TECHNOLOGIES FOR IMAGING WITH BIOLUMINESCENTLY LABELED
PROBES

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF BIOENGINEERING

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Andreas Markus Loening

June 2006

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

Sanjiv Sam Gambhir   Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

Christopher H. Contag

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

Jennifer R. Cochran

Approved for the University Committee on Graduate Studies.

Abstract

Luciferases, which have seen expansive employment as reporter genes in biological research, could also be used in applications where the protein itself is fused or conjugated to ligands in order to create bioluminescently-labeled imaging probes. In the context of small animal imaging, bioluminescent labeling can be expected to yield greater sensitivity than more traditional fluorescent or radioisotopic labeling approaches. This dissertation focuses on developing the underlying technologies to enable bioluminescent labeling as a routine methodology.

For a variety of practical reasons, the luciferase from *Renilla reniformis* is generally the most appropriate for use as a bioluminescent label. However, the native enzyme is overly labile in serum, and a consensus sequence driven mutagenesis screen was employed to improve its properties. The result of this mutagenesis was an 8 mutation variant of *Renilla* luciferase (RLuc8) that, compared to the parental enzyme, was 200-fold more resistant to inactivation in murine serum and exhibited a 4-fold improvement in light output. Also generated were variants optimized for use as reporter genes that showed 5-fold greater light output while exhibiting greater responsiveness to transient gene expression. An additional impediment with *Renilla* luciferase is that its bioluminescence emission spectrum is not ideal for *in vivo* imaging. Through a combination of random mutagenesis and site-directed mutations in the substrate-binding pocket, variants of *Renilla* luciferase with up to 60 nm red-shifts were created that are more optimal for imaging applications.

To further facilitate the analysis of *Renilla* luciferase, its crystallographic structure was determined to a resolution of 1.5 angstroms. This represents the first structure of a coelenterazine-dependent luciferase, and should aid future studies involving this enzyme. Additional work presented that extends beyond bioluminescent labeling includes: AMIDE, an open source software tool for displaying and analyzing multimodality volumetric image data sets; the crystallographic structure of *Renilla reniformis* green fluorescent protein; and bacterial expression and analysis of *Gaussia princeps* luciferase.

Finally, as an example implementation of a bioluminescently-labeled imaging probe, an epidermal growth factor/luciferase fusion protein was created and its utility analyzed both *in vitro* and *in vivo*. In conclusion, the work presented in this dissertation advances bioluminescent labeling as a practical addition to the molecular imaging toolbox.

Acknowledgments

First and foremost, I'd like to take this opportunity to thank my adviser Dr. Sanjiv Sam Gambhir, who's support I'll be eternally grateful for. Sam took a gamble when he allowed me to begin working in his lab before I started medical school at UCLA, and he's been putting up with that fateful decision for the past 7 years. In all that time, Sam's always given me both the resources as well as the independence to pursue a multitude of projects of interest to me, the more successful of which have become this dissertation.

In addition to Sam, many other people have given advisory support over the past couple of years and deserve acknowlegement. My dissertation committee has been a great help, and I've been appreciative of Dr. Chris Contag's knowledge of all things concerning bioluminescence and Dr. Jennifer Cochran's expertise with both EGF as well as random mutagenesis. And of course, how can I ever thank Dr. Anna Wu enough. Anna gave me countless hints, pushes, and prods on everything ranging from protein engineering to being a real scientist. If it were not for her crucial little nudges of encouragement at critical times during this work, very little of this dissertation would ever have come to fruition.

I've had the great opportunity to collaborate with and learn from Tim Fenn during the course of the work covered within Chapter 4 and Appendix B. Dr. Dennis O'Kane provided advice and Jelena Levi assisted with the luminometer calibration presented in Appendix A. I need to thank Dr. Bruce Bryan, who has been extremely giving with both his knowledge as well as his reagents (plasmids, purified protein, substrates, bioluminescent squirt guns, bioluminescent lollipops, etc.). In addition to providing much needed training in the beginning of the work, Tove Olafsen is also responsible for the ^{124}I labeling work presented in Chapter 6. The friendly aid of Drs. Brad Rice, Tamara Troy, and Olivier Coquoz of Xenogen Corporation made possible the bioluminescence spectral measurements presented in Chapters 2 and 3. Zac Walls served as grammar hound and provided many helpful comments on this manuscript. Additional people have guided me during my training, and I'd like to thank Tim Doyle, Mark

Breidenbach, Steve Kaiser, Pavel Strop, Arion Chatziioannou, Andrew Goertzen, David Stout, Khoi Nguyen, and Meera Iyer. I'd also like to thank Elizabeth Gill and all of the members of the Gambhir lab for tolerating me for the past couple of years.

I'd like to thank my parents, both for instilling in me a healthy appreciation for hard work and for being patient as I've postponed my medical career to do this research thing. Finally, I'd like to acknowledge various friends over the years that have helped me maintain my sanity. Greg, Vicki, Zollars, Yee-li, Susy, Michelle, Zac, Mark, Kaiser, Tim, Olivier, Julia, and countless others, thank you all.

I'll leave you now, with several pearls of wisdom that I've accumulated over the years.

- "We gotta beat the morals out of you" - Sandip Biswal

- "Twenty years ago, graduate students were happy if they were not beaten" - Frank Berger

- "I'm busy, I have to fight cancer" - Frank Berger

- "That's because, inside you're already dead" - Z. Floyd Walls

- "Well of course you can't understand your data, you have too many controls" - Anna M. Wu

- "Thanks to [my adviser], who gave me enough rope to hang myself and then stood by patiently as I gasped and flailed about in the early stages of this work." - Thomas Quinn

Colophon

This manuscript was written as a collection of LaTeX [107] files, making use of the suthesis-2e.sty style file written by Joseph Pallas and maintained by Emma Pease. Portable Document Format (PDF) output was made from the LaTeX source files using pdfLaTeX. The document was typeset with 10 point New Century Schoolbook for text, and 10 point Times for equations.

The sequence alignments shown in Figures 2.3 and 2.4 were made using the software tool Jalview [41] and editing the resultant postscript file by hand. Other sequence alignments (e. g. Figure 3.7) were made using TeXshade [16].

Graphs were made using GNU Octave [54], diagrams were made using XFig [203], chemical structures were drawn using gchempaint [25], protein structural figures were made using PyMOL [49], and protein topology cartoons were made using TopDraw [21]. The R1924P quantum efficiency data and the luminol chemiluminescence emission spectrum utilized in Appendix A were digitized from their respective source data using Engauge [146].

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Molecular imaging is an emerging field at the intersection of engineering, biology, and chemistry, that seeks to non-invasively interrogate the biological processes underlying complex diseases and physiological conditions. Its targets range from gene expression to metabolic pathways, and it utilizes a variety of tools within a plethora of modalities (e.g. Positron Emission Tomography, Magnetic Resonance, Bioluminescence Imaging).

One of the best known examples of molecular imaging in the research setting, is the use of reporter genes to indirectly monitor levels of gene expression in living subjects [43, 135]. In reporter gene imaging, a genetic construct is designed such that the reporter's expression and resultant activity levels will track the expression of a gene of interest. This construct is then transferred into the target cells through a variety of techniques, and monitored using the appropriate imaging modality (e.g. bioluminescence imaging for a luciferase reporter gene). In practice, researchers are usually interested in target protein levels rather than gene transcription levels, and reporter gene imaging represents a circuitous route for monitoring a protein that may not accurately track the desired variable.

In order to truly monitor protein levels, unique imaging probes must be developed to detect each specific protein. Traditionally these probes have been constructed by coupling a radioactive or fluorescent moiety to a ligand that is known to bind to the protein of interest. A new concept is to use a luciferase as a bioluminescent label in place of the more traditional imaging moieties. As discussed below, the advantage of using a bioluminescent label over similar fluorescent or radioactive approaches, is that in the context of small animal imaging the bioluminescence approach has the potential to be more sensitive [242].

(a) Ligand/Luciferase   (b) Antibody/Luciferase
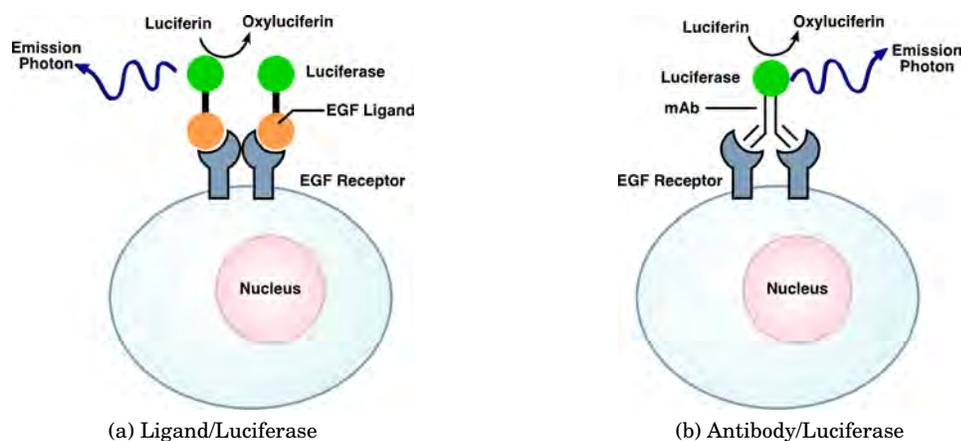
Figure 1.1: Schematic diagrams of bioluminescent imaging probes. Panel (a) shows an example of a light emitting epidermal growth factor (EGF)/luciferase fusion protein. Panel (b) shows a similar scheme using an antibody in place of a native receptor ligand.

The focus of this dissertation is on developing the underlying technologies needed to make bioluminescent labeling a reality.

1.1   Concept of Bioluminescent Labeling

A schematic explaining the basic concept of a bioluminescently labeled imaging probe is shown in Figure 1.1. In such an approach, a fusion would need to be made between a luciferase and a ligand specific for the receptor of interest. A variety of ligands appropriate for bioluminescent labeling have been suggested, including engineered antibodies [218, 216] as well as the small protein growth factors that are the native ligands for various receptors [125, 169, 168]. Following construction, these luciferase/ligand fusion proteins could be used for studying receptor targets both in *in vitro* assays and *in vivo* imaging. For the imaging case, such studies would be conducted by injecting the fusion imaging probe into the subject and allowing it time to distribute. The imaging probe could then be detected at multiple later time points by injecting the substrate for the luciferase and imaging the subject.

1.2   Comparison of Bioluminescent and Traditional Labels

The primary reason for developing luciferases as bioluminescent labels is the increase in sensitivity they may yield, especially in the context of small animal imaging.

Bioluminescence has been noted to be more sensitive than radioactivity in a number of *in vitro* [140] and *in vivo* [242] assays. As an example for the *in vitro* case, ^{125}I assays have sensitivities in the range of 10 amole (10×10^{-18} mole) [140]. This is ∼3 orders of magnitude worse than the lower limit of ∼0.1 amole found for an unoptimized assay based on *Renilla* luciferase (Figure A.4), and ∼4 orders of magnitude worse than the ∼0.01 amole sensitivity of commercial firefly luciferase assays [187].

As a sensitivity comparison for the *in vivo* case, imagine a million ligands labeled with ^{111}In that are directed to a particular receptor [180]. Assuming a high specific activity (44,000 MBq/-μmol) [180], if all million probes bound to their target receptor they could be expected to yield ∼4 photons over the course of a one minute image acquisition. A similar probe labeled bioluminescently, such as discussed in Chapter 6, would give off ∼1 million photons in the same interval at subsaturating levels of substrate (10% of peak output). This is over 5 orders of magnitude more photons than the radioactive case. Of course, the optical wavelength photons of bioluminescence are much more greatly attenuated in biological tissues than the ^{111}In spawned gamma rays, but at the superficial locations normally utilized in small animal research models this attenuation is insufficient to annul the 5 orders of magnitude greater photon flux from bioluminescence.

Imaging with bioluminescent labels would have other advantages over radioactivity based approaches. For one, at many imaging centers equipment appropriate for small animal bioluminescence imaging is more readily available than equivalent instrumentation for studies utilizing radioactivity. Additionally, in bioluminescence the researcher has the capability of reinjecting substrate, allowing high photon fluxes at all imaging times. Radioactivity based imaging is fundamentally limited by the radioactive decay half-life of the isotope in use. A final advantage of bioluminescence is that radioactivity and associated radioactive wastes are a nuisance to work with. While possible hazards with the use of radioactivity are generally overblown, the regulatory hurdles erected to ensure the safety of workers have achieved their goals by making the use of radioactivity decidedly unattractive.

Comparing bioluminescence and fluorescence is a harder conjecture. Few head-to-head comparisons have been made between these two modalities, and results depend crucially on the particular assay being performed as well as the equipment in use. In the case of small animal imaging, bioluminescence has historically won for two reasons. First, fluorescence requires an excitation photon of light to go in as well as the emission photon to travel out, effectively doubling the path length over which photons must travel compared to bioluminescence. Second,

all those excitation photons have the ability to excite endogenous fluorophores, and the resulting background autofluorescence strongly limits the sensitivity of fluorescence techniques. As an example of the advantages of bioluminescence, a comparison of fluorescent quantum dots and bioluminescent quantum dots indicated that the bioluminescent versions could be as much as 10^5-fold more sensitive due to the reduced path length issue alone [194]. Another example, using commonly available imaging equipment and comparing commonly available bioluminescent (firefly luciferase) and fluorescent (DsRed) reporter genes appropriate for *in vivo* imaging, found the bioluminescent reporter was 30 to 750-fold more sensitive [214].

As a final example, a rough calculation is illustrative for a comparison between a bioluminescently labeled imaging probe and one labeled with a near-infrared imaging dye. For this case, we'll compare the commercially available fluorophore Cy7 to the stabilized *Renilla* luciferase variant developed in Chapter 2. Imagine 1 million Cy7 labeled probes, lying at a depth of 0.5 cm of rat liver tissue and spread over an area of 1 cm^2. A standard continuous wave fluorescence imaging system can be expected to emit $\sim 1 \times 10^{13}$ excitation photons/s over this amount of area [214]. Assuming an excitation wavelength of 715 nm, 5×10^{11} photons/s will penetrate to this depth (Figure 3.17). Given that the extinction coefficient for Cy7 is 250,000 M^{-1} cm^{-1} and the quantum yield is 0.28, it can be calculated that ~ 60 photons/s will result from fluorescence. In turn, 1 million of the stabilized *Renilla* luciferase variant operating at just 10% of the enzyme's peak output value (Section 2.2.3) will output 3.2×10^4 photons/s (this is of course assuming the luciferase is in a physiological compartment accessible to its substrate). Even though just 1% of these bioluminescent photons will be in the desired "optical window" (> 650 nm) for tissue penetration (Figure 2.7), the bioluminescently labeled imaging probe will still be 5-fold brighter than the Cy7 labeled probe. And this is all before taking into account the autofluorescence background signal resulting from endogenous chromophores that will further limit the sensitivity of the fluorescent imaging probe. Additionally, the red-shifted variants of *Renilla* luciferase developed in this dissertation (Chapter 3) will lead to a several-fold increase in the precentage of bioluminescent photons in the "optical window", further increasing the advantage of bioluminescence versus fluorescence.

There is, of course, a major disadvantage to bioluminescent labeling that radioisotopes and for the most part fluorophores are immune to; namely, probe stability. Bioluminescent labeling necessarily utilizes a luciferase as the imaging moiety, and the resistance of the protein to inactivation in the environments to which it is exposed is of paramount concern. As will be shown in Chapter 6, this can be the limiting factor for a bioluminescently labeled imaging probe. An

additional caveat to bioluminescent labeling is that luciferases are of much greater size than radioisotopes and most conventional fluorophores, with the result that bioluminescent labeling will have a greater propensity to interfere with the activity and distribution of the labeled item than a fluorescent or radioactive moiety would.

1.3   Choice of Luciferase

Choosing an appropriate luciferase to use for bioluminescent labeling is mostly an act of exclusion. The first selection criterion is that the gene for the luciferase needs to be available. This is due to the expectation that bioluminescent labeling will often be achieved by creating fusion proteins at the genetic level.

The luciferase from the North American firefly *Photinus pyralis* is both the first eukaryotic luciferase cloned [233] and the most commonly used luciferase in molecular biology. In addition to firefly luciferase a number of other beetle luciferases have been cloned, most notably the green and red click beetle luciferases from *Pyrophorus plagiophthalamus* [240]. These beetle luciferases are commonly used in molecular imaging with great success [43, 246], but they are not optimal for employment as bioluminescent labels. First, these luciferases are not particularly small (∼62 kDa). More importantly, in addition to their substrate they are dependent on ATP, molecular oxygen, and magnesium for activity. This dependence on ATP would critically hinder the application of beetle luciferases as bioluminescent labels *in vivo*, as serum ATP concentrations are generally below 10 nM [243].

Other classes of luciferases that have been cloned include the bacterial, dinoflagellate, *Vargula* luciferin utilizing, and coelenterazine utilizing luciferases. Bacterial luciferases were quickly dismissed as a viable option for bioluminescent labeling, as the active luciferases are actually heterodimers of two proteins [236].

Dinoflagellate luciferases, in turn, were a relative unknown when the work here started and were not serious entertained. Recently, a 46 kDa active fragment of dinoflagellate luciferase has been developed [149] and utilized as a reporter in mammalian cells [204]. Future workers may wish to examine this enzyme, although its applicability to small animal imaging has not yet been tested.

The *Vargula* luciferin utilizing luciferases consist of two highly similar proteins cloned from the Ostracod species, *Vargula hilgendorfii* [209] and *Cypridina noctiluca* [153]. They were removed from consideration partly due to their size (61 kDa), but mainly because their luciferin

is both extraordinarily expensive and atrociously unstable.

This process of exclusion left for consideration the luciferases that utilize coelenterazine (a type of imidazolopyrazine). These luciferases are generally appropriate for use as bioluminescent labels as they are not ATP dependent, and typically require only molecular oxygen and their substrate for luminescence. A handful of these coelenterazine luciferases have been cloned. The most attractive from this class, due to its 18 kDa size and high enzymatic activity, is the luciferase from *Gaussia princeps*. This luciferase however, is relatively new, inadequately characterized, and has been recalcitrant to production in bacteria (as shown in Appendix C). Similarly, the luciferase from *Oplophorus gracilirostris* is attractive due to its 19 kDa size [89], but it expresses poorly in bacteria, it has not been well studied, and it may potentially exist as a dimer.

This process of exclusion leaves us with the luciferase from the sea pansy *Renilla reniformis*. This luciferase has been well characterized [85, 86, 138, 139, 71, 128, 120], is available commercially as both a gene and as protein, exists as a monomer, is of an appropriate size for use as a label (36 kDa), and has been successfully produced in bacteria [127]. Additionally, existing work [119] has already identified a variant with enhanced resistant to inactivation at 37°C.

1.4 *Renilla reniformis* and its Luciferase

Renilla reniformis, also known as the "sea pansy', is an animal that consists of a sessile collection of polyps (Figure 1.2) and lives along the South Atlantic coast of the United States of America. When disturbed, it reacts with waves of bright green bioluminescence. This bioluminescence display is mediated by a primitive neural network and is thought to be a mechanism for warding off predators.

The light emitting cells (photocytes) of *Renilla reniformis* are gastrodermal cells containing large membrane-bound intracellular structures termed luminelles [196]. These luminelles, in turn, contain large numbers of small (0.2 μm average diameter) membrane-bound particles called lumisomes [5] that can be stimulated to bioluminesce by the application of calcium (Ca^{++}) [6]. The lumisomes house the three proteins most important for bioluminescence in *Renilla reniformis*: a luciferase (RLuc[1]), a green fluorescent protein (RrGFP), and a Ca^{++} activated luciferin binding protein (RrLBP).

---

[1]It would be more appropriate to label this luciferase RrLuc, to distinguish it from the almost identical luciferase from *Renilla mülleri* (RmLuc) that is also now commercially available. However, to remain consistent with the previous

(a) Superior View (b) Closeup
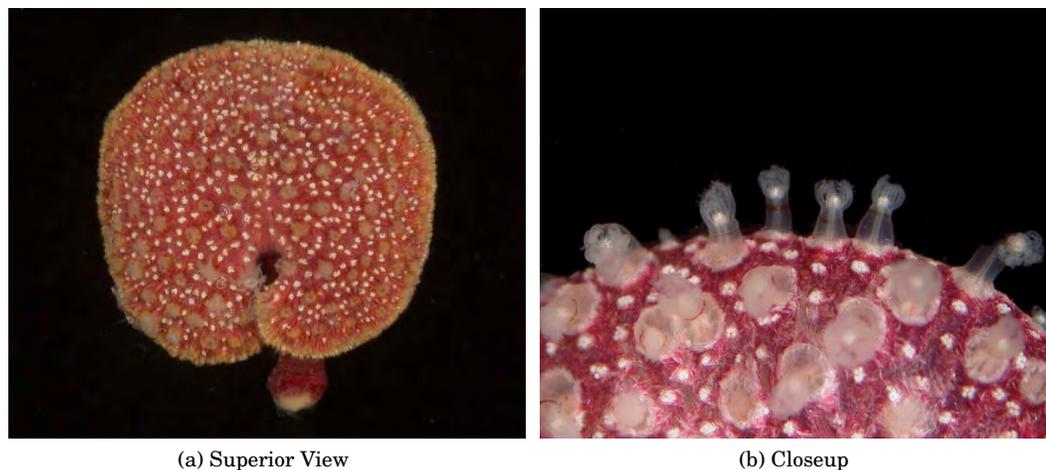
Figure 1.2: Photographs of *Renilla reniformis*. Panel (a) is a superior view of the animal, showing the surface that is exposed to the marine environment. The peduncle, a specialized polyp just visible toward the bottom of the picture, is used to anchor the animal in the sand. Panel (b) is a closeup of the superior surface of the animal, showing the individual polyps that compose the animal. These photos are used by permission of the Southeastern Regional Taxonomic Center/South Carolina Department of Natural Resources.

RLuc was first purified to homogeneity in the mid 1970's [138], with the associated gene for the luciferase cloned in 1991 [128] (GenBank Identifier GI:126502). RLuc has 311 amino acids, is 36 kDa in size, contains a characteristic α/β-hydrolase fold sequence [82], and is highly similar to a number of bacterial haloalkane dehalogenases (see Chapter 2).

Interestingly, in the animal it is not RLuc that is the light emitter for bioluminescence but rather RrGFP (GenBank Identifier GI:14161475). The energy released by RLuc mediated coelenterazine degradation is passed via resonance energy transfer to RrGFP and is emitted as a green photon of light (Discussed further in Appendix B). In *Renilla*, the luciferase is always associated intracellularly with the GFP [196].

The coelenterazine for the bioluminescence reaction is stored by the Ca^{++} triggered protein RrLBP [103] (GenBank Identifier GI:125988). RrLBP is a 184 amino acid long protein, and is a member of the EF-hand superfamily of Ca^{++}-binding proteins. The EF-hand superfamily also includes Aequorin [73], although RrLBP and Aequorin do not show much similarity [103] (50% similarity over 50% of the protein).

The final bioluminescence associated protein that has been studied from *Renilla reniformis* is the luciferin sulfokinase. In *Renilla reniformis*, the coelenterazine is stored as a sulfate

literature the abbreviation RLuc will be maintained. Similarly, the use of the expression "*Renilla* luciferase" in this dissertation implies *Renilla reniformis* luciferase.

derivative [84]. Luciferin sulfokinase is used to convert this luciferyl-sulfate storage form to the active substrate coelenterazine, which in turn allows recharging of the luciferin binding protein [45].

Its interesting to think how much biology has changed in the interleaving decades between Matthews' initial purification of RLuc in the mid 1970's [137] and the work performed for this manuscript. Matthews's protocol, employing standard biochemistry techniques of the day, started with 5 kg of animal and would yield $\sim$5 mg of pure RLuc following a month of work. In this manuscript, modern biological techniques allowed purification of greater amounts of luciferase from less than 100 ml of bacterial culture in a day.

1.5 Epidermal Growth Factor Receptor as a Model System

As a model system for the validation of a bioluminescently labeled imaging probe, the choice was made to focus on studying the epidermal growth factor receptor (EGFR) through the development of epidermal growth factor ligand (EGF)/luciferase fusion proteins.

The choice of EGF/EGFR was made for a variety of reasons, included the following. Active forms of EGF are known to be relatively stable (on the order of hours) in serum at 37°C [7], an obvious necessity for an imaging probe that will be injected into the bloodstream. EGF is a relatively small (6 kDa) protein ligand; useful from the standpoint of biodistribution kinetics as well as simplifying cloning and protein expression. The existence of previous work in which EGF has been fused to a variety of other constructs without compromising the ligand's binding affinity [244, 226] was considered auspicious. EGF is known to readily express in bacteria using a variety of methods [244, 112, 213]. Well established cell lines exist that either overexpress EGFR (A431 cells) or show minimal EGFR expression (NIH 3T3 cells).

Finally, EGFR makes for an interesting target due to its clinical relevance. This receptor is overexpressed in a large number of cancers [150], at the level of 100-fold overexpression in >30% of human breast cancers [99]. As such, EGFR has recently become a target for cancer therapy, with the development of novel anti-tumor drugs such as cetuximab (Erbitux) and gefitinib (Iressa). One can imagine that a bioluminescently labeled EGFR probe would be a useful tool for studying a variety of *in vitro* and *in vivo* cancer models.

1.6 Outline of the Dissertation

This dissertation consists of several projects that are superficially disparate. The following outline is a guide to the Chapters and Appendices that follow, and gives some explanation of how the different elements fit together.

Chapter 2 addresses one limiting factor to the use of RLuc as a bioluminescent label, namely, its rapid ($\tau_{1/2}$=0.5-1 hour) inactivation in murine serum at 37°C (see Table 2.1). A single point mutation of RLuc (C124A) that increases the enzyme's resistance to inactivation approximately 6-fold has been reported [119]. However, even this level of resistance is insufficient for the tagging of large proteins (e.g. antibodies) that may require time scales on the order of days to sufficiently distribute. For this reason, a semi-rational, consensus sequence driven mutagenesis strategy [113] was pursued to identify mutations that could increase the resistance of RLuc to inactivation or increase the enzyme's light output. By combining these mutations, a variant of RLuc (RLuc8) was created with properties appropriate for use as a bioluminescent label. Compared to the native enzyme, RLuc8 exhibited a 200-fold enhancement in resistance to serum inactivation as well as a 4-fold increase in light output.

Chapter 3 covers mutagenesis studies performed on RLuc8 that were mainly focused on altering its bioluminescence emission spectrum. As blue wavelength light is strongly attenuated in biological tissues, a red shifted variant of *Renilla* luciferase (normal peak of 481 nm) would be a great advantage in imaging studies. Through a combination of rational and random mutagenesis approaches, the bioluminescence spectrum of the luciferase was successfully red shifted ∼50 nm.

Chapter 4 describes collaboratory work performed with Dr. Timothy Fenn on generating a crystallographic structure of *Renilla* luciferase. This structural information should allow future rational mutagenesis approaches to further improve the properties of *Renilla* luciferase for *in vivo* imaging.

Chapter 5 describes a software tool called AMIDE that was developed for multimodality medical imaging. One can easily imagine situations where it would be advantageous to combine the acquired data from the use of a bioluminescently labeled probe with other imaging modalities such as computerized tomography (CT) and positron emission tomography (PET). As a dearth of such software was available, a new application was written to facilitate the merging and analysis of such multi-modality imaging data.

Chapter 6 describes an attempt to demonstrate the creation and use of a bioluminescently

labeled imaging probe directed at EGFR. Although the probe was functional in a variety of *in vitro* assays, it was eventually proven to be unsuccessful for *in vivo* imaging purposes.

Chapter 7 concludes the dissertation by presenting instances were the stabilized *Renilla* luciferase variant (RLuc8) has proven successful as a bioluminescent label in the imaging of small animal research models. It also highlights some future research paths that would further aid the development of bioluminescent labeling technology.

In addition to the above chapters, several appendices are included with information less central to the main thrust of the dissertation. Appendix A covers calibration of a standard luminometer to allow converting the data to absolute units. Appendix B presents the structure of the *Renilla reniformis* green fluorescent protein, which was produced as an offshoot of the work done in Chapter 4. Finally, Appendix C describes work done on *Gaussia* luciferase in order to produce this protein successfully in bacterial expression systems.

Chapter 2

Optimization of *Renilla* Luciferase for Serum Conditions

The work in this chapter owes its existence to a conversation during a beer social, in which I was trying to pique Mark Breidenbach's interest in crystallizing *Renilla* luciferase (RLuc). Although little of productive merit was achieved on that particular evening, Mark did later run the RLuc primary sequence through the PredictProtein server [181] and concluded that the structure of RLuc would not be overly interesting as fairly similar proteins (haloalkane dehalogenases) had already been crystallized. Later, on pouring through the PredictProtein results, I noticed in the sequence alignment that the cysteine at position 124 of RLuc was generally an alanine in the similar proteins. This deviation from the consensus was interesting, as work by Liu and Escher [119] on a C124A mutant (denoted C152A in Liu and Escher) had already shown that the mutation enhances RLuc's resistance to inactivation.

The authors behind the C124A mutation had come upon it while sequentially mutating the three cysteines in *Renilla* luciferase to alanine, in an attempt to engineer a version of RLuc appropriate for mammalian cell secretion. This single point mutation increases RLuc's resistance to inactivation to a $\sim$7 h half-life for the condition of murine serum at 37°C (Table 2.1). However, even this level of resistance is insufficient for the bioluminescent tagging of large proteins (e. g. antibodies) that may require time scales on the order of days to sufficiently distribute. For this reason, RLuc variants with even greater stability[1] were desired.

---

[1]The word "stability" is generally used in this manuscript to connote an enzyme's resistance to inactivation in serum. This differs from the more common connotation of the word, in which it implies thermodynamic stability.

Based on the alignment data, I speculated that one interpretation of the C124A results was that alanine is the preferred amino acid for this location within the tertiary fold of the enzyme and results in better thermodynamic stability of the protein's hydrophobic core [118]. This hypothesis was predicated upon the idea that mutations have accumulated in the RLuc gene that are either neutral only in the specific context of the *Renilla* lumisomes, or required for the association with and resonance energy transfer to the green fluorescent protein (GFP) homo-dimer that is the normal light emitter in *Renilla* [228]. This interpretation led to a mutation strategy, in which candidate mutations would be picked at locations where RLuc clearly diverged from the consensus sequence and then screened for their contribution to inactivation resistance.

In hindsight, this mutagenesis strategy is not at all novel. A number of previous studies have utilized similar consensus guided semi-rational mutagenesis strategies [199, 237, 113, 4], albeit with differing experimental methods utilized as measures of stability (e. g. thermal denaturation, chemical denaturing, proteolytic cleavage). The common rationale behind all these studies, equally applicable to the study here, is that evolution tends to disfavor amino acids that destabilize a protein [96]. By looking at a family of similar proteins, one can pick candidate mutations that have already been screened by nature to be tolerated within the context of the protein fold, and are therefore less likely to be deleterious to the protein than a residue picked at random. These point mutations can then be screened and, if viewed as favorable in the context of the particular study, be combined to a create an *über*-protein that achieves the desired properties.

In this chapter, a semi-rational, consensus sequence driven mutagenesis strategy [113] is pursued in order to identify mutations that increase the resistance of RLuc to inactivation in serum. Periplasmic expression of the proteins is used here, and it was chosen for the simple reason that it was already being used for the work discussed in Chapter 6.

2.1 Methods

2.1.1 Materials

Coelenterazine was from Prolume (Pinetop, AZ). Benzyl-coelenterazine (coelenterazine-*h*) was a generous gift from Dr. Bruce Bryan. Coelenterazine-*n* and coelenterazine-*cp* were from Biotium (Hayward, CA). Bisdeoxycoelenterazine (also known as coelenterazine-*400a*, di-dehydro coelenterazine, or DeepBlueC) was from Perkin Elmer (Boston, MA). The chemical structures
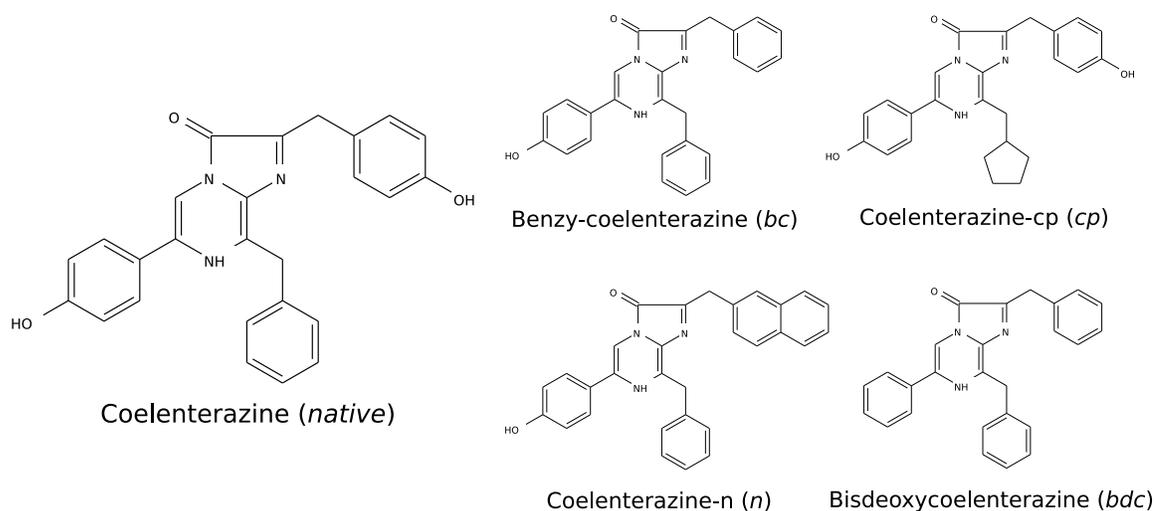
Figure 2.1: Chemical structures of coelenterazine and several analogs. Abbreviations used in the text are given in parenthesis.

of these compounds are shown in Figure 2.1. Coelenterazine and the analogs were dissolved in propylene glycol at 0.5 μg/ml and stored in small aliquots at -80°C.

2.1.2 Luminometer Calibration

Light measurements were made using a Turner 20/20 and later a Turner 20/20n luminometer (Turner Biosystems, Sunnyvale, CA). The luminometers were calibrated to absolute units (photons/s) using the luminol light standard performed in dimethyl sulfoxide (DMSO) [110, 165]. This calibration is described in detail in Appendix A. As the spectral peak of luminol chemiluminescence in DMSO (486 nm) is close to the spectral peak of *Renilla* luciferase bioluminescence with coelenterazine (481 nm), no corrections were applied for the spectral sensitivity of the luminometer when assays were performed using coelenterazine.

2.1.3 Computational Prediction

A PSI-BLAST search [2], performed in December 2003 using the PredictProtein server [181], identified a number of sequences similar to RLuc. An alignment between RLuc and the 9 most similar sequences ($\geq 46\%$ similarity) was then generated using CLUSTAL W [211] .

