100

Figure 7: Gel filtration. (A) Elution profile of the Superose 12 column. Fractions of 0.25 ml were collected and protein cincentrations were minitored by absorbtion at OD_{280}. (B) Gel retardation assay using fragment 1 and 3 μl of a 1:20 dilution of the indicated fractions. Fractions 16 to 20 were pooled.


Gel Filtration
OD(280)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
10
20
30
40
50
60
70
Fraction Number

B

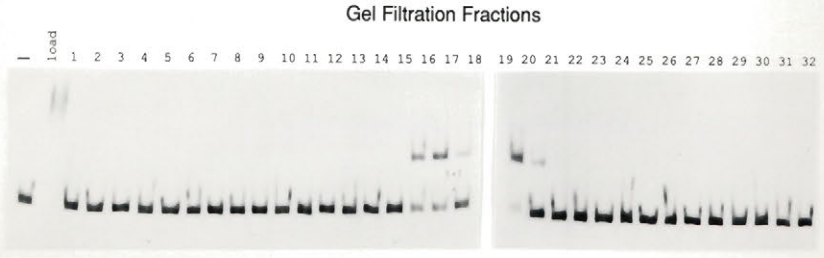

Gel Filtration Fractions
—
load
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32

Figure 8: SDS-polyacrylamide gel analysis of fractions obtained during purification of DNA-binding activity. Lane 1, 25 μg of cell lysate; lane 2, 55 μg of dialysed DEAE pool; lane 3, 21 μg of concentrated DNA cellulose pool; lane 4, 1.8 μg of concentrated mono Q pool. The molecular weight standards (Sigma) are BSA (66 kDa); ovalbumin (45 kDa); glyceraldehyde dehydrogenase (36 kDa); carbonic anhydrase (29 kDa); trypsinogen (24 kDa); trypsin inhibitor (20 kDa); α-lactalbumin (14 kDa).

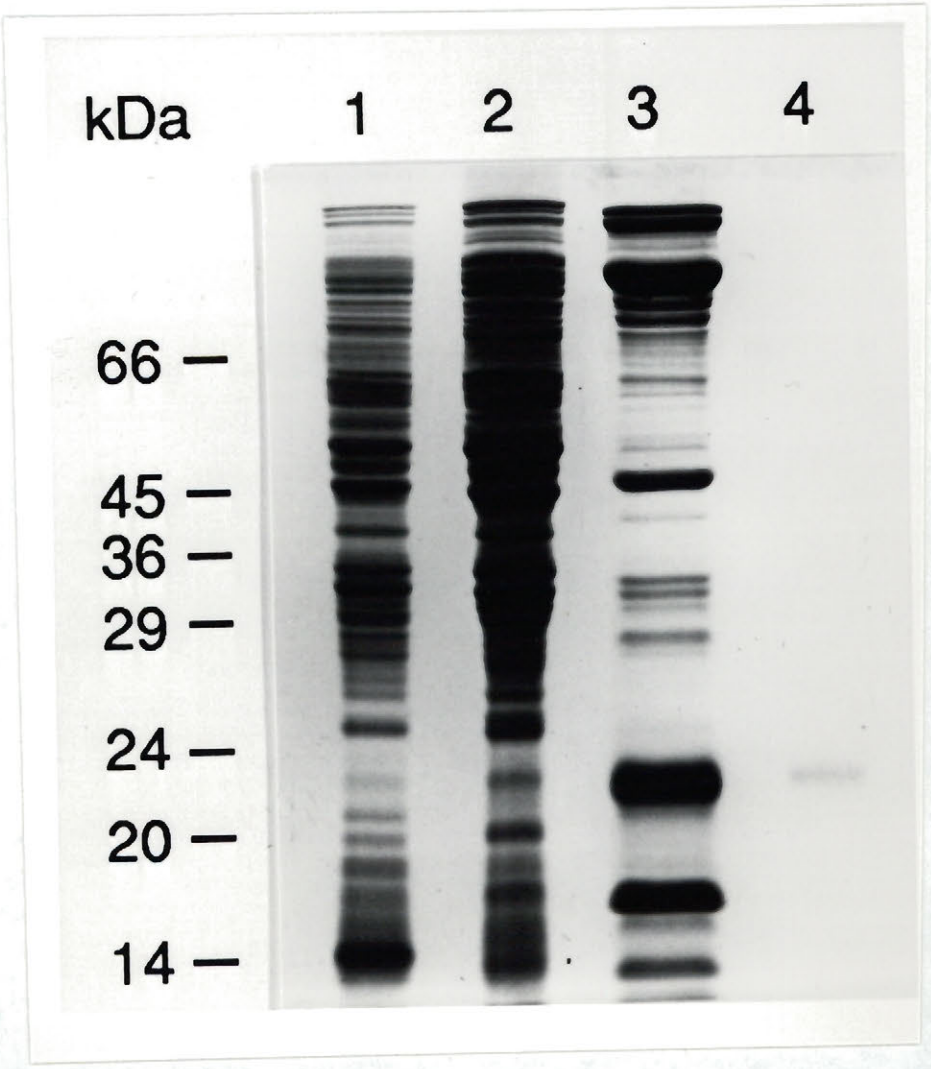

kDa
1
2
3
4
66
45
36
29
24
20
14

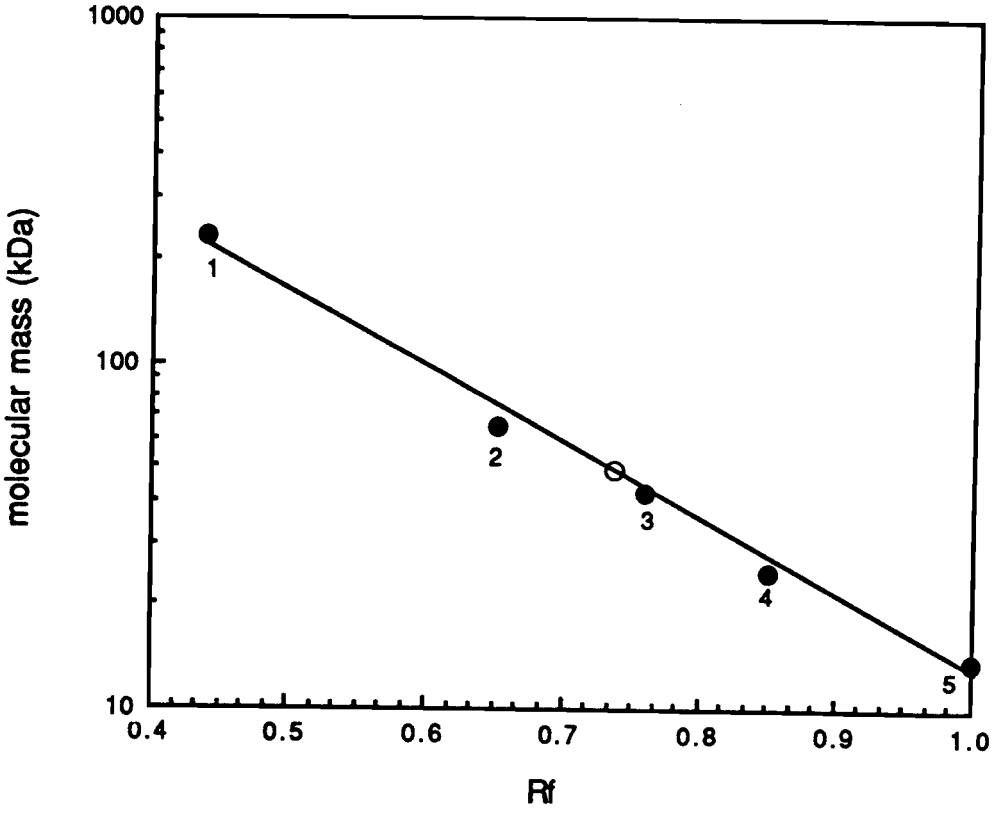
Table 1

Summary of purification of DNA-binding activity from *V. harveyi*

Fraction	Volume (ml)	mg/ml	Total Protein (mg)	Total Units	Specific Activity[1]	Yield (%)	Fold Purification
Cell lysate	150	10	1500	16,500	11	-	-
DEAE-Sepharose	31	22	682	8,184	12	49	1.1
dsDNA-cellulose	0.2	21	4.2	387	90	2.3	8.2
Gel Filtration	1.1	0.4	0.44	319	725	1.1	66
Mono Q	0.24	0.6	0.14	174	1216	1.0	110

[1] One unit of activity corresponds to 1 pmole of bound DNA/mg of protein in the standard gel retardation assay.

