12-2003

Optimization of Bacterial Luciferase for Expression in Mammalian Cells

Stacey Swiger Patterson
University of Tennessee - Knoxville

To the Graduate Council:

I am submitting herewith a dissertation written by Stacey Swiger Patterson entitled "Optimization of Bacterial Luciferase for Expression in Mammalian Cells." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Microbiology.

Dr. Gary S. Sayler, Major Professor

We have read this dissertation and recommend its acceptance:

Dr. Robert Moore, Dr. Jeff Becker, Dr. Patricia Tithof, Dr. Yun You

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Stacey Swiger Patterson entitled "Optimization of Bacterial Luciferase for Expression in Mammalian Cells." I have examined the final electronic copy of the dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Microbiology.

Dr. Gary S. Sayler

Dr. Gary S. Sayler, Major Professor

We have read this dissertation
and recommend its acceptance:

Dr. Robert Moore

(Dr. Robert Moore)

Dr. Jeff Becker

(Dr. Jeff Becker)

Dr. Patricia Tithof

(Dr. Patricia Tithof)

Dr. Yun You

(Dr. Yun You)

Accepted for the Council:

Dr. Anne Mayhew

Vice Provost and Dean of
Graduate Studies

(Original signatures are on file with official student records.)

OPTIMIZATION OF BACTERIAL LUCIFERASE FOR EXPRESSION IN MAMMALIAN CELLS

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

STACEY SWIGER PATTERSON

December, 2003

DEDICATION

This dissertation is dedicated to my husband

Brian Christopher Patterson

and my son

Andrew Michael Patterson

who have loved and supported me through it all.

They give my life meaning.

ACKNOWLEDGEMENTS

I would like to thank my major professor, Dr. Gary S. Sayler for his patience and guidance throughout this work. Although the outcome was sometimes questionable, he always supported my ideas and allowed me to find my own direction. I would also like to thank the other members of this committee, Dr. Robert Moore, Dr. Patricia Tithof, Dr. Jeff Becker and Dr. Yun You for their advice and input on the completion of this research project.

I would like to thank my friends and colleges, Dr. Hebe Dionisi, Dr. Rakesh Gupta and Dr. Steve Ripp, for their guidance and practical advice. They were all invaluable resources throughout and I will always be indebted to them.

I would like to give a special thanks to my good friend, Kim Cook, who was always there for me. Her friendship has helped me in more ways than she will ever know.

I want to thank my all of my friends and coworkers at the CEB, especially Kim Cook, Hebe Dionisi, Rakesh Gupta, Steve Ripp, Pat Jagier, Leslie Saidak and Jacque Young, Jack McPherson and Liz Mitchell, whose friendship have made this an enjoyable chapter in my life. I will never forget you guys!

I would like to express my sincere thanks to my mom and step-dad, Sharon and Ed Morgan, my dad and step-mom, Arlen and Stephanie Swiger, my grandmother, Mary E. Swiger and my in-laws, Michael and Marian Patterson for there support and understanding. I also have enjoyed the company of my sister Sheena and niece Nakiyaah since they moved to Knoxville. I will really miss them.

Finally, I would like to thank my husband, Brian, who has been there through this long journey and stood by me every step of the way. Without his love and support this would have been much more difficult. Honey, we finally made it!

ABSTRACT

This study describes the optimization of expression of the bacterial luciferase enzyme in mammalian cells. Previous attempts to express this heterodimeric enzyme complex in mammalian cells have been met with only modest success. In this research effort, several vector formats were evaluated to fully determine the optimal format for the expression of these genes. It was determined that the bacterial luciferase enzyme produced optimal bioluminescence in mammalian cells when the genes were cloned and expressed as a bicistronic transcript fused with an internal ribosomal entry site (IRES).

To optimize the enzyme expression further, a novel approach to codon optimize the genes was performed. To accomplish this task, completely synthetic versions of the codon optimized sequences were generated. This codon optimization, led to an increase in bioluminescence levels greater than two orders of magnitude versus the wild type genes.

Additionally, the availability of the $FMNH_2$ substrate was evaluated and determined to be a limiting substrate for the reaction. In an attempt to alleviate this limitation, a flavin oxidoreductase gene (*frp*) from *Vibrio harveyi* was cloned and expressed along with the codon optimized *luxA* and *luxB* genes. Although the expression of this enzyme enhanced the bioluminescence significantly, $FMNH_2$ remains the limiting substrate for optimal bioluminescence.

To produce a usable reporter cell line, the reporter must remain stable within the cells for long periods of time. The overall stability of the engineered cells was assessed to determine the persistence of the reporter for long-term monitoring applications. These data revealed that the luciferase genes were stable in HEK293 cells for more than

forty passages (five months) in culture in the absence of antibiotic, indicating that these cell lines would be stable enough for relatively long term monitoring projects and applications.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Background and Research Objectives

Bacterial luciferase is a powerful reporter protein system since it allows for the development of real-time autonomous sensors that the invasive manipulations required by other reporter proteins do not permit. Mammalian cell lines expressing reporter proteins have been widely used in both basic and applied research for the investigation of a variety of cellular functions. These applications include, but are not limited to, promoter analysis (Guignard et al., 1998; Zhang et al., 1997), identification of transcription factors (Ichiki et al., 1998; Schwechheimer et al., 1998), discovery of genes that are potential targets for disease (Watson et al., 1998) evaluation of cross talk mechanisms (Naylor, 1999), and *in vivo* sensing of tumor and/or disease progression (Contag et al., 1998). However, current mammalian bioreporter technology is limited due to its inability to function as a stand-alone, real-time reporter *in vivo*. Current methodologies that use firefly luciferase (Luc) and green fluorescent protein (GFP) reporter systems in mammalian cells require lysis and substrate addition or exogenous excitation, respectively, to produce a measurable response. Consequently, these cells cannot serve as continuous on-line monitoring devices. Bacterial luciferase is unique in that it is the only bioreporter system available that generates its own substrate, thus eliminating the need for cell destruction or exogenous substrate addition.

Extensive work has been published using the bacterial *lux* system in prokaryotic organisms for the development of whole cell biosensors (Simpson et al., 1998; Sayler et

al., 2001; Ripp et al., 2000; Corbiser et al., 1999; Kohler et al., 2000; King et al., 1990, and VanDyk et al., 2001), advancements in biocomputing applications (Simpson et al., 2001) and *in vivo* imaging models (Contag et al., 1998 and Francis et al., 2001). Recently, a *lux*-based yeast reporter cell line has been developed for the detection of estrogenic compounds (Gupta et al., 2003). This research was the first successful attempt to express the complete *lux* operon required for autonomous bioluminescence in a eukaryotic organism. Unfortunately, this technology has yet to be successfully implemented into mammalian cells. Several attempts by various groups have been made to express bacterial luciferase enzyme in mammalian cells. These efforts have been met with only modest success as numerous obstacles have been encountered preventing efficient expression of the *lux* proteins. A major effort and the first step required to realize the ultimate potential of this technology is to achieve efficient expression of the heterodimeric luciferase (*luxA* and *luxB*) protein.

The bacterial luciferase enzyme is a heterodimeric protein complex made up of an α and β subunit encoded by the genes *luxA* and *luxB*, respectively. Because it is not possible to express multiple genes as a polycistronic operon in eukaryotes, alternate expression platforms are needed to obtain optimal thermostability and proper folding which should aid in obtaining an adequate bioluminescent signal from mammalian cells. Furthermore, the availability of the co-factors required for the *lux* reaction including $FMNH_2$ and O_2 in mammalian cells has been suspected to be inadequate and levels need to be evaluated. Other strategies for bioluminescence optimization and possible gene amplification have not been previously pursued and the potential is unknown. This avenue of research may result in a mammalian cell line able to produce the

bioluminescent levels required for sensitive monitoring of target analytes. Also, for an autonomously driven bioluminescent mammalian cell line to realize its full technological potential, the expression system must remain stable for long periods of time without the need for selective pressure. The stability of mammalian cell lines harboring the luciferase protein remains a question and needs to be evaluated. In response to these questions, the following hypotheses are tested in this research:

- **Hypothesis 1**: Expression of the bacterial luciferase (*lux*) subunits as individual proteins rather than as a monomeric translational fusion results in efficient folding and thermostability resulting in a higher bioluminescent signal in mammalian cells.

- **Hypothesis 2**: Codon optimization of the bacterial luciferase (*lux*) genes is required to significantly enhance translation of the message and ultimately result in greater bioluminescence levels from mammalian cells harboring these optimized genes.

- **Hypothesis 3**: Stably integrated constructs will be persistent in the absence of selective pressure for long periods of time.

- **Hypothesis 4**: Mammalian cells possess or can be engineered to express adequate available concentrations of the required co-factor $FMNH_2$ for efficient bioluminescence

Literature Review

Reporter Gene Technology

Reporter genes are defined as genes that produce a measurable phenotype that can be distinguished from other proteins within a cell (Alam and Cook, 1990). The use of reporter genes for applied technologies requires that the signal be easily detected and respond rapidly, possess a linear response to the target analyte for quantification and produce a reproducible signal (Naylor, 1999 and Wood, 1995). Several reporter proteins have been shown to be valuable tools for advancing both basic and applied research. Examples of such basic applications include the development of reporter fusions for the identification and analysis of promoter regions (Guignard et al., 1998; Zhang et al., 1997), identification of transcription factors and induction/repression schemes (Ichiki et al., 1998; Schwechheimer et al., 1998), as well as the discovery of genes as potential targets for disease (Watson et al., 1998) and evaluation of cross talk and signal transduction mechanisms (Naylor, 1999). Furthermore, reporter gene fusions have been utilized for the creation of whole cell biosensors for environmental monitoring (King et al., 1990; Ripp et al., 2000 and Kohler et al., 2000), advancement of biocomputing applications (Simpson et al., 2001) and *in vivo* imaging of disease onset and progression (Francis et al., 2001) as well as drug efficacy screening (Contag et al., 1998).

Reporter genes can be used to study any pathway that is controlled on a transcriptional level. The signals produced are responses to alterations in either gene regulation or expression within the cell (Wood, 1995). Eukaryotic reporter systems require receptor proteins for sensing and shuttling of analyte compounds. These

analytes can be extracellular signals that are either detected on the cell surface and trigger intracellular signal transduction pathways or bind an intracellular receptor which can subsequently bind to specific response elements resulting in transcriptional activation (Naylor, 1999). In either case, by fusing reporter proteins to promoter elements within the target pathway, when promoter induction occurs, the reporter protein is generated and a detectible phenotypic change occurs within the cell (Figure 1). This reporter signal is then measured and provides a simple way to determine if and when a particular analyte affects gene expression (Levitzki, 1996). In prokaryotic reporter cell schemes, generally the target analyte can bind directly to the promoter or repressor element and induce transcription directly (Figure 2) (Kohler et al., 2000). Nevertheless, in either case the ultimate outcome is the same and a detectable and often times quantifiable signal is produced.

Reporter Genes and Proteins

Several reporter proteins have been shown to be valuable tools in various areas of research. In order for a reporter protein to be useful, the generated signal must have a low endogenous background level in the host cell and produce an easily detectable response. Choice of the optimal reporter protein for individual applications is essential for success. There are several criteria that should be considered in the selection process of a reporter gene. (1) The reporter protein should be absent from the host to prevent complications in distinguishing signal from background noise. (2) The assay for signal detection should either be established or easily measured in a rapid, simple, and cost-

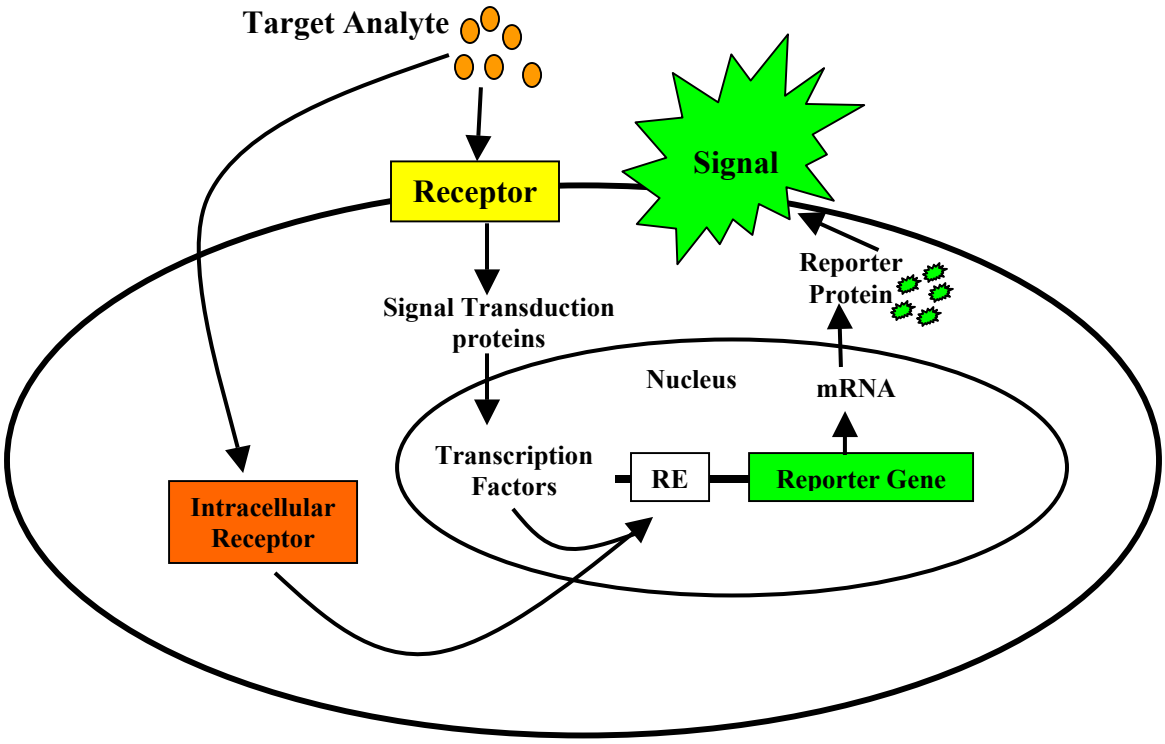


Figure 1. Schematic representation of a generalized eukaryotic bioreporter cell. Response elements (RE) are specific for the target analyte.

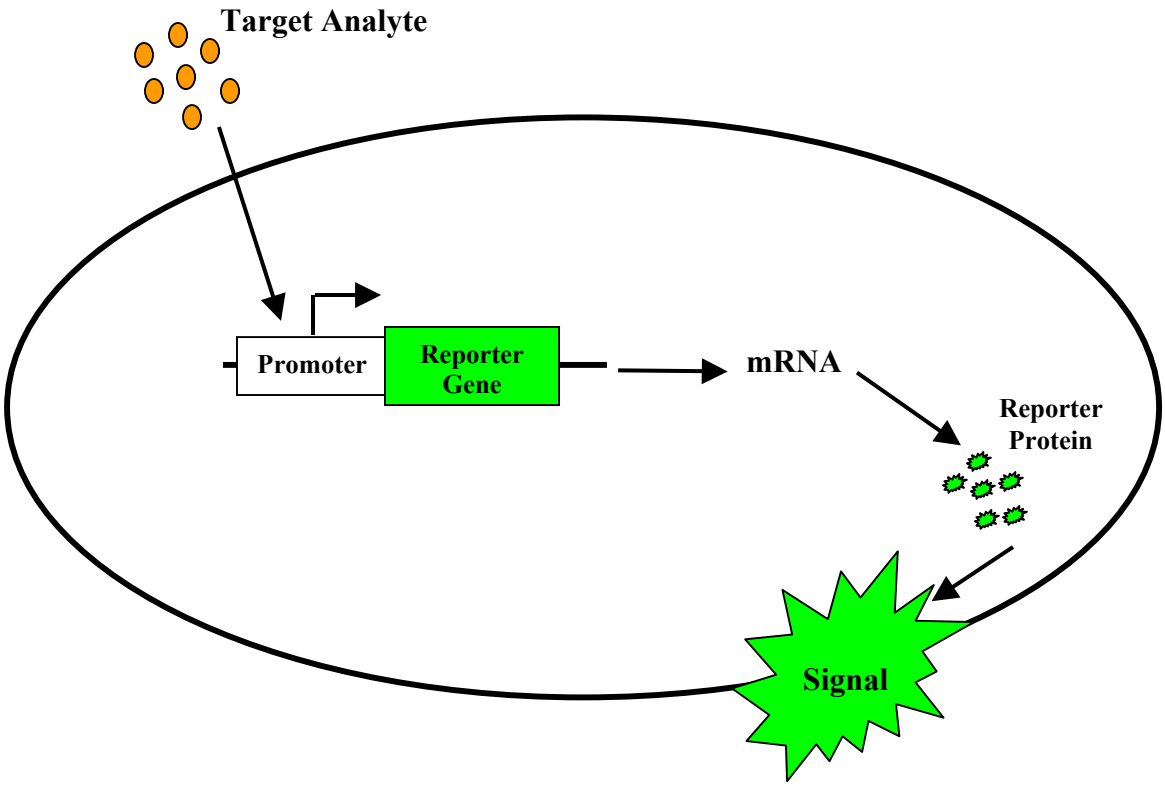


Figure 2. Schematic diagram of a generalized prokaryotic bioreporter cell. The promoter region is specific for the target analyte.

effective manor. (3) In order to analyze induction of both small and large changes in transcriptional activation, the reporter should have a large linear dynamic range for its response (4) The expression of the reporter gene must not effect the overall physiology of the cell (Ausubel et al., 1997). Each reporter protein identified to date has distinct advantages and disadvantages for application. Table 1 summarizes some of these differences between a few of the more commonly used reporter proteins available.

There are two main classes of reporter assays available, *in vitro* and whole-cell bioassays. *In vitro* reporter applications refer to the protocols in which the transcriptional activation is quantified in cell lysates or in the media from excreted proteins. The measurement can be a direct quantification of the protein or an indirect response to enzymatic or immunological stimulation (Alam and Cook, 1990). Although these methodologies may be useful under certain circumstances, *in vivo* or whole cell assays provide more reliable data for studies comparing promoter strengths, enhancer regions and determining other cell requirements. Of all of the known reporter proteins, the bacterial luciferase (*lux*) has the distinct advantage in that it is the only bioreporter system available able to make its own substrate and generate an autonomous signal. This property has made the bacterial bioluminescence reporter system an invaluable tool for the creation of whole cell biosensors for remote sensing in prokaryotic organisms (Sayler et al., 2001). Unfortunately, the one caveat in this technology to date is that it has not been efficiently expressed in mammalian cells limiting its full potential (Meighen, 1991 and Naylor, 1999).

Table 1. Comparison of commonly used reporter proteins used for sensing applications.

Reporter	Advantage	Disadvantage
Chloramphenicol acetyltransferase (CAT)	No endogenous activity. Automated detection (ELISA)	Requires the addition of substrate and separation of substrate and product.
β-galactosidase	Stable, Simple colorimetric and chemiluminescent assay available.	Endogenous activity (mammalian cell). Requires the addition of substrate.
Firefly luciferase (LUC)	High specific activity, no endogenous activity, easily detectable	Requires addition of substrate (luciferin), O_2 and ATP.
Green Fluorescent Protein (GFP)	Autofluorescent Mutants with altered spectral qualities available.	Moderate sensitivity. Background fluorescence may interfere. Requires exogenous excitation
Bacterial luciferase (*lux*)	Broad dynamic range, easily measured, no exogenous substrate addition required.	Requires O2 only expressed in prokaryotes

Bacterial Bioluminescence

Organisms able to generate light have intrigued researchers for centuries. Species able to produce bioluminescence are diverse, ranging from fireflies and mushrooms to dinoflagellates and bacteria (Harvey, 1952). The luciferase enzymes involved in the catalysis of the bioluminescence reaction are also evolutionarily very diverse with the only one true commonality being that they all require O_2 as a co-factor (Fisher et al., 1996).

Bioluminescent bacteria are the most abundant of the light emitting organisms (Meighen, 1991). They can be found in a high abundance in marine, freshwater and terrestrial environments (Hastings et al., 1985). Most bioluminescent bacteria have been classified into three genera: *Vibrio*, *Photobacterium* and *Photorhabdus* (formerly *Xenorhabdus*). Organisms belonging to the first two genera generally can be found in marine environments. These organisms have been identified as free-living planktonic bacteria and symbionts with a variety of fish and squid species (Wilson and Hastings, 1998). The *Photorhabdus* genus contains strains that can colonize terrestrial organisms and tend to be found acting in symbiosis with worms and caterpillars (Farmer et al., 1989 and Colepicolo et al., 1989).

Biochemistry of Bacterial Bioluminescence

In all bioluminescent organisms, the enzymes that catalyze the luminescent reaction are referred to as luciferases, while the required substrates are luciferins (Wilson and Hastings, 1998). Further, this light producing reaction requires molecular O_2, the reducing power of $FMNH_2$ and the energy of ATP as co-factors and substrates.

The bacterial luciferase enzyme is a heterodimeric protein encoded by the *luxA* and *luxB* genes. The heterodimeric complex forms a 77 kDa enzyme comprised of α (40 kDa) and β (37 kDa) subunit polypeptides. Because the subunits are related (>30% amino acid homology) they are thought to be products of a gene duplication event (Baldwin et al., 1979). The complete luciferase enzyme is a flavin monooxygenase that binds a reduced flavin molecule as a specific substrate. However, only the α subunit carries the active center (Fisher et al., 1995). A specific role of for the β subunit has not become clear, but its presence is essential for a high quantum yield reaction (Baldwin et al., 1995). Nevertheless, the β subunit has been shown to have some impact on the enzyme's thermostability (Meighen et al., 1971; Cline and Hastings, 1972 and Szittner and Meighen, 1990), binding of $FMNH_2$ (Cline, 1973; Meighen and Bartlett, 1980; Welch and Baldwin, 1981 and Watanabe et al., 1982) as well as efficient binding of aldehyde (Tu and Henkin, 1983).

All bacterial luciferases studied to date catalyze the same overall reaction:

$$FMNH_2 + O_2 + RCHO \rightarrow FMN + RCOOH + H_2O + hv \ (\lambda_{max} = 490nm)$$

The natural aldehyde for the reaction is thought to be tetradecanal in most species of luminescent bacteria, however, the more thermostable forms of luciferase (*Vibrio harveyi* and *Photorhabdus luminescens*) tend to produce higher bioluminescence in the presence of dodecanal and decanal (Schmidt et al., 1989). The general 1:1 stoichiometry of the luciferase subunits is conserved throughout all species of bioluminescent bacteria known (Meighen, 1991). However, the amino acid sequence

of each polypeptide is somewhat diverse. Certain motifs within the luciferase enzyme are conserved throughout. There is a 60-85% identity between amino acid sequences of the α subunit of reported various *Vibrio, Photobacterium* and *Photorhabdus* species while the β subunit is less conserved with a 50-65% identity (Szittner and Meighen, 1990). The higher conservation of the α subunit sequence may be a direct reflection for the need to conserve the active center and catalytic properties of luciferase (Meighen et al., 1971).

Although the aldehyde substrate is not necessary for the luciferase reaction itself, its presence significantly increases the light output kinetics (Volkova et al., 1999). The genes required for synthesis of aldehydes are catalyzed by a multienzyme fatty acid reductase and synthase (Rodriduez et al., 1983). These genes are all located within the *lux* operon of all bioluminescent bacteria. Generally, the reductase (*luxC*) and transferase (*luxD*) are located upstream of the *luxA* and *luxB* luciferase genes while the synthase (*luxE*) gene is located immediately downstream. The primary reaction catalyzed by this system is the reduction of fatty acids by the reductase and synthase enzymes. The synthase acts to activate the fatty acid, which results in a fatty acyl-AMP intermediate that remains bound to the enzyme. The acyl group is then transferred to the synthase and then further transferred to the reductase, where it becomes reduced by NAD(P)H to the corresponding aldehyde. The transferase subunit is responsible for the transfer of activated fatty acyl groups. The fatty acid is then recycled. Each multienzyme complex responsible for this reaction has been found to consist of a central tetramer of reductase subunits bound to one synthase and one weakly associated transferase subunit (Li et al., 2000).

To produce light, the bioluminescence reaction appears to go through several intermediate steps. With a stoichiometry of 1:1:1, the α and β subunits bind with a reduced flavin molecule ($FMNH_2$) to form a C4a hydroxyflavin. As this hydroxyflavin becomes dehydrated to FMN, a blue-green light is emitted (Baldwin et al., 1979). A schematic diagram of the overall reaction is shown in Figure 3. This enzymatic reaction has a relatively slow turnover rate.

The luciferase subunits have been shown to fold independently and interact during the folding process. They then form an active heterodimeric complex following isomerization (Ziegler et al., 1991). However, because of the relatedness of the two subunits, if they are present individually (in the absence of the other subunit) they tend to form inactive homodimers that cannot refold into the active heterodimeric form (Waddle et al, 1987 and Ziegler et al., 1991). Because of this complex stoichiometric requirement for folding, there is a lag time of at least three to four minutes to complete an active enzyme after translation of the subunit polypeptides (Ziegler et al., 1993). A schematic diagram of the folding pattern of the luciferase subunits is shown in Figure 4.

Thermostability of Bacterial Luciferase

Bacterial luciferase (*lux*) genes cloned from various species of luminescent bacteria have been used to create a myriad of reporter constructs. Given that the *lux* operons from *V. fischeri* and *V. harveyi* were the first to be cloned (Engebrecht et al., 1983 and Cohn et al., 1983), the vast majority of these clones are derived from these sequences. However, just as selection of the appropriate reporter for individual

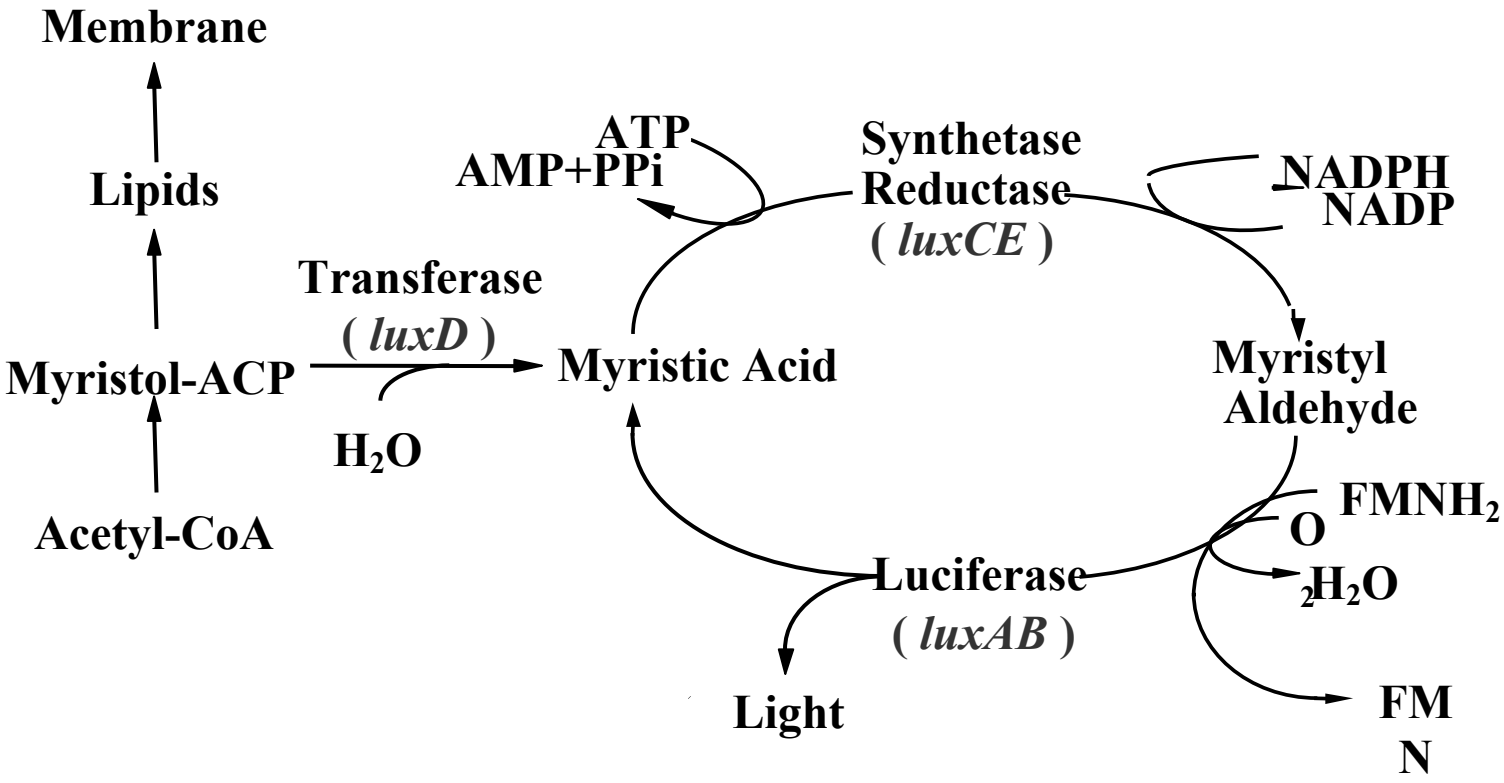


Figure 3. Biochemistry of the bacterial bioluminescence reaction. (Figure Courtesy of A. Heitzer.)

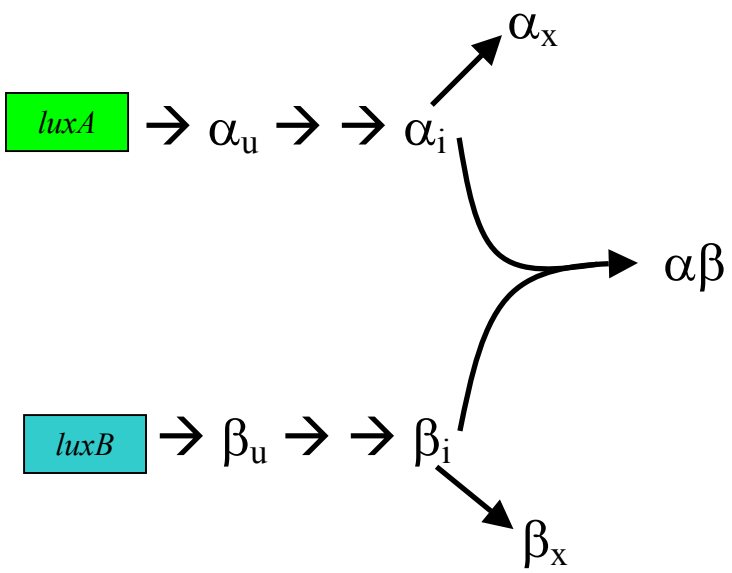


Figure 4. Schematic representation of the folding of an active luciferase (αβ). *luxA* and *luxB* represent the individual genes, u denotes the unfolded form of the polypeptides, i denotes the inactive form before dimerization, and x denotes the homodimeric form that can no longer form an active luciferase. Figure adapted from Zeigler et al., 1993.

applications is important, the selection of the appropriate luciferase is equally critical. A key difference between luciferase enzymes from different species is their overall thermostability *in vivo*. The luciferase from *V. fischeri* has been shown to be heat labile at temperatures above 30°C losing almost all catalytic properties at 37°C (Meighen, 1991). The *V. harveyi* luciferase remains relatively stable at 37°C and luciferase enzymes from *P. luminescens* are quite stable at 42°C (Szittner and Meighen, 1990). Furthermore, the luciferase enzyme from *P. luminescens* is optimally bioluminescent at 37°C. For selection of the application appropriate enzyme, the optimal growth temperature for the host should be considered. For the ultimate expression of the bacterial luciferase in mammalian cells, *P. luminescens* would appear to be the appropriate choice and therefore, is the enzyme that was chosen in this research.

Use of the Bacterial Luciferase (*luxCDABE*) Reporter System

Various *lux*-based reporter systems have been constructed mostly by the insertion of a specific promoter in front of the *lux* cassette on either a plasmid or transposon and then mobilizing the plasmid into the appropriate strain of bacteria. The various constructs that have been designed are too numerous to completely review in this document. Briefly, whole cell bioreporters have been generated to monitor the catabolic genes involved in degradation pathways including but not limited to; naphthalene (Burlage et al., 1990), toluene (Applegate et al., 1997), and *m*-toluate (deLorenzo et al., 1993). Van Dyk et al. (1995) used *lux* fusions to monitor

heat shock gene expression and then later expanded this technology to monitor genome wide expression patterns in *E. coli* (Van Dyk et al., 2001). Lux-based reporters have also been used to monitor DNA damage (Vollmer et al., 1997), oxidative stressors (Wallace et al., 1994) and in the creation of countless whole-cell biosensors for monitoring compounds like nitrate (Prest et al., 1997), arsenic (Cai et al., 1997), nickel (Tibazarwa et al., 2000), lead (Corbiser et al., 1996), 2,4-D (Hay et al, 2000) and iron (Khang et al., 1997). Lux fusions have been further used for the *in vivo* monitoring of pathogenic infection in whole mouse models (Contag et al., 1995; Francis et al., 2000 and Francis et al., 2001).

Advantages and Disadvantages of Using Bacterial Luciferase as a Reporter

Use of the bacterial luciferase gene system has several advantages over other bioreporter systems available. First, the absence of background luminescence in a nonluminescent host makes this a very attractive system because the lower end of the signal detection is only limited by the noise within the detector itself. With the development of new, more sensitive detectors and noise reduction schemes, very low concentrations of luciferase activity can be detected and quantified. Secondly, the light intensity has been shown to be a direct measurement of the amount of luciferase present (Meighen, 1991). This linear detection range is very wide relative to other reporter proteins available (Meighen, 1991). Furthermore, the luminescent signal can be detected within a matter of seconds making the assay relatively quick and easy for the user.

The ability to directly measure *in vivo* gene function without the disruption of the cell or loss of cell viability is perhaps the most obvious advantage of the bacterial luciferase enzyme system. This property has made the *lux* system sought out by many for various research applications. As a result, numerous prokaryotic biosensors have been developed. These sensors have been employed in a variety of applications ranging from environmental pollutant monitoring (Sayler et al., 2001; Ripp et al., 2000; Corbiser et al., 1999; Kohler et al., 2000; King et al., 1990) to visualizing infections *in vivo* (Francis et al., 2001). The biggest limitation to date is the inability to efficiently express the *lux* system in eukaryotic organisms. Recently, it was shown that the complete *lux* operon from *P. luminescens* can be expressed in the yeast *S. cerevisiae* as proof in principle for the further application into mammalian cells (Gupta et al., 2003).

However, the exact interpretation of light levels from intact cells is complicated because the intensity depends not only on luciferase concentrations within the cell, but also the availability of the aldehyde and $FMNH_2$ substrates. As this system is moved into higher eukaryotes this measurement may become more complicated (Meighen, 1991).

Expression of Bacterial Luciferase in Mammalian Cells

Unlike polycistronic expression of multiple genes often found in bacterial systems, eukaryotic gene expression requires that each individual gene be preceded by its own promoter. This has limited the expression of the *lux* genes in eukaryotes to this point. To overcome this, several researchers have generated a monocistronic

version of the *V. harveyi* luciferase protein by creating a translational fusion of the individual subunits. This fusion protein allows for the expression of both genes from a single promoter element. Unfortunately, these fusion proteins have not produced adequate bioluminescence *in vivo* to generate a reliable mammalian sensor for gene expression analysis. This loss of activity has been attributed to improper folding and low thermostability of the fusion (Kirchner et al., 1989; Olsson et al., 1989; Almashanu et al., 1990; Escher et al., 1989; Costa, 1991; Pazzagli et al., 1992; Gelmini et al., 1993). Bioluminescence levels were significantly increased if the host cells were grown at lower temperatures (Escher et al., 1989; Costa et al., 1991 and Almashanu et al., 1990). Based on these data, it was determined that the fusion was unable to properly fold into its active heterodimeric form at 37°C. This was thought to be caused, in part, by the short polylinker region that separates the two subunits. It was hypothesized that a short linker between the two genes may impose an unnatural strain on the dimerization process and limit the amount of active heterodimer able to form. To overcome this limitation, several attempts have been made to alter this polylinker region and allow for a more natural folding of the two subunits. The number of linker codons tested has ranged from one to twenty-two. The relative activities (expressed in *E. coli*) of the enzyme are lowest with a short (one amino acid) linker ranging from 0.04% (Boylan et al., 1989) to 19% (Almashanu et al., 1990). The highest activities were obtained with a ten amino acid polylinker, which produced 90% activity when grown at 23°C but only 8% at 37°C. None of the *Vibrio harveyi* fusion proteins reported to date have shown the ability to remain stable at the optimal mammalian growth temperature, 37°C.

Interestingly, if cells, prokaryotic or eukaryotic, harboring the *lux* fusion genes were grown at 37°C and then cooled to 23°C, a significant increase in bioluminescence levels were shown to occur (Costa, 1991 and Escher et al., 1989). On the other hand if chloramphenicol was added to the cells prior to the cooling step, the bioluminescence remained low (Esther et al., 1989). These data indicated that the light levels produced in the absence of the antibiotic were a result of *de novo* synthesis and suggest a problem in folding of the protein at higher temperatures that can not be recovered after folding is complete (Esther et al., 1989). These data further support the model of *luxA* and *luxB* folding proposed earlier by Ziegler et al. (1991).

Koncz et al. (1987) reported the expression of the heterodimeric bacterial luciferase protein from *V. harveyi* as individual proteins expressed in a dual promoter vector format. The expression levels, although difficult to compare were said to be adequate for monitoring chimeric genes in plant extracts (Koncz et al., 1987). These data also showed for the first time that individual subunits of the protein could be transcribed and translated separately and subsequently assemble to form a functional luciferase enzyme in a eukaryotic cell.

Internal Ribosomal Entry Sites (IRES)

In prokaryotes, translation of multiple adjacent genes within a single operon is common. In these cases, the entire operon is transcribed as a single mRNA regulated by the upstream promoter region. The translation of the mRNA is then initiated by direct complementary base pairing between the 16S rRNA and mRNA Shine-

Dalgarno sequence. Upon binding, the 30S subunit is brought into an internal position around the start codon (AUG) where translation of the protein begins. Translation of several proteins from one mRNA is possible (Jackson, 2000). This type of translation scheme is referred to as polycistronic translation.

Until recently, it was believed that eukaryotic translation was limited to cap dependent initiation. This mechanism involves a methyl-7-G(5′)pppN structure (cap) at the 5′ end of the mRNA that is recognized by a protein complex initiation factor to begin translation (Hershey and Merrick, 2000). This initiation complex scans the mRNA for the first AUG triplet downstream of the terminal 5′ cap usually within 50 to 100 bases where translation begins (Hennecke, 2001). In this type of initiation, the simultaneous translation of multiple proteins from one mRNA is not possible and monocistronic translation is the only option. Typically, each open reading frame is transcribed and translated independently from its own promoter.

More recently, alternative translation mechanisms have been identified that have been shown to initiate translation in a cap-independent manor in eukaryotic organisms and their viral pathogens. These alternative initiation schemes were first identified within the genomes of poliovirus and encephalomyocarditis virus (EMCV) (Jackson, 1988 and Jang et al., 1988). These viral sequences are naturally uncapped at their 5′ ends. They possess several complex features that would be predicted to impair efficient ribosome binding (Vagner, 2001). Nevertheless, protein translation was shown to be initiated at these sites, both *in vitro* and *in vivo* after viral infection

(Martinez-Salas et al., 2001). Furthermore, these sequences, termed internal ribosomal entry sites (IRES) were found in all genera of picnovirus (Vagner, 2001).

IRES elements can be defined as specific nucleotide sequences that allow for ribosomal entry and translation initiation directly at the start codon (AUG) rather than requiring scanning from the 5′ end, cap structure, of the mRNA (Pestova et al., 2001 and Kozac, 2001). IRES activity is based on the secondary structure of the mRNA and has been shown to be extremely sensitive to even point mutations that may alter the integrity of this structure (Haller and Semler, 1992). IRES elements from various sources, however, have been shown to lack conservation of primary sequences (Pestova et al., 1991). Known IRES elements also vary greatly in their overall length, ranging from 200nt in insect RNA viruses (Wilson et al., 2000) to as large as 600nt in picnovirus IRES elements (Nicholson et al., 1991). Along with these variations in nucleotide sequence and size, IRES elements have been shown to have varying mechanisms from translation initiation (Martinez-Salas, 1999). However, certain secondary structures remain constant and have been shown to be important for the initial physical contact with the 40S ribosomal subunit for translation initiation. Examples of these specific regions include, double stranded mRNA segments and hairpin loop structures (Honda et al., 1996 and Honda et al., 1999).

The IRES element isolated from EMCV has been shown to initiate translation by ribosomal binding at codons close to the 3′ border of the IRES sequence (Kalinski et al., 1990). Unlike binding to the ribosomes to a Shine-Dalgarno (linear) sequence, IRES binding to the 40S subunit is determined by several noncontiguous sequences (Pestova et al., 2001). Whether the IRES and 18S rRNA physically bind is still yet

to be determined. Once the ribosome binds an elF (elongation) initiation factor stimulates the formation of the 48S complex and forms internal loops on the IRES sequence (Jubin, 2000 and Pestova et al., 1996). Then, in an ATP dependent manor, translation begins directly at the AUG start codon.

Most IRES elements identified to date, represent an evolutionary survival scheme for viruses upon infection. Once the virus infects the host cell, the cap dependent translation machinery is shut down and only the viral proteins are made (Vagner, 2001). Interestingly, several eukaryotic cellular IRES elements have also been identified. The first cellular IRES was a 220nt 5′ untranslated region (UTR) of the immunoglobulin heavy chain binding protein (BiP). This protein was shown to be highly translated after viral infection and thus in a cap-independent manor (Macejak and Scarnow, 1991). Other cellular IRES elements have since been identified and shown to be related to various stress responses. For example, anti-apoptotic genes have been shown to use IRES elements for translation initiation of proteins, like Apaf-1 (Coldwell et al., 2000). Furthermore, translation initiation factors have been shown to become translated in this fashion, as well including DAP5 (Henis-Korenblit et al., 2000) and ELF4G (Johannes and Sarnow, 1998). It has been hypothesized that translation from IRES elements may have been selected for as a last stitch effort to survive harsh conditions by providing a failsafe method to ensure synthesis of certain proteins under specific physiological conditions (Pestova et al., 2001). A list of viral and cellular IRES elements that have been identified to date are listed in Table 2.

Several IRES elements have been used to create bicistronic expression vectors for the co-expression of multiple genes from the same promoter (Wong et al., 2002

Table 2. Examples of viral and cellular IRES elements identified to date.

Virus/gene Type	Virus/gene	Reference
Viral RNAs		
Picnoviruses	Poliovirus (PV)	Pelletier and Sonenberg, 1988
	Encephalomyocarditis virus (EMCV)	Jang et al., 1988
	Foot and mouth disease virus (FMDV)	Kuhn et al., 1990
Flavivirus	Hepatitis C virus (HCV)	Reynolds et al., 1995
Pestivirus	Classical Swine fever virus	Pestova et al., 1998
Retrovirus	Murine leukemia virus	Berlioz and Darlix, 1995
Lentivirus	Simian immunodeficiency virus	Ohlmann et al., 2000
Insect RNA virus	Cricket paralysis virus	Wilson et al, 2000
Cellular mRNAs		
Translation initiation factors	ElF4G	Johannes and Sarnow, 1998
	DAP5	Henis-Korenblit et al., 2000
	Initiation factor G4	Wong et al., 2002
Transcription factors	c-Myc	Stoneley et al., 2000
	NF-*k*B-repressing factor	Oumard et al., 2000
Growth Factors	Vascular endothelial growth factor	Huez et al., 1998
	Fibroblast growth factor	Creancier et al., 2000
	Platelet-derived growth factor B	Bernstein et al., 1997
Homeotic genes	Antennapedia	Oh et al., 1992
Survival Proteins	X-linked inhibitor of apoptosis	Holick and Korneluk, 2000
	Apaf-1	Coldwell et al., 2000
Miscellaneous	BiP	Macejak and Sarnow, 1991
Yeast	p150	Zhou et al., 2001

and deFelipe, 2002). These new expression vectors have gained much attention due to their potential impact as valuable tools for drug and gene delivery vectors for multi-drug combined therapies for treating diseases such as cancer and AIDS (deFelipe, 2002). Consequently, for the expression of multigene enzyme systems like bacterial luciferase in eukaryotes, IRES based bicistronic vectors may prove to be an invaluable tool.

Several IRES elements have been tested for implementation into this format and improved vector development. Some IRES elements, like the IRES isolated from poliovirus, are vulnerable to adjacent gene placement (Mosser et al., 2000). The EMCV IRES element has been shown to be immune to these types of effects and upstream genes have little effect on the downstream gene expression (Gorski and Jones, 1999). This property has made the EMCV IRES the most frequent choice for creating reliable, high expression bicistronic vectors (Meilke et al., 2000 and Harries et al., 2000). However, by expressing two genes in a promoter-gene1-IRES-gene2 format, the expression of the second gene has been shown to possess lower overall expression levels ranging from 6-100% activity when compared to the first gene (Mizuguchi et al., 2000). This has been regarded as typical expression levels from these vectors and therefore, to determine optimal expression, several clones must be tested (Clontech Corporation, personal communication).

New IRES elements are being frequently discovered and better options for optimal expression of multiple genes will become available. Wong et al., 2002 report that a newly isolated IRES element from eukaryotic initiation factor G4 produces more than 100 fold higher overall expression of the genes that follow it compared to

other IRES elements. This element possesses enhancer abilities as well as acting as a ribosomal entry site. By using this IRES element or others identified in the future, it will be possible to create bicistronic and possibly polycistronic eukaryotic expression vectors with enhanced expression of each gene from a single promoter.