   A later BLAST search [2, 1] was performed in February 2006 using the nonredundant ("nr") subset of the GenBank database, all sequences with an E value of $< 1^{-10}$ were initially retained. Redundant sequences were then removed, as well as sequences that did not contain a

conserved catalytic triad (corresponding to D120, E144, H285 in the RLuc sequence) known to be required for activity in bacterial haloalkane dehalogenases. Following this, an alignment between RLuc and the 14 remaining sequences was generated using CLUSTAL W.

A homology model of RLuc was built with SWISS-MODEL (v3.5) [185] using the default parameters (Figure 2.5). In generating this homology model, SWISS-MODEL utilized several crystal structures of the haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* (PDB files 1iz8, 1k63, 1k6e, 1iz7, and 1mj5) [201, 161].

2.1.4   Construction of *Renilla* Luciferase Mutants

The gene *hrluc* from the plasmid phRL-CMV (Promega, Madison, WI) was used as the initial template for cloning. This gene is a human codon usage optimized version of *rluc*, and encodes a protein identical to RLuc with the exception of a T2A substitution. To construct a bacterial expression plasmid, PCR was used to replace the stop codon with a HindIII restriction site and the N-terminal methionine codon with a pelB leader sequence. The pelB leader sequence, consisting of the first 22 codons of the pectate lyase B gene from *Erwinia carotovora* [114], directs protein expression into the bacterial periplasm and is cleaved from the final protein product. Using NcoI and HindIII restriction sites, the PCR product was inserted into the pBAD/Myc-His A plasmid[2] (Invitrogen, Carlsbad, CA), which adds a Myc epitope, a 6xHis tag, and a stop codon to the C-terminus of the gene. In the RLuc8 construct described later, as well as mutants based off of it, the plasmid's SalI site was used for insertion in order to remove the Myc epitope from the construct. Site directed mutagenesis in this chapter was performed using a QuikChange II XL kit (Stratagene, La Jolla, CA), please note that a more economical protocol used for later work is presented in Section 3.1.1. When needed, cytoplasmic expression plasmids were generated by reverting the pelB leader sequence back to a methionine codon using PCR. All constructs and mutations were confirmed by sequencing. When plasmids are referenced in this manuscript, the presence of "pelB" in the plasmid name indicates whether the plasmid is the periplasmic (e.g. pBAD-pelB-RLuc8) or cytoplasmic (e.g. pBAD-RLuc8) expressing version.

---

[2]One potential caveat to using an arabinose inducible expression system (e. g. pBAD), is that graded induction using subsaturating levels of arabinose requires special host cells. In contrast to initial reports, the arabinose promoter in wild type *E. coli* cells acts in essentially a binary capacity [147].

2.1.5   Protein Production and Purification

All proteins were produced from the periplasmic (pelB containing) expression plasmids unless otherwise noted. To express proteins periplasmically, plasmid containing *E. coli* LMG 194[3] cells were grown at 32°C in Terrific Broth [182] with 50 μg/ml ampicillin. Cultures were allowed to reach an OD_{600} of 0.7 and then induced by addition of L-(+)-arabinose to a final concentration of 0.2%. 12-14 h later, cells were harvested and the periplasm extracted by osmotic shock. The 12-14 h growth period was chosen entirely for convenience.

The osmotic shock protocol was based on the protocol described by Neu and Heppel [156], with the alterations that the initial wash steps were removed, and that the volumes of sucrose solution and ice cold water used were half of the initial culture volume. Briefly, the culture was pelleted, resuspended in room temperature 20% sucrose, 30 mM Tris-HCl, 1 mM EDTA, pH 8, shaken for 10 min at room temperature, pelleted at $13,000 \times g$ for 10 min at 4°C, resuspended in ice cold water, shaken at 4°C for 10 min, and centrifuged at $13,000 \times g$ for 10 min at 4°C. The supernatant after this step contains the periplasmic fraction.

For the purposes of purification, the periplasmic fraction was brought to the same concentration as the nickel affinity chromatography wash buffer (WB: 300 mM NaCl, 20 mM HEPES, 20 mM imidazole, pH 8) using a 10x stock, and phenylmethylsulphonylfluoride (PMSF) was added to 1 mM. The solution was then clarified by 0.2 μm filtration and passed over a nickel affinity column (Ni-NTA Superflow, Qiagen, Valencia, CA). The 0.2 μm filtration was found to be critical for insuring loading on the nickel affinity resin occurred in a reasonable time period and led to an increased yield of recovered activity. The column was washed with WB and eluted with elution buffer (EB: 300 mM NaCl, 20 mM HEPES, 250 mM imidazole, pH 8). Protein concentration measurements were made using the Bradford assay [24] with human serum albumin (HSA: Baxter Healthcare Corporation, Glendale, CA) as the standard. Aliquots were taken at this point for gel electrophoresis (Figure 2.2). To the remainder of the elution, HSA was added to 1% as a carrier protein. All samples were stored at 4°C.

For those instances in which cytoplasmic expression plasmids were used, the cells were grown as above. After harvesting, the cell pellet was frozen; thawed for 1 h with gentle agitation in WB containing 1 mg/ml lysozyme, 10 μg/ml RNase A, and 5 μg/ml DNase I; sonicated; and centrifuged to remove debris. These lysates were then cleared by 0.2 μm filtration and purified by nickel affinity chromatography as above.

---

[3]LMG 194 cells contain a Δara714 genotype. This deletion of most of the araBAD operon makes the cells unable to utilize arabinose as an energy source, allowing the arabinose inducible promoter in the pBAD plasmid to remain activated by a single dose of arabinose.
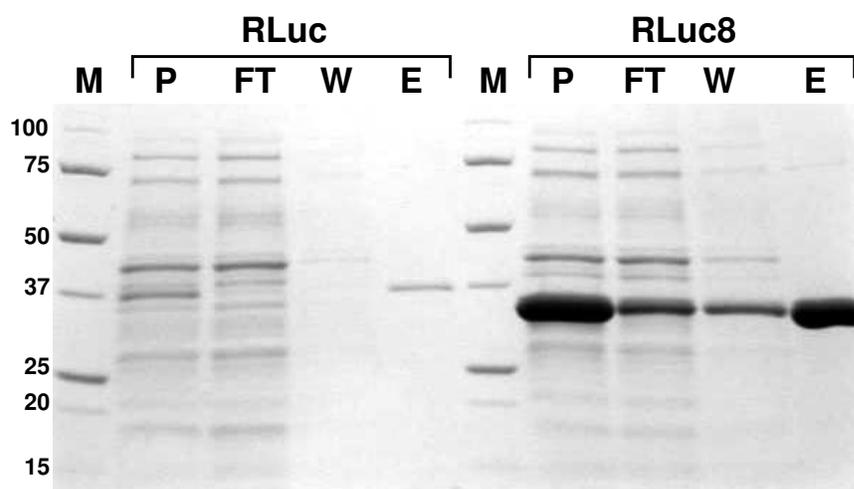
Figure 2.2: Coomassie stained SDS-PAGE gel of RLuc and RLuc8 at several points during the periplasmic purification process. The lanes are labeled as follows: M - Marker, P - Periplasmic fraction, FT - Flow through from nickel affinity column, W - Wash from column, E - Elution from column. As the elution volume was 5% of the periplasmic fraction, the periplasmic fraction, flow through, and wash were concentrated 20-fold using 3 kDa cut-off centrifugal concentrators (Pall, Ann Arbor, Michigan). The expected sizes for RLuc and RLuc8 are 38.7 kDa and 36.9 kDa, respectively, with the difference in size arising from a Myc epitope added by the expression vector used for RLuc. These protein masses were confirmed by MALDI-TOF.

2.1.6 Characterization of *Renilla* Luciferase Mutants

Luciferase activity was measured by adding 1 μl of sample (diluted as necessary in EB containing 1% HSA) to 100 μl room temperature 100 mM sodium phosphate buffer pH 7.0 (NaPB) [195, 68], adding 1 μl of 0.5 μg/μl coelenterazine or analog, manually mixing, and reading for 10 s in a luminometer. The time between the addition of the luciferin and the start of measurement was approximately 4 s. Some notes on the linearity and other aspects of this assay can be found in Appendix A.

Measurements of inactivation in serum were done by mixing 0.5 μl dilute luciferase with either 20 μl mouse serum or 50 μl rat serum (Equitech-Bio, Kerrville, TX), placing the sample in a 37°C incubator, and removing aliquots over time for activity testing. To calculate the serum inactivation half-life, a mono-exponential decay model was fit to the data using a Nelder/Mead Simplex non-linear least squares minimization algorithm provided by the GNU Octave numerical programming language [54]. Emission spectra were measured at ambient temperature in NaPB using a Triax 320 (Horiba Jobin Yvon, Edison, NJ). This device incorporates an optical grating device with a liquid N_2 cooled CCD detector, and as the entire emission spectrum is

acquired concurrently no correction for kinetic decay of the light output needs to be performed. The acquired data were filtered as necessary, and normalized so that the total area under the curve was equalized.

For lyophilization, pure protein (without HSA) was exchanged from EB into lyophilization buffer (50 mM ammonium bicarbonate, pH 7.5) using a PD-10 column (GE Healthcare Life Sciences, Piscataway, NJ), snap frozen in liquid N_2, and lyophilized. Lyophilized protein was then stored at room temperature.

For quantum yield measurements, separate 1 μl drops of protein ($\geq$2 pmole) and substrate (0.2 pmole) were placed in a tube, 100 μl of NaPB was injected by the luminometer to mix, and the total light output was integrated (generally 5-10 min). For coelenterazine-n, the protein amount was increased 10-fold and the acquisition time lengthened to ensure the reaction approached completion.

2.1.7 Kinetics

Kinetics were assessed by injecting 100 μl of NaPB already containing coelenterazine onto 1 μl of protein (diluted appropriately in EB containing 1% HSA), and recording the light output for 20 min. The coelenterazine concentrations tested were at final concentrations of 118, 24, 4.7, 0.94, 0.19, and 0.038 μM. The final luciferase concentrations were in the range of 1-7 pM. Oxygen concentration was assumed to be 0.49 mM [230]. Coelenterazine absorbance was corrected for, although this was only significant for the highest concentration (10% attenuation). The values were converted from photons/s to molecules/s using the data from the quantum yield measurements, converted from flux units to mass units via integration, and processed using the kinetic curve fitting program Dynafit [105].

2.1.8 Mammalian Expression

In order to construct mammalian expression vectors, bacterial expression vectors containing the desired mutations were used as templates for PCR, with primers designed such that a methionine codon replaced the N-terminal pelB sequence and a C-terminal stop codon replaced the Myc epitope and 6xHis tag. The primers also contained appropriate NheI and HindIII restriction sites to allow insertion of the product into the pcDNA 3.1 plasmid (Invitrogen). The resultant plasmids were transiently transfected using SuperFect (Qiagen) into 293T cells [53] or Chinese Hamster Ovarian (CHO, ATCC# CCL-61) cells growing in 24-well plates following

the manufacturer's protocol. Briefly, 0.5 μg/well of DNA was mixed with 5 μl/well SuperFect and 60 μl/well serum-free medium (Dulbecco's modified Eagle medium - DMEM). After 10 min of incubation at room temperature, an additional 350 μl/well of serum supplemented medium (DMEM with 10% fetal bovine serum) was added, and this mixture was applied to wells that had been plated the day before at 50,000 cells/well.

First Round

For the initial mammalian expression studies, the cells were maintained in the transfection medium for 24 h. Following this period, 0.5 μg/well of coelenterazine was added, and the plates were imaged in an IVIS 50 imaging system (Xenogen, Alameda, CA).

Second Round

For the second set of mammalian expression studies, the transfection medium was replaced with fresh medium after 3 h. At several time points following the transfection, cells were lysed using passive lysis buffer (Promega), measured for total protein content using the Bradford assay, and assessed for luciferase activity using coelenterazine in the same manner as described above for bacterially expressed luciferase. Intracellular stability of the luciferases was assessed by adding cycloheximide to the wells at a concentration of 100 μg/ml, and lysing cells at several time points thereafter. Westerns were run on lysates with a monoclonal antibody to RLuc (MAb 4400, Chemicon, Temecula, CA) in order determine the luciferase protein content, with bacterially produced purified RLuc8 used as the standard.

2.2 Results

2.2.1 Computational Predictions for *Renilla* Luciferase

Via sequence similarity searches, RLuc was predicted to contain a characteristic α/β-hydrolase fold from residues 71 to 301 [133]. It was also found to have a high level of similarity to a number of proteins, most of which are known or putative haloalkane dehalogenases. An alignment between the sequences found in a December 2003 database search is shown in Figure 2.3. A later search performed in February 2006 pulled up several new sequences, and the corresponding alignment is shown in Figure 2.4. A homology model of RLuc, based on the crystal structure of the haloalkane dehalogenase LinB, is presented in Figure 2.5.
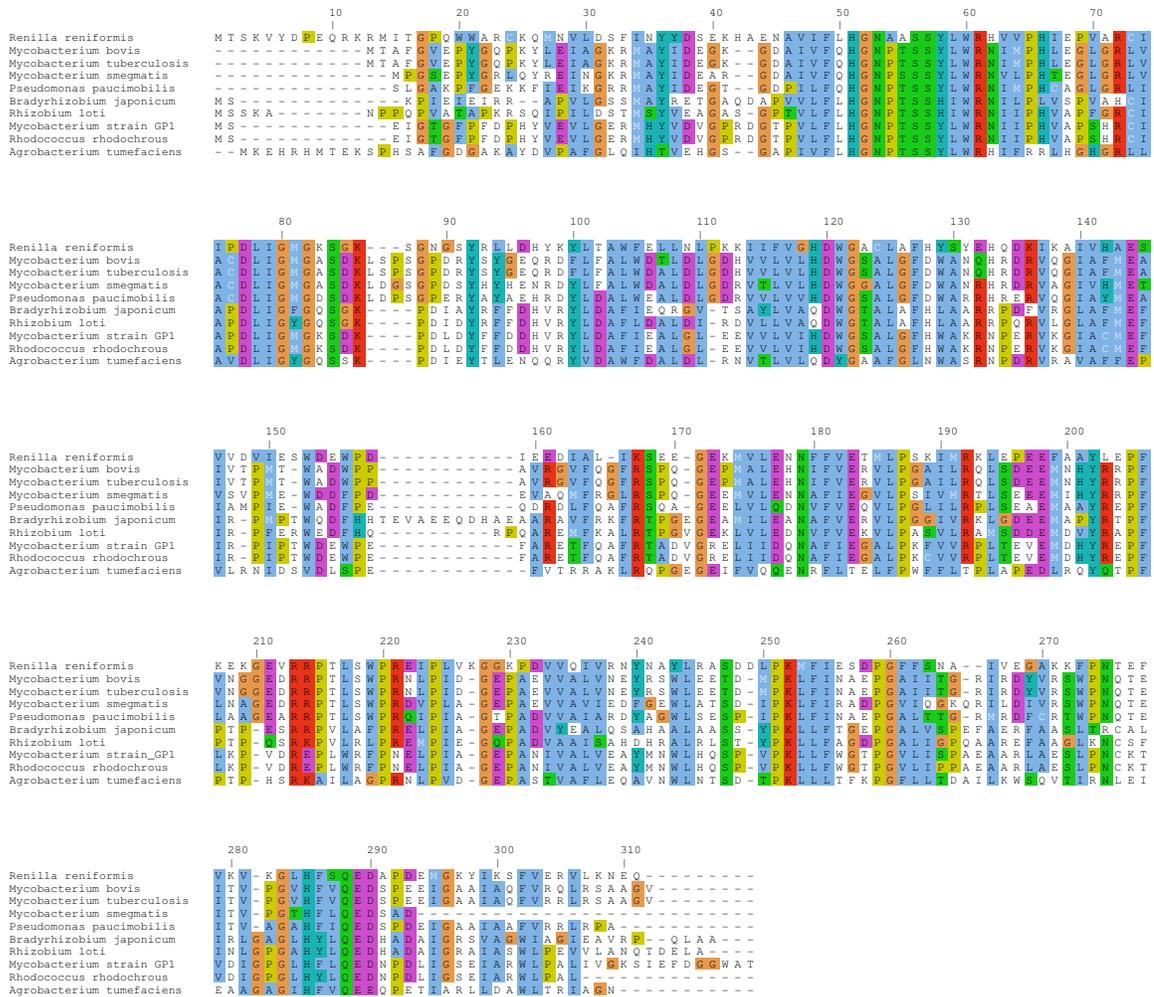
Figure 2.3: Initial alignment of *Renilla reniformis* luciferase with several haloalkane dehalogenases from various bacterial species. This is the sequence alignment (December 2003) off of which most of the consensus-guided candidate mutations were picked. The residue numbering is with respect to the RLuc sequence, and the alignment is depicted with a ClustalX color scheme [210].

Figure 2.4: Later alignment of *Renilla reniformis* luciferase with proteins from 14 other species. This is an improved alignment done in February 2006 that includes several sequences not available when the original alignment (Figure 2.3) was performed. When possible, the names of the proteins are given in parenthesis. With the exception of RLuc and the protein of unknown function predicted from the *Strongylocentrotus purpuratus* (Purple Sea Urchin) genome, all the aligned sequences are from bacterial species. Species 1-5, 12, and 13 are marine organisms. Proteins 6 and 8-11 are known haloalkane dehalogenases. Proteins 3-5, 7, 14, and 15 are putative haloalkane dehalogenases. Amino acids in the sequences corresponding to gaps in the aligned *R. reniformis* sequence have been removed. The bacterial dehalogenases have a conserved catalytic triad [82] corresponding to D120, E144, and H285 in the RLuc sequence. Residues that were selected for mutagenesis in RLuc are shown in boxes. RLuc was 40% identical and 72% similar to the protein from *S. purpuratus*, and 29-43% identical and 49-62% similar to the bacterial proteins. The GenBank Identifiers for sequences 1-15 are 160820, 72160391, 40062609, 83943988, 85826216, 34810153, 16508080, 50399582, 28558081, 14025217, 27349338, 68181465, 50082962, 29608437, and 15163540, respectively.

Figure 2.5: Homology model of *Renilla* luciferase based on its similarity to the haloalkane dehalogenase LinB. The region of the enzyme from residue 35 to 309 was successfully modeled using SWISS-MODEL and is shown. The N-terminus (N) is in blue and the C-terminus (C) is in red. The presumptive catalytic triad of D120, E144, and H285 is marked, along with the mutation site M185.

2.2.2   Mutagenesis of *Renilla* Luciferase and Screening

In the hope of further enhancing RLuc's resistance to inactivation beyond that achieved with the C124A mutation, a number of further mutations were explored. Candidate mutations were chosen from the alignment data at positions where RLuc most clearly diverged from the consensus sequence. For instance, the candidate mutation A55T was chosen because RLuc harbors the aliphatic amino acid alanine at position 55, while nearly all the other proteins harbor a hydroxylic residue of either threonine or serine at this site. Similarly, S287L was chosen as a candidate because RLuc contains a hydroxylic residue at this position, differing from the consensus aliphatic residue. Some of the candidates, such as M253L, are less obvious. This mutation substitutes an aliphatic residue for another aliphatic, but brings the RLuc sequence into consensus with the highly conserved local sequence near this position.

Complete results with respect to activity, inactivation in serum, and emission spectra are summarized in Table 2.1 for 25 initial mutations done on a background of RLuc with the C124A mutation. Note that activity was defined as a 10 s integration of the light output curve in order to disfavor mutations that merely increased the burst value at the expense of total light output. Representative data for inactivation in serum and emission spectra are shown in Figures 2.6 and 2.7, respectively.

The assayed values for RLuc reported in Table 2.1 corresponded well with previous values reported in the literature. In terms of resistance to inactivation under serum-like conditions, the values reported here for recombinant RLuc ($\tau_{1/2} = 0.4 - 0.9\ h$) are in line with Liu *et al.*, who reported a half-life of 0.6 h for recombinant RLuc in hamster blood at 37°C [120], as well as Lorenz *et al.*, who reported a half-life of 0.5 h for recombinant RLuc in a high ionic strength buffer [127].

The measured emission peak for RLuc with coelenterazine (481 nm) corresponded almost exactly with the previously published value of 482 nm for RLuc purified directly from *Renilla reniformis* [71]. Since previous work in the literature has often used benzyl-coelenterazine instead of the native substrate, the emission spectra for RLuc was assessed with this analog. Using benzyl-coelenterazine, an emission peak of 483 nm was recorded for RLuc, which corresponded closely with the previous value of 480 nm reported for both recombinant RLuc [128] and RLuc purified directly from *Renilla reniformis* [138].

Peak light flux from recombinant RLuc was determined to be $(1.2 \pm 0.2) \times 10^{23}$ photons/s/mole enzyme when in the presence of 24 μM coelenterazine. This value corresponds acceptably with the value[4] of 6.5×10^{22} photons/s/mole enzyme reported for RLuc purified directly from *Renilla*

	Activity (photons/s/mole enzyme)					Serum Inact. $\tau_{1/2}$ (h)		Wavelength (nm)			
	native	bc	cp	n	bdc	mouse	rat	peak	mean	fwhm	%>600
Native RLuc	$(3.2\pm0.3)\times10^{22}$	5.4×10^{22}	1.7×10^{22}	8.3×10^{21}	3.5×10^{19}	0.9±0.1	0.4±0.1	481	497	93	3
Initial Mutations	Activity (relative to RLuc)										
C124A	1.2±0.1	0.75	0.79	0.63	0.68	7.1±0.4	6.6±0.5	482	498	96	3
C124A-ΔMyc	1.3±0.1	0.91	1.1	0.87	1.0	4.0	4.5	481	499	96	3
F33R/I34M/C124A	0.15	0.15	0.16	0.12	0.20	0.3	0.3	481	497	96	3
E44G/C124A	0.94	0.78	0.74	0.66	0.98	2.6	3.3	486	502	94	4
A54G/A55G/C124A	0.12	0.10	0.06	0.15	0.19	2.4	3.0	476	492	98	3
A54P/A55T/C124A	0.21	0.15	0.11	0.38	0.22	119	129	470	483	96	2
A54P/C124A	0.05	0.04	0.05	0.08	0.06	14	13	468	482	96	2
A55T/C124A	1.7	1.2	0.58	1.4	2.4	30	29	486	504	91	4
F116L/C124A	1.3	1.0	1.3	0.88	1.8	11	9.4	486	502	97	4
C124A/S130A	1.7	1.4	1.7	1.4	2.6	18	14	482	498	96	3
C124A/K136R	2.5±0.3	2.1	1.9	1.9	2.6	12	11	482	498	95	3
C124A/A143M	1.7	1.3	0.95	1.5	1.6	30	29	480	497	97	3
C124A/F180A	0.02	0.01	0.03	0.01	0.01	1.6	1.6	488	504	104	5
C124A/M185V	3.4	3.0	15	7.8	44	5.7	3.7	485	500	97	4
C124A/M191L	1.1	0.99	0.97	1.0	1.2	6.5	5.1	480	496	97	3
C124A/E195S/P196D	0.12	0.10	0.12	0.10	0.15	1.0	0.7	482	498	96	3
C124A/F199M	0.58	0.44	0.53	0.49	0.46	6.7	6.0	480	495	97	3
C124A/L203R	0.55	0.55	0.52	0.41	0.43	2.7	2.2	484	501	95	4
C124A/G229E	0.02	0.01	0.03	0.03	0.01	1.9	1.8	473	490	92	3
C124A/Q235A	1.2	1.1	1.1	1.0	1.2	3.3	3.6	473	489	95	2
C124A/M253L	1.9	1.4	1.6	1.6	1.7	15	10	471	488	95	3
C124A/S257G	1.1	0.95	1.3	1.1	3.0	1.3	1.4	477	493	95	3
C124A/F261L/F262L	0.00	0.00	0.00	0.00	0.00	ND		ND			
C124A/F262L	0.03	0.03	0.01	0.06	0.03	5.8	6.4	478	495	95	3
C124A/S287L	3.9	2.8	3.4	5.0	9.5	28	20	478	496	93	2
C124A/M295I	1.0	0.83	0.57	0.72	0.86	5.0	4.9	480	497	93	3
C124A/K300A	1.1	1.0	1.1	1.0	1.3	3.5	3.9	481	497	96	3
										97	3
Inactivation Resistant											
RLuc8	4.3±0.2	3.0	5.8	8.8	59	281 ±49	86 ±9	486	503	94	4

Table 2.1: Mutations of RLuc altered light output and rates of inactivation in serum. Activity values are the result of integrating over 10 s and are not peak burst values. "Native" indicates the native substrate, while "*bc*", "*cp*", "*n*", and "*bdc*" indicate the analogs benzyl-coelenterazine, coelenterazine-*cp*, coelenterazine-*n*, and bisdeoxycoelenterazine, respectively. The results for the native enzyme are reported in absolute units, while the values for the mutants are reported as relative to the native enzyme for the given substrate. Since bisdeoxycoelenterazine's emission spectrum is significantly blue shifted from the other substrates, a multiplication correction of 0.6 was applied to correct for the luminometer's enhanced spectral sensitivity at these shorter wavelengths (see Table A.1). The wavelength measurements shown are for native coelenterazine, and the mean and peak wavelengths differ due to the non-symmetrical distribution of the emission spectrum. C124A-ΔMyc differs from C124A in that the Myc epitope introduced by the bacterial expression plasmid has been removed in order to make it directly comparable to RLuc8. RLuc8 contains the mutations A55T, C124A, S130A, K136R, A143M, M185V, M253L, and S287L. In cases where a particular protein was produced, purified, and assayed independently three or more times, the standard error of the mean is reported. ND - Not Determined. Inact. - inactivation.
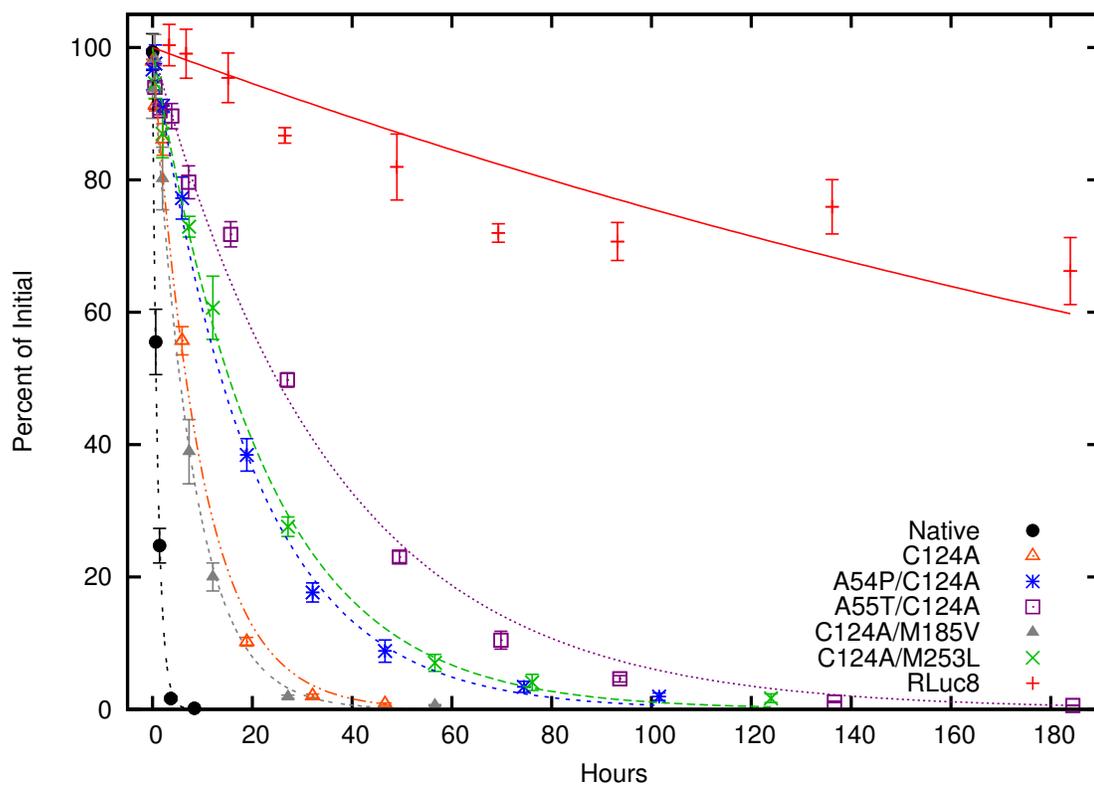
Figure 2.6: Mouse serum inactivation data for RLuc, RLuc8, and several other mutations. Protein was incubated in mouse serum at 37°C in triplicate, with aliquots removed at various times to determine the remaining luciferase activity. The error bars represent the standard error of the mean, and the lines drawn between points are from the fit to a mono-exponential decay model.
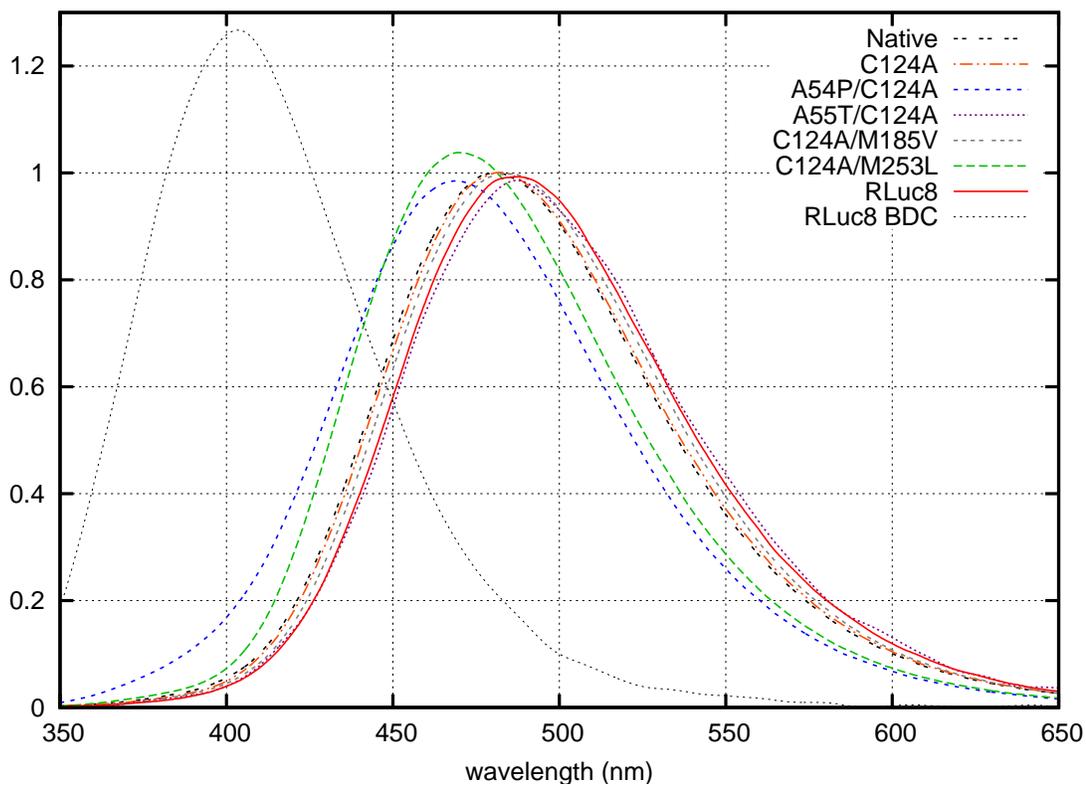
Figure 2.7: Normalized bioluminescence emission spectra for RLuc, RLuc8, and several other mutations. Most of the consensus guided mutations resulted in only small shifts from the emission spectra of the native luciferase. Coelenterazine was used for obtaining all the emission spectra except for the curve labeled RLuc8 BDC, which was obtained using bisdeoxycoelenterazine and is included for the purpose of comparison. The normalization equalized the total area under the curve.

reniformis [138], and 9×10^{22} photons/s/mole enzyme reported for recombinant RLuc [127].

2.2.3 Combining Mutations for a Luciferase Resistant to Serum Inactivation

For the purpose of generating a mutant RLuc more appropriate for use as a bioluminescent label in small animal imaging applications, the initial mutations were judged for their ability to confer resistance to serum inactivation as well as their light output. After excluding A54P due to its strong negative effect on light output and F116L as it yielded only a small increase in resistance to serum inactivation, the 7 remaining mutations that exhibited either increased light output or enhanced resistance to serum inactivation were combined along with the C124A mutation into a single protein designated as "RLuc8". The 8 mutations present in RLuc8 are A55T, C124A, S130A, K136R, A143M, M185V, M253L, and S287L. Since the Myc epitope was removed during the cloning of RLuc8, a C124A mutant was constructed without the Myc epitope to facilitate a valid comparison (C124A-ΔMyc). The activity, resistance to serum inactivation, and spectra peak values for these two enzymes are shown in Table 2.1.

When compared to the native enzyme, RLuc8 exhibited a greater than 4-fold enhancement in activity, a 200-fold increased resistance to serum inactivation, and a small but measurable 5 nm red shift in the emission spectrum. Compared to the C124A mutant, RLuc8 showed a 3-fold increase in activity and at least a 13-fold improved resistance to inactivation in murine serum.

Peak light flux from RLuc8 was determined to be $(1.9 \pm 0.1) \times 10^{23}$ photons/s/mole enzyme when in the presence of 24 μM coelenterazine. As this is only 50% greater than the corresponding value for RLuc, the 4-fold enhancement in specific activity (measured over 10 s) must have contributions from other kinetic parameters in addition to the increase in the initial reaction velocity.

2.2.4 Protein Production and Storage

Final recovery of periplasmically produced purified protein ranged from 5 μg/ml of culture (RLuc) to 50 μg/ml of culture (RLuc8), with the mutants more resistant to inactivation generally falling toward the high end of this range. Protein assessed by electrophoresis indicated

---

[4]The description of the assay in the associated reference to this value is misleading, as the coelenterazine concentration used sounds unreasonably low for the amount of light output that was recorded. The authors did measure activity at low concentrations of substrate, but neglected to state that they then multiplied their acquired values by 9.25 to "correct" for not using saturating amounts of coelenterazine [137].

that the periplasmically expressed protein was >95% pure after the one step purification process (Figure 2.2). Following purification, the proteins were stored in EB+1%HSA at 4°C, and were not observed to lose activity under these conditions for at least 3 months (data not shown). Cytoplasmically expressed RLuc8 produced significantly more protein, and generally resulted in a luciferase yield on the order of 500 mg/L of culture.

As an additional test of stability under storage conditions, RLuc8 was assessed for its ability to tolerate lyophilization. Following lyophilization, storage for 1 day at room temperature, and rehydration, RLuc8 regained full activity. This is in marked contrast to lyophilized RLuc, which loses 90% of its activity following storage for 1 day at room temperature [137]. Following lyophilization and storage for 6 months at room temperature, RLuc8 exhibited 50% of its initial activity upon rehydration.