Figure 9: Molecular weight determination of DNA-binding protein. Elution position of purified protein (open circle) in superose 12 gel filtration chromatography as compared to protein standards (closed circles). 1) catalase (232 kDa), 2) BSA (66 kDa), 3) ovalbumin (45 kDa), 4) chymotrypsinogen A (25 kDa), ribonuclease A (13.7 kDa).


1000
100
10
molecular mass (kDa)
1
2
3
4
5
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Rf

Identification of the DNA-Binding Protein as LuxR

N-terminal sequence analysis resulted in the detection of low levels of amino acids at each cycle suggesting that the N-terminal might be blocked. However, the amino acids corresponding to that found in the N-terminal sequence of LuxR of *V. harveyi* were present at each cycle. This result along with its molecular mass of 23 kDa suggested that the purified protein was LuxR. Application of the gel retardation and footprint assays to cell extracts containing LuxR confirmed directly that the DNA binding protein was LuxR. The mobility shift patterns using extracts of *V. harveyi* wild type and transposon-generated mutants were compared to those of the purified protein and an extract of recombinant *E. coli* containing *luxR* (Figure 10). The transposon-generated mutants were derived from *V. harveyi* strain BB7 (Martin et al., 1989), which differs from the strain used to identify and isolate the DNA-binding protein (*V. harveyi* strain B392). Nonetheless, both wild type *V. harveyi* strains (lanes 2 and 3), plus the LuxC$^-$ and luxE$^-$ mutants (lanes 4 and 5), produced the characteristic mobility shift pattern as seen for recombinant LuxR (*E. coli* containing pMR1403; lane 8) and the purified protein (lane 9). Transposon insertions into two different regions of the *luxR* gene resulted in failure of the mutants to exhibit the typical DNA-binding pattern (lanes 6 and 7), supporting the evidence that the purified protein is LuxR. These mutants were able to generate a weak mobility shift that differed from the LuxR-generated shift pattern. This could be important, as other proteins may be involved in *lux* gene regulation (see Discussion). The footprint pattern caused by recombinant *E. coli* containing *luxR* was identical to that of wild type *V. harveyi* (Figure 11), with regions A and B being protected from DNase I digestion. Protein from the LuxR$^-$ mutant failed to bind, establishing that LuxR binds to both region A (-253 to -290) and region B (-116 to -170).

In Vitro *Transcription*

The purified LuxR protein was used in a transcription assay in an effort to reconstitute *lux* gene regulation *in vitro* (Figure 12). In the *in vitro* transcription

Figure 10: Gel retardation assay of *V. harveyi* wild type and mutants. Two μg of the following cell extracts were combined with fragment 1 (see figure 4): lane 1, no protein; lane 2, *V. harveyi* strain B392; lane 3, *V. harveyi* strain BB7; lane 4, *V. harveyi* MR1124 ($LuxC^-$); lane 5, *V. harveyi* MR1108 ($LuxE^-$); lane 6, *V. harveyi* MR1130 ($LuxR^-$); lane 7, *V. harveyi* MR1101 ($LuxR^-$); lane 8, *E. coli* DH5α containing pMR1403. Lane 9 is a gel retardation assay using 0.06 μg of purified LuxR.

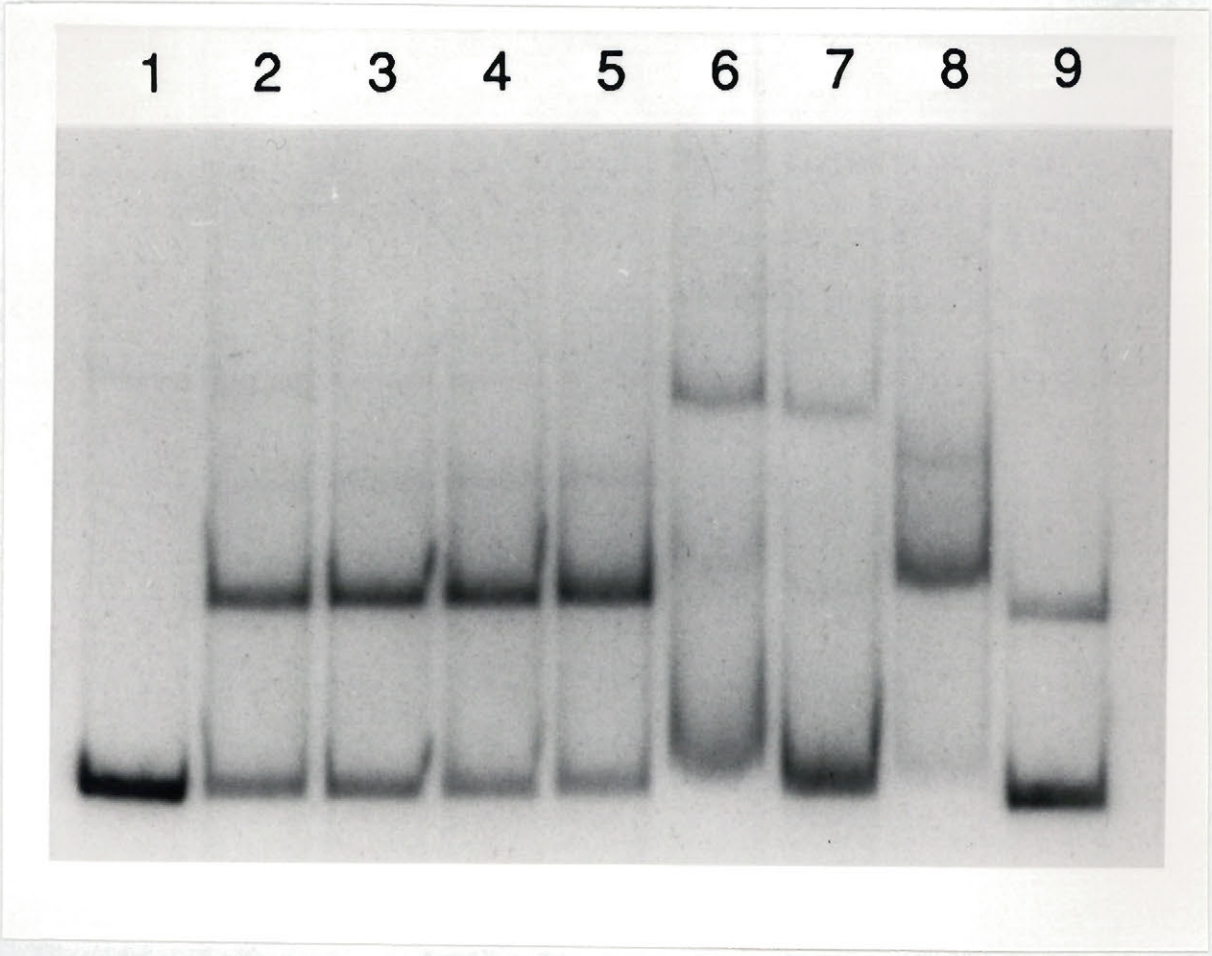

1
2
3
4
5
6
7
8
9

Figure 11: DNase I protection analysis. DNase I digestion of 5' end-labelled fragment 2 (see figure 4) with: 10 μg of BSA (lane 1), 10 μg of *V. harveyi* cell extract (lane 2), 10 μg of *V. harveyi* MR1130 ($LuxR^{-}$) cell extract (lane 3), and 10 μg of *E. coli* containing pMR1403 cell extract (lane 4). The protein-binding regions are indicated by the brackets marked A and B.