Codon Usage

The standard term "universal genetic code" comes from the fact that there are 64 possible codons coding for only 20 amino acids. Although the genetic code is degenerate, the alternate synonymous codons are not used with equal frequency (Sharp et al., 1988). In fact, it has been shown that in multivariate analysis, that each species has a major trend in codon usage among genes (Schultz and Yarus, 1996). This trend has also been shown to differ from highly expressed versus lower expressed genes within the same species (Aota and Ikemura, 1986 and Sharp et al., 1986). Several more distinct patterns in codon usage become apparent when genes are sorted into the top and bottom 10% of protein activity within the cell (Sharp et al., 1988). Within these general trends, it has been shown that there is not only a selective difference but also a preference for certain codons in highly expressed proteins (Sharp et al., 1993). The genes that encode for these proteins have a highly biased codon usage pattern with a higher frequency of optimal codons used and a lower frequency or absence all together of the other possible codons (Grantham et al., 1981; Ikemura, 1985 and Sharp et al., 1993). Lowly expressed genes have been shown to possess a more random pattern of codon usage (Hoekema et al., 1987). In fact, Gouy and Gauter (1982) showed that the frequency of optimal codons in a

particular gene was highly correlated with its expression in *E. coli*. The same scenario was seen with genes expressed from *S. cerevisiae* (Sharp and Cowe, 1991).

However, it should be noted that a gene is not necessarily expressed at a low level simply because it is made up of low frequency codons (Kurland, 1991). Codon usage is not the only factor involved in gene expression. Several other factors have profound impacts on the expression levels of genes in all organisms. Non-inhibitory flanking sequences that surround the gene and optimal ribosomal binding sites are critical (Nassal et al., 1987). Nevertheless, codon bias does play a key role in the expression efficiency in all species tested to date (Amicis and Marchetti, 2000).

The question has arisen, to whether or not this increase in gene expression is due to a more efficient translation of the protein or some other factor. The answer to this questions still remains unclear, however, evidence is building that it may be a combination of factors. It has been shown, that optimal codons are codons to which the species in question possesses an overabundance of that particular tRNA molecule. Furthermore, these optimal codons are translated faster than their lower frequency counterparts (Sorensen et al., 1989), which is thought to lead to a more efficient translation (Anderson and Kurland, 1990). However, the speed of overall protein translation has not been shown to be significantly affected (Kurland, 1991).

Other hypotheses for codon optimization having a direct impact on protein expression have been set forth including a reduction of *cis* acting inhibitory elements (AU rich regions) (Kofman et al., 2003a) and an overall increase in mRNA stability (Kofman et al., 2003b). Kofman et al. (2003b) also proposes that there may be an inefficient processing and transport from the nucleus of mRNAs possessing lower

frequency codons in eukaryotes. This idea was supported by the fact that if the genes were expressed on a plasmid (in the cytoplasm) the activity of the protein was somewhat recovered.

It is unclear to this point if codon usage has a specific regulatory function in cells. Rare codons are present more frequently in the 5′ end of lowly expressed genes in *E. coli* (Goldman et al., 1995). It is not known, however, if the optimization of only 5′ sequences would be enough to significantly enhance expression to reach maximal protein activity (Vervoort et al., 2000).

Codon optimization is the term given to a synthetic creation of a gene sequence to possess the optimal codon usage patterns for the host organism. Several examples of codon optimization have been recently published. These optimized proteins have been primarily designed for expression in mammalian hosts, as mammalian expression of foreign genes is often times limited (Narum et al., 2001). Some codon optimization schemes have also been designed to optimize human genes for expression in yeast or bacteria to provide for a simpler protocol for investigation or to generate large quantities of individual proteins (Baev et al., 2001).
Disbrow et al. (2003) and Arregui et al. (2003) have used codon optimization to optimize the expression of the poorly expressed E5 and E7 proteins from human papillomaviruses (HPV) that have been shown to have early transformation activity on infected cells. They have shown that codon optimization was able to increase expression as much as 100 fold versus the wild type. This overexpression of the protein resulted in cell death to much of the population expressing the codon optimized protein. Based on this increase in cell mortality, they further hypothesized

that the virus may have selectively evolved to use a different coding pattern from the host as a way to survive *in vivo* for longer periods of time (Arregui et al., 2003).

Other groups have used codon optimization to increase expression of viral or bacterial proteins for the efficient generation of antibodies. The amount of DNA required to produce a high titer of antibody is significantly reduced if the codon usage patterns of the genes are optimized (Narum et al., 2001 and Deml et al., 2001).

Codon optimization has also been used to increase the efficiency of reporter proteins for expression in mammalian cells. Zhang et al., 2002 optimized the green fluorescent protein (GFP) from *Aequrea victoria* for enhanced expression in mammalian cells. This enhanced GFP gene, EGFP, was shown to make the protein 35 times brighter than the wild type version (Zhang et al., 2002). This same idea has been used to optimize the expression of *Renilla* luciferase proteins in mammalian cells (Gruber and Wood, 2000). In both of these instances, the reporter protein became a stronger reporter for gene expression and reliable monitoring formats.

CHAPTER 2

DETERMINING AN OPTIMAL EXPRESSION FORMAT FOR EXPRESSION OF THE BACTERIAL LUCIFERASE GENES (*luxA* AND *luxB*)

Introduction

Mammalian cell lines expressing reporter proteins are commonly used in both basic and applied research. Current methodologies that depend on firefly luciferase (Luc) and green fluorescent protein (GFP) reporter constructs in mammalian cells are limited due to the required cell lysis, substrate addition and/or exogenous excitation to evoke a measurable response. Consequently, these reporter constructs cannot be implemented into continuous, real-time, on-line monitoring devices or strategies. Bacterial luciferase is unique in that it is the only bioreporter system available that generates its own substrate, thus allowing for autonomous signal generation. Unfortunately, the bacterial luciferase system's potential has not been realized in mammalian cells because of difficulties encountered with efficient expression of this multi-enzyme system.

Unlike polycistronic expression of multiple genes often found in bacterial systems, eukaryotic gene expression generally requires that each individual gene be preceded by its own promoter. This has limited the expression of the *lux* genes in eukaryotes to this point. In an attempt to overcome this, several researchers have generated a monocistronic version of the *V. harveyi* luciferase protein by creating a translational fusion of the individual subunits. Unfortunately, these efforts have been met with only modest successes. The loss of bioluminescence activity has been attributed to improper folding and low thermostability of the fusion protein (Kirchner et

al., 1989; Olsson et al., 1989; Almashanu et al., 1990; Esther et al., 1989; Costa, 1991; Pazzagli et al., 1992; Gelmini et al., 1993). None of the fusion proteins reported to date have shown the ability to remain stable at the optimal mammalian growth temperature, 37°C.

Koncz et al. (1987) reported the successful expression of the heterodimeric bacterial luciferase protein from *V. harveyi* as individual proteins expressed in a dual promoter vector format. The expression levels, although difficult to compare were said to be adequate for monitoring chimeric genes in plant extracts (Koncz et al., 1987). These data also showed for the first time that individual subunits of the protein could be transcribed and translated separately and subsequently assemble to form a functional luciferase enzyme in a eukaryotic cell. Gupta et al. (2003) showed that by linking the *lux* genes transcriptionally with IRES elements, the complete *lux* operon could be efficiently expressed in the yeast, *S. cerevisiae*. However, no one expression format has been shown to be the optimal choice for expression of the bacterial luciferase genes in mammalian cells and therefore no mammalian reporter systems are currently available that utilize this uniquely powerful reporter system. Therefore, further research is needed to identify the optimal expression format for the heterodimeric luciferase protein in mammalian cells. In this research effort the specific objectives are:

- To construct and evaluate the overall bioluminescence potential from a constitutively expressed *luxAB* fusion protein, a dual promoter vector harboring both the *luxA* and *luxB* genes, expression from co-transfected plasmids harboring the *luxA* and *luxB* genes independently and

expression from a transcriptionally fused *luxA* and *luxB* that are independently translated via an IRES element.

- To establish if episomal expression of the *lux* genes provides a higher bioluminescent signal than constructs integrated into the host's chromosome.
- To determine the stability of an episomal plasmid in mammalian cell lines without selective pressure.
- To evaluate if $FMNH_2$ is a limiting substrate for efficient bioluminescence in mammalian cells.

Materials and Methods

Cell Culture and Plasmid Maintenance

All relevant constructs and strains, bacterial and mammalian, used in this study are outlined in Table 3. *E. coli* cells were routinely grown in Luria Bertani (LB) (Fisher Scientific, Pittsburgh, PA) broth containing the appropriate antibiotic selection with continuous shaking (200rpm) at 37°C. Kanamycin and Ampicillin were used at a final concentration of 50μg/ml and 100 μg/ml, respectively.

All cell culture reagents and media were obtained from Sigma Aldrich, (St. Louis, MO) unless otherwise stated. Mammalian cells were grown in the appropriate complete growth media containing 10% heat-inactivated horse or fetal bovine serum, 0.01mM non-essential amino acids and 0.1mM sodium pyruvate in a Dubelco's

Table 3. Strains and plasmids used in the determination of the optimal expression format for bacterial luciferase in mammalian cells.

Plasmid/Strain Designation	Relevant Genotype/ Characteristics	Source
Strains		
E. coli		
DH5α	Φ80d*lacZ*ΔM15, *recA*1, *endA*1, *gyrA*96, *thi*-1, *hsdR*17 (r_K-, m_K+), *supE*44, *relA*1, *deoR*, Δ(*lacZYA-argF*)U169	Gibco, BRL
TOP 10	F-, mcrA Δ(mrr-hsdRMS-mcrBC) Φ80/lacZ ΔlacX74 deoR recA1 araD139 Δ(ara-leu) 7697 ga/K rpsL endA1 nupG	Invitrogen
Mammalian Cell Lines		
HEK293	Permanent line of primary human embryonal kidney transformed by sheared human adenovirus type 5 (Ad 5) DNA. ATCC# CRL-1573	ATCC
COS-7	Monkey Kidney cells transformed with an origin defective mutant of SV40 which codes for wild-type T antigen. ATCC# CRL-1651	ATCC
HeLa	Human cervical cancer cell line ATCC# CCL-2	ATCC
Plasmids		
pCR2.1-TOPO	TOPO TA cloning vector for easy cloning of PCR products generated with 3′ A overhangs Km^r, Amp^r	Invitrogen
pCR4-TOPO	TOPO TA cloning vector for easy cloning of PCR products generated with 3′ A overhangs designed for sequencing Km^r, Amp^r	Invitrogen
pcDNA3.1	Mammalian expression vector, constitutive CMV promoter, contains a Neomycin G418 antibiotic selection and a ColEI and Amp^r for replication in *E. coli*	Invitrogen
pCEP4	Mammalian episomally maintained expression vector, constitutive CMV promoter, Hygromycin antibiotic selection marker and a ColEI and Amp^r for replication in *E. coli*	Invitrogen
pREP9	Mammalian episomally maintained expression vector, constitutive RSV promoter, Neomycin G418 antibiotic selection marker and a ColEI and Amp^r for replication in *E. coli*	Invitrogen

Table 3. Continued

Plasmid/Strain Designation	Relevant Genotype/ Characteristics	Source
pIRES2-DsRed2	Mammalian expression vector containing the internal ribosomal entry site (IRES) of the encephalomyocarditis virus between the multi-cloning site and a DsRed reporter protein, a constitutive CMV promoter, Neomycin G418 antibiotic selection marker and a pUC ori and Km^r for replication in *E.coli*	Clontech
pIRES	Mammalian expression vector containing the internal ribosomal entry site (IRES) of the encephalomyocarditis virus between two multi-cloning sites which allows for the expression of two genes under the control of a single constitutive CMV promoter, Neomycin G418 antibiotic selection marker and a pUC ori and Km^r for replication in *E.coli*	Clontech
pCR4PLluxCDABE	pCR4 harboring a 6.1 kb *luxCDABE* cassette from *Photorhabdus luminescens*	This Study
pCR4luxA	pCR4 TA cloning vector harboring the *luxA* from *Photorhabdus luminescens*	This Study
pCR4luxB	pCR4 TA cloning vector harboring the *luxB* from *Photorhabdus luminescens*	This Study
pCR2luxA	pCR2.1 TA cloning vector harboring the *luxA* from *Photorhabdus luminescens*	This Study
pCR2luxB	pCR2.1 TA cloning vector harboring the *luxB* from *Photorhabdus luminescens*	This Study
pCRluxAf	pCR2.1 TA cloning vector harboring the *luxA* from *Photorhabdus luminescens* amplified with the reverse fusion primer	This Study
pCRluxBf	pCR2.1 TA cloning vector harboring the *luxB* from *Photorhabdus luminescens* amplified with the *luxB* forward fusion primer	This Study
pCRluxABf	pCR2.1 TA cloning vector harboring the *luxAB* generated by ligating the *luxAf* and *luxBf* together	This Study
pcDNABf	pcDNA3.1 harboring the *luxAB*f	This Study

Table 3. Continued

Plasmid/Strain Designation	Relevant Genotype/ Characteristics	Source
pcDNABfKoz	pcDNA3.1 harboring the *luxAB*f with the Kozac sequence modification	This Study
pcDNluxA	pcDNA3.1 harboring the *luxA* from *P. luminescens*	This Study
pCEPluxA	pCEP4 harboring *luxA* from *P.luminescens*	This Study
pREPluxB	pREP9 harboring *luxB* from *P.luminescens*	This Study
pCRSVluxBpA	pCR4 harboring the RSV promoter, *luxB* and the SV40 pA with introduced *ClaI* and *BglII* sites on both the 5′ and 3′ ends of the gene	This Study
pREPABf	pREP9 harboring *luxAB* fusion from *P.luminescens*	This Study
pCEPluxARluxB	pCEP4 harboring the *luxA* cloned into the MCS and the RSV-*luxB*-SV40pA into a unique *ClaI* restriction site within the vector	This Study
pcDNARB	pcDNA3.1 harboring *luxA* cloned into the MCS and the RSV-*luxB*-SV40pA into a unique *BglII* restriction site within the vector	This Study
pluxAIEGFP	pIRES-EGFP harboring the *luxA* from *P. luminescens*	This Study
pluxBIDsRed	pIRES-DsRed harboring the *luxB* from *P. luminescens*	This Study
pCR4NotIluxA	pCR4 harboring the *luxA* from *P. luminescens* with introduced NotI sites on both the 5′ and 3′ ends of the gene	This Study
pluxAIRES3	pIRES harboring the *luxA* from *P. luminescens* cloned into the MCS(A)	This Study
pluxAIRESluxB	pIRES harboring the *luxA* cloned into the MCS(A) and *luxB* cloned into MCS(B) from *P. luminescens*	This Study

minimal essential media base (DMEM) (M4655). Cells were routinely grown at 37°C in a 5% CO_2 atmosphere to confluency and split every three to four days by trypsinization at a 1:4 ratio into fresh complete growth media. Appropriate concentrations of antibiotic were used to maintain constructs after transfection according to susceptibility kill curve analysis. Kill curves were completed for each cell line and lot of antibiotic. A range of typical concentrations used for each antibiotic is found in Table 4.

Antibiotic Kill Curves

Kill curve experiments were performed to determine the antibiotic susceptibility for each cell line to each lot of antibiotic. Cells were plated into six well tissue culture plates and grown to 50-60% confluency. Varying concentrations of antibiotic were mixed into each well along with one control well (no antibiotic). The plates were incubated at 37°C in a 5% CO_2 atmosphere for seven to ten days with the media being refreshed every three days. Cells were checked daily by microscopic analysis and changes in cell morphology and viability were recorded. The minimum concentration of antibiotic that was toxic to the cells within eight days was used for selection of stable cell lines (Table 4).

Construction of a *luxA- luxB* Fusion Protein

To create a monocistronic version of the heterodimeric luciferase protein encoded by the *luxA* and *luxB* genes, a translational fusion of the two polypeptides was

Table 4. Concentrations of antibiotics used for each cell line based on results of kill curves.

Antibiotic	HEK293	COS-7	HeLa
Neomycin G418	450-650μg/ml	400-700μg/ml	500-750μg/ml
Zeocin	250-400μg/ml	Not Tested	Not Tested
Hygromycin	400-600μg/ml	Not Tested	Not Tested

constructed. Synthetic oligonucleotides were designed to fuse the 3′ end of the *luxA* to the 5′ end of the *luxB*. Both oligonucleotide sequences were complementary to sequences at the 3′ end of the *luxA*, the intergenic region between the *luxA* and *luxB*, and the start codon of *luxB*. These oligos were synthesized with the following modifications: ***luxA Reverse Primer:*** the *luxA* stop codon (TAG) was removed by substituting a Cytosine nucleotide for the Guanine resulting in a tyrosine codon, a single base addition of Guanine was also added to place the *luxB* in the same reading frame to create a fusion protein, further, a Guanine was substituted to generate an *AvrII* restriction site at the 3′ end of the *luxA*. ***luxB Forward Primer:*** the primer was the exact complement of the *luxA* reverse primer. The *luxA* and *luxB* genes were individually amplified using *Taq* polymerase (to generate 3′ A overhangs) and TA cloned into the pCR2.1 TOPO cloning vector (Invitrogen Corporation, Carlsbad, CA) to construct pCRluxAf and pCRluxBf. These plasmids were then digested with *EcoRI* and *AvrII* and the products were gel purified by electroelution (Sambrook et al., 1999). The fragments were then ligated in equal molar concentrations for 2 h. The ligation product (1 μl) was then used as template for a PCR reaction using the outermost *luxA* forward and *luxB* reverse primer pair to generate a *luxAB* fusion. The resultant PCR products of the correct size were TA TOPO cloned into the pCR2.1 TOPO cloning vector to generate pCRluxABf. The construct was then digested with *AvrII* and *EcoRI* restriction enzymes and sequenced to ensure its integrity. The sequence of the wild type and modified intergenic region of the fusion are shown in Figure 5.

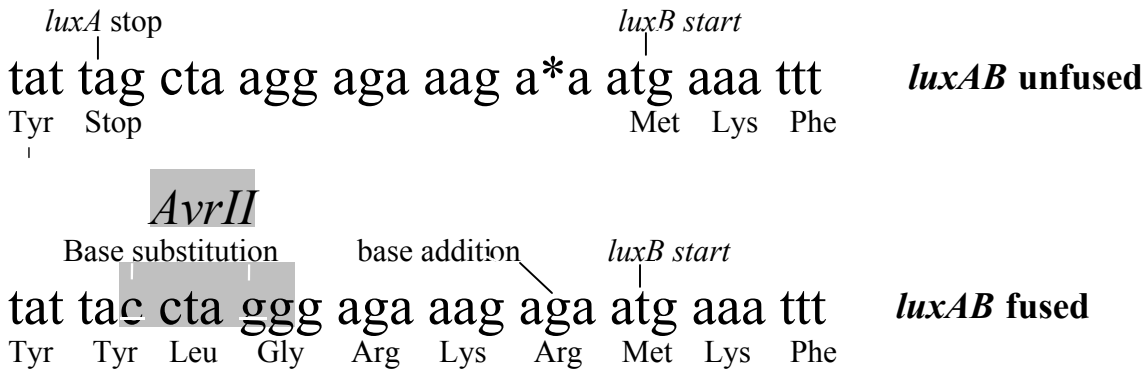


Figure 5. Sequence of the wild type *luxAB* and *luxABf*. Introduced *AvrII* site is in the shaded area. Base substitutions and base additions are noted and * represents the absence of the base in the wild type *luxAB*.

Episomal Expression

For the expression fusion protein as an episome, the *luxABf* was cloned into the pREP9 mammalian expression vector (Clontech Corporation, Palo Alto, CA). To create this clone, the *luxABf* insert was cleaved from pCRABf with *KpnI* and *XhoI* restriction sites located within the vector and then ligated into the pREP9 vector that had been digested with the same enzymes. The ligation of these two fragments generated pREPABf (Figure 6A).

Chromosomal Expression

The *luxAB* fusion was cleaved from pCRABf via 3′ and 5′ *EcoRI* sites and non-directionally cloned into pcDNA3.1 to generate pcDNABf (Figure 6B) for chromosomal insertion and expression in mammalian cells.

To facilitate high levels of expression in eukaryotes, Kozac sequences are often inserted at the 5′ ends of genes. To accomplish this addition to the *luxAB* fusion a *luxA* forward primer was modified to insert a Kozac sequence by substituting an Adenine at the –3 position and a Guanine at the +4 position around the start site of the *luxA*. The *luxAB* fusion was then amplified from the pCRABf plasmid construct and the resultant PCR product was subsequently TA cloned into pCR2.1 to generate pCRABfKoz. This insert was then cleaved by *EcoRI* and cloned into the pcDNA3.1 mammalian expression vector to create pcDNABfKoz.

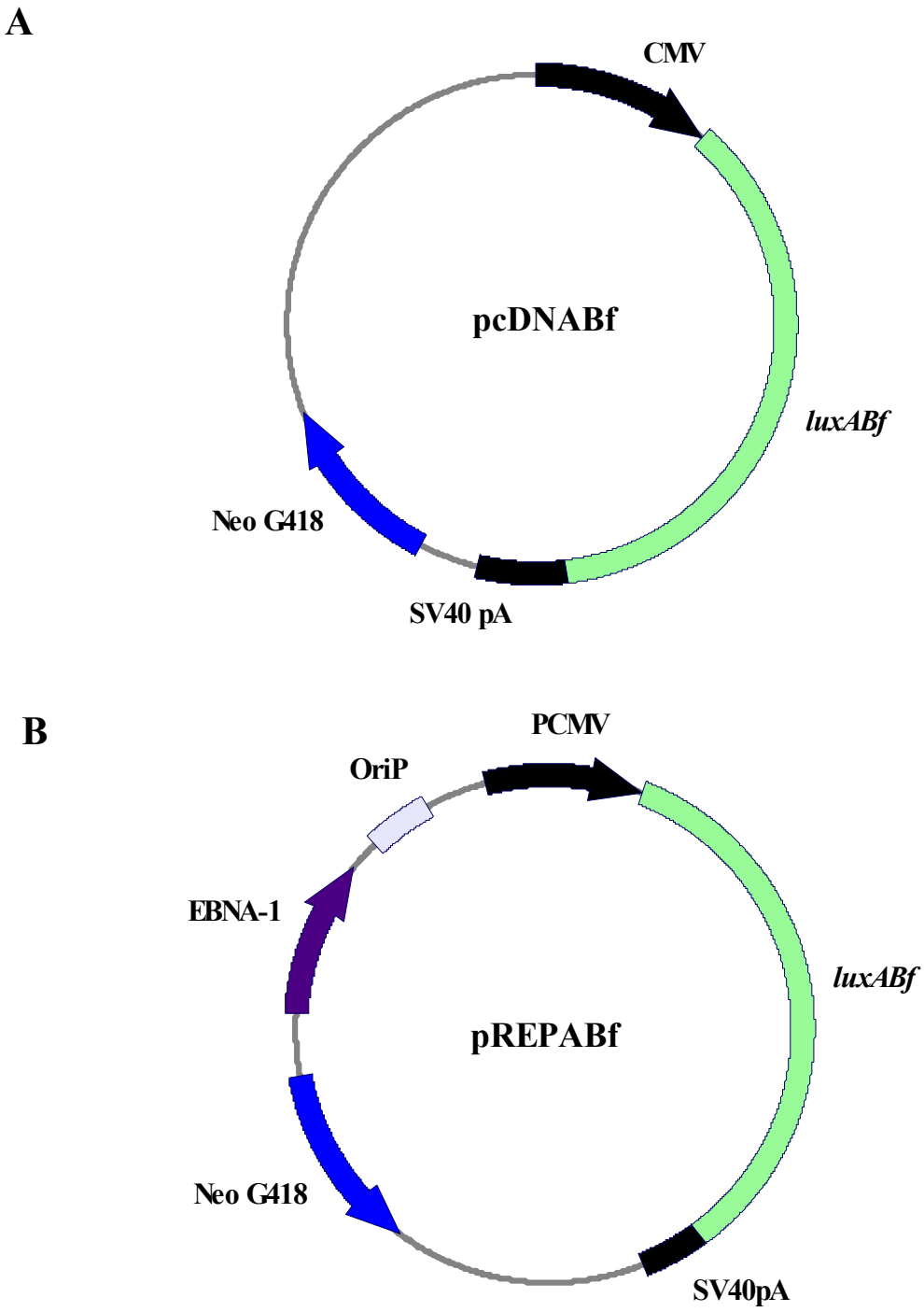


Figure 6. Diagram of vector plasmids for the expression of the *luxAB* fusion in mammalian cells. pcDNABf construct provides chromosomal expression and neomycin G418 selection. B. pREPABf construct allows for episomal expression and neomycin G418 selection.

Construction of a Dual Promoter Vector for Expression of *luxA* and *luxB*

Episomal Expression

Pasting pieces from pCEP4 and pREP9 episomal mammalian expression vectors together generated a dual promoter vector that allowed for the expression of both genes (*luxA* and *luxB*) from their own promoters. First, *luxA* was cleaved from pCRluxA with introduced 5′ *NheI* and a 3′ *XhoI* sites and cloned into pCEP4 to generate pCEPluxA. The *luxB* gene was then cloned into pREP9 to generate pREPluxB using the same clone strategy described above for cloning *luxA* into pCEP4. Oligonucleotide primers were designed and synthesized to amplify the RSV promoter region of pREP9, *luxB* gene, and the SV40 ployA region from pREPluxB with the introduction of 5′ and 3′ *ClaI-BglII* restriction sites. The resultant PCR product was then TA TOPO cloned into pCR4 TOPO (Invitrogen Corporation, Carlsbad, CA) to generate pCRSVluxBpA. This construct was then cleaved via *ClaI* and the insert cloned into a complementary site on pCEPluxA to generate pCEPluxARluxB (Figure 7A).

Chromosomal Expression

To chromosomally express both the *luxA* and *luxB* from individual promoters with only one selection marker required, a dual promoter vector was constructed. This expression vector was generated on the plasmid backbone of the pcDNA3.1 mammalian expression vector (Invitrogen Corporation, Carlsbad, CA). First, the *luxA* gene was cleaved from pCRluxA with *EcoRI* and nondirectionally ligated into pcDNA3.1. The clones were then checked for insert presence and orientation by restriction digestion and sequencing to generate pcDNluxA. Once verified the RSV, *luxB* and SV40 polyA

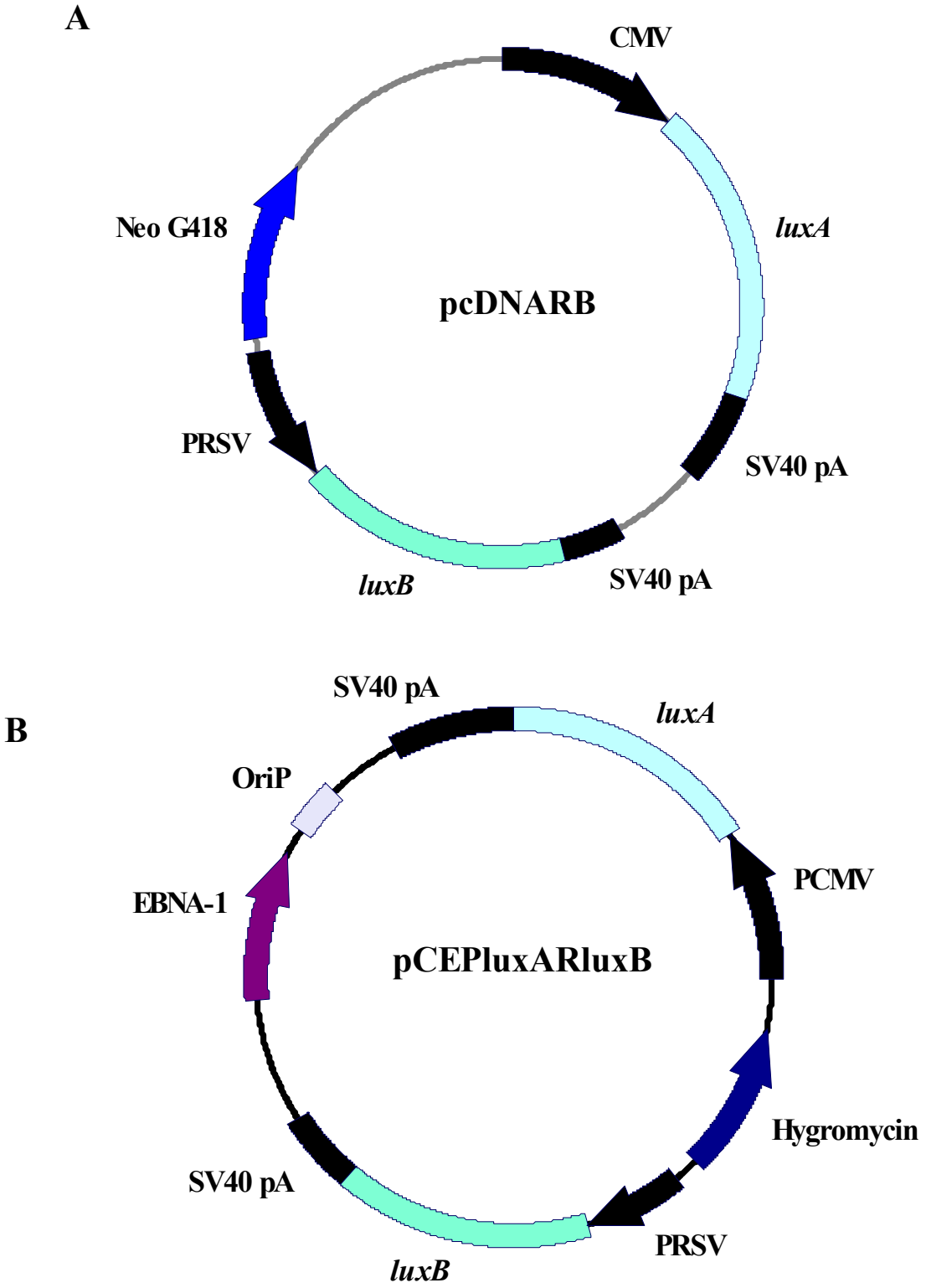


Figure 7. Diagram of vector plasmids for the expression of *luxA* and *luxB* from individual promoters within the same vector in mammalian cells. A. pcDNARB construct allows for chromosomal expression and hygromycin selection. B. pCEPluxARluxB construct provides episomal expression and neomycin G418 selection.

region were cleaved from pCRSVluxBpA via a unique 5′ and 3′ *BglII* sites and cloned into pcDNluxA to generate pcDNARB (Figure 7B).

Construction of GFP vectors for co-transfection Experiments

Episomal Expression

Co-transfection of each gene (*luxA* and *luxB*) on separate plasmids would allow the proteins to be expressed independently. For episomal expression, the pREP9 and pCEP4 (both from Clontech Corporation, Palo Alto, CA) expression vectors were used. The *luxA* and *luxB* genes were cloned into these vectors as previously described to generate pCEPluxA (Figure 8A) and pREPluxB (Figure 8B).

Chromosomal Expression

Separate EGFP and DsRed reporter vectors were constructed to allow for co-transfection of *luxA* and *luxB* genes on separate plasmids and integration within the host's chromosome. *luxA* was cloned into pIRES-EGFP and *luxB* was cloned into pIRES2-DsRed2 (both from Clontech Corporation, Palo Alto, CA). These reporter vectors were chosen to allow for the monitoring of co-transfection efficiency and plasmid maintenance. Each plasmid contains a GFP variant reporter gene under the translational control of an encephalomyocarditis virus (EMCV) internal ribosomal entry site (IRES) that immediately follows the multi-cloning site within the vector. To generate these constructs, each vector was digested with *EcoRI* and the *luxA* or *luxB* was cleaved from pCRluxA or pCRluxB using the same enzyme. The vector and inserts were then ligated via the complementary ends to generate pluxAIEGFP (Figure 9A) and pluxBIDsRed (Figure 9B).

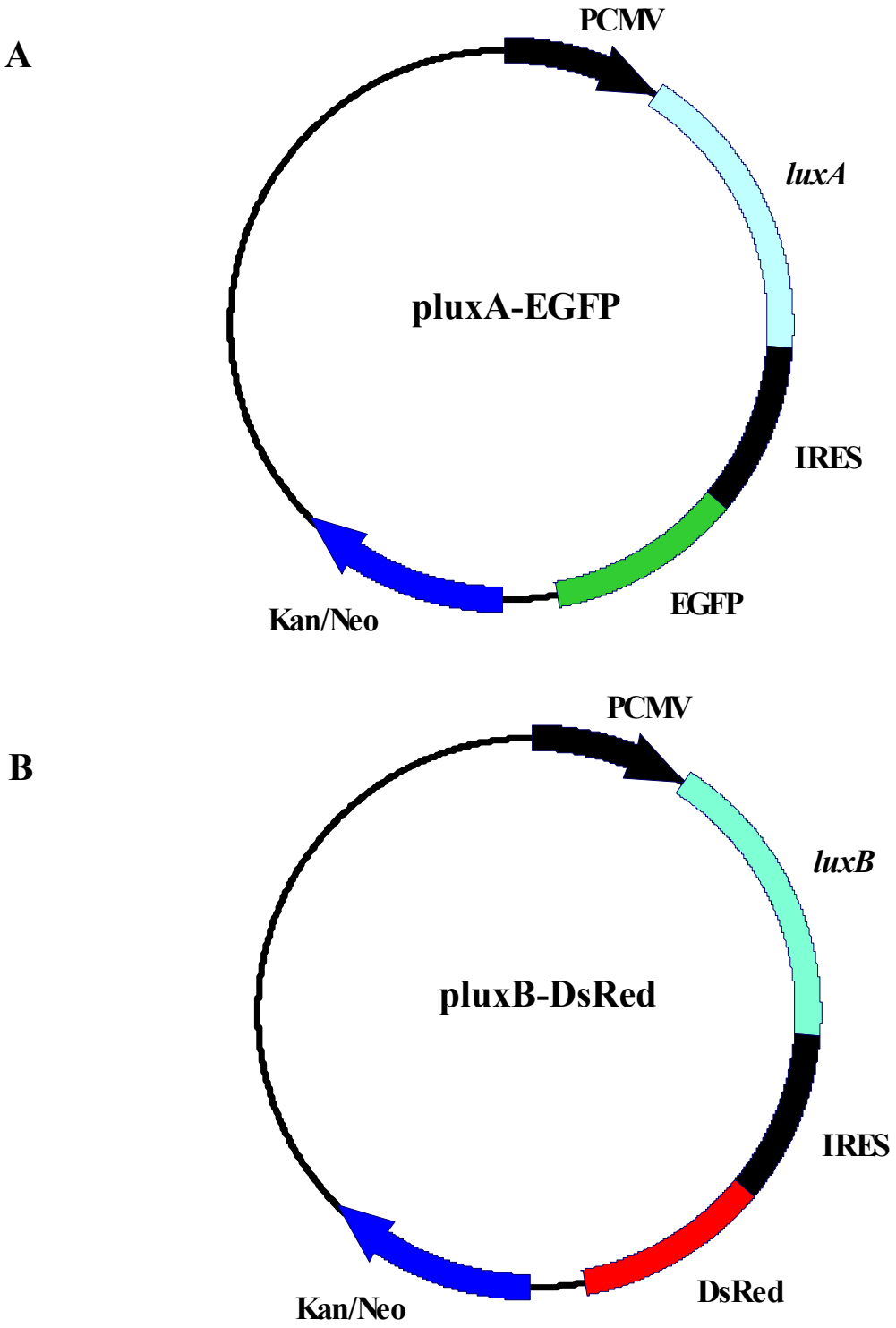


Figure 8. Diagram of vector plasmids for the expression of *luxA* and *luxB* from individual plasmids along with a GFP reporter protein in mammalian cells. A. pluxA-EGFP expresses the *luxA* gene and an EGFP reporter protein from a single bi-cistronic transcript. B. pluxB-DsRed construct expresses the *luxB* gene and a DsRed reporter protein from a single bi-cistronic transcript.

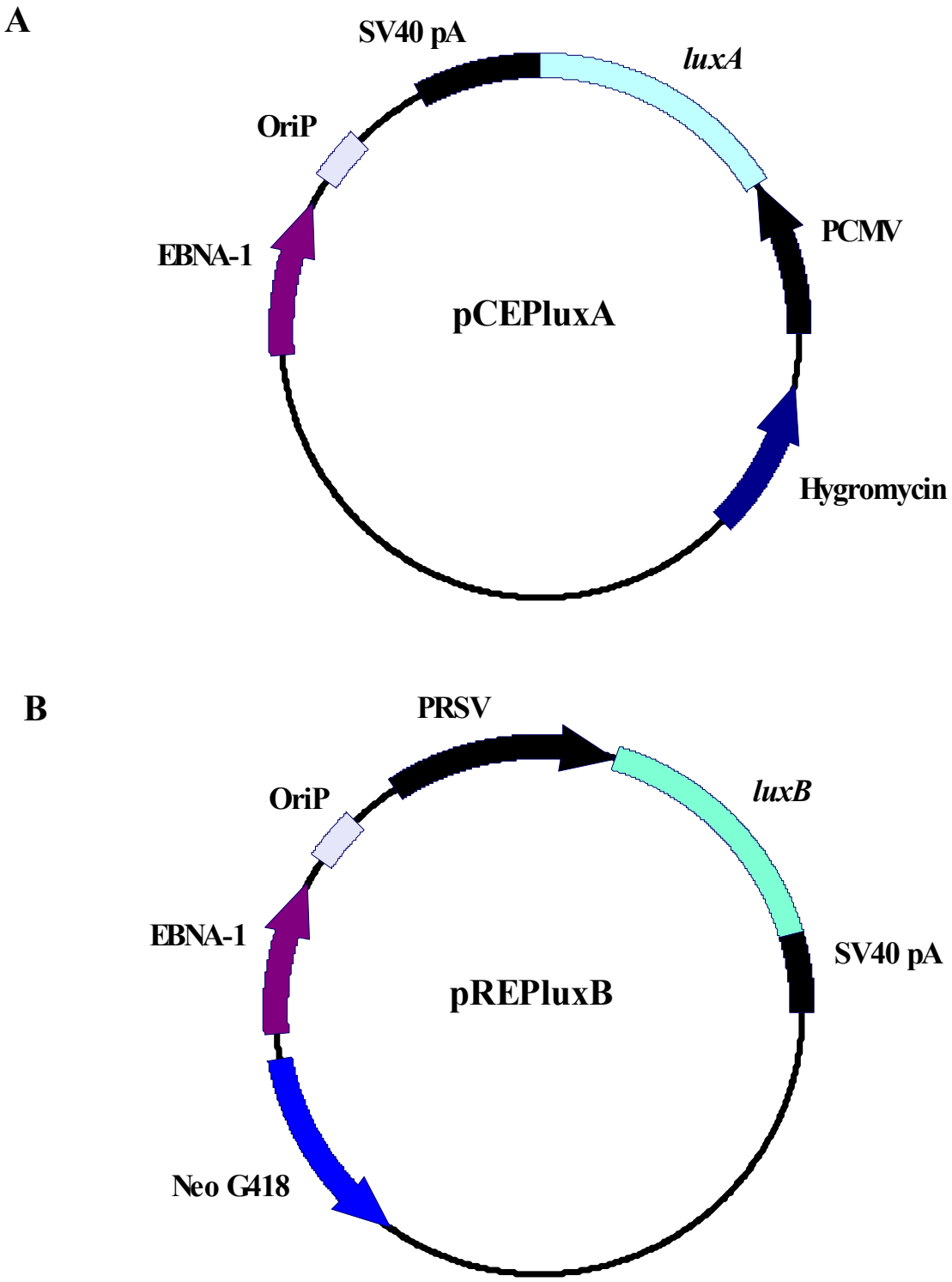


Figure 9. Diagram of vector plasmids for the expression of *luxA* and *luxB* from individual plasmids in mammalian cells. A. pCEPluxA allows for expression of the *luxA* gene as an episome with hygromycin selection. B. pREPluxB construct allows fro the expression of the *luxB* gene as an episome with neomycin G418 selection.

Construction of a Bi-cistronic Vector to Express *luxA* and *luxB*

To co-express both *luxA* and *luxB* genes from a single promoter, the pIRES mammalian expression vector was chosen (Clontech Corporation, Palo Alto, CA). This expression vector contains two multi-cloning sites separated by an internal ribosomal entry site (IRES) from encephalomyocarditis virus (EMCV). The IRES element allows for the expression of two genes (one cloned into each multi-cloning site) from a single constitutive CMV promoter. To create this construct, the *luxA* gene from *P. luminescens* was amplified from pCR4PLluxCDABE plasmid that harbors the complete *luxCDABE* cassette (Table 3) with the introduction of unique *NotI* restriction sites on both the 5′ and 3′ ends of the *luxA* gene. The resultant PCR product was TA TOPO cloned into pCR4 TOPO to generate pCR4NotIluxA. The *luxA* gene was then cloned into the MCS(A) of pIRES via the unique *NotI* restriction sites to generate pluxAIRES3. Once this construct was confirmed by sequencing, the plasmid was purified using the Wizard midi-prep plasmid purification kit according to the manufacturer's instructions (Promega Corporation, Madison, WI). The *luxB* gene was cleaved via a 5′ *XbaI* and 3′ *SpeI* site from pCRluxB and cloned into the *XbaI* site within the MCS(B) of pluxAIRES3 to generate pluxAIRESB (Figure 10).

Ligation Reactions

Plasmid vectors and inserts were digested (2-6 h) with the appropriate enzymes (Promega Corporation, Madison, WI). Linearized vectors were dephosphorylated using a calf intestine alkaline phosphatase enzyme according to the manufacturer's

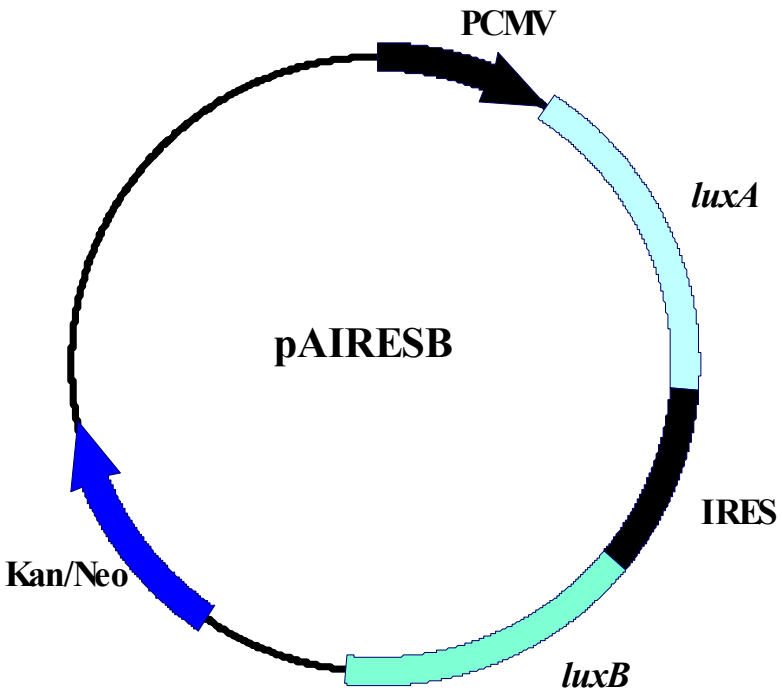


Figure 10. Diagram of vector construct for the expression of *luxA* and *luxB* as a single bi-cistronic transcript, pAIRESB.

instructions (Promega Corporation, Madison, WI). Both vector and insert DNA were gel purified from 1% agarose gels using the Geneclean gel extraction kit (Bio101, Carlsbad, CA). The recovered DNA was then quantified using a Dyna Quant 200 fluorometer (Hoefer Pharmacia Biotech Incorporated, San Francisco, CA) and ligations were set up as 20µl reactions using a 3:1 molar ratio of insert to vector DNA. The ligation reactions were then incubated at 17°C overnight.

Electroporation

Electrocompetent cells were prepared as outlined by the manufacturer (BTX, San Diego, CA). Electroporations were performed using the BTX Electroporator 600 with the following conditions: 40µl cells, 1-2µl ligation mixture (above), a 2.5kV pulse for 4.7ms using a 2mm gap cuvette. After the pulse, cells were immediately resuspended in 1ml of sterile LB and allowed to recover for 1 h at 37°C (200 rpm). Cells were then plated on selective media containing the appropriate antibiotic.

Selection of Bacterial Clones

Resistant colonies were picked after 24 h and expanded to patches on grid plates. To test for proper insert presence and orientation, rapid boil plasmid mini-preps (Promega Corporation, Madison, WI) were done followed by the digestion of the plasmid with the appropriate restriction enzyme mixture according to the manufacturer's instructions (Promega Corporation, Madison, WI). Products were run on 1% agarose gels to determine if the banding pattern indicated the insert presence and

proper orientation. Upon identifying correct clones, the plasmids were further purified using the Wizard midiprep plasmid purification system according to the manufacture's protocol (Promega Corporation, Madison, WI) and sequenced.

Sequencing

All constructs were sequenced to ensure their integrity. Sequencing was done in the University of Tennessee Molecular Biology Service Facility using an Applied Biosystems 3100 Genetic Analyzer sequencer (Foster City, CA).