2.2.5 Comparison with Cytoplasmically Expressed Protein, C24 Mutants, and A2T Mutants

In order to assess what effects the oxidative environment of the bacterial periplasm may be having on the proteins, RLuc and RLuc8 were expressed in the reducing environment of the bacterial cytoplasm using expression plasmids that did not include the pelB leader sequence. Compared to the periplasmically expressed RLuc presented in Table 2.1, cytoplasmically expressed RLuc had 25% greater activity with similar serum inactivation half-lives (0.8 h and 0.4 h for mouse and rat serum, respectively). Cytoplasmically expressed RLuc8 was also purified and had a similar activity (within 10%) and serum inactivation half-lives (290 h and 65 h in mouse and rat serum, respectively) to the periplasmically expressed RLuc8.

As a further test of whether inadvertent disulfide bonds may be formed in the periplasmic environment, C24A and C24S mutants of RLuc8 were produced and analyzed (Table 2.2). These mutants exhibited one third the activity of RLuc8, as well as small drops in resistance to serum inactivation.

The parental *Renilla* luciferase used in all these experiments was the hRL construct from Promega, which differs from the true native sequence due to the presence of a T2A mutation. The reason for this mutation is not clear, but it may have been introduced in order to incorporate an NcoI restriction site straddling the start codon of the gene. To check the effects of this mutation, A2T mutations were incorporated in both the RLuc and RLuc8 constructs. The periplasmic expression yields from these constructs were extraordinarily low, presumably reflecting that a hydrophobic residue is preferred over a hydroxylic residue for the position

	Specific Activity (relative to RLuc)					Serum Inact. $\tau_{1/2}$ (h)		Wavelength (nm)			
	native	bc	cp	n	bdc	mouse	rat	peak	mean	fwhm	%>600
RLuc	1.0	1.0	1.0	1.0	1.0	0.9±0.1	0.4±0.1	481	497	93	3
RLuc8	4.3±0.2	3.0	5.8	8.8	59	281 ±49	86 ±9	486	503	94	4
A2T Mutations											
A2T	1.5	1.2	1.3	0.8	1.0	ND			ND		
RLuc8/A2T	3.9	2.4	4.2	4.6	41	ND			ND		
C24 Mutations											
RLuc8/C24A	1.4	0.91	1.9	0.50	18	63	29		ND		
RLuc8/C24S	1.4	1.3	2.0	2.8	25	111	53	480	498	92	3
Catalytic Triad											
RLuc8/D120A	0.000	0.001	0.001	0.003	0.21	> 100	> 100		ND		
RLuc8/D120N	0.023	0.016	0.050	0.34	5.1	> 100	> 100	483	501	94	4
RLuc8/E144A	0.000	0.000	0.000	0.000	0.000	57	13		ND		
RLuc8/E144Q	0.000	0.000	0.000	0.000	0.002	> 100	> 100		ND		
RLuc8/H285A	0.023	0.020	0.046	0.028	0.20	> 100	21		ND		
Inactivation Prone											
M185V	4.4	2.6	12	4.1	20	0.8	0.3		ND		
M185V/Q235A	4.8	2.7	14	7.1	20	0.5	0.2		ND		

Table 2.2: Additional mutations of *Renilla* luciferase to test for the effects of the T2A mutation, inappropriate disulfide bond formation, conservation of the catalytic triad, and decreased resistance to inactivation. The data for RLuc and RLuc8 is repeated from Table 2.1 for the purpose of comparison. Coelenterazine analog abbreviations and bisdeoxycoelenterazine emission corrections are as stated in the caption for Table 2.1. The wavelength measurements shown are for native coelenterazine, and the mean and peak wavelengths differ due to the non-symmetrical distribution of the emission spectrum. ND - Not Determined. Inact. - inactivation.

immediately C-terminal to the signal peptidase cut site. Due to the low expression, the values reported in Table 2.2 can only be taken as approximate, but they seem to indicate that the A2T mutation has little effect on the specific activity of the luciferase.

2.2.6 Mutations to Test Proposed Active Site

Based on the catalytic triad of residues known to be critical for activity in the haloalkane dehalogenases, it was predicted that D120, E144, and H285 would be required for activity in *Renilla* luciferase as well. The locations of these residues in a homology model of *Renilla* luciferase are shown in Figure 2.5. To test the hypothesis that these residues comprise a portion of the enzyme's active site, further mutations were made at these sites on the RLuc8 construct, with the results shown in Table 2.2. With respect to maintaining luciferase activity, mutations at these proposed active site residues were deleterious.

2.2.7 Combining Mutants for a Luciferase Prone to Inactivation

In many cases where a reporter gene is employed, the goal is to follow the dynamics of gene activation and repression. In these contexts, it is desirable for the reporter protein's activity to be labile so that levels of the reporter more accurately track the current state of gene activation. For this reason, brighter mutants that were more susceptible to serum inactivation were constructed, with the idea that these constructs would be more labile intracellularly as well. To accomplish this, the initial double mutants were compared to the single mutant C124A in order to identify mutations that led to increased activity without increasing resistance to serum inactivation (e. g. M185V) or increased serum inactivation without affecting initial light output (e. g. Q235A, S257G). Combining these mutations in the absence of C124A resulted in the mutants M185V and M185V/Q235A (Table 2.2) that showed more rapid inactivation in serum as well as enhanced activity in comparison to RLuc.

2.2.8 Quantum Yield and Kinetic Parameters of the Mutants

To understand the basis for RLuc8's higher activity compared to RLuc, both quantum yield and kinetic measurements were undertaken. The results shown in Table 2.3 indicated that RLuc8 had a 30% improvement in quantum yield for native coelenterazine, and a ∼30-fold increase in quantum yield for bisdeoxycoelenterazine. The values for native RLuc with native coelenterazine agree favorably with previously published values between 5.5 and 6.9% [139, 71]. The native RLuc quantum yield values for bisdeoxycoelenterazine, however, are 1/3 the value from previous reports [71]. The reason for this discrepancy is not clear.

Modeling the complete kinetics of *Renilla* luciferase turned out to be a rather complex undertaking. In order to make the task tractable, a standard Michaelis-Menten model was first fit to the initial velocity data for coelenterazine concentrations from 0.038 to 24 μM. Substrate binding was assumed to occur at a much faster rate than the enzymatic step. The parameter results for this fitting are shown in the first two columns of Table 2.4. Since the second Michaelis-Menten step is irreversible in this case, the substrate dissociation constant (K_s) is equal to the standard Michaelis-Menten constant (K_m). The values shown in Table 2.4 are roughly consistent with a previously published K_m value of 2 μM for RLuc in the presence of benzyl-coelenterazine [139].

For RLuc and C124A, the Michaelis-Menten model was inadequate to explain the initial velocity data at the highest coelenterazine concentration used (118 μM), nor could it satisfac-

	Quantum Yield (%)				
	native	*bc*	*cp*	*n*	*bdc*
Native RLuc	5.3±0.1	3.2±0.04	4.7±0.03	6.1±0.2	$(3.6\pm0.5)\times10^{-3}$
C124A	5.4±0.3	3.6±0.1	5.2±0.1	6.4±0.01	$(4.6\pm0.3)\times10^{-3}$
A55T/C124A	5.7±0.2	3.9±0.1	4.5±0.1	5.7±0.1	$(6.0\pm0.5)\times10^{-3}$
C124A/S130A	5.3±0.1	3.4±0.04	5.0±0.1	5.9±0.2	$(4.0\pm0.2)\times10^{-3}$
C124A/K136R	5.4±0.1	3.3±0.1	5.1±0.1	6.0±0.1	$(4.3\pm0.2)\times10^{-3}$
C124A/A143M	5.2±0.3	3.5±0.1	4.8±0.1	5.8±0.2	$(3.8\pm0.5)\times10^{-3}$
C124A/M185V	6.9±0.3	6.3±0.1	10.1±0.2	9.4±0.4	$(104.4\pm4.0)\times10^{-3}$
C124A/M253L	5.5±0.1	3.5±0.1	5.1±0.2	5.8±0.1	$(4.6\pm0.2)\times10^{-3}$
C124A/S287L	6.1±0.2	5.0±0.1	7.2±0.3	7.7±0.2	$(12.5\pm0.3)\times10^{-3}$
RLuc8	6.9±0.1	6.1±0.1	8.9±0.1	9.6±0.4	$(118.9\pm5.1)\times10^{-3}$

Table 2.3: Mutations of RLuc altered quantum yield. Since bisdeoxycoelenterazine's emission spectrum is significantly blue shifted from the other substrates, a multiplication correction of 0.6 was applied to correct for the luminometer's enhanced spectral sensitivity at these shorter wavelengths (see Table A.1). Standard errors of the mean are reported.

	K_s (μM)	k_{lum} (s^{-1})	K_o (nM)	k_{inact} (s^{-1})	K_p (nM)
Native RLuc	2.9±1.0	3.9±0.4	58±1	$(8.5\pm0.2)\times10^{7}$	0.62±0.01
C124A	2.7±0.8	4.7±0.4	26±2	$(5.7\pm0.2)\times10^{7}$	2.2±0.03
RLuc8	1.6±0.2	4.9±0.1	0	0	Infinity

Table 2.4: Parameter fits to the kinetic data for RLuc, the C124A mutant, and RLuc8 in the presence of coelenterazine. K_s and k_{lum} were derived from fitting the initial velocity data to the Michaelis-Menten model. For RLuc and C124A, the remaining parameters were then obtained by fitting the complete model (Figure 2.8) using these previously obtained values for K_s and k_{lum}. For RLuc8 the Michaelis-Menten model could satisfactorily fit the full progress curve data in all cases. Because of this, the higher order model was not fit to the RLuc8 data and the corresponding parameters (K_o, k_{inact}, and K_p) have not been determined. Instead, appropriate values for these three parameters have been entered in the above table to allow the full model to fit the RLuc8 data given the Michaelis-Menten parameters. K_s, K_o, and K_p are expressed as dissociation constants. Data from coelenterazine concentrations in the range of 0.038 to 24 μM were used. The errors shown are the formal standard errors of the fitted parameters.
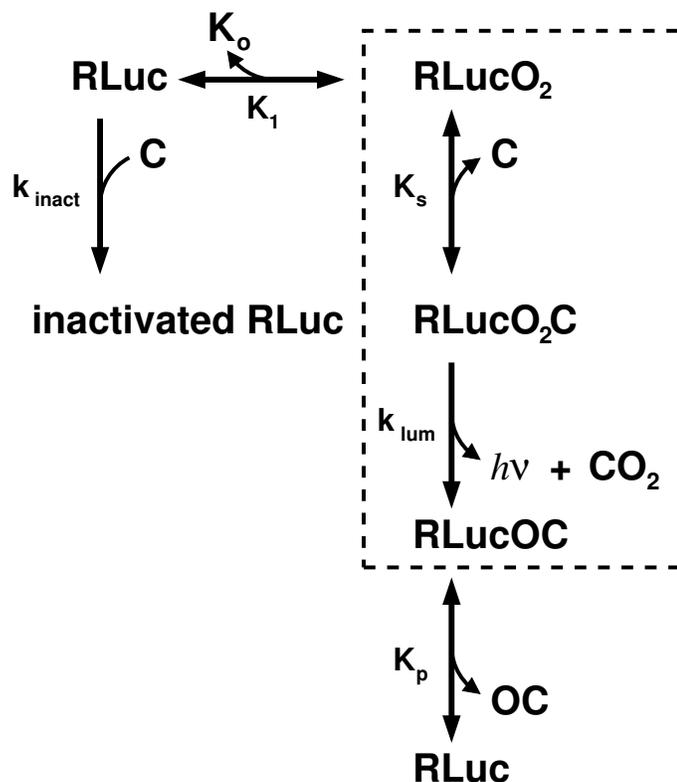
Figure 2.8: A proposed model for the kinetics of *Renilla* luciferase. The portion of the model enclosed by the dashed box is a Michaelis-Menten subset of the model that is appropriate for modeling the initial velocity of the reaction at lower substrate concentrations. C designates coelenterazine and OC designates oxidized coelenterazine (coelenteramide). Reversible and irreversible steps are designated by double arrow and single arrow lines, respectively.

torily model the kinetic progress curves. On account of this, the higher order model shown in Figure 2.8 was proposed, with the subset of the model enclosed in the dashed box corresponding to the Michaelis-Menten model. This higher order model adds a product release step and an irreversible inactivation pathway. The inactivation pathway was based on the known inactivation of RLuc when combined with μM levels of coelenterazine in anoxic conditions [139], and that product inhibition alone was insufficient to explain the observed kinetic data. Using the kinetic parameters from the Michaelis-Menten model, the additional parameters were fit using the complete model with the results shown in Table 2.4.

	293T	CHO
RLuc	22.5±0.7	11.7±1.2
C124A	62.8±12.9	64.0±4.9
C124A/M185V	174.1±25.7	172.4±8.4
RLuc8	328.8±34.5	281.8±42.1

Table 2.5: Transient transfection of RLuc and several mutants into 293T or CHO cells. The luciferases for this study were in pcDNA 3.1 plasmids under the control of the constitutive promoter from cytomegalovirus (CMV). Samples were measured in triplicate, data is in units of photons$\times 10^6$/s/cm^2/steradian, and the reported error is the standard error of the mean. For a given cell line, all differences between groups were significant at $p \leq 0.06$ using a two-tailed t-test with the incorporation of a Bonferroni-Holms correction for multiple comparisons [81].

2.2.9 Testing of Mutants in Mammalian Expression - First Round

In order to determine whether the *in vitro* data gathered for the RLuc mutants and RLuc8 would translate into the context of a mammalian reporter gene, expression vectors were constructed for RLuc, C124A, C124A/M185V, and RLuc8 in a pcDNA 3.1 backbone. These mammalian expression plasmids were then transiently transfected into 293T or CHO cells. The results 24 h post transfection, shown in Table 2.5, demonstrated increased light output for the mutants consistent with the *in vitro* data.

2.2.10 Testing of Mutants in Mammalian Expression - Second Round

A second round of mammalian expression experiments were performed utilizing 293T cells to assess the intracellular half-lives of the proteins and to test the RLuc variants prone to inactivation (M185V, M185V/Q235A). For this work, additional mammalian expression plasmids were constructed for the M185V and M185V/Q235A variants in a pcDNA 3.1 backbone. These mammalian expression plasmids, along with the ones utilized in the previous section, were transiently transfected into 293T cells. Measurements of light output over time following transfection, as shown in Figure 2.9 with respect to the RLuc plasmid, demonstrated that the mutations conferred increased light output following transfection in mammalian cells. A cycloheximide study was performed to assess the enzymatic stability for the luciferase variants in the context of the mammalian cytoplasm. As shown in Figure 2.10, the relative differences in inactivation resistance, but not the absolute differences, were consistent with the serum inactivation experiments. Through densitometry measurements of western blots performed on the cell lysates, the amount of luciferase was estimated for the different conditions and used to calculate the specific activity values shown in Table 2.6. These values were roughly consistent

	photons/s/mole enzyme
RLuc	$(4.2\pm0.2) \times 10^{22}$

Mutant	Activity (relative to RLuc)
C124A	1.6±0.1
C124A/M185V	4.2±0.6
M185V	2.5±0.2
M185V/Q235A	2.6±0.2
RLuc8	4.4±0.2

Table 2.6: Estimated specific activity values for *Renilla* luciferase and several variants expressed in mammalian cells. 48 h following transfection into 293T cells, the cells were lysed and analyzed for luciferase activity. Luciferase protein mass in the lysates was estimated via western blotting. Values were measured in quadruplicate, and standard errors of the mean are given. The estimated activity of RLuc is given in absolute values, with the remaining conditions given as relative to that of the RLuc condition.

with the *in vitro* data from bacterially expressed protein.

2.3   Discussion

The amino acid sequence of *Renilla reniformis* luciferase, along with that of the closely related *Renilla mülleri* luciferase (96% identical), contains a characteristic α/β-hydrolase fold sequence [166, 82]. This fold pattern is found in enzymes that catalyze a diverse range of chemical reactions in all kingdoms of life. Interestingly, within the α/β-hydrolase family RLuc shows a high level of similarity to the bacterial haloalkane dehalogenases, enzymes that catalyze the hydrolytic detoxification of halogenated compounds. This similarity even extends to the conserved haloalkane dehalogenase catalytic triad, present as D120, E144, and H285 in RLuc, being required for appreciable luciferase activity. The level of similarity is unexpected as RLuc is an oxygenase and *Renilla reniformis* is not a bacterium.

The evolution of a coelenterazine using oxygenase from an enzyme that catalyzes an unrelated reaction would not, in fact, present a considerable challenge for evolution. Coelenterazine chemiluminesces easily in aprotic solutions, and an initial enzyme would have to provide little more than a hydrophobic environment for coelenterazine to achieve some low level of bioluminescence [177]. As an example of this, serum albumin, which binds to a variety of lipophilic compounds but to the best of my knowledge has experienced no selective pressure to act as a bioluminescent enzyme, exhibits a low level of bioluminescence in the presence of coelenterazine [33]. That no similarity is seen between most of the coelenterazine using luciferases
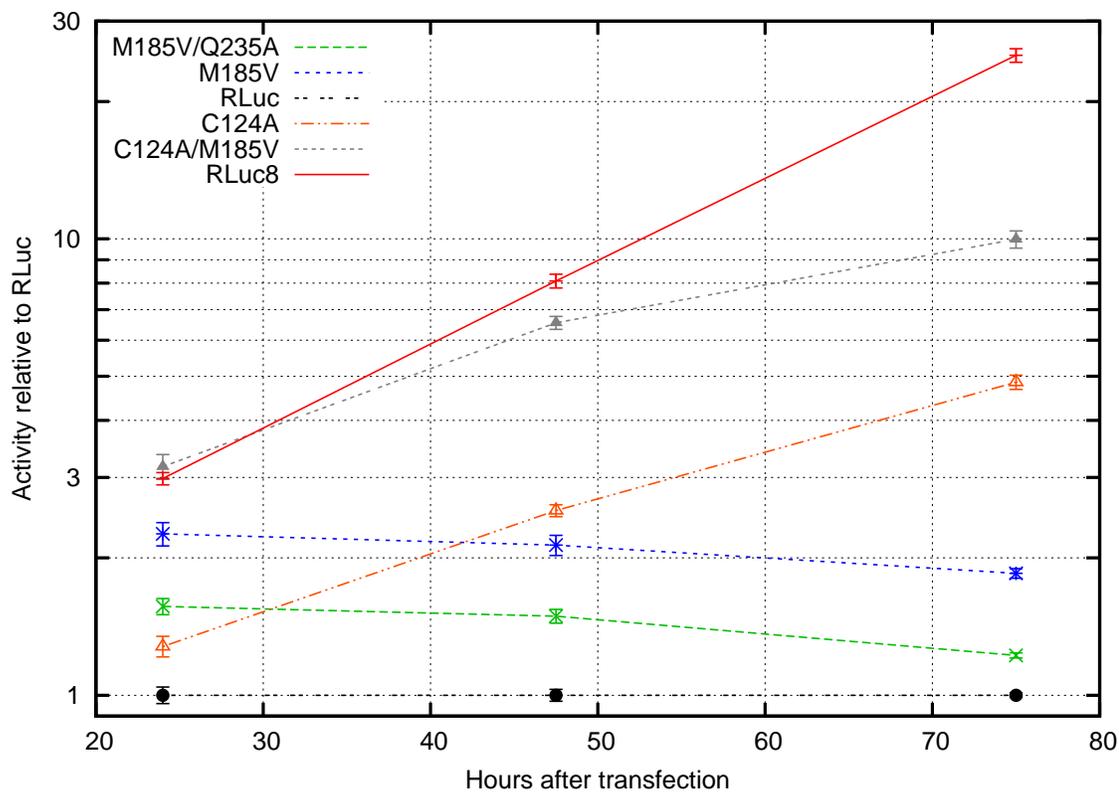
Figure 2.9: Mammalian cell expression of native RLuc and several mutants following transient transfection into 293T cells. The luciferases for this study were in pcDNA 3.1 plasmids under the control of the constitutive promoter from cytomegalovirus (CMV). Light output per total cellular protein was recorded for each condition, and is reported as relative to the value of the RLuc condition at the given time point. The absolute values for the RLuc condition at the given time points were (8.6 ± 0.3), (6.7 ± 0.2), and $(1.3\pm0.02)\times10^9$ photons/s/mg of total protein. Samples were measured in quadruplicate and error bars represent standard error of the mean. With the exception of the C124A/M185V and RLuc8 conditions at the 24 h time point, all differences between groups for a given time point were significant at $p \leq 0.05$ using a two-tailed t-test with the incorporation of a Bonferroni-Holms correction for multiple comparisons [81].
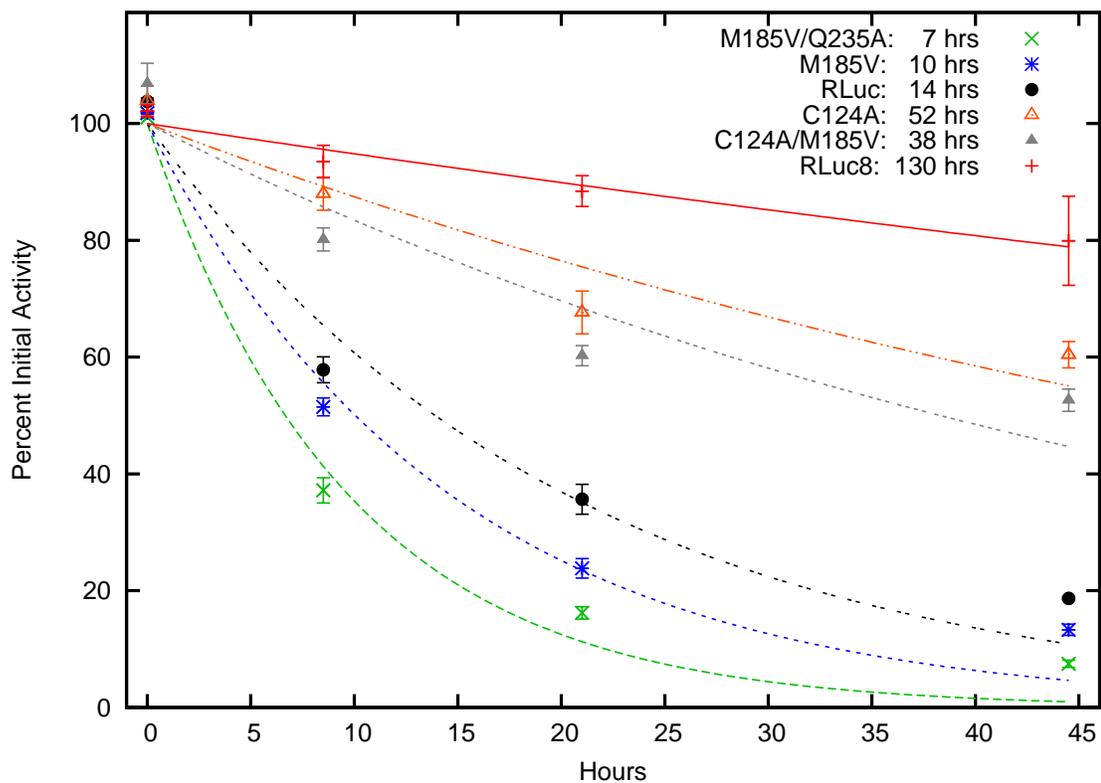
Figure 2.10: Measurements of the intracellular inactivation of luciferase activity. 293T cells, 48 h after being transiently transfected with the indicated luciferases, were exposed to 100 μg/ml of cycloheximide to inhibit new protein synthesis. Cells were assayed for light output per total cellular protein, with the data fit to a mono-exponential decay model. Samples were measured in quadruplicate, and the error bars represent the standard error of the mean. The estimated intracellular activity half-lives are given in the figure key.

identified so far (e. g. *Gaussia*, *Renilla*, *Pleuromamma*, *Oplophorus*) [134], even when the luciferases are from species within the same family (*Gaussia* versus *Pleuromamma*), lends further credence for the ease with which evolution can generate coelenterazine using luciferases from a variety of parental proteins.

The high level of primary sequence similarity between RLuc, the bacterial haloalkane dehalogenases, and the protein sequence derived from the *Strongylocentrotus purpuratus* genome is harder to explain. As the bacteria, sea pansies (e. g. *Renilla* species), and sea urchins (e. g. *S. purpuratus*) diverged at least 500 million years ago, horizontal gene transfer could be one possible explanation for this similarity. However, this would seem to imply two separate gene transfer events for the ancestral luciferase gene, which seems rather unlikely. On another note, it has not escaped my attention that the gene from *S. purpuratus* could potentially be a luciferase, a finding that might explain the blue bioluminescence emission that has been observed in the fertilized eggs of this species [78]. A sequence alignment in which residues encircling the putative active pocket are marked (Figure 3.7) is consistent with this hypothesis.

As discussed in the introduction, the C124A mutation was come upon by Liu and Escher while engineering a *Renilla* luciferase variant more appropriate for mammalian cell secretion [119]. Their hypothesis was that, in a secreted version of *Renilla*, these cysteines would have a propensity to form inactivating disulfide bonds in the oxidizing environment of the protein secretion pathway. They reported a complete loss of activity when mutating the second cysteine residue (C73) to an alanine. Notably, many of the aligned proteins contain a cysteine near this location (Figure 2.4). They also reported enhanced stability of the enzyme when the third cysteine (C124) was replaced with an alanine, and suggested that this was due to the blocking of unintended disulfide bond formation in oxidative environments. The homology model, however, indicates that the third cysteine is buried and removing the capability to form an inactivating disulfide bond is unlikely to explain the increased stability seen for this mutant after protein folding. More likely, the C124A mutation increases stability by allowing better packing of the hydrophobic core [118]. Interestingly, the alignment data shows that an alanine is favored at this position, coincidentally the amino acid chosen to substitute by Liu and Escher.

Liu and Escher also suggested that a disulfide bond formed under oxidative conditions between C24 and C73 could increase the resistance of the enzyme to inactivation, and showed a number of experiments consistent with this hypothesis. In contrast, the experiments done here did not reveal any significant changes with respect to inactivation resistance between

protein produced in the reducing environment of the cytoplasm and the oxidative environment of the periplasm. To further test for the presence of a stabilizing C24-C73 disulfide bond, C24A and C24S mutants of RLuc8 were also purified. Although these proteins were more prone to inactivation (RLuc8/C24S had ~75% of the resistance to inactivation of RLuc8), the increase in inactivation does not appear to be at the level one would expect from the removal of a stabilizing disulfide bond. As additional evidence disfavoring a C24-C73 disulfide bond, the three crystal structures presented in Chapter 4 (Sections 4.3, 4.4, and 4.5) from periplasmically expressed protein do not contain a C27-C73 disulfide bond.

The discrepancies between these results and those of Liu and Escher regarding the potential C24-C73 disulfide bond formation have not been reconciled, but may be due to a number of reasons:

1. Liu and Escher used a mammalian expression system whereby RLuc was secreted from the cell. The protein could conceivably be glycosylated[5] when expressed from this system, leading to some of the differences observed between their secreted RLuc and bacterially expressed RLuc. A potential human N-glycosylation site, predicted using the NetNGlyc 1.0 server `http://www.cbs.dtu.dk/services/NetNGlyc/`, is located at N89/G90/S91. Alternatively, the bacterial periplasmic expression system used here may not be sufficiently oxidative, or does not have the proper folding chaperones, to allow a C24-C73 disulfide bond to form.

2. In the Liu and Escher paper, when dithiothreitol (DTT) is added to the sample to measure the stability under reducing conditions, the authors also purged the sample of oxygen. In a 1977 paper [139] rapid inactivation and aggregation of RLuc was observed when the protein was combined with its substrate in the absence of oxygen. Conceivably, the removal of oxygen by Liu and Escher could have inadvertently altered the stability of the enzyme, even in the absence of substrate.

3. Since the experiments in the Liu and Escher paper were done in cell culture medium with fetal bovine serum (FBS), it could be that upon addition of DTT to the sample, disulfide bonds in native serum proteins were reduced and these proteins were in turn aggregating with RLuc. This inactivation could be fast enough to mask other processes going on.

---

[5]It was originally reported that purified *Renilla* luciferase was 3% carbohydrate by weight [138]. Assuming glycosylation in Coelenterates is equivalent to other eukaryotes, the intracellular location of RLuc in the lumisomes of *Renilla reniformis* would seem to be consistent with glycosylation of the protein. Later reports however failed to show any differences, as assessed by SDS-PAGE gel electrophoresis, when comparing recombinant non-glycosylated RLuc obtained from bacterial expression with potentially glycosylated RLuc either purified directly from *Renilla reniformis* [128] or expressed as a secreted protein in mammalian cells [119].

It is not yet clear how exchanging particular amino acids from *Renilla*'s sequence to the consensus yields an enzyme with greater resistance to inactivation. The packing of the protein core is crucial for a protein's evolutionary fitness [36], and enhanced protein packing is an oft-cited reason for stability inducing effects [118, 132, 36]. It may well be the mechanism through which many of the mutations increase the luciferase's resistance to inactivation in the serum environment. In support of this, of the 7 stabilizing mutations that went into RLuc8, only K136R is a surface residue. This is in contrast with previous studies that indicated a majority of stabilizing mutations identified using consensus guided mutagenesis strategies were surface residues [113].

An alternative hypothesis is that some of the mutation sites conferring resistance to inactivation may have previously been important in the aforementioned *in vivo* interaction between RLuc and *Renilla* GFP. Now, in the environment of *in vitro* assays absent the GFP, the wild type amino acids may be detrimental.

The exact mechanism for the luciferase's inactivation in serum was never determined, and it is not currently obvious if the inactivation is due to unfolding/aggregation, oxidation, proteolysis, or some other process. Western blots performed on time points taken during the serum inactivation experiments were inconclusive, as the secondary antibodies directed at the mouse anti-*Renilla* primary antibody exhibited to much background binding to endogenous immunoglobulins. An anti-C-terminal-6xHis-tag antibody directly conjugated to horse radish peroxidase (Invitrogen) was tried as well, but this proved to bind too non-specifically to be useful.

Three of the mutations, K136R, M185V, and S287L, showed sizable increases in the light output of the enzyme. In the case of M185V a portion of the increase in light output can be explained by enhanced quantum yield, particularly for the coelenterazine analogs, with the difference assumed to arise from enhanced kinetics. The disproportionate increase in light output seen with M185V for several of the coelenterazine analogs leads us to speculate that the light increase and specificity decrease are related, and that M185 may be positively selected for in RLuc and *Renilla mülleri* luciferase to ensure specificity of the reaction. This trade off between substrate recognition and light output could arise if the residue at position 185 was important for substrate recognition. The alignment data gives some credence to the hypothesis as M185 is located in the "cap" of the enzyme, a domain often used for substrate specificity in the haloalkane dehalogenases [82]. The location of M185 in the crystal structures presented in Chapter 4 lends further support, as its placement atop the presumptive catalytic site (formed

by D120, E144, and H285) would be an appropriate position for conferring substrate specificity to the luciferase.

S287L, while not located in the active site, is close enough to the active site residue H285 that slight alterations in the structure of the enzymatic pocket induced by S287L could explain the modest increase in quantum yield and the larger increase in light output seen with this mutation.

K136, on the other hand, is located on the surface of the protein approximately 20 Å distant from the presumptive active site. It is still possible that the increased light output seen with the K136R mutation could arise through long range perturbations in the folding of the active site, especially as K136 lies in a loop between the active site residues D120 and E144. Incidentally, *Renilla mülleri* luciferase contains an arginine at this position.

An alternative hypothesis for the increases in light output, especially for M185 and S287, is that these residues are somehow involved in the transfer of energy to the GFP homo-dimer that interacts with RLuc in nature. As such, these residues have been selected for their role in energy transfer to the GFP fluorophore and not for their ability to optimize the quantum yield of the luciferase on its own.

Using the luciferase assay, RLuc8 displayed a $\sim$4-fold increase in light output versus RLuc. For RLuc8, combining the increases in quantum efficiency with the enhanced Michaelis-Menten parameters would only predict a $\sim$70% increase in light output over RLuc. Additionally, for RLuc and the C124A mutant, the Michaelis-Menten model could not satisfactorily fit the reaction progress curves nor the initial reaction velocity at the highest coelenterazine concentration tested (118 μM), even when product inhibition was incorporated into the model. This is in line with previous failed attempts at elucidating a satisfactory kinetic model for RLuc [137]. A speculation is that a complete kinetic model would need to factor in coelenterazine-dependent inactivation of RLuc, a phenomenon known to occur rapidly with coelenterazine but not coelenteramide in anoxic conditions [139]. Such a model is presented in Figure 2.8, and it was able to fit the data. Whether or not this model is actually correct, or is just able to fit the data, has not been satisfactorily assessed.

A couple caveats bear mentioning regarding the values reported in this study. As discussed in Appendix A, absolute calibration for bioluminescence is a difficult undertaking. This difficulty is evidenced by the large number of studies that choose to report their results in terms of relative measures. Although I am comfortable with the precision of the results presented here, I am less so with their accuracy, and additional differences in substrates, substrate pu-

rity, buffers, and variations in the luciferase constructs used (e.g. 6xHis tags, Myc epitopes) complicate direct comparison with previous work. The C-terminal 6xHis tag may be especially problematic, as the activity levels of the cytoplasmically expressed variants of RLuc and RLuc8 with cleavable 6xHis tagged utilized later in this manuscript (Table 4.3) were several fold higher than the 6xHis tagged, periplasmically expressed proteins presented in this Chapter. This finding correlates with previous experience in our laboratory that has found that peptide or protein fusions on either end of *Renilla* luciferase invariably lead to some decrease in the enzyme's light output. Surprisingly, the values reported here for RLuc matched within 30% to previous bioluminescence values for both recombinant RLuc [127] and RLuc purified directly from *Renilla reniformis* [138], and may indicate that the "purified" protein in these previous reports actually contained some level of heterogeneity.

Additional caveats must be mentioned for interpreting the enzymatic activities and resistance to serum inactivation half-life measurements reported in Tables 2.1 and 2.2. Since the inactivation of the RLuc mutants can occur at any point during protein production and purification, it should be expected that the activity values for the more labile proteins will be down weighted with respect to their true values. As an example of this, RLuc expressed cytoplasmically in bacteria resulted in 25% higher specific activity compared to RLuc that had been purified through the more time consuming periplasmic extraction protocol. With regard to the resistance to serum inactivation half-life values, these experiments were performed using an incubator in order to avoid evaporation and condensation issues when assessing the higher stability mutants. The slow heat transfer in air, however, means that the reported values for the shorter half-life mutants (< 1 h) are overestimations.