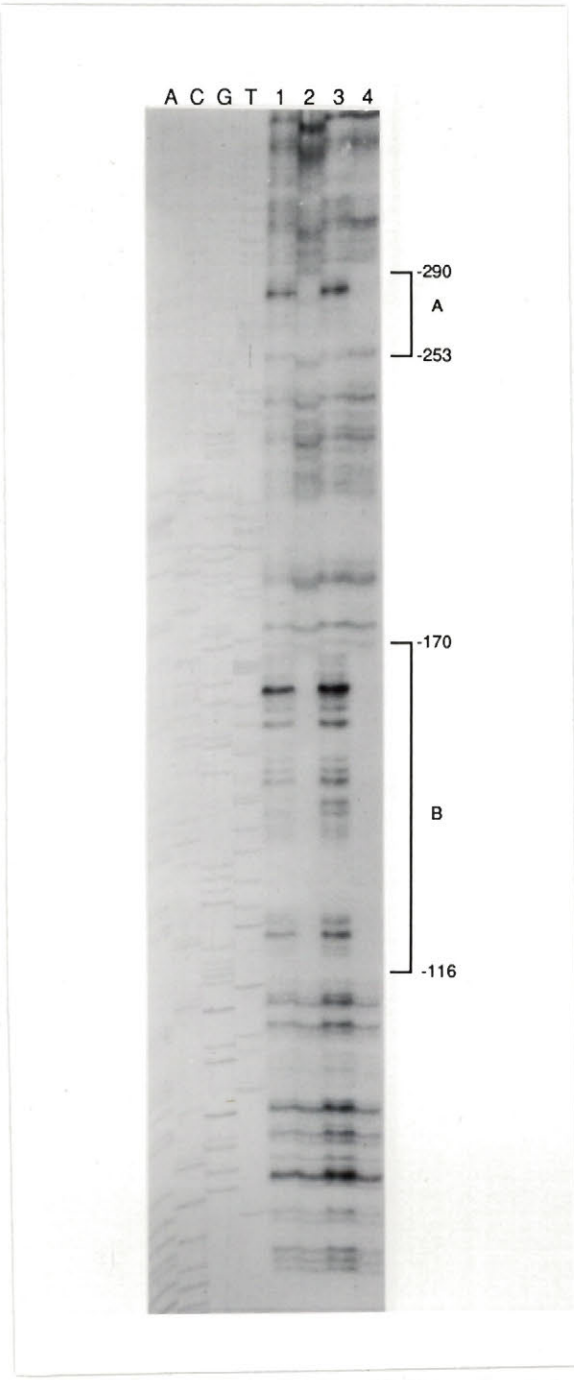
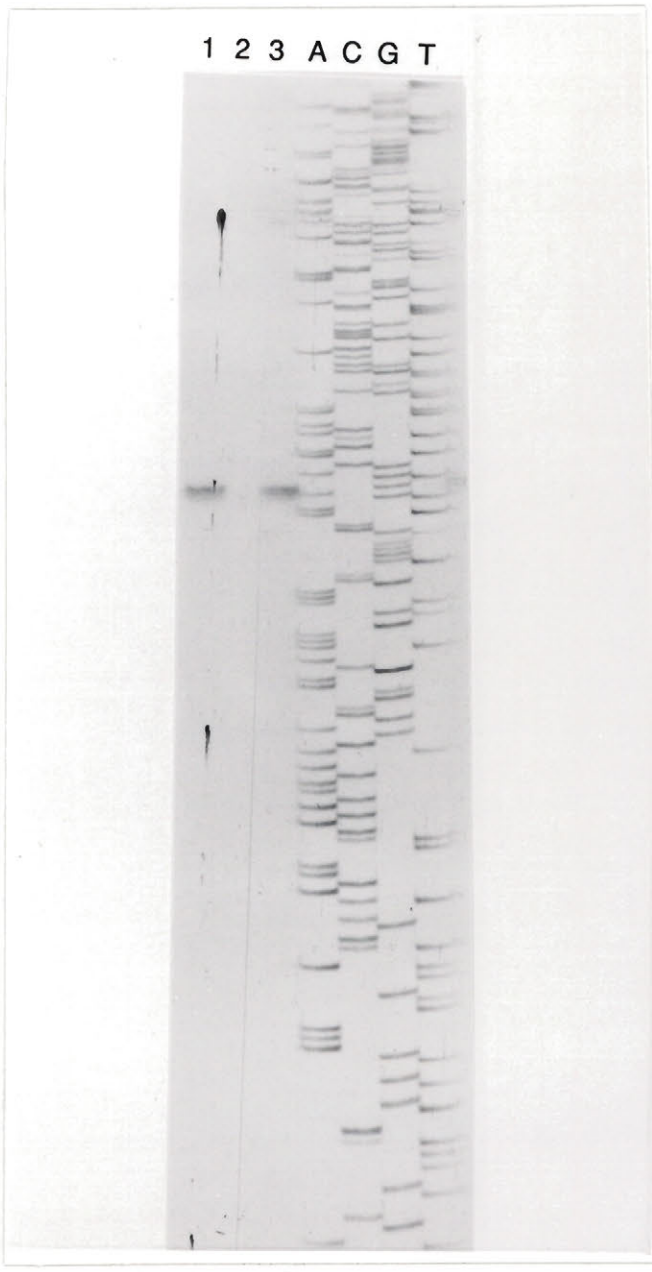

A C G T 1 2 3 4
-290
A
-253
-170
B
-116

Figure 12: *In vitro* transcription with *V. harveyi* RNA polymerase and purified LuxR. *In vitro* transcription was performed using fragment 2 and the following additions: lane 1, *V. harveyi* RNA polymerase; lane 2, *V. harveyi* RNA polymerase and 0.6 μg of LuxR, added before heparin addition; lane 3 *V. harveyi* RNA polymerase and 0.6 μg of LuxR, added after heparin addition. Lanes A, C, G, and T contain an M13mp18 sequence ladder and serve as size markers.


1 2 3 A C G T

assay, heparin is added after addition of RNA polymerase but before nucleotide addition in order to prevent non-specific initiation events by binding to free RNA polymerase. Without addition of LuxR, transcription initiation occurs at -123, as previously described (Swartzman et al, in press). When LuxR is added before addition of heparin, transcription is blocked at -123 and initiation is not seen at -26 (data not shown). However, transcription initiation at -123 occurs when LuxR is added following addition of heparin. Considering that protein-binding region B contains the -123 start site, the results suggest that LuxR and RNA polymerase compete for binding at the same site. Without LuxR in *E. coli* or the *in vitro* transcription assay, transcription initiation is permitted to occur at -123. Although LuxR is responsible for shifting the transcriptional start site from -123 to -26 *in vivo*, this was not evident *in vitro*, indicating the involvement of another factor or factors in *lux* gene regulation.

DISCUSSION

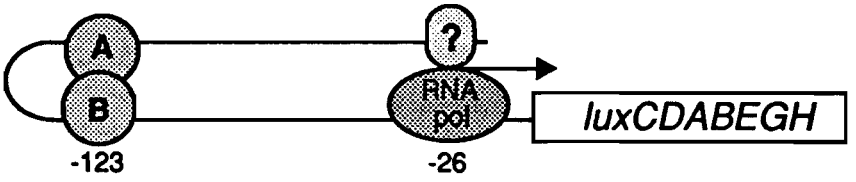
The DNA-binding protein detected in extracts of *V. harveyi* was identified as LuxR, the protein required for expression of *luxCDABEGH* mRNA and for accurate transcription of the *lux* promoter (Martin et al.,1989; Swartzman et al., in press). From the mobility shift and footprinting studies, it appears that LuxR is induced with cellular growth, which correlates with the 3-fold increase in *luxR* mRNA observed in post-induced cells (Martin et al., 1989). This increase may not be due to autoinducer-related cell-density regulation, as autoinducer-independent cell-density modulation of *lux* gene expression in *V. fischeri* has been reported (Dunlap and Kua, 1992). Binding by LuxR can be abolished when poly(dA-dT), but not poly(dA)-poly(dT), is used as competitor DNA. Several eukaryotic proteins have been found to bind specifically to A + T rich DNA, including: two nuclear DNA binding proteins of *Dictyostelium discoideum* (Garreau and Williams, 1983), the nuclear protein D1 of *Drosophila melanogaster* (Levinger and Varshavsky, 1982), datin from the yeast *Saccharomyces cerevisiae* (Winter and Varshavsky, 1989), and α-protein from mammalian cells (Solomon et al., 1986). Although the function of these proteins *in vivo* is largely unknown, α-protein (Solomon et al., 1986) and datin (Winter and Varshavsky, 1989) may be associated with the nuclear scaffold. Of the known A + T DNA-binding proteins, only the gene encoding datin has been cloned and sequenced. A comparison of amino acid sequences revealed no similarities between the LuxR and datin proteins.

LuxR binds to two regions of DNA upstream of the *luxC* gene: region A (-290 to -253) and region B (-116 to -170). Both regions are A + T rich and both contain the octamer AGTGTTAC. Experiments are currently being conducted to determine the functional relevance of this sequence. It is curious that the size of the protected regions in the footprint assay are not similar (37 bp for region A versus 55 bp for region B). It is possible that region A is bound by one subunit or molecule while region B is bound by two subunits or more of LuxR, perhaps through cooperative interactions.