Determination of Thermostability of the *luxAB* Fusion Protein

To determine the thermostability of the *luxAB* fusion protein, *E. coli* cells harboring the pCRABf construct were grown at 23°C, 30°C and 37°C overnight with 50 μg Kanomycin/ml in LB. The bioluminescence levels for each temperature condition were taken in triplicate. Bioluminescence measurements were done using the FB14 luminometer (Zylux Corporation, Pforzheim, Germany) at a 1 s integration and reported as relative light units (RLU). To normalize the data each bioluminescence reading was divided by absorbance $O.D._{600}$ for the culture and reported as relative bioluminescence.

Transfection of Mammalian Cells

Transfection of all cell lines was done in six well poly-D-lysine coated tissue culture plates (Fisher Scientific, Pittsburgh, PA). Cells were split from stock cultures and inoculated into each well at approximately $1X10^5$ cells per well in complete growth

media. The plate was then placed at 37°C in a 5% CO_2 atmosphere for 1-2 days until the cells became 80-90% confluent. The day of transfection, the medium was refreshed. DNA for transfections was purified from 100ml overnight *E. coli* cultures using the Wizard Purefection plasmid purification kit to remove endotoxins according to the manufacturer's instructions (Promega Corporation, Madison, WI). For chromosomal integration, the plasmid DNA was linearized before transfection to increase proper integration. For episomal expression, plasmids were transfected as circular DNA.

HEK293 Cells

Purified plasmid DNA (3.2 μg) was mixed into 200 μl of serum free DMEM in a 1.5 ml tube. In a second tube, 8 μl of Lipofectamine 2000 reagent (Invitrogen Corporation, Carlsbad, CA) was added to 200 μl of serum free DMEM. The lipofectamine mixture was added to the DNA mixture within 5 min and incubated at room temperature for 20 min. The entire mixture (400 μl total) was added directly to the appropriate well on the plate and rocked back and forth to ensure adequate mixing. Twenty-four hours post transfection, the complexes were removed and the media was replaced with fresh complete growth media supplemented with the appropriate antibiotic for selection.

COS-7 Cells

Purified plasmid (1.5 μg) was mixed with 100 μl of serum free DMEM in a 1.5 ml tube. In a second tube, 5 μl of lipofectin (Invitrogen Corporation, Carlsbad, CA) was mixed with 100 μl of serum free DMEM. The two mixtures were then mixed together and incubated at room temperature for 45 min. After incubation, 0.8 ml of

serum free DMEM was added to the mixture and then directly placed on the cells in the plate that had been previously rinsed two times with serum free DMEM to remove any residual serum. The complexes were incubated on the cells for 5 h at 37°C in a 5% CO_2 atmosphere. After incubation DMEM supplemented with 20% fetal calf serum was added. Forty-eight hours after transfection, complete DMEM plus antibiotic was applied to the cells for selection.

HeLa Cells

Purified plasmid DNA (1.6 μg) was added to 200 μl of serum free DMEM along with 10 μl of the PLUS Reagent (Invitrogen Corporation, Carlsbad, CA) and incubated at room temperature for 15 min. In a second tube 1 μl of Lipofectin (Invitrogen Corporation, Carlsbad, CA) reagent was added to 50 μl of serum free DMEM and incubated at room temperature for 30 min. The two tubes were then mixed gently and incubated further for 15 min. The growth media was removed from the cells and replaced with serum free DMEM and the DNA-Lipofectin complexes were added directly to the wells and gently mixed. The plates were then incubated at 37°C in a 5% CO_2 atmosphere for 3 to 5 h. After incubation, 15% fetal bovine serum was added to each well. Twenty-four hours post transfection, growth media supplemented with the appropriate antibiotic was applied to the cells for selection.

Selection of Mammalian Cell Clones

Twenty-four to forty-eight hours post transfection, selective medium was added to all wells and refreshed every three to four days. Within two weeks all control cells

were dead and the transected cells were forming small colonies on the plate surface. Colonies were separated from the rest of the well by placing a sterile chamber around the cell mass and sealing it with silicon (Fisher Scientific, Pittsburgh, PA). The media could then be removed and each colony could be trypsinized and transferred to individual tissue culture flasks. To accomplish this, after washing with a PBS solution, 200 µl of a 1X Trypsin-EDTA solution (Sigma Aldrich, St. Louis, MO) was added directly to the chamber and incubated at 37°C for 3 to 5 min. The trypsin-EDTA solution was then replaced with complete growth media and the cells were transferred to a 25cm^2 tissue culture flask for propagation. Each clone was given a number and expanded to individual cell lines. Each line was split and maintained as described earlier with the addition of selective media. At between nine and twenty cell lines were propagated in this manner for each plasmid tested.

Bioluminescence Assays from Mammalian Cells

To determine bioluminescence potential from each cell line clone, total proteins were extracted and *in vitro* enzyme (bioluminescence) assays performed. To extract the proteins, the cells were trypsinized from the plate or flask surface using standard protocols and resuspended into 2.0 ml Sarstedt tubes (Fisher Scientific, Pittsburgh, PA). The cells were then spun down and washed two times in sterile phosphate buffered saline (PBS) to remove any residual medium (Sigma Aldrich, St. Louis, MO). Cell pellets were then resuspended into 1 ml 0.1M potassium phosphate buffer pH 7.8 and disrupted by three consecutive cycles of freeze (30 s liquid N_2) thaw (5 min at 37°C) extraction. After disruption, the cell debris was pelleted by spinning the samples at

14,000Xg for 5 min and the supernatant was used in the bioluminescence assay. To determine light intensity, the protein extract was mixed with 0.1 mM NAD(P)H, 4 μM FMN, 0.2% (w/v) BSA, 0.002% (w/v) n-decanal. Bioluminescence was measured using the FB14 luminometer (Zylux Corporation, Pforzheim, Germany) at a 1 s integration and reported as relative light units (RLU). To determine if $FMNH_2$ was a limiting factor for the bioluminescence reaction, a flavin oxidoreductase enzyme (1U) isolated from *V. harveyi* (Roche Scientific, Indianapolis, IN) was added to the bioluminescence assay and the light levels were measured again for comparison (Table 5).

Bioluminescence signals were normalized between samples and cell lines by dividing the RLU measurement by the total protein and reporting the bioluminescence as RLU/μg total protein. Protein concentrations were determined using the Coomassie Plus protein assay according to the manufacture's instructions (Biorad, Hercules, CA).

In Vitro Transcription/Translation

To determine if the *lux* genes could be generated *in vitro* in rabbit reticulocyte lysate (mammalian translation machinery), pCR2.1 TOPO vectors harboring *luxA*, *luxB* and *luxABf* were transcribed and translated. First, the plasmid DNA containing the genes was digested at a unique *SpeI* restriction site at the 3′ end of the gene within the vector. This digestion linearized the plasmid and allowed for the generation of run-off transcript from the vector derived T7 promoter. Each gene was transcribed via T7 polymerase using the RiboMax large-scale transcription system (Promega Corporation, Madison, WI). Three individual transcription reactions were set up along with

Table 5. Final constructs transfected into mammalian cell lines.

Plasmid	Expression	Cell Line Tested	Selection Marker
pcDNABf	Chromosomal	HEK293, HeLa, COS-7	G418
pcDNABfKoz	Chromosomal	HEK293, HeLa, COS-7	G418
pREPABf	Episomal	HEK293	G418
pCEPluxA	Episomal	HEK293	Hygromycin
pREPluxB	Episomal	HEK293	G418
pluxA-EGFP	Chromosomal	HEK293	G418, EGFP
pluxB-DsRed	Chromosomal	HEK293	G418, DsRed
pCEPluxARluxB	Episomal	HEK293	Hygromycin
pcDNARB	Chromosomal	HEK293	G418
pluxAIRESluxB	Chromosomal	HEK293	G418

a positive T7 control and a negative control containing no template DNA. Each reaction was set up according to the manufacturer's protocol and then incubated at 37°C for 1 h. Transcripts were quantified by absorbance (260/280) measurements (Beckman Coulter, Fullerton, CA). Ten micrograms per ml of total RNA transcript was then added to 50 µl (total volume) rabbit reticulocyte lysate translation reactions. Each reaction was gently mixed on ice according to the manufacturer's protocol for ^{35}S labeled protein generation and then incubated at 30°C for 90 min (Promega Corporation, Madison, WI). Once translation was complete, 15 µl of each reaction was loaded onto a 12% SDS-PAGE mini-gel and run at 30mA for 1 h. The gel was removed and dried at 60°C with vacuum pressure using a model 443 Slab Dryer (BioRad, Hercules, CA) onto 3MM filter paper (Fisher Scientific, Pittsburgh, PA). To visualize the generated proteins, the gel was placed onto x-ray film overnight and specific activity was measured upon film development.

Genomic DNA Isolation

Genomic DNA from each cell line clone was accomplished using the Wizard genomic DNA extraction kit according to the manufacture's protocols (Promega Corporation, Madison, WI).

RNA Isolation and Blotting

At passage six post transfection, selected cell line clones were expanded to 75cm^2 tissue culture flasks. When the cells became 80-95% confluent, the cells were

trypsinized to remove the cells from the surface and transferred to 2.0 ml Sarstedt tubes (Fisher Scientific, Pittsburgh, PA). Cells were spun down and washed two times in PBS (Sigma Aldrich, St. Louis, MO). Total RNA was then isolated from the cells using the RNeasy kit (Quiagen, Valencia, CA) according to the manufacturer's instructions for isolation of total RNA from mammalian cells. To remove any contaminating DNA, the RNA was digested for 30 min with DNaseI (Promega Corporation, Madison, WI). To remove the DNaseI enzyme, the clean-up procedure from the RNeasy kit was used (Quiagen, Valencia, CA). Total RNA was then quantified using the Beckman DU-640 spectrophotometer absorbance at 260/280 (Beckman Coulter, Fullerton, CA).

RNA Slot Blotting

Ten micrograms of total RNA were loaded onto a Biotrans^{TM} nylon membrane (ICN, Irvine, CA) using a Bioslot blot apparatus (Biorad, Hercules, CA) according to the manufacturer's protocol. A ^{32}P labeled probe was generated complementary to the 300 bp of the *luxA* or *luxB* gene from *P. luminescens* using standard PCR protocols with the incorporation of a [^{32}P] labeled dCTP nucleotide. The free nucleotides were removed and the probe purified by column purification according to the manufacture's instructions (Stratagene, La Jolla, CA). The specific activity of each probe was measured by scintillation counting (Beckman Coulter, Fullerton, CA). Before use, the dsDNA probe was boiled for 10 minutes to denature the DNA and directly added to the pre-hybridization solution (SLIME). The blot was incubated with the probe at 48°C overnight. After probe hybridization, the blot was washed 4 times in 20X SSC to remove any unbound activity. The wash temperatures were determined experimentally to achieve optimal probe binding without excess background activity. The blot was air

dried and then placed on a phosphorescence intensifier screen (Molecular Dynamics, Piscataway, NJ). Specific activity was measured using the STORM 840 phosphoanalyzer and the data analyzed using the ImageQuant data analysis software (Molecular Dynamics, Piscataway, NJ).

Protein Isolation and Western Blotting

To extract the proteins, cells were trypsinized from a plate or flask surface and resuspended into 2.0 ml Sarstedt tubes (Fisher Scientific, Pittsburgh, PA). The cells were then spun down and washed two times in sterile phosphate buffered saline (PBS) to remove any residual media (Sigma Aldrich, St. Louis, MO). Cell pellets were resuspended into 1 ml 0.1M potassium phosphate buffer pH 7.8 and disrupted by three consecutive cycles of freeze (30 s liquid N_2) thaw (5 min at 37°C) extraction. After disruption, the cell debris was pelleted by spinning the samples at 14,000Xg for 5 min and the supernatant was used as total soluble protein for Western blot analysis.

Protein concentrations were determined using the Coomassie Plus protein assay according to the manufacturer's instructions (Pierce, Rockford, IL). Equal amounts (100 – 250 μg) of protein were loaded onto a 12% SDS-PAGE gel. Minigels were run at 30 mA for approximately 2 h and larger slab gels were run at 30 mA overnight. The proteins were then electroblot transferred to a PDVF membrane (Biorad, Hercules, CA) using a semi-dry electroblotter according to the manufacture's instructions (CBS Scientific Company, Incorporated, Del Mar, CA). Blots were then blocked overnight in 5% nonfat dry milk and hybridized with a polyclonal antibody raised against a 16 amino acid *luxA* polypeptide (′N′ - FDDSDQTRGYDFNKGC - ′C′) or a 16 amino acid *luxB*

polypeptide (′N′ - CMILVNYNEDSNKAKQ - ′C′) (Genemed Synthesis, Incorporated, San Francisco, CA). Antibodies were diluted in T-TBS (Tris Buffered Saline + 3% Tween 20) at a 1:500 dilution and applied to the membrane at room temperature for 5 h to overnight. The blot was then washed several times in T-TBS and incubated with a Goat Anti-Rabbit second antibody that has been conjugated to alkaline phosphatase. The blot was then developed according to the manufacture's protocol (Biorad, Hercules, CA).

Statistics

Statistical analysis of the data presented here was conducted using either the JMP (SAS Institute, Incorporated, Pacific Grove, CA) or Microsoft Excel (Microsoft, Seattle, WA) statistical software packages. Graphs were made using Sigma Plot software (SPSS, SAS Institute, Incorporated, Pacific Grove, CA) or Microsoft Excel (Microsoft, Seattle, WA). All error bars on graphs indicate one standard deviation of the mean from triplicate samples. Significant differences were determined using either t-test or 1 way ANOVA analysis at a level of α=0.05.

Results

LuxAB Fusion Protein

Creation of the LuxAB Fusion

A LuxAB fusion protein from *P. luminescens* was generated by the elimination of the stop codon at the 3′ end of the *luxA* gene and the addition of one base within the intergenic region to place the two genes into the same reading frame. Although the

fusion protein was functional, light levels were significantly lower ($p = 0.05$) when expressed constitutively in *E. coli*. As shown in Figure 10, the LuxAB fusion protein had bioluminescence activity levels only approximately 30% of the wild type unfused LuxA and LuxB. The addition of a Kozak sequence further reduced the bioluminescence level to approximately 5% of the wild type protein (Figure 11).

Thermostability of the LuxAB Fusion

To determine if temperature had an effect on the folding and activity of the fusion protein, the fused and wild type versions of the *luxAB* constitutively expressed in *E. coli* were grown as 100ml liquid cultures overnight at 23˚C, 30˚C and 37˚C. Bioluminescence measurements were taken in triplicate and reported as specific bioluminescence (RLU/O.D. $_{600}$). Results are shown in Table 6A. In *E. coli*, there was no statistical difference ($p = 0.05$) between bioluminescence activities of cells grown at varied temperatures. To further evaluate this in a eukaryote, the *luxAB* fusion construct was cloned into a yeast expression vector (pYES-TOPO) and transformed into *S. cerevisiae.* Surprisingly, contradictory to the data from *E. coli,* light levels significantly decreased when the cells were grown at 37˚C (Table 6B).

***In Vitro* Expression**

To mimic mammalian cell translation machinery, the *luxA*, *luxB* and *luxAB* fusion genes were transcribed and translated *in vitro*. Although equal molar amounts of RNA transcript were added to each translation reaction and equal volumes of the

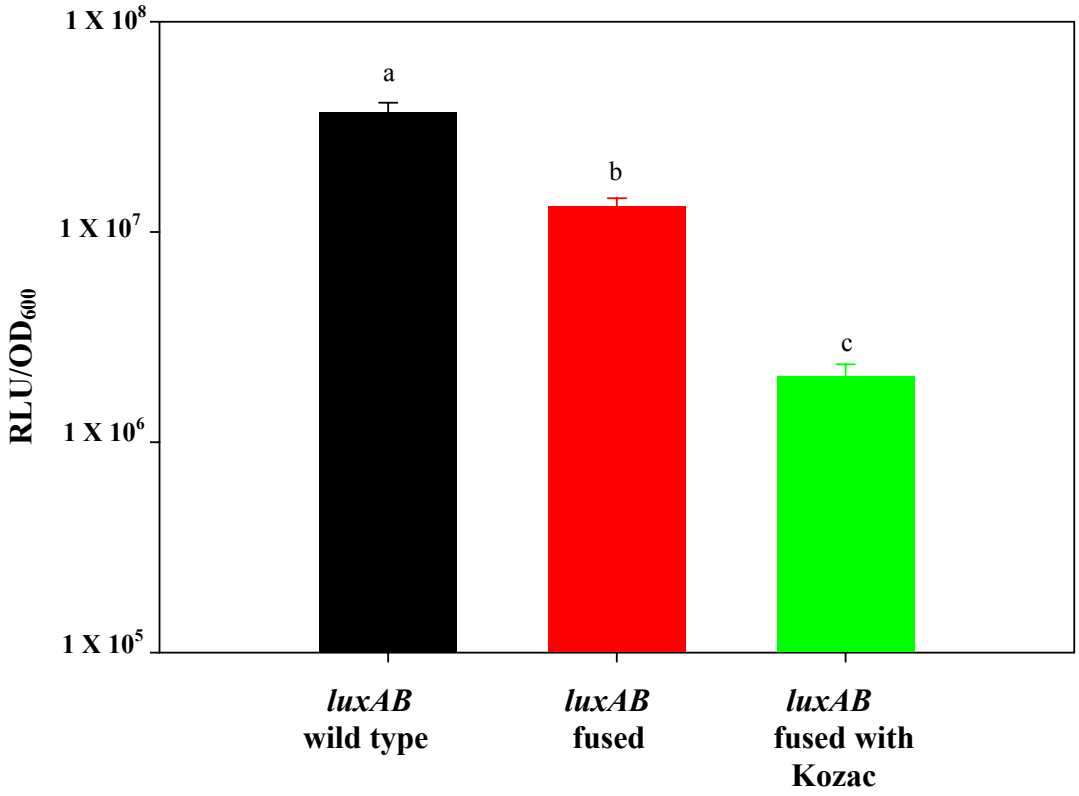


Figure 11. Bioluminescence values from *E. coli* cells harboring either wild type *luxA* and *luxB*, the *luxAB* fusion protein, or the *luxAB* fusion protein with the Kozac sequence modifications grown at 37°C. Bioluminescence values were taken in triplicate with the addition of 20 µl of a 1% n-decanal emulsion. (a, b and c signify significant differences at $p<0.05$)

Table 6. Effect of temperature on the LuxAB fusion protein's activity. A. Light emission from *E. coli* clones harboring the *luxAB* fusion genes when grown at 23°C, 30°C and 37°C. B. Specific bioluminescence of pYES2.1-TOPO with *luxAB* fusion in *S. cerevisiae* grown on galactose inducing media at 30˚C and 37˚C.

A

Temperature	23°C	30°C	37°C
Specific bioluminescence ± SD	$1.56*10^7$ $\pm 0.12*10^7$	$1.47*10^7$ $\pm 0.12*10^7$	$1.52*10^7$ $\pm 0.11*10^7$

B

Temperature °C	(+) Aldehyde - Specific bioluminescence (light/O.D.) ± SD
30°C	64,534 (±1,545)
37°C	16,223 (±1,018)

translation reaction were loaded onto the gel, the individual LuxA and LuxB proteins were not detected while the LuxAB fusion protein was easily detectable (approximately 78 kDa) (Figure 12).

Expression in Mammalian Cell Lines

Three mammalian cell lines were chosen to evaluate the expression of the LuxAB fusion protein. These cell lines included; HeLa, COS-7 and HEK293. To determine if the plasmid was present within each cell type after antibiotic selection of clones, PCR was performed on the genomic DNA from each clone with a *luxA* specific primer set. All cell line clones that resulted in a positive PCR product were further investigated for *luxA* message (mRNA), protein and bioluminescence activity. All clones tested had *luxA* mRNA levels higher than background vector controls, but the levels varied greatly between cell lines and individual clones (data not shown). The bioluminescence values obtained from cell extracts also varied between cell types and clones. The exogenous addition of an oxidoreductase enzyme to produce $FMNH_2$ increased the bioluminescence levels more than ten fold. These data suggest that $FMNH_2$ is a limiting substrate for the bioluminescence in mammalian cells. Figure 13 shows the average bioluminescence (RLU/mg total protein) from the brightest three clones from each cell type harboring pcDNABf within its chromosome. HEK293 cells consistently produced the highest bioluminescence levels, however these differences were not statistically significant ($p = 0.05$). Polyclonal antibodies to peptide epitopes within the LuxA and LuxB proteins were obtained (Genemed Synthesis, San Francisco, CA). Unfortunately, even though the bioluminescence levels were quantifiable the *lux*

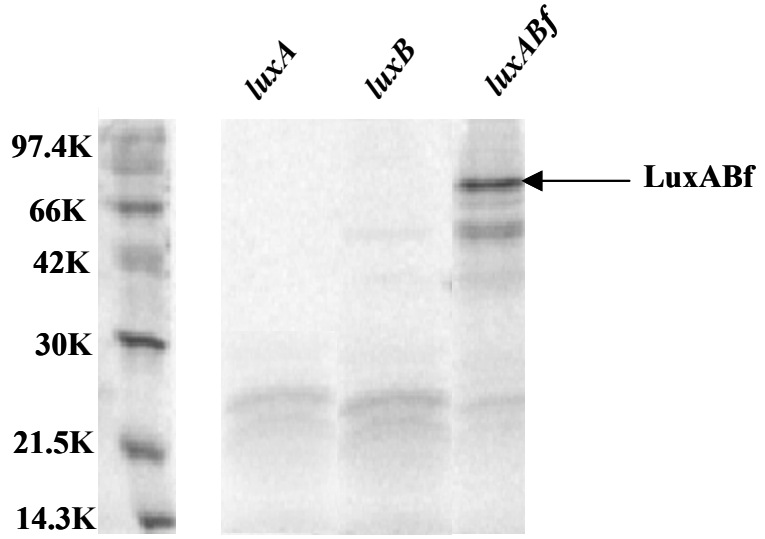


Figure 12. LuxA, LuxB and LuxAB fusion proteins translated *in vitro* in rabbit reticulocyte lysate with the incorporation of ^{35}S methionine. The molecular weight marker is labeled with [^{14}C] methylated protein.

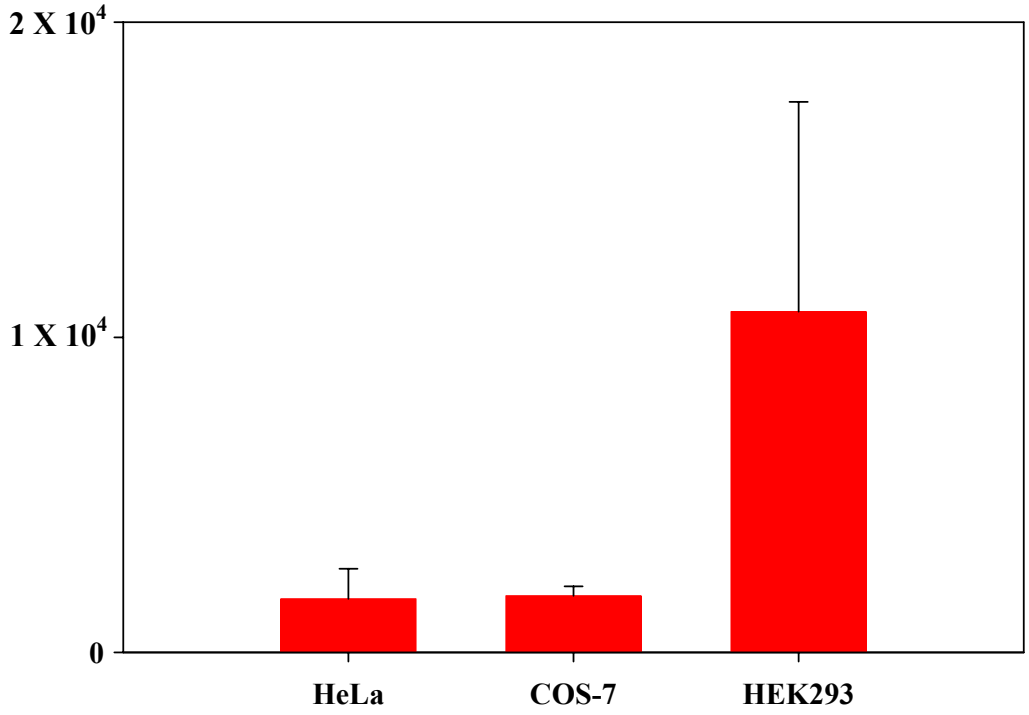


Figure 13. Relative bioluminescence (RLU/mg total protein) from stably integrated pcluxABf harboring the *luxAB* fusion gene. Bioluminescence was measured from cell extracts upon the addition of 0.002% n-decanal and $FMNH_2$. Bioluminescence reported as an average of the top three clones from each cell line.

protein levels were not adequate for detection by immunoblot analysis.

LuxAB Fusion with Kozac Modifications

Although the Kozac sequence modifications on the 5′ end of the *luxA* gene was shown to drastically diminish bioluminescence levels in *E. coli* (Figure 11), the Kozac sequence is a mammalian ribosomal binding site and therefore to evaluate its true effectiveness the construct was tested in HEK293 cells. Stable cell lines harboring pcDNABfKoz were obtained and tested for bioluminescence activity. The light levels were significantly reduced (>90%) compared to HEK293 cells expressing the fusion protein without the Kozac modifications (Figure 14).

Alternative Expression Formats

Although detectable bioluminescence levels were obtained from mammalian cell lines harboring the LuxAB fusion protein, these levels were not sufficient for the creation of a reliable biosensor. Therefore, other expression formats were evaluated. The bioluminescence activity from the LuxAB fusion protein in HEK293 cells was compared to the expression of the individual *luxA* and *luxB* genes on either a single plasmid in a dual promoter format or by co-transfecting cells with separate plasmids carrying the genes. Stable cell lines expressing the *lux* genes in a dual promoter or co-transfected format were obtained. Furthermore, each vector format (fusion, dual promoter and co-transfection) was evaluated when the constructs were maintained as episomal plasmids or were integrated into the host's chromosome. The average bioluminescence levels (RLU/mg total protein) of the three brightest clones for each

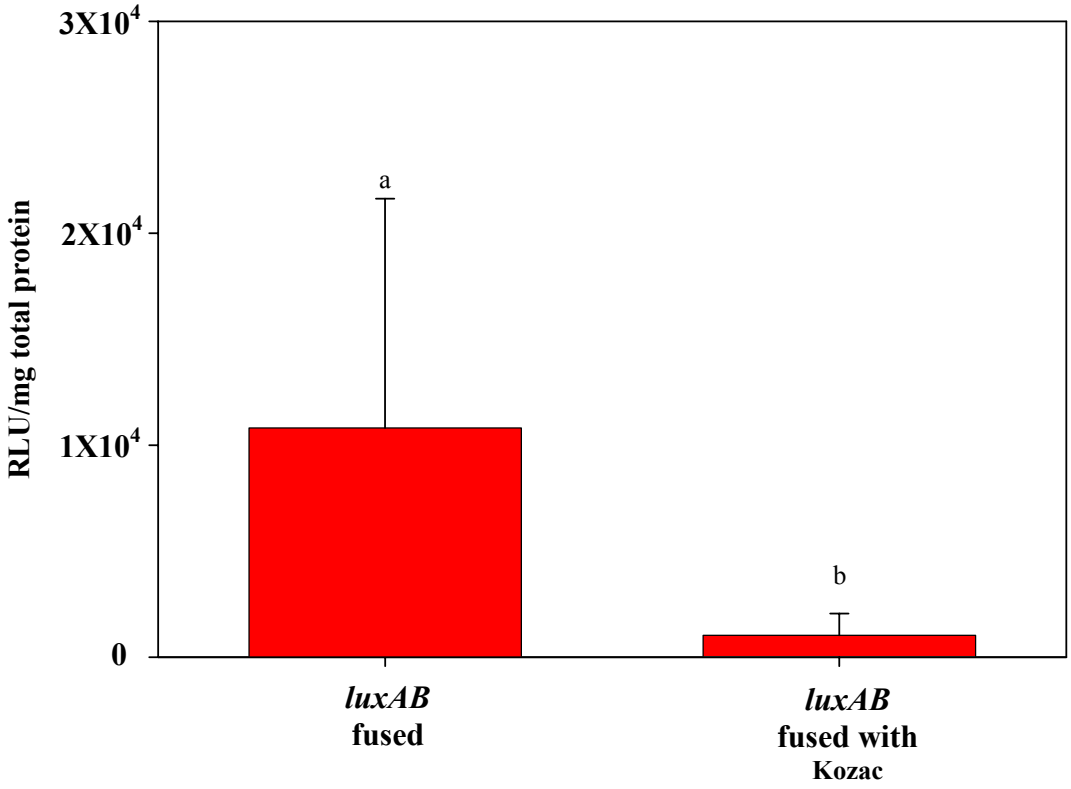


Figure 14. Relative bioluminescence (RLU/mg total protein) from stably integrated pcluxABf harboring the *luxAB* fusion gene or pcluxABfKoz harboring the luxAB fusion with the Kozac modification. Bioluminescence was measured from cell extracts upon the addition of 0.002% n-decanal and $FMNH_2$. Bioluminescence reported as an average of the top three clones from each cell line.

construct tested are shown in Figure 15. When the constructs were expressed as episomes within the nucleus of the cells, the bioluminescence was on average higher for the LuxAB fusion protein and for co-transfected plasmids. However, these differences were not statistically significant ($p<0.05$). The episomally expressed dual promoter plasmid carrying *luxA* and *luxB* genes (pCEPARB, Figure 6B) resulted in a significantly lower bioluminescence level than its integrated counterpart (pcDNARB, Figure 6A). This reduced amount of bioluminescence from cells episomally expressing pCEPluxARluxB was surprising. Further experiments were conducted in an attempt to identify the expression problems with this construct.

Since the Lux proteins were not detectable by immunoblot, mRNA levels were evaluated. RNA slot blot analysis revealed that clone DE4 had significantly higher amounts of *luxA* message compared to the *luxB* message. The other two clones tested harboring the construct (DE8 and DE9) had the exact opposite trend for message quantities (Figure 16). The backbone vector for this construct had only minimal hybridization with either the *luxA* or *luxB* probes.

Stability of Constructs in Mammalian Cells

To evaluate the stability of each construct in the absence of antibiotic selection, each HEK293 cell line clone was grown in complete growth media without antibiotic for twenty passages. The bioluminescence levels were obtained every fifth passage for comparison. In general, all clones were stable for at least five passages after the antibiotic removal. However, the constructs that were maintained as episomes began to decline in bioluminescence activity by passage ten. The co-transfected cell line

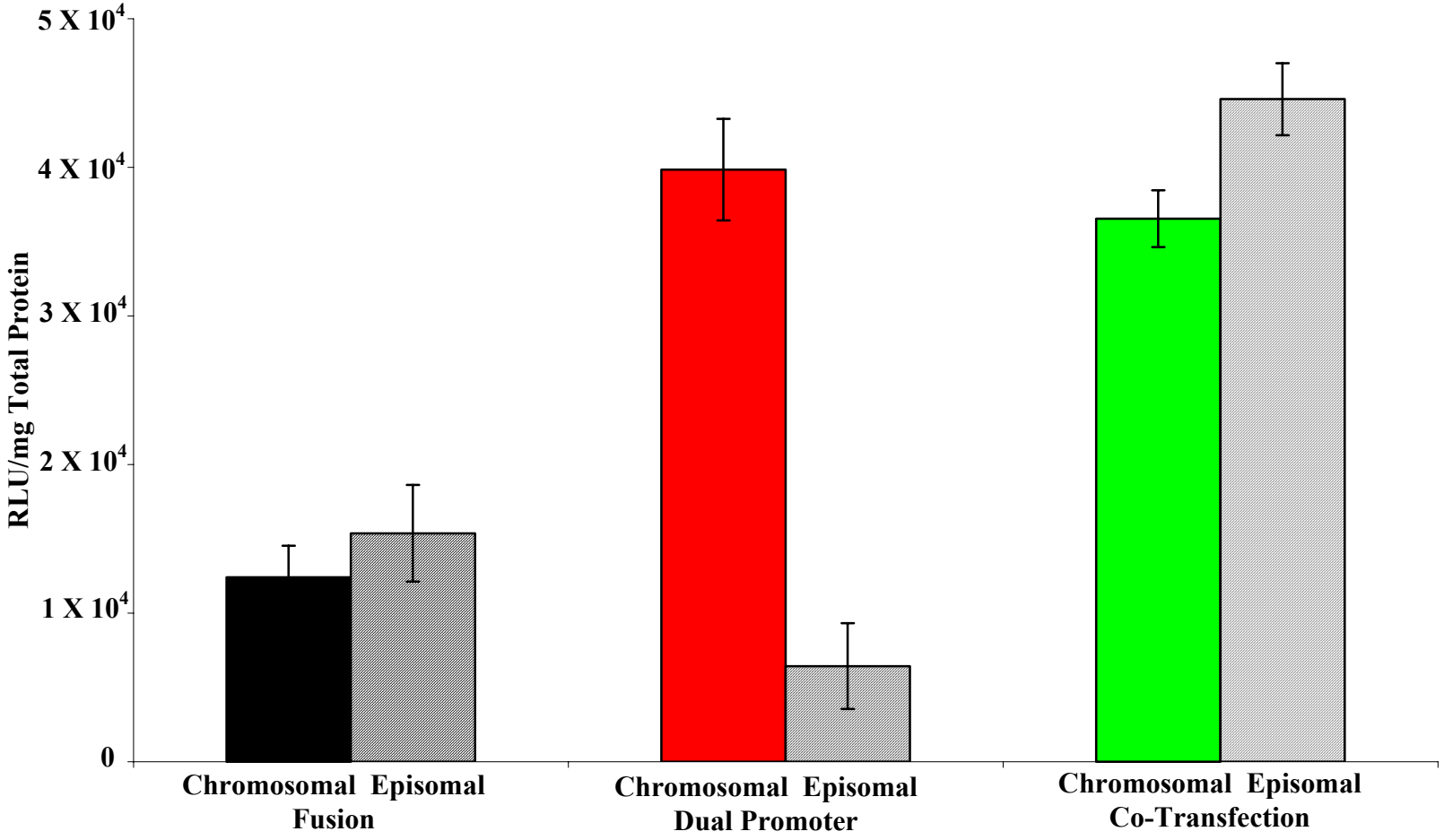


Figure 15. Relative bioluminescence (RLU/mg total protein) from either chromosomally integrated or episomal plasmids expressing *luxA* and *luxB* as a fusion protein (black), on one plasmid with separate promoters (red) or co-transfected on separate plasmids (green).

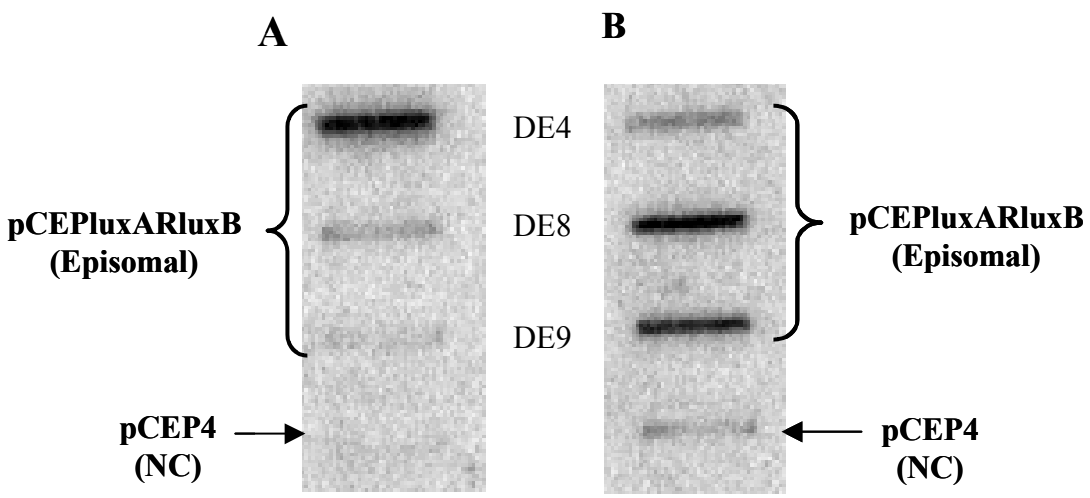


Figure 16. RNA slot blot analysis of HEK293 clones harboring the pCEPluxARluxB construct as an episome. **A.** Probed with a 300 bp region of the *luxA* gene from *P. luminescens*. **B.** Probed with a 300 bp region of the *luxB* from *P. luminescens*. A negative vector control (NC) of pCEP4 was added to determine background hybridization levels.

harboring two plasmids both maintained episomally had the fastest bioluminescence decline rate. The bioluminescence levels from this cell line declined approximately 10% per generation. The constructs that were integrated into the host's chromosome remained relatively stable throughout the twenty passage (approximately 2.5 months) evaluation (Figure 17).

Bicistronic Expression

To evaluate the expression of the *luxA* and *luxB* genes as a single bicistronic transcript, an IRES element from EMCV virus was used to transcriptionally fuse the two genes together (Figure 9). From each of the nine stable cell line clones obtained harboring this construct, the bioluminescence (RLU/mg total protein) was at least an order of magnitude greater than the average levels obtained from the brightest clones with any of the other expression formats tested (Figure 18). On average, there was no significant difference between bioluminescence levels obtained from HEK293 cells expressing the *luxAB* fusion, the *luxA* and *luxB* in a dual promoter format or as co-transfected separate plasmids. However, the bioluminescence levels from HEK293 cells harboring the *luxA* and *luxB* as a single bicistronic transcript consistently produced significantly higher light levels (Figure 19).

Because the Lux proteins were not detectable from stable cell lines, mRNA from the three brightest clones from each construct was isolated and probed with a [^{32}P] labeled probe complementary to the *luxA* sequence. *luxA* mRNA was detectable from every cell line tested above background vector control levels (Figure 20). However, the amount of transcript (determined as intensity values from autoradiougraphy) varied

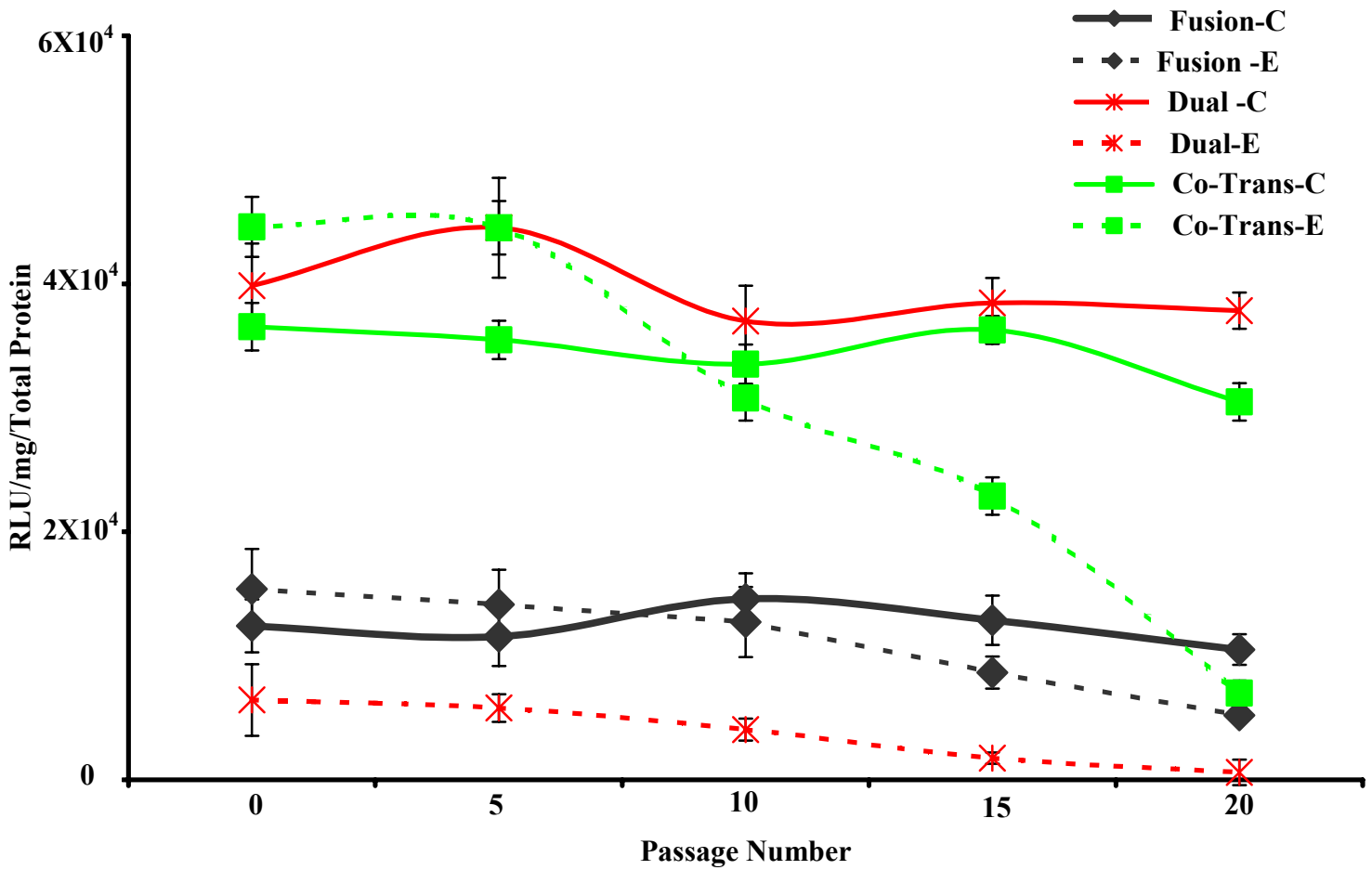


Figure 17. Relative bioluminescence levels (RLU/mg total protein) from stable HEK293 cell line clones grown in the absence of antibiotic. Solid lines indicate bioluminescence from cell lines that the constructs integrated into the host chromosome. Dashed lines indicate clones expressing the genes episomally.

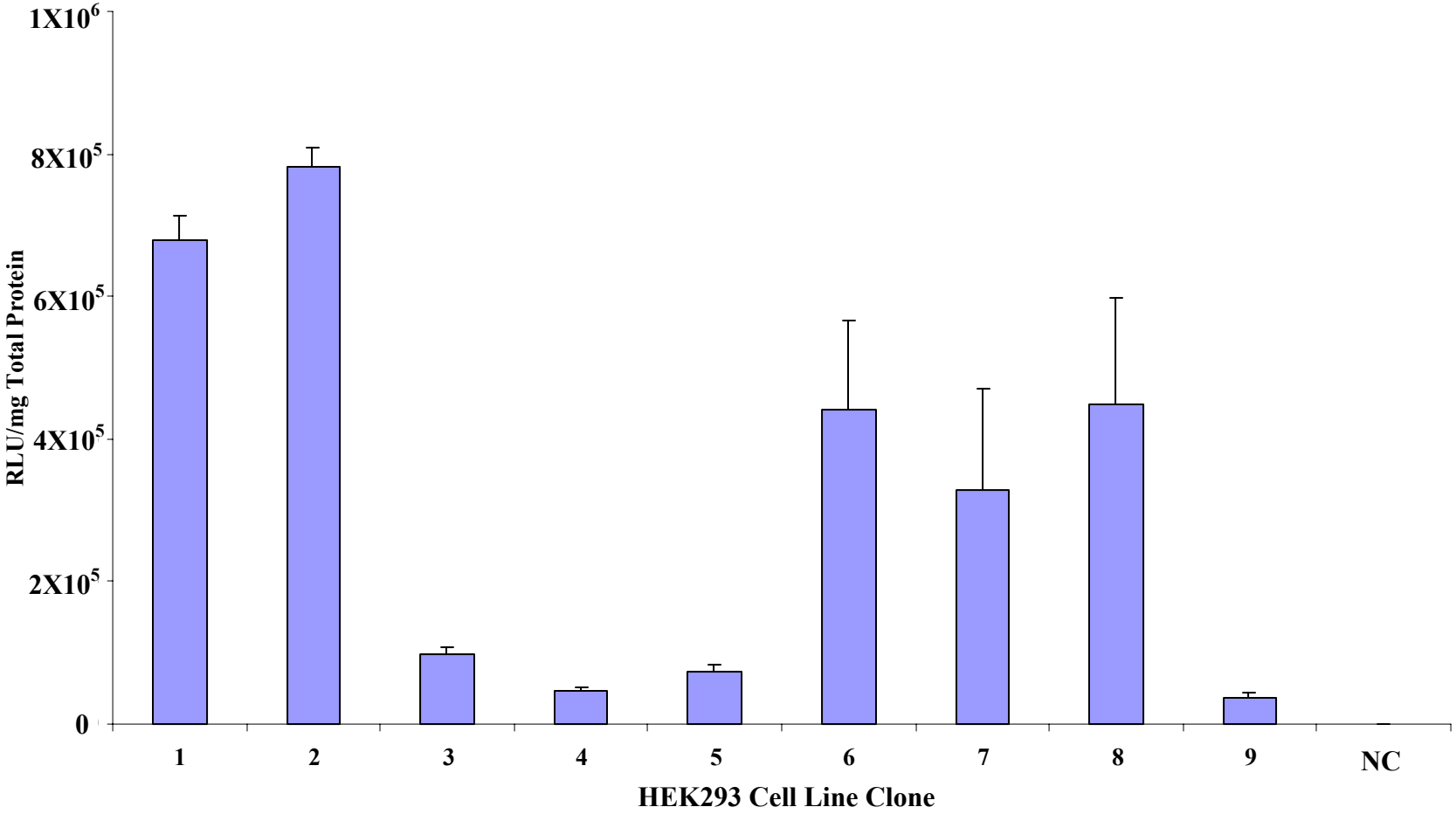


Figure 18. Bioluminescence levels (RLU/mg total protein) from stable HEK293 cell line clones expressing the *luxA* and *luxB* as a bicistronic transcript via an IRES element. The negative control is an HEK293 line transfected with the pIRES vector only.