Some attention must also be given to the presence of the Myc epitope, which was present in most of the bacterial expression constructs (Tables 2.1 and 2.2), but was dropped in C124A-ΔMyc, RLuc8, and the RLuc8 based mutations. From comparison of the C124A protein (which contains a Myc epitope) and the C124A-ΔMyc protein, it would appear that the presence of the Myc epitope does cause a slight decrease in the light output of the luciferase. In any case, the effect here is small and does not confound the results that RLuc8 emits significantly more light than the native luciferase. In terms of resistance to inactivation in serum, removal of the Myc epitope did cause a small decrease in this value for the C124A protein. This effect, however, only strengthens the assertion that RLuc8 is substantially more resistant to inactivation than the native enzyme.

Earlier reports have shown that native *Renilla* luciferase irreversibly loses activity due to

self-association when stored at concentrations greater than 0.5 mg/ml at 4°C [138]. While RLuc has not been tested for this effect, RLuc8 has been concentrated to 300 mg/ml in 10 mM NaCl, 10 mM HEPES pH 7.4 without any signs of activity loss or aggregation (See Chapter 4).

Transient transfection of plasmids containing RLuc, C124A, C124A/M185V, M185V, M185V-/Q235A, and RLuc8 into mammalian cells demonstrated that the basic trends derived from the *in vitro* mutation analysis are applicable in the context of a mammalian cell as well. There were some differences in terms of the absolute values of the results between the first and second round of mammalian transfection studies. These differences may have arisen from differences in methodology, namely, the cells were assayed directly in the 24-well culture trays in the first round experiments. This method most likely overly highlights mutants that are resistant to serum inactivation. Some amount of cell death inevitably occurs in cell culture studies, and the protein contents of these cells are released to the medium. When coelenterazine is applied to the wells, the luciferase in the medium has direct access to the substrate, while coelenterazine must traverse the plasma membrane to reach the luciferase contained in the living cells. The end effect would be that a combination of activity was measured from intracellular luciferase and from luciferase exposed to the serum containing medium in these experiments. The second round of experiments was done by aspirating the medium and dead cells, lysing the remaining adherent cells, and performing all the measurements on these lysates. These experiments are felt to be more representative and the results are used for the discussion here.

The absolute rate for RLuc inactivation in the mammalian cells (Figure 2.10) was an order of magnitude slower than that observed in serum (Tables 2.1 and 2.2). This should not be particularly surprising, given that RLuc is an intracellular eukaryotic enzyme in its native environment. On the other hand, the relative inactivation trends between the different luciferase variants were consistent between the two sets of data. For instance, M185V/Q235A inactivated approximately twice as fast as RLuc whether the experiment was performed in serum or in mammalian cells.

The specific activities measured from the mammalian transfection experiment (Table 2.6) were roughly consistent (within 30-50%) with the activities measured for the bacterially expressed proteins. The differences in the absolute values between the two sets of experiments are most likely due to errors in estimating the amounts of luciferase in the mammalian cellular lysates, and that the mammalian cellular lysates include a mix of active and inactivated protein accumulated during the course of the experiment.

Luciferases are extraordinarily useful in a variety of experiments that require reporter

genes. In instances where the reporter gene is constitutively expressed (e. g. cell trafficking studies [15]), RLuc8 should be advantageous because of its greatly increased light output compared to RLuc in mammalian cells (Figure 2.9). This increased light output stems from both RLuc8's increased specific activity (Table 2.6) and its decreased rate of inactivation (Figure 2.10).

In most reporter gene experiments, however, the investigator wishes to follow the dynamics of gene induction and suppression. In these contexts, a slow rate of inactivation of the reporter would be a detriment to the experiment, as the stability of the signal would obscure transient changes in gene expression. The single mutant M185V and the double mutant M185V/Q235A should be of utility in these cases, as both these mutants show an increase in specific activity as well as an increase in the rate of protein inactivation relative to RLuc.

One issue with the use of coelenterazine catalyzing luciferases for reporter gene assays in mammalian cells is that coelenterazine is a substrate for MDR1 P-glycoprotein (Pgp) [173]. While the resultant transport of coelenterazine out of mammalian cells can be used to measure levels of Pgp, in most studies this phenomenon leads to an inadvertent modulation of signal intensity. For this reason, there has been interest in the coelenterazine analogs coelenterazine-*cp* and coelenterazine-*n* as they are not substrates for Pgp [173]. These analogs, however, suffer from reduced light output when used with RLuc (2-fold for coelenterazine-*cp*, 4-fold for coelenterazine-*n*, see Table 2.1) as well as ~4-fold higher background rates of auto-chemiluminescence [247]. Combined, these factors lead to a drop in the signal to background ratio of 8 and 16-fold for coelenterazine-*cp* and coelenterazine-*n*, respectively. The M185V mutation greatly reduces the disadvantages of these alternative substrates. In the case of coelenterazine-*cp*, the M185V mutation leads to a signal to noise ratio that it is only a factor of two lower than that achieved by the native substrate. RLuc8 is not as effective at using coelenterazine-*cp* as the M185V mutation alone, most likely because the A55T mutation present in RLuc8 decreases its ability to use this substrate.

Bisdeoxycoelenterazine has been proposed as a better analog to use with bioluminescence resonance energy transfer (BRET) studies because of the increased separation between the bioluminescence and the fluorescence spectra [91]. Bisdeoxycoelenterazine, however, suffers from extraordinary low light output when used with native RLuc (Table 2.1) because of poor quantum yield (Table 2.3). Although low bioluminescence quantum efficiency does not necessarily imply low light output from BRET[6], and increased bioluminescence quantum efficiency may not translate into a corresponding increase in BRET output, preliminary data in the lab-

oratory indicates that bioluminescence quantum yield and BRET output are indeed related, at least when the acceptor moiety is a variant of *Aequorea* GFP [46] (Note that, unlike *Renilla* GFP, *Aequorea* GFP does not naturally interact with *Renilla* luciferase [228]). Both RLuc8 and the M185V mutation should be of great utility in these BRET assays, as they confer a 20-60 fold increase in light output with bisdeoxycoelenterazine. Interestingly, although C124A alone does not improve utilization of bisdeoxycoelenterazine, it appears to facilitate the M185V mutation, as C124A/M185V has a ∼2-fold better light output with this substrate compared to M185V alone.

2.4 Conclusion

In summary, mutants of RLuc have been characterized with respect to their resistance to serum inactivation as well as their light output. These results, in turn, have been utilized in order to develop luciferases optimized for different purposes. An 8 mutation form of RLuc (RLuc8) was created that has greatly improved characteristics for use as a bioluminescent label. Compared to the native enzyme, RLuc8 exhibited a 200-fold improvement in resistance to murine serum inactivation, a 4-fold improvement in light output, and a 5 nm red shift in the emission spectrum. The enhancement in light output arises from a combination of increases in quantum yield and improved kinetics. A double mutant of RLuc (M185V/Q235A) was created that has improved performance as a reporter gene. Compared to the native enzyme it has twice the rate of inactivation, as measured in murine serum, while incorporating a close to 5-fold improvement in light output. These optimized *Renilla* luciferases represent significant improvements that will increase the sensitivity of luciferase-based assays for both *in vitro* experiments and *in vivo* imaging.

Although the work done here was performed by generating single mutants via site specific mutagenesis and screening the resultant proteins, a more efficient method could potentially have been used based on the rough correlation between protein expression levels and the serum stability values (data not shown). This method [4] would still have involved consensus mutations, but would utilize a multiple site-directed mutagenesis protocol (e. g. QuikChange Multi Kit, Stratagene) to incorporate random assortments of the proposed mutations in a single PCR

---

[6]An early report utilizing bisdeoxycoelenterazine noted that while RLuc had an extremely low quantum yield with this substrate, the quantum yield would go up 200-fold when *Renilla* GFP was added to the mixture [71]. The authors speculated that the presence of the resonance energy transfer pathway for the excited state coelenteramide to traverse was reducing the amount going into the quenching pathway. An alternative hypothesis, consistent with the M185V mutation results, is that binding of the GFP to RLuc leads to a conformational change in the enzymatic pocket such that there is a reduction in the tendency for excited state coelenteramide to quench.

step. Following transformation into bacteria and growth at 37°C[7], the resulting colonies could be screened in a high throughput fashion (as utilized for the random mutagenesis in Chapter 3) in order to select for mutations that lead to higher levels of protein expression and presumably greater protein stability.

---

[7]For increased levels of stringency, the plates could potentially be incubated at elevated temperatures (e. g. 60°C) for short periods immediately prior to screening. This however may kill the bacterial cells, and necessitate purifying the plasmid directly from the bacterial colony.

Chapter 3

RLuc8 Mutagenesis Studies

Luciferases are commonly used as reporter genes, and hopefully in the future as bioluminescent labels, in a variety of biological assays performed both *in vitro* and *in vivo*. For *in vitro* assays such as cell culture transfection studies, the wavelength of light that a luciferase yields is usually of little consequence. In contrast, for small animal imaging it is highly advantageous for a luciferase to emit a large percentage of its photons in the red to near-infrared wavelengths (600-900 nm), as tissue attenuation of optical photons is minimized in this region of the spectrum [231].

For the various reasons outlined in Section 1.3, *Renilla* luciferase (RLuc) was chosen as the imaging moiety to use for bioluminescent labeling. A major limitation in the use of any of the known coelenterazine utilizing luciferases, however, is that the spectral peaks of these luciferases lie in the blue region of the visible spectrum. In the case of RLuc, the spectral peak is at 481 nm, with only $\sim$3% of the photons of wavelengths above 600 nm. As pointed out later in the text (Figure 3.18), for luciferase locations at anything deeper than superficial depths, the majority of the photons that actually make it out of the animal are these few >600 nm wavelength photons [246]. Clearly, a *Renilla* luciferase with a bathochromic (red) shifted emission spectrum and therefore a greater number of >600 nm wavelength photons would be advantageous for use in small animal imaging.

The theory behind shifting the bioluminescence emission spectrum of RLuc starts with an understanding of the luminescence reaction. As shown in Figure 3.1, the reaction starts with coelenterazine and molecular oxygen, and yields carbon dioxide, coelenteramide, and a photon of light. After some early confusion as to the exact mechanism, it was convincingly shown that
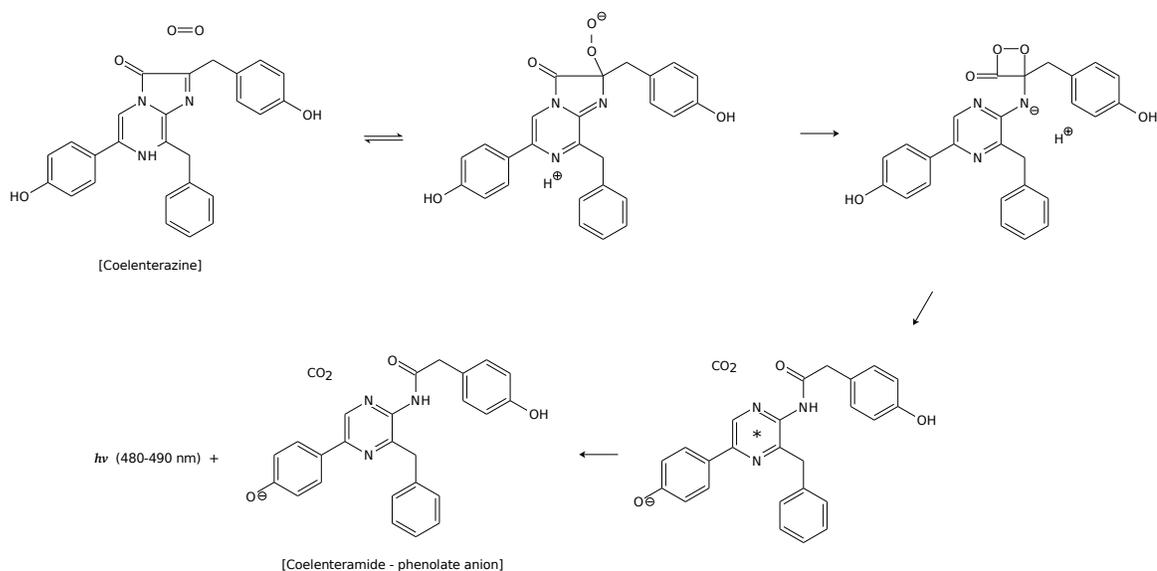
Figure 3.1: The luminescence reaction catalyzed by *Renilla* luciferase, showing the dioxetane intermediate step. The phenolate anion of coelenteramide is shown as the emitting species in this diagram, although assignment of the emitting species' exact state is a matter of contention (see main text).

the reaction goes through a dioxetane (also called dioxetanone or cyclic peroxide) intermediate step [72].[1] This same dioxetane intermediate mechanism has been confirmed for other coelenterazine using luciferases such as *Cypridina* luciferase and *Oplophorus* luciferase [189, 190], and a dioxetane intermediate is involved in firefly luciferase's reaction with D-luciferin [188] as well. The dioxetane is a high energy bond, and the break down of this structure leaves an electron in the resultant coelenteramide in an excited electronic state. At this point, the electron/coelenteramide system can lose energy and return to the ground electronic state through a number of processes, with the desired transition in bioluminescence being the conversion of this energy into a photon of light. The important thing to gather from all of this, is that the emitted photon's wavelength depends directly on the energy difference between the excited and ground states. In turn, the difference in energy between these two states will depend on the local chemical environment in which the coelenteramide finds itself.

The bioluminescence emission spectrum of *Renilla* luciferase would be expected to be directly related to the fluorescence spectrum of coelenteramide when it is bound to the luciferase.

---

[1]These elegant experiments were done with the use of $^{17}O_2$ and a coelenterazine analog containing alternatively an ^{18}O or ^{14}C labeled carbonyl. Interestingly, the ^{18}O labeled coelenterazine analog demonstrated that the carbonyl oxygen is almost completely exchanged with oxygen from H_2O while the analog is in the presence of the luciferase but prior to completion of the enzymatic process.

This is not what is actually observed, and is believed to indicate that the chemical environment the excited state product experiences is not retained for any significant period of time after the ground state is reached [138]. This is in contrast to *Cypridina* luciferase, where the bioluminescence emission spectrum is identical to the fluorescence spectrum of the coelenteramide/luciferase conjugate.

In any case, the chemical state that the excited coelenteramide is in can be investigated by exposing coelenteramide to a variety of solvents, so as to achieve different anions of the compound, and then measuring its fluorescence. Through such work, it has been suggested that the blue light emission (481 nm peak for RLuc) associated with coelenterazine bioluminescence is due to the excited state coelenteramide existing in its amide anion form (Figure 3.2b) [192] when it is in the protein's enzymatic pocket. For RLuc, the literature has historically agreed with this assignment of the emitting species [86, 138]. However, at least in the case of photoproteins such as obelin and aequorin, more recent literature has strongly favored assigning the phenolate anion (Figure 3.2c) as the blue emitting species in bioluminescence [87, 222, 123]. If the phenolate anion is the blue light emitter in RLuc, it would be in agreement with work utilizing analogs of coelenterazine, where replacing the phenol group's hydroxyl with a methyl group, a methyl ester, or a hydrogen (e. g. bisdeoxycoelenterazine) maintains the molecule in its neutral state and leads to a ~400 nm emission peak [71].

An interesting observation from this work, is that coelenteramide can be shown to emit a green fluorescence (535-550 nm) when it is in particular chemical environments. While early reports suggested that this green fluorescence is due to the phenolate/amide dianion [85], more recent reports [192] have indicated that the green fluorescent coelenteramide is in fact the pyrazine anion (Figure 3.2d). Recognition that coelenteramide can achieve an anionic state in which it fluoresces green led to the hypothesis that, by proper alteration of the enzymatic pocket of the luciferase, the necessary chemical environment could be attained within RLuc so as to favor the green emitting anion form. If the blue light emitter in RLuc bioluminescence is indeed the phenolate anion and the green fluorescent form is the pyrazine anion, one can imagine that favoring the pyrazine anion resonance form of the coelenteramide molecule will lead to a bathochromic shift in the emission spectrum.

This chapter begins with further mutagenesis work done using the consensus mutagenesis approach, in which the aim was to see if any "easy gains" could be made in improving the properties of RLuc8. This is followed by predictions regarding the active pocket of RLuc, along with site specific mutagenesis studies of the predicted catalytic pocket residues in an attempt

400 nm

(a) Neutral Species

435-458 nm

(b) Amide Anion

480-490nm

(c) Phenolate Anion

535-550nm
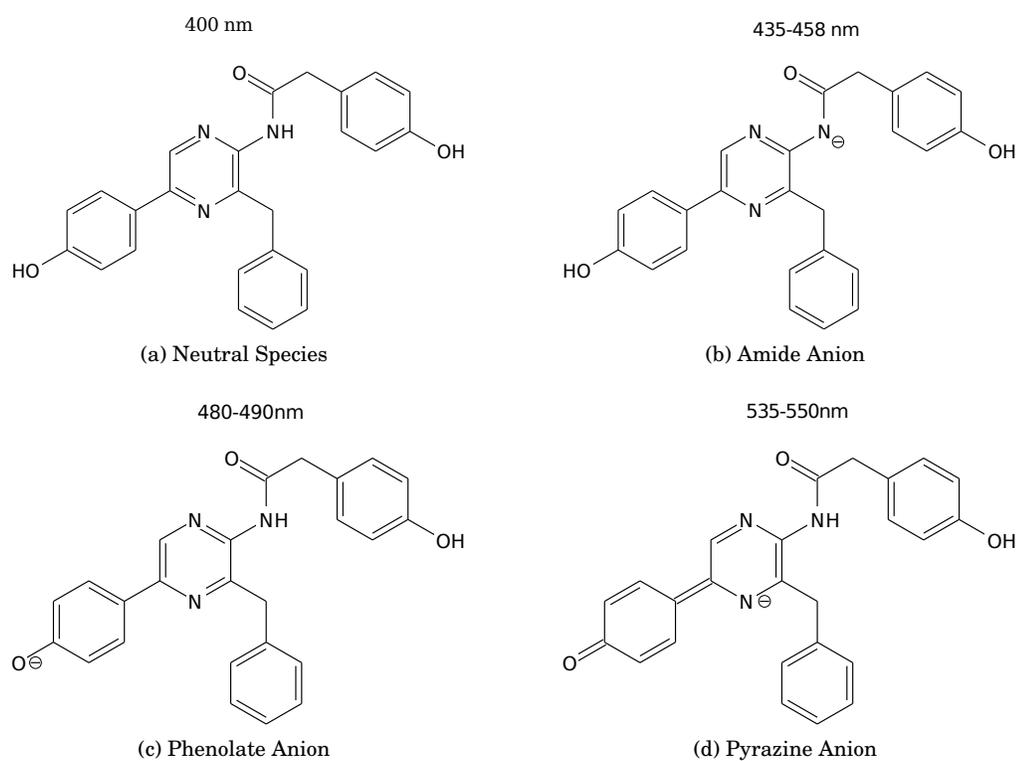
(d) Pyrazine Anion

Figure 3.2: Different anionic states of coelenteramide. The fluorescence emission peak thought to be associated with each species is given above the structure [192, 50]. Note that the phenolate and pyrazine anions are different resonant structures of the same molecule.

to red-shift the emission spectrum of the protein. The chapter concludes with iterations of random and saturation mutagenesis done in an attempt to achieve the goal of a viable red-shifted *Renilla* luciferase. In essence, this chapter encompasses a directed evolution strategy [79].

3.1 Methods

Please note that many of the methods used in this chapter are identical to those from Chapter 2 and are not repeated here. Only those methods that are new or have been modified are explained below. Also note that all experiments in this chapter were based off of the pBAD-pelB-RLuc8 periplasmic expression plasmid. This means that all constructs were expressed with a pelB leader sequence on the N-terminus (which is removed during export to the periplasm), and contain a VDHHHHHH sequence on their C-terminus.

3.1.1 Site Directed Mutagenesis

As site-directed mutagenesis is such a commonly used protocol and few good references are available, the method is described in some detail here. The following protocol is based largely on the commercial QuikChange kit (Stratagene), and the basic reaction scheme is shown in Figure 3.3. The important thing to keep in mind is that the concentration of primer versus template determines whether primer-primer or template-primer interactions dominate. If a reaction fails due to the formation of primer-primer dimers (as seen on gel electrophoresis of the reaction product), dropping the primer concentration, increasing the template concentration, decreasing the annealing temperature, or adding 5% DMSO can help. Additionally, a high fidelity polymerase is fairly important for the PCR reaction, as the length of DNA being duplicated is quite long. In the work here, PfuUltra Hotstart (Stratagene) was used.

Oligonucleotide primers should be designed with a T_m of ~78 °C, with the intended mutation in the middle of the primer. A primer design web site is available to automate this process at `http://bioinformatics.org/primerx`. For more complicated primers, such as for multiple amino acid changes, please see Zheng *et al.* [248]. While the QuikChange (Stratagene) protocol specifies that the oligonucleotides need to be PAGE purified, unpurified primers (Stanford PAN Facility) have been used with equal success and were used extensively in this chapter.

The reaction mixture consisted of 10% 10x PfuUltra buffer, 2.5 pmoles of each primer, 10 nmoles of dNTP (2.5 nmoles of each), 2 ng of template DNA, and 0.5 U of PfuUltra Hotstart,
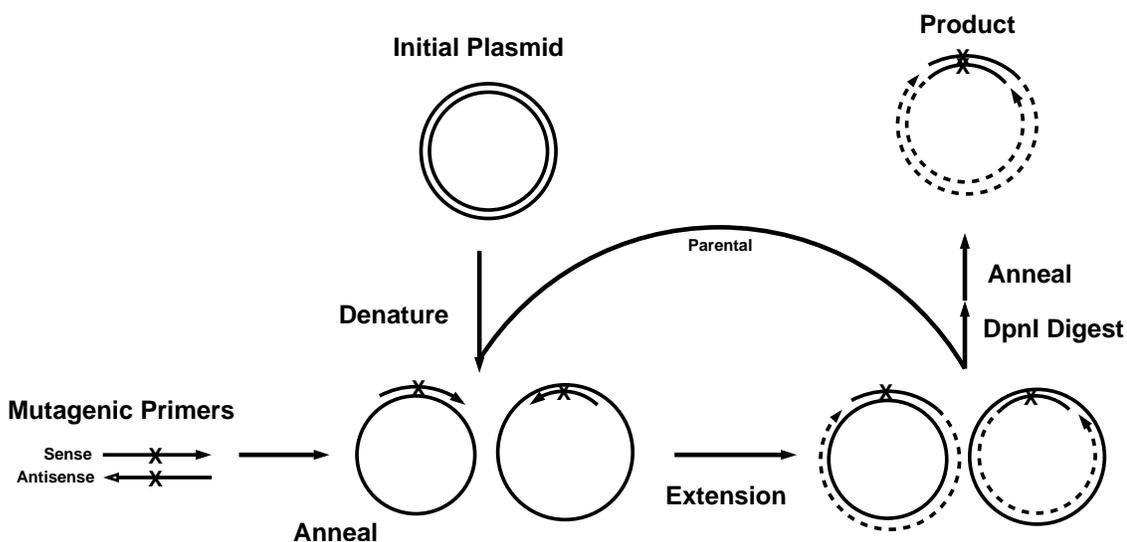
Figure 3.3: Reaction scheme for site directed mutagenesis. Note that only the parental strand is recycled in the reaction loop, so the reaction is a linear rather than exponential amplification technique.

in a total volume made up to 12.5 μl with water. The template DNA needs to be methylated, so a bacterial strain deficient in DNA methylation such as JM101 cannot be used as the plasmid source. The reaction protocol was 5 min at 95°C, followed by 18 cycles of 50 s at 95°C, 50 s at 60°C, 1 min+1 min/1 kb of template at 68°C, and finished with 7 min at 68°C.

Following the reaction the product was digested with 5 U of DpnI (New England Biolabs, Ipswich, MA), a restriction enzyme that cuts only at methylated sites. This digests the template plasmid while leaving the PCR product intact. Since the transformation efficiency of circular template plasmid is several orders of magnitude better than linear PCR products, the DpnI digest is crucial to avoid large numbers of contaminating colonies containing the parental template. Following digestion the DNA was transformed directly into a cell line proficient in accepting unmethylated DNA (e. g. XL-1).

3.1.2 Random Mutagenesis

Random mutagenesis was accomplished using Mutazyme II (Stratagene). Mutazyme II is a mix of Mutazyme DNA polymerase (Stratagene) and a *Taq* DNA polymerase mutant that is touted to produce a more even blend of mutations than normal *Taq* polymerase under error-prone conditions. Mutazyme II proved to be less tolerant of poor primer design than standard polymerases. Additionally, the instructions from Stratagene appear to be misguided as to the

amount of template to begin with to achieve a given mutation rate. After many iterations, the following conditions were found to achieve a mutation rate of ∼5 base pairs/kb.

5 μl	10x Mutazyme buffer
2.5 μl	dimethyl sulfoxide (DMSO)
1 μl	dNTP mix (10 mM each)
1 μl	forward primer (100 pmoles/μl)
1 μl	reverse primer (100 pmoles/μl)
3 pg	template DNA
1 μl	Mutazyme
x μl	H$_2$O
50 μl	total

Note that template DNA refers only to the portion of DNA that is being mutated. For instance, to mutate the pBAD-pelB-RLuc8 plasmid, 15 pg of plasmid would be used (∼3 pg of *rluc8*).

The PCR protocol was:

Segment	# Cycles	Temperature	Duration
1	1	95°C	2 min
2	50	95°C	1 min
		Primer T$_m$	1 min
		72°C	1 min/1 kb template + 1.5 min
3	1	72°C	10 min

Following PCR, the product was purified and digested overnight in a 37°C bacterial incubator with DpnI (to remove parental template) along with the appropriate restriction enzymes (NcoI/SalI) for insertion into the plasmid backbone. After gel purification, a total of 200 ng of mutated insert and plasmid backbone (SalI/NcoI digested pBAD) at a 2:1 insert/backbone molar ratio were ligated overnight at 16°C in a 20 μl reaction. 5 μl of the ligated product was then transformed into 50 μl Top10 bacteria cells (Invitrogen), and spread on 8 150 mm diameter Terrific Broth/50 μg/ml ampicillin (TB/Amp) agar plates containing 0.2% L-(+)-arabinose. Following 20 h of incubation at 32°C, the plates were airbrushed with a phosphate buffered saline (PBS) solution containing 1% 0.5 mg/ml coelenterazine in propylene glycol and imaged immediately using an IVIS 200 bioluminescence imaging system (Xenogen). Three 5 s acquisitions were made using a DsRed, a GFP, and an open filter. Acquired images were processed in GNU Octave [54] using a collection of custom scripts. Colonies were selected both automatically with these scripts as well as manually for brightness and/or spectral shifts. An example of the results of this image processing and colony selection is shown in Figure 3.4. Colonies were then screened further as described in section 3.1.4.

(a) Photograph

(b) Open Emission

(c) Synthetic Image
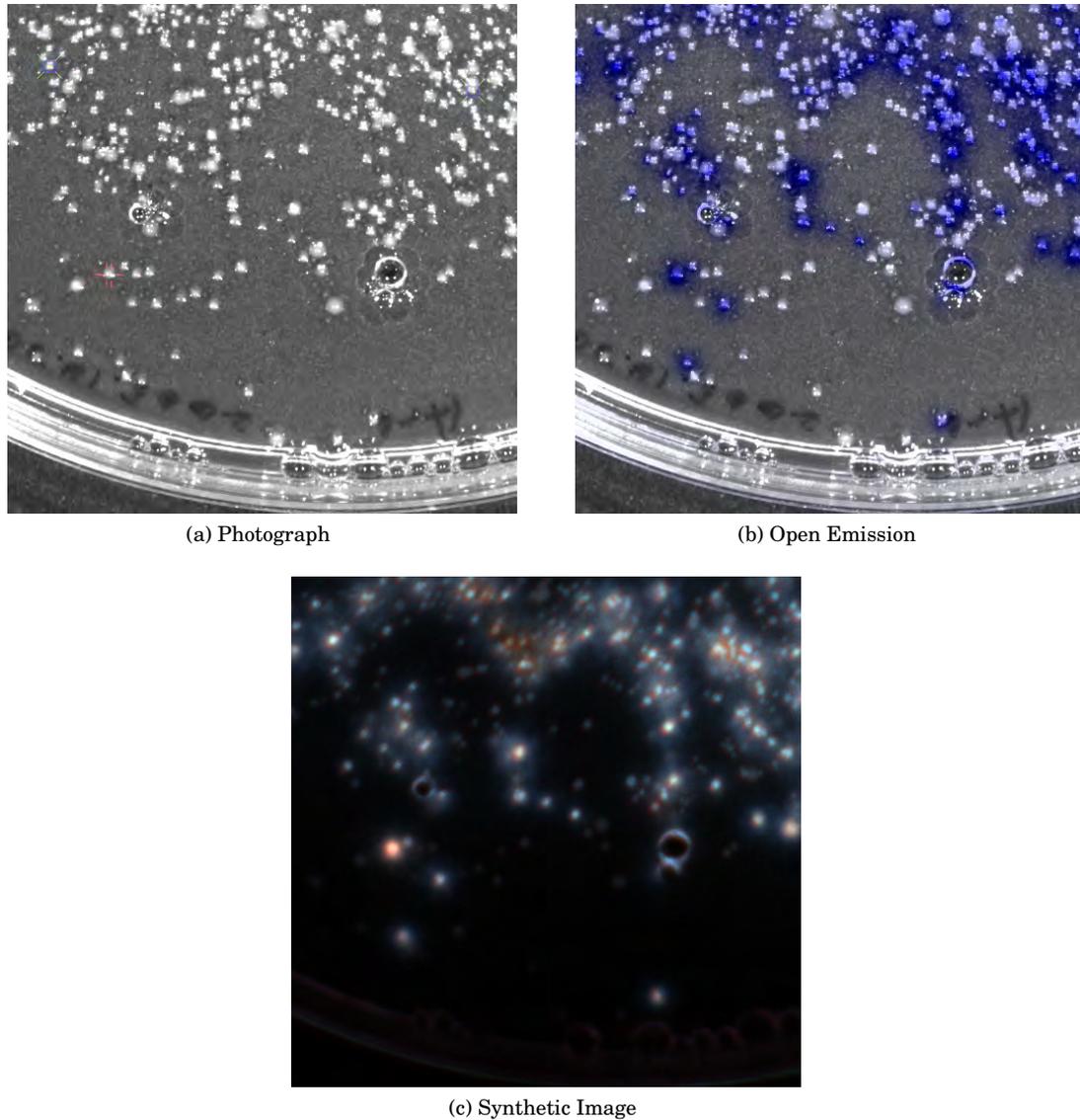
Figure 3.4: An example of bacterial colonies being screened for color shifts and increases in brightness. Panel (a) shows a portion of a plate that contained several thousand colonies. After airbrushing with coelenterazine, emission images were taken using open, GFP, and DsRed filters. Panel (b) shows the results of super-imposing the open emission filter on the photograph. Panel (c) shows a synthetic image, where the open, GFP, and DsRed acquisitions have been combined as the image's blue, green, and red channels, respectively. A computer algorithm was used for ranking the brightness of the colonies in each of the acquisitions. Colonies ranked at the top in the open emission, the GFP filter, and the DsRed filter acquisitions are marked in the photograph shown in Panel (a) with blue squares, green diagonals, and red crosses, respectively. The colony marked with the red crosses in Panel (a) (erroneously detected as two colonies by the computer algorithm) is the red-shifted variant RLuc8/A22P/D162N/F261W (Table 3.4).

3.1.3 Saturation Mutagenesis

Saturation mutagenesis at specific locations was performed by making use of Type IIs restriction enzymes and primers containing a randomized codon sequence. The method presented here is a modification of a previously published protocol [101] and differs mainly in that the entire plasmid is used as the template for PCR obviating the need for a second ligation step. The basic scheme is outlined in Figure 3.5.

Standard PCR conditions were used, with the exception that the extension time was increased to 2 min/kb, 5% DMSO was included in the reaction, template was used at 10% of usual concentrations, and primers were used at 20% of usual concentrations. Restriction digests using BpiI (Fermentas, Hanover, MD) and DpnI were performed overnight in a 37°C bacterial incubator. Following gel purification, ligation, transformation, and plating was done as in the random mutagenesis case.

3.1.4 Small Scale Protein Expression

For random and saturation mutagenesis experiments, the clones initially selected from the agar plates were further screened by small scale expression experiments. Selected colonies were picked into 2 ml TB/Amp each and grown to saturating conditions at 37°C (~12 h). 2 ml TB/Amp containing 0.2% L-(+)-arabinose was then added to each tube and the cultures were grown an additional 12 h at 32°C. Following this, half of each culture was spun down and submitted to the osmotic shock protocol as described in Section 2.1.5.

The periplasmic fractions were assayed for specific activity, assayed for bioluminescence color shifts visually, and stored at 4 °C. Bright and/or color shifted variants were then submitted for sequencing. For variants identified as having novel mutations, the periplasmic fraction was brought to 1xWB (Section 2.1.5) from a 10x stock, and further purified using nickel affinity spin columns (Ni-NTA Spin Kit, Qiagen) with 1xEB as the elution buffer. The elution was brought to 1% HSA, and then assayed for specific activity.

Interesting color shifts were confirmed by spectrophotometry as described in Section 2.1.6. As before, emission spectra were filtered as necessary and normalized to equalize the total area under the curve.
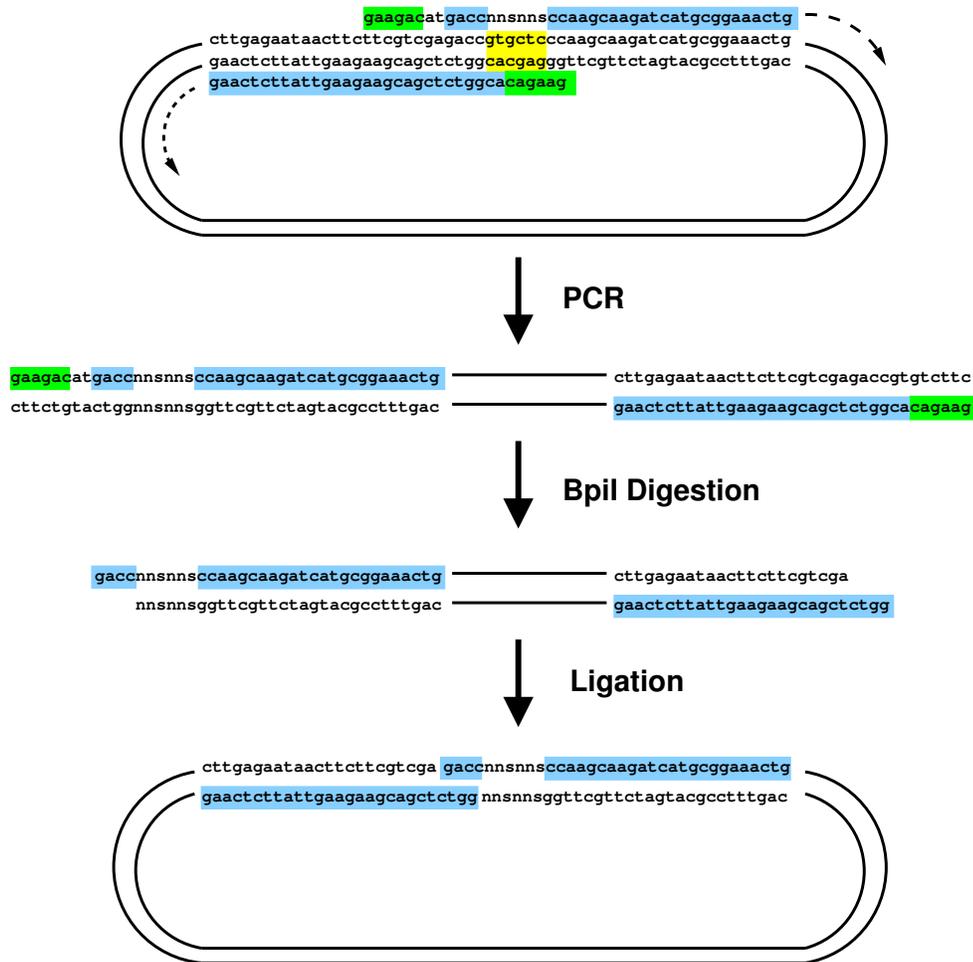
Figure 3.5: A schematic diagram of a saturation mutagenesis scheme using a type IIs restriction enzyme. In this particular example the saturation target, shown in yellow, is V185/L186. The restriction enzyme recognition sites for BpiI are shown in green. Random nucleotides are represented as follows: N=A/T/C/G, and S=C/G. For an NNS sequences, amino acids with only one codon have a 1/32 chance of being used. To have at least a 99% probably of hitting a given pair of codons, ~7000 colonies need to be screened.