The *in vitro* transcription assay using LuxR demonstrated that LuxR blocks transcription from the -123 promoter utilized *in vitro* and in recombinant *E. coli* containing *luxCDEBEGH*. Transcription at the -26 promoter is restored *in vivo* when *luxR* is supplied *in trans* in recombinant *E. coli* (Swartzman et al., in press). Because LuxR was unable to effect this result *in vitro*, it is likely that the *in vitro* transcription assay is incomplete and requires one or several other factors. Because luminescence in *V. harveyi* is known to be catabolite repressed, and a CRP binding site has been located 455 bp upstream of *luxC* (Miyamoto et al., 1988b), it is possible that CRP is involved in transcription initiation. Currently, there is no evidence that autoinducer and LuxR of *V. harveyi* are functionally related, as LuxR stimulates *luxCDEABEGH* mRNA expression in an autoinducer-independent fashion (Showalter et al., 1990). There may be an unidentified autoinducer-controlled transcriptional regulator analogous to that of *V. fischeri* required for activation of the -26 promoter. Finally, there is a distinct possibility that supercoiling may play a role in transcriptional activation. Bioluminescence expression in marine bacteria has been linked to gyrase activity (Watanabe et al., 1991) and an excess of gyrase A copurifies with *V. harveyi* RNA polymerase (Swartzman and Meighen, in press). A model for transcription initiation of the *lux* promoter is presented in Figure 13. LuxR binds to regions A and B causing a loop to form, perhaps with the aid of gyrase activity. This loop brings another regulatory factor into close proximity of the -26 promoter, which increases RNA polymerase open or closed form occupancy of the promoter and activates transcription of the *lux* promoter. Elucidation of the mechanism and consequences of LuxR-DNA binding (ie: recognition sites, cooperative binding, looped DNA, bent DNA, etc.) will be critical for our understanding of *V. harveyi lux* gene regulation.

Although the general regulators CRP and LexA have been shown to bind to the *V. fischeri lux* promoter region (Shadel et al., 1990; 1990a), specific binding of *V. fischeri* LuxR to its operator has not been observed (Kaplan and Greenberg, 1987). LuxR of *V. harveyi* is therefore the first DNA-binding protein to be purified and identified from a luminescent bacterium. Regulation of the *V. harveyi lux* operon

Figure 13: Model of transcription activation of the *V. harveyi lux* promoter. LuxR binds to two sites of DNA upstream from the *V. harveyi lux* promoter, marked A and B. It is proposed that the LuxR proteins interact to form a loop of DNA, allowing another regulatory factor to come into close proximity of the -26 promoter and activate transcription of the *V. harveyi lux* operon.


A
B
-123
?
RNA pol
-26
luxCDABEGH

appears to be a complicated process requiring several different regulatory elements. LuxR is essential but is not the only factor involved *lux* gene regulation. As genetic manipulations have failed to unambiguously identify other loci required for bioluminescence, biochemical approaches must be explored to provide the missing elements. This can be accomplished with the addition of LuxR in the *in vitro* transcription assay to search for factors required for accurate transcription of the *lux* promoter, as well as by examining the properties and function of the *lux* DNA-binding protein in the *V. harveyi* LuxR$^-$ mutant.

CHAPTER 7

GENERAL DISCUSSION

When the work presented in this thesis was first initiated, little was known about the regulation of *lux* genes of *V. harveyi*. The structural *lux* genes were cloned and identified, yet the pattern of cell density-dependant expression of bioluminescence was absent in recombinant *E. coli*, even when 18 kbp of continuous *lux* DNA was used (Miyamoto et al., 1987; 1988a). It was unknown whether the regulatory factors were present but nonfunctional in *E. coli*, or perhaps missing altogether from the 18 kbp fragment of *lux* DNA. It was also possible that *E. coli* lacked the transcriptional machinery necessary for *V. harveyi lux* gene regulation.

It was initially thought that the *V. harveyi lux* system would be comparable to that of *V. fischeri*, as the function and organization of the *lux* structural genes, as well as the cell-density development of bioluminescence, appeared to be well conserved between the two genera of bacteria. However, in contrast to *V. harveyi*, recombinant *E. coli* containing *V. fischeri lux* DNA was able to duplicate the cell-density modulation of light production seen in luminescent *V. fischeri* (Engebrecht et al., 1983; Engebrecht and Silverman, 1984). Two *lux* genes, *luxR* and *luxI*, identified as the putative receptor and autoinducer synthesis genes, respectively, were required for transcriptional regulation of the *V. fischeri lux* genes.

The nucleotide sequence upstream of the *V. harveyi luxC* gene provided the first concrete evidence demonstrating the difference in the *V. harveyi* and *V. fischeri lux* operons (Miyamoto et al., 1988b). In *V. fischeri*, the *luxI* gene immediately precedes *luxC* and is the first gene of the right *lux* operon. *luxR* is adjacent to *luxI*, but is transcribed in the opposite direction and constitutes the left *lux* operon. However, there is no open reading frame analogous to *luxI* and *luxR* upstream of *luxC* in *V. harveyi*. Instead, there is an A + T rich sequence spanning 600 nucleotides before reaching an open reading frame of unknown function located and transcribed in the same direction as *luxR* relative to the *lux* structural genes. Furthermore, although the *V. harveyi* and *V. fischeri* autoinducers are structurally similar, they are chemically distinct and display no species cross-reactivity (Cao and Meighen, 1989). These and other observations indicated that the model of *lux* gene regulation presented for *V.*

fischeri does not adequately explain *lux* gene regulation in *V. harveyi*. The nucleotide sequences of *lux* systems from other luminescent bacteria examined to date have revealed that the existence of *luxI* and *luxR* in close proximity to the structural *lux* genes is unique to *V. fischeri* (Meighen, 1991). A model of *lux* gene regulation based on that of *V. harveyi* may therefore be more representative of the development of bioluminescence in a large subset of marine bacteria.

Two sets of polycistronic mRNAs were found to code for the proteins of the *lux* operon (Miyamoto et al., 1985; 1988b). The first set, consisting of four mRNA species, were induced with the development of luminescence and extended from the 5' ends of the *luxD* and *luxA* genes to the 3' end of *luxB* and 4 kb further downstream. The second set of polycistronic mRNAs were so low in abundance their regulation could not be studied on Northern blots. All four mRNA species of this second set extend from the 5' end of *luxC* to various points downstream. It remains uncertain whether these polycistronic messages arose from specific RNA processing or from internal transcription initiation. Although a functional *in vitro* transcriptional start site has been located upstream of *luxD* (unpublished results), it fails to allow for expression of a reporter gene *in vivo* when transconjugated into *V. harveyi* (Miyamoto, personal communication). There is some evidence demonstrating that internal initiation is not a factor in generating the polycistronic messages. Inactivation of target genes by transposon mutagenesis with mini-MU disrupts transcription of the downstream genes (Engebrecht et al., 1983). For the *V. harveyi* Lux$^-$ transposon-generated mutants, aldehyde was able to restore bioluminescence for the LuxE$^-$, but not the LuxC$^-$ mutant. Addition of aldehyde bypasses the requirement for the fatty acid reductase complex encoded by *luxC*, *luxD*, and *luxE*. Since *luxE* is located downstream of *luxA* and *luxB*, its disruption did not interfere with luciferase synthesis. In contrast, transposon insertion of *luxC*, the first gene of the *lux* operon, abolished expression of luciferase, indicating that internal transcription initiation is not occurring.

In order to examine *V. harveyi lux* gene regulation in detail, it was necessary to

first locate the transcriptional start site of the *V. harveyi lux* operon. As some of the polycistronic messages terminated 2-3 kb downstream of *luxE*, the last known gene of the operon, it was also of interest to search for possible open reading frames corresponding to the excess mRNA sequences. As such, both the 5' and 3' ends of the *V. harveyi lux* operon were mapped. Transcription initiation occurred 26 bp upstream (-26) of the first ATG codon of *luxC*, and was found to be influenced by cell-density as well as by glucose and arginine levels in the growth medium. Although the mechanism by which arginine stimulates bioluminescence is unknown (Nealson et al., 1970), the results suggest that it functions at a transcriptional level. As exogenously added autoinducer to *V. harveyi* grown in minimal media lacking arginine does not stimulate luminescence (Sun, personal communication), it would appear that arginine activates transcription by an autoinducer-independent mechanism.