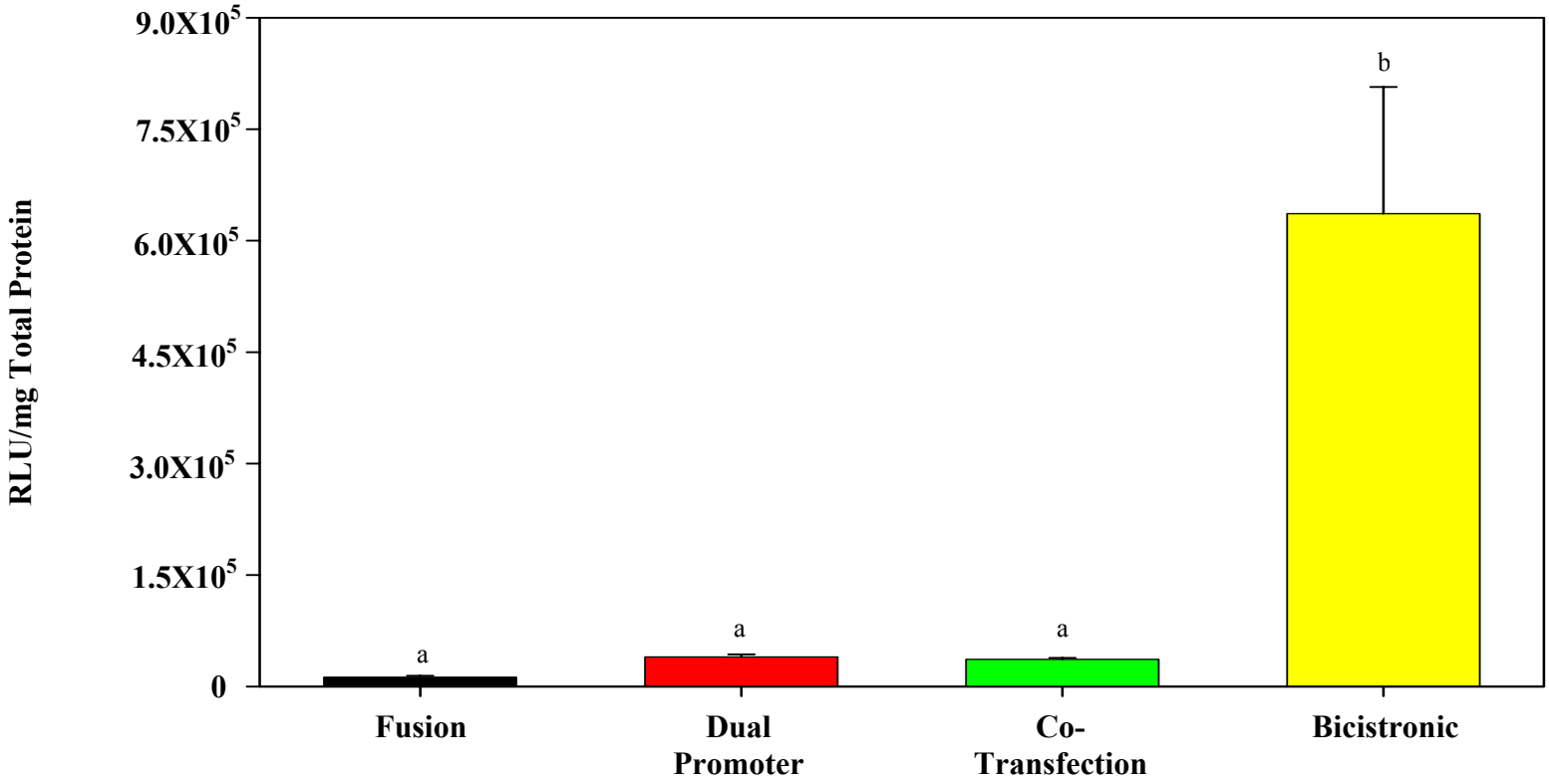


Figure 19. Relative bioluminescence levels (RLU/mg total protein) from stable HEK293 cell line clones expressing *luxAB* fusion (black), *luxA* and *luxB* from a dual promoter vector (red), co-transfection of the *luxA* and *luxB* on separate plasmids (green) or as a bicistronic transcript via an IRES element (yellow).

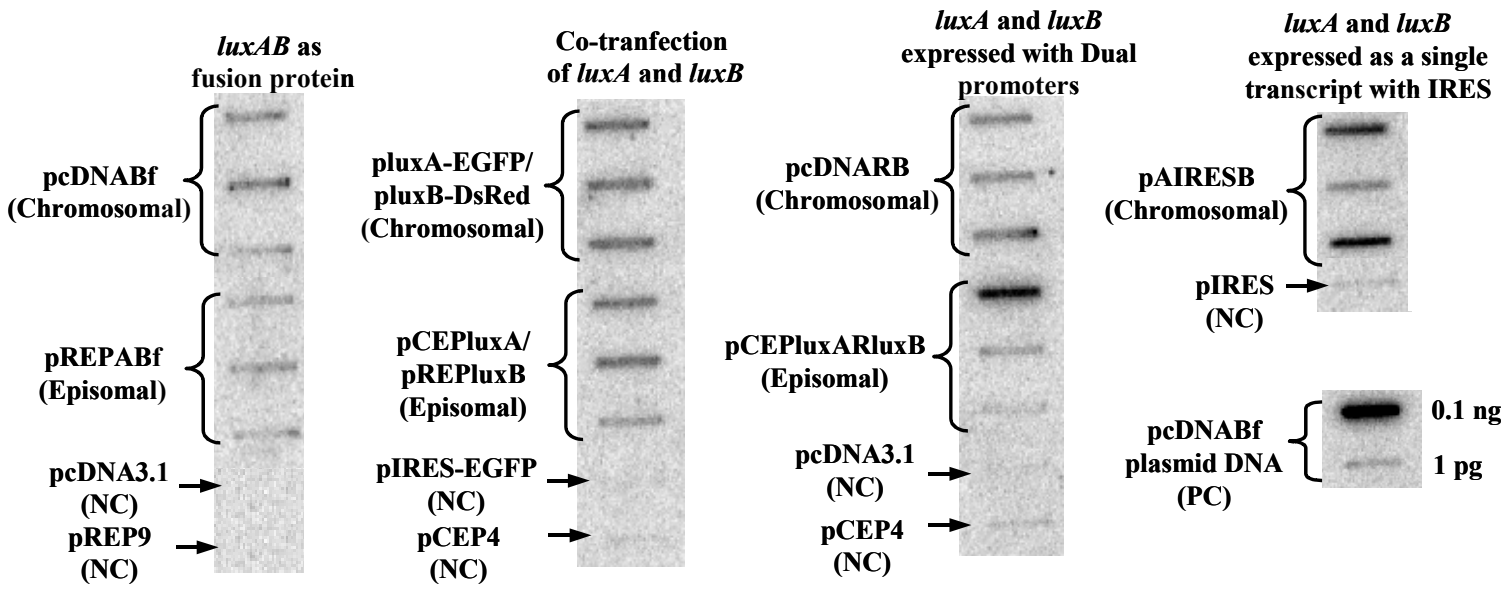


Figure 20. RNA slot blot analysis of HEK293 clones expressing the *luxA* and *luxB* genes either as a fusion protein, on separate plasmids (co-transfected), within the same plasmid but with separate promoters (dual promoter) or a transcriptional fusion linked with an IRES element. Negative vector controls (NC) were added to determine background hybridization levels. Plasmid (pcDNABf) DNA was added as a positive control (PC).

greatly between cell lines and clones and did not correlate with the bioluminescence levels obtained (R^2=0.15) (Figure 21).

Discussion

Previous work by Escher et al. (1989) showed that the *luxAB* fusion, using genes from *Vibrio harveyi,* was stable at elevated temperatures if initially expressed in *E. coli* at 23°C. However, when the fused protein was grown and expressed at 37°C there was a greater than 99% reduction in light. These data suggest that the fused *luxAB* does not fold properly at elevated temperatures. The luciferase from *P. luminescens* has a higher thermal stability ($t_{1/2}$ >3 h at 45°C) than *V. harveyi* ($t_{1/2}$ 5 min. at 45°C) (Meighen, 1991). Therefore, a translational fusion generated from the *P. luminescens luxA* and *luxB* genes was generated and evaluated. Although the *luxAB* fusion was functional in *E. coli*, bioluminescence activity was significantly reduced (70%) compared to the wild type unfused genes (Figure 11). In the unfused *luxAB* the α and β subunits are individually translated and are free to fold into their specific conformation (Tu and Mager, 1995). Therefore, the reduction in bioluminescence may be due to steric hindrance involved in the way the subunits form the heterodimer when expressed as a protein monomer. The addition of a Kozak sequence further reduced the bioluminescence level to approximately 5% of the wild type protein. Nevertheless, the Kozac sequence is a mammalian ribosomal binding site and therefore to evaluate its true effectiveness the construct was tested in HEK293 cells. The light levels were significantly reduced (>90%) compared to HEK293 cells expressing the fusion protein without the Kozac modifications (Figure 13). The addition of a Kozak sequence (G at

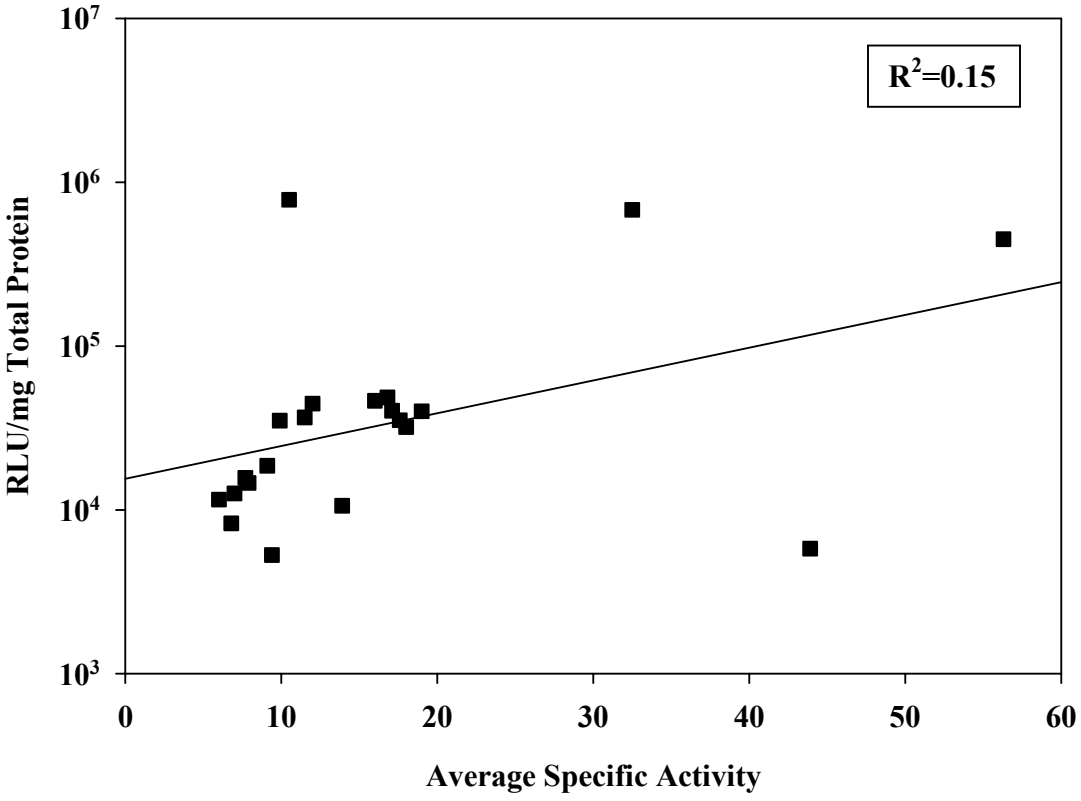


Figure 21. Relative bioluminescence (RLU/mg total protein) versus average intensity values from RNA slot blot analysis of each clone.

the +4 position) specifically changes the second amino acid of *luxA* from Lysine to Glutamic acid. These two amino acids have opposite net charges, which could result in a modification of the protein's secondary structure ultimately altering the protein's function resulting in decreased bioluminescence.

To determine temperature effect on the folding of the fusion protein, the fused and unfused versions of the *luxAB* were grown at 23˚C, 30˚C and 37˚C. In *E. coli*, there was not a statistical difference ($p < 0.05$) associated with temperature on bioluminescence as seen by Escher et al. This suggested that the folding problems in the *V. harveyi* LuxAB fusion protein were not present in the *P. luminescens* LuxAB fusion protein. However, when the fusion construct was expressed in the yeast, *S. cerevisiae* the bioluminescence levels significantly decreased as temperature increased to 37˚C (Table 6B). The differences seen in these two systems may be a result of the bacterial system's ability to transcribe the *luxB* independently due to the ribosomal binding site and *luxB* start codon still present in the fusion. When the fusion is expressed in the yeast system, the *luxB* is no longer independently expressed resulting in a true fusion protein that is unable to properly fold at 37˚C. The independent expression of the *luxB* in bacteria may have resulted in the unfused LuxB subunit forming the heterodimeric conformation with the LuxA within the LuxAB fusion resulting in the unaffected bioluminescence observed when the construct was expressed in *E. coli*.

In an attempt to mimic mammalian translation machinery, *in vitr*o transcription and translation of the *luxA, luxB* and *luxAB* fusion were performed in a rabbit reticulocyte lysate system. Although the 77kDa fusion protein was easily detected, the individual proteins, LuxA and LuxB, were not. This result was unexpected because

equal amounts of RNA transcript were added to each translation reaction. These data suggest that the formation of the heterodimeric (α and β subunit) complex may be required for not only efficient bioluminescence activity, but also for the overall stability of the protein.

Although detectable amounts of bioluminescence were obtained from mammalian cell lines harboring the LuxAB fusion protein, these levels were not sufficient for the creation of a reliable biosensor. Therefore, other expression formats were evaluated in an attempt to optimize bioluminescence activity. It was thought that by expressing the *lux* genes separately, the subunits would be able to form a more natural heterodimeric conformation. Human embryonic kidney cells (HEK293) were used for these evaluations. Cells were transfected with a dual promoter vector construct that that was developed to constitutively express each gene from a separate promoter or co-transfected with two plasmids each harboring either the *luxA* or *luxB* gene. Furthermore, to evaluate the differences in protein expression from genes integrated in the host's chromosome versus those constructs maintained as episomal plasmids, each expression format (fusion, dual promoter and co-transfection) was constructed on a plasmid backbone able to replicate episomally in HEK cells. The bioluminescence levels from stable cell lines harboring each expression variation were determined. Although there were slight variations in activity the differences were not statistically significant ($p=0.05$). The only exception was the reduced bioluminescence activity obtained from cells harboring a dual promoter vector episomally (Figure 15). The low light levels from these clones were somewhat surprising considering that the average bioluminescence from the fusion protein and from cells co-transfected with two

plasmids were slightly higher when the constructs were maintained as episomes. Upon further analysis, it was determined that the mRNA levels from the individual *lux* genes were not equal and therefore in this expression format, one promoter is inducing transcription at a higher rate than the other. This type of promoter occlusion where the transcription of one of the two promoters was significantly dampened has been seen previously (Horlick et al., 2000). The unequal availability of one of the *lux* subunits at a level over the other may prevent the proper formation of the heterodimeric active luciferase protein and may result in inactive homodimer formation.

In order for a bacterial *lux*-based mammalian bioreporter to be useful, the constructs need to remain stable in the absence of antibiotic selection for long periods of time. Efficient maintenance and stability of foreign genes requires that the DNA replicate once per cell cycle and be retained (integrated or episomally) in the nucleus. Expression plasmids harboring the *luxA* and *luxB* genes in three individual expression formats were created on both the traditional integration vectors and on Epstein-Barr virus (ori-P) based episomal plasmid vectors. To determine the stability of these constructs in HEK293 cells, the cell line clones were grown for twenty passages in complete growth media without antibiotic. In general, all clones (chromosomal and episomal) were stable for at least five passages after the antibiotic removal. However, the constructs that were maintained as episomes began to lose bioluminescence activity by passage ten with episomal co-transfected cells resulting in the fastest bioluminescence decline rate (Figure 17). Although there was a significant decline in bioluminescence activity from episomally based constructs over time, the light was not completely lost from any of the cells lines. Therefore, this reduction in

bioluminescence may be the result of natural plasmid loss and generation of a plasmid equilibrium within the cells. Immediately following transfection the DNA molecules within each cell can be very high and a natural decline in this concentration occurs to a steady state for plasmid maintenance (Middleton and Sugen, 1994 and Horlick et al.,1997). This number can vary, but the average is between 50 and 100 copies per cell with the further loss of approximately 5% per generation in the absence of selection (Yates and Guan, 1991). The constructs that were integrated into the host's chromosome remained relatively stable throughout the twenty passage (approximately 2.5 months) evaluation. These data indicate that integration of the *lux* genes within the host's chromosome may be the most suitable way to express the genes in mammalian cells for long-term gene maintenance and stable bioluminescence activity.

In order to optimize the bioluminescence potential from mammalian cells the *lux* genes need to be processed and expressed much in the way they are in bacteria. To establish a more natural expression format for the heterodimeric luciferase protein, the *luxA* and *luxB* genes were cloned into a bicistronic mammalian expression vector. This vector was developed to allow for the expression of two genes of interest under the control of a single constitutive promoter with the use of an internal ribosomal entry site (IRES). IRES elements can be defined as specific nucleotide sequences that allow for ribosomal entry and translation initiation directly at the start codon (AUG) rather than requiring scanning from the 5′ end, cap structure, of the mRNA (Pestova et al., 2001 and Kozac, 2001). Since the *lux* genes are naturally found in a polycistronic operon, it was thought that by expressing the genes in this format a more natural production and formation of the heterodimer could be obtained. From each of the stable cell line

clones obtained harboring *lux* genes expressed as a bicistronic transcript, the bioluminescence (RLU/mg total protein) was at least an order of magnitude greater than levels obtained with any of the other expression formats tested (Figure 18). On average, there was no significant difference between bioluminescence levels obtained from HEK293 cells expressing the *luxAB* fusion, the *luxA* and *luxB* in a dual promoter format or as co-transfected separate plasmids. However, the bioluminescence levels from HEK293 cells harboring the *luxA* and *luxB* as a single bicistronic transcript constitutively produced significantly higher light levels (Figure 19).

Based on these data it was determined that of the four expression formats evaluated that the bicistronic expression of the *luxA* and *luxB* genes was by far the best choice. Furthermore, although in general, the bioluminescence levels were slightly less, the stability of the construct when integrated into the host's chromosome makes this a more suitable choice for the development of bacterial *lux*-based mammalian biosensors.

CHAPTER 3

CODON OPTIMIZATION OF THE BACTERIAL LUCIFERASE FOR EXPRESSION IN MAMMALIAN CELLS

Introduction

The standard term "universal genetic code" comes from the fact that there are sixty-four possible codons for only twenty amino acids. Although the genetic code is degenerate, the alternate synonymous codons are not used with equal frequency (Sharp et al., 1988). In fact, it has been shown that there is not only a selective difference, but also a preference for certain codons in highly expressed genes (Sharp et al., 1993). This obvious codon bias has been shown to play a key role in the gene expression efficiency in all species tested to date (Amicis and Marchetti, 2000). Furthermore, codon usage patterns are not conserved between organisms of different species. This is especially true between genes from prokaryotes and eukaryotes.

An obvious first step in developing a mammalian cell line that utilizes the potential benefits of the bacterial luciferase enzyme system is to optimize the expression of the heterodimeric luciferase protein. The bioluminescence levels obtained through the expression of the wild type genes in various expression formats, although promising are not adequate for the development of reliable mammalian biosensors. Based on these data, the *lux* genes need further optimization in order to realize their full potential as mammalian reporter proteins.

Codon optimization is the term given to the synthetic creation of a gene sequence to possess the optimal codon usage patterns for the host organism. Several examples of successful codon optimization have been recently published. These

optimized proteins have been primarily designed for increased expression in mammalian hosts, as mammalian expression of foreign genes is often times limited (Narum et al., 2001). Several reporter proteins have been codon optimized in an attempt to increase expression in mammalian cells including the optimization of green fluorescent protein (GFP) (Zhang et al., 2002) and *Renilla* luciferase (Gruber and Wood, 2000) reporter genes. In both of these instances, codon optimization resulted in these reporter proteins becoming stronger reporters for gene expression and reliable monitoring tools in mammalian cells.

Based on this knowledge and the fact that further optimization is needed for the efficient expression of the bacterial luciferase in mammalian cells, it is hypothesized that by codon optimizing the *luxA* and *luxB* genes from *P. luminescens* that the bioluminescence activity from mammalian cell lines harboring these genes would be enhanced. The specific objectives of the this research are:

- To evaluate the *luxA* and *luxB* gene sequences from *P. luminescens* for codon usage pattern differences compared with optimal mammalian codon usage.
- To design a codon optimized sequence for the *luxA* and *luxB* genes to potentially allow for enhanced expression in mammalian cells.
- Compare the codon optimized sequences to the wild type genes using prediction analysis programs for mammalian expression.
- Synthesize complete codon optimized genes from oligonucleotides.

- To determine if codon optimization significantly improves the expression of bacterial luciferase in mammalian cells.
- To evaluate on what level of expression (transcription or translation) an increase in activity is derived.

Materials and Methods

Cell Culture and Plasmid Maintenance

All relevant constructs and strains, bacterial and mammalian, used in this study are outlined in Table 7. *E. coli* cells were routinely grown in Luria Bertani (LB) (Fisher Scientific, Pittsburgh, PA) broth containing the appropriate antibiotic selection with continuous shaking (200rpm) at 37°C. Kanamycin and Ampicillin were used at a final concentration of 50μg/ml and 100 μg/ml, respectively.

All cell culture reagents and media were obtained from Sigma Aldrich, (St. Louis, MO) unless otherwise stated. Mammalian cells were grown in the appropriate complete growth media containing 10% heat-inactivated horse serum, 0.01mM non-essential amino acids and 0.1mM sodium pyruvate in a Dubelco's minimal essential media base (DMEM) (M4655). Cells were routinely grown at 37°C in a 5% CO_2 atmosphere to confluency and split every three to four days by trypsinization at a 1:4 ratio and transfer into fresh complete growth media. Appropriate concentrations of antibiotic were used to maintain constructs after transfection according to susceptibility kill curve analysis. Kill curves were completed for each lot of antibiotic. The range of typical concentrations used for the selection of HEK293 cell line clones was between

Table 7. Strains and plasmids used in the determination of the optimal expression format for bacterial luciferase in mammalian cells.

Plasmid/Strain Designation	Relevant Genotype/ Characteristics	Source
Strains		
E. coli		
DH5α	Φ80d*lacZ*ΔM15, *recA*1, *endA*1, *gyrA*96, *thi*-1, *hsdR*17 (r_K-, m_K+), *supE*44, *relA*1, *deoR*, Δ(*lacZYA-argF*)U169	Gibco, BRL
TOP 10	F-, mcrA Δ(mrr-hsdRMS-mcrBC) Φ80/lacZ ΔlacX74 deoR recA1 araD139 Δ(ara-leu) 7697 ga/K rpsL endA1 nupG	Invitrogen
Mammalian Cells		
HEK293	Permanent line of primary human embryonal kidney transformed by sheared human adenovirus type 5 (Ad 5) DNA. ATCC# CRL-1573	ATCC
Plasmids		
pCR2.1-TOPO	TOPO TA cloning vector for easy cloning of PCR products generated with 3′ A overhangs Km^r, Amp^r	Invitrogen
pCR4-TOPO	TOPO TA cloning vector for easy cloning of PCR products generated with 3′ A overhangs designed for sequencing Km^r, Amp^r	Invitrogen
pIRES	Mammalian expression vector containing the internal ribosomal entry site (IRES) of the encephalomyocarditis virus between two multi-cloning sites which allows for the expression of two genes under the control of a single constitutive CMV promoter, Neomycin G418 antibiotic selection marker and a pUC ori and Km^r for replication in *E.coli*	Clontech
pCR4luxB	pCR4 TA cloning vector harboring the *luxB* from *Photorhabdus luminescens*	This Study
pNotIluxA	pCR4 harboring the *luxA* from *P. luminescens* with introduced NotI sites on both the 5′ and 3′ ends of the gene	This Study
pWTAI3	pIRES harboring the *luxA* from *P. luminescens* cloned into the MCS(A)	This Study

Table 7. Continued

Plasmid/Strain Designation	Relevant Genotype/ Characteristics	Source
pCOA#1	pCR4 vector harboring the codon optimized *luxA* gene with errors at base 365 and 1003.	This Study
pCOA#11	pCR4 vector harboring the codon optimized *luxA* gene with errors at bases 11, 28 and 365	This Study
pCOB#6	pCR4 vector harboring the codon optimized *luxB* gene with errors at base 321 and 829.	This Study
pCOB#7	pCR4 vector harboring the codon optimized *luxB* gene with errors at base 287 and 569.	This Study
pWTA-I-WTB	pIRES harboring the *luxA* (WTA) cloned into the MCS(A) and *luxB* (WTB) cloned into MCS(B) from *P. luminescens*	This Study
pCOA-I-WTB	pIRES harboring the codon optimized *luxA* (COA) into the MCS (A) and wild type *luxB* (WTB) into the MCS (B) from *P. luminescens*	This Study
pCOA-I-COB	pIRES harboring the codon optimized *luxA* (COA) into the MCS (A) and codon optimized *luxB* (COB) into the MCS (B) from *P. luminescencs*	This Study
WTA-I-WTB(1-20)	HEK293 cell lines stably transfected with the pWTA-I-WTB plasmid and selected by G418.	This Study
COA-I-WTB(1-20)	HEK293 cell lines stably transfected with the pCOA-I-WTB plasmid and selected by G418.	This Study
COA-I-COB(1-20)	HEK293 cell lines stably transfected with the pCOA-I-COB plasmid and selected by G418.	This Study

450µg and 650µg/ml.

Determining Codon Optimized Sequence of *Photorhabdus luminescens luxA* and *luxB* Genes

To determine a codon optimized sequence for *P. luminescens luxA* and *luxB* genes, the codon ratios within the wild type genes were analyzed and compared to optimal codon usage patterns from highly expressed (top 10%) mammalian genes. The optimal codon ratios were determined by information tabulated in Genbank. The overall ratio for usage of each codon within the wild type genes was altered to more closely match mammalian codon usage (Table 8A and 8B). In general, low frequency codons were used rarely or not at all and higher frequency codons were used more often. The codons were replaced within the wild type sequences in a random fashion. The sequence was further analyzed for any potential splice sites or other regulatory regions using the NetGene2 algorithm for prediction of potential acceptor and donor splice sites (www.cbs.dtu.dk/cgi-bin/nph-webface?jobid=netgene2). Any potential splice sites were removed. Transcription factor binding sites were also identified, however, these sequences were too numerous to successfully eliminate. After the final codon optimized sequence was determined, it was compared to the wild type sequence using the Genescan prediction algorithm (http://genes.mit.edu) to evaluate the potential expression of the new sequence versus the wild type.

Table 8. Codon usage of wild type versus codon optimized genes. (A) *luxA* (B) *luxB* WTA= wild type *luxA*, WTB= wild type *luxB*, COA= codon optimized *luxA* COB= codon optimized *luxB*.

A

Amino Acid	Codon	WTA	COA	Amino Acid	Codon	WTA	COA
Phe	TTT	14	6	Tyr	TAT	12	5
	TTC	5	13		TAC	5	12
Leu	TTA	9	-	Ter	TAA	0	-
	TTG	5	2		TAG	1	1
	CTT	9	4	His	CAT	10	3
	CTC	1	10		CAC	1	8
	CTA	2	-	Gln	CAA	11	6
	CTG	3	13		CAG	3	8
Ile	ATT	14	4	Asn	AAT	14	5
	ATC	4	20		AAC	6	15
	ATA	6	-	Lys	AAA	17	4
Met	ATG	9	9		AAG	6	19
Val	GTT	3	3	Asp	GAT	15	6
	GTC	2	6		GAC	8	17
	GTA	10	-	Glu	GAA	13	2
	GTG	6	12		GAG	9	20
Ser	TCT	5	-	Cys	TGT	4	3
	TCC	1	11		TGC	4	5
	TCA	6	1	ter	TGA	0	-
	TCG	2	-	Trp	TGG	6	6
Pro	CCT	2	5	Arg	CGT	4	-
	CCC	4	4		CGC	5	9
	CCA	2	2		CGA	2	-
	CCG	3	-		CGG	2	3
Thr	ACT	6	5	Ser	AGT	3	1
	ACC	1	14		AGC	1	5
	ACA	9	1	Arg	AGA	2	3
	ACG	4	-		AGG	0	-
Ala	GCT	13	15	Gly	GGT	7	4
	GCC	3	9		GGC	5	11
	GCA	3	-		GGA	9	5
	GCG	5	-		GGG	5	6

Table 8. Continued

B

Amino Acid	Codon	WTB	COB	Amino Acid	Codon	WTB	COB
Phe	TTT	13	5	Tyr	TAT	16	4
	TTC	5	13		TAC	1	13
Leu	TTA	9	-	Ter	TAA	1	1
	TTG	6	-		TAG	0	-
	CTT	3	3	His	CAT	8	2
	CTC	2	2		CAC	3	9
	CTA	2	-	Gln	CAA	6	3
	CTG	3	18		CAG	5	8
Ile	ATT	13	5	Asn	AAT	18	4
	ATC	5	21		AAC	6	20
	ATA	8	-	Lys	AAA	16	4
Met	ATG	8	8		AAG	6	20
Val	GTT	13	3	Asp	GAT	17	3
	GTC	3	7		GAC	3	17
	GTA	4	-	Glu	GAA	23	6
	GTG	1	11		GAG	8	25
Ser	TCT	3	-	Cys	TGT	4	1
	TCC	1	8		TGC	2	5
	TCA	4	-	ter	TGA	0	-
	TCG	0	-	Trp	TGG	2	2
Pro	CCT	6	7	Arg	CGT	1	-
	CCC	1	2		CGC	2	3
	CCA	3	2		CGA	0	-
	CCG	1	-		CGG	1	1
Thr	ACT	4	4	Ser	AGT	7	2
	ACC	2	12		AGC	2	7
	ACA	6	-	Arg	AGA	3	3
	ACG	4	-		AGG	0	-
Ala	GCT	8	14	Gly	GGT	6	1
	GCC	3	6		GGC	3	9
	GCA	4	-		GGA	3	2
	GCG	5	-		GGG	2	2

Synthesizing the Codon Optimized luxA and luxB Genes

Once the codon optimized sequence had been determined, oligonucleotides for each gene were designed that covered the complete sequence (Table 9& 10). Each oligo was designed with an 18 – 23 base pair overlap on the 5′ and 3′ ends with its adjacent oligos. These overlapping regions were designed with Tm values of 53°C - 56°C. Once the oligos were designed they were synthesized by Sigma Genosys (Sigma, St. Louis, MO) and polyacrylamide gel (PAGE) purified to ensure full-length products. Each oligonucleotide was placed into a PCR reaction with the following conditions; internal oligos (0.25 pmol), the two outermost oligos (25 pmols), dNTP mixture (200 nm), 1X *Pfu* buffer, 1X *Pfu* Enhancer solution, $MgCl_2$ (concentration determined experimentally) and 1U of *Pfu* DNA polymerase (Stratagene, La Jolla, CA).

All PCR reactions were performed in 0.2 ml thin walled PCR tubes using a PTC-225 DNA Engine (MJ Research, Waltham, MA). For gene synthesis the following program was used; (1) initial denaturation 95°C for 5 min, (2) 30 cycles of 94°C for 1 min, 50°C for 1 min and 68°C for 2 min followed by (3) final extension 68°C for 10 min. Resultant PCR products were run on 1% agarose gels in 1X TBE. Unfortunately, there were no detectable products of the correct size. As an alternative strategy, four separate reactions were set up with four adjacent oligos in each reaction (Figure 22). The two innermost primers were added at a final concentration of 0.25 pmols and the two outermost oligos were used as both template and primers at a concentration of 25 pmols. Each piece was then amplified using the parameters outlined above with the exception of the extension step was reduced from 2 min to 45 sec. The resultant PCR

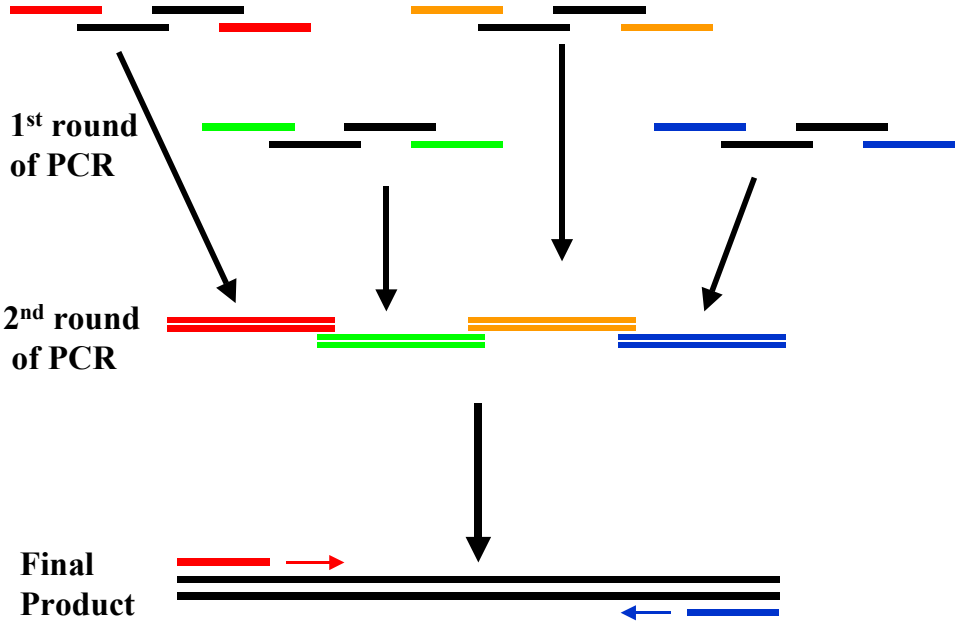


Figure 22. Schematic diagram of the recursive PCR method used to construct the synthetic *luxA* and *luxB* genes. Outside oligonucleotides were added at 25 pmol final concentration while the inside oligonucleotides were added at 0.25 pmol final concentration.

Table 9. Oligonucleotide primer sequences used to synthesize the codon optimized *luxA* gene.

Primer Name	Sequence 5′ to 3′
COA1 (Sense)	5′CGAAACCATGAAGTTCGGCAACTTCCTGCTCACATATCAGCC TCCCCAGTTTTCCCAAACCGAGGTCATGAAGCGGCTGGTTA AGCTCGGCCGCATCTC C 3′
COA2 (Antisense)	5′ AAGCAGCAGCGACATAAGGGTTACCAAGCAGGCCGAA CTCGGTGAAGTGGTGCTCCAGCAGCCACACGGTGTCGAAAC CGCACTCCTCGGAGATGCGGCCGAGCTTA 3′
COA3 (Sense)	5′ CCCTTATGTCGCTGCTGCTTATCTGCTCGGCGCCAC CAAGAAACTGAACGTCGGCACTGCCGCTATCGTTCTC CCCACCGCCCATCCAGTCCGCCAGCTT 3′
COA4 (Antisense)	5′ GAAGTCCTTGTTGTAAAGCCCGCGGCAGATGCCGAA CCGAAAGCGCCCCT TGGACATTTGATCCAGCAAGTTC ACGTCCTCAAGCTGGCGGACTGGATGG 3′
COA5 (Sense)	5′ CGGGCTTTACAACAAGGACTTCCGCGTGTTCGGCA CCGACATGAACAACAGCCGCGCCCTGGCCGAGTGTT GGTACGGGCTGATCAAGAATGGCATGA 3′
COA6 (Antisense)	5′ GAGCGCCACCTCTGCTGTAAGCGGCGGGGTTCACTT TGACTTTGTGGAACTTGATGTGCTCATTGTCGGCTTCC ATGTATCCCTCGGTCATGCCATTCTTGATCAGCC 3′
COA7 (Sense)	5′ ACAGCAGAGGTGGCGCTCCTGTTTATGTGGTGGCTG AGTCAGCTAGTACCACTGAGTGGGCTGCTCAATTTGG CCTCCCTATGATCCTGTCCTGGATCATCAACAC 3′
COA8 (Antisense)	5′ CAGGCAGTGGTCGATGTTATGAATGTCGTGCCCG TACTCTTGAGCCACTTCGTTGTAAAGCTCGAGCTGG GCCTTCTTCTCATTAGTG TTGATG ATCCAGGACAGG 3′
COA9 (Sense)	5′ CATAACATCGACCACTGCCTGTCCTACATCACCTC CGTGGACCACGACTCCATCAAGGCCAAGGAGATTTG CCGGAAGTTTCTCGGGCATTGGTATGATAG 3′
COA10 (Antisense)	5′ AACACGAAATCGCGCCACTGCCCCTTGTTGAAGTC GTAACCTCTGGTCTGGTCGGAGTCGTCAAAGATAGTG GTAGCATTCACG TAGCTAT CATACCAATGCCCGAG 3′

Table 9. Continued

Primer Name	Sequence
COA11 (Sense)	5′ AGTGGCGCGATTTCGTGTTGAAAGGACACAAGG ATACTAACAGACGCATCGACTACAGCTACGAGATCAA TCCCGTGGGCACCCCTCAG GAGTGCATTGACATCATCC 3′
COA12 (Antisense)	5′ ATGGAAGCGATGATCTCGTCCACGGTTCCGTTAGCCTCA AATCCACAACAGATGTTGGAGATTCCGGTAGCATCAA TGTCCTTTTGGATGATG TCAATGCACTCCTG 3′
COA13 (Sense)	5′ GACGAGATCATCGCTTCCATGAAGCTCTTCCAGTCCG ATGTCA TGCCATTCCTCAAGGAGAAGCAACGCA GCCTCCTGTACTAGGGATCC 3′
COA14 (Antisense)	5′ GGATCCCTAGTACAGGAGGCTGC 3′

Table 10. Oligonucleotide primer sequences used to synthesize the codon optimized *luxB* gene.

Primer Name	Sequence 5′ to 3′
COB1 (Sense)	5 ′CGAAACCATGAAGTTCGGACTGTTCTTCCTTAACTTCATCA ACTCCACCACTGTGCAGGAGCAAAGCATCGTGCGCATGCAG GAGATCACCGAGTATGTGGAC 3′
COB2 (Antisense)	5 ′CACAGTCAGAGGAGCGCCGACAACGCCATTGTCGGAAAA GTGGTTCTCGTACACCAGGATCTGCTCGAAGTTCAGCTTG TCCACATAC TCGGTGATCTCC 3′
COB3 (Sense)	5 ′ GGCGCTCCTCTGACTGTGTCCGGCTTCCTGCTCGGCCT GACCGAGAAGATCAAAATTGGCTCCCTGAACCACATCAT CACCACTCATCATCCTGTCGCCATCGCT 3′
COB4 (Antisense)	5 ′ GTGCATCTCGTCCTTCTTCTCGCAATCGCTGAACCCC AGGATGAATCTCCCCTCGCTCAGCTGATCCAGCAGGCA AGCCTCCTCAGCGATGGCGACAGGATG 3′
COB5 (Sense)	5 ′GAGAAGAAGGACGAGATGCACTTTTTCAACCGCCCTGT GGAATATCAGCAG CAACTGTTTGAAGAGTGCTACGAGAT CATTAACGACGCTCTGACCACCGGCTACTGC 3′
COB6 (Antisense)	5 ′AGCGGTGACATACTTCCGAGGGCCGCCTGGGGTGTAA GCGTGGGGGTTGACGGAGATTTTAGGGAAGCTGTAG AAGTCATTGTCGGGGTTGCAGTAGCCGGTGGTCAG 3′
COB7 (Sense)	5 ′TCGGAAGTATGTCACCGCTACCAGTCATCACATCGTGG AGTGGGCTGCCAAG AAAGGCATCCCTCTCATCTTTAAGT GGGATGACTCCAACGACGTGAGATACGAGTA 3′
COB8 (Antisense)	5′ TAACCAGGATCATCAGCTGGTGGTCGATTTCGGACAG GTCAACGTCATATTTGTCAGCCACGGCCTTGTATCTC TCAGCGTACTCGTATCTCACGTCGTTGG 3′
COB9 (Sense)	5′ CCAGCTGATGATCCTGGTTAACTACAACGAAGACAGC AACAAGGCTAAG CAGGAGACCCGCGCCTTCATTAGCGA CTACGTGCTTGAAATGCACCCTAAC 3′
COB10 (Antisense)	5′ CCAGCTTAGCAGCAGTGATACACTCGGTGTAGTTTCCG ACAGCGTTCTCGGCGATGATTTCCTCAAGCTTGTTCTCGA AGTTCTCGTTAGGGTGCATTTCAAGCAC 3′

Table 10. Continued

Primer Name	Sequence
COB11 (Sense)	5′ TGTATCACTGCTGCTAAGCTGGCCATCGAGAAGTGCGGT GCTAAGAGTGTCCTGCTGTCCTTTGAGCCAATGAATGAC CTGATGAGCCAAAAGAACGTCAT 3′
COB12 (Antisense)	5′ GGATCCTTAGGTGTACTCCATGTGGTACTTCTTAATATTG TCGTCCACAATGTTGATGACGTTCTTTTGGCTCATCAG 3′

products were then gel purified using the Geneclean gel extraction kit according to the manufacturer's instructions (Bio101, Carlsbad, CA). The extracted products were quantified using a Dyna Quant 200 fluorometer (Hoefer Pharmacia Biotech Incorporated, San Francisco, CA) and placed into a second PCR reaction at equal molar concentrations (0.25 pmols). The two outermost (5′ and 3′) oligos were used as primers at a final concentration of 25 pmols. After the second PCR reaction, the products of the correct size were again gel purified as previously described. Because *Pfu* polymerase produces blunt end products, 3′ A overhangs were added to allow for TA TOPO cloning of the products. To accomplish this, the gel-extracted product was mixed with dATP (200nM) 1X amplitaq buffer and 1U of *Taq* polymerase (Amersham Pharmacia, San Francisco, CA) and placed at 72°C for 20 – 30 min. Immediately following the addition of the A's, the product was TA TOPO cloned into the pCR4 TOPO cloning vector (Invitrogen Corporation, Carlsbad, CA). Resultant colonies were then checked for insert by an *Eco*RI restriction digest and sequenced to ensure their integrity.

Site Directed Mutagenesis

Although the oligos were successfully joined into a double stranded synthetic gene, several point mutations were determined by sequencing. A number of clones for each gene were completely sequenced in an attempt to identify a flawless clone without success. To correct these errors, site directed mutagenesis was done. First, for the codon optimized *luxA* gene, two separate clones pCOA#1 and pCOA#11 were used as template. Site directed mutagenesis primers were designed to introduce the necessary changes. The complete *luxA* sequence was amplified in two separate sections (365 bp

from pCOA#1 and 719bp from pCOA#11) that overlapped between the bases where the necessary changes were required. Each segment was gel purified and then linked back together by a second round of PCR as described for the original gene synthesis. Subsequently, 3′ A overhangs were added and the product TA TOPO cloned into pCR4. Upon sequencing, a construct with the correct sequence was identified and termed pPA2. Site directed mutagenesis was also performed on the codon optimized *luxB* sequence using overlapping primers designed to introduce the proper changes. The complete *luxB* sequence was amplified in three segments (324 bp from pCOB#7, 340 bp from pCOB#6 and 319 bp pCOB#7) from two separate clones (pCOB#7 and pCOB#6) and subsequently linked by PCR as previously described. A construct of the correct sequence was produced and termed pPB2.