	Specific Activity (relative to RLuc)				Serum Inact. $\tau_{1/2}$ (h)		Wavelength (nm)			
	native	*bc*	*cp*	*n*	*bdc*	mouse	rat	peak	mean	fwhm	%>600
RLuc8	4.3±0.2	3.0	5.8	8.8	59	281±49	86 ±9	486	503	94	4
Consensus Mutations											
RLuc8/I75A	8.0	4.4	9.0	15	89	580	277	485	503	93	4
RLuc8/N109D	6.5	3.6	6.0	10	71	364	76	482	501	94	3
RLuc8/P111G	6.8	4.2	7.4	13	82	31	8.9		ND		
RLuc8/Y131K	5.6	3.4	6.1	10	72		ND		ND		
RLuc8/E132R	5.5	3.2	6.3	11	57		ND		ND		
RLuc8/I137V	7.3	4.5	7.6	13	83	202	33	484	502	93	4
RLuc8/V149P	0.66	0.62	0.92	0.90	8.6		ND		ND		
RLuc8/T184G	6.1	3.2	5.9	1.2	84		ND	486	505	93	4
RLuc8/T184R	4.6	2.9	4.7	6.0	53		ND	483	502	92	4
RLuc8/S188GK189A	1.3	1.5	0.42	0.46	25		ND		ND		
RLuc8/V212D	6.4	3.8	6.6	12	78	308	54		ND		
RLuc8/N264T/A265GR	0.03	0.02	0.02	0.03	0.05		ND		ND		
RLuc8/V267R/E268D/G269I	3.1	1.4	5.3	7.4	54		ND		ND		
RLuc8/K271R/K272S	6.2	3.3	7.7	11	70		ND		ND		
RLuc8/K282P	5.2	3.2	5.2	8.7	75		ND		ND		
Combined Mutations											
RLuc8/N109D/I137V	4.7	2.6	4.8	7.6	47		ND		ND		
RLuc8/I75A/N109D/I137V	5.7	3.2	6.5	9.5	69		ND		ND		

Table 3.1: Results of further consensus guided mutagenesis on RLuc8. The data for RLuc8 is repeated from Table 2.1 for the purpose of comparison. Substrate abbreviations are as shown in Figure 2.1. Spectra measurements were made using coelenterazine. Specific activities are relative to that of RLuc (Table 2.1). FWHM - full width at half maximum. ND - not determined. Inact. - inactivation.

3.2 Results

3.2.1 Consensus Driven Mutagenesis of RLuc8

Further consensus sequence directed mutagenesis, with residues targeted in a manner identical to what was done in Section 2.2.2, was performed using RLuc8 as the parental template. These results (Table 3.1) showed that some improvements could still be achieved in terms of light output and/or resistance to inactivation in murine serum. However, combining these different identified mutations (e.g. N109D/I137V) led to reductions in the parameters being optimized.

3.2.2 Probing of the Active Site of RLuc8

In an attempt to rationally alter the emitted wavelength of *Renilla* luciferase, the location and orientation of the substrate in the active pocket was conjectured. This was done by assuming the catalytic triad was used for coordinating the oxygen, that the orientation of the substrate

would be similar to that seen with other α/β-hydrolases (e. g. Figure 7.6b), and that the vary-
ing affinities of the different mutations in Chapter 2 for the various coelenterazine analogs
was due to close interactions between the mutation and the altered side chain of the analog.
This conjecture was initially formulated using the Swiss-Model derived RLuc structure (Sec-
tion 2.2.1), but was switched over to the RLuc8 crystallographic structure (Section 4.3) once
it became available. The results of this exercise are shown in Figure 3.6. The correspond-
ing amino acids of the putative active pocket are marked in the primary sequence shown in
Figure 3.7.

Using this model of coelenterazine/coelenteramide in the active pocket as a guide, a total of
74 site specific mutations were made at the 22 residues thought to interact with the substrate.
With the exception of the I223 location at which mutagenesis was saturating, the subset of
possible mutations done at each residue was selected based on what would be considered "safe"
with respect to the tertiary fold of the enzyme [22]. The results of this screen are shown in
Tables 3.2 and 3.3, with the resultant color shifts shown in Figure 3.8. From this screen,
a total of 21 mutations at 10 different residue locations resulted in observable shifts in the
emission spectrum. The variants with bathochromic shift mutations presumably have active
pockets that favor the green fluorescing anion form of coelenteramide. Unsurprisingly, given
that the enzymatic pocket of RLuc8 is already evolved for the reaction at hand, nearly all these
mutations led to significant reductions in the light output of the luciferase.

3.2.3 Random Mutagenesis on RLuc8

Round 1 - Random Mutagenesis on RLuc8/F261W, RLuc8/F262W

The original purpose of the random mutagenesis study was to take some of the red-shifted mu-
tations identified in the active pocket site directed mutagenesis study and improve their cat-
alytic abilities. The RLuc8/F261W and RLuc8/F262W variants were picked as starting points,
as they both gave rise to appreciable bathochromic shifts while not overly compromising the
light output of the luciferase. Random mutagenesis was performed on these templates, and
screened in bacteria for both increases in light output as well as emission color shifts. The re-
sults of this study are given in Table 3.4, and the emission spectra of several of these variants
are shown in Figure 3.9. Interestingly, several mutation locations were overrepresented. Of
these mutations, those at residues E155 and G269 lead increases in light output, and those at
D162 lead to further bathochromic shifts in the emission spectra.
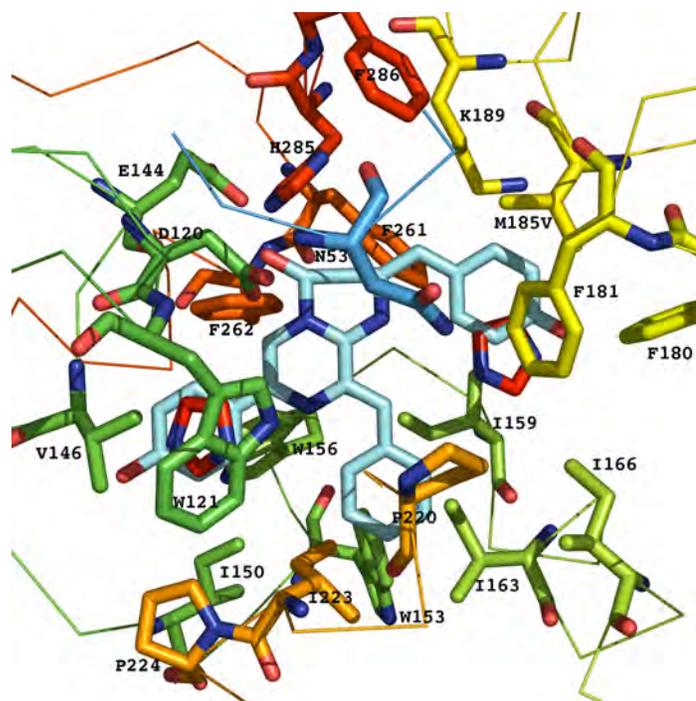
(a) Estimated Coelenterazine Orientation
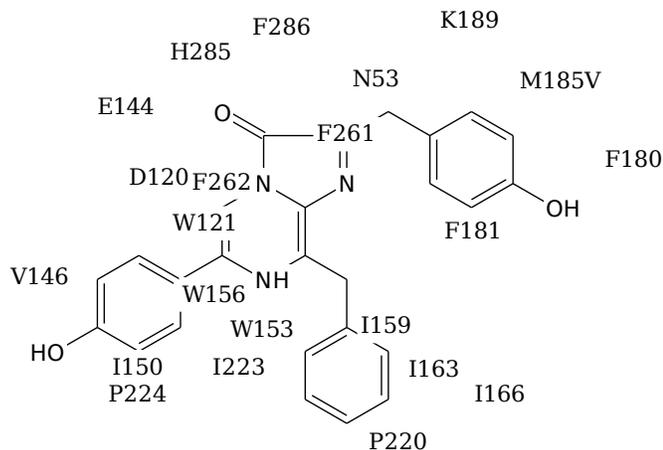
(b) Putative Interacting Residues

Figure 3.6: A guess for the coelenterazine orientation in the RLuc8 site. The orientation of coelenterazine in the context of the RLuc8 structure shown in Panel (a) is guessed based primarily on the location of the presumptive catalytic triad residues and the effect that the M185V mutation had on the various coelenterazine analogs tested (Figure 2.1). The RLuc8 structure used in this figure is described in section 4.3. The two imidazoles that were present in this structure are shown in red. Panel (b) diagrammatically shows the residues that putatively interact with the substrate.

```
RLuc8       MASKVYDPEQRKRMITGPQWWARCKQMNVLDSFINYYDSE...KH   42
RLuc        MTSKVYDPEQRKRMITGPQWWARCKQMNVLDSFINYYDSE...KH   42
RmLuc       MTSKVYDPELRKRMITGPQWWARCKQMNVLDSFINYYDSE...KH   42
XP_787066   AAAMGSRNQSTIPLVTADEWWGKCKKVDVLGEKMSYYDSDPQNSS   75
XP_794218.1 ...MASRNQATIPLVTADEWWGKCKKVDVLGSKMSYYDSDPQNRS   42


RLuc8       AENAVIFLHGNATSSYLWRHVVPHIEPVARCIIPDLIGMGKSGKS   87
RLuc        AENAVIFLHGNAASSYLWRHVVPHIEPVARCIIPDLIGMGKSGKS   87
RmLuc       AENAVIFLHGNAASSYLWRHVVPHVEPVARCIIPDLIGMGKSGKS   87
XP_787066   SKHAVFLHGNPTSSYLWRNVMPQVEPIARCLAPDLIGQGRSNKL   120
XP_794218.1 GKHTAVFLHGNPTSSYLWRNVIPQVEPIARCLAPDLIGMGRSDKL   87


RLuc8       GNGSYRLLDHYKYLTAWFELLNLPKKIIFVGHDWGAALAFHYAYE   132
RLuc        GNGSYRLLDHYKYLTAWFELLNLPKKIIFVGHDWGACLAFHYSYE   132
RmLuc       GNGSYRLLDHYKYLTEWFKHLNLPKKIIFVGHDWGACLAFHYCYE   132
XP_787066   ANHSYRFVDHYRYLSAWFDSVNLPEKVCIVCHDWGSGLGFHWCNE   165
XP_794218.1 ASRSYRFLDHYRYLSAWFDALKLPEKITVVCHDWGTALGLHWCSE   132


RLuc8       HQDRIKAIVHMESVVDVIESWDEWPDIEEDIALIK.SEEGEKMVL   176
RLuc        HQDKIKAIVHAESVVDVIESWDEWPDIEEDIALIK.SEEGEKMVL   176
RmLuc       HQDRIKAVVHAESVVDVIESWDEWPDIEEDIALIK.SEEGEKMVL   176
XP_787066   HRDRIEGLIHMESVVAPVPGWDRFPDMAKDFFQVLRSEAGDDLVL   210
XP_794218.1 HRDRLEAIVHMEGVLKPMT.WDIFPDSMRDIFLALRSDAGEEMIL   176


RLuc8       ENNFFVETVLPSKIMRKLEPEEFAAYLEPFKEKGEVRRPTLSWPR   221
RLuc        ENNFFVETMLPSKIMRKLEPEEFAAYLEPFKEKGEVRRPTLSWPR   221
RmLuc       ENNFFVETMLPSKIMRKLEPEEFAAYLEPFKEKGEVRRPTLSWPR   221
XP_787066   QKNYFIELLLPRAIMRELRPEEMDAYREPFKNPGEDRRPTLTWPR   255
XP_794218.1 KKNMFIETILPLAIKRKLRQEEMDAYREPFKNPGEDRRPLLTFPR   221


RLuc8       EIPLVKGGKPDVVQIVRNYNAYLRASDDLPKLFIESDPGFFSNAI   266
RLuc        EIPLVKGGKPDVVQIVRNYNAYLRASDDLPKMFIESDPGFFSNAI   266
RmLuc       EIPLVKGGKPDVVEIVRNYNAYLRASHDLPKMFIESDPGFFSNAI   266
XP_787066   EIPIKGDGPDDVIAIASSYNAWLKESADLPKLYIHAKPGFFSEGI   300
XP_794218.1 QIPIQGEGPEETVAIATAYHAWIKGTADLPKFCILATPGIFSEWG   266


RLuc8       VEGAKKFPNTEFVKVKGLHFLQEDAPDEMGKYIKSFVERVLKNEQ   311
RLuc        VEGAKKFPNTEFVKVKGLHFSQEDAPDEMGKYIKSFVERVLKNEQ   311
RmLuc       VEGAKKFPNTEFVKVKGLHFSQEDAPDEMGNYIKSFVERVLKNEQ   311
XP_787066   KKGIANWPNQKTVESEGLHFLQEDSPIQIGDHVKDFLSALYK...   342
XP_794218.1 TGITKDWPNHKVVQVEGSHFFQEDSPIQTGDYIKEFLSSVFK...   308
```

X non conserved
X similar
X conserved
X all match

Figure 3.7: Sequence of RLuc8 with residues surrounding the putative active pocket marked. The putative active pocket residues are indicated by the black bar above the sequence. RLuc8 has been aligned with RLuc, the luciferase from *Renilla mülleri* (RmLuc), and two proteins of unknown function predicted from the *Strongylocentrotus purpuratus* genome (XP_787066 - GenBank Identifiers GI:72160391, XP_794218.1 - GI:72149470). Note that RLuc8 and associated variants used in this chapter have had the N-terminal methionine replaced by the pelB leader sequence, and have a valine/aspartate (SalI restriction site) and a 6x-His tag on the C-terminus.

| | Specific Activity (relative to RLuc) | | | | | Wavelength (nm) | | | |
	native	h	cp	n	bdc	peak	mean	fwhm	%>600 nm
RLuc8	4.3±0.2	3.0	5.8	8.8	59	486	503	94	4
Active Pocket Mutations									
RLuc8/N53D	0.002	0.004	0.002	0.001	0.01			ND	
RLuc8/N53Q	0.10	0.16	0.52	0.04	0.31	475	491	92	3
RLuc8/N53S	0.001	0.002	0.003	0.005	0.03			ND	
RLuc8/W121F	0.05	0.02	0.02	0.03	0.15	478	496	94	3
RLuc8/W121H	0.003	0.002	0.004	0.01	0.02			ND	
RLuc8/W121Y	0.003	0.007	0.01	0.01	0.01			ND	
RLuc8/V146I	1.1	1.1	0.60	0.50	21	484	502	95	4
RLuc8/V146M	1.0	0.66	0.51	0.47	0.43	481	498	94	3
RLuc8/V146W	0.000	0.000	0.000	0.000	0.000			ND	
RLuc8/I150F	0.51	0.56	0.50	1.0	5.3	485	504	95	4
RLuc8/I150H	0.53	0.25	0.19	2.6	8.1	494	514	98	6
RLuc8/I150M	2.9	1.8	3.1	2.2	24	488	508	95	5
RLuc8/I150W	0.94	0.65	0.62	1.3	8.2	485	503	94	4
RLuc8/I150Y	0.02	0.01	0.02	0.29	0.15	487	506	97	4
RLuc8/W153F	4.9	3.1	4.9	7.9	104	484	502	95	4
RLuc8/W153Y	1.2	0.75	1.1	2.1	15	485	503	96	4
RLuc8/W156F	3.9	2.7	7.2	12	81	486	504	93	4
RLuc8/W156H	0.46	0.48	0.53	2.3	1.2	490	510	96	5
RLuc8/W156Y	3.0	2.6	5.2	9.0	91	483	501	94	4
RLuc8/I159F	0.60	0.43	0.56	1.7	1.4	491	510	101	5
RLuc8/I159H	0.04	0.04	0.02	0.11	0.53	506	526	108	10
RLuc8/I159W	0.12	0.13	0.08	0.44	0.28	490	508	104	6
RLuc8/I159Y	0.003	0.002	0.01	0.05	1.1	513	536	113	13
RLuc8/I163F	0.57	0.80	0.39	0.61	7.6	483	502	95	4
RLuc8/I163H	0.16	0.21	0.12	0.19	6.1	499	519	102	7
RLuc8/I163W	0.16	0.26	0.15	0.26	2.1	498	517	103	7
RLuc8/I163Y	0.13	0.16	0.10	0.14	1.2	502	521	103	8
RLuc8/I166F	1.3	1.2	1.1	1.8	11	483	501	96	4
RLuc8/I166H	0.04	0.08	0.03	0.05	0.52	483	502	100	4
RLuc8/I166L	4.4	2.4	5.3	9.5	55	486	506	92	4
RLuc8/I166W	0.004	0.01	0.003	0.01	0.20	498[†]	508	110	7
RLuc8/I166Y	0.23	0.35	0.16	0.34	2.5	493	508	99	5
RLuc8/F180I	0.63	0.65	0.62	0.70	8.7	486	504	101	5
RLuc8/F180W	4.0	2.4	3.3	4.6	45	485	502	93	4
RLuc8/F180Y	3.0	2.2	2.6	3.4	52	484	499	105	4
RLuc8/F181W	0.07	0.05	0.05	0.13	4.8	479	494	95	3
RLuc8/F181Y	0.07	0.11	0.03	0.06	1.2	497	515	103	6
RLuc8/K189E	4.4	2.6	3.8	6.7	61	484	501	95	4
RLuc8/K189H	3.6	2.0	1.8	6.1	44	485	502	94	4
RLuc8/K189I	1.1	1.0	4.4	1.9	19	484	500	96	4
RLuc8/K189R	0.70	0.86	0.58	0.45	1.1	484	502	93	4
RLuc8/P220H	0.003	0.003	0.003	0.01	0.05			ND	
RLuc8/P224H	0.08	0.03	0.07	0.004	0.29	484	500	95	4
RLuc8/Y240F	5.5	2.0	4.4	5.3	48	484	502	92	4
RLuc8/F261W	0.20	0.38	0.38	0.02	0.76	504	524	98	8
RLuc8/F261Y	0.07	0.93	0.53	0.01	16	487	506	97	4
RLuc8/F261W/F262W	0.000	0.000	0.00	0.002	0.003	512	531	115	11
RLuc8/F262W	0.60	0.20	0.23	0.07	0.02	500	521	99	7
RLuc8/F262Y	0.01	0.01	0.01	0.001	0.04	511	532	104	10
RLuc8/F286W	0.08	0.11	0.07	0.04	0.24	481	499	92	3
RLuc8/F286Y	0.07	0.08	0.23	0.05	0.24	482	501	93	4

Table 3.2: Results of site-directed mutagenesis in the active pocket of RLuc8. The data for RLuc8 is repeated from Table 2.1 for the purpose of comparison. Results for I223 are shown in Table 3.3. Y240 is not believed to be in the active pocket, but is proximal to it. Substrate abbreviations are as shown in Figure 2.1. Spectra were measured using coelenterazine. Specific activities are relative to that of RLuc (Table 2.1) and were not corrected for the luminometer's wavelength dependent sensitivity. [†]RLuc8/I166W showed a shoulder peak at 415 nm that was 28% of the height of the main peak at 498 nm. FWHM - full width at half maximum. ND - not determined.

| | Specific Activity (relative to RLuc) | | | | | Wavelength (nm) | | | |
	native	h	cp	n	bdc	peak	mean	fwhm	%>600 nm
RLuc8	4.3±0.2	3.0	5.8	8.8	59	486	503	94	4
I223 Mutations									
RLuc8/I223A	0.68	0.37	0.94	0.51	2.0			ND	
RLuc8/I223C	3.0	1.7	6.2	3.7	12	503	524	103	9
RLuc8/I223D	0.01	0.01	0.01	0.06	0.10	503	524	106	9
RLuc8/I223E	0.01	0.01	0.01	0.11	0.21	497	517	104	7
RLuc8/I223F	2.7	2.1	1.8	2.5	10	486	505	92	4
RLuc8/I223G	0.14	0.07	0.12	0.33	1.3	498	518	105	7
RLuc8/I223H	0.07	0.05	0.09	0.43	1.9	508	527	105	9
RLuc8/I223K	0.002	0.002	0.001	0.003	0.26	491	509	97	5
RLuc8/I223L	1.3	1.8	1.2	1.4	16	483	502	95	4
RLuc8/I223M	0.19	0.37	0.25	0.64	14	501	521	98	7
RLuc8/I223N	0.30	0.31	0.33	0.48	1.7	505	527	102	9
RLuc8/I223P	0.01	0.01	0.01	0.03	0.13	486	505	96	5
RLuc8/I223Q	0.06	0.06	0.06	0.31	2.3	505	527	102	9
RLuc8/I223R	0.003	0.002	0.002	0.01	0.24	505	528	108	10
RLuc8/I223S	0.61	0.45	0.49	0.71	1.7	499	517	104	7
RLuc8/I223T	0.43	0.36	0.33	0.48	0.54	499	517	102	7
RLuc8/I223V	2.5	1.9	2.8	3.0	31	490	509	97	5
RLuc8/I223W	0.003	0.004	0.01	0.02	0.01	484	508	104	7
RLuc8/I223Y	0.02	0.02	0.02	0.04	0.07	486	505	97	4
Double Mutants									
RLuc8/I223C/F261W	0.12	0.12	0.43	0.01	0.03	511	529	104	8
RLuc8/I223C/F262W	0.07	0.02	0.09	0.01	0.01	511	529	100	8
RLuc8/I223H/F261W	0.002	0.003	0.005	0.01	0.02			ND	
RLuc8/I223H/F262W	0.001	0.001	0.001	0.01	0.01			ND	
RLuc8/I223M/F261W	0.02	0.01	0.01	0.01	0.05			ND	
RLuc8/I223M/F262W	0.002	0.001	0.001	0.01	0.001			ND	
RLuc8/I223Q/F261W	0.003	0.005	0.003	0.01	0.02			ND	

Table 3.3: Results of saturation mutagenesis on RLuc8 at the putative active pocket residue of I223, along with some double mutants. The data for RLuc8 is repeated from Table 2.1 for the purpose of comparison. Substrate abbreviations are as shown in Figure 2.1. Spectra were measured using coelenterazine. Specific activities are relative to that of RLuc (Table 2.1) and were not corrected for the luminometer's wavelength dependent sensitivity. FWHM - full width at half maximum. ND - not determined.

(a)

(b)
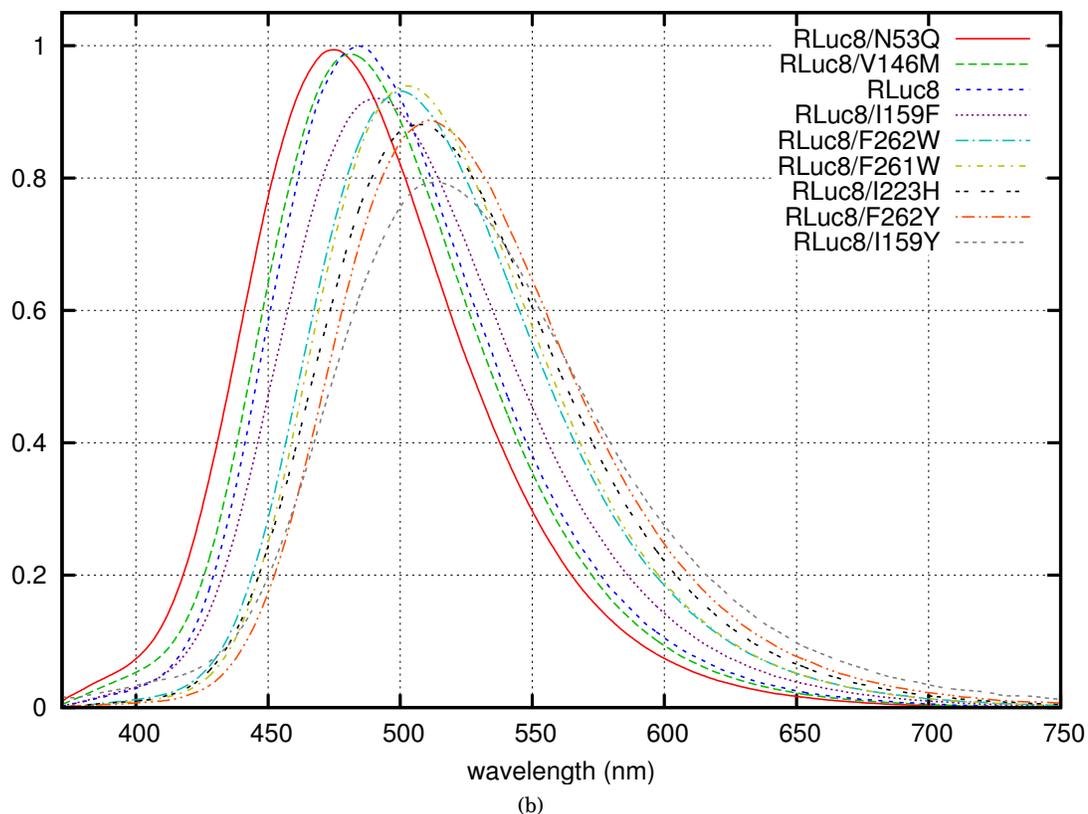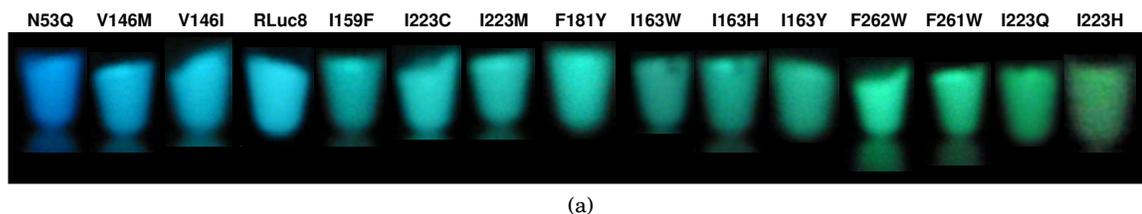
Figure 3.8: Normalized bioluminescence emission spectra and color photographs for several of the active site mutants of RLuc8. The images shown in Panel (a) were made by mixing purified protein with coelenterazine, and photographing the resultant emission with a standard consumer-grade digital camera. Normalized emission spectra resulting from the catalysis of coelenterazine are shown in Panel (b). The normalization equalized the total area under the curve.

F261W Mutants	Specific Activity	Wavelength (nm)			%>600 nm
		peak	mean	fwhm	
RLuc8/F261W	0.20	505	524	98	8
RLuc8/R11P/F261W/V267I	0.25	501	522	99	7
RLuc8/A22P/**D162**N/F261W	0.10	526	547	97	13
RLuc8/V63I/**L94**F/F261W/F278I	0.36	501	522	98	7
RLuc8/R93L/**D162**E/F261W	0.05	535	551	112	17
RLuc8/**L94**F/F261W	0.15	501	521	97	7
RLuc8/K113R/**E155**K/F261W	0.54	503	523	98	7
RLuc8/**A123**S/F261W	0.48	504	523	98	7
RLuc8/M143T/F261W	0.26	503	523	98	7
RLuc8/**D162**N/F261W/S188N	0.05	525	544	97	13
RLuc8/A164T/D248E/F261W/K297N	0.12	502	521	98	7
RLuc8/F261W/N264S	0.27	503	523	99	7
RLuc8/F261W/K271R	0.14	505	526	99	8
F262W Mutants					
RLuc8/F262W	0.60	500	521	99	7
RLuc8/Q26K/**E155**K/F262W	0.75	500	521	99	7
RLuc8/P65H/A130T/F262W	0.12	501	522	99	8
RLuc8/F105V/E151K/**D162**E/F262W	0.07	535	551	119	18
RLuc8/**A123**S/F262W	0.50	499	519	98	7
RLuc8/**E155**G/E183D/F262W	0.89	501	522	98	7
RLuc8/**E155**K/E169D/F262W	0.60	501	521	98	7
RLuc8/K167M/K173N/F262W	0.54	498	519	98	7
RLuc8/V234I/F262W/**G269**R	0.86	501	523	99	8
RLuc8/F262W/**G269**E	0.57	500	520	98	7
RLuc8/F262W/**G269**R	1.3	502	523	100	8
RLuc8/F262W/M295V	0.47	499	520	98	7

Table 3.4: Results of random mutagenesis on RLuc8/F261W and RLuc8/F262W. Mutation locations that showed up multiple times are designated by bold text. The data for RLuc8/F261W and RLuc8/F262W is repeated from Table 3.2 for the purpose of comparison. ~45,000 and ~30,000 colonies were screened for the random mutagenesis on RLuc8/F261W and RLuc8/-F262W, respectively. Coelenterazine was used for measuring the spectra and specific activity. Specific activities are relative to that of RLuc (Table 2.1) and were not corrected for the luminometer's wavelength dependent sensitivity. FWHM - full width at half maximum.
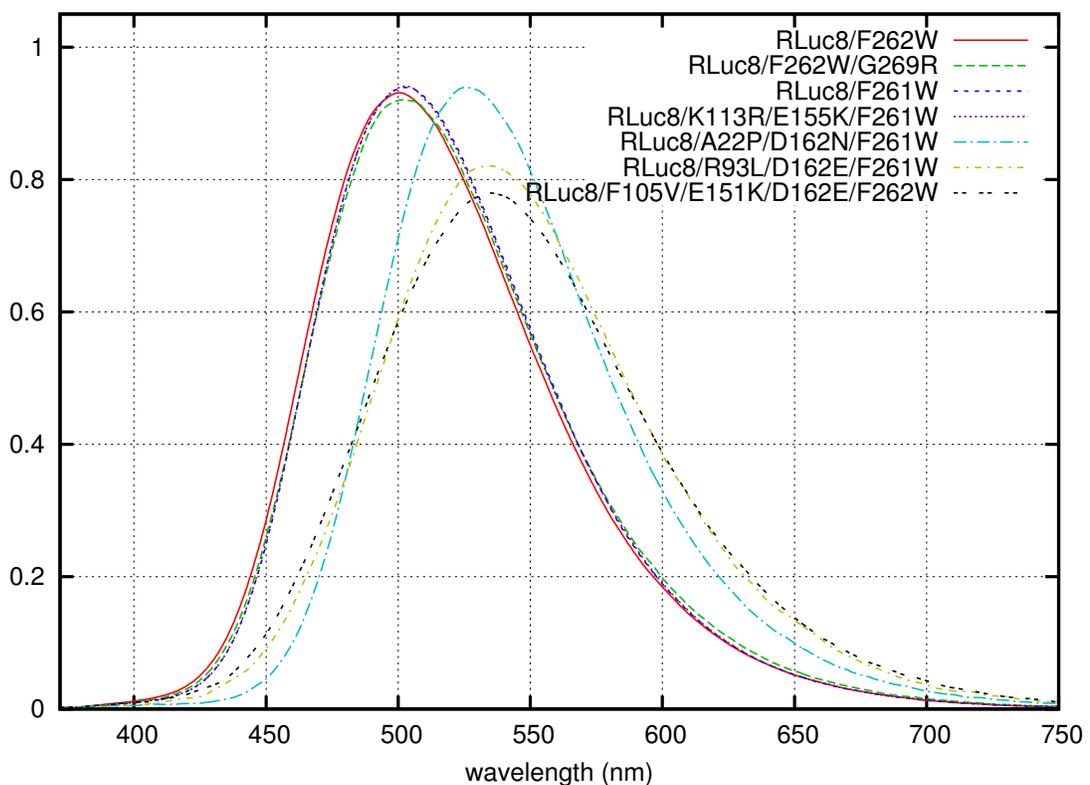
Figure 3.9: Normalized bioluminescence emission spectra for several variants of RLuc8/F261W and RLuc8/F262W obtained by random mutagenesis. Coelenterazine was used for obtaining all spectra. The RLuc8/F261W and RLuc8/F262W curves are repeated from Figure 3.8b.

Round 2 - Saturation Mutagenesis at D162/I163

As the A123S mutation showed some potential for increasing the light output of RLuc8/F261W, it was incorporated into RLuc8, RLuc8/F261W, and RLuc8/F262W by site specific mutation. Although this mutation is somewhat detrimental for RLuc8 and RLuc8/F262W, it was retained with the idea that a later round of mutagenesis would be done at the F261 site.

These three A123S containing templates, along with RLuc8, were used for saturation mutagenesis at the D162 residue identified in the previous random mutagenesis screen. As the D162 residue borders the active pocket residue I163, this residue was incorporated into the saturation mutagenesis screen as well. Results of saturation mutagenesis at D162/I163 performed on the templates RLuc8, RLuc8/A123S, RLuc8/A123S/F261W, and RLuc8/A123S/F262W are given in Table 3.5. Emission spectra for several of these variants are shown in Figures 3.10, 3.11, and 3.12, with color photographs of select variants given in Figure 3.13.

Quite surprisingly, a single point mutation (D162E) could lead to a significant red-shift of the emission spectra of the luciferase without a severe compromise in the luciferase's ability to output light. Also interesting, is that several of the selected mutants had significant side peaks around 410 nm. This side peak is presumptively emanating from the neutral species of coelenteramide (Figure 3.2a). For RLuc8/A123S/D162L/I163V, it gives the variant a whitish-purple color when the bioluminescence is visualized.