The 3' end of the *lux* mRNA was located after a typical rho-independent transcription terminator and before a similar terminator for a gene of unknown function transcribed in the opposite direction. The nucleotide sequence upstream from the terminator demonstrated the presence of two previously undetected genes within the *lux* operon. LuxG and LuxH are either essential for *V. harveyi* survival or nonessential for light production, as transposon mutagenesis in *V. harveyi* has failed to detect any transposon insertions in *luxG* or *luxH* generating a Lux^- phenotype (Martin et al., 1989). It must also be considered that recombinant *E. coli* containing the *V. harveyi lux* operon with or without *luxG* and *luxH* produce the same amount of light. *LuxG*, but not *luxH*, was found to be present in the *V. fischeri lux* operon, located immediately upstream of a bidirectional rho-independent terminator. Subsequently, it was shown that *luxG* is present in the *lux* operons of two other luminescent marine bacteria, *P. phosphoreum* and *P. leiognathi* (Lee et al., 1991). The function of *luxG* in bioluminescence may be related to the habitat of the microorganisms, as *luxG* does not appear to be within the *lux* operon of the terrestrial luminescent bacteria *X. luminescens* (Meighen, 1991).

It was recently shown that LuxG is member of the FNR (ferredoxin $NADP^+$ reductase)-family, and contains both a FAD-containing and an $NADP^+$-binding domains (Andrews et al., 1992). It has been previously considered that a *lux* gene coding for a flavin reductase would be identified (Meighen, 1991), as NAD(P)H flavin reductases have been purified from luminescent bacteria (Jablonski and DeLuca, 1977; Michaliszyn et al., 1977) and would be required to provide the $FMNH_2$ substrate for luciferase. It is very likely then that LuxG is an $NADP^+$-dependant flavin reductase that appears to assist, but is nonessential for bioluminescence under laboratory conditions.

It has also been recently shown that the *luxH* gene product is homologous to 3,4-dihydroxy-2-butanone 4-phosphate synthase, an enzyme involved in the biosynthesis of riboflavin (Richter et al., 1992). This gene was previously reported as *htrP*, which was shown to be essential for bacterial growth at high temperatures (Raina et al., 1991) and is identical to *ribB* of *E. coli*, which is homologous to the N-terminal sequence of an open reading frame within the riboflavin synthesis operon of *B. subtilis* (Richter et al., 1992). The occurrence of only one of six riboflavin synthesis genes in the *V. harveyi lux* operon suggests that LuxH may catalyse a rate-limiting reaction for the biosynthesis of riboflavin in order to maintain the high levels of $FMNH_2$ required by luciferase (Richter et al., 1992). Three open reading frames downstream of *luxG* in *P. leiognothi* were shown to share identity with the gene products of three riboflavin synthesis genes of *B. subtilis* (Lee and Meighen, 1992). As no transcriptional terminator has been located within the sequence, it is most likely that they are co-transcribed with the *lux* genes.

The inclusion of *luxG*, *luxH*, and other genes related to flavin metabolism with the *lux* operons of various light-producing bacteria indicates that they may have evolved with the bioluminescence system. While they may serve to enhance $FMNH_2$ production, the disappearance of many of these genes from the *lux* operons of different luminescent genera suggests that their function is extraneous. Conversely, the bacteria may require one or more of these genes for their survival and/or

luminescence in their various symbiotic relationships and environmental habitats.

As the examination of *V. harveyi lux* gene regulation in recombinant *E. coli* proved to give unsatisfactory results (Miyamoto et al., 1987), it was decided that *in vitro* transcription studies may overcome the difficulties encountered *in vivo*. Because of the uncertainty of the ability of *E. coli* RNA polymerase to initiate transcription at a *V. harveyi* promoter, the *V. harveyi* RNA polymerase was purified and its activity was compared to that of *E. coli* in the *in vitro* transcription assay. Both enzymes initiated transcription at the same site, demonstrating that they share similar promoter specificities.

An abundance of gyrase A or the A subunit of topoisomerase II copurified with the RNA polymerase up until the final stages of purification and was estimated to constitute over 30% of the DNA-binding proteins eluting from the DNA-cellulose column. It is unclear why just one subunit of this heterdimeric protein would be present in this fraction, since both subunits are required for complete activity. As the genes coding for the A and B subunits are unlinked in *E. coli* (Gellert, 1981) it is possible that they are differentially expressed in *V. harveyi*. Alternatively, both subunits may be present in equal amounts but only weakly associated, resulting in only one subunit binding to the column. Indeed, the A subunit is known to contain the DNA-binding domain, while subunit B contains the ATPase activity (Maxwell and Gellert, 1986). Topoisomerase II creates negative supercoils in DNA and is expected to stimulate transcription initiation since it creates a favourable environment for DNA strand-separation during open-complex formation (Pruss and Drlica, 1989). In addition, positive supercoils (relaxed DNA) are generated before and negative supercoils are generated behind the moving transcription complex, necessitating their removal by topoisomerases II and I, respectively, to maintain the proper helical tension of the genome (Liu and Wang, 1987). The topoisomerase enzymes are also linked to DNA replication, recombination and repair (Drlica, 1984). It is possible that the *V. harveyi* genome requires a high degree of negative supercoiling for transcription, which may help explain why we have been unable to reconstitute

regulated light production in recombinant *E. coli* containing the *lux* DNA. The link found between gyrase activity and bioluminescence in *P. phosphoreum* (Watanabe et al., 1991) may reflect the requirement for topoisomerase II in *lux* gene regulation.

The *in vitro* transcription studies indicated that the start site used *in vitro* differed from that used *in vivo*, suggesting there was a crucial element missing in the assay. It was of interest to determine if a similar pattern is seen in recombinant *E. coli*, also suspected of missing *lux* regulatory factors. Both *in vivo* and *in vitro* transcriptional start sites were mapped to -123 bp in front of *luxC*, almost 100 bp upstream of the -26 start site found *in vivo* for *lux* mRNA isolated from *V. harveyi*. It was concluded that one or several factors were missing in the *in vitro* assay as well as in *E. coli* for proper transcription initiation at the -26 promoter. Transposon mutagenesis in *V. harveyi* has previously identified *luxR* as a transcriptional activator of the *lux* operon that was able to stimulate expression of *lux* mRNA in an autoinducer-independent manner (Martin et al., 1989; Showalter et al., 1990). Accordingly, the effect of *luxR* on the transcriptional start site in recombinant *E. coli* was examined. It was discovered that *luxR* is the missing element in *E. coli* required for transcriptional activation from the -26 *lux* promoter. Although *luxR* greatly increases the expression of bioluminescence in recombinant *E. coli*, this stimulation does not exhibit the characteristic growth- and autoinducer-dependancy (Showalter et al., 1990). The nucleotide sequence of *luxR* demonstrated the presence of a cro-like DNA-binding domain (helix-turn-helix motif) in the LuxR protein, however it did not exhibit any similarity to LuxR of *V. fischeri* (Showalter et al., 1990). It is likely, therefore, that another factor is required for autoinducer-dependant regulation of the *V. harveyi lux* operon.

Initial attempts to identify factors involved in *lux* gene regulation led to the detection of a DNA-binding protein found exclusively in cell extracts of *V. harveyi* that bound to a region of DNA upstream from *luxC* in a gel retardation assay (Swartzman et al., 1991). Further investigation showed that this protein bound specifically to regions -253 to -290 (A) and -116 to -170 (B) upstream of the *luxC*

gene. Both regions contain the octamer AGTGTTAC and are rich in A + T nucleotides. While the relevance of the octamer remains uncertain, it was established that the A + T sequences are important for binding. Although several other eukaryotic DNA-binding proteins with A + T specificity have been identified, to my knowledge this is the first indication of a prokaryotic DNA-binding protein that is subject to poly(dA-dT) competition. α-Protein of African green monkey cells, a poly(dA-dT) binding protein, as well as IHF (integration host factor) and HU (histone-like protein) of prokaryotes, contact the minor groove of B-form DNA (Solomon et al., 1986; Harrison, 1991), while most sequence specific DNA-binding proteins are known to interact with the major groove of the DNA helix (Pabo and Sauer, 1984).