Construction of a Bicistronic Expression Vector

To compare the expression of the codon optimized *luxA* and *luxB* genes to the wild type, the pIRES vector was used (Clontech Corporation, Palo Alto, CA). This expression vector contains two multi-cloning sites separated by an internal ribosomal entry site (IRES) from encephalomyocarditis virus (EMCV). The IRES element allows for the expression of two genes (one cloned into each multi-cloning site) from a single constitutive CMV promoter. For comparison purposes, a wild type *luxA* and *luxB* (pWTA-I-WTB) construct, a codon optimized *luxA* and wild type *luxB* (pCOA-I-WTB) construct and a codon optimized *luxA* and codon optimized *luxB* (pCOA-I-COB) construct were generated.

pWTA-I-WTB

To create this construct, the *luxA* gene from *P. luminescens* was amplified from pPLluxCDABE plasmid that harbors the complete *luxCDABE* cassette (Table 7) and unique *NotI* restriction sites were introduced on both the 5′ and 3′ ends of the *luxA* gene. The resultant PCR product was TA TOPO cloned into pCR4 TOPO to generate pNotIluxA. The *luxA* gene was then cloned into the MCS(A) of pIRES via the unique *NotI* restriction sites to generate pWTAI. Once this construct was confirmed by sequencing, the plasmid was purified using the Wizard midi-prep plasmid purification kit according to the manufacturer's instructions (Promega Corporation, Madison, WI). The *luxB* gene was cleaved via a 5′ *XbaI* and 3′ *SpeI* site from pCRluxB and cloned into the *XbaI* site within the MCS(B) of pWTAI to generate pWTA-I-WTB (Figure 23A).

pCOA-I-WTB

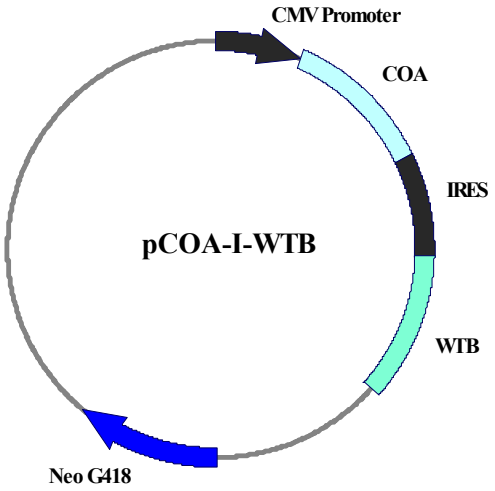
To generate this construct, the codon optimized *luxA* gene (COA) was cleaved from pPA2 via unique *NotI* restriction sites and cloned into the MCS(A) of the pIRES vector (Clontech Corporation, Palo Alto, CA) to generate pCOAI. Once this construct was confirmed by sequencing, the plasmid was purified using the Wizard midi-prep plasmid purification kit according to the manufacturer's instructions (Promega Corporation, Madison, WI). The wild type *luxB* gene was cleaved via a 5′ *XbaI* and 3′ *SpeI* site from pCRluxB and cloned into the *XbaI* site within the MCS(B) of pCOAI to generate pCOA-I-WTB (Figure 23B).

pCOA-I-COB

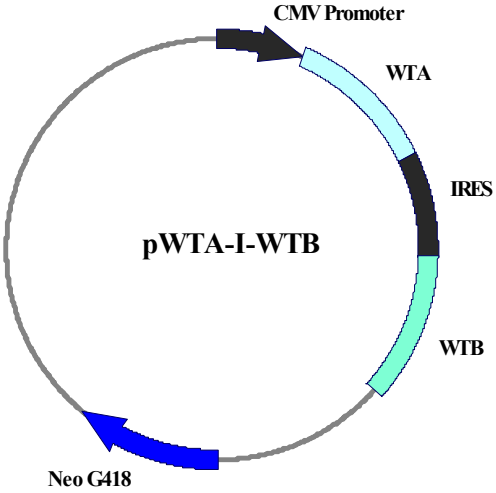
To generate this construct, the codon optimized *luxB* (COB) gene was cleaved from pPB2 via introduced 5′ and 3′ *XbaI* sites and cloned into the MCS(B) from pCOAI to create pCOA-I-COB (Figure 23C).

Figure 23. Schematic diagram of the final constructs used to compare the wild type *luxA* and *luxB* to the codon optimized genes. A. Wild type *luxA* and wild type *luxB* B. Codon optimized *luxA* and wild type *luxB* C. Codon optimized *luxA* and codon optimized *luxB*.

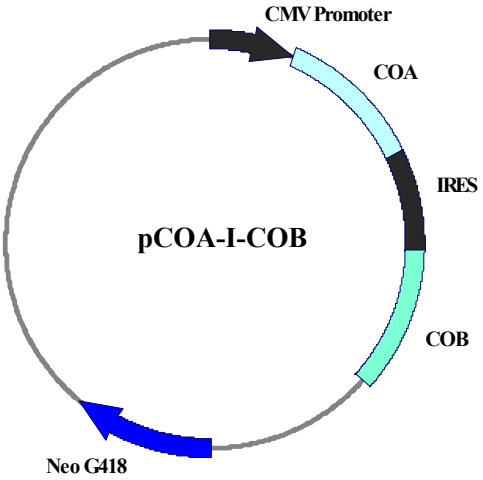
A


CMV Promoter
COA
IRES
pCOA-I-WTB
WTB
Neo G418

B


CMV Promoter
WTA
IRES
pWTA-I-WTB
WTB
Neo G418

C


CMV Promoter
COA
IRES
pCOA-I-COB
COB
Neo G418

Ligation Reactions

Plasmid vectors and inserts were digested (2-6 h) with the appropriate enzymes (Promega Corporation, Madison, WI). Linearized vectors were dephosphorylated using a calf intestine alkaline phosphatase enzyme according for the manufacturer's instructions (Promega Corporation, Madison, WI). Both vector and insert DNA were gel purified from 1% agarose gels using the Geneclean gel extraction kit (Bio101, Carlsbad, CA). The recovered DNA was then quantified using a Dyna Quant 200 fluorometer (Hoefer Pharmacia Biotech Incorporated, San Francisco, CA) and ligations were set up as 20µl reactions using a 3:1 molar ratio of insert to vector DNA. The ligation reactions were then incubated at 17°C overnight.

Electroporation

Electrocompetent cells were prepared as outlined by the manufacturer (BTX, San Diego, CA). Electroporations were performed using the BTX Electroporator 600 with the following conditions: 40µl cells, 1-2µl ligation mixture, a 2.5kV pulse for 4.7ms using a 2mm gap cuvette. After the pulse, cells were immediately resuspended in 1ml of sterile LB and allowed to recover for 1 h at 37°C (200 rpm). Cells were then plated on selective media containing the appropriate antibiotic.

Selection of Bacterial Clones

Resistant colonies were picked after 24 h and expanded to patches on grid plates. To test for proper insert presence and orientation, rapid boil plasmid mini-preps (Promega Corporation, Madison, WI) were done followed by the digestion of the

plasmid with the appropriate restriction enzyme mixture according to the manufacturer's instructions (Promega Corporation, Madison, WI). Products were run on 1% agarose gels to determine if the banding pattern indicated the insert presence and proper orientation. Upon identifying correct clones, the plasmids were further purified using the Wizard midiprep plasmid purification system according to the manufacturer's protocol (Promega Corporation, Madison, WI) and sequenced.

Sequencing

All constructs were sequenced to ensure their integrity. Sequencing was done in the University of Tennessee Molecular Biology Service Facility using an Applied Biosystems 3100 Genetic Analyzer sequencer (Foster City, CA).

Transfection of Mammalian Cells

Transfection of mammalian cell lines was done in six well poly-D-lysine coated tissue culture plates (Fisher Scientific, Pittsburgh, PA). Cells were split from stock cultures and inoculated into each well at approximately $1X10^5$ cells per well in complete growth media. The plate was then placed at 37°C in a 5% CO_2 atmosphere for 1-2 days until the cells became 80-90% confluent. The day of transfection, the media was refreshed. DNA for transfections was purified from 100ml overnight *E. coli* cultures using the Wizard Purefection plasmid purification kit to remove endotoxins according to the manufacturer's instructions (Promega Corporation, Madison, WI). For chromosomal integration, the plasmid DNA was linearized before transfection to increase proper integration.

HEK293 Cells

Purified plasmid DNA (3.2 μg) was mixed into 200 μl of serum free DMEM in a 1.5 ml tube. In a second tube, 8 μl of Lipofectamine 2000 reagent (Invitrogen Corporation, Carlsbad, CA) was added to 200 μl of serum free DMEM. The lipofectamine mixture was added to the DNA mixture within 5 min and incubated at room temperature for 20 min. The entire mixture (400 μl total) was added directly to the appropriate well on the plate and rocked back and forth to ensure adequate mixing. Twenty-four hours post transfection, the complexes were removed and the media was replaced with fresh complete growth media supplemented with the appropriate antibiotic for selection.

Selection of Mammalian Cell Clones

Twenty-four hours post transfection, selective media was added to all wells and refreshed every three to four days. Within two weeks all control wells were dead and the transfected cells were forming small colonies on the plate surface. Colonies were separated from the rest of the well by placing a sterile chamber around the cell mass and sealing it with silicon (Fisher Scientific, Pittsburgh, PA). The media could then be removed and each colony could be trypsinized and transferred to individual tissue culture flasks. To accomplish this, after washing with a PBS solution, 200 μl of a 1X Trypsin-EDTA solution (Sigma Aldrich, St. Louis, MO) was added directly to the chamber and incubated at 37°C for 3 to 5 min. The trypsin-EDTA solution was then replaced with complete growth media and the cells were transferred to a 25cm^2 tissue culture flask for propagation. Each clone was given a number and expanded to

individual cell lines. Each line was split and maintained as described earlier with the addition of selective media. Twenty cell lines were propagated in this manner for each plasmid tested.

Bioluminescence Assays from Mammalian Cells

To determine bioluminescence potential from each cell line clone, total proteins were extracted and *in vitro* enzyme (bioluminescence) assays performed. To extract the proteins, the cells were trypsinized from the plate or flask surface using standard protocols and resuspended into 2.0 ml Sarstedt tubes (Fisher Scientific, Pittsburgh, PA). The cells were then spun down and washed two times in sterile phosphate buffered saline (PBS) to remove any residual media (Sigma Aldrich, St. Louis, MO). Cell pellets were then resuspended into 1 ml 0.1M potassium phosphate buffer pH 7.8 and disrupted by three consecutive cycles of freeze (30 s liquid N_2) thaw (5 min at 37°C) extraction. After disruption, the cell debris was pelleted by spinning the samples at 14,000Xg for 5 min and the supernatant was used in the bioluminescence assay. To determine light intensity, the protein extract was mixed with 0.1 mM NAD(P)H, 4 μM FMN, 0.2% (w/v) BSA, 0.002% (w/v) n-decanal. Bioluminescence was measured using the FB14 luminometer (Zylux Corporation, Pforzheim, Germany) at a 1 s integration and reported as relative light units (RLU). To determine if $FMNH_2$ was a limiting factor for the bioluminescence reaction, a flavin oxidoreductase enzyme (1U) isolated from *V. harveyi* (Roche Scientific, Indianapolis, IN) was added to the bioluminescence assay and the light levels were measured again for comparison.

Bioluminescence signals were normalized between samples and cell lines by dividing the RLU measurement by the total protein and reporting the bioluminescence as RLU/μg total protein. Protein concentrations were determined using the Coomassie Plus protein assay according to the manufacture's instructions (Biorad, Hercules, CA).

In Vitro Transcription/Translation

To determine if the *lux* genes could be translated *in vitro* in rabbit reticulocyte lysate (mammalian translation machinery), pIRES vector harboring the wild type *luxA* (WTA), and codon optimized *luxA* (COA) were transcribed and translated. First, the plasmid DNA containing the genes was digested at a unique *Xba*I restriction site at the 3′ end of the gene within the vector. This digestion linearized the plasmid and allowed for the generation of run-off transcript from the vector derived T7 promoter. Each gene was transcribed via T7 polymerase using the RiboMax large-scale transcription system (Promega Corporation, Madison, WI). Three individual transcription reactions were set up along with a positive T7 control and a negative control containing no template DNA. Each reaction was set up according to the manufacturer's protocol and then incubated at 37°C for 1 h. Transcripts were quantified by absorbance (260/280) measurements (Beckman Coulter, Fullerton, CA). Ten micrograms per ml of total RNA transcript was then added to 50 μl (total volume) rabbit reticulocyte lysate translation reactions. Each reaction was gently mixed on ice according to the manufacturer's protocol for S^{35} labeled protein generation and then incubated at 30°C for 90 min (Promega Corporation, Madison, WI). Once translation was complete, 15 μl of each reaction was

loaded onto a 12% SDS-PAGE mini-gel and run at 30 mA for 1 h. The gel was removed and dried at 60°C with vacuum pressure using a model 443 Slab Dryer (BioRad, Hercules, CA) onto 3MM filter paper (Fisher Scientific, Pittsburgh, PA). To visualize the generated proteins, the gel was placed onto an intensifier screen overnight and specific activity was measured using the STORM 840 phosphoanalyzer (Molecular Dynamics, Piscataway, NJ).

Genomic DNA Isolation and Southern Blotting

Genomic DNA from each clone was accomplished using the Wizard genomic DNA extraction kit according to the manufacturer's protocols (Promega Corporation, Madison, WI). After isolation each preparation was quantified using a Dyna Quant 200 fluorometer (Hoefer Pharmacia Biotech Incorporated, San Francisco, CA). In two separate reaction tubes restriction digestions were set up with 2.5μg of DNA each using a *Bam*HI restriction enzyme according to the manufacturer's instructions (Promega Corporation, Madison, WI). Digestions were carried out in a 37°C water bath for four hours. After digestion the products were loaded and run on a 1% agarose gel at 30V for 6 hours. The gel was then stained with ethidium bromide and photographed before the transfer. The gel was then soaked for 15 min in a depurination solution (250mM HCl) and 30 min in a denaturation solution (0.5M NaOH and 1M NaCL), rinsed with dH2O and then neutralized two times for 15 min in (0.5M Tris/ 1.5M NaCl) before a final equalization in 20X SSC. The DNA was then transferred to Biotrans^{TM} nylon membrane (ICN, Irvine, CA) using the Turbo blotter apparatus according to the manufacturer's instructions (Schleicher and Schuell, Keene, NH).

Double stranded DNA probes were generated complementary to a 300 bp portion of the codon optimized and wild type *luxA* genes using standard PCR protocols with the incorporation of a [^{32}P] labeled dCTP nucleotide. The probe was purified by column purification according to the manufacturer's instructions (Stratagene, La Jolla, CA). The specific activity of the each probe was measured by scintillation counting (Beckman Coulter, Fullerton, CA). Double stranded probes were boiled for 10 min to denature the DNA and directly added in equal amounts of specific activity to each blot. The blot was incubated with the probe at 65°C overnight. After probe hybridization, the blot was washed 4 times in 20X SSC to remove any unbound activity. The wash temperatures were determined experimentally to achieve optimal probe binding without excess background activity. The blot was air dried and then placed on a phosphorescence intensifier screen (Molecular Dynamics, Piscataway, NJ). Specific activity was measured using the STORM 840 phosphoanalyzer and the data analyzed using the ImageQuant data analysis software package (Molecular Dynamics, Piscataway, NJ).

RNA Isolation and Blotting

At passage six, post transfection, selected cell line clones were expanded to 75cm^2 tissue culture flasks. When the cells became 80-95% confluent, they were trypsinized to remove the cells from the surface and transferred to 2.0 ml Sarstedt tubes (Fisher Scientific, Pittsburgh, PA). Cells were spun down and washed two times in sterile PBS (Sigma Aldrich, St. Louis, MO). Total RNA was then isolated from the cells using the RNeasy kit (Quiagen, Valencia, CA) according to the manufacturer's

instructions for isolation of total RNA from mammalian cells. To remove any contaminating DNA, the RNA was digested for 30 min with DNaseI (Promega Corporation, Madison, WI). To remove the DNaseI enzyme, the clean-up procedure from the RNeasy kit was used (Quiagen, Valencia, CA). Total RNA was then quantified using the Beckman DU-640 spectrophotometer absorbance at 260/280 (Beckman Coulter, Fullerton, CA).

Northern Blotting

Ten micrograms of total RNA were loaded onto a 0.8% agarose formaldehyde gel and run at 100V for 2 hrs. The gel was then stained in an ethidium bromide solution and visualized. The RNA was then transferred to a Biotrans^{TM} nylon membrane (ICN, Irvine, CA) using a semi-dry electroblot transfer apparatus according to the manufacturer's instructions (CBS Scientific, San Francisco, CA).

A 26 base pair oligonucleotide was designed to specifically hybridize to the codon optimized and wild type *luxA* sequences. This oligonucleotide was then 3′ end labeled with a γ [^{32}P] dATP by T4 polynucleotide kinase according to the manufacturer's protocol (Promega Corporation, Madison, WI). The oligonucleotide probe was then purified by column purification as outlined by the manufacturer (Stratagene, La Jolla, CA). The specific activity of the probe was measured by scintillation counting (Beckman Coulter, Fullerton, CA) and added directly to the blot.

Double stranded DNA probes were generated complementary to a 300 bp portion of the codon optimized and wild type *luxA* genes using standard PCR protocols with the incorporation of a [^{32}P] labeled dCTP nucleotide. The probe was purified by column purification according to the manufacturer's instructions (Stratagene, La Jolla,

CA). The specific activity of the each probe was measured by scintillation counting (Beckman Coulter, Fullerton, CA). Double stranded probes were boiled for 10 min to denature the DNA and directly added in equal amounts of specific activity to each blot.

The blot was incubated with the probe at 50°C overnight. After probe hybridization, the blot was washed 4 times in 20X SSC to remove any unbound activity. The wash temperatures were determined experimentally to achieve optimal probe binding without excess background activity. The blot was air dried and then placed on a phosphorescence intensifier screen (Molecular Dynamics, Piscataway, NJ). Specific activity was measured using the STORM 840 phosphoanalyzer and the data analyzed using the ImageQuant data analysis software (Molecular Dynamics, Piscataway, NJ).

Protein Isolation and Western Blotting

To extract the proteins, cells were trypsinized from a plate or flask surface and resuspended into 2.0 ml Sarstedt tubes (Fisher Scientific, Pittsburgh, PA). The cells were then spun down and washed two times in sterile phosphate buffered saline (PBS) to remove any residual media (Sigma Aldrich, St. Louis, MO). Cell pellets were resuspended into 1 ml 0.1M potassium phosphate buffer pH 7.8 and disrupted by three consecutive cycles of freeze (30 s liquid N_2) thaw (5 min at 37°C) extraction. After disruption, the cell debris was pelleted by spinning the samples at 14,000Xg for 5 min and the supernatant was used as total soluble protein for Western blot analysis.

Protein concentrations were determined using the Coomassie Plus protein assay according to the manufacturer's instructions (Pierce, Rockford, IL). Equal amounts (100 – 250 μg) of protein were loaded onto a 12% SDS-PAGE gel. Minigels were run

at 30 mA for approximately 2 h and larger slab gels were run at 30 mA overnight. The proteins were then electroblot transferred to PDVF membrane (Biorad, Hercules, CA) using a semi-dry electroblotter according to the manufacturer's instructions (CBS Scientific Company, Incorporated, Del Mar, CA). Blots were then blocked overnight in 5% nonfat dry milk and hybridized with a polyclonal antibody raised against a 16 amino acid *luxA* polypeptide (′N′ - FDDSDQTRGYDFNKGC - ′C′) or a 16 amino acid *luxB* polypeptide (′N′ - CMILVNYNEDSNKAKQ - ′C′) (Genemed Synthesis, Incorporated, San Francisco, CA). Antibodies were diluted in T-TBS (Tris Buffered Saline + 3% Tween 20) at a 1:500 dilution and applied to the membrane at room temperature for 5 h to overnight. The blot was then washed several times in T-TBS and incubated with a Goat Anti-Rabbit second antibody that has been conjugated to alkaline phosphatase. The blot was then developed according to the manufacturer's protocol (Biorad, Hercules, CA).

Statistics

Statistical analysis of the data presented here was conducted using either the JMP (SAS Institute, Incorporated, Pacific Grove, CA) or Microsoft Excel (Microsoft, Seattle, WA) statistical software packages. Graphs were made using Sigma Plot software (SPSS, SAS Institute, Incorporated, Pacific Grove, CA) or Microsoft Excel (Microsoft, Seattle, WA). All error bars on graphs indicate one standard deviation of the mean. Significant differences were determined using either t-test or 1 way ANOVA analysis at a level of α=0.05.

Results

Determining a Codon Optimized Sequence of *P. luminescens luxA* and *luxB* for Expression in Mammalian Cells

The ratio of codons in the wild type *luxA* and *luxB* nucleotide sequences was compared to codon usage patterns of highly expressed (top 10%) mammalian genes according to the Genbank sequence database. It was determined that the codon usage patterns between *P. luminescens* and human genes were extremely different. Therefore, to create an optimized version of the *lux* genes, the codon ratios were altered to more closely follow codon usage patterns within the human genome. Higher frequency codons were used more often while rare codons were eliminated from the sequence entirely. Changes were made within the nucleotide sequence in a random fashion. This codon optimized sequence was further analyzed for potential regions that may act as target splice sites or other regulatory signals. The sequence was then modified until all potential splice sites and the more obvious regulatory sequences were removed. A comparison of the final codon optimized and wild type *lux* sequences was made. Once the codon optimized sequence was finalized it was tested using the GENSCAN online algorithm that predicts protein expression levels of gene sequences in human cells by comparing the sequence to known highly expressed genes within the matrix specified (http://genes.mit.edu). The results of this analysis were encouraging and a predicted a significant increase in expression on both transcriptional and translational levels. Further, although verification was not possible, GENSCAN predicted a cleavage of the first twenty amino acids of the wild type LuxA protein when expressed in mammalian

cells (Table 11). This cleavage was eliminated in the codon optimized sequence and a full length product was predicted to form. A seqence alignment of the wild type and codon optimized genes is shown in Figures 24 and 25. The wild type and codon optimized *luxA* and *luxB* ratios for codon usage is shown in Table 8A and 8B.

Construction of the Codon Optimized *luxA* and *luxB* Genes

To evaluate the potential impact of codon optimization on the expression of the bacterial luciferase genes in mammalian cells, codon optimized versions of each gene were synthesized *in vitro*. To generate functional genes, single stranded oligonucleotides (80-106 bp) were designed that spanned the entire gene sequence with overlapping (18-23 bp) regions. Four oligonucleotides were placed into a single PCR reaction to amplify segments of the genes individually (Figure 22). The two outside oligonucleotides were used as both template and primers for the amplification reaction and the internal oligos as template. Resultant PCR products of the appropriate size were placed into a second PCR reaction and the fragments were then amplified to link the pieces together using the two outermost oligonucleotides as primers (Figure 22). Products of the correct size were again purified and TA TOPO cloned to generate pPA2 and pPB2. Complete sequence analysis was performed and revealed several introduced errors that were subsequently corrected by site directed mutagenesis.

Table 11. GENSCAN transcription and translation prediction scores for expression of the *luxA* and *luxB* in a human host. (http://genes.mit.edu)

Gene	Type	Begin	End	Length	I	T	CodRg	P	Trans.
luxA(wt)	1	61	1083	1023	45	42	791	0.7	67.01
luxA(op)	1	1	1083	1083	66	42	1910	0.88	181.78
luxB(wt)	1	1	984	984	51	38	585	0.97	46.37
luxB(op)	1	1	984	984	66	41	1952	0.99	185.60

I = initiation signal T = termination signal CodRg = Coding Region score
P = probability of an exon Trans. = exon score
*Score interpretation: 0-50 = weak 50-100 = moderate >100 = strong

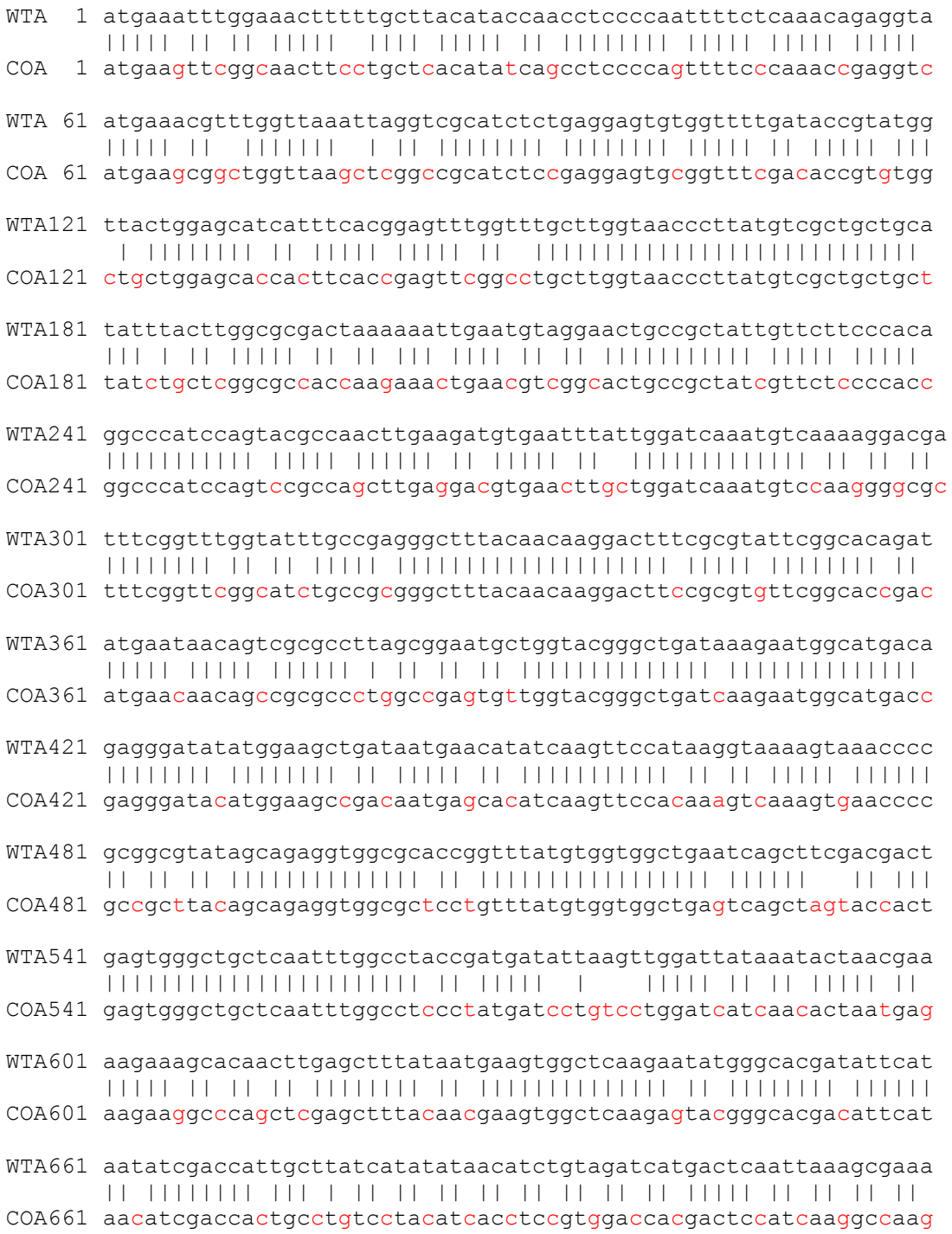


Figure 24. Wild type and codon optimized *luxA* sequence alignment.

```
WTA721 gagatttgccggaaatttctggggcattggtatgattcttatgtgaatgctacgactatt
       |||||||||||||| ||||| ||||||||||||||   || |||||||||| |||||
COA721 gagatttgccggaagtttctcgggcattggtatgatagctacgtgaatgctaccactatc

WTA781 tttgatgattcagaccaaacaagaggttatgatttcaataaagggcagtggcgtgacttt
       ||||| || || ||||| || |||||||| || ||||| || |||||||||| || ||
COA781 tttgacgactccgaccagaccagaggttacgacttcaacaaggggcagtggcgcgatttc

WTA841 gtattaaaaggacataaagatactaatcgccgtattgattacagttacgaaatcaatccc
       || || |||||||| || |||||||  | || || || ||||| ||||| |||||||||
COA841 gtgttgaaaggacacaaggatactaacagacgcatcgactacagctacgagatcaatccc

WTA901 gtgggaacgccgcaggaatgtattgacataattcaaaaagacattgatgctacaggaata
       ||||| || || ||||| || |||||||| || ||||| |||||||||||||| |||||
COA901 gtgggcaccctcaggagtgcattgacatcatccaaaaggacattgatgctaccggaatc

WTA961 tcaaatatttgttgtggatttgaagctaatggaacagtagacgaaattattgcttccatg
       || || || |||||||||||||| ||||| ||||| || ||||| || || |||||||||
COA961 tccaacatctgttgtggatttgaggctaacggaaccgtggacgagatcatcgcttccatg

WTA1021 aagctcttccagtctgatgtcatgccatttcttaaagaaaaacaacgttcgctattatat
        |||||||||||||| |||||||||||||| || || || || |||||    ||  | ||
COA1021 aagctcttccagtccgatgtcatgccattcctcaaggagaagcaacgcagcctcctgtac

WTA1081 tag
        |||
COA1081 tag
```

Figure 24. Continued

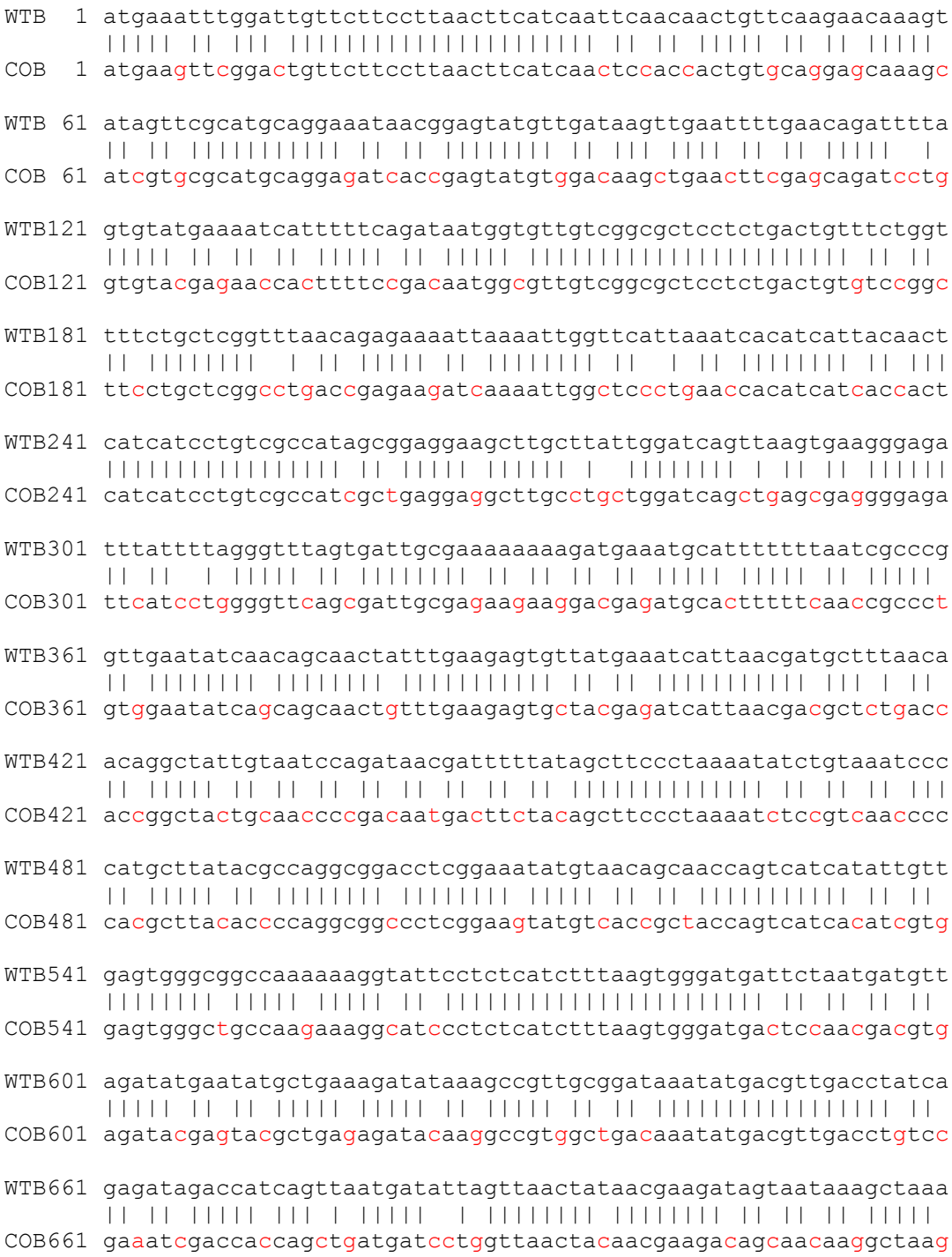


Figure 25. Wild type and codon optimized *luxB* sequence alignment.

```
WTB721 caagagacgcgtgcatttattagtgattatgttcttgaaatgcaccctaatgaaaatttc
       || ||||| || || || ||||| || || || |||||||||||||||| || || |||
COB721 caggagacccgcgccttcattagcgactacgtgcttgaaatgcaccctaacgagaacttc

WTB781 gaaaataaacttgaagaaataattgcagaaaacgctgtcggaaattatacggagtgtata
       || || || ||||| ||||| || || || ||||||||||||| || || ||||||||
COB781 gagaacaagcttgaggaaatcatcgccgagaacgctgtcggaaactacaccgagtgtatc

WTB841 actgcggctaagttggcaattgaaaagtgtggtgcgaaaagtgtattgctgtcctttgaa
       ||||| |||||| |||| || || ||||| ||||| || |||||  ||||||||||||
COB841 actgctgctaagctggccatcgagaagtgcggtgctaagagtgtcctgctgtcctttgag

WTB901 ccaatgaatgatttgatgagccaaaaaaatgtaatcaatattgttgatgataatattaag
       |||||||||||  |||||||||||| || || ||||| ||||| || || |||||||||
COB901 ccaatgaatgacctgatgagccaaaagaacgtcatcaacattgtggacgacaatattaag

WTB961 aagtaccacatggaatatacctaa
       |||||||||||||| || ||||||
COB961 aagtaccacatggagtacacctaa
```

Figure 25. Continued

In Vitro Transcription and Translation of the Wild Type and Codon Optimized *luxA*

To quickly evaluate the translation efficiency in a mammalian cell system of the codon optimized and wild type *luxA* genes, *in vitro* transcription and translation analysis was performed. The pIRES expression vector contains a bacteriophage T7 promoter region upstream of the MCS (A). This promoter was used to generate runoff transcripts of the wild type and codon optimized *luxA* sequences. The transcript was then translated *in vitro* using rabbit a reticulocyte lysate system that incorporates a ^{35}S methoinine into the polypeptide sequence and allows for easy detection. The codon optimized LuxA protein (COA) was determined to be produced by this system approximately twenty fold over the wild type LuxA protein (Figure 26).

In Vivo Expression of the Wild Type Versus Codon Optimized *luxA* and *luxB* Genes

To evaluate the optimized genes *in vivo*, wild type and codon optimized versions of the *luxA* and *luxB* were cloned into the pIRES mammalian expression vector to allow for bicistronic expression of both genes with only one selection marker. Twenty stable clones (HEK293 cells) were selected for each construct along with one negative vector control. At passage three post transfection, each clone was tested *in vitro* for bioluminescence upon the addition of n-decanal and $FMNH_2$. These data revealed that each clonal cell line varied in its bioluminescence levels (Figure 27). The average bioluminescence from each gene combination is shown in Figure 28. Based on these

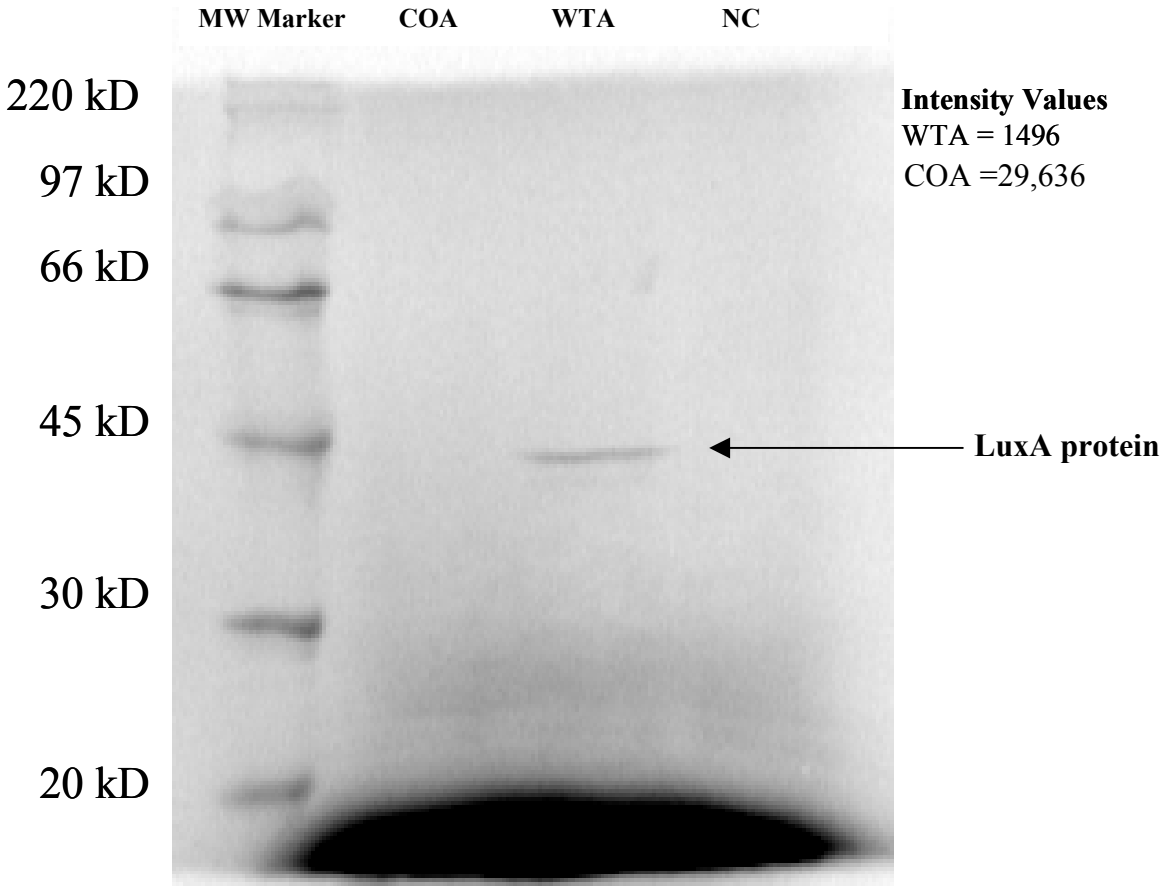
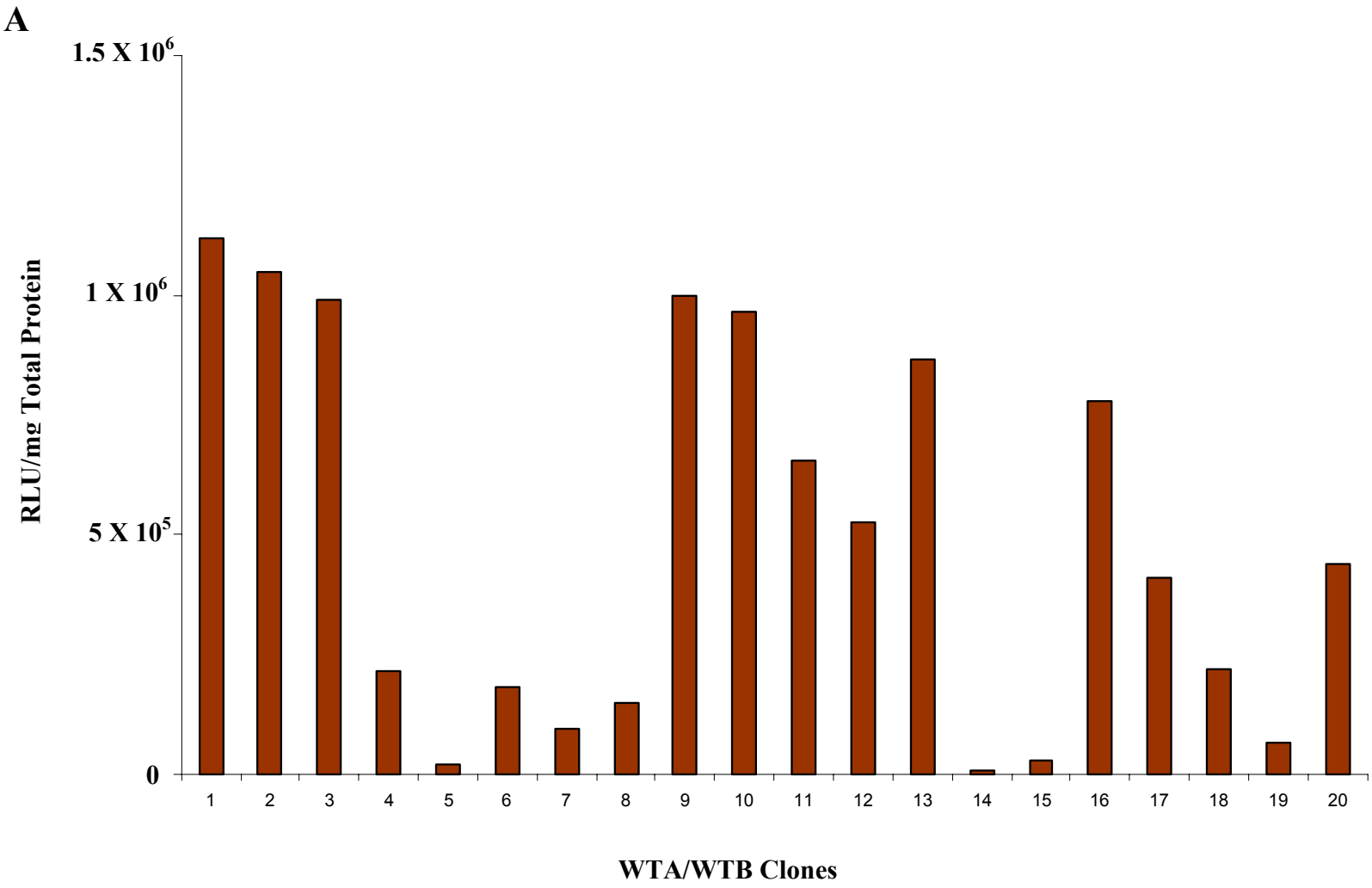


Figure 26. *In vitro* translation products of the wild type *luxA* and codon optimized *luxA* genes. Products were labeled by the incorporation of [^{35}S] methionine.

Figure 27. Bioluminescence measurements taken at passage three post transfection for the twenty clones for each construct. A. WTA/WTB clones B. COA/WTB clones C. COA/COB Clones.