Round 3 - Saturation Mutagenesis at F261/F262

Using RLuc8/A123S and RLuc8/A123S/D162E/I163L as templates, saturation mutagenesis was done at the F261/F262 residues in an attempt to find the best color shift residues at this location. Out of ~15,000 colonies screened from the mutagenesis with RLuc8/A123S as the template, only parental, RLuc8/A123S/F261W, and RLuc8/A123S/F262W clones were selected by the screen. Out of ~9000 colonies screened from the mutagenesis performed using RLuc8/A123S/D162E/I163L as the template, the only non-parental clone selected by the screening process was RLuc8/A123S/D162E/I163L/F261W. The data for this mutant is included in Table 3.5, with the corresponding emission spectrum shown in Figure 3.9. The results of this screen would indicate that no further improvements in either light output or color shift can be made at the F261/F262 position.

RLuc8 Mutants	Specific Activity	peak	mean	fwhm	shoulder	Shoulder/ Peak Ratio	%>600 nm	Clones
RLuc8	4.3	486	503	94			4	
RLuc8/D162E	1.4	522	537	108			12	1
RLuc8/D162E/I163M	1.2	519	530	108			10	1
RLuc8/D162E/I163T	0.08	539	547	125			19	1
RLuc8/D162N	2.1	510	526	96	408	0.05	8	2
RLuc8/D162N/I163V	1.6	516	531	103	408	0.08	10	6
RLuc8/D162P/I163L	0.28	525	515	103	406	0.50	10	1
RLuc8/D162S/I163V	4.7	485	504	92			4	1
RLuc8/A123S Mutants								
RLuc8/A123S	2.8	484	502	92			4	
RLuc8/A123S/D162C/I163V	0.95	520	539	94			11	1
RLuc8/A123S/D162E	1.5	522	536	107			12	2
RLuc8/A123S/D162E/I163L	2.4	523	538	102			12	2
RLuc8/A123S/D162L/I163V	0.29	532	515	124	409	0.75	13	1
RLuc8/A123S/D162N	2.0	509	526	96	407	0.05	8	1
RLuc8/A123S/D162N/I163L	2.4	507	523	93	404	0.07	7	2
RLuc8/A123S/D162N/I163S	0.19	523	535	110	407	0.18	13	2
RLuc8/A123S/D162T/I163C	0.10	527	514	119	409	0.69	12	1
RLuc8/A123S/F261W Mutants								
RLuc8/A123S/F261W	0.48	504	523	98			7	
RLuc8/A123S/D162T/F261W	0.23	526	547	102			14	1
RLuc8/A123S/D162E/F261W	0.21	533	547	116			16	1
RLuc8/A123S/D162E/I163L/F261W	0.12	538	553	107			17	
RLuc8/A123S/D162N/I163M/F261W	0.29	520	539	95			11	3
RLuc8/A123S/D162N/I163V/F261W	0.21	531	551	102			16	3
RLuc8/A123S/F262W Mutants								
RLuc8/A123S/F262W	0.50	499	519	98			7	
RLuc8/A123S/D162E/I163V/F262W	0.06	541	558	113			21	1
RLuc8/A123S/D162N/F262W	0.23	527	544	100			13	1

Table 3.5: Results from saturation mutagenesis at the D162/I163 residues of RLuc, RLuc8/-A123S, RLuc8/A123S/F261W, and RLuc8/A123S/F262W. If the spectrum had a shoulder peak, it is noted above along with the ratio of the height of the shoulder peak to the main peak. Clones indicates how many colonies coded for the same protein sequences (but not necessarily the same nucleotide sequence), and is an indicator for how well the search space was covered. ~8000, ~4,000, ~8,000, and ~10,000 colonies were screened from the RLuc8, RLuc8/A123S, RLuc8/A123S/F261W, and RLuc8/A123S/F262W saturation mutagenesis reactions, respectively. The data for RLuc, RLuc8/A123S/F261W, and RLuc8/A123S/F262W is repeated from Tables 2.1 and 3.4 for the purpose of comparison. RLuc8/A123S/D162E/I163L/-F261W was actually identified during the work described in Section 3.2.3, but is included here for simplicity. Coelenterazine was used for measuring the spectra and specific activity. Specific activities are relative to that of RLuc (Table 2.1), and were not corrected for the luminometer's wavelength dependent sensitivity (see Table A.1). FWHM - full width at half maximum.
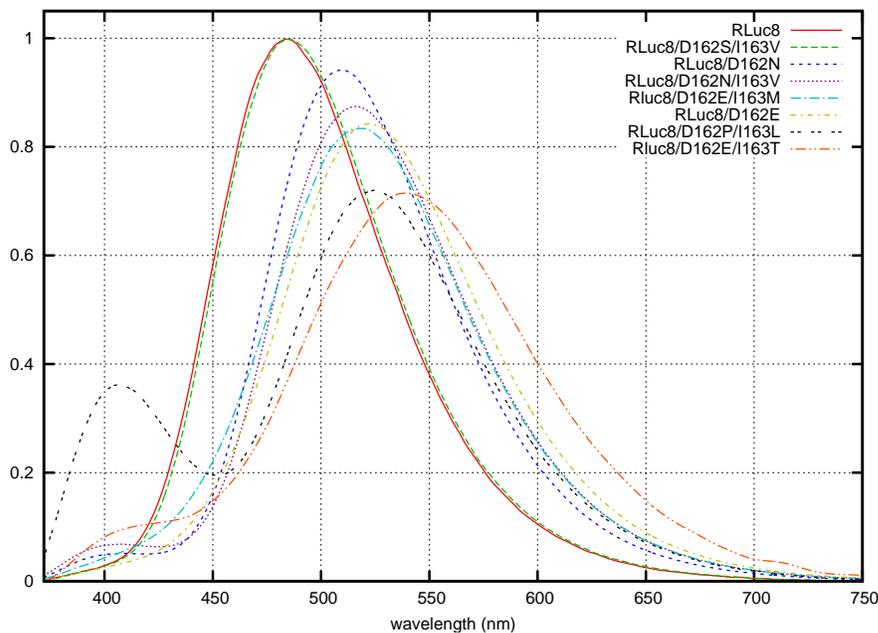
Figure 3.10: Normalized bioluminescence emission spectra for the variants obtained from saturation mutagenesis of RLuc8 at the D162/I163 positions. Coelenterazine was used for obtaining all spectra.
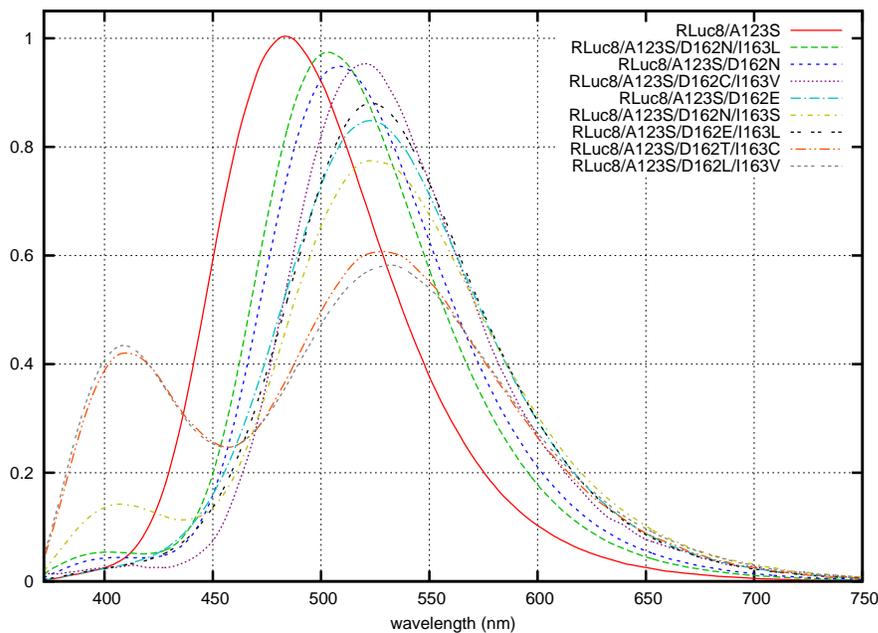
Figure 3.11: Normalized bioluminescence emission spectra for the variants obtained from saturation mutagenesis of RLuc8/A123S at the D162/I163 positions. Coelenterazine was used for obtaining all spectra.
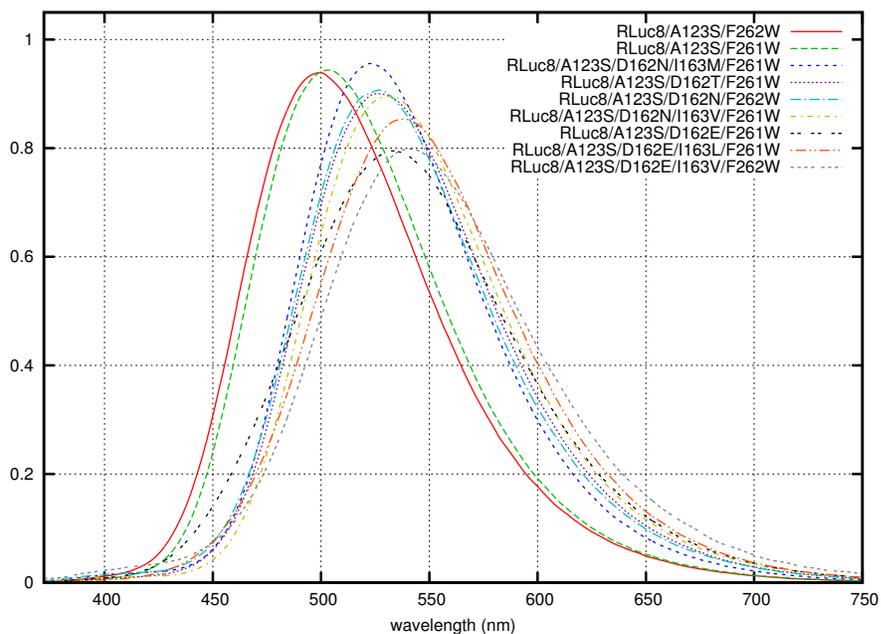
Figure 3.12: Normalized bioluminescence emission spectra for the variants obtained from saturation mutagenesis of RLuc8/A123S/F261W and RLuc8/A123S/F262W at the D162/I163 positions. Coelenterazine was used for obtaining all spectra.
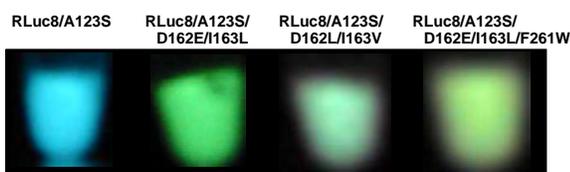
Figure 3.13: Color photographs for several of the variants screened in the D162/I163 saturation mutagenesis step. The images shown were made by mixing purified protein with coelenterazine, and photographing the resultant emission with a standard consumer-grade digital camera.

Round 4 - Saturation Mutagenesis at I223/P224

As saturation mutagenesis at I223 yielded several bathochromic shift mutations (Table 3.3), saturation mutagenesis was performed on the I223/P224 location, as well as on the 5 residues P220/R221/E222/I223/P224. The templates used in this screen were RLuc8 and RLuc8/A123S/-D162E/I163L. For each of these 4 conditions, ∼4000 colonies were screened for light output and/or color shifts. All the clones selected from this mutagenesis were the parental sequence, and were either parental template that had escaped the DpnI digestion, or templates containing silent mutations. This would seem to indicate that no further optimization can be done in this region of the protein.

Round 5 - Saturation Mutagenesis at V185/L186

With respect to increases in light output, the M185V mutation was the most interesting one that arose from the consensus sequence driven mutagenesis of Chapter 2. For this reason, saturation mutagenesis at V185/L186 and V185/L186/P187/S188/K189 was performed on the templates RLuc8 and RLuc8/A123S/D162E/I163L. The results of this screen are shown in Table 3.6, with corresponding emission spectra shown in Figure 3.14. Interestingly, the V185L mutation coupled to A123S/D162E/I163L could lead to a further ∼8 nm bathochromic shift in the emission spectrum, but when V185L was present alone in the RLuc8 background it led to no observable shift in the emission.

Round 6 - Saturation Mutagenesis at D154/E155

At this point, it was decided to take some of the more promising candidates from the previous rounds of mutagenesis, and try saturation mutagenesis at the E155 position. E155 had been identified in the initial random mutagenesis screen (Table 3.4) as a residue that could be mutated to increase the light output from the RLuc8/F261W and RLuc8/F262W constructs, and it was hoped that mutagenesis at this location would lead to improvements in RLuc8/A123S/-D162E/I163L/V185L, RLuc8/A123S/D162E/I163L/F261W, and RLuc8/A123S/D162E/I163V/F-262W. The neighboring D154 position was included in this saturation mutagenesis screen as well. The results of this screen, shown in Table 3.7 with corresponding spectra for several of the mutants in Figure 3.15, demonstrated that significant improvements of all three parental constructs could be achieved by mutagenesis at these two positions.

RLuc8 Mutants	Specific Activity	Wavelength (nm)			%>600 nm
		peak	mean	fwhm	
RLuc8	4.3±0.2	486	503	94	4
RLuc8/V185L	3.3	485	504	95	4
RLuc8/V185Q	4.5	482	500	93	3
RLuc8/V185K/L186M/P187A/S188A/K189L	0.85	480	497	102	4
RLuc8/A123S/D162E/I163L Mutants					
RLuc8/A123S/D162E/I163L	2.4	523	538	102	12
RLuc8/A123S/D162E/I163L/V185L	2.1	532	545	106	15
RLuc8/A123S/D162E/I163L/V185L/L186F	0.72	530	541	110	14
RLuc8/A123S/D162E/I163L/V185A/P187V/S188K/K189M	0.31	510	525	113	10

Table 3.6: Results from saturation mutagenesis at V185/L186, as well as random mutagenesis over V185/L186/P187/S188/K189. The templates used for mutagenesis were RLuc8 and RLuc8/A123S/D162E/I163L, and for the purposes of comparison the data for these variants is repeated from Tables 2.1 and 3.5, respectively. Approximately 20,000 colonies were screened for each of the V185/L186 saturation mutagenesis reactions, and 2500-6500 colonies screened for the V185/L186/P187/S188/K189 random mutagenesis reactions. Coelenterazine was used for measuring the spectra and specific activity. Specific activities are relative to that of RLuc (Table 2.1), and were not corrected for the luminometer's wavelength dependent sensitivity (see Table A.1). FWHM - full width at half maximum.
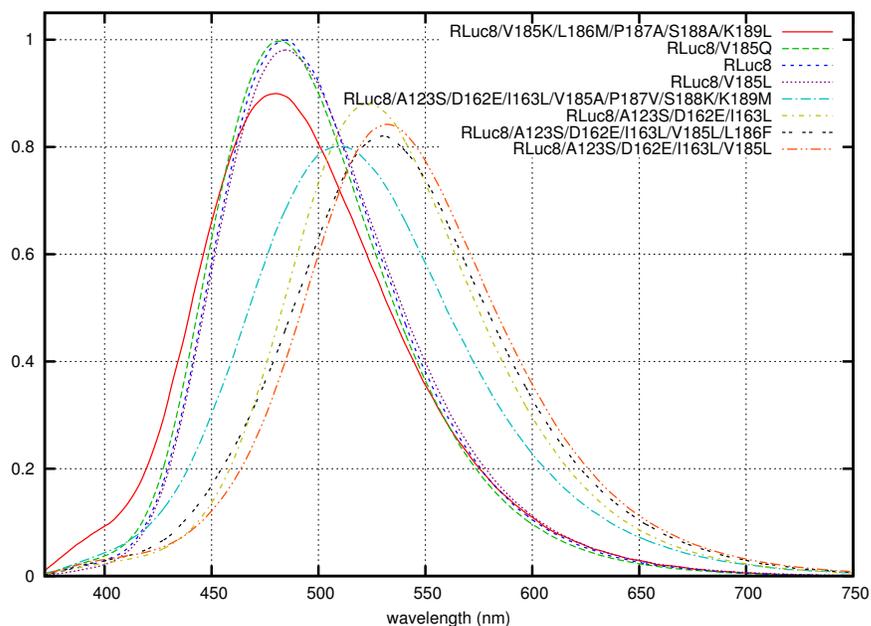
Figure 3.14: Normalized bioluminescence emission spectra for the variants obtained from saturation mutagenesis at the V185/L186 positions, as well as random mutagenesis between V185-K189. Spectra for RLuc8 and RLuc8/A123S/D162E/I163L are repeated here for the purposes of comparison. Coelenterazine was used for obtaining all spectra.
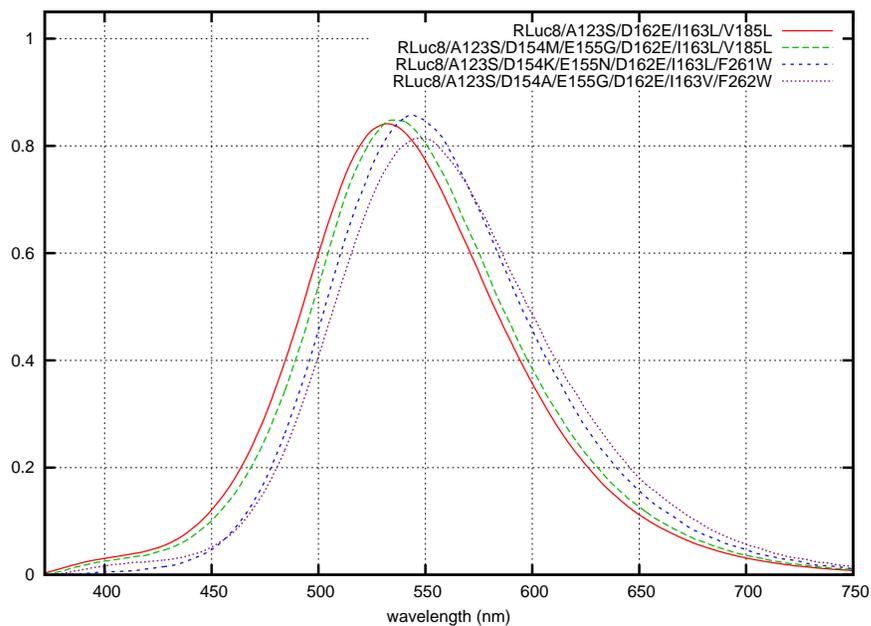
Figure 3.15: Normalized bioluminescence emission spectra for several variants obtained from saturation mutagenesis at the D154/E155 positions. Coelenterazine was used for obtaining all spectra.

RLuc8/A123S/D162E/I163L/V185L Mutants	Specific Activity	Wavelength (nm)			%>600 nm
		peak	mean	fwhm	
RLuc8/A123S/D162E/I163L/V185L	2.1	532	545	106	15
RLuc8/A123S/D154M/E155G/D162E/I163L/V185L	3.5	535	550	104	17
RLuc8/A123S/D154R/E155T/D162E/I163L/V185L	2.9	531	546	104	15
RLuc8/A123S/E155G/D162E/I163L/V185L	3.4	532	545	104	15
RLuc8/A123S/D162E/I163L/F261W Mutants					
RLuc8/A123S/D162E/I163L/F261W	0.12	538	553	107	17
RLuc8/A123S/D154K/E155N/D162E/I163L/F261W	0.97	545	560	106	21
RLuc8/A123S/D154R/E155G/D162E/I163L/F261W	0.39	537	554	106	18
RLuc8/A123S/E155G/D162E/I163L/F261W	0.66	537	554	107	17
RLuc8/A123S/E155K/D162E/I163L/F261W	0.46	541	556	107	18
RLuc8/A123S/D162E/I163V/F262W Mutants					
RLuc8/A123S/D162E/I163V/F262W	0.06	541	558	113	21
RLuc8/A123S/D154A/E155G/D162E/I163V/F262W	0.60	547	564	111	23
RLuc8/A123S/D154T/E155G/D162E/I163V/F262W	0.54	544	560	112	21
RLuc8/A123S/D154V/E155G/D162E/I163V/F262W	0.87	543	560	112	21
RLuc8/A123S/E155G/D162E/I163V/F262W	0.80	543	560	111	21

Table 3.7: Results from saturation mutagenesis at the D154/E155 residues. The templates used for mutagenesis were RLuc8/A123S/D162E/I163L/V185L, RLuc8/A123S/D162E/I163L/-F261W, and RLuc8/A123S/D162E/I163V/F262W. The data for these parental constructs are repeated from Tables 3.5 and 3.6 for the purpose of comparison. Approximately 3000-9000 colonies were screened from each template. Coelenterazine was used for measuring the spectra and specific activity. Specific activities are relative to that of RLuc (Table 2.1), and were not corrected for the luminometer's wavelength dependent sensitivity (see Table A.1). FWHM - full width at half maximum.

Round 7 - Saturation Mutagenesis at G269/A270

In an attempt to further improve the green emitting luciferase variants, the results of the previous screen at the D154/E155 positions were applied to an additional saturation mutagenesis screen at the G269/A270 positions. The G269 residue was identified in the initial mutagenesis screen (Table 3.4) as a position at which mutations could lead to significant improvements in the light output of RLuc8/F262W. With this in mind, saturation mutagenesis of G269 and the neighboring residue A270 was performed on the parental constructs RLuc8/A123S/E155G/-D162E/I163L/V185L, RLuc8/A123S/D154K/E155N/D162E/I163L/F261W, and RLuc8/A123S/-D154V/E155G/D162E/I163V/F262W. Between 3000-6000 colonies were screened for each condition, but no improved variants (with respect to light output or red-shift) were identified in the selection process.

Round 8 - Random Mutagenesis on RLuc8/A123S/D154V/E155G/D162E/I163V/F262W

In an attempt to identify locations that may yield further red-shifts in the bioluminescence emission spectrum, the RLuc8/A123S/D154V/E155G/D162E/I163V/F262W construct was subjected to random mutagenesis. In a small screen of ∼15,000 colonies, no further improvements in either light output or emission spectrum red-shifts were observed.

3.2.4 The Bioluminescence Emission Spectra of the Green-Emitting Variants with Bisdeoxycoelenterazine

In order to relate to previous studies the mechanism whereby these RLuc8 variants exhibited green-peaked bioluminescence emission spectra with the substrate coelenterazine, the bioluminescence spectra for several of these variants were assessed with the substrate bisdeoxycoelenterazine. The results, presented in Table 3.8, showed that the emission spectra with bisdeoxycoelenterazine remained largely unchanged for the different variants. This finding agrees with previous studies of the fluorescent characteristics of coelenteramide analogs [192], in that removal of the hydroxyl (replaced with a hydrogen in bisdeoxycoelenterazine) from the phenol group in resonance with the pyrazine ring led to an inability to form the green emitting pyrazine anion.

	Wavelength (nm)			
	peak	mean	fwhm	%>600 nm
RLuc	401	426	ND	1
RLuc8	401	419	66	1
RLuc8/A123S/D162E/I163L	397	429	56	2
RLuc8/A123S/D162E/I163L/V185L	396	427	58	2
RLuc8/A123S/E155G/D162E/I163L/V185L	394	418	56	2

Table 3.8: Emission spectra with bisdeoxycoelenterazine for RLuc and several variants. FWHM - full width at half maximum. ND - not determined.

3.3 Discussion

It was unclear before starting this work whether a protein could present the necessary chemical environment to favor the green emitting anion form of coelenteramide (presumably the pyrazine anion, Figure 3.2d). The main evidence against being able to achieve this anionic form in the context of a protein is that no known coelenterazine using luciferases are able to emit a green peaked emission spectrum. There are many cases where green bioluminescence is observed from organisms with coelenterazine using luciferases, but the green light emission has always been found to be due to bioluminescence resonance energy transfer (BRET) from a blue emitting luciferase to a green fluorescent protein (GFP) (further details on GFP and BRET are discussed in Appendix B). If a green emission spectrum from a coelenterazine using luciferase was possible, why would the more complex solution of having an accessory fluorescent protein have evolved?

The data in this chapter shows that a green emitting coelenterazine using luciferase is indeed possible, presumably due to favoring the pyrazine anion of coelenteramide in the enzymatic pocket. Further more, this shift in spectrum can come with little loss in the ability of the luciferase to emit light. So why then have many marine organisms evolved to exhibit a luciferase/GFP BRET pair? One hypothesis arises from a comparison of the GFP and green emitting luciferase spectra shown in Figure 3.16. This comparison highlights the extraordinarily broad emission spectra of the luciferases, and the relatively narrow peak emission of the GFP. Presumably, there is an evolutionarily advantageous reason for *Renilla reniformis* to emit green light, and the use of a GFP allows a greater number of these photons to be emitted at these desired wavelengths.

Alternatively, the GFP may simply have evolved first. With the GFP already in existence, it may have been easier from an evolutionary standpoint to co-opt the pre-existing GFP rather
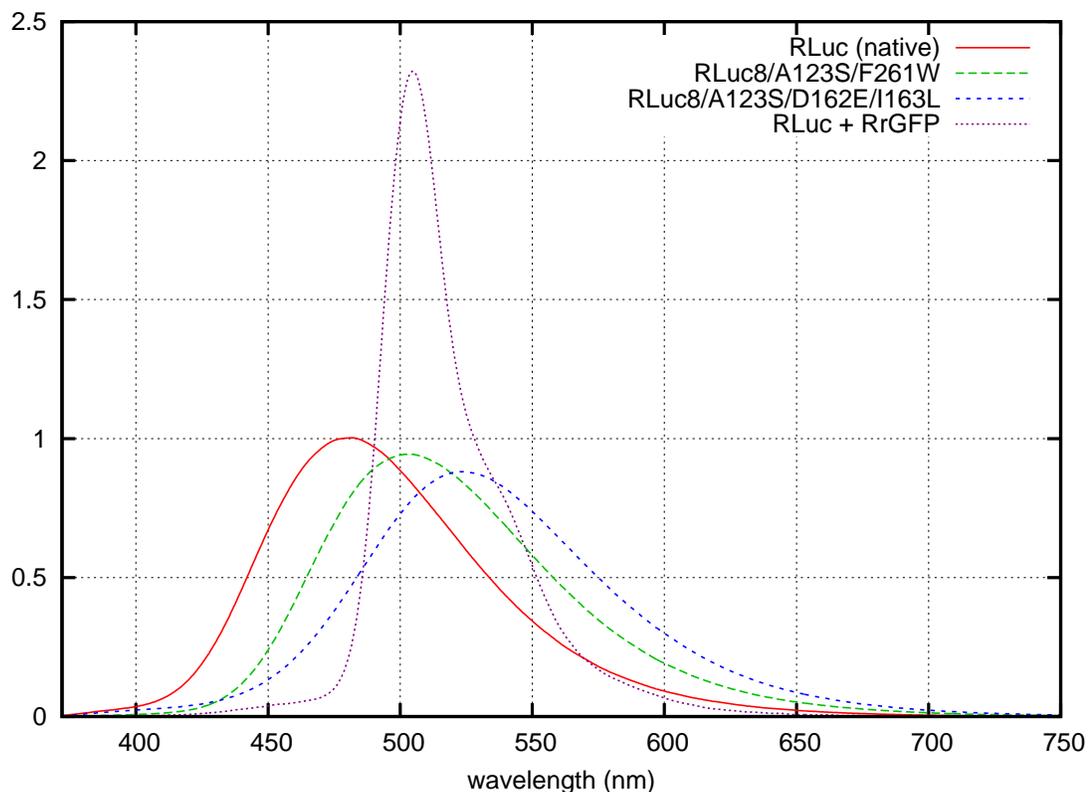
Figure 3.16: A comparison of the *Renilla reniformis* Green Fluorescent Protein (RrGFP) emission spectrum and the emission spectra of RLuc and several of its variants. The RrGFP emission spectrum was made by combining RrGFP dimer with RLuc at a 1:1 molar ratio, adding coelenterazine, and measuring the emission in the same manner as for the RLuc variants. RrGFP is characterized further in Appendix B.

than to go through a sequence of mutations to achieve a green emitting luciferase.

An interesting question to ask, is how much of an effect on *in vivo* imaging capabilities could one expect from the ∼25-65 nm emission shifts that the luciferase variants present here. The expected benefit of these bathochromic (red) shift variants is complicated by the fact that these shifts are moving the peak emission squarely into a local maximum around 550 nm in the hemoglobin absorption curve.

To answer these questions, rough calculations of light attenuation were made using rat liver absorption values (Figure 3.17). Rat liver was chosen as a model organ to study merely because the absorption values are available in the range of wavelengths that are of interest here. The liver may not in fact be the most appropriate model organ for this work due to its relatively high hemoglobin content. Note that in this absorption data, the effects of the 550 nm
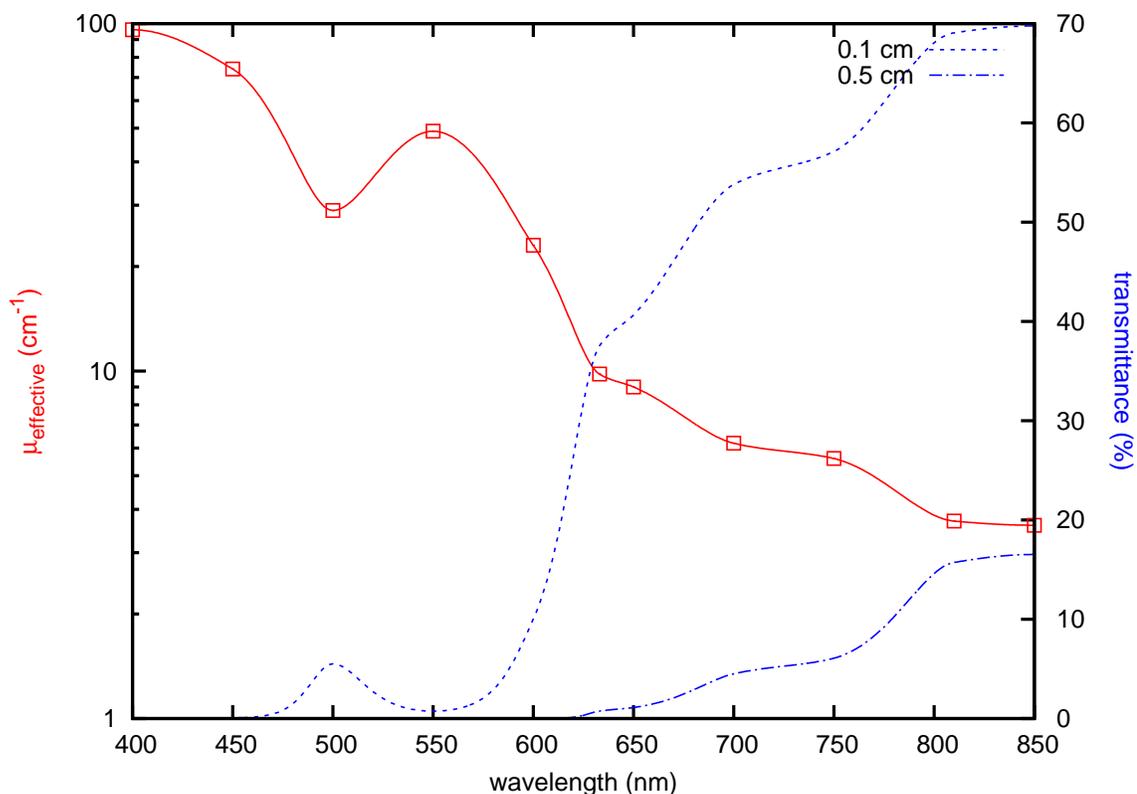
Figure 3.17: The effective absorbance ($\mu_{effective}$) of rat liver for optical wavelengths. Due to the sparseness of the source data [37, 38], the interleaving points were interpolated using a piece-wise cubic hermite polynomial. The resulting transmittance values calculated for depths of 0.1 cm and 0.5 cm are shown.

hemoglobin absorption peak can be clearly seen.

Multiplying the rat liver transmittance values by the normalized emission spectra for several of the luciferase variants gives the emission spectra shown in Figure 3.18. At 0.1 cm of tissue depth, the main spectral peak of *Renilla* luciferase is severely diminished, and at 0.5 cm depth one can observe that it is really only the photons with >600 nm wavelength that are able to escape from the tissue. This underscores the importance of these few long wavelength photons in allowing the detection of this luciferase in small animal imaging.

A more quantitative comparison is given in Table 3.9, where predictions are made as to the relative gain in light output versus RLuc for the various luciferase variants at 0.1 and 0.5 cm depth of liver tissue. Again, these results underscore the advantageousness of having a red-shifted *Renilla* luciferase for small animal imaging applications. They also point out that the benefits of red-shifting the emission spectrum outweigh any penalties from the local
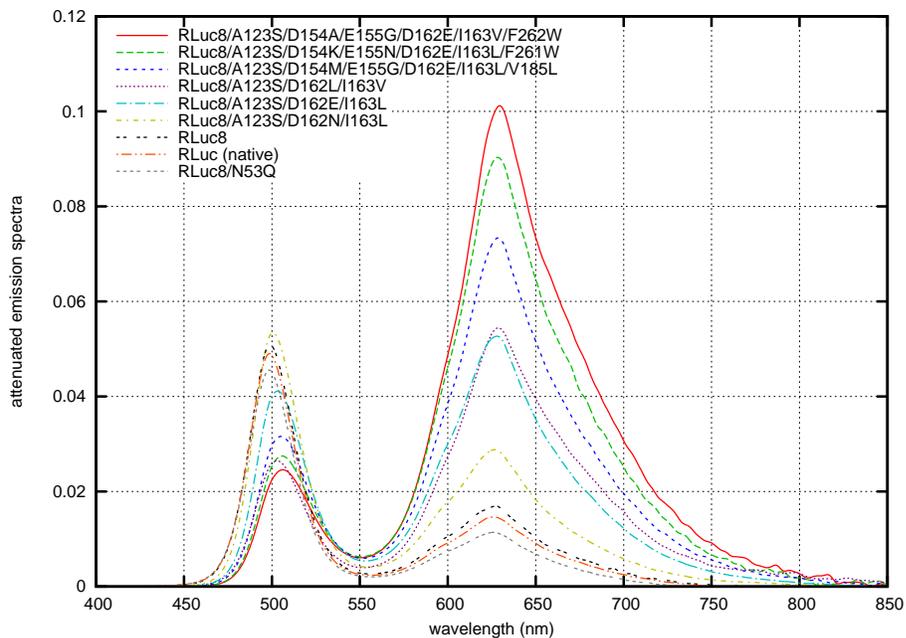
(a) 0.1 cm depth
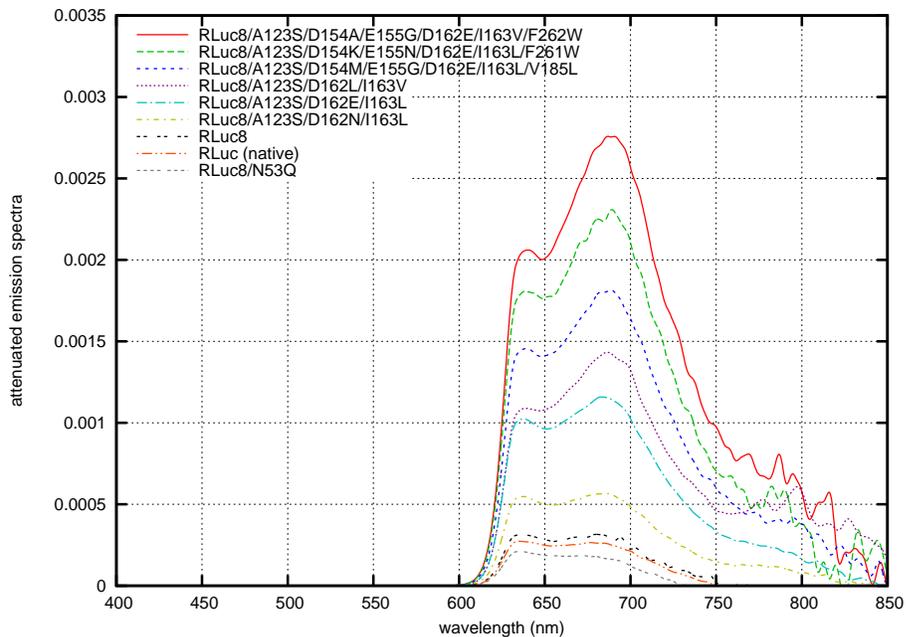
(b) 0.5 cm depth

Figure 3.18: Estimated effects to the normalized emission spectra of various mutants after passing through either 0.1 cm or 0.5 cm of rat liver tissue. The curves were made by multiplying the normalized emission spectra by the calculated absorbance values for the different thicknesses of rat liver shown in Figure 3.17.