The DNA-binding activity was purified to homogeneity and was shown to be the protein encoded by the *luxR* gene of *V. harveyi*. When used in the *in vitro* transcription assay, LuxR failed to stimulate transcription initiation at the -26 promoter site. This was somewhat surprising, as the presence of *luxR* in recombinant *E. coli* shifted the *in vivo* start site from -123 to -26. Different assay conditions, such as a modulation of protein concentrations or the use of a supercoiled DNA template, may be required to reconstitute functional activation of the -26 promoter. It is also possible that other factors (such as CRP) present in *E. coli* have to be included in the *in vitro* transcription assay. Under certain conditions, LuxR blocked transcription initiation at the -123 promoter. As the -123 promoter lies within binding region B, it is likely that LuxR competes with RNA polymerase for binding and thus blocks its access to this promoter site.

A model for transcription initiation was presented that envisioned the formation of a loop with a contact point consisting of the two bound LuxR proteins, placing a distant DNA-bound regulator within reach of the -26 promoter. This regulator may be CRP, since the *lux* operon is known to be catabolite repressed (Nealson et al, 1970), and there is a putative DNA-binding site for CRP upstream of the LuxR binding sites (Miyamoto et al., 1988b). However, *E. coli* recombinants containing

luxR and the *lux* genes do not exhibit the characteristic cell-density dependant development of bioluminescence and *luxR* of *V. harveyi* does not appear to be functionally related to the *luxR* gene of *V. fischeri*. Consequently, the other factor may be analogous to *luxR* of *V. fischeri* and stimulate transcription initiation in an autoinducer-dependant fashion.

Transposon mutagenesis of the *V. harveyi* genome has resulted in the generation of a series of dark mutants with transposon insertions within the *lux* operon and *luxR* (Martin et al., 1989; Showalter et al., 1990). No other genes were found to be absolutely required for light production. It is possible that relatively highly levels of constitutive *lux* expression can be achieved with just LuxR, while an autoinducer- and cell density-dependant mechanism of transcriptional activation serves only to enhance the level of light production. This is a distinct possibility as recombinant *E. coli* containing *luxR* and the *lux* operon produces only 10 fold less light than wild type *V. harveyi* (Showalter et al., 1990). Thus it would be difficult to select for an autoinducer/regulator mutant emitting intermediate levels of light amid bright colonies on a petri dish.

One strategy to search for an inducer-dependant regulator protein would be to assay for a protein that would bind to radioactive autoinducer. This has already been attempted for the *V. fischeri* LuxR and autoinducer, with negative results (Kaplan and Greenberg, 1987). However, LuxR proved to be insoluble when overexpressed and had to be solubilized in 6 M guanadine HCl. The protein may have remained inactive even after removal of the denaturant. A similar attempt in our laboratory with ^{3}H-labelled *V. harveyi* autoinducer has given similar negative results (Sun, personal communication). It is even possible that a metabolite of the autoinducer is responsible for regulator-binding. Gel retardation and DNase I footprinting assays readily detected the presence of LuxR and should be able to do likewise for an autoinducer-receptor complex that binds to DNA. A LuxR^{-} mutant, the cell extract of which exhibited some binding activity to DNA upstream of *luxC* in the gel retardation assay, may provide the means to search for this factor.

Further analysis of the mechanism of transcriptional activation mediated by LuxR will greatly enhance our knowledge of *V. harveyi lux* gene regulation. Aside from the helix-turn-helix DNA binding motif, the *luxR* gene product shares no significant homology with any other protein in the data base (Showalter et al., 1990). As it also appears to be a novel prokaryotic poly(dA-dT) binding protein, LuxR may represent a unique class of transcriptional activators. It should be mentioned that the presence of *luxR* on a medium or high copy-number plasmid in *E. coli* results in poor viability and growth. The toxicity of LuxR may be attributed to its ability to bind to random A + T sequences. Although the number of LuxR molecules present in *V. harveyi* is unknown, it does not appear to have the same effect in this bacterium since *V. harveyi* transconjugated with *luxR* present on a medium copy-number plasmid does not impede cellular growth (unpublished data).

There is still much work to be done regarding the charaterization LuxR-DNA binding. The putative consensus sequence has yet to be confirmed and the concept of cooperative binding to the two LuxR binding sites has to be tested. The ability of LuxR to bend DNA or cause loop formation should also be investigated. Finally, a map of the functional domains of LuxR would help elucidate its mechanism of transcriptional activation.

As mentioned earlier, the presence of *lux* regulatory genes located in close proximity to the *lux* structural genes has been found only in the *V. fischeri* luminescent system. Although LuxR-type binding to the *V. harveyi* LuxR binding sites has not been detected with extracts from any other luminescent species examined to date (Swartzman et al, 1991), it is possible that these microorganisms contain a protein with a function similar to that of LuxR of *V. harveyi* but specific for its own DNA. Indeed, sequences upstream of *P. leiognathi* (Szittner, personal communication) and *P. phosphoreum* (Lee et al, 1991) are rich in A + T residues. Examination of the details of *lux* gene activation in *V. harveyi* may therefore provide a general model for the regulation of bioluminescence in many luminescent bacteria.

GENERAL REFERENCES

Andrews, S. C., Shipley, D., Keen, J. N., Findlay, J. B. C., Harrison, P. M., and Guest, J. R. (1992) *FEBS Lett.* **3**:247-252.

Baldwin, T. O., Berends, T., Bunch, T. A., Holzman, T. F., Rausch, S. K., Shamansky, L., Treat, M. L., and Zeigler, M.M. (1984) *Biochemistry* **23**:3663-3667.

Baldwin, T. O., J. H. Devine, R. C. Heckel, J.-W. Lin, and Shadel, G. S. (1989) *J.Biolumin. Chemilumin.* **4**:326-341.

Baumann, P., and Baumann, L. (1981) In *The Pokaryotes: a Handbook on Habitats, Isolation and Identification of Bacteria.* (M. P. Starr, H. Stolp, H. G. Truper, A. Balows, and H. G. Schlegal, Eds) pp. 1302-1331. Springer-Verlag, New York.

Baumann, P., Baumann, L., Woolkalis, M., and Bang, S. (1983) *Ann. Rev. Microbiol.* **37**:369-398.

Bear, D. G., and Peabody, D. S. (1988) *Trends Biochem. Sci.* **13**:343-347.

Belas, R., Mileham, A., Cohn, D., Hilmen, M., Simon, M., and Silverman, M. (1982) *Science* **218**:791-793.

Berk, A. J., and Sharp, P. A. (1978) *Proc. Natl. Acad. Sci. USA* **75**:1274-1278.

Boylan, M., A. F. Graham, and Meighen E. A. (1985) *J. Bacteriol* **163**:1186-1190.

Boylan, M., Miyamoto, C., Wall, L., Graham, A., and Meighen, E. (1989) *Photochem. Photobiol* **49**:681-688.

Boyle, R. (1668) *Philos. Trans. R. Soc. London* **2**:581-600.

Bracco, L., Kotlarz, D., Kolb, A., Diekmann, S., and Buc, H. (1989) *EMBO J.* **8**:4289-4296.

Burgess, R. R., Travers, A. A., Dunn, J. J., and Bautz, E. K. F. (1969) *Nature* **221**:43.

Burgess, R. R., and Jendrisak, J.J. (1975) *Biochemistry* **14**:4634-4638.

Buttner, M. J., Chater, K. F., and Bibb, M. J. (1990) *J. Bacteriol.* **172**:3367-3378.

Cao, J.-G., and Meighen, E. A. (1989) *J. Biol. Chem.* **264**:21670-21676.

Carlomagna, M. S., A. Riccio, and Bruni, C. B. (1985) *J. Bacteriol.* **163**:362-368.

Chamberlin, M. J. (1974) *Ann. Rev. Biochem.* **43**:721-75.

Chen N.-Y., F.-M. Hu, and Paulus, H. (1987) *J. Biol. Chem.* **262**:8787-8798.

Choi, S. H., and Greenberg, E. P., (1991) *Proc. Natl. Acad. Sci. USA* **88**:11115-11119.

Choi, S. H., and Greenberg, E. P., (1992) *J. Bacteriol.* **174**:4064-4069.

Christianson, T. W., and Clayton, D. A. (1986) *Proc. Natl. Acad. Sci. USA* **83**:6277-6281.

de Cromrugghe, B., Busby, S., and Buc, H. (1984) *Science* **224**:831-838.

Devine, J. H., Countryman, C., and Baldwin, T. O. (1988) *Biochemistry* **27**:837-842.

Devine, J. H., Shadel, G. S., and Baldwin, T. O. (1989) *Proc. Natl. Acad. Sci. USA* **86**:5688-5692.