A
1.5 X 10^6
1 X 10^6
5 X 10^5
0
RLU/mg Total Protein
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
WTA/WTB Clones

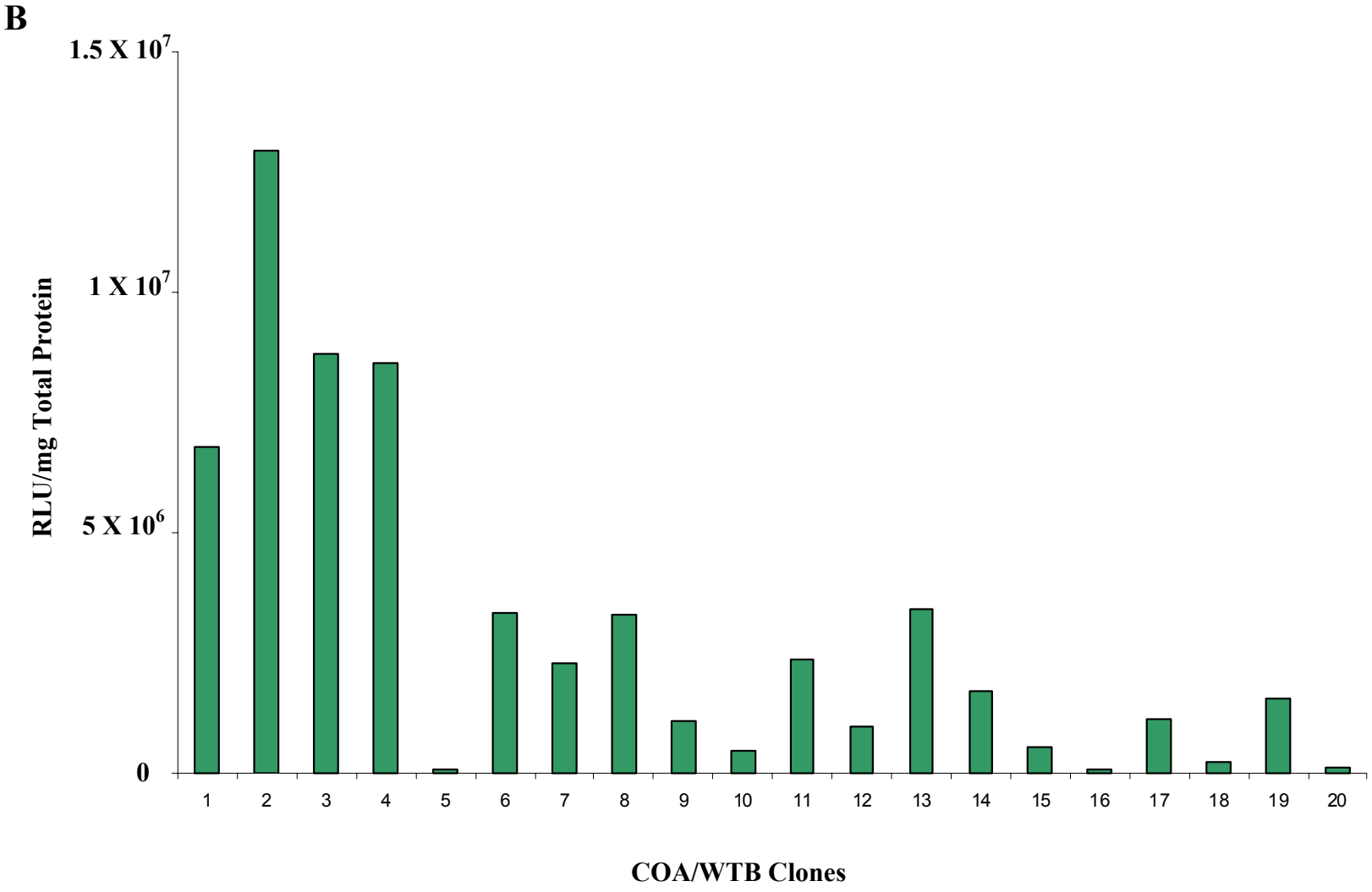


Figure 27. Continued

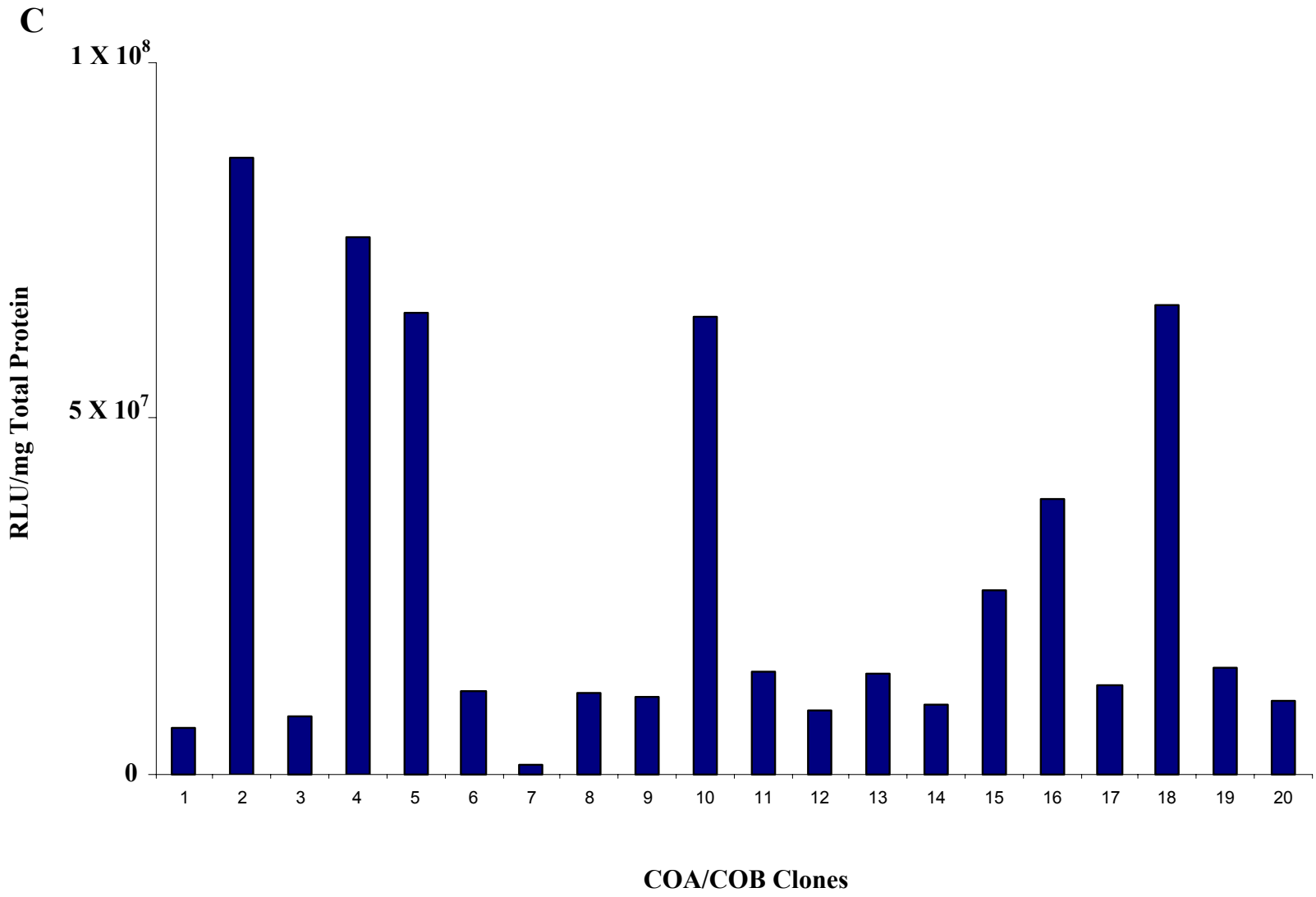


Figure 27. Continued

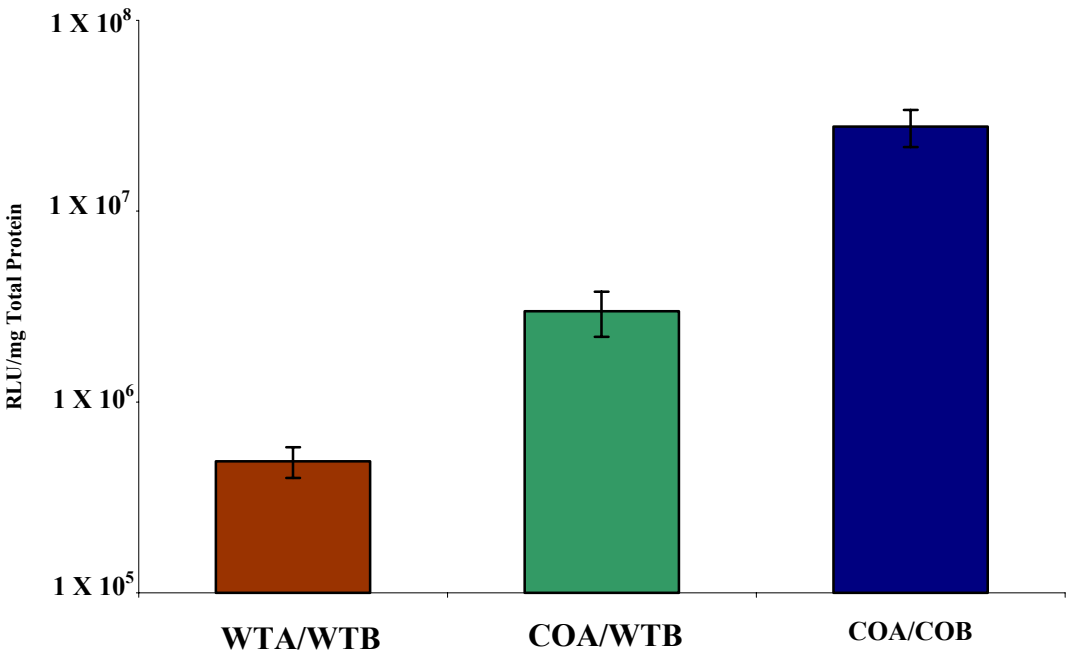


Figure 28. Average bioluminescence from stably transfected HEK293 cell lines. (20 clones tested for each clone type in triplicate).

data, the two or three clones producing the highest bioluminescence levels were chosen for further study. At passage six, each clone selected was expanded into triplicate 75cm^2 tissue culture flasks. From these cells, total genomic DNA, total RNA and soluble proteins were extracted for further analysis.

Determining Insertion Number in HEK293 Clones

To determine gene insertion number in each clone, a southern blot was performed using *luxA* probes generated to both the wild type and codon optimized *luxA* sequences. As shown in Figure 29, all of the cell lines tested had either one or two copies of the gene inserted with the exception of the COA/COB3 clone. To simplify further measurements, this clone was then disregarded for further bioluminescence comparisons.

Determination of *luxA* Message Levels in HEK293 Clones

To determine transcript levels, total RNA was extracted and northern blot analysis was performed. The same probes that were used for Southern blot analysis were used in these experiments as well. Transcript levels were determined to be approximately equal with the exception of the WTA/WTB1 clone that had a lower amount of *luxA* transcript (Figure 30). The vector (NC) control had little to no background hybridization (Figure 30). The ethidium bromide stained 28S was included as an RNA loading reference.

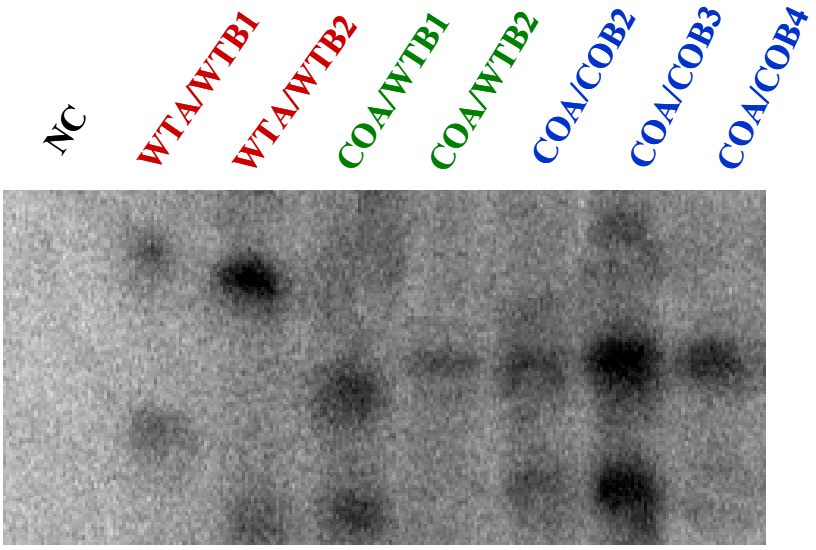


Figure 29. Southern blot analysis on the stable HEK293 clones harboring either wild type *luxA* and *luxB*, codon optimized *luxA* and wild type *luxB* or codon optimized *luxA* and *luxB*. The blot was probed with a 300 bp [^{32}P] labeled probe of both the wild type and codon optimized *luxA* sequence.

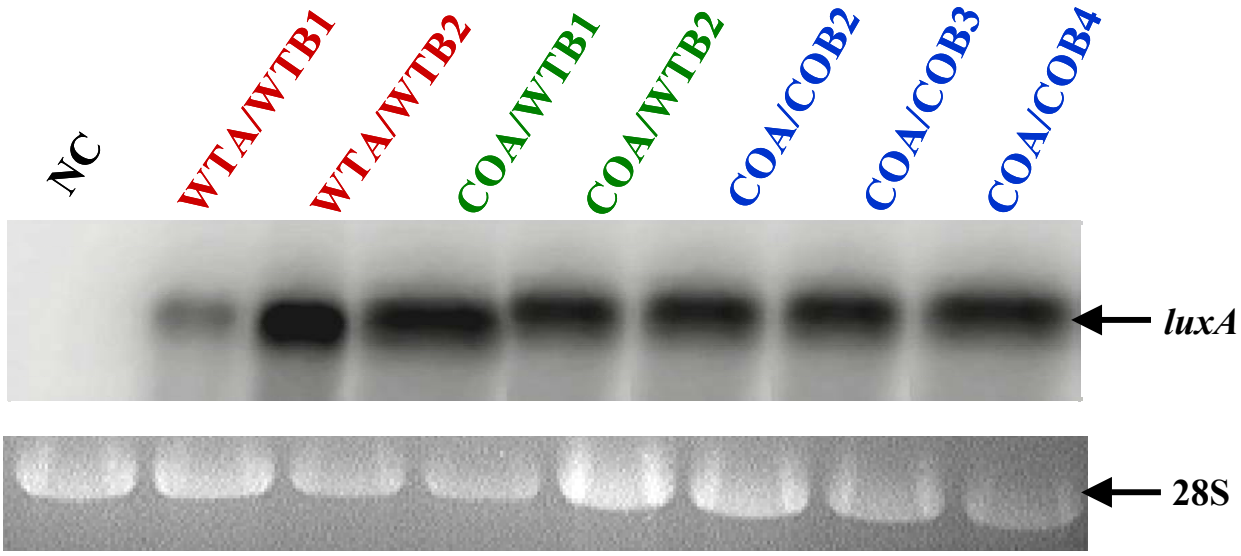


Figure 30. Northern blot analysis of the stable HEK293 clones harboring either wild type *luxA* and *luxB*, codon optimized *luxA* and wild type *luxB* or codon optimized *luxA* and *luxB*. Ethidium bromide stained 28S rRNA was used to ensure RNA quality and loading controls.

Determination of LuxA Protein Levels in HEK293 Clones

Total soluble proteins from each clone were isolated by a series of freeze (liquid N_2) thaw (37°C) cycles. Two hundred fifty micrograms of total protein were run on an SDS-PAGE gel and Western blot analysis was performed using a polyclonal *luxA* antibody (Figure 31). LuxA protein was not detected in any of the wild type *luxA* and *luxB* clones, only detected at very low levels in codon optimized *luxA* with wild type *luxB* clones, but readily detectable when both genes were codon optimized (Figure 31). This increase in LuxA protein concentration was observed despite the fact that the levels of *luxA* mRNA transcript were relatively equivalent for all of the clones tested (Figure 32).

Bioluminescence Levels from Wild Type Versus Codon Optimized Luciferase Genes

Bioluminescence levels were evaluated on whole cell extracts upon the addition of n-decanal and $FMNH_2$. Each clone was tested in triplicate from individual $35cm^2$ wells. Bioluminescence values were found to be greater than two orders of magnitude higher in cell lines harboring both a codon optimized *luxA* and *luxB* (COA/ COB) over that of the cell lines harboring the wild type genes (WTA/WTB) (Figure 33). The bioluminescence levels obtained increased in the order WTA/WTB < COA/WTB < COA/COB. Based on these data it was determined that codon optimization had a significant effect ($p<0.05$) on the bioluminescence potential from HEK293 cells.

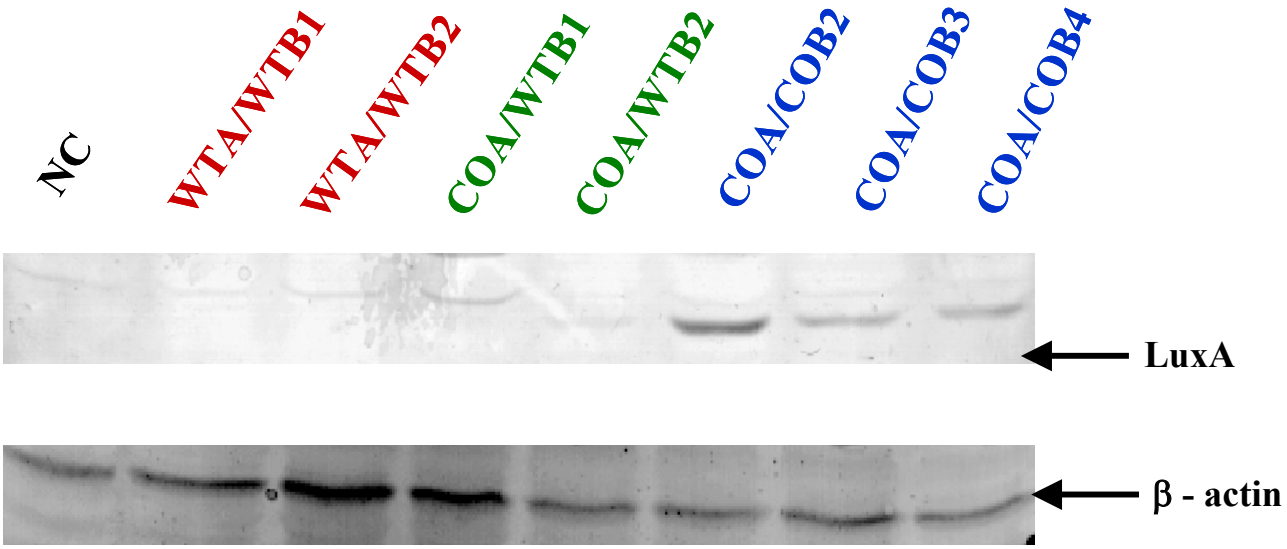


Figure 31. Western blot analysis of HEK293 clones harboring either wild type *luxA* and *luxB*, codon optimized *luxA* and wild type *luxB* or codon optimized *luxA* and *luxB*. The β-actin protein was used as a loading control.

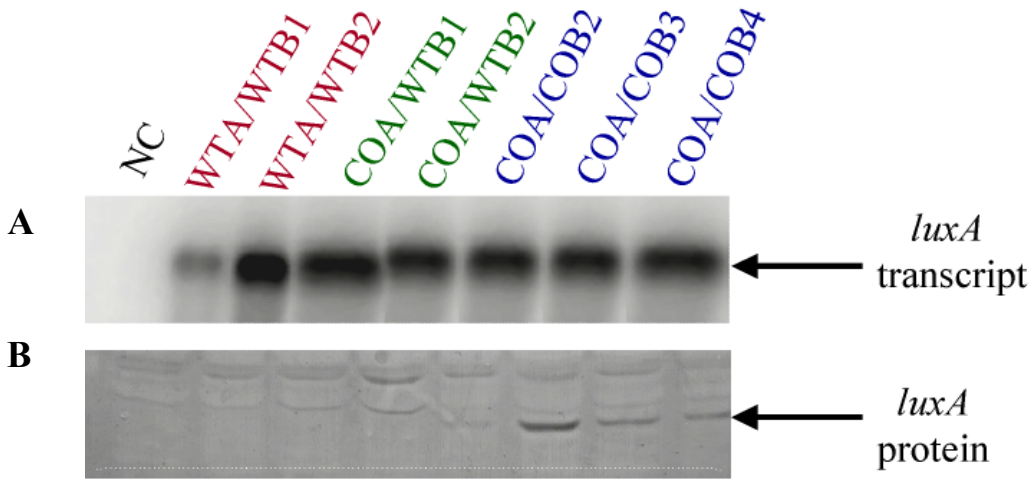


Figure 32. Comparison of mRNA levels and protein levels in each of the stable HEK293 cell line clones. A. Northern blot of total RNA (20µg) from stably transfected HEK293 cells probed with ^{32}P labeled complimentary *luxA* probes. B. Western blot of total soluble protein (250 µg) from stably transfected HEK293 cells immunoblotted with a polyclonal *luxA* antibody.

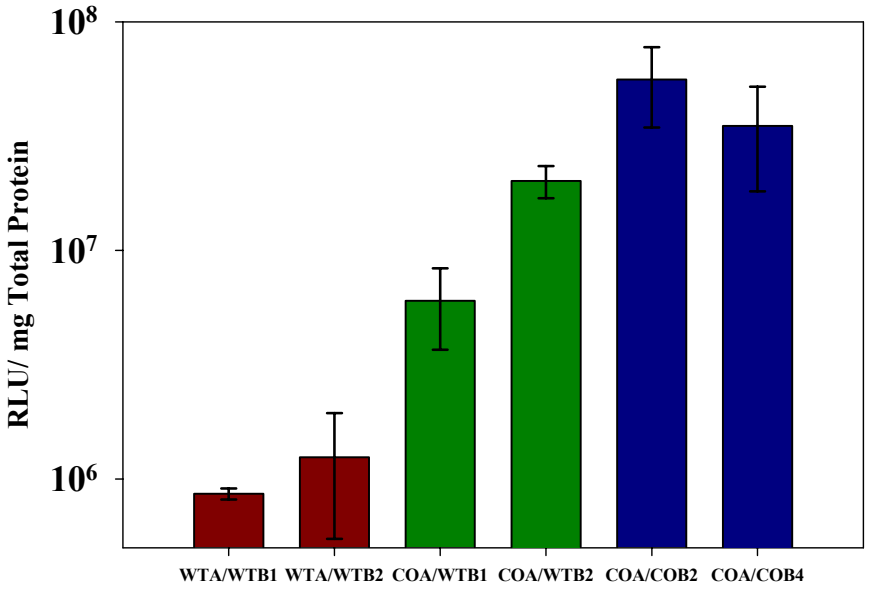


Figure 33. Average bioluminescence from individual HEK293 clones stably transfected with WTA/IRES/WTB, COA/IRES/WTB or COA/IRES/COB.

Discussion

Codon usage regulates gene expression at the level of translation and the usage patterns between species are not conserved (Kurland, 1991). This is especially true between genes derived from eukaryotes versus those from prokaryotes. Therefore, to efficiently express the bacterial *lux* genes in mammalian cells, the nucleotide sequence was altered in such a way as to create a "humanized" form of the gene without altering the amino acid sequence. This approach has been used previously to optimize the expression of both GFP and *Renilla* luciferase proteins for expression in mammalian cells (Zhang et al., 2002 and Gruber and Wood, 2000). The design of this new sequence was carefully determined, removing all potential splice sites and most regulatory regions. After the final codon optimized sequence was determined, it was evaluated using the GENSCAN prediction algorithm to determine the potential expression efficiency in a human cell. According to the output from this program the overall expression of the codon optimized *lux* genes would be significantly improved versus the wild type. The increase in expression was predicted to be caused by an increase in both transcription and translation efficiency. Furthermore, it was predicted that the first sixty bases (20 amino acids) of the wild type *luxA* gene would be completely eliminated when expressed in mammalian cells. Considering that this region of the LuxA protein holds most of the catalytic properties (active site) for the bacterial luciferase enzyme, this would be devastating for its expression. If this were the case, the low expression levels observed for the LuxAB fusion protein, shown earlier, may be better explained in part by a nonfunctional protein being formed rather than inefficient folding or heat liability.

To test the expression of the codon optimized genes, modified versions were required. However, because the necessary changes were to numerous to achieve by site directed mutagenesis, a complete *in vitro* gene synthesis protocol was pursued. Large oligonucleotides (80-106 bp) were designed with overlapping (18-22 bp) regions. The original plan was to amplify all of the oligonucleotides together in one PCR reaction according to methods set by Prodromou and Pearl (1992). However, because of the larger size of the *lux* genes (approximately 1000 bp each), this was not possible. As an alternative, the gene was synthesized in parts and the subsequently linked by a second round of PCR. The two outside oligonucleotides were used as both template and primers for each reaction. After some experimental effort, it was determined that for optimal amplification the internal oligonucleotides (template) should be added at a concentration that equaled 100 fold less than the outside oligos. Amplification products of the correct size were cloned and sequenced. Unfortunately, sequence analysis revealed several base substitution mutations within all clones tested. These mutations were present despite the fact that care was taken by using a *Pfu* polymerase that has proof reading abilities. This finding was disappointing while not surprising given that two consecutive PCR reactions were required to obtain the final gene product resulting in > 60 cycles of amplification. To eliminate these errors and produce the proper sequence, site directed mutagenesis was performed.

It was determined previously through work accomplished in *S. cerevisiae* and mammalian cells for the expression of the bacterial luciferase genes that IRES elements may be an efficient way to express independent proteins as single bicistronic transcripts. This expression format provides the most natural expression of the genes, most closely

mimicking the polycistronic form found in the bacterial operon. Therefore, experiments were set up to compare the codon optimized and wild type *luxA* and *luxB* genes in mammalian cells using an IRES based expression vector. The expression vector used was designed to highly express two independent genes under the control of a single constitutive encephalomyocarditis virus (ECMV) promoter region by linking two multicloning sites fused to either side of an internal ribosomal entry site (IRES). The IRES element allows for the translation of two consecutive open reading frames from one messenger RNA (Jang et al., 1990; Jackson et al., 1990; Rees et al., 1996). By constructing plasmids with different combinations of the codon optimized *luxA* and *luxB* with their wild type counterparts, a direct comparison of the genes was made.

To quickly determine if a difference in translation efficiency could be detected between the optimized and wild type *luxA* genes, *in vitro* transcription and translation analysis was performed. The codon optimized *luxA* gene (COA) was detected approximately twenty fold over wild type (Figure 26). This finding supported the results that shown earlier with the *in vitro* generation of the wild type LuxA protein. Since the rabbit reticulocyte lysate translation system is used to mimic mammalian translation machinery *in vitro*, these results indicated that the codon optimization would indeed make a significant impact on the translation efficiency of the *lux* proteins in mammalian systems.

HEK293 cells were transfected with the WTA/ WTB, WTA/ COB or COA/COB constructs and stable cell line clones were selected by antibiotic resistance. Twenty stable clones for each *luxA* and *luxB* combination were selected and bioluminescence levels were determined upon the exogenous addition of n-decanal and $FMNH_2$. The

bioluminescence significantly increased in the order WTA/WTB < COA/WTB < COA/COB. These data indicated that codon optimization had made a significant impact on the potential bioluminescence levels obtained from mammalian cells. To analyze this data further, the two or three brightest clones were chosen for further study. From these cells, total genomic DNA, total RNA and total soluble proteins were extracted.

Foreign gene integration in mammalian cells is a random event, therefore it is possible to have more than one insertion of the construct occur during each transfection. Since integration is fairly inefficient, the copy number per cell is generally very low. However, because of this possibility, it was important to determine the copy number of the inserted genes for at true comparison. To accomplish this, Southern blot analysis was performed on each of the seven clones that produced the highest bioluminescence levels. All of the cell lines tested had either one or two copies of the gene inserted with the exception of the COA/COB3 clone which had three. To simplify further measurements, this clone was then disregarded for further bioluminescence comparisons. Nevertheless, it should be noted that increased copy number does not correlate with increased expression levels. Numerous other factors have been shown to impact from gene expression at a greater level.

The overall amount of *luxA* mRNA transcript was determined by Northern blot analysis. Transcript levels were relatively equal with the exception of the WTA/WTB1 clone that produced significantly lower amounts of *luxA* mRNA. Since each construct contained the same promoter (ECMV) element and initiation signals, it would be expected that each clone would have approximately equal amounts of transcript for the

introduced genes. However, *cis* acting regulatory elements could potentially interfere with transcription initiation and overall transcript levels *in vivo*. These types of interferences would vary based on where within the chromosome the genes were integrated. Therefore, the position effect of various clones could explain the lower amount of transcript detected with the WTA/WTB1 clone. Other factors that can potentially impact the amount of RNA transcript would be a direct result of increased RNA degradation of certain mRNA sequences that can occur. This type of RNA instability would be less likely after codon optimization because of the removal of several AU rich target degradation regions. However, because the lower amount of transcript was not seen in both the WTA/WTB clones, this scenario is unlikely.

To evaluate the overall protein concentrations and determine translation efficiency of each construct, Western blot analysis was performed. Total soluble proteins from each clone were isolated and quantified. Western blot analysis was performed using a polyclonal *luxA* antibody (Figure 31). The LuxA protein was not detectable from WTA/WTB clones and faintly visible in the COA/WTB clones. However, large amounts of LuxA protein were detected from the COA/COB clones which harbored a construct carrying codon optimized versions of both genes. This finding was intriguing and unexpected. Since the only available antibody was raised against a polypeptide of LuxA, it was expected that the constructs harboring the COA/WTB and COA/COB would produce equal amounts of LuxA protein. Since this was not the case, these data indicated that the codon optimization of both genes might infer stability on the heterodimeric complex that makes up the luciferase enzyme. This

increased stability of the complex may have allowed the proteins to be detected in the Western blot while the other construct was not detected.

Perhaps the most valuable measurement to determine if codon optimization was a success is the amount of enzyme activity that could be obtained from each construct. Bioluminescence levels were evaluated on whole cell extracts upon the addition of n-decanal and $FMNH_2$. Average bioluminescence values were found to be greater than two orders of magnitude higher in cell lines harboring both a codon optimized *luxA* and *luxB* (COA/COB) over that of the cell lines harboring the wild type genes (WTA/WTB) (Figure 33). While bioluminescence levels were significantly higher in clones expressing COA and WTB versus WTA and WTB, the optimal bioluminescence was obtained from clones harboring optimized versions of both genes (Figure 33). These data further support the stabilization conclusion of the heterodimeric protein. Based on these data it was determined that codon optimization had a significant effect on the protein expression in HEK293 cells.

In conclusion, the codon optimization of the *luxA* and *luxB* genes was successful in increasing the overall expression levels of the individual proteins. This increase in protein quantities resulted in a significant increase in bioluminescence from cell lines harboring these constructs. Furthermore, the bioluminescence levels from codon optimized *luxA* and *luxB* provide adequate bioluminescence for the proof in principle data needed for the future development of reliable reporter constructs for analyte sensing in mammalian cells.

CHAPTER 4

EVALUATION OF MAMMALIAN CELLS FOR $FMNH_2$ AVAILABILITY AND STABILITY OF BACTERIAL LUCIFERASE IN STABLE MAMMALIAN CELL LINES

Introduction

The bacterial luciferase enzyme system has several advantages over other bioreporter systems available. Of these advantages, the ability to directly measure *in vivo* gene expression without the disruption of the reporter cell or loss of cell viability is perhaps the most obvious benefit. This property has made the *lux* system sought out by many for various research applications using prokaryotic cells. As a result, numerous biosensor systems utilizing the bacterial luciferase system have been developed and are currently in use. However, as discussed earlier, the bacterial luciferase system has not yet been efficiently expressed in mammalian cell lines and therefore the full potential of this technology is not yet realized. The obstacles encountered by researchers trying to employ the bacterial *lux* system in eukaryotes have included low expression levels of the Lux proteins and limited amounts of substrates and cofactors required for the reaction within the cells.

The bacterial luciferase enzyme system consists of a multi-enzyme complex encoded by five genes that provide the luciferase enzyme as well as the luciferin (aldehyde) substrate for the reaction. However, this reaction additionally requires the host cell metabolism to provide adequate amounts of molecular O_2, the reducing power of $FMNH_2$ and the energy of ATP as co-factors and substrates. To produce bioluminescence, the reaction goes through several intermediate steps. With a

stoichiometry of 1:1:1, the α and β subunits bind with a reduced flavin molecule ($FMNH_2$) to form a C4a hydroxyflavin. As this hydroxyflavin becomes dehydrated to FMN, a blue-green light is emitted (Baldwin et al., 1979). Because of the stoichiometry of the reaction and its absolute requirement, the $FMNH_2$ molecule is considered to be an additional substrate for the bioluminescence reaction rather than a co-factor. On the other hand, the luciferase reaction itself can proceed in the absence of the decal-aldehyde substrate, but its presence significantly increases the light output kinetics and therefore is required for bioreporter applications (Volkova et al., 1999).

In bacteria, the availability of $FMNH_2$ is not a limiting factor for the bioluminescence reaction. Nevertheless, in several bioluminescent strains of bacteria, a flavin oxidoreductase gene (*frp*) has been identified in close proximity to the *lux* operon. This enzyme has the ability to reduce pools of FMN within the cell to $FMNH_2$ and then recycle itself to catalyze further reductions. Although the availability of $FMNH_2$ in yeast was shown to be a limiting substrate for the bioluminescence reaction in these cells, it was subsequently shown that yeast could be further engineered to express the flavin oxidoreductase gene (*frp*) from *Vibrio harveyi* to overcome this limitation (Gupta et al., 2003). This was the first illustration of the use of a flavin reductase enzyme to improve the bioluminescence output from an engineered *lux* bioreporter cell. Although the availability of the $FMNH_2$ substrate has not been completely evaluated in mammalian cells, it has been hypothesized to be one of the major problems leading to inefficient expression of the bacterial luciferase system in eukaryotes. Furthermore, to this point, all mammalian cell line experiments have been conducted with the exogenous addition of $FMNH_2$ or the enzymes required to produce

this substrate. In order to efficiently express the bacterial luciferase system in mammalian cells, the availability of $FMNH_2$ needs to be fully evaluated and the possibilities available to overcome this limitation need to be explored.

For the future generation of a useful mammalian bioreporter cell line for gene expression analysis or target analyte monitoring, the engineered cell line not only needs to be able to efficiently express the *lux* genes but, it must also remain stable for long periods of time in the absence of antibiotic selection. Some of the proposed applications for this technology may require that the cells remain stable for extended periods with very little to no intervention. Therefore, the overall stability of engineered mammalian cell lines harboring the *lux* genes needs to be evaluated to completely understand the limitations of this technology. Therefore, in this research effort, the following objectives will be met:

- Determine to what extent $FMNH_2$ limits the bioluminescence reaction in mammalian cell lines expressing the bacterial luciferase enzyme.
- Obtain a bioluminescent cell line that can overexpress the flavin oxidoreductase enzyme from *Vibrio harveyi*.
- Evaluate the ability of this engineered cell line to produce adequate levels of $FMNH_2$ for the bioluminescence reaction.
- Evaluate the overall stability of the constructs within stably integrated mammalian cell lines engineered with the bacterial luciferase genes for long term maintenance without antibiotic pressure.

Materials and Methods

Cell Culture and Plasmid Maintenance

All relevant constructs and strains, bacterial and mammalian, used in this study are outlined in Table 12. *E. coli* cells were routinely grown in Luria Bertani (LB) (Fisher Scientific, Pittsburgh, PA) broth containing the appropriate antibiotic selection with continuous shaking (200 rpm) at 37°C. Kanamycin and Ampicillin were used at a final concentration of 50μg/ml and 100μg/ml, respectively.

All cell culture reagents and media were obtained from Sigma Aldrich (St. Louis, MO) unless otherwise stated. Mammalian cells were grown in the appropriate complete growth media containing 10% heat-inactivated horse serum, 0.01mM non-essential amino acids and a Dubelco's minimal media base (DMEM) (M4655). Cells were routinely grown at 37° C in a 5% CO_2 atmosphere to confluency and split every three to four days by trypsinization at 1:4 ratio and transferred into fresh complete growth media. Appropriate concentrations of antibiotic were used to maintain constructs after transfection according to susceptibility kill curve analysis. Kill curves were completed for each lot of antibiotic. The range of typical concentrations used for the selection of HEK293 cell line clones was between 450μg/ml and 650μg/ml of Neomycin G418 and 250μg/ml and 400μg /ml of Zeocin (Invitrogen Corporation, Carlsbad, CA).

Table 12. Constructs and Strains used in this study.

Plasmid/Strain Designation	Relevant Genotype/ Characteristics	Source
Strains		
E. coli		
DH5α	Φ80d*lacZ*ΔM15, *recA*1, *endA*1, *gyrA*96, *thi*-1, *hsdR*17 (r_K-, m_K+), *supE*44, *relA*1, *deoR*, Δ(*lacZYA-argF*)U169	Gibco, BRL
TOP 10	F-, mcrA Δ(mrr-hsdRMS-mcrBC) Φ80/lacZ ΔlacX74 deoR recA1 araD139 Δ(ara-leu) 7697 ga/K rpsL endA1 nupG	Invitrogen
Mammalian Cells		
HEK293	Permanent line of primary human embryonal kidney transformed by sheared human adenovirus type 5 (Ad 5) DNA. ATCC# CRL-1573	ATCC
Plasmids		
pCR4-TOPO	TOPO TA cloning vector for easy cloning of PCR products generated with 3′ A overhangs designed for sequencing Km^r, Amp^r	Invitrogen
pIRES	Mammalian expression vector containing the internal ribosomal entry site (IRES) of the encephalomyocarditis virus between two multi-cloning sites which allows for the expression of two genes under the control of a single constitutive CMV promoter, Neomycin G418 antibiotic selection marker and a pUC ori and Km^r for replication in *E.coli*	Clontech
pcDNAHISMAX	Mammalian expression vector allows for constitutive expression under the control of a constitutive CMV promoter contains a SPC163 translational enhancer, Zeocin antibiotic selection marker and pUC ori and Amp^r for replication in *E. coli*.	Invitrogen
pcDNA3.1Zeo	Mammalian expression vector allows for constitutive expression under the control of a constitutive CMV promoter, Zeocin antibiotic selection marker and pUC ori and Amp^r for replication in *E. coli*.	Invitrogen
pCR4frp	pCR4 TA cloning vector harboring the *frp* from *Vibrio harveyi.*	This Study

Table 12. Continued

Plasmid/Strain Designation	Relevant Genotype/ Characteristics	Source
pMaxfrp	pcDNAHISMAX vector harboring the frp gene from *V. harveyi.*	This Study
pcfrpZeo	pcDNA vector harboring the *frp* gene from *V. harveyi.*	This Study
Mammalian Cell Line Constructs		
COA/COB(2)	HEK293 cell lines stably transfected with the pCOA-I-COB plasmid and selected by G418.	Chapter 3
frp (1-9)	COA/COB2 cell line stably transfected with the pcDNAfrp plasmid and selected by G418 and Zeocin.	This Study

Construction of a Mammalian Expression System for *frp*

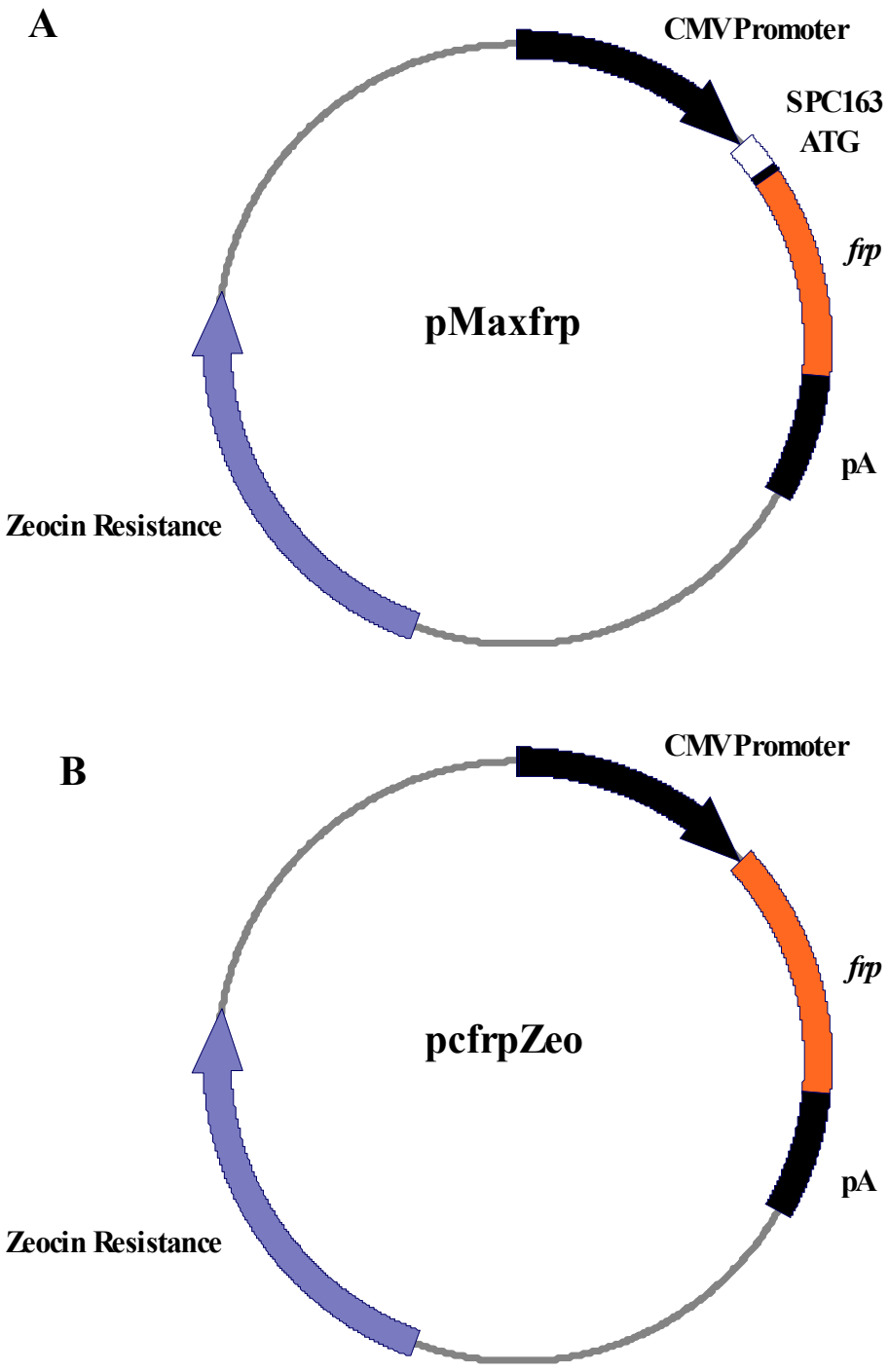
To generate a strain to overexpress the flavin oxidoreductase enzyme in mammalian cells, the *frp* gene was amplified from *V. harveyi* strain VHU08996 DNA. The gene was then TA TOPO cloned into the pCR4-TOPO cloning vector according to the manufacturer's instructions to generate pCR4frp (Invitrogen Corporation, Carlsbad, CA) and subsequently cut and ligated into the pcDNAHISMAX mammalian expression vector using introduced unique 5`*Bam*HI and 3` *Not*I restriction sites to generate pMaxfrp (Figure 34A) (Invitrogen Corporation, Carlsbad, CA). This expression vector possesses an SPC163 untranslated sequence upstream of the gene insert. This sequence has been shown to enhance translation between four and five fold over expression without the enhancer.

A second plasmid was generated to express the *frp* gene from *V. harveyi* by cloning the gene via the introduced unique 5`*Bam*HI and 3`*Not*I restriction sites into the pcDNA3.1Zeo mammalian expression vector to generate pcfrpZeo (Figure 34B).

Ligation Reactions

Plasmid vectors and inserts were digested (2-6 h) with the appropriate enzymes (Promega Corporation, Madison, WI). Linearized vectors were dephosphorylated using a calf alkaline phosphatase enzyme according to the manufacturer's instructions (Promega Corporation, Madison, WI). Both vector and insert DNA were gel purified from 1% agarose gels using the Geneclean gel extraction kit (Bio101, Carlsbad, CA). The recovered DNA was then quantified using the Dyna Quant 200 fluorometer (Hoefer Pharmacia Biotech Incorporated, San Francisco, CA) and ligations were set

Figure 34. Schematic diagram of the expression vectors used to express the flavin oxidoreductase enzyme (*frp*) from *V. harveyi* in mammalian cells. A. Expression in the pcDNAHISMAX that provides an SPC163 translational enhancer region for enhanced translation and protein expression driven from a constitutive CMV promoter region. B. Expression in the pcDNA3.1Zeo provides high constitutive expression from a CMV promoter.


A
CMV Promoter
SPC163
ATG
frp
pA
pMaxfrp
Zeocin Resistance
B
CMV Promoter
frp
pA
pcfrpZeo
Zeocin Resistance

up as 20µl reactions using a 3:1 molar ratio of insert to vector DNA. The ligation reactions were then incubated at 17ºC overnight.

Electroporation

Electrocompetent cells were prepared as outlined by the manufacturer (BTX, San Diego, CA). Electroporations were performed using the BTX Electroporator 600 with the following conditions: 40µl cells, 1-2µl ligation mixture, a 2.5kV pulse for 4.7 ms using a 2 mm gap cuvette. After the pulse, cells were immediately resuspended in 1 ml of sterile LB and allowed to recover for 1 h at 37ºC (200 rpm). Cells were then plated on selective media containing the appropriate antibiotic.

Selection of Bacterial Clones

Resistant colonies were picked after 24 h and expanded to patches on grid plates. To test for proper insert presence and orientation, rapid boil mini-preps were done followed by the digestion of the plasmid with the appropriate restriction enzymes according to the manufacturer's instructions (Promega Corporation, Madison, WI). Digestions products were run on 1% agarose gels to determine if the banding pattern indicated the insert presence and proper orientation. Upon identification of correct clones, the plasmids were further purified using the Wizard midiprep plasmid purification system according to the manufacturer's protocol (Promega Corporation, Madison, WI) and sequenced.

Sequencing

All constructs were sequenced to ensure their integrity. Sequencing was done in the University of Tennessee Molecular Biology Service Facility using an Applied Biosystems 3100 Genetic Analyzer sequencer (Foster City, CA).

Transfection of Mammalian Cells

Transfection of mammalian cell lines was done in six well poly-D-lysine coated tissue culture plates (Fisher Scientific, Pittsburgh, PA). Cells were split from stock cultures and inoculated into each well at approximately 1 X 10^5 cells per well in complete growth media. The plate was then placed at 37ºC in a 5% CO_2 atmosphere for 1-2 days until the cells became 80-90% confluent. The day of transfection, the media was refreshed. DNA for transfections was purified from 100ml overnight *E. coli* cultures using the Wizard Purefection plasmid purification kit to remove endotoxins according to the manufacturer's instructions (Promega Corporation, Madison, WI). For chromosomal integration, the plasmid DNA was linearized before transfection to increase proper integration.

HEK293 Cells

Purified plasmid DNA (3.2µg) was mixed into 200µl of serum free DMEM in a 1.5 ml tube. In a second tube, 8µl of Lipofectamine 2000 reagent (Invitrogen Corporation, Carlsbad, CA) was added to a 200µl of serum free DMEM. The lipofectamine mixture was added to the DNA mixture within 5 min and incubated at room temperature for 20 min. The entire mixture (400µl total) was added directly to the appropriate well on the plate and rocked back and forth to ensure adequate mixing.

Twenty-four hours post transfection, the complexes were removed and the media was replaced with fresh complete growth media supplemented with the appropriate antibiotic for selection of the two co-transfected plasmids.

Selection of Mammalian Cell Clones

Twenty-four hours post transfection, selective media was added to all wells and refreshed every three to four days. Because these transfections were conducted on a cell line that already harbored a Neomycin G418 resistance plasmid, the G418 was added at a concentration to maintain the plasmid and Zeocin was added to select for the second plasmid. When the COA/COB2 clone was co-transfected with the pMaxfrp, resistant clones never appeared within the transfected wells. Therefore, the pcfrp clone was generated to determine if the overexpression of the gene was causing a lethal product for the cells. Within two weeks after co-transfection with this plasmid construct, all control wells were dead and the transfected cells were forming small colonies on the plate surface. Colonies were separated from the rest of the well by placing a sterile chamber around the cell mass and sealing it with silicon (Fisher Scientific, Pittsburgh, PA). The media could then be removed and each colony could be trypsinized and transferred to individual tissue culture flasks. To accomplish this, after washing with a PBS solution, 200µl of a 1X Trypsin-EDTA solution (Sigma Aldrich, St. Louis, MO) was added directly to the chamber and incubated at 37°C for 3 to 5 min. The trypsin-EDTA solution was then replaced with complete growth media and the cells were transferred to a $25cm^2$ tissue culture flask for propagation. Each clone was given a number and expanded to individual cell lines. Each line was split and

maintained as described earlier with the addition of selective media. Nine cell lines were propagated in this manner.