	Specific Activity	Corrected Activity	Wavelength mean (nm)	% Transmitted		Effective Output	
				0.1 cm	0.5 cm	0.1 cm	0.5 cm
RLuc	1.0	1.0	497	2.8	0.025	1.0	1.0
RLuc8	4.3	4.3	503	3.1	0.029	4.7	5.0
RLuc8/A123S/D162L/I163V	0.29	0.35	515	5.7	0.17	0.71	2.4
RLuc8/A123S/D162N/I163L	2.4	3.1	523	4.2	0.065	4.7	8.1
RLuc8/A123S/D162E/I163L	2.4	3.6	538	5.8	0.13	7.4	19
RLuc8/A123S/D162E/I163L/V185L	2.1	3.4	545	6.9	0.18	8.4	25
RLuc8/A123S/D154M/E155G/D162E/I163L/V185L	3.5	6.0	550	7.4	0.20	16	48
RLuc8/A123S/D154K/E155N/D162E/I163L/F261W	0.97	1.9	560	8.7	0.26	5.9	20
RLuc8/A123S/D154A/E155G/D162E/I163V/F262W	0.60	1.2	564	9.7	0.31	4.2	15

Table 3.9: Effects of tissue depth on the relative light output of several *Renilla* luciferase variants. Note that the spectral sensitivity of the luminometer's detector penalizes the red-shifted variants; the "Corrected Activity" takes this into account using the rough compensation factors from Table A.1. "% Transmitted" is the percent of photons that are transmitted through the given depth of rat liver tissue, as calculated based on the spectra data and the transmittance data shown in Figure 3.17. "Effective Output" is the corrected specific activity of the enzyme multiplied by the percent of photons transmitted for the given depth. The effective output values as well as the specific activity values have been normalized to those of RLuc.

hemoglobin absorption peak at 550 nm.

When bioluminescence is utilized in small animal imaging, the beetle luciferases have been more commonly employed than *Renilla* luciferase. The primary reason for this has been that the green to red peaked emission spectra of the beetle luciferases is transmitted through biological tissue more readily than the blue peaked emission spectrum of *Renilla* luciferase. A comparison of the emission spectra from the Click Beetle luciferases and the RLuc variants developed here, shown in Figure 3.19, demonstrates that the red-shifted RLuc variants have entered the range of spectra previously only available with the beetle luciferases. Note that the emission spectrum of Firefly luciferase is variable, depending upon both temperature and its chemical environment [221], and would lie somewhere intermediate to the spectra for Click Beetle Green luciferase and Click Beetle Red luciferase.

An estimate for the transmittance of these spectra through rat liver can be made as has been done previously for the RLuc variants, and is shown in Figure 3.20. This estimate demonstrates two points. First, for an equal number of generated photons several of the RLuc variants should perform as well as Click Beetle Green luciferase. Second, after filtering by biological tissues, the emitted bioluminescence spectra are all essentially identical in form.

All mutagenesis experiments performed in this chapter utilized protein that was expressed periplasmically. Utilizing periplasmic expression was a great aid in this work, as protein recovered from the osmotic shock protocol was of sufficient purity to allow meaningful screening to be done at this stage. Additionally, close to 100% purity can be achieved for the periplasmi-
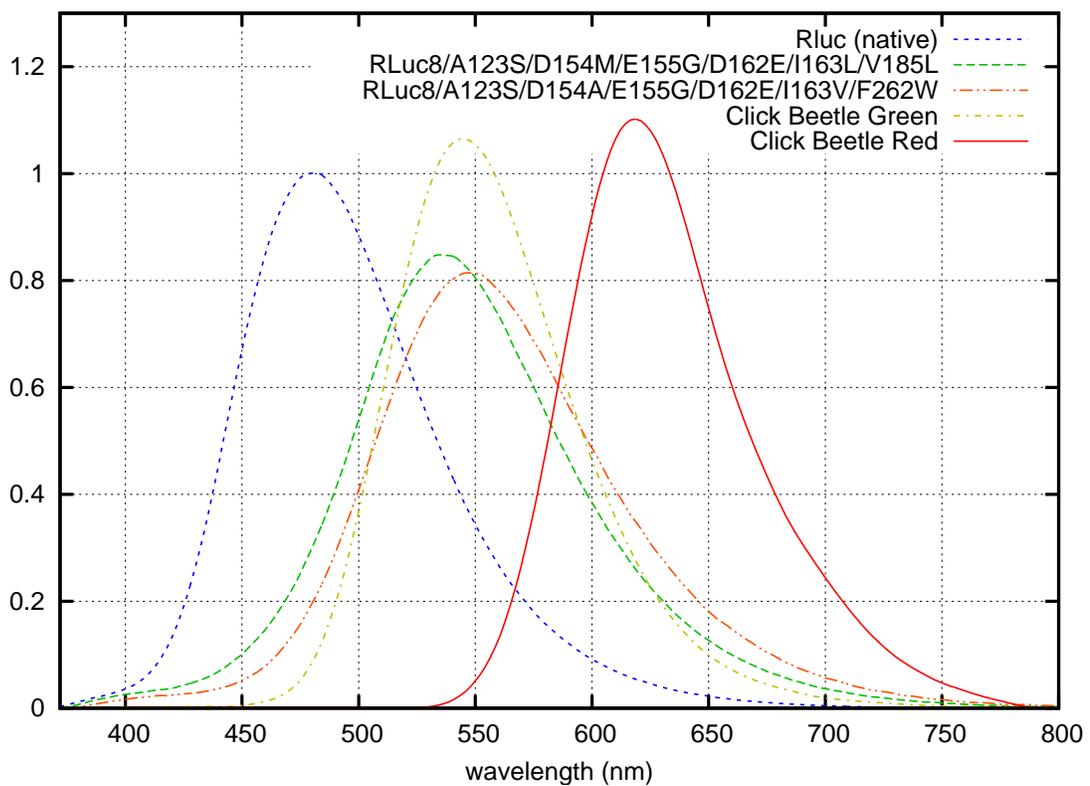
Figure 3.19: Normalized bioluminescence emission spectra for the Click Beetle luciferases and several RLuc variants. The spectra for the Click Beetle luciferases are from Zhao *et al.* [246]. Coelenterazine was used for obtaining the spectra from the RLuc variants. Normalization was to the total area under the curve, and not the emission peak.
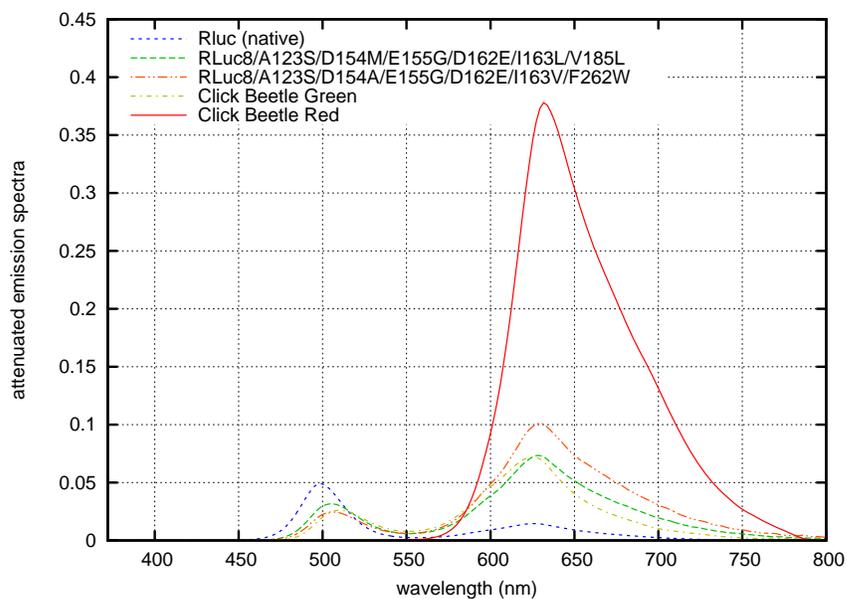
(a) 0.1 cm depth
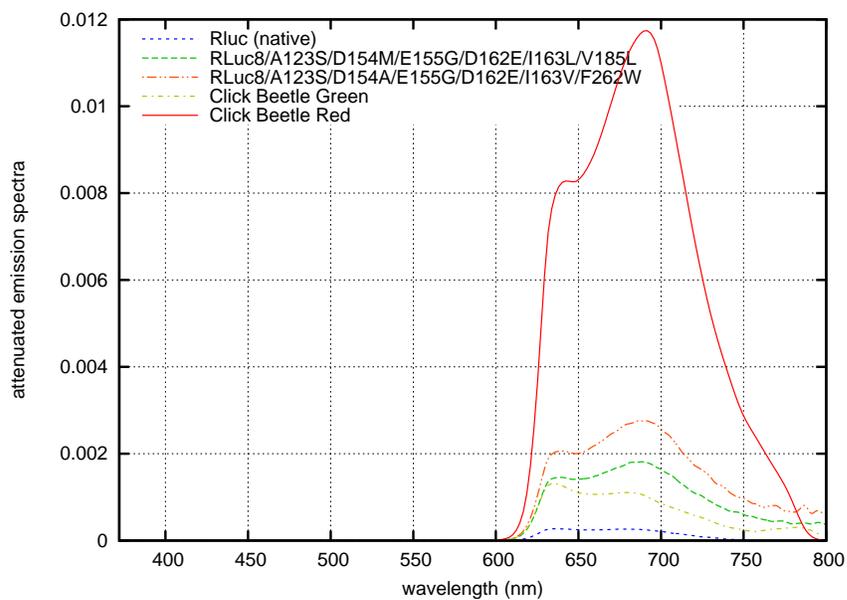
(b) 0.5 cm depth

Figure 3.20: Estimated effects to the normalized emission spectra from the Click Beetle luciferases or various RLuc mutants after passing through either 0.1 cm or 0.5 cm of rat liver tissue. The curves were made by multiplying the normalized emission spectra by the calculated absorbance values for the different thicknesses of rat liver shown in Figure 3.17. This comparison assumes that all the luciferases emit an equal number of photons. The Click Beetle luciferase emission spectra are from Figure 3.19.

cally expressed protein after a single nickel affinity chromatography step (Figure 2.2). Efficient periplasmic expression, however, requires the protein to be loosely folded before transloca- tion [12], which is accomplished through molecular chaperones that maintain the protein in an unfolded state [47]. This raises the question of whether the use of periplasmic expression biases the mutations seen, especially when only 1 or 2 mutations can substantially decrease the efficiency of periplasmic expression (Examples in Sections 2.2.5 and 4.2.4).

The IVIS 200 system used for screening the mutant containing bacterial plates proved surprisingly unsuited for the task at hand. This system is optimized for imaging the low light levels encountered in animal imaging of bioluminescent reporter genes. As one of these optimizations to reduce noise, the read-out speed of the system's CCD is extraordinarily slow. For a 1 s image acquisition, the read out can take an additional 15-30 s depending on the image resolution being acquired. This long delay period led to serious issues with substrate depletion when multi-spectral imaging was attempted.

Two other annoyances were encountered when using this system. The first issue, was that images acquired using different filters were not aligned and generally required shifting of a pixel or two in each direction to get proper overlap. The other, more serious annoyance, was that the photographic images contained specular artifacts due to the 4 point light sources used for illumination. These specular artifacts limited the effectiveness of the automated image processing for colony detection and scoring.

An ideal technique for luciferase mutagenesis studies would allow mutating and screening cells at the single cell level, as has been done for engineering GFP [227]. However, the standard single cell sorting method used in modern biology, fluorescent activated cell sorting (FACS), is inappropriate for selecting on bioluminescence. As each cell is only in front of the machine optics for $\sim$10 μs, there is inadequate time to collect sufficient photons from bioluminescence to do meaningful cell sorting. An alternative idea would be to add the reaction product and attempt to select on the fluorescent characteristics of the luciferin/luciferase complex. While this may potentially work for other luciferases, for the particular case of *Renilla* luciferase this is not an option. As mentioned previously, the chemical environment within the luciferase during photon emission is not maintained for any appreciable time after coelenteramide has reached its ground state [138]. For sorting directly on bioluminescence, future work could potentially make use of stream imaging cell sorters that are currently being developed (Amnis Corp., Seattle, WA). Alternatively, if a single cell sorting microfluidic device [63] were available, it could be coupled to a sensitive spectrophotometer (e. g. Triax 320: Horiba Jobin Yvon, Edison

NJ) to allow this selection to be accomplished.

The work in this chapter has generated *Renilla* luciferase variants with green peaked emission spectra. The obvious question is, can the emission spectrum be further red-shifted? As the amount of interaction between the luciferase and the luciferin is not well understood a conclusive answer cannot be stated at this time, but it may be that the pyrazine anion of coelenteramide represents a limit as to the bathochromic shift that this luciferin/luciferase system can accomplish. Further bathochromic shifts of the emission spectrum may not be achievable from mutagenesis of the luciferase, and it may be more fruitful at this point to consider altering the structure of the luciferin as discussed in Section 7.5.

Chapter 4

The Crystallographic Structure of *Renilla* Luciferase

The work in this chapter is the result of a collaboration with Dr. Timothy D. Fenn and the laboratory of Dr. Axel Brunger. Dr. Fenn assisted with the crystallization screens, acquired the X-ray diffraction data, and calculated the structures. Data interpretation was done jointly.

Renilla luciferase (RLuc) is very similar to a number of haloalkane dehalogenases (Section 2.2.1). As the structures of several of these haloalkane dehalogenases have already been determined [201, 157], it was not expected that obtaining a crystal structure of RLuc by itself would reveal anything overtly interesting about its tertiary fold.

The real hope of the work presented in this chapter, was to obtain a crystal of RLuc in complex with either its substrate or product. The structure of such a complex would be interesting for two reasons. First, RLuc is an oxidase and not a hydrolase like the haloalkane dehalogenases. It would be interesting to explore the basis for this difference in catalytic effect, especially in the face of the high level of primary structure similarity and the sharing of active site residues with the haloalkane dehalogenases. Second, an appreciation of the manner in which the substrate (or product) binds in the active pocket would lend insight into how the protein could be rationally mutated so as to alter properties dependent on this complex formation. Such properties include the protein's bioluminescence emission spectrum and its substrate specificity.

Rather than focusing on the native luciferase, the work here has instead focused on crystallizing the 8 mutation stabilized variant of RLuc (RLuc8) generated in Chapter 2. This choice

was made for the pragmatic reason that periplasmic expression of RLuc8 yields at least an order of magnitude more protein than periplasmic expression of the native luciferase.

4.1 Methods

4.1.1 Periplasmic Constructs

The pBAD-pelB-RLuc8 construct used here has already been described in Section 2.1.4. This construct has an N-terminal PelB signal peptide, and a C-terminal non-cleavable 6xHis tag.

A version of RLuc8 for periplasmic expression with an N-terminal, cleavable 6xHis tag was generated as follows. PCR was used to replace the second and third codons (A2, S3) with an NcoI site, a 6xHis tag, and a thrombin site (LVPR/GS). In the same PCR step, the C-terminal SalI site and 6xHis tag were replaced with a stop codon followed by a PmeI site. After NcoI/PmeI digestion, this product was inserted into a correspondingly digested pBAD-pelB backbone to make pBAD-pelB-6xH-thr-S3RLuc8, where thr indicates the thrombin protease site. Note that after thrombin digestion of the expressed protein, the residues GS of the thrombin site will remain attached to the protein, so the luciferase sequence will begin at S3. For this reason, the expressed protein from this construct is labeled as S3RLuc8.

4.1.2 Cytoplasmic Constructs

The pBAD-RLuc8 construct described in Section 2.1.4 was used for cytoplasmically expressing RLuc8. It was also the basis for constructing the various surface mutation constructs utilized in Section 4.2.4 that were generated by site-directed mutagenesis.

The truncation constructs used in Section 4.2.6 were made by first creating a pBAD plasmid with an N-terminal 6xHis tag and a thrombin site. The genetic construct NcoI-6xHis-thr-HindIII-PmeI was synthetically made using appropriate primers and PCR. This construct was then NcoI/PmeI digested and inserted into a similarly digested pBAD backbone to make pBAD-6xHis-thr. This construct contains a SexAI/DraIII site at the start of the thrombin site, a BamHI site in the final two codons of the thrombin site, as well as a HindIII site 3' to the thrombin site. Note that use of the SexAI site requires the DNA to be unmethylated, while the BamHI site is not unique in the plasmid. To make the protein expression constructs, appropriate primers were synthesized, PCR was performed using pBAD-RLuc8 as the template, and the products were digested and ligated into the pBAD-6xHis-thr backbone. An example

of the nomenclature for the resulting constructs is the plasmid pBAD-6xHis-thr-I15RLuc8Δ5 that was used to express the protein I15RLuc8Δ5. The final purified protein from this construct consists of the residues GS (left from the thrombin cut site), followed by the RLuc8 sequence starting at I15 and missing its last 5 residues. The only exception to this naming scheme, are the "S3" plasmids (e. g. pBAD-6xHis-thr-S3RLuc8, pBAD-6xHis-thr-S3RLuc), were the S3 is both part of the thrombin site and the first residue of the luciferase.

4.1.3   Expression and Purification

Periplasmic expression and initial nickel affinity purification for the pBAD-pelB-RLuc8 construct was identical to what has already been described in section 2.1.5. For the N-terminal cleavable 6xHis tag construct, expression and nickel affinity purification were similar, with the alterations that incubation of the culture following induction was done at 30°C for 6 h, and thrombin digestion was done immediately following nickel affinity purification by incubating with 1 μg calf α-thrombin per mg protein overnight at 4°C.

Expression of cytoplasmic constructs was performed as follows. Plasmid containing *E. coli* LMG194 cells were grown to an OD_{600} of 0.7 in Terrific Broth, induced with 0.2% arabinose, grown for 12 h at 32°C, pelleted, and frozen. The cells were then lysed by thawing in nickel affinity wash buffer (WB: 300 mM NaCl, 20 mM HEPES, 20 mM imidazole, pH 8) containing 1 mg/ml lysozyme, 10 μg/ml RNase A, and 5 μg/ml DNase I, slowly mixing for 1 hr, and sonicating. The volume of WB used was 5% of the initial culture volume. The lysates were clarified by centrifugation at 10K RCF 4°C 30 min, and allowed to bind to nickel affinity resin (Ni-NTA Superflow, Qiagen) for 1 hr at 4°C with gentle mixing. The amount of resin used for this step was based on the yield of protein expected, with a binding capacity of 40 mg protein/ml resin assumed. Following binding, the resin was washed extensively with WB. For non-cleavable constructs, the protein was eluted at this point with nickel affinity elution buffer (EB: 300 mM NaCl, 20 mM HEPES, 250 mM imidazole, pH 8). For cleavable constructs, 1 μg calf α-thrombin per mg protein was adding to a slurry of 50% resin in WB, the mixture was incubated overnight at 4°C, and the protein was then eluted with WB.

For both periplasmic and cytoplasmic constructs, purification continued with a combination of anion exchange and gel filtration chromatography. All purification steps were done at 4°C. For anion exchange chromatography, the protein (at this point in either EB or WB) was diluted with anion exchange start buffer (10 mM NaCl, 10 mM Tris pH 8) to a NaCl concentration of <60 mM. The diluted protein was then bound to an anion exchange column

(Source 15Q, GE Healthcare, Piscataway, NJ) and eluted with a gradient of NaCl. Elution occurred at ~100 mM NaCl. The final stage of purification was gel filtration chromatography, performed with a 320 ml volume Sephacryl S-100 column (GE Healthcare) and a running buffer of 10 mM NaCl, 10 mM HEPES, pH 7.4.

This purified protein (in 10 mM NaCl, 10 mM HEPES, pH 7.4) was concentrated as necessary using 10 kDa cut-off, Amicon Ultra centrifugal concentrators (Millipore, Billerica, MA).

4.1.4 Characterization

Protein size and monodispersity were confirmed using a Superdex 200 analytical grade gel-filtration column (GE Healthcare) followed by in-line multi-angle light scattering (MALS) and refractive index detectors (DAWN EOS and Optilab DSP, Wyatt Technologies, Santa Barbara, CA). A dn/dc value of 0.185 ml/g was assumed in all calculations, and all processing was performed using the ASTRA software package (Wyatt Technologies).

Appropriate molecular weights were confirmed using a Matrix Assisted Laser Desorption/-Ionization Time of Flight (MALDI-TOF) mass spectrometry (Bruker Daltonics, Billerica MA). All samples were spotted using a sinapinic acid matrix (Ciphergen Biosystems, Fremont CA) and analyzed in positive ion mode.

4.1.5 Crystallization

Crystallography trials were done in either hanging-drop or sitting-drop formats. For hanging-drop screening, drop sizes were generally 3 μl and consisted of 50% mother liquor and 50% of the protein solution. Sitting-drop setups sere utilized for 96-well plate high-throughput screening, with 0.5 μl drops consisting of 50% mother liquor and 50% protein. Unless otherwise listed, protein concentrations prior to drop set up can be assumed to be between 20-30 mg/ml. Please note that protein concentrations quoted in the text refer to the concentration before set up, so the initial concentration in the drop is half the given value. The same holds true for the mother liquor constituents. Crystallization trays were kept at 20°C, unless specifically mentioned otherwise, and were generally observed for at least a year.

4.1.6 X-Ray Diffraction and Structural Determination

A homology model of RLuc8 was created using SWISS-MODEL [185] and the crystal structures of of the haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* (PDB files

1iz8, 1k63, 1k6e, 1iz7, and 1mj5) [201, 161]. The loop regions of this homology model were then removed, and the resulting homology model was used to bootstrap the phasing process via molecular replacement. Matthews coefficient calculations [136] suggested the presence of a monomer in the $P6_1$ asymmetric unit that was located during molecular replacement with Phaser [141].

Following initial phasing, simulated annealing refinement against a maximum likelihood target function was carried out as implemented in the Crystallography and NMR System [27]. The loop regions were then rebuilt using ARP/wARP [172, 171], followed by additional simulated annealing refinement. Final rounds of refinement included the addition of water molecules and discrete side-chain disorder using Coot [55], followed by conjugate gradient refinement using a maximum likelihood target function as implemented in Refmac [152].

4.2 Crystallography Condition Screening

What follows is a synopsis of the different constructs and conditions that were screened while trying to crystallize the luciferase. This information is included here for those who would like to replicate the crystallization conditions, or who would be interested in screening for new crystallization conditions and would like to know what has already been attempted. The readers who are interested in the results of the crystallization rather than the path are encouraged to skip to Section 4.3.

4.2.1 RLuc8 - Periplasmic

Purified RLuc8 that had been expressed periplasmically was characterized by light scattering and MALDI-TOF mass spectrometry. Light scattering results indicated that RLuc8 existed as a monomer in solution, as molar mass moment calculations based on the multi-angle scattering indicated a molecular weight of 33.8 kDa (error: 7%) with a relatively low polydispersity across the gel filtration elution profile (∼11%).

By mass spectrometry, RLuc8 was measured as 36.8 kDa - within error of the expected size of 36.9 kDa. A minor peak around 38.5 kDa was noticed on some preparations, and may indicate that the PelB signal peptide isn't consistently processed and removed. This potential issue was addressed later with the N-terminal cleavable 6xHis tag construct discussed in Section 4.2.2.

Initial screening trials, utilizing the Hampton Screen and Hampton Screen II (Hampton Research, Aliso Viejo, CA), were done at 4°C and 20°C with an RLuc8 concentration of 20 mg/ml. These screens failed to reveal a crystallization condition. Additional rounds of screening were performed using the Wizard Screens I and II (Emerald BioSystems, Bainbridge Island, WA) and the Clear Strategy Screen 2 (Molecular Dimension, Apopka, FL) at 20°C with a protein concentration of 20 mg/ml. Two crystallization conditions were identified and are discussed below.

Diammonium Phosphate

The Wizard Screen II showed a potential hit in the condition 0.2 M NaCl, 1.0 M diammonium phosphate [$(NH_4)_2HPO_4$], 0.1 M imidazole pH 8.0. In hindsight, this condition was actually a salt crystal, but fortunately a grid screen around the "hit" condition was set up. After 8 months, a cluster of crystals was observed in the condition 0.3 M NaCl, 1.25 M diammonium phosphate, 0.1 M imidazole pH 8.0, 20°C (Figure 4.1a). The cluster easily separated into single shards with mechanical stimulation. After passing through a cryoprotectant consisting of a saturated solution of malonate [83], the single shards were frozen and submitted to X-ray diffraction. The resulting structure is described in Section 4.3.

A variety of screens were set up around this condition in an effort to replicate the crystallization and speed crystallization. This included screening over protein concentrations (450-10 mg/ml), macroseeding (protein at 20 and 60 mg/ml), microseeding (protein at 20 and 120 mg/ml), grid screens around the condition (protein at 20 and 100 mg/ml), addition of coelenterazine and benzyl-coelenterazine to the drops (protein at 20 and 100 mg/ml), the inhibitors of *Renilla* luciferase dibenzylamine and 4-benzyloxyaniline [139] (protein at 20 and 100 mg/ml), and the Hampton Additive Screens 1-3 (protein at 20 and 100 mg/ml). These screens were entirely unsuccessful in speeding up the crystallization process, and for the most part unsuccessful in replicating the crystallization condition. The grid screens did develop a couple additional crystals (e. g. Figure 4.1b). Perhaps most interesting is the condition that was set up with an effective protein concentration of 450 mg/ml (Figure 4.1d). This condition developed a large crystal ($\sim$100 μm) 8 months after set up that diffracted to 1.6 $\mathring{A}$.

Potassium Thiocyanate

After $\sim$8 months, the Clear Strategy Screen 2 grew crystals in the condition 0.15 M potassium thiocyanate (KSCN), 15% w/v polyethylene glycol (PEG) 6000, 0.1 M Tris-HCl pH 7.4. A pho-
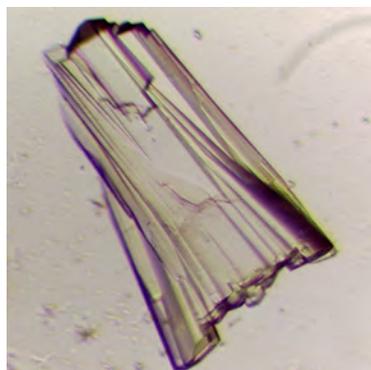
(a) 8 months, 0.3 M NaCl, 1.25 M di-ammonium phosphate, 0.1 M imidazole pH 8.0

(b) 6 months, 0.3 M NaCl, 1.45 M di-ammonium phosphate, 0.1 M imidazole pH 8.0

(c) 8 months, 0.3 M NaCl, 1.25 M di-ammonium phosphate, 0.1 M imidazole pH 8.0
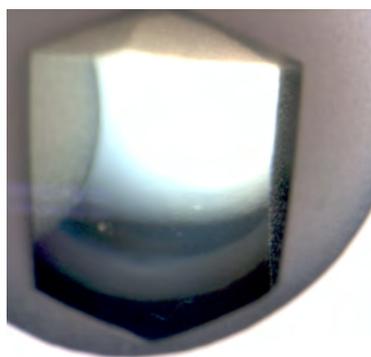
(d) 8 months, 0.3 M NaCl, 1.25 M di-ammonium phosphate, 0.1 M imidazole pH 8.0
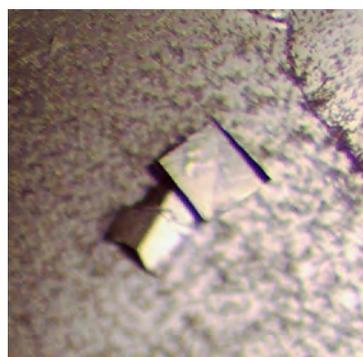
(e) 8 months, 0.15 KSCN, 15% w/v PEG 6000, 0.1 M Tris-HCl pH 7.4

Figure 4.1: Photographs of various crystals formed from RLuc8. All conditions were incubated at 20°C. The length of time between set up of the tray and when the image was taken, along with the mother liquor condition, are shown below each photograph. Panels (a) and (e) used RLuc8 at 20 mg/ml. Panel (b) used RLuc8 at 100 mg/ml. Panel (c) used RLuc8 at 300 mg/ml. Panel (d) was set up such that the effective protein concentration used was 450 mg/ml (225 mg/ml in final drop).

tograph of the crystal is shown in Figure 4.1e, and the corresponding structure is presented in section 4.4. Grid screens around this condition as well as microseeding experiments, with protein concentrations of 20 and 100 mg/ml, were unable to reproduce this condition.

4.2.2   S3RLuc8 - Periplasmic

As mentioned above, the RLuc8 preparations occasionally showed a small level of heterogeneity on mass spectrometry, presumably because of variable cleavage of the PelB signal peptide. The heterogeneity was suspected as a possible reason why the screens utilizing periplasmic RLuc8 did not consistently yield crystals. For this reason, the pBAD-pelB-6xH-thr-S3RLuc8 construct was created, featuring an N-terminal cleavable 6xHis tag. As the PelB leader sequence precedes the N-terminal tag, any heterogeneity in the signal peptide processing is irrelevant as the remaining peptide will be removed during the thrombin digestion.

Periplasmic expression of this construct exhibited a large pool of activity in the culture medium. Although periplasmic proteins can be released into the culture medium by deliberate permeabilization of the outer membrane [10, 131, 12], the effect seen here is more likely due to some amount of toxicity from blocking of the periplasmic export machinery [197]. Purification from a rich culture medium was not an option due to chelators present in such media that would strip the nickel from the affinity resin. Instead, the amount of protein lost to the culture medium was reduced by optimizing the culture conditions to an incubation of 30°C for 6 h following induction. Even under these optimized conditions, roughly half the measurable activity was present in the culture medium.

Following thrombin digestion and purification, mass spectrometry of the S3RLuc8 protein showed the desired single peak. Screening utilizing the Hampton Screen and Hampton Screen II set up at both 20°C and 4°C did not yield a crystallization condition. Fine needle clusters were seen in the 4°C plate with the condition 0.1 M HEPES pH 7.5, v/v 10% isopropanol, w/v 20% PEG 4000, but further screening was unable to reproduce this condition. Note that this condition reappears in Section 4.2.7. Another potential hit was seen in the 4°C plate with the condition 0.2 M Lithium Sulfate, 0.1 M Tris-HCl, pH 8.5, 30% w/v PEG 4000, but this too was irreproducible.

PEG 3350

A screen using the TOPAZ microfluidics chip (Fluidigm, South San Francisco, CA) was performed. The TOPAZ system utilizes free interface diffusion, and enables sampling of a large

portion of chemical space in a single experiment. A promising condition was seen in the condition 0.5 M sodium acetate, 25% w/v PEG 3350, and a grid screen using 20 mg/ml protein was done at 20°C and 4°C. After two weeks, small clusters of needle crystals were seen at 0.3-0.4 M sodium acetate with 20% w/v PEG 3350 at 4°C. Extensive further screening using Hampton Additive Screens 1-3, Hampton Detergent Screens 1-3, a pH screen, and microseeding yielded only a single crystal. This crystal grew after 5 months in a detergent condition containing 1.8 mM n-Decyl-β-D-thiomaltoside (Figure 4.2a). The corresponding structural information is presented in Section 4.5.

Diammonium Phosphate

A grid screen using S3RLuc8 at 30 mg/ml was performed around the previously identified diammonium phosphate condition (0.3 M NaCl, 1.25 M diammonium phosphate, 0.1 M imidazole pH 8.0, 20°C). Crystals formed under various conditions between 2 and 8 months following set up. Photographs of these crystals are shown in Figures 4.2b- (h). Various concentrations of the inhibitors dibenzylamine and 4-benzyloxyaniline [139] were tried as additives to the diammonium phosphate condition with no success. Microseeding was attempted as well, but this failed to speed up the crystallization process.

4.2.3   Additional Periplasmically Expressed Proteins

Two additional periplasmically expressed proteins, RLuc/A54P/A55T/C124A (Table 2.1) and RLuc8/N109D (Table 3.1), were tried for crystallography. The protein RLuc/A54P/A55T/C124A was screened using the Hampton Screen and Hampton Screen II at 20°C with no success. The protein RLuc8/N109D was utilized in a grid screen around the 0.5 M sodium acetate, 25% w/v PEG 3350 condition at 20°C and protein concentrations of 60 mg/ml and 300 mg/ml, but no crystals were formed.