Dolan, K. M., and Greenberg, E. P. (1992) *J. Bacteriol.* **174**:5132-5135.

Drlica, K. (1984) *Microbiol. Rev.* **48**:273-289.

Dubois, R. (1887) *C. R. Soc. Biol. Ser. 8* 3564-568.

Dunlap, P. V. (1989) *J. Bacteriol.* **171**:1199-1202.

Dunlap, P. V., and Greenberg, E. P. (1985) *J. Bacteriol.* **171**:3549-3552.

Dunlap, P. V., and Greenberg, E. P. (1988) *J. Bacteriol.* **170**:4040-4046.

Dunlap, P. V., and Greenberg, E. P. (1991) in *Microbial Cell-Cell Interactions* (M. Dwarkin, Ed) pp. 219-253. Amer. Soc. Microbiol.

Dunlap, P. V., and Kuo, A. (1992) *J. Bacteriol.* **174**:2440-2448.

Dunlap, P. V., and Ray, J. M. (1989) *J. Bacteriol.* **171**:3549-3552.

Eberhard, A. (1972) *J. Bacteriol.* **109**:1101-1105.

Eberhard, A., Burlingame A. L., Eberhard, C., Kenyon, G. L., Nealson, K. H., and Eberahrd, A., Widrig, C. A., McBath, P., and Schineller, J. B. (1986) *Arch. Microbiol.* **146**:35-40.

Eberhard, A., Longin, T., Widrig, C. A., Stranick, S. J. (1991) *Arch. Microbiol.* **155**:294-297.

Engebrecht, J., Nealson K. H., and Silverman, M. (1983) *Cell* **32**:773-781.

Engebrecht, J., and Silverman, M. (1984) *Proc. Natl. Acad. Sci. USA* **81**:4154-4158.

Engebrecht, J., and Silverman, M. (1986) *Genetic Eng.* **8**:31-44.

Engebrecht, J., and Silverman, M. (1987) *Nucleic Acids Res.* **15**:10455-10467.

Eschenlauer, A. C., and Reznikoff, W. S. (1991) *J. Bacteriol.* **173**:5024-5029.

Farmer, J. J., III, Jorgensen, J. H., Grimont, P. A. D., Akhurst, R. J., Poinar, Jr., G. O., Ageron, E., Pierce, G. V., Smith, J. A., Carter, G. P., Wilson, K. L., and Hickman-Brenner, F. W. (1989) *J. Clin Microbiol.* **27**:1594-1600.

Ferri, S. R., and Meighen, E. A. (1991) *J. Biol. Chem.* **266**:12852-12857.

Fried, M., and Crothers, D. M. (1981) *Nucleic Acids Res.* **9**:6505-6525.

Friedrich, W. F., and Greenberg, E. P. (1983) *Arch. Microbiol.* **134**:87-91.

Gambello, M. J., and Iglewski, B. H. (1991) *J. Bacteriol.* **173**:3000-3009.

Garreau, H., and Williams, J. G. (1983) *Nucleic Acids Res.* **11**:8473-8484.

Gellert, M. (1981) *Ann Rev. Biochem.* **50**:879-910.

Gething, M. J., and Sambrook, J. (1992) *Nature* **355**:33-45.

Harrison, S. C. (1991) *Science* **353**:715-719.

Harveyi, E. N. (1952) *Bioluminescence* Academic Press, Inc., New York.

Hastings, J. W., Potrikas, C. J., Gupta, S. C., Kurfast, M., and Makemson, J. C. (1985) *Adv. Microbiol. Physiol.* **26**:235-291.

Haygood, M. G., and Nealson, K. H. (1985a) *Symbiosis* 1:39-51.

Haygood, M. G., and Nealson, K. H. (1985b) *J. Bacteriol.* **162**:209-216.

Helmann, J. D., and Chamberlin, M. J. (1988) *Ann. Rev. Biochem.* **57**:839-872.

Henikoff, S., Wallace, J. C., and Brown, J. P. (1990) *Methods Enzymol.* **183**:111-132.

Higgens, N. P., Peebles, C. L., Sugino, A., and Cozzarelli, N. R. (1978) *Proc. Natl. Acad. Sci. USA* **75**:1773-1777.

Hoover, T. R., Santero, E., Porter, S., and Kustu, S. (1990) *Cell* **63**:11-22.

Hudson, G. S., and Davidson, B. E. (1984) *J. Mol. Biol.* **180**:1023-1051.

Igarashi, K., and Ishihama, A. (1991) *Cell* **65**:1015-1022.

Jablonski, E., and Deluca, M. (1977) *Biochemistry* **16**:2932-2936.

Jerlstrom, P. G., Bezjak, D. A., Jennings, M. P., and Beacham, I. R. (1989) *Gene* 78:37-46.

Kahn, D., and Ditta, G. (1991) *Mol. Microbiol.* **5**:987-997.

Kaplan, H. B., and Greenberg, E. P. (1985) *J. Bacteriol.* **163**:1210-1214.

Kaplan , H. B., and Greenberg, E. P. (1987) *Proc. Natl. Acad. Sci. USA.* **84**:6639-6643.

Kunst, F., Debarbouille, M., Msadek, T., Young, M., Mauel, C., Karamata, D., Klier, A., Rapoport, G., and Dedoner, R. (1988) *J. Bacteriol.* **170:**5093-5101.

Laemmli, U. K. (1970) *Nature* (London) **227**:680-685.

Lagacé, M., Howell, B. W., Burak, R., Lusty, C. J., and Shore, G. C. (1987) *J. Biol. Chem.* **262**:10415-10418.

Lange, R., and Hengge-Aronis, R. (1991) *Mol. Microbiol.* **5**:49-59.

Lee, D-. H., and Schleif, R. F. (1989) *Proc. Natl. Acad. Sci. USA* **86**:476-480.

Lee, C. Y., Szittner, R. B., and Meighen, E. A. (1991) *Eur. J. Biochem.* **201:**161-167.

Lee, C. Y., and Meighen, E. A. (1992) *Biochem. Biophys. Res. Commun.* **155**:351-358.

Levinger, L., and Varshavsky, A. (1982) *Proc. Natl. Acad. Sci. USA* **79**:7152-7156.

Liu, L. F., and Wang, J. C. (1987) *Proc. Natl. Acad. Sci. USA* **84**:7024-7027.

Lloubes, R., D. Baty, and Lazdunski, C. (1988) *Nucleic Acids Res.* **16**:3739-3749.

Lonetto, M., Gribskov, M., and Gross, C. A. (1992) *J. Bacteriol.* **174:**3843-3849

Losick, R. and Pero, J. (1981) *Cell* **25**:582-584.

Malan, T. P., and McClure, W. R. (1984) *Cell* **39**:173-180.

Maniatis, T., and Efstratiadis, A. (1980) *Methods Enzymol.* **65**:299-305.

Mancini, J. A., M. Boylan, R. R. Soly, A. F. Graham, and Meighen, E. A. (1988) *J. Biol. Chem.* **263**:14308-14314.

Markwell, M. A., Hass, S. M., Tolbert, N. E., and Bieber, L. L. (1981) *Methods Enzymol.* **171**:296-303.

Martin, M., Showalter, R., and Silverman, M. (1989) *J. Bacteriol.* **171**:2406-2414.

Maxwell, A. and M. Gellert. (1986) *Adv. in Prot. Chem.* **38**:69-107.

McClure, W. R., (1985) *Ann. Rev. Biochem.* **54**:171-204.

McElroy, W. D., and Seliger, H. H. (1962) *Sci. Am.* **207(6)**:76-89.

Meighen, E. A. (1988) *Ann. Rev. Microbiol.* **42**:151-176.

Meighen, E. A. (1991) *Microbiol. Rev.* **55**:123-142.

Meighen, E. A., Slessor, K. N., and Grant, G. G. (1982) *J. Chem. Ecol.* **8**:911-921.

Michaliszyn, G., Wing, S., and Meighen, E. (1977) *J. Biol. Chem.* **252**:7495-7499.

Miyamoto, C. M., Graham, A. D., Boylan, M., Evans, J. F., Hasel, K. W., Meighen,

E. A., and Graham, A. F. (1985) *J. Bacteriol.* **161**:995-1001.

Miyamoto, C. M., Byers, D., Graham, A. F., and Meighen, E. A. (1987) *J. Bacteriol.* **169**:247-253.