In Vitro Bioluminescence Assays

To evaluate the bioluminescence potential from each cell line clone, total proteins were extracted and *in vitro* enzyme (bioluminescence) assays were performed. To extract the proteins, the cells were first trypsinized from the plate or flask surface and resuspended into 2.0 ml Sarstedt tubes (Fisher Scientific, Pittsburgh, PA). The cells were then spun down and washed two times in sterile phosphate buffered saline (PBS) to remove any residual media (Sigma Aldrich, St. Louis, MO). Cell pellets were then resuspended into 1 ml 0.1M potassium phosphate buffered pH 7.8 and disrupted by three consecutive cycles of freeze (30 s liquid N2) thaw (5 min at 37ºC) extraction. After disruption, the cell debris was pelleted by spinning the samples at 14,000 X g for 5 min and the supernatant was used in the bioluminescence assay. To determine light intensity, the protein extract was mixed with 0.1mM NAD(P)H, 4μM FMN, 0.2% (w/v) BSA, 0.002% (w/v) n-decanal. Bioluminescence was measured using the FB14 luminometer (Zylux Corporation, Pforzheim, Germany) at a 1 s integration and reported as relative light units (RLU). To evaluate the limitation of $FMNH_2$ for the bioluminescence reaction, a flavin oxidoreductase enzyme (1U) isolated and purified from *V. harveyi* (Roche Scientific, Indianapolis, IN) was added to the mixture and light levels were measured again for comparison.

Bioluminescence levels were determined by taking measurements from triplicate 35 cm^2 wells and recording the values as relative light units (RLU) using a 1 s integration time.

Whole Cell Bioluminescence Assays

To evaluate the bioluminescence levels from intact cells, the adherent cell lines were first trypsinized from the flask or plate surface and resuspended in 2.0 ml Sardstedt tubes (Fisher Scientific, Pittsburgh, PA). The cells were then spun down and washed two times in sterile phosphate buffered saline (PBS) to remove any residual media (Sigma Aldrich, St. Louis, MO). Cell pellets were resuspended into 1 ml 0.1M potassium phosphate (pH 7.8). The cells were then mixed 1:1 volume of the enzyme mix that consisted of: 0.2% BSA and 0.002% (w/v) n-decanal. Bioluminescence was measured immediately using the FB14 luminometer (Zylux Corporation, Pforzheim, Germany) at a 1 s integration and reported as relative light units (RLU).

Bioluminescence levels were normalized between samples and cell lines by dividing the RLU measurement by the number of cells in the assay and reporting the bioluminescence as relative light units (RLU) per 1 X 10^5 cells. The total cell count was determined by direct counting the samples by standard methods. The cells were first mixed with a Trypan Blue stain (Sigma Aldrich, St. Louis, MO) and individual cells were counted under light microscopy with a hemocytometer.

Stability of the Bacterial Luciferase Constructs

In order to determine the overall stability of the HEK293 cell line clones expressing the bacterial luciferase genes, each line was grown in culture for > 40 passages in the presence and absence of antibiotic selection. Cell extract bioluminescence measurements were taken every fifth passage to compare light levels over time.

Statistics

Statistical analysis of the data presented here was conducted using either the JMP (SAS Institute, Incorporated, Pacific Grove, CA) or Microsoft Excel (Microsoft, Seattle, WA) statistical software packages. Graphs were made using Sigma Plot software (SPSS, SAS Institute, Incorporated, Pacific Grove, CA) or Microsoft Excel (Microsoft, Seattle, WA). All error bars on graphs indicate one standard deviation of the mean. Significant differences were determined using either t-test or 1 way ANOVA analysis at a level of α=0.05. Statistically different groups were indicated on graphs by letter.

Results

Evaluation of $FMNH_2$ Bioavailability in Mammalian Cells

To determine the overall bioavailability of the $FMNH_2$ substrate in mammalian cells, bioluminescence assays were performed and light measurements were taken before and after the addition of a purified flavin oxidoreductase enzyme. This enzyme in the presence of FMN and NAD(P)H reduces the FMN to the required $FMNH_2$ for the

reaction. Bioluminescence levels from each of the cell line clones increased at least an order of magnitude after the addition of $FMNH_2$ (Figure 35, maroon bars). These data illustrated that $FMNH_2$ was extremely limiting for the bioluminescence reaction from these engineered human cell lines. However, in every case, the clones harboring the *luxA* and *luxB* genes alone were able to produce bioluminescence levels above background without the addition of the flavin oxidoreductase enzyme indicating that some $FMNH_2$ was available within the cells for the reaction (Figure 35, blue bars). Nevertheless, to achieve optimal bioluminescence values and generate a useful bioreporter, the lack of available $FMNH_2$ within mammalian cells needs to be addressed.

Expression of the Flavin Oxidoreductase Enzyme

In an attempt to overcome this limitation, the COA/COB2 clone (brightest clone) was co-transfected with an *frp* gene that was isolated from *V. harveyi* and cloned into a mammalian expression vector containing a translational enhancer region upstream of the multi-cloning site. Unfortunately, when COA/COB2 clones were co-transfected with this plasmid, the HEK293 cells were not able to survive and as a result no clones were obtained from this construct. As an alternative approach, the *frp* gene was cloned into and expressed constitutively from the pcDNA3.1Zeo vector that allows for high constitutive expression but does not contain the SPC163 enhancer region. Nine stable cell line clones were obtained by resistance to toxic concentrations of both Neomycin G418 and Zeocin antibiotics simultaneously. Resultant clones were expanded to individual cell lines and tested for bioluminescence potential.

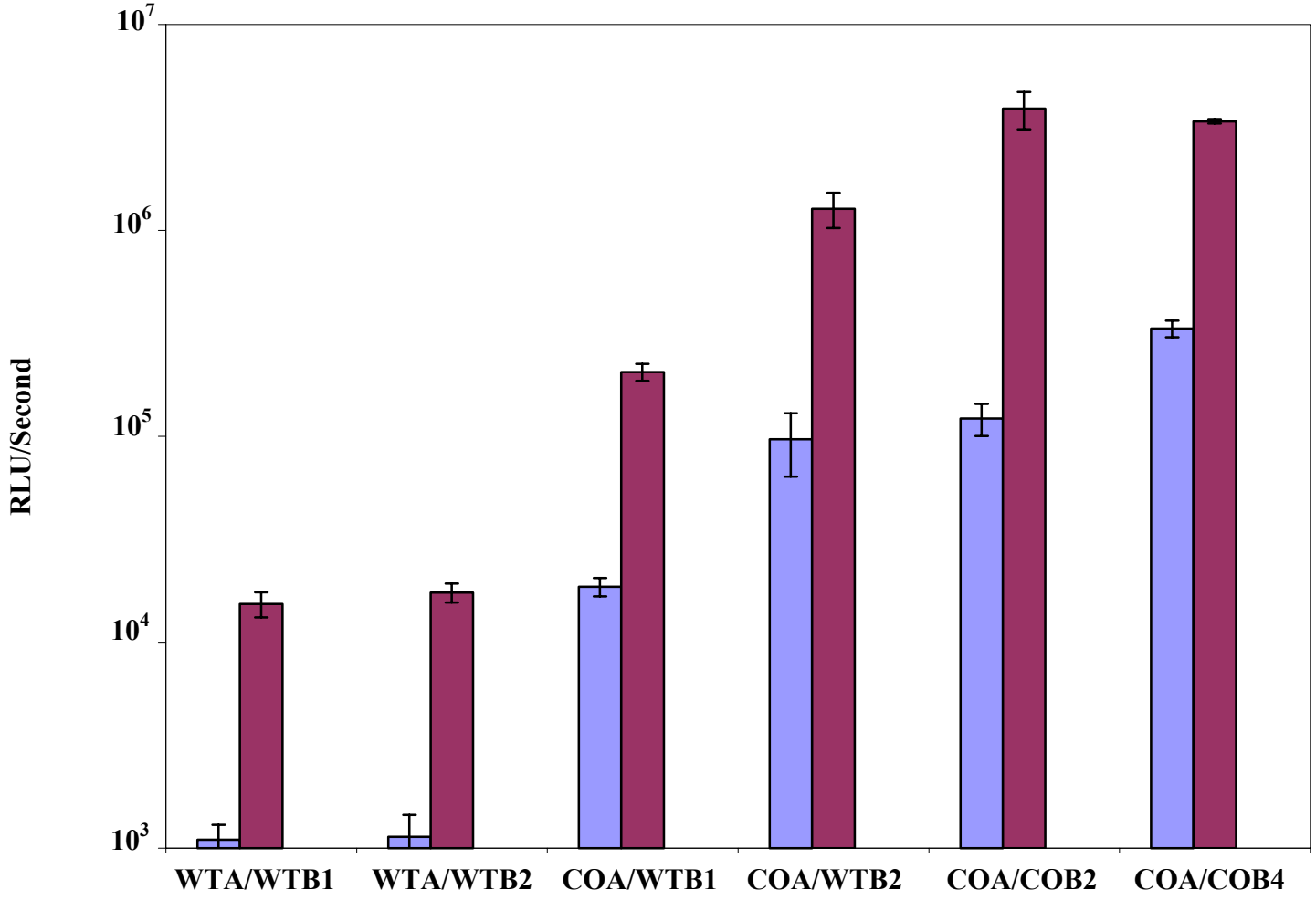


Figure 35. Bioluminescence levels (RLU) from stable HEK293 clones before (blue bars) and after (maroon bars) the addition of $FMNH_2$.

Bioluminescence levels were evaluated from total cell protein extracts and in whole cell bioluminescence assays.

In Vitro Bioluminescence Assays

From *in vitro* bioluminescence assays (total protein extracts), the overall light levels increased with the expression of the *frp* gene at least an order of magnitude in both the absence and after the addition of exogenous flavin oxidoreductase versus the COA/COB2 clone without the *frp* gene tested under the same conditions (Figure 36). These data indicated that the expression of the *frp* gene was successful in producing an excess of available $FMNH_2$ within HEK293 cells. The further increase in bioluminescence after the exogenous addition of the purified oxidoreductase enzyme however, indicates that the system has yet to reach saturation.

The bioluminescence levels obtained from the cell extract, *in vitro*, assays remained stable for several minutes before gradually declining to background levels. The light intensity could be increased back to peak levels upon exogenous addition of additional NAD(P)H to provide the reducing power for the flavin oxidoreductase enzyme and generate more $FMNH_2$. Thus the luciferase complex itself remained stable throughout the assay and bioluminescence levels were correlated to availability or decay of reduced FMN.

Whole Cell Bioluminescence Assays

Whole cell bioluminescence assays were performed to determine if these cell lines could produce adequate bioluminescence levels for use in gene expression

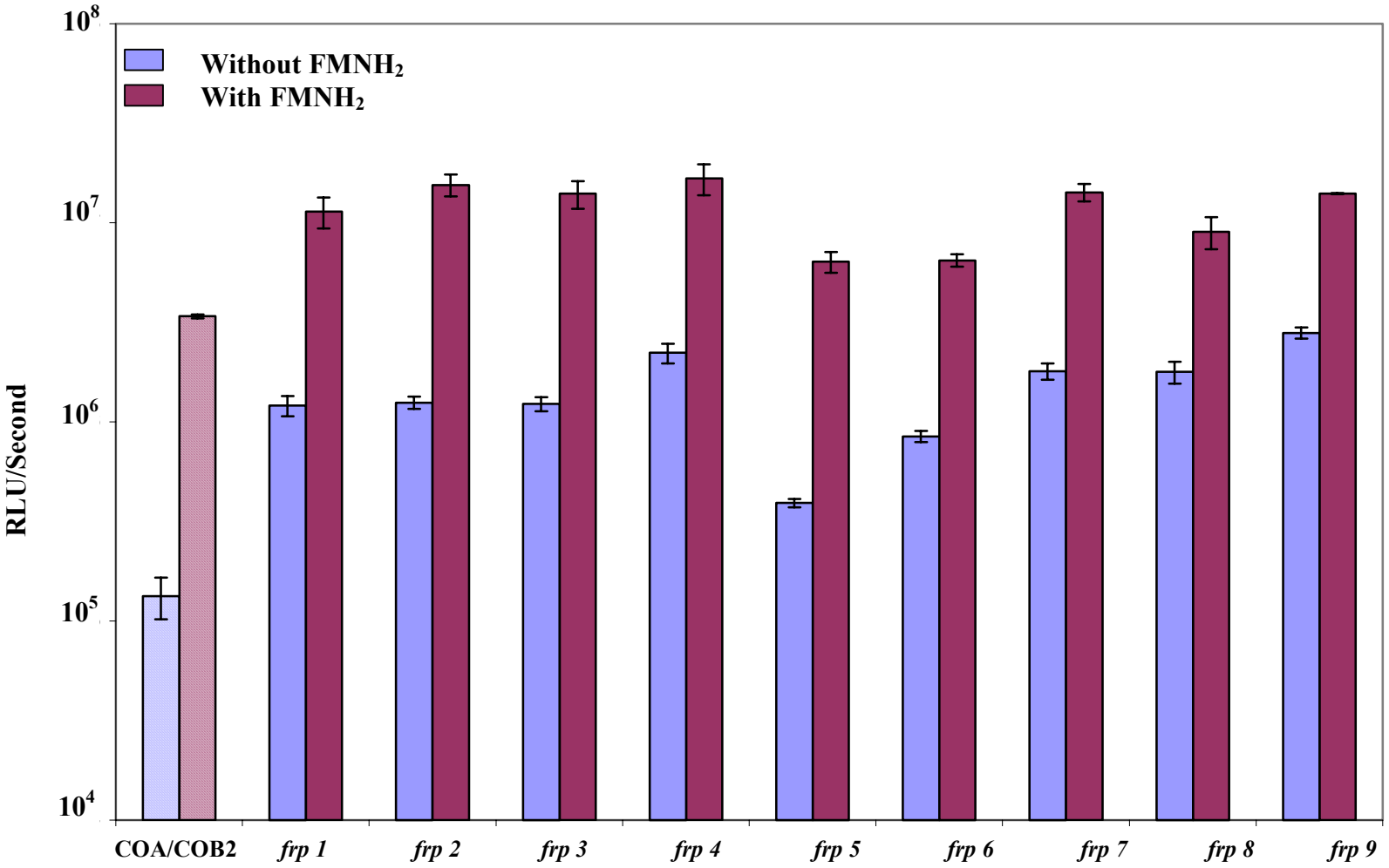


Figure 36. Bioluminescence levels from clone COA/COB2 versus the COA/COB2 clone co-expressed with a *V. harveyi* flavin oxidoreductase enzyme. Bioluminescence measurements were taken before the addition of $FMNH_2$ (blue bars) and after (maroon bars).

analysis, much in the same way that firefly luciferase (Luc) is currently used today in several reporter applications. Average bioluminescence levels from the COA/COB2 clone were obtained that were at least two orders of magnitude greater than background levels (4×10^4 RLU/s versus 380 RLU/s). The bioluminescence was further increased at least another order of magnitude when the *frp* gene was co-expressed along with the luciferase genes. All clones co-transfected to express the *frp* gene produced significantly more light than without the enzyme being expressed ($p<0.05$). Furthermore, there were significant differences between the nine *frp* clones as well (Figure 37). However, unlike the relatively stable nature of the bioluminescence signal from *in vitro* bioluminescence assays, the light levels from these whole cell clones resulted in a flash bioluminescent response. The maximum light output was obtained within 1 s of n-decanal addition and returned to background levels within five seconds. These levels could not be induced with the further addition of n-decanal or $FMNH_2$ to achieve a second peak in bioluminescent activity.

In Vitro versus *In Vivo* Light Measurements

To better determine cytoplasmic concentrations of $FMNH_2$, whole cell bioluminescence measurements were compared to levels obtained from *in vitro* assays. In the whole cell assays, the light levels were obtained immediately upon the addition of n-decanal and recorded as relative light units (RLU) for 1×10^6 cells. The cells were then lysed and bioluminescence was remeasured upon the addition of n-decanal and $FMNH_2$. The cell number was determined by direct counting. Overall, the light levels

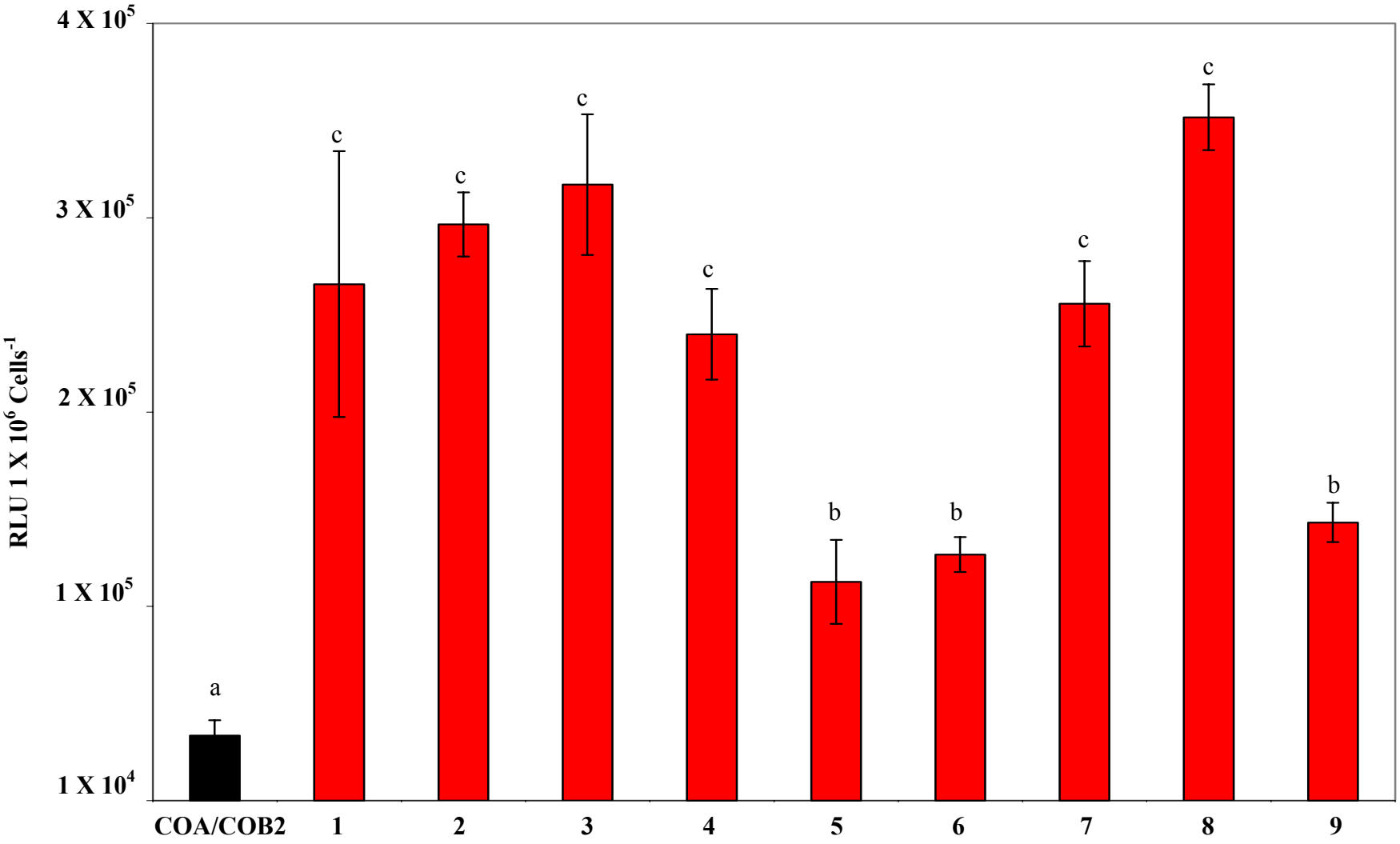


Figure 37. Bioluminescence levels from HEK293 cells upon the addition of 0.002% n-decanal. Bioluminescence measurements were taken from triplicate 35cm^2 wells.

from whole cells were much lower (as much as 20 fold) than those obtained from protein extracts (Table 13).

Stability of Bacterial Luciferase in Mammalian Cells Over Long Periods of Time

The stability of mammalian cell lines engineered to stably express the bacterial luciferase genes was monitored by performing bioluminescence assays over time. The bioluminescence levels remained relatively constant for forty passages, for every clone except WTA/WTB2 where the level radically deteriorated after passage thirty (Figure 38). Although the light levels for the other clones remained relatively stable during this time, other phenotypic changes occurred within the cells, including a lower binding affinity to the flask surface.

Discussion

The expression of the bacterial luciferase enzyme system in eukaryotic cells has long been desired. Unfortunately, several obstacles have been encountered that resulted in only modest success when trying to employ this technology in eukaryotes. Among the problems associated with the ultimate development of this technology, the lack of available $FMNH_2$, the reduced flavin molecule that is required for the bioluminescence reaction is perhaps one of the most obvious. The $FMNH_2$ limitation in mammalian cells engineered with the *luxA* and *luxB* genes was shown to hamper potential light outputs significantly (Figure 35). These data clearly showed that light levels were significantly enhanced upon the exogenous addition of a flavin reductase enzyme.

Table 13. Bioluminescence levels from HEK293 cells engineered to express *luxA* and *luxB* genes with and without the co-expression of the flavin oxidoreductase (*frp*) enzyme. Each bioluminescence measurement was made in triplicate from approximately 1 X 10^6 cells and reported as relative light units (RLU).

Clone	Whole Cell Assay	*In Vitro* (without $FMNH_2$)	*In Vitro* (with $FMNH_2$)
COA/COB2	4.3 X 10^4 (±7,913)	2.5 X 10^5 (±31,201)	6.77 X 10^6 (±170,098)
frp1	2.75 X 10^5 (±68,373)	2.42 X 10^6 (±161,278)	22.8 X 10^6 (±4,078,998)
frp2	3.06 X 10^5 (±16,526)	2.50 X 10^6 (±98,006)	31.0 X 10^6 (±3,909,260)
frp3	3.27 X 10^5 (±36,193)	2.46 X 10^6 (±101,456)	28.0 X 10^6 (±4,433,714)
frp4	2.5 X 10^5 (±23,355)	4.46 X 10^6 (±457,293)	33.4 X 10^6 (±5,975,079)
frp5	1.22 X 10^5 (±21,595)	7.86 X 10^5 (±36,862)	12.74 X 10^6 (±1,518,130)
frp6	1.37 X 10^5 (±9,004)	1.74 X 10^6 (±112,367)	12.96 X 10^6 (±912,574)
frp7	2.66 X 10^5 (±21,971)	3.65 X 10^6 (±346,997)	28.4 X 10^6 (±2,864,018)
frp8	3.62 X 10^5 (±16,907)	3.47 X 10^6 (±440,620)	18.02 X 10^6 (±2,286,558)
frp9	1.54 X 10^5 (±10,151)	5.64 X 10^6 (±381,295)	28.56 X 10^6 (±197,408)

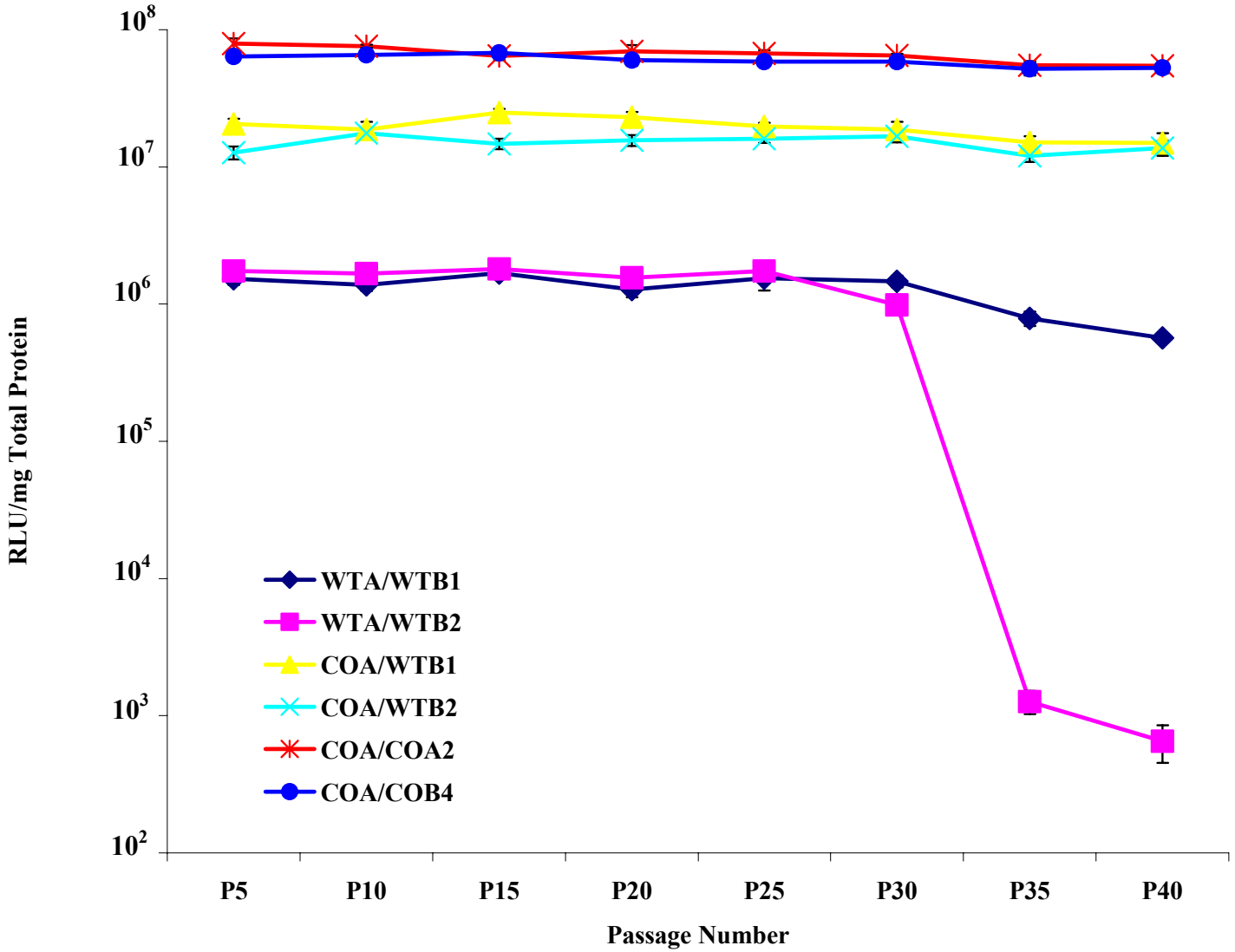


Figure 38. Average bioluminescence levels of HEK293 clones harboring *luxA* and *luxB* cultured without antibiotic versus passage number. Bioluminescence values are the average of triplicate measurements.

Therefore, efforts were explored to overcome the $FMNH_2$ limitation by attempting to express the reductase enzyme *in vivo* to provide adequate levels of this substrate.

Recently, this approach was taken in yeast cells by the overexpression of an NAD(P)H –dependent FMN reductase enzyme from *V. harveyi* (Gupta et al., 2003 and Szittner et al., 2003). In these studies, the bioluminescence levels from whole yeast cells were significantly increased by the expression of the *frp* gene. These data illustrated that the limiting substrate in yeast was indeed a lack of the reduced $FMNH_2$ and that the FMN and NAD(P)H precursors were available within cytoplasm of yeast cells. Based on these results, it was hypothesized that the overexpression of the *frp* gene in mammalian cells may also provide adequate pools of $FMNH_2$ to catalyze bioluminescence. Therefore, HEK293 cells harboring a codon optimized *luxA* and *luxB* gene pair were co-transfected with the *frp* gene and stable cell lines obtained.

From *in vitro* bioluminescence assays of clones expressing *frp*, the overall light levels increased at least an order of magnitude in both the absence and after the addition of exogenous flavin oxidoreductase versus the COA/COB2 clone without the *frp* gene, tested under the same conditions (Figure 36). These data indicated that the expression of the *frp* gene was successful in producing higher concentrations of $FMNH_2$ within HEK293 cells. However, because the light levels increased further upon the addition of exogenous flavin oxidoreductase, it was determined that $FMNH_2$ remains the limiting factor for the bioluminescence potential from these mammalian cells and has yet to reach saturation.

Furthermore, it is believed that much of the available $FMNH_2$ within mammalian cells would be found at much higher concentrations within the mitochondria of the cell. As a result, the $FMNH_2$ would remain sequestered from the bioluminescence (Lux) proteins within the cytoplasm and not available for the reaction. By monitoring light levels as total protein extracts, the location of available $FMNH_2$ within the cytoplasm is unknown. To better determine cytoplasmic concentrations of $FMNH_2$, whole cell bioluminescence assays were performed. Overall, the light levels from whole cells were much lower (as much as 20 fold) than those obtained from protein extracts (Table 13). These data indicate that the cytoplasmic concentrations of $FMNH_2$ are limiting and that much of the cellular pool of the reduced flavin molecule is compartmentalized within organelles, like the mitochondria. Therefore, future experiments should be conducted to allow for the expression of the Lux proteins in the mitochondria of the cell to allow for their interaction with other necessary substrates for the reaction. Targeting of the Lux proteins can be easily accomplished with the addition of a signaling peptide on the N- terminus of the proteins that will shuttle the enzymes into the mitochondria of the cell.

The bioluminescence levels obtained from the cell extract, *in vitro*, assays remained stable for several minutes before gradually declining to background levels. The light intensity could be increased back to peak levels upon exogenous addition of additional NAD(P)H to provide the reducing power for the flavin oxidoreductase enzyme and generate more $FMNH_2$. Thus, the luciferase complex itself remained stable throughout the assay and the light output was directly correlated to the available $FMNH_2$. On the other hand, the whole cell *in vivo* assays provided a flash

luminescence response that could not be induced further to give a second peak upon the addition of n-decanal or $FMNH_2$. These data indicate that the n-decanal substrate was toxic to the cells. Once the cells were lysed and proteins extracted, high bioluminescence levels could be obtained upon the addition of n-decanal and $FMNH_2$. These data further support the idea that the addition of n-decanal to whole cells resulted in cell mortality and not the loss of stable Lux proteins.

The stability of mammalian cell lines engineered to express the bacterial luciferase genes was monitored to determine if the constructs could maintain their bioluminescence levels over time in the absence of antibiotics. Although the bioluminescence outputs remained relatively constant for > 40 passages, one of the clones (WTA/WTB2) bioluminescence levels radically deteriorated after passage thirty (Figure 38). Furthermore, during this culture period, other phenotypic changes occurred within the cells, including a lower binding affinity to the flask surface. It was further determined that the loss of light from the WTA/WTB2 clone was not a result of the loss of the genes within the cells and perhaps a change in the cell's physiology or some unknown regulatory mechanism that caused the loss of light. Even though the WTA/WTB2 clone resulted in a complete loss of bioluminescence activity, PCR analysis revealed that the *luxA* and *luxB* genes were still present. Nevertheless, the bioluminescence levels did remain stable for more than five months in culture without the need for antibiotic selection, indicating that these cell lines would be stable enough for relatively long term monitoring projects and applications as long as the proper control experiments were also included.

CHAPTER 5

SUMMARY AND CONCLUSIONS

This investigation achieved optimized expression of the bacterial luciferase enzyme in mammalian cells. Previous attempts to express this heterodimeric enzyme complex in mammalian cells have been met with only modest success. In this research effort, several vector formats were evaluated and a novel approach to codon optimize the genes was performed. Additionally, the limited availability of the $FMNH_2$ substrate was evaluated and steps were taken to overcome this limitation. The overall stability of the engineered cells was also assessed to determine the persistence of the reporter for long-term monitoring applications. Based on the findings of this study, the following conclusions were drawn:

- A translational fusion of the two luciferase subunits is not an efficient way to express this enzyme in eukaryotes likely due to thermal instability and the inability of the subunits to properly fold forming an active heterodimer.

- Integrated expression of the luciferase provides a more stable expression format for long-term persistence of the luciferase genes.

- The bacterial luciferase enzyme produced optimal bioluminescence in mammalian cells when the individual genes were expressed as a bicistronic transcript fused with an internal ribosomal entry site (IRES).

- Codon optimization of the *luxA* and *luxB* genes significantly increased the translation efficiency of the proteins *in vitro* and *in vivo* when expressed in HEK293 cells. This increase in translation in turn resulted in significant increases in bioluminescence output from the cells.

- $FMNH_2$ is a limiting substrate for the bioluminescence reaction in mammalian cells.

- The expression of a flavin oxidoreductase gene in HEK293 provides additional $FMNH_2$ for the bioluminescence reaction. However, this substrate remains limiting for the reaction.

- The available $FMNH_2$ within mammalian cells may be compartmentalized in organelles and not readily available to interact with the luciferase enzyme.

- The expression of the bacterial luciferase genes in HEK293 cells remains relatively stable for more than 40 passages (5 months) in culture without the selective pressure of an antibiotic.

Determining an Optimal Expression Format for Expression of the Bacterial Luciferase Genes (*luxA* and *luxB*)

As a first attempt to express the bacterial luciferase enzyme in mammalian cells, a LuxAB fusion protein was designed. Previous work by Escher et al. (1989) showed that the *luxAB* fusion, using genes from *Vibrio harveyi,* was stable at elevated temperatures if initially expressed in *E. coli* at 23°C. However, when the fused protein was grown and expressed at 37°C there was a greater than 99% reduction in light. These data suggest that the fused *luxAB* does not fold properly at elevated temperatures. The luciferase from *P. luminescens* has a higher thermal stability ($t_{1/2}$ >3 h at 45°C) than *V. harveyi* ($t_{1/2}$ 5 min. at 45°C) (Meighen, 1991). Therefore, a translational fusion of the *P. luminescens luxA* and *luxB* genes was generated in this work and evaluated. Although the *luxAB* fusion was functional in *E. coli*, bioluminescence activity was significantly reduced (70%) compared to the wild type unfused genes. In the unfused *luxAB* the α and β subunits are individually translated and are free to fold into their specific conformation (Tu and Mager, 1995). Therefore, the reduction in bioluminescence may be due to steric hindrance involved in the way the subunits form the heterodimer when expressed as a protein monomer.

The addition of a Kozak sequence further reduced the bioluminescence level to approximately 5% of the wild type protein. Nevertheless, the Kazak sequence is a mammalian ribosomal binding site and therefore to evaluate its true effectiveness the construct was tested in HEK293 cells. The light levels were significantly reduced (>90%) compared to HEK293 cells expressing the fusion protein without the Kazak modifications. The addition of a Kozak sequence (G at the +4 position) specifically

changes the second amino acid of *luxA* from Lysine to Glutamic acid. These two amino acids have opposite net charges, which could result in a modification of the protein's secondary structure ultimately altering the protein's function resulting in decreased bioluminescence. Therefore, future studies were conducted in the absence of the Kazak sequence or with the addition of an external Kazak for ribosomal binding.

To determine temperature effect on the folding of the fusion protein, the fused and unfused versions of the *luxAB* were grown at 23˚C, 30˚C and 37˚C. In *E. coli*, there was not a statistical difference ($p = 0.05$) associated with temperature on bioluminescence as seen by Escher et al. (1989). This suggested and led us to believe that the folding problems in the *V. harveyi* LuxAB fusion protein were not present in the *P. luminescens* LuxAB fusion protein. However, when the fusion construct was expressed in the yeast, *S. cerevisiae* the bioluminescence levels significantly decreased as temperature increased to 37˚C. The differences seen in these two systems may be a result of the bacterial system's ability to transcribe the *luxB* independently due to the ribosomal binding site and *luxB* start codon still present in the fusion. When the fusion is expressed in the yeast system, the *luxB* is no longer independently expressed resulting in a true fusion protein that is unable to properly fold at 37˚C. The independent expression of the *luxB* in bacteria may have resulted in the unfused LuxB subunit forming the heterodimeric conformation with the LuxA within the LuxAB fusion resulting in the unaffected bioluminescence observed when the construct was expressed in *E. coli*.

In an attempt to mimic mammalian translation machinery, *in vitr*o transcription and translation of the *luxA, luxB* and *luxAB* fusion were performed in a rabbit

reticulocyte lysate system. Although the 77kDa fusion protein was easily detected, the individual proteins, LuxA and LuxB, were not. This result was unexpected because equal amounts of RNA transcript were added to each translation reaction. These data suggest that the formation of the heterodimeric (α and β subunit) complex may be required for not only efficient bioluminescence activity, but also for the overall stability of the protein.

Although detectable amounts of bioluminescence were obtained from mammalian cell lines harboring the LuxAB fusion protein, these levels were not sufficient for the creation of a reliable biosensor. Therefore, other expression formats were evaluated in an attempt to optimize bioluminescence activity. It was thought that by expressing the *lux* genes separately, the subunits would be able to form a more natural heterodimeric conformation. Human embryonic kidney cells (HEK293) were used for these evaluations. Cells were transfected with a dual promoter vector construct that that was developed to constitutively express each gene from a separate promoter or co-transfected with two plasmids each harboring either the *luxA* or *luxB* gene. Furthermore, to evaluate the differences in protein expression from genes integrated in the host's chromosome versus those constructs maintained as episomal plasmids, each expression format (fusion, dual promoter and co-transfection) was constructed on a plasmid backbone able to replicate episomally in HEK cells. The bioluminescence levels from stable cell lines harboring each expression variation were determined. Although there were slight variations in activity, the differences were not statistically significant ($p=0.05$). The only exception was the reduced bioluminescence activity obtained from cells harboring a dual promoter vector episomally. The low light levels

from these clones were somewhat surprising considering that the average bioluminescence from the fusion protein and from cells co-transfected with two plasmids were slightly higher when the constructs were maintained as episomes. Upon further analysis, it was determined that the mRNA levels from the individual *lux* genes were not equal and therefore in this expression format, one promoter is inducing transcription at a higher rate than the other. This type of promoter occlusion where the transcription of one of the two promoters was significantly dampened has been seen previously (Horlick et al., 2000). The unequal availability of one of the *lux* subunits at a level higher than the other, may prevent the proper formation of the heterodimeric active luciferase protein and may result in inactive homodimer formation.

In order for bacterial *lux*-based mammalian bioreporter to be useful, the constructs need to remain stable in the absence of antibiotic selection for long periods of time. Efficient maintenance and stability of foreign genes requires that the DNA replicate once per cell cycle and be retained (integrated or episomally) in the nucleus. Constructs harboring the *luxA* and *luxB* genes in three individual expression formats were created on both the traditional integration vectors and on Epstein-Barr virus (ori-P) based episomal plasmid vectors. To determine the stability of these constructs in HEK293 cells, the cell line clones were grown for twenty passages in complete growth media without antibiotic. In general, all clones (chromosomal and episomal) were stable for at least five passages after the antibiotic removal. However, the constructs that were maintained as episomes began to lose bioluminescence activity by passage ten with episomal co-transfected cells resulting in the fastest bioluminescence decline rate. Although there was a significant decline in bioluminescence activity from episomally

based constructs over time, the light was not completely lost from any of the cells lines. Therefore, this reduction in bioluminescence may be the result of natural plasmid loss and generation of a plasmid equilibrium within the cells. Immediately following transfection the DNA molecules within each cell can be very high and a natural decline in this concentration occurs to a steady state for plasmid maintenance (Middleton and Sugen, 1994 and Horlick et al.,1997). This number can vary, but the average is between 50 and 100 copies per cell with the further loss of approximately 5% per generation in the absence of selection (Yates and Guan, 1991). The constructs that were integrated into the host's chromosome remained relatively stable throughout the twenty passages (approximately 2.5 months) evaluation. These data indicate that integration of the *lux* genes within the host's chromosome may be the most suitable way to express the genes in mammalian cells for long-term gene maintenance and stable bioluminescence activity.

In order to optimize the bioluminescence potential from mammalian cells the *lux* genes need to be processed and expressed much in the way they are in bacteria. To establish a more natural expression format for the heterodimeric luciferase protein, the *luxA* and *luxB* genes were cloned into a bicistronic mammalian expression vector. This vector was developed to allow for the expression of two genes of interest under the control of a single constitutive promoter with the use of an internal ribosomal entry site (IRES). IRES elements can be defined as specific nucleotide sequences that allow for ribosomal entry and translation initiation directly at the start codon (AUG) rather than requiring scanning from the 5′ end, cap structure, of the mRNA (Pestova et al., 2001 and Kozak, 2001). Since the *lux* genes are naturally found in a polycisitronic operon, it

was thought that by expressing the genes in this format a more natural production and formation of the heterodimer could be obtained. From each of the stable cell line clones obtained harboring *lux* genes expressed as a bicistronic transcript, the bioluminescence (RLU/mg total protein) was at least an order of magnitude greater than levels obtained with any of the other expression formats tested. On average, there was no significant difference between bioluminescence levels obtained from HEK293 cells expressing the *luxAB* fusion, the *luxA* and *luxB* in a dual promoter format or as co-transfected separate plasmids. However, the bioluminescence levels from HEK293 cells harboring the *luxA* and *luxB* as a single bicistronic transcript constitutively produced significantly higher light levels.

Based on these data it was determined that of the four expression formats evaluated that the bicistronic expression of the *luxA* and *luxB* genes was by far the best choice. Furthermore, although in general, the bioluminescence levels were slightly less, the stability of the construct when integrated into the host's chromosome makes this a more suitable choice for the development of bacterial *lux*-based mammalian biosensors. Therefore, the first hypothesis set forth in this research can be accepted which stated that the expression of the bacterial luciferase (*lux*) subunits as individual proteins rather than as a monomeric translational fusion will result in more efficient folding and thermostability resulting in a higher bioluminescent signal in mammalian cells.

Codon Optimization of the Bacterial Luciferase for Expression in Mammalian Cells

Codon usage regulates gene expression at the level of translation and the usage patterns between species are not conserved (Kurland, 1991). This is especially true between genes derived from eukaryotes versus those from prokaryotes. Therefore, to efficiently express the bacterial *lux* genes in mammalian cells, the nucleotide sequence was altered in such a way as to create a "humanized" form of the gene without altering the amino acid sequence. This approach has been used previously to optimize the expression of both GFP and *Renilla* luciferase proteins for expression in mammalian cells (Zhang et al., 2002 and Gruber and Wood, 2000). The design of this new sequence was carefully determined, removing all potential splice sites and most regulatory regions. After the final codon optimized sequence was determined, it was evaluated using the GENSCAN prediction algorithm to determine the potential expression efficiency in a human cell. According to the output from this program the overall expression of the codon optimized *lux* genes would be significantly improved versus the wild type. The increase in expression was predicted to be caused by an increase in both transcription and translation efficiency. Furthermore, it was predicted that the first sixty bases (20 amino acids) of the wild type *luxA* gene would be completely eliminated when expressed in mammalian cells. Considering that this region of the LuxA protein holds most of the catalytic properties (active site) for the bacterial luciferase enzyme, this would be devastating for its expression. If this were the case, the low expression levels observed for the LuxAB fusion protein, shown

earlier, may be better explained in part by a nonfunctional protein being formed rather than inefficient folding or heat liability.

To test the expression of the codon optimized genes, modified versions were required. However, because the necessary changes were to numerous to achieve by site directed mutagenesis, a complete *in vitro* gene synthesis protocol was pursued. Large oligonucleotides (80-106 bp) were designed with overlapping (18-22 bp) regions. The original plan was to amplify all of the oligonucleotides together in one PCR reaction according to methods set by Prodromou and Pearl (1992). However, because of the larger size of the *lux* genes (approximately 1000 bp each), this was not possible. As an alternative, the gene was synthesized in parts and the subsequently linked by a second round of PCR. The two outside oligonucleotides were used as both template and primers for each reaction. After some experimental effort, it was determined that for optimal amplification the internal oligonucleotides (template) should be added at a concentration that equaled 100 fold less than the outside oligos. Amplification products of the correct size were cloned and sequenced. Unfortunately, sequence analysis revealed several base substitution mutations within all clones tested. These mutations were present despite the fact that care was taken by using a *Pfu* polymerase that has proof reading abilities. This finding was disappointing while not surprising given that two consecutive PCR reactions were required to obtain the final gene product resulting in > 60 cycles of amplification. To eliminate these errors and produce the proper sequence, site directed mutagenesis was performed.

It was determined previously through work accomplished in *S. cerevisiae* and mammalian cells for the expression of the bacterial luciferase genes that IRES elements

may be an efficient way to express independent proteins as single bicistronic transcripts. This expression format provides the most natural expression of the genes, most closely mimicking the polycistronic form found in the bacterial operon. Therefore, experiments were set up to compare the codon optimized and wild type *luxA* and *luxB* genes in mammalian cells using an IRES based expression vector. The expression vector used was designed to highly express two independent genes under the control of a single constitutive encephalomyocarditis virus (ECMV) promoter region by linking two multicloning sites fused to either side of an internal ribosomal entry site (IRES). The IRES element allows for the translation of two consecutive open reading frames from one messenger RNA (Jang et al., 1990; Jackson et al., 1990; Rees et al., 1996). By constructing plasmids with different combinations of the codon optimized *luxA* and *luxB* with their wild type counterparts, a direct comparison of the genes was made.

To quickly determine if a difference in translation efficiency could be detected between the optimized and wild type *luxA* genes, *in vitro* transcription and translation analysis was performed. The codon optimized *luxA* gene (COA) was detected approximately twenty fold over wild type. This finding supported the results that shown earlier with the *in vitro* generation of the wild type LuxA protein. Since the rabbit reticulocyte lysate translation system is used to mimic mammalian translation machinery *in vitro*, these results indicated that the codon optimization would indeed make a significant impact on the translation efficiency of the *lux* proteins in mammalian systems.