Miyamoto, C. M., Boylan, M., Graham, A. F., and Meighen, E. A. (1988a) *J. Biol. Chem.* **263**:13393-13399.

Miyamoto, C. M., A. F. Graham, and Meighen, E. A. (1988b) *Nucleic Acids Res.* **16**:1551-1562.

Miyamoto, C. M., E. A. Meighen, and Graham, A. F. (1990) *J. Bacteriol.* **172**:2046-2054.

Mizuno, T., M.-Y. Chou, and Inouye, M. (1984) *Proc. Natl. Acad. Sci. USA* **81**:1966-1970.

Neal, R. J., and Chater, K. F. (1987) *Gene* **58**:229-241.

Nealson, K. H. (1977) *Arch. Microbiol.* **112**:73-79.

Nealson, K. H. (1979) *Trends Biochem. Sci.* **4**:105-110.

Nealson, K. H., and Hastings, J. W. (1977) *Arch. Microbiol.* **112**:9-16.

Nealson, K. H., and Hastings, J. W. (1979) *Microbiol. Rev.* **43**:496-518.

Nealson, K. H., Platt, T., and Hastings, J. W. (1970) *J. Bacteriol.* **104**:313-322.

Nealson, K. H., Eberhard, A., and Hastings, J. W. (1972) *Proc. Natl. Acad. Sci. USA* **69**:1073-1076.

O'Kane, D. J., Vervoort, J., Muller, F., and Lee, J. (1987) In *Flavins and Flavoproteins* (D. E. Edmondson and B. B. McCormick Eds.) pp. 641-645. Walter de Gruyer & Co., Berlin.

Oppenheimer, N. J. (1981) *Biochemistry* **20**:2444-2449.

Pabo, P. T., and Sauer, R. T. (1984) *Ann. Rev. Biochem.* **53**:293-321.

Paces, V., Vlcek, C., and Urbanek, P. (1986) *Gene.* **44**:107-114.

Platt, T. (1986) *Ann. Rev. Biochem.* **55**:339-372.

Postle, K. and Good, R. F. (1985) *Cell* **41**:577-585.

Pruss, G. J., and Drlica, K. (1989) *Cell* **56**:521-523.

Raibaud, O., and Schwartz, M. (1984) *Ann. Rev. Genet.* **18**:173-206.

Raibaud, O., and Richet, E. (1987) *J. Bacteriol.* **169**:3059-3061.

Raina, S., Mabey, L., and Georgopoulos, C. (1991) *J. Bacterol.* **173**:5999-6008.

Reznikoff, W. S., Siegele, D. A., Cowing, D. W., and Gross, C. A. (1985) *Ann. Rev. Genet.* **19**:355-87.

Reznikoff, W. S. (1992) *J. Bacteriol.* **174**:655-658.

Richter, G., Volk, R., Krieger, C., Lahm, H-. W., Röthlisberger, U., and Bacher, A. (1992) *J. Bacteriol.* **174**:4050-4056.

Riendeau, D., and Meighen, E. A. (1979) *J. Biol. Chem.* **154**:7488-7490.

Rosenberg, M., and Court, D. (1979) *Ann. Rev. Genet.* **13**:319-53.

Ronson, C. W., Nixon, B. T., and Ausubel, F. M. (1987) *Cell* **49**:570-581.

Rosson, R. A., and Nealson, K. H. (1981) *Arch Microbiol.* **129**:299-304.

Ruby, E. -G., and Nealson, K. H. (1976) *Biol. Bul.* **151**:574-586.

Ruby, E. -G., and Nealson, K. H. (1977) *Science* **196**:432-434.

Sanger, F., Nicklen, S., and Culson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.

Schleif, R. (1988) *Science* **240**:127-128.

Schollmeier, K., D. Gartner, and Hillen, W. (1985) *Nucleic Acids Res.* **13**:4227-4237.

Schultz, S. C., Sheilds, G. C., and Steitz, T. A. (1991) *Science* **252**:1001-1007.

Shadel, G. S., Devine, J. H., and Baldwin, T. O. (1990a) *J. Biolum. Chemilum.* **5**:99-106.

Shadel, G. S., Young, R., and Baldwin, T. O. (1990b) *J. Bacteriol.* **172**:3980-3987.

Shadel, G. S., and Baldwin, T. O. (1991) *J. Bacteriol.* **173**:568-574.

Shadel, G. S., and Baldwin, T. O. (1992a) *J. Biol. Chem.* **267**:7690-7695.

Shadel, G. S., and Baldwin, T. O. (1992b) *J. Biol. Chem.* **267**:7696-7702.

Shine, J., and Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**:1342-1346.

Showalter, R. E., Martin, M. O., and Silverman, M. R. (1990) *J. Bacteriol.* **172**:2946-2954.

Siebenlist, U., Simpson, R. B., and Gilbert, W. (1980) *Cell* **20**:269-281.

Silverman, M., Martin, M., and Engebrecht, J. (1989) In *Genetics of Bacterial*

Diversity (D. Hopwood, and K. F. Chater, Eds.) pp. 71-85. Academic Press, Inc., New York.

Slock, J., VanRiet, D., Kolibachuk, D., and Greenberg, E. P. (1990) *J. Bacteriol.* **172**:3974-3979.

Solomon, M. J., Strauss, F., and Varshavsky, A. (1986) *Proc. Natl. Acad. Sci. USA* **83**:1276-1280.

Soly, R. R., J. A. Mancini, S. R. Ferri, M. Boylan, and Meighen, E. A. (1988) *Biochem. Biophys. Res. Comm.* **155**:351-358.

Soly, R. R., and Meighen, E. A. (1991) *J. Mol. Biol.* **219**:69-77.

Stout, V., Torres-Cabassa, A., Maurizi, M. R., Gutnick, D., and Gottesman, S. (1991) *J. Bacteriol.* **173**:1738-1747.

Su, W., Porter, S., Kutsu, S., and Echols, H. (1990) *Proc. Natl. Acad. Sci. USA.* **87**:5504-5508.

Swartzman, E., Kapoor, S., Graham, A. F., and Meighen, E. A. (1990a) *J. Bacteriol.* **172**:6797-6802.

Swartzman, E., Miyamoto, C., Graham, A., and Meighen, E. A. (1990b) *J. Biol. Chem.* **265**:3513-3517.

Swartzman, E., Cao, J. G., Miyamoto C., Graham A., and Meighen, E. A. (1991) in *Bioluminescence and Chemiluminescence: Current Status* (P. E. Stanley and L. J. Kricka, eds.) pp. 27-30. John Whiley and Sons.

Thomas, G. M., and Poinar, Jr., G. O. (1979) *Int. J. Syst. Bacteriol.* **29**:352-360.

Ulitzer, S. (1989) J. Biolumin. Chemilumin. **4**:317-325.

Ulitzer, S., and Hastings, J. W. (1979) *Proc. Natl. Acad. Sci. USA* **76**:265-267.

Ulitzer, S., and Kuhn, J. (1988) *J. Biolumin. Chemilumin.* **2**:81-93.

Vidal-Ingigliardi, D., Richet, E., and Raibaud, O. (1991) *J. Mol. Biol.* **218**:323-334.

von Hippel, P. H., Bear, D. G., Morgan, W. D., and McSwiggen, J. A. (1984) *Ann. Rev. Biochem.* **53**:389-446.

Wall, L. A., and Meighen, E. A. (1986) *Biochemstry* **25**:4315-4321.

Wall, L. A., Byers, D. M., and Meighen, E. A. (1984) *J. Bacteriol.* **159**:720-724.

Wall, L., Rodriguez, A., and Meighen, E. (1986) *J. Biol. Chem.* **261**:16018-16025.

Watenabe, H., Inaba, H., and Hastings, J. W. (1991) In *Bioluminescence and Chemiluminescence: Current status.* (P. E. Stanley and L. J. Kricka, eds.) pp. 43-46. John Wiley and Sons.

Westhoff, P., and Gitterman, R. G. (1988) *Eur. J. Biochem.* **171**:551-564.

Winkelman, J. W., and Hatfield, G. W. (1990) *J. Biol. Chem.* **265**:10055-10060.

Winter, E., and Varshavsky, A. (1989) *EMBO* **8**:1867-1877.

Wu, H-. M., and Crothers, D. M. (1984) *Nature* **308**:509-519.

Yoshida, H., Tsuyoshi, K., Yamagishi, J., and Nakamura, S. (1988) *Mol. Gen. Genet.* **211**:1-7.