HEK293 cells were transfected with the WTA/ WTB, WTA/ COB or COA/COB constructs and stable cell line clones were selected by antibiotic resistance. Twenty

stable clones for each *luxA* and *luxB* combination were selected and bioluminescence levels were determined upon the exogenous addition of n-decanal and $FMNH_2$. The bioluminescence significantly increased in the order WTA/WTB < COA/WTB < COA/COB. These data indicated that codon optimization had made a significant impact on the potential bioluminescence levels obtained from mammalian cells. To analyze this data further, the two or three brightest clones were chosen for further study. From these cells, total genomic DNA, total RNA and total soluble proteins were extracted.

Foreign gene integration in mammalian cells is a random event, therefore it is possible to have more than one insertion of the construct occur during each transfection. Since integration is fairly inefficient, the copy number per cell is generally very low. However, because of this possibility, it was important to determine the copy number of the inserted genes for at true comparison. To accomplish this, Southern blot analysis was performed on each of the seven clones that produced the highest bioluminescence levels. All of the cell lines tested had either one or two copies of the gene inserted with the exception of the COA/COB3 clone which had three. To simplify further measurements, this clone was then disregarded for further bioluminescence comparisons. Nevertheless, it should be noted that increased copy number does not correlate with increased expression levels. Numerous other factors have been shown to impact from gene expression at a greater level.

The overall amount of *luxA* mRNA transcript was determined by Northern blot analysis. Transcript levels were relatively equal with the exception of the WTA/WTB1 clone that produced significantly lower amounts of *luxA* mRNA. Since each construct

contained the same promoter (ECMV) element and initiation signals, it would be expected that each clone would have approximately equal amounts of transcript for the introduced genes. However, *cis* acting regulatory elements could potentially interfere with transcription initiation and overall transcript levels *in vivo*. These types of interferences would vary based on where within the chromosome the genes were integrated. Therefore, the position effect of various clones could explain the lower amount of transcript detected with the WTA/WTB1 clone. Other factors that can potentially impact the amount of RNA transcript would be a direct result of increased RNA degradation of certain mRNA sequences that can occur. This type of RNA instability would be less likely after codon optimization because of the removal of several AU rich target degradation regions. However, because the lower amount of transcript was not seen in both the WTA/WTB clones, this scenario is unlikely.

To evaluate the overall protein concentrations and determine translation efficiency of each construct, Western blot analysis was performed. Total soluble proteins from each clone were isolated and quantified. Western blot analysis was performed using a polyclonal *luxA* antibody. The LuxA protein was not detectable from WTA/WTB clones and faintly visible in the COA/WTB clones. However, large amounts of LuxA protein were detected from the COA/COB clones which harbored a construct carrying codon optimized versions of both genes. This finding was intriguing and unexpected. Since the only available antibody was raised against a polypeptide of LuxA, it was expected that the constructs harboring the COA/WTB and COA/COB would produce equal amounts of LuxA protein. Since this was not the case, these data indicated that the codon optimization of both genes might infer stability on the

heterodimeric complex that makes up the luciferase enzyme. This increased stability of the complex may have allowed the proteins to be detected in the Western blot while the other construct was not detected.

Perhaps the most valuable measurement to determine if codon optimization was a success is the amount of enzyme activity that could be obtained from each construct. Bioluminescence levels were evaluated on whole cell extracts upon the addition of n-decanal and $FMNH_2$. Average bioluminescence values were found to be greater than two orders of magnitude higher in cell lines harboring both a codon optimized *luxA* and *luxB* (COA/COB) over that of the cell lines harboring the wild type genes (WTA/WTB). While bioluminescence levels were significantly higher in clones expressing COA and WTB versus WTA and WTB, the optimal bioluminescence was obtained from clones harboring optimized versions of both genes. These data further support the stabilization conclusion of the heterodimeric protein. Based on these data it was determined that codon optimization had a significant effect on the protein expression in HEK293 cells.

In conclusion, the codon optimization of the *luxA* and *luxB* genes was successful in increasing the overall expression levels of the individual proteins. This increase in protein quantity resulted in a significant increase in bioluminescence from cell lines harboring these constructs. Furthermore, the bioluminescence levels from codon optimized *luxA* and *luxB* provide adequate bioluminescence for the proof in principle data needed for the future development of reliable reporter constructs for analyte sensing in mammalian cells. Based on these data, the second hypothesis statement proposed in this research that stated that codon optimization of the bacterial

luciferase (*lux*) genes will significantly enhance translation of the proteins and ultimately result in greater bioluminescence levels from mammalian cells harboring these optimized genes is accepted.

$FMNH_2$ Availability

Among the problems associated with the ultimate development of a mammalian *lux* bioreporter cell line, the lack of available $FMNH_2$, the reduced flavin molecule that is required for the bioluminescence reaction is perhaps one of the most obvious. The $FMNH_2$ limitation in mammalian cells engineered with the *luxA* and *luxB* genes was shown to hamper potential light outputs significantly. These data clearly showed that light levels were significantly enhanced upon the exogenous addition of a flavin reductase enzyme. Therefore, efforts were explored to overcome the $FMNH_2$ limitation by attempting to express the reductase enzyme *in vivo* to provide adequate levels of this substrate.

Recently, this approach was taken in yeast cells by the overexpression of an NAD(P)H –dependent FMN reductase enzyme from *V. harveyi* (Gupta et al., 2003 an Szittner et al., 2003). In these studies, the bioluminescence levels from whole yeast cells were significantly increased by the expression of the *frp* gene. These data illustrated that the limiting substrate in yeast was indeed a lack of the reduced $FMNH_2$ and that the FMN and NAD(P)H precursors were available within cytoplasm of yeast cells. Based on these results, it was hypothesized that the overexpression of the *frp* gene in mammalian cells may also provide adequate pools of $FMNH_2$ to catalyze

bioluminescence. Therefore, HEK293 cells harboring a codon optimized *luxA* and *luxB* gene pair were co-transfected with the *frp* gene and stable cell lines obtained.

From *in vitro* bioluminescence assays of clones expressing *frp*, the overall light levels increased at least an order of magnitude in both the absence and after the addition of exogenous flavin oxidoreductase versus the COA/COB2 clone without the *frp* gene, tested under the same conditions. These data indicated that the expression of the *frp* gene was successful in producing higher concentrations of $FMNH_2$ within HEK293 cells. However, because the light levels increased further upon the addition of exogenous flavin oxidoreductase, it was determined that $FMNH_2$ remains the limiting factor for the bioluminescence potential from these mammalian cells and has yet to reach saturation. Based on these data, further experiments will be necessary to accept or reject the third hypothesis set forth in this work, but at this point $FMNH_2$ remains the limiting substrate for bioluminescence.

Nevertheless, it is believed that much of the available $FMNH_2$ within mammalian cells would be found at much higher concentrations within the mitochondria of the cell. As a result, the $FMNH_2$ would remain sequestered from the bioluminescence (Lux) proteins within the cytoplasm and not available for the reaction. By monitoring light levels as total protein extracts, the location of available $FMNH_2$ within the cytoplasm is unknown. To better determine cytoplasmic concentrations of $FMNH_2$, whole cell bioluminescence assays were performed. Overall, the light levels from whole cells were much lower (as much as 20 fold) than those obtained from protein extracts. These data indicate that the cytoplasmic concentrations of $FMNH_2$ are limiting and that much of the cellular pool of the reduced flavin molecule is

compartmentalized within organelles, like the mitochondria. Therefore, future experiments should be conducted to allow for the expression of the Lux proteins in the mitochondria of the cell to allow for their interaction with other necessary substrates for the reaction. Targeting of the Lux proteins can be easily accomplished with the addition of a signaling peptide on the N- terminus of the proteins that will shuttle the enzymes into the mitochondria of the cell.

The bioluminescence levels obtained from the cell extract, *in vitro*, assays remained stable for several minutes before gradually declining to background levels. The light intensity could be increased back to peak levels upon exogenous addition of additional NAD(P)H to provide the reducing power for the flavin oxidoreductase enzyme and generate more $FMNH_2$. Thus, the luciferase complex itself remained stable throughout the assay and the light output was directly correlated to the available $FMNH_2$. On the other hand, the whole cell *in vivo* assays provided a flash luminescence response that could not be induced further to give a second peak upon the addition of n-decanal or $FMNH_2$. These data indicate that the n-decanal substrate was toxic to the cells. Once the cells were lysed and proteins extracted, high bioluminescence levels could be obtained upon the addition of n-decanal and $FMNH_2$. These data further support the idea that the addition of n-decanal to whole cells resulted in cell mortality and not the loss of stable Lux proteins.

Stability of the Luciferase Constructs in Mammalian Cells

The stability of mammalian cell lines engineered to express the bacterial luciferase genes was monitored to determine if the constructs could maintain their

bioluminescence levels over time in the absence of antibiotics. Although the bioluminescence outputs remained relatively constant for > 40 passages, one of the clones (WTA/WTB2) bioluminescence levels radically deteriorated after passage thirty. Furthermore, during this culture period, other phenotypic changes occurred within the cells, including a lower binding affinity to the flask surface. It was further determined that the loss of light from the WTA/WTB2 clone was not a result of the loss of the genes within the cells and perhaps a change in the cell's physiology or some unknown regulatory mechanism that caused the loss of light. Even though the WTA/WTB2 clone resulted in a complete loss of bioluminescence activity, PCR analysis revealed that the *luxA* and *luxB* genes were still present. Nevertheless, the bioluminescence levels did remain stable for the other five clones tested for more than five months in culture without the need for antibiotic selection, indicating that these cell lines would be stable enough for relatively long term monitoring projects and applications given that the proper control experiments were also included to ensure the integrity of the reporter strain.

LITERATURE CITED

Alam, J. and J. K. Cook. 1990. Reporter genes: Application to the study of mammalian gene transcription. *Anals of Biochemistry*. 188:245–254.

Almashanu, S., B. Musafia, R. Hader, M. Suissa, and J. Kuhn. 1990. Fusion of *luxA* and *luxB* and its expression in Escherichia coli, *Sacchromyces cervisiae* and *Drosophila melanogaster*. *Journal of Bioluminescence and Chemiluminescence.* 5(2):89-98.

Anderson, S. G. and C. G. Kurland. 1990. Codon preferences in free-living microorganisms. *Microbiological Reviews*. 54(2):198-210.

Aota, S. and T. Ikemura. 1986. Diversity in G + C content at the third position of codons in vertebrate genes and its cause. *Nucleic Acids Research.* 14(16):6345-6355.

Applegate, B., C Kelley, L. Lackey, J. McPherson, S. Kerhmeyer, F-M. Menn, P. Bienkowski and G. Sayler.1997. *Pseudomonas putida* B2: a *tod-lux* bioluminescent reporter for toluene and trichloroethylene co-metabolism. *Journal of Industrial Microbiology and Biotechnology*. 18: 4-9.

Arregui, C., V. Juarez, and H. zur Hausen. 2003. A Synthetic E7 Gene of Human Papillomavirus Type 16 that Yields Enhanced Expression of the protein in Mammalian Cells and is Useful for DNA Immunization Studies. *Journal of Virology.* 77(8):4928-4937.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. D. Seidman, J. A. Smith and K. Struhl. 1997. *Short Protocols in Molecular Biology*, eds. 7. John Wiley & Sons, New York, N. Y.

Baev, D. X. Li and M. Edgerton. 2001. Genetically engineered human salivary histatin genes are functional in *Candida albicans;* development of a new system for studying histatin candidacidal activity. *Microbiology*. 147(12):3323-3334.

Baldwin, T. O., M. M. Ziegler and D. A. Powers. 1979. Covalent structure of subunits of bacterial luciferase NH_2-terminal sequence demonstrates subunit homology. *PNAS*. 76:4887-4889.

Baldwin, T. O., J. A. Christopher, A. K. Forrest, C. J. Shofield.1995. Structure of bacterial luciferase. *Current Opinion in Structural Biology.* 5(6):798-809.

Berlioz, C. and J. L. Darlix. 1995. An internal ribosomal entry mechanism promotes translation of murine leukemia virus gag polyprotein precursors. *Journal of Virology*. 69(4):2214-2222.

Bernstein, J., O. Sella, S.-Y. Le and O. Elroy-Stein.1997. PDGF2/c-*sis* mRNA leader contains a differentiation-linked internal ribosomal entry site (D-IRES). *Journal of Biological Chemistry.* 272:9356-9362.

Boylan, M., Pelletier, J. & Meighen, E.A.1989. Fused bacterial luciferase subunits catalyze light emission in eukaryotes and prokaryotes. *Journal of Biological Chemistry*. 264: 1915-1918.

Burlage, R.S., Bemis, L.A., Layton, A.C., Sayler, G.S. & Larimer, F. 1990. Comparative genetic organization of incompatibility group-P degradative plasmids. *Journal of Bacteriology.* 172: 6818-6825.

Cai, J. and M. S. DuBow. 1997. Use of a luminescent bacterial biosensor for biomonitoring and characterization of arsenic toxicity of chromated copper arsenate (CCA). *Biodegradation.* 8: 105-111.

Cline, T. W. and J. W. Hastings. 1972. Mutationally altered bacterial luciferase. Implications for subunit functions. *Biochemistry*. 11: 3359-3370.

Cline, T. W. and J. W. Hastings.1973. Mutationally altered bacterial luciferase. Implications for subunit functions. 11(18): 3359-3370.

Cohn, D. H., R. C. Ogden, J. N. Ableson, T. O. Baldwin, K. H. Nealson, M. I. Simon and A. J. Mileham.1983. Cloning of the *Vibrio harveyi l*uciferase genes: use of a synthetic oligonucleotide probe. *PNAS.* 80:120-123.

Coldwell, M. J., S. A. Mitchell, M. Stoneley, M. MacFarlane, A. E. Willis. 2000. Initiation of Apaf-1 translation by internal ribosome entry. *Oncogene*. 19(7): 899-905.

Colepicolo, P., K. W. Cho, G. O. Poinar, J. W. Hastings. 1989. Growth and luminescence of the bacterium *Xenorhabdus luminescens* from a human wound. A*pplied and Environmental Microbiology*. 55:2601-2606.

Contag,C. H., P. R. Contag, J. I Mullins, S. D. Stevenson and D. A. Benaron. 1995. Photonic detection of bacterial pathogens in living hosts. *Molecular Microbiology*. 18:593-603.

Contag, P.R., Olomu, I.N., Stevenson, D.K. & Contag, C.H. 1998. Bioluminescent indicators in living mammals. *Nature Medicine*. 4:245-247.

Contag, C. H., D. K. Stevenson. 2001. *In vivo* patterns of heme oxygenase-1 transcription. *Journal of Parintology*. S1:119-127.

Corbiser, P. 1996. Bacterial biosensors for the toxicity assessment of solid wastes. *Environmental Toxicology and Water Quality*. 11:171-177.

Corbiser, P. et al. 1999. Whole cell- and protein-based biosensors for the detection of bioavailable heavy metals in environmental samples. *Anal. Chim. Acta.* 387:235-244.

Costa, S., E. Douville, J. Bell and E. Meighen. 1991. Expression of fused bacterial luciferase in mammalian cells. *Bioluminescence and Chemiluminescence*. 1991:31-34.

Creancier, L, D. Morello, P. Mercier, A. C. Prats.2000. Fibroblast growth factor 2 internal ribosome entry site (IRES) activity *ex vivo* and in transgenic mice reveals a stringent tissue-specific regulation. *Journal of Cell Biology*. 150(1):275-281.

Daunert, S. et al. 2000. Genetically engineered whole-cell sensing systems: Coupling biological recognition with reporter genes. *Chemical Reviews.* 100:2705-2738.

De Amicis, F. and S. Marchetti. 2000. Intercodon dinucleotides affect codon choice in plant genes. *Nucleic Acids Research*. 28(17):3339-3375.

deFilpe, P. 2002. Polycistronic viral vectors. *Current Gene Therapy.* 2(3):355-378.

deLorenzo, V., S. Fernandez, M. Herrero, U. Jakubik and K. Timmis. 1993. Engineering of alkyl-and haloaromatic-responsive gene expression with mini-transposons containing regulated promoters of biodegradative pathways of *Pseudomonas. Gene*. 130:41-46.

Deml, L., A. Bojak, S. Steck, M. Graf, J. Wild, R. Schirmbeck, H. Wolf, R. Wagner. 2001. Multiple effects of codon usage optimization on expression and immunogenicity of DNA candidate vaccines encoding the human immunodeficiency virus type 1 Gag protein. *Journal of Virology*. 75(22):10991-11001.

Disbrow, G.L., I. Sunitha, C. C. Baker, J. Hanover, R. Schlegel. 2003. Codon optimization of the HPV-16 E5 gene enhances protein expression. *Virology*. 311(1):105-114.

Engebrecht, J., K. H. Nealson and M. Silverman. 1983. Bacterial bioluminescence: isolation and genetic analysis of functions from *Vibrio fischeri*. *Cell*. 32: 773-781.

Escher, A., D.J. O'Kane, J. Lee, and A.A. Szalay. 1989. Bacterial luciferase αβ fusion protein is fully active as a monomer and highly sensitive *in vivo* to elevated temperature. *PNAS*. 86:6528-6532.

Farmer, J. J., J. H. Jorgensen, P. A. D. Grimont, R. J. Akhurst, G. O. Poinar, Jr., E. Ageron, G. V. Pierce, J. A. Smith, G. P. Carter, K. L. Wilson and F. W. Hickman-Brenner. 1989. *Xenorhabdus luminescens* (DNA hybridizations group 5) from human clinical specimens. *Journal of Clinical Microbiology*. 27:1594-1600.

Fisher, A. J., F. M. Raushel T. O. Baldwin and I. Rayment. 1995. Three-dimensional structure of bacterial luciferase from *Vibrio harveyi* at 2.4 A resolution. *Biochemistry*. 34:6581-6586.

Fisher, A. J., T. B. Thompsom, J. B. Thoden, T. O. Baldwin. 1996. The 1.5 A resolution crystal structure of bacterial luciferase in low salt conditions. *The Journal of Biological Chemistry*. 271(36): 21956-21968.

Francis, K. P., D. Joh, C. Bellinger-Kawahara, M. J. Hawkinson, T. F. Purchio and P. G. Contag. 2000. Monitoring bioluminescent *Staphylococcus* infections in living mice using a novel *luxCDABE* construct. *Infection and Immunity*. 68(6):3594-3600.

Francis, K.P., J. Yu, C. Bellinger-Kawahara, D. Joh, M. J. Hawkinson, G. Xiao, T. F. Purchio, M. G. Caparon, M. Lipstitch and P. R. Contag. 2001. Visualizing pneumococcal infections in the lungs of live mice using bioluminescent *Streptococcus pneumoniae* transformed with a novel gram-positive *lux* transposon. *Infection and Immunity.* 69:3350-3358.

Gelmini, S., P. Pinzani, C. Orlando, R. Sestini, T. O. Baldwin and M. Pazzagli. 1993. Chloramphenicol acetyl-transferase, firefly and bacterial luciferase as reporter genes in transfection of mammalian cells. *Bioluminescence and Chemiluminescence*. 1993: 200-206.

German, M.S., L.G. Moss, and W.J. Rutter. 1990. Regulation of insulin gene expression by glucose and calcium in transfected primary islet cultures. J*ournal of Biological Chemistry*. 265:22063-22066.

Goldman, E., A. H. Rosenberg, G. Zubay, F. W. Studier. 1995. Consecutive low-usage leucine codons block translation only when near the 5' end of a message in *Escherichia coli. Journal of Molecular Biology.* 245(5):467-473.

Gorski, J.A., K. R. Jones. 1999. Efficient bicistronic expression of *cre* in mammalian cells. *Nucleic Acids Research*. 27(9):2059-2061.

Gouy, M. and C.Gautier. 1982. Codon usage in bacteria: correlation with gene expressivity. *Nucleic Acids Research*. 10(22):7055-7074.

Gough, D.A., J.C. Armour, and D.A. Baker. 1997. Advances and prospects in glucose assay technology. *Diabetologia*. 40:S102-S107.

Grantham, R., C. Gautier, M. Gouy, M. Jacobzone, R. Mercier. 1981. Codon catalog usage is a genome strategy modulated for gene expressivity. *Nucleic Acids Research.* 9(1):43-74.

Gruber, M.G., and Wood, K.V. 2000. Design strategy for synthetic luciferase reporter genes. *Abstracts for the International Symposium for ISBC*, Monterey, CA.

Guignard, I., C. Combadiere, H.L. Tiffany and P.M. Murphy.1998. Gene organization and promoter function for CC chemokine receptor 5 (CCR5). *Journal of Immunology*. 160:985–992.

Gupta, R. K., S. S. Patterson, S. A. Ripp and G. S. Sayler. 2003. Expression of the *Photorhabdus luminescens lux* genes (*luxA, B, C, D* and *E*) in *Saccharomyces cerevisiae. FEMS Yeast Research*. In Press.

Haller, A. A. and B. L. Semler. 1992. Linker scanning mutagenesis of the internal ribosome entry site of poliovirus RNA. *Journal of Virology*. 66(8):5075-5086.

Harries, M., N. Phillipps, R. Anderson, G. Prentice and M. Collins. 2000. Comparison of bicistronic vectors containing internal ribosomal entry sites (IRES) using expression of human interlukin-12 (IL-12) as a readout. *Journal of Gene Medicine*. 2(4):243-249.

Harvey, E. N. 1952. *Bioluminescence*. Academic Press, Incorporated, New York, NY.

Hastings, J. W., C. J. Potrikas, S. C. Gupta, M. Kurfurst and J. C. Makemson. 1985. Biochemistry and physiology of bioluminescent bacteria. *Advances in Microbial Physiology*. 26:235-291.

Hay, A. G., B. M. Applegate, N. G. Bright, and G. S. Sayler. 2000. A bioluminescent whole-cell reporter for detection of 2,4-dichlorophenoxyacetic acid and 2,4-dichlorophenol in soil. *Applied and Environmental Microbiology*. 66:4589-4594.

Hellen, C.U.T. and P. Sarnow. 2001.Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes and Development.* 15:1593-1612.

Henis-Korenblit, S., N. L. Strumpf, D. Goldstaub, A. Kimchi. 2000. A novel form of DAP5 protein accumulates in apoptotic cells as a result of caspase cleavage and internal ribosome entry site-mediated translation. *Molecular and Cellular Biology*. 20(2): 496-506.

Hennecke, M., M. Kwissa, K. Metzger, A. Oumard, A. Kroger, R. Schirmbeck, J. Reimann and H. Hauser. 2001. Composition and arrangement of genes define the strength of IRES-driven translation in bicistronic mRNAs. *Nucleic Acids Research*. 29(16): 3327-3334.

Henstrand, J.M., A. Schaller, M. Braun, N. Amrhein, and J. Schmid.1996. *Saccharomyces cerevisiae* chorismate synthase has a flavin reductase activity. *Molecular Microbiology*. 22: 859-866.

Hershey, J.W.B. and W. C. Merrick. 2000. The pathway and mechanism of initiation of protein synthesis. In Sonenberg, N., Hershey,J.W.B. and Mathews,M.B. (eds), *Translational Control of Gene Expression*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 33–88.

Hoekema, A., R. A. Kastelein, M. Vasser, H. A. de Boer. 1987. Codon replacement in the PGK1 gene of *Saccharomyces cerevisiae*: experimental approach to study the role of biased codon usage in gene expression. *Molecular and Cellular Biology*. 7(8):2914-2924.

Horlick, R. A., A. E. Schilling, P. Samama, R.N. Swanson, V. D. Fitzpatrick, Robbins AK, B. Damaj. 2000. Combinatorial gene expression using multiple episomal vectors. *Gene.* 243(1-2):187-194.

Holick, M. and R. G. Korneluk. 2000. Functional characterization of the X-linked inhibitor of apoptosis (XIAP) internal ribosomal entry site element: role of La autoantigen in XIAP translation. *Molecular and Cellular Biology*. 20:4648-4657.

Hollis, R.P. et al. 2001. Toxicity of the bacterial luciferase substrate, n-decyl aldehyde, to *Saccharomyces cerevisiae* and *Caenorhabditis elegans*. *FEBS Letters*. 506: 140-142.

Honda, M., E. A. Brown and S. M. Lemon. 1996. Stability of a stem-loop involving the initiator AUG controls the efficiency of internal initiation of translation on hepatitis C virus RNA. *RNA*. 2: 955-968.

Honda, M., R. Rijnbrand, G. Abell, D. Kim, S. M. Lemon. 1999. Natural variation in translational activities of the 5' nontranslated RNAs of hepatitis C virus genotypes 1a and 1b: evidence for a long-range RNA-RNA interaction outside of the internal ribosomal entry site. *Journal of Virology*. 73(6):4941-4951.

Huez, I., L. Creancier, S. Audigier, M. C. Gensac, A. C. Prats, H. Prats. 1998. Two independent internal ribosome entry sites are involved in translation initiation of vascular endothelial growth factor mRNA. *Molecular and Cellular Biology.* 18(11):6178-6190.

Ikemura, T. 1985. Codon usage and tRNA content in unicellular and multicellular organisms. *Molecular Biology of Evolution*. 2(1):13-34.

Ichiki, T., M. Usui, M. Kato, Y. Funakoshi, K. Ito, K. Egashira and A. Takeshita. 1998. Down regulation of angiotensin II type 1 receptor gene transcription by nitric oxide. *Hypertension*. 31: 342–348.

Jackson, R. J. 1988. RNA Translation. Picornaviruses break the rules. *Nature*. 334:292-293.

Jackson, R. J. 2000. Comparative view of initiation site selection mechanism. *Translational control of gene expression.* P. 151. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, N. Y.

Jang, S. K., H. G. Krausslich, M. J. Nicklin, G. M. Duke, A. C. Palmenberg, E. Wimmer E. 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *Journal of Virology*. 62(8):2636-2643.

Johannes, G. and P. Sarnow. 1998. Cap-independent polysomal association of natural mRNAs encoding c-myc, BiP, and eIF4G conferred by internal ribosome entry sites. *RNA*. 4(12):1500-1513.

Jubin, R., N. E.Vantuno, J. S. Kieft, M. G. Murray, J. A. Doudna, J. Y. Lau, B. M. Baroudy. 2000. Hepatitis C virus internal ribosome entry site (IRES) stem loop IIId contains a phylogenetically conserved GGG triplet essential for translation and IRES folding. *Journal of Virology.* 74(22):10430-10437.

Kalinski, A., J. M. Weisemann, B. F. Matthews, E. M. Herman. 1990. Molecular cloning of a protein associated with soybean seed oil bodies that is similar to thiol proteases of the papain family. *Journal of Biological Chemistry*. 265(23):13843-13848.

Khang, Y. H. et al. 1997. Measurement of iron-dependence of *pupA* promoter activity by a *pup-lux* bioreporter. *Journal of Microbiology and Biotechnology*. 7: 352-355.

King, J.M.H. et al. 1990. Rapid, sensitive bioluminescent reporter technology for naphthalene exposure and biodegradation. *Science* 249: 778-781.

Kirchner, G., Roberts, J.L., Gustafson, G.D. & Ingolia, T.D. 1989. Active bacterial luciferase from a fused gene expression of a *Vibrio harveyi luxAB* translational fusion in bacteria, yeast and plant cells. *Gene*. 81: 349-354.

Kofman, A., M. Graf, L. Deml, H. Wolf, R. Wagner. 2003a. Codon usage-mediated inhibition of HIV-1 gag expression in mammalian cells occurs independently of translation. *Tsitologiia.* 45(1):94-100.

Kofman, A., M. Graf, A. Bojak, L. Deml, K. Bieler, A. Kharazova, H. Wolf, R. Wagner. 2003b. HIV-1 gag expression is quantitatively dependent on the ratio of native and optimized codons. *Tsitologiia*. 45(1):86-93.

Kohler, S., Belkin, S. & Schmid, R.D. 2000. Reporter gene bioassays in environmental analysis. *Fresenius Journal of Analitical Chemistry*. 366:769-779.

Koncz, C., O. Olsson, W.H.R. Langridge, J. Schell, and A.A. Szalay. 1987. Expression and assembly of functional bacterial luciferase in plants. *PNAS*. 84:131-135.

Kozak, M. 1987. An Analysis of 5'-Noncoding Sequences from 699 Vertebrate Messenger RNAs. *Nucleic Acids Research*. 15:8125-8148.

Kozak, M. 1990. Downstream Secondary Structure Facilitates Recognition of Initiator Codons by Eukaryotic Ribosomes. *PNAS*. 87:8301-8305.

Kozak, M. 1991. An Analysis of Vertebrate mRNA Sequences: Intimations of Translational Control. *Journal of Cell Biology*. 115:887-903.

Kozak, M. 2001. New ways of initiating translation in eukaryotes? *Molecular and Cellular Biology.* 21:8241-8246.

Kurland, C.G. 1991. Codon bias and gene expression. *FEBS Letters*. 285:165-169.

Kurland, C. G. 1993. Major Codon Preference: theme and variations. *Biochemical Society Transactions*. 1993:21.

Kuhn, R., N. Luz, E. Beck. 1990. Functional analysis of the internal translation initiation site of foot-and-mouth disease virus. *Journal of Virology*. 64(10):4625-4631.

Levitzki, A. 1996. Targeting signal transduction for disease therapy. *Current Opinions in Cell Biology.* 8(2):239-244.

Li, J., R. Szittner and E. Meighen. 2000. Hyperacitivity and interactions of a chimeric myristoryl-ACP thioesterase from the *lux* system of luminescent bacteria. *Biochimica et Biophysica Acta.* 1481:237-246.

Li, Z. and E. Meighen. 2000. The turnover of bacterial luciferase is limited by a slow decomposition of the ternary enzyme-product complex of luciferase, FMN, and fatty acid. *Journal of Biological Chemistry.* 269:6640-6644.

Macejak, D. G. and P. Scarnow. 1991. Internal initiation of translation mediated by the 5' leader of a cellular mRNA. *Nature.* 353(6339):90-94.

Martinez-Salas, E. 1999. Internal ribosome entry site biology and its use in expression vectors. *Current Opinions in Biotechnology*. 10(5):458-464.

Martinez-Salas, E., R. Ramos, E. Lafuente, Lopez de Quinto. 2001. Functional interactions in internal translation initiation directed by viral and cellular IRES elements. *Journal of General Virology*. 82(5):973-984.

Meighen, E. A., M. Z. Nicoli and J. W. Hastings. 1971. Functional differences of the nonidentical subunits of bacterial luciferase, properties of hybrids of native and chemically modified bacterial luciferase. *Biochemistry*. 10:4069-4073.

Meighen, E. A. and I. Bartlett. 1980. Complementation of subunits from different bacterial luciferases. Evidence for the role of the β subunit in the bioluminescent mechanism. *Journal of Biological Chemistry*. 255:11181-11187.

Meighen, E.A. 1991. Molecular biology of bacterial bioluminescence. *Microbiological Reviews*. 55:123-142.

Meighen, E.A. & Szittner, R.B. 1992. Multiple repetitive elements and organization of the *lux* operons of luminescent terrestrial bacteria. *Journal of Bacteriology.* 174:5371-5381.

Meighen, E.A. & Dunlap, P.V. 1993. Physiological, biochemical and genetic control of bacterial bioluminescence. *Advances in Microbial Physiology*. 34:1-67.

Meilke, C., M. Tummler, D. Schubeler, I. Von Hoegen, H. Hauser. 2000. Stabilized long-term expression of the heterodimeric proteins from tricistronic mRNA. *Gene*. 254:1-8.

Middleton, T., B. Sugden. 1994. Retention of plasmid DNA in mammalian cells is enhanced by binding of the Epstein-Barr virus replication protein EBNA1. *Journal of Virology*. 68(6):4067-4071.

Mizuguchi, H., Z. Xu, A. Ishii-Watabe, E. Uchida, T. Hayakawa. 2000. IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector. *Molecular Therapy*. 1(4):376-382.

Mosser, D.D., A.W. Caron, L. Bourget, P. Jolicouer and B. Massie. 1997. Use of a dicistronic expression cassette encoding the green fluorescent protein for the screening and selection of cells expressing inducible gene products. *Biotechniques*. 22:150–152.

Narum, D. L., S. Kumar, W. O. Rogers, S. R. Fuhrmann, H. Liang, M. Oakley, A. Taye, B. K. Sim, S. L. Hoffman. 2001. Codon optimization of gene fragments encoding *Plasmodium falciparum* merzoite proteins enhances DNA vaccine protein expression and immunogenicity in mice. *Infection and Immunity*. 69(12): 7250-7253.

Nassal, M., T. Mogi, S. S. Karnik, H. G. Khorana. 1987. Structure-function studies on bacteriorhodopsin. III. Total synthesis of a gene for bacterio-opsin and its expression in *Escherichia coli. Journal of Biological Chemistry*. 262(19):9264-9270.

Naylor, L. H. 1999. Reporter gene technology: the future looks bright. Biochemical Pharmacology. 58:749-757.

Nicholson, R., J. Pelletier, S. Y. Le, N. Sonenberg. 1991. Structural and functional analysis of the ribosome landing pad of poliovirus type 2: *in vivo* translation studies. *Journal of Virology.* 65(11):5886-5894.

Oh, S. K., M. P. Scott, P. Sarnow. 1992. Homeotic gene Antennapedia mRNA contains 5'-noncoding sequences that confer translational initiation by internal ribosome binding. *Genes and Development.* 6(9):1643-1653.

Ohlmann, T., M. Lopez-Lastra, J. L. Darlix. 2000. An internal ribosome entry segment promotes translation of the simian immunodeficiency virus genomic RNA. *Journal of Biological Chemistry*. 275(16):11899-11906.

Olsson, O. et al. 1989. Engineering of monomeric bacterial luciferases by fusion of *luxA* and *luxB* genes in *Vibrio harveyi. Gene*. 81:335-347.

Oumard, A., M. Hennecke, H. Hauser, M. Nourbakhsh. 2000. Translation of NRF mRNA is mediated by highly efficient internal ribosome entry. *Molecular and Cellular Biology*. 20(8):2755-2759.

Pazzagli, L. J. H. Devine, D. O. Peterson and T. O. Baldwin. 1992. Use of bacterial and firefly luciferase as reporter genes in DEAE-dextran-mediated transfection in mammalian cells. *Biochemistry.* 204:315-323.

Pelletier, J. and N. Soneberg. 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature.* 334:320-325.

Pestova, T. V., C. U. T. Hellen and E. Wimmer. 1991. Translation of poliovirus RNA: Role of an essential cis-acting oligopyrimidine element within the 5'-nontranslated region and involvement of a cellular 57-kilodaton protein. *Journal of Virology*. 65:6194–6204.

Pestova, T.V., I. N. Shatsky, C. U. Hellen. 1996. Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Molecular and Cellular Biology.* 16(12):6870-6878.

Pestova, T. V., S. I. Borukhov and C. U. T. Hellen. 1998. Eukaryotic ribosomes require eIFs 1 and 1A to locate initiation codons. *Nature*. 394: 854–859.

Pestova, T. V., V. G. Kolupaeva, I. B. Lomakin, E. V. Pilipenko, I. N. Shatsky, V. I. Agol and C. U.T. Hellen. 2001. Molecular mechanisms of translation initiation in eukaryotes. *PNAS*. 98(13):7029-7036.

Prest, A. G., M. K. Winson, J. R. M. Hammond and G. S. A. B. Stewart. 1997. The construction and application of a lux-based nitrate biosensor. *Letters of Applied Microbiology.* 24:355-360.

Prodromou, C., and L.H. Pearl. 1992. Recursive PCR: a novel technique for total gene synthesis. *Protein Engineering*. 5:827-829.

Reynolds, J. E., A. Kaminski, A. R. Carroll, B. E. Clarke, D. J. Rowlands, R. J. Jackson. 1996. Internal initiation of translation of hepatitis C virus RNA: the ribosome entry site is at the authentic initiation codon. *RNA*. 2(9): 867-878.

Ripp, S. et al. 2000. Controlled field release of a bioluminescent genetically engineered microorganism for bioremediation process monitoring and control. *Environmental Science and Technology.* 34:846-853.

Rodriduez, A., L. Wall, D. Riendeau and E. A. Meighen. 1983. Fatty acid acylation of proteins in bioluminescent bacteria. *Biochemistry*. 22:5604-5611.

Sachs, A.B., Sarnow, P. & Hentze, M.W. 1997. Starting at the beginning, middle, and end: Translation initiation in eukaryotes. *Cell.* 89:831-838.

Sayler, G.S., Fleming, J.T. & Nivens, D.E. 2001. Gene expression monitoring in soils by mRNA analysis and gene *lux* fusions. *Current Opinions in Biotechnology*. 12:455-460.

Schmidt, T. M., K. Kopecky and K. H. Nealson. 1989. Bioluminescence of the insect pathogen *Xenorhabdus luminescens*. *Applied and Environmental Microbiology.* 55: 2607-2612.

Schultz D. W. and M.Yarus. 1996. On malleability of the genetic code. *Journal of Molecular Evolution.* 42:597-601.

Schwechheimer, C.. C. Smith and M.W. Bevan. 1998. The activities of acidic and glutamine-rich transcriptional activation domains in plant cells: Design of modular transcription factors for high-level expression. *Plant Molecular Biology*. 36:195–204.

Selam, J.L. 1997. Management of diabetes with glucose sensors and implantable insulin pumps. *ASAIO Journal*. 137-142.

Sharp, P. M., T. M. Tuohy, K. R. Mosurski. 1986. Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acid Research.* 14(13):5125-5143.

Sharp, P. M., E. Cowe, D. G. Higgins, D. C. Shields, K. H. Wolfe and F. Wright. 1988. Codon usage patterns in *Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster* and *Homo sapiens;* a review of the considerable within species diversity. *Nucleic Acids Research*. 16(17):8207-8211.

Sharp, P. M., M. Stenico, J. F. Peden and A. T. Lloyd. 1993. Codon usage: mutational bias, translational selection, or both? *Biochemical Society Transactions*. 21.

Sharp, P. M. and E. Cowe. 1991. Synonymous codon usage in *Saccharomyces cerevisiae. Yeast.* 7(7):657-678.

Simpson, M.L., G.S. Sayler, B.M. Applegate, S. Ripp, D.E. Nivens, M.J. Paulus, and G.E. Jellison Jr. 1998. Bioluminescent bioreporter integrated circuits: novel whole-cell biosensors. *Trends in Biotechnology*. 16:332-338.

Simpson, M.L., Sayler, G.S., Fleming, J.T. & Applegate, B. 2001. Whole-cell biocomputing. *Trends in Biotechnology*. 19:317-323.

Sorensen, M. A., C. G. Kurland, S. Pedersen. 1989. Codon usage determines translation rate in *Escherichia coli. Journal of Molecular Biology*. 207(2):365-377.

Stoneley, M., T. Subkhankulova, J. P. Le Quesne, M. J. Coldwell, C. L. Jopling, G. J. Belsham, A. E. Willis. 2000. Analysis of the c-myc IRES; a potential role for cell-type specific trans-acting factors and the nuclear compartment. *Nucleic Acid Research*. 28(3):687-694.

Szittner, R. & Meighen, E. 1990. Nucleotide sequence, expression and properties of luciferase coded by *lux* genes from a terrestrial bacterium. *Journal of Biological Chemistry.* 265:16581-16587.

Tibazarwa, C., S. Wuertz, M. Mergeay, L. Wyns, D. van Der Lelie. 2000. Regulation of the *cnr* cobalt and nickel resistance determinant of *Ralstonia eutropha (Alcaligenes eutrophus)* CH34. *Journal of Bacteriology*. 182(5):1399-1409.

Tu S. C. and J. Henkin. 1983. Characterization of the aldehyde binding site of bacterial luciferase by photoaffinity labeling. *Biochemistry.* 22(2):519-523.

Tu, S.C., and H.I.X. Mager. 1995. Biochemistry of bacterial bioluminescence. *Photochemistry and Photobiology*. 62:615-624.

Vagner, S., B. Galy and S. Pyronnet. 2001. Irresistible IRES. *EMBO reports.* 21(101):893-898.

Van Dyk, T.K., Y. Wei, M. K. Hanafey, M. Dolan, M. J. Reeve, J. A. Rafalski, L. B. Rotham-Denes and R. A. LaRossa. 2001. A genomic approach to gene fusion technology. *PNAS* U. S. A. 98:2555-2560.

Vervoort, E. B., A. van Ravenstein, N.N. M. E. van Peij, J. C. Heikoop, P. J. M. van Haastert, G. F. Verheijden and M. H. K. Linskens. 2000. Optimizing heterologous expression in *Dictyostelium*: importance of 5' codon adaptation. *Nucleic Acids Research.* 28(10):2069-2074.

Volkova, N., R. Szittner, and E. Meighen. 1999. Control of Luminescence decay and flavin binding by the LuxA carboxyl-terminal regions in chimeric bacterial luciferases. *Biochemistry.* 38:13820-13828.

Vollmer, A. C., S. Belkin, D. R., Smulski, T. K. VanDyk and R. A. LaRossa. 1997. Detection of DNA damage by use of *Escherichia coli* carrying a *recA'::lux, uvrA'::lux,* or *alkA'::lux* reporter plasmids. *Applied and Environmental Microbiology.* 63:2566-2571.

Waddle, J. J., T. C. Johnston, T. O. Baldwin. 1987. Polypeptide folding and dimerization in bacterial luciferase occur by a concerted mechanism *in vivo. Biochemistry*. 26(16):4917-4921.

Wallace, W. H., J. T. Fleming, D. C. White, and G. S. Sayler. 1994. An algD-bioluminescent reporter plasmid to monitor alginate expression in biofilms. Microbial Ecology. 27; 224-240.

Watanabe, H., J. W. Hastings. 1982. Specificities and properties of three reduced pyridine nucleotide-flavin mononucleotide reductases coupling to bacterial luciferase. *Molecular and Cellular Biology*. 44(3):181-187.

Watson M. A., C. Darrow, D.B. Zimonjic, N.C. Popescu and T.P. Fleming. 1998. Structure and transcriptional regulation of the human mammaglobin gene, a breast cancer associated member of the uteroglobin gene family localized to chromosome 11q13. *Oncogene.* 16: 817–824.

Welches, W. R., T. O. Baldwin. 1981. Active center studies on bacterial luciferase: modification of the enzyme with 2,4-dinitrofluorobenzene. *Biochemistry*. 20(3):512-517.

Wilkins, E. and P Atanasov. 1995. Integrated implantable device for long-term glucose monitoring. *Biosensors and Bioexlectronics*. 10:485-494.

Wilson, J. E., T. V. Pestova, C. U. Hellen and P. Sarnow. 2000. Initiation of protein synthesis from the A site of the ribosome. *Cell*. 102:511-520.

Wilson, T. and W. Hastings. 1998. Bioluminescence. *Annual Review of Cell and Developmental Biology.* 14:197-230.

Wood, K. V.1995. Marker proteins for gene expression. *Current Opinions in Biotechnology.* 6:50-58.

Wang, J., Z. Chen, X. Xia, D. Lu, J. Xue, C. Ruan. 2002. A bicistronic retroviral vector to introduce drug resistance genes into human umbilical cord blood CD34+ cells to improve combination chemotherapy tolerance. *Chinese Medical Journal (English)*. 114(1):25-29.

Yates, J. L., N. Guan. 1991. Epstein-Barr virus-derived plasmids replicate only once per cell cycle and are not amplified after entry into cells. *Journal of Virology*. 65(1):483-488.

Zhou, W., Edelman, G.M. & Mauro, V.P. 2001. Transcript leader regions of two *Saccharomyces cerevisiae* mRNAs contain internal ribosome entry sites that function in living cells. *PNAS* U. S. A. 98:1531-1536.

Zhang, J., R. Kuvelkar, B. Cheewatrakoolpong, S. Williams, R.W. Egan and M.M. Billah. 1997. Evidence for multiple promoters of the human IL-5 receptor α subunit gene. A novel 6-base pair element determines cell-specific promoter function. J*ournal of Immunology*. 159:5412–5421.

Zhang, T., V. Kruys, G. Huez, C. Gueydan. 2002. AU-rich element-mediated translational control: complexity and multiple activities of trans-activating factors. *Biochemical Society Transactions.* 30(6):952-958.

Ziegler, M. M., A. C. Clark, J. F. Sinclair, A. F. Chaffotte, M. E. Goldburg and T. O. Baldwin. 1991. Folding and assembly of the subunits of bacterial luciferase. *Bioluminescence and Chemiluminescence*. 1991:178-182.

Ziegler, M. M., M. E. Goldberg, A. F. Chaffotte, T. O. Baldwin. 1993. Refolding of luciferase subunits from urea and assembly of the active heterodimer. Evidence for folding intermediates that precede and follow the dimerization step on the pathway to the active form of the enzyme. *Journal of Biological Chemistry.* 268(15):10760-10765.

VITA

Stacey Swiger Patterson was born in Clarksburg, WV on July 23, 1974. She grew up in Salem, WV and attended Harden Grade School, Salem Junior High and Liberty High School. She moved to Knoxville, TN in January 1991 and finished high school at West High School in Knoxville, TN. She entered The University of Tennessee, Knoxville as a freshman in the fall of 1991. She received a Bachelor of Science in Biology in August 1995.

After working for one year, she began her graduate studies at East Tennessee State University in the Department of Environmental Health Sciences in the fall of 1996. Under the direction of Dr. Phillip Scheuerman, her thesis work focused on the detection of the protozoan pathogens, *Cryptosporidium* and *Giardia* in rural Tennessee. She received a Master of Science in Environmental Health in May 1999. She continued her graduate education at The University of Tennessee, Knoxville at the Center for Environmental Biotechnology in May 1999 with Dr. Gary Sayler.

The author will begin working as a post-doctoral fellow at The University of South Florida in the College of Marine Science in St. Petersburg, FL in October, 2003.