RLuc8/N109D was also used in a high throughput Ozma PEG Ion Screen (Emerald Biosystems). After 4 months, small cylindrical crystals grew in 200 mM magnesium sulfate, 20% w/v PEG 8000. After 8 months, long hexagonal crystals were observed in a 200 mM sodium fluoride, 30% w/v PEG 1000 condition, and small cylindrical crystals were observed in a 200 mM sodium fluoride, 20% w/v PEG 8000 condition. These conditions were not pursued further due to the amount of time that they required for crystallization. Additionally, by this point it was realized that a condition containing an organic solvent would be needed to aid in the crystallization of luciferase in complex with its substrate, as coelenterazine is sparingly soluble in
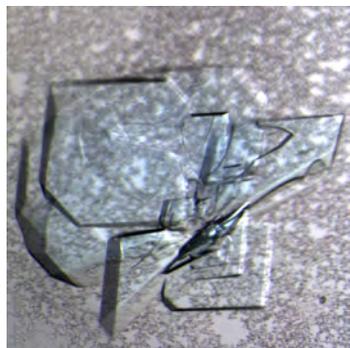
(a) 5 months, 0.4 M sodium acetate, 20% w/v PEG 3350, 1.8 mM n-Decyl-β-D-thiomaltoside
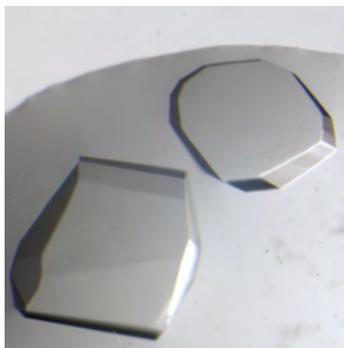
(b) 2 months, 0.5 M NaCl, 1.5 M diammonium phosphate, 0.1 M imidazole pH 8.0
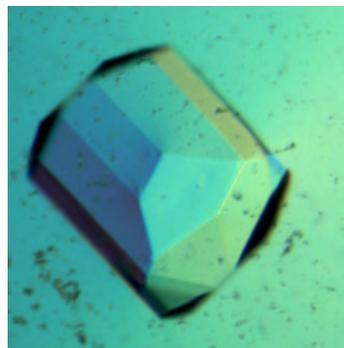
(c) 8 months, 0.3 M NaCl, 1.5 M diammonium phosphate, 0.1 M imidazole pH 8.0
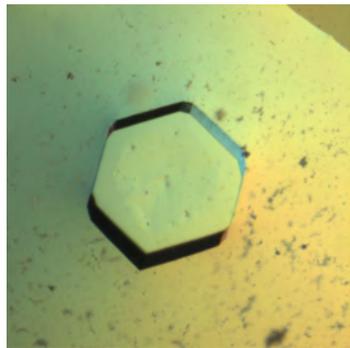
(d) 8 months, 0.5 M NaCl, 1.25 M diammonium phosphate, 0.1 M imidazole pH 8.0
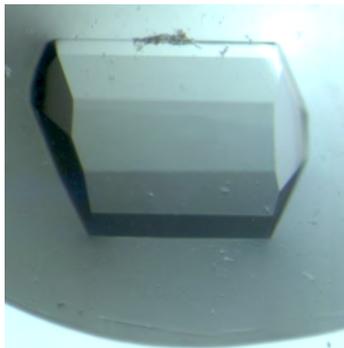
(e) 4 months, 0.5 M NaCl, 1.55 M diammonium phosphate, 0.1 M imidazole pH 8.0
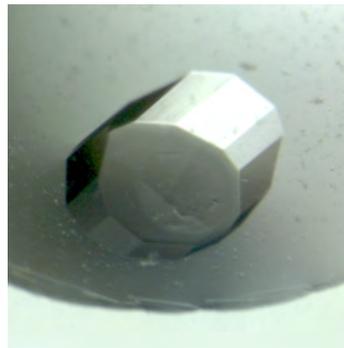
(f) 4 months, 0.5 M NaCl, 1.60 M diammonium phosphate, 0.1 M imidazole pH 8.0
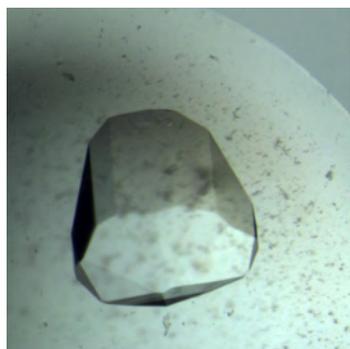
(g) 4 months, 0.6 M NaCl, 1.45 M diammonium phosphate, 0.1 M imidazole pH 8.0
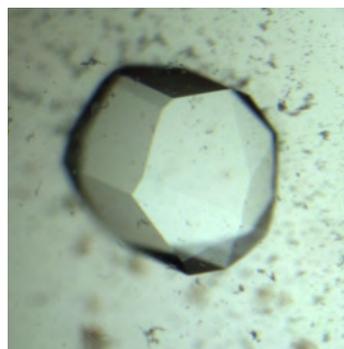
(h) 4 months, 0.7 M NaCl, 1.40 M diammonium phosphate, 0.1 M imidazole pH 8.0

Figure 4.2: Photographs of various crystals formed from S3RLuc8. All the conditions were done using periplasmically expressed S3RLuc8 at 30 mg/ml and stored at 20°C. The length of time following set up of the tray that the image was taken, along with the mother liquor condition, are shown below each photograph.

aqueous solution.

4.2.4 Surface Mutations

It has been estimated that ~80% of proteins that will crystallize will do so in one (or more) of 50 different conditions [90]. The converse statement to this, is that if a protein doesn't crystallize readily using standard screens, further crystallization screening is of diminishing utility.

When initial crystallization trials fail, a common semi-rational method employed to aid in crystallization of a protein is to switch a small number of charged surface residues to alanines [51]. The rationale behind this method, is that removal of hydrophilic residues reduces the amount of entropy loss due to the desolvation associated with making protein-protein contacts. As the tertiary structure of the protein in question is generally not known, glutamates and lysines are usually targeted as these residues can be assumed to be on the surface of a protein [14], especially if more than one of these residues are located in close proximity within the primary sequence.

In our case, the structure of RLuc8 was already known (Figure 4.6), and the goal was not to crystallize the protein but rather to find a new crystallization condition that included an organic solvent in which the protein's substrate could be dissolved. The following pairs of K/A and E/A mutations were chosen "intelligently" based on the crystal contacts of the structure: K12A/E106A, K25A/E277A, and E183A/K227A.

All 7 possible combinations of these three mutations were made. Periplasmic expression was attempted for the initial constructs containing a single pair of mutations, but this proved to yield very low amounts of protein. The success of periplasmic expression is highly dependent on the protein that is being expressed [35, 131], and in this case a single pair of mutations was capable of interfering with the protein production. For this reason, new cytoplasmic expression constructs were made based on the pBAD-RLuc8 (non-pelB containing) plasmid. The yield and specific activity of these cytoplasmically expressed proteins is given in Table 4.1. The emission spectrum of RLuc8/E183A/K227A/K12A/E106A was measured with the use of coelenterazine, and gave a peak of 482 nm and a mean of 503 nm.

Hampton Crystal Screen and Crystal Screen II done at 20°C and a protein concentration of 20 mg/ml yielded nothing for RLuc8, RLuc8/K12A/E106A, RLuc8/K25A/E277A and RLuc8/E183A/K227A. The previous diammonium phosphate (Section 4.2.1) and potassium thiocyanate (Section 4.2.1) were tried as well with no success. The remaining 4 proteins listed in Table 4.1 (5-8) were unsuccessfully screened using the Hampton Index Screen HT.

	Protein	Yield (μg/ml)	Specific Activity (relative to RLuc)
1	RLuc8	528	4.8
2	RLuc8/K12A/E106A	392	5.0
3	RLuc8/K25A/E277A	312	6.1
4	RLuc8/E183A/K227A	418	5.6
5	RLuc8/K12A/K25A/E106A/E277A	140	4.0
6	RLuc8/K25A/E183A/K227A/E277A	40	3.0
7	RLuc8/812A/E106A/E183A/K227A	232	3.4
8	RLuc8/K12A/K25A/E106A/E183A/K227A/E277A	20	4.0

Table 4.1: Yield and specific activity of cytoplasmically expressed RLuc8 and the surface mutation constructs. Yield represents μg of recovered protein per ml of culture. Note that the yield of cytoplasmically expressed RLuc8 is ~10-fold greater than the yield from a similar culture size of periplasmically expressed RLuc8 (Section 2.2.4). Specific activities were measured using coelenterazine, and are reported as relative to that of periplasmically expressed, C-terminal 6xHis tagged RLuc (Table 2.1).

At this point, it was noticed that a minor band was appearing in denaturing gels under non-reducing conditions performed on proteins 5-8 (Table 4.1). This minor band ran at almost the same position as the major band, and disappeared under reducing conditions. This observation was taken to indicate that these particular proteins had some propensity to form intramolecular disulfide bonds between the two cysteines remaining in the sequence. The proteins RLuc8 and RLuc8/K25A/E277A were carefully checked as well, but no such indications of oxidation were seen. In any case, a Hampton Index Screen HT was set up with RLuc8 and RLuc8/K12A/-K25A/E106A/E277A at 20 mg/ml in the presence of 2 mM TCEP, pH 7. This additional screen failed to yield crystals. Additional screens of the proteins in Table 4.1 were done around the diammonium phosphate condition with and without 2 mM TCEP, but with no success.

As discussed in Section 4.2.7 below, crystals were eventually obtained from RLuc8/K25A/-E277A and the structure is presented in Section 4.6. Interestingly, the surface mutations (K25A, E277A) were not involved in the contacts between the proteins in this crystal.

4.2.5   Cocrystallization with RrGFP

As the green fluorescent protein from *Renilla reniformis* (RrGFP) is known to associate with RLuc, it was hoped that a combination of RrGFP and RLuc8 would crystallize under conditions in which RLuc8 alone would not. Additionally, it was hoped that a cocrystal of RLuc8 and RrGFP would elucidate the details of the bioluminescence resonance energy transfer (BRET) that is known to occur between these two proteins. Further details on RrGFP and its role in
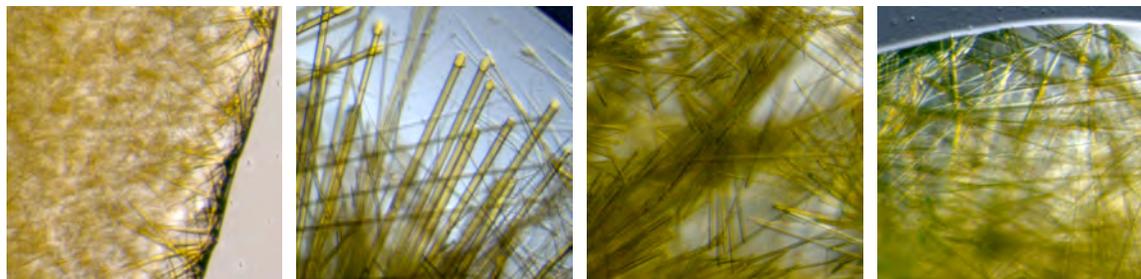
(a) 1 day, 0.2 M magnesium acetate tetrahydrate, 0.1 M sodium cacodylate pH 6.5, 20% PEG 8000

(b) 3 days, 0.1 M MES pH 6.5, 12% w/v PEG 20,000

(c) 3 days, 0.1 M HEPES pH 7.5, 10% w/v PEG 6000, 5% v/v MPD

(d) 3 days, 0.1 M HEPES pH 7.5, 10% w/v PEG 8000, 8% v/v ethylene glycol

Figure 4.3: Results of screening RrGFP/RLuc8 for crystallization conditions. RLuc8 and freshly purified RrGFP were combined in a 2:1 molar ratio (one RrGFP dimer per RLuc8), with the resulting concentration of 38 mg/ml RrGFP and 26 mg/ml RLuc8. Screening was with the Hampton Crystal Screen and Crystal Screen II at 20°C. The length of time following set up before crystals were noticed, along with the mother liquor condition, is shown below each photograph. MPD - 2-methyl-2,4-pentanediol.

BRET are given in Appendix B, along with details of the expression, purification, and structural determination of RrGFP. As things turn out, RrGFP crystallizes amazingly easily in a number of different conditions, yielding crystals of high quality.

The combination of RrGFP and RLuc8 (cytoplasmically expressed) has been more problematic. The Hampton Screen and Hampton Screen II were utilized for identifying the conditions shown in Figure 4.3, and running these crystals on denaturing gels indicated that the needles did indeed contain both RrGFP and RLuc8. However, none of these crystals were of sufficient size to achieve diffraction below 5 Å, and further crystallization trials around these conditions, including macro seeding, have failed to improve upon the size of the crystals. The Hampton Additive Screens 1-3 were attempted on the 0.1 M MES pH 6.5, 12% w/v PEG 20,000 condition, but only pure GFP crystals were achieved. The 0.1 M MES pH 6.5, 12% w/v PEG 20,000 was also tried with various proteins discussed later (e. g. RLuc8/K25A/E277A) taking the place of RLuc8, but to no avail.

4.2.6   N and C-terminal Truncated Variants

At this point, a total of 3 crystal structures had been obtained. From a comparison of these different structures, a feature that stood out was the variability in placement of the first ∼10-15 amino acids of the protein's N-terminus. A representation of this variability can be observed

Protein	Yield (μg/ml)	Specific Activity (relative to RLuc)
S3RLuc8	60	11.9
K12RLuc8	14	6.5
I15RLuc8	34	3.1
S3RLuc8Δ5	37	4.3
I15RLuc8Δ5	~4	ND

Table 4.2: Yield and specific activity of cytoplasmically expressed N and C-terminal truncation variants of RLuc8. Yield represents μg of recovered protein per ml of culture. All constructs had N-terminal cleavable 6xHis tags. Specific activities were measured using coelenterazine, and are reported as relative to that of periplasmically expressed, C-terminal 6xHis tagged RLuc (Table 2.1). ND - Not determined.

in Figure 4.4. As the N-terminal's placement seemed to vary considerably, it was thought that removing portions of the N-terminus might aid in crystallization by making the resultant protein more structurally homogeneous. For this reason, a series of N-terminal truncation mutants of RLuc8 were pursued.

Following the favorable yields of the cytoplasmically expressed RLuc8 construct, it was decided to pursue cytoplasmic expression for the truncation mutants as well. Additionally, these constructs were designed to have an N-terminal cleavable 6xHis tag, and a number of them included a deletion of the last 5 amino acids of the protein. In total, 12 plasmids were made to express the following proteins: S3RLuc8, S3RLuc8Δ5, Y6RLuc8, Y6RLuc8Δ5, E9RLuc8, E9RLuc8Δ5, K12RLuc8, K12RLuc8Δ5, I15RLuc8, I15RLuc8Δ5, P18RLuc8, and P18RLuc8Δ5. Screening bacterial plates containing the constructs showed that the proteins P18RLuc8 and P18RLuc8Δ proteins were not capable of emitting light. Only four of these proteins were actually purified, and the results are given in Table 4.2. No signs of oxidation were seen with these constructs (K12RLuc8, I15RLuc8, S3RLuc8Δ5) on denaturing gel electrophoresis under reducing and non-reducing conditions.

No crystallization conditions were revealed using a Hampton Crystal Screen HT at 20°C with S3RLuc8Δ5 at 25 mg/ml and 6 mg/ml, I15RLuc8 at 25 mg/ml and 6 mg/ml, and I15RLuc8 at 25 mg/ml with 2 mM TCEP pH 7. A Hampton Index Screen, similarly unproductive, was performed at 20°C with S3RLuc8Δ5 at 25 mg/ml and 6 mg/ml, K12RLuc8 at 12 mg/ml and 6 mg/ml, and I15RLuc8 at 25 mg/ml and 6 mg/ml. A Hampton Crystal Screen Classics Light HT was used at 20°C with S3RLuc8Δ5 (25 mg/ml) and I15RLuc8 (24 mg/ml), but failed to produce any crystallization conditions. The previous diammonium phosphate (Section 4.2.1) and potassium thiocyanate (Section 4.2.1) were tried as well with no success.

(a) Diammonium phosphate, (4-308)

(b) Potassium thiocyanate, (4-308)

(c) n-Decyl-β-D-thiomaltoside - 1, (4-309)

(d) n-Decyl-β-D-thiomaltoside - 2, (10-308)
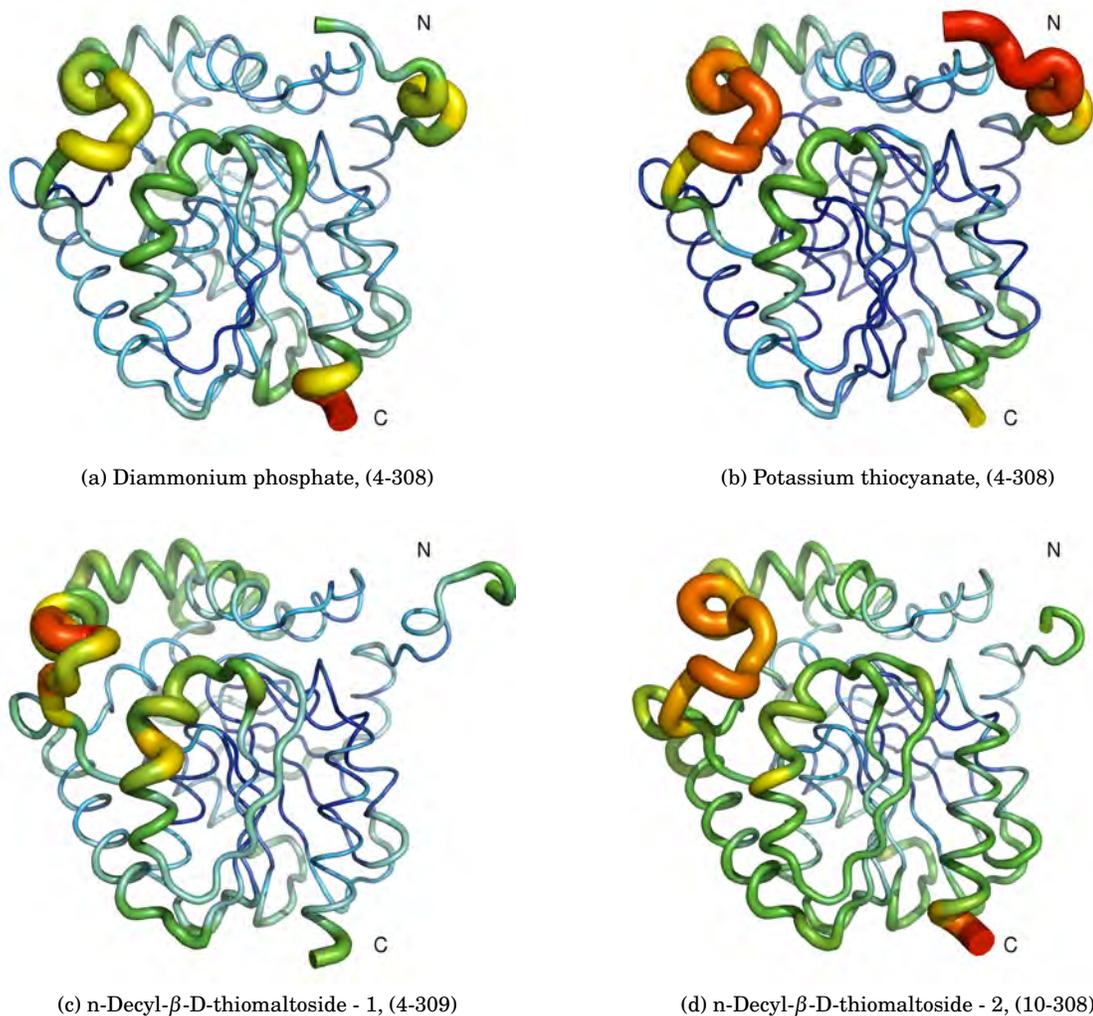
Figure 4.4: Cartoon representations of the B-factors for several of the structures. In theory, the B-factor (or temperature factor) indicates the mobility of a residue. In reality, the B-factor can also indicate errors in model building. The diammonium phosphate structure is from Figure 4.6. The panels labeled N-Decyl-β-D-thiomaltoside are the two proteins in the unit cell of the structure described in Figure 4.9. The potassium thiocyanate structure is described in Section 4.4. The residues visible in each structure, out of a total of 311 possible, are given in parenthesis.

K12RLuc8 was screened using the Hampton Crystal Screen HT at 20°C and a protein concentration of 12 mg/ml. Following, a number of days, needle crystal clusters were seen in the following condition: 0.1 M HEPES pH 7.5, 10% v/v isopropanol, 20% w/v PEG 4000 (Figure 4.5a). As this condition contained a solvent that was amenable to dissolving coelenterazine, extensive further screens were pursued as described in the following section.

4.2.7   Polyethylene Glycol/Isopropanol Condition

The needle clusters of K12RLuc8 that formed in the condition 0.1 M HEPES pH 7.5, 10% v/v isopropanol, 20% w/v PEG 4000, 20°C, proved difficult to duplicate. A sequence of further grid screens around this condition and utilizing the various other previously purified luciferase constructs were done. Through these trials, the following condition was found to give highly reproducible crystals in a ~2 day time period: RLuc8/K25A/E277A at 21 mg/ml in 0.1 M HEPES pH 7.5, 10% isopropanol containing 6 mg/ml benzyl-coelenterazine, w/v 15% PEG 3350, at 20°C. Both benzyl-coelenterazine and coelenterazine were tested with this condition, with benzyl-coelenterazine giving more consistent results. As shown in Figure 4.5, this condition produced long (~100 μm) rectangular crystals that were planar and rather thin. While the Hampton Additive Screens 1-3 were unsuccessful in improving the quality of these crystals, macroseeding was able to improve the thickness of the plates. For crystal freezing, the cryoprotectant used was the mother liquor condition with 35% MPD (2-methyl-2,4-pentanediol). The structure resulting from this condition is presented in Figure 4.10.

A similar condition was found to yield crystals from cytoplasmically expressed RLuc8 after about a month as shown in Figure 4.5c, with the corresponding structure presented in Figure 4.12. Again, 35% MPD in the mother liquor condition was used as the cryoprotectant. This condition was identical to the condition found for RLuc8/K25A/E277A, with the exception that coelenterazine was found to yield better crystals then benzyl-coelenterazine. Crystals obtained using these conditions were squarish, planar, and somewhat more consistent in size and quality than the RLuc8/K25A/E277A crystals. Microseeding of this condition was successful in reducing the length of time required for crystallization.

4.2.8   Additional Cytoplasmic Constructs

A series of additional constructs, presented in Table 4.3, have been expressed for further screening. Interestingly, the variants of RLuc and RLuc8 with cleavable 6xHis tags (S3RLuc

(a) 5 days, 0.1 M HEPES pH 7.5, 10% v/v isopropanol, 20% w/v PEG 4000

(b) 1 day, 0.1 M HEPES pH 7.5, 10% v/v isopropanol with 6 mg/ml benzyl-coelenterazine, 15% w/v PEG 3350

(c) 2 days, 0.1 M HEPES pH 7.5, 10% v/v isopropanol with 6 mg/ml coelenterazine, 15% w/v PEG 3350, microseeded

(d) 2 days, 0.1 M HEPES pH 7.5, 10% v/v isopropanol with 6 mg/ml benzyl-coelenterazine, 15% w/v PEG 3350, microseeded

(e) 1 day, 0.1 M HEPES pH 7.5, 10% v/v isopropanol with 6 mg/ml benzyl-coelenterazine, 15% w/v PEG 3350, microseeded

(f) 1 day, 0.1 M HEPES pH 7.5, 10% v/v isopropanol with 6 mg/ml benzyl-coelenterazine, 15% w/v PEG 3350, microseeded

(g) 2 days, 0.1 M HEPES pH 7.5, 10% v/v isopropanol with 6 mg/ml benzyl-coelenterazine, 15% w/v PEG 3350, microseeded

(h) 1 day, 0.1 M HEPES pH 7.5, 10% v/v isopropanol with 6 mg/ml benzyl-coelenterazine, 15% w/v PEG 3350, microseeded

Figure 4.5: Photographs of crystals produced under various PEG/isopropanol conditions. All conditions were at 20°C. Panel (a) used K12RLuc8 at a concentration of 12 mg/ml. Panel (b) used RLuc8/K25A/E277A at a concentration of 18 mg/ml Panel (c) used RLuc8 at a concentration of 21 mg/ml. Panel (d) used S3RLuc8/K25A/E277A at 23 mg/ml. Panel (e) used S3RLuc8/-K25A/D120A/E277A at 21 mg/ml. Panel (f) used S3RLuc8/K25A/E144A/E277A at 20 mg/ml. Panel (g) used S3RLuc8/K25A/E144Q/E277A at 22 mg/ml. Panel (h) used S3RLuc8/K25A/-E277A/H285A at 18 mg/ml. The length of time following set up of the tray that the image was taken, along with the mother liquor condition, is shown below each photograph. With the exception of Panel (a), all photographs were acquired with the aid of a polarizer to highlight the crystals.

Protein	Yield (μg/ml)	Specific Activity (relative to RLuc)
S3RLuc	10	4.0
S3RLuc8	60	11.9
S3RLuc8/C24S	60	6.8
S3RLuc8/K25A/E277A	60	7.1
S3RLuc8/K25A/D120A/E277A	24	0.002
S3RLuc8/K25A/E144A/E277A	4	7.0×10^{-6}
S3RLuc8/K25A/E144Q/E277A	5	0.01
S3RLuc8/K25A/E277A/H285A	7	0.1

Table 4.3: Yield and specific activity of additional proteins expressed for crystallography screening. All constructs were cytoplasmically expressed, and had N-terminal cleavable 6xHis tags. Yield represents μg of recovered protein per ml of culture following nickel affinity purification. Specific activities were measured using coelenterazine, and are reported as relative to that of periplasmically expressed, C-terminal 6xHis tagged RLuc (Table 2.1). Data for S3RLuc8 is repeated from Table 4.2.

and S3RLuc8) were several fold more active than previous periplasmically expressed 6xHis tagged RLuc and RLuc8 (Table 2.1), although the relative ratio between RLuc and RLuc8 activity remained relatively constant. S3RLuc was set up in a Hampton Index HT Screen at 20°C and 21 mg/ml, but no crystals resulted.

The protein S3RLuc8/C24S was expressed due to lingering concerns regarding the possibility that the preceding constructs may occasional incur a disulfide bond between C24 and C73 and therefore be structurally heterogeneous. A Hampton Index HT Screen done with this protein at 24 mg/ml and 20°C failed to produce any crystallization conditions. The protein S3RLuc8/K25A/E277A was expressed with the hope that it would produce better quality crystals in the PEG 3350/isopropanol condition than RLuc8/K25A/E277A. Microseeding this protein from the RLuc8/K25A/E277A crystals proved successful, and visually better quality crystals were formed (Figure 4.5d).

Finally, S3RLuc8/K25A/E277A based constructs containing the catalytic triad mutations (D120A, E144A, E144Q, H285A) were produced. This work was done with the hope that these potentially inactivating mutations would allow crystallization of the protein in complex with the substrate rather than the product. All these proteins were initially microseeded from the RLuc8/K25A/E277A condition in the PEG 3350/isopropanol condition containing 0.6 mg/ml benzyl-coelenterazine, and then successfully microseeded off themselves in later crystallography setups. However, neither substrate nor product was observed in the diffraction data from the D120A, E144A, or E144Q mutation containing crystals. The H285A crystals, in turn, have

not yet been used successfully for X-ray diffraction.

For the D120A, E144Q, and H285A mutations, the color of the hanging-drops could be observed to change from a deep yellow to a pale yellow color within approximately one day. As benzyl-coelenterazine is a strong absorber at blue wavelengths (resulting in a yellow color) but coelenteramide is not [86], the observed color change was taken to indicate that sufficient enzymatic activity remained with these mutants so that coelenterazine was still being degraded to coelenteramide. The E144A drops, however, remained a deep yellow color during the crystallization process.

4.3 Structure of RLuc8 from the Diammonium Phosphate Condition

The structure derived from the RLuc8 crystal grown in the diammonium phosphate condition is presented in Figure 4.6, with the corresponding statistics presented in Table 4.4. Residues 4-308 of the protein were successfully identified in the electron density data. Not identified were two residues from the N-terminus, along with 3 residues on the C-terminus and the 6xHis tag.

Interestingly, two imidazole molecules, apparently from the mother liquor, were located in the presumptive catalytic pocket of the molecule. A close up of these imidazoles and associated water present in the active pocket are shown in Figure 4.7. Previous reports [137] have reported enhanced enzymatic activity of RLuc in the presence of imidazole, with a maximal activity enhancement of 2-fold at $\sim$4 mM. While the reason for this potentiation remains unclear, a plausible explanation is that imidazole maintains the enzymatic pocket in a conformation appropriate for the binding of coelenterazine.

Much like the similar bacterial haloalkane dehalogenase enzymes [82], *Renilla* luciferase has a characteristic α/β-hydrolase fold sequence at its core [166]. A topological map of the RLuc8 α/β-hydrolase fold is shown in Figure 4.8a, along with the locations of the presumptive catalytic residues D120/E144/H285A within this diagram. α/β-hydrolases have their nucleophile (D120 in RLuc) immediately after the fifth β-sheet ($\beta5$) in what is termed the "nucleophile elbow". The sequence pattern for this elbow is generally G-X-Nuc-X-G [75], and corresponds to GHDWG (residues 118-122) in *Renilla*. For unknown reasons, many of the haloalkane dehalogenases that RLuc is most similar to do not contain the first glycine in the nucleophile elbow sequence (Figure 2.4).

The most interesting feature observable in Figure 4.6 that was not apparent in the pre-

	RLuc8 diammonium	RLuc8 KSCN	S3RLuc8 thiomaltoside	RLuc8/K25A/E277A PEG/isopropanol	RLuc8 PEG/isopropanol
Cell Parameters					
space group	P 6_1	P 6_1	P $2_1 2_1 2_1$	P 2_1	P 2_1
protomers/asymmetric unit	1	1	2	2	2
dimensions (Å)					
a	119.468	119.443	81.178	51.776	51.329
b	119.468	119.443	82.279	75.672	74.473
c	47.995	48.048	90.379	89.185	89.249
angles (degrees)					
α	90.00	90.00	90.00	90.0	90.00
β	90.00	90.00	90.00	76.48	103.45
γ	120.00	120.00	90.00	90.00	90.00
Data Collection Statistics					
number of reflections	76,992	12,067	117,436	60,992	58,532
possible reflections	77,128	13,522	119,443	62,054	60,685
completeness (%)	99.8	89.2	98.3	98.3	96.5
Model Statistics					
resolution range (Å)	50-1.4	50-2.5	50-1.4	50-1.8	50-1.8
R_{free} (%)	18.3	32.4	30.5	22.9	22.4
$R_{working}$ (%)	16.5	27.2	28.7	19.7	19.2

Table 4.4: Statistics for the crystallographic structures of RLuc8 and related proteins. RLuc8:diammonium corresponds to the structure presented in Figure 4.6. RLuc8:KSCN is the potassium thiocyanate condition discussed in Section 4.4. S3RLuc8:thiomaltoside is the crystal shown in Figure 4.9, note that refinement of the diffraction data from this condition was not completed. RLuc8/K25A/E277A:PEG/isopropanol is the structure presented in Figure 4.10. RLuc8:PEG/isopropanol is the coelenteramide containing structure presented in Figure 4.12. The cross-validation statistic R_{free} was computed from a randomly chosen subset (5%) of the diffraction data that had been excluded from the model refinement process [26]. $R_{working}$ was calculated as $\sum_{hkl} ||F_{obs}| - |F_{calc}||/\sum_{hkl} |F_{obs}|$.
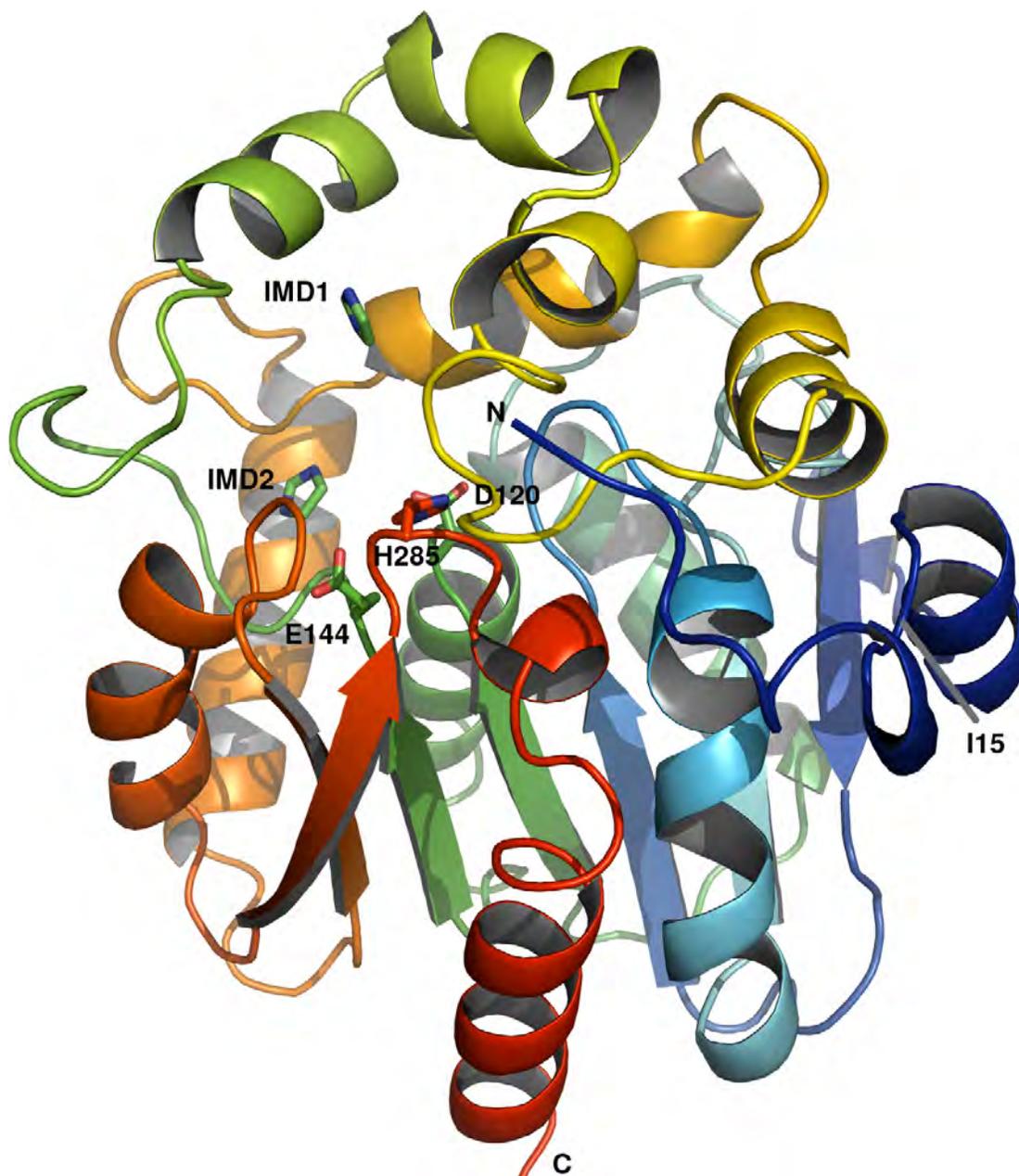
Figure 4.6: A cartoon representation of the structure of RLuc8 crystallized in the diammonium phosphate condition. Residues 4-308 of RLuc8 are shown, with the N-terminus (N) in blue and the C-terminus (C) in red. The presumptive catalytic triad residues of D120, E144, and H285 are marked, along with the two imidazole molecules (IMD1, IMD2) that were present in the structure and the location of I15. The condition was 0.3 M NaCl, 1.25 M diammonium phosphate, 0.1 M imidazole pH 8.0, 20°C, and the corresponding photograph of the crystal can be seen in Figure 4.1a. The orientation of the structure here has been matched with that of Figure 2.5 to facilitate comparisons with the Swiss-Model derived homology model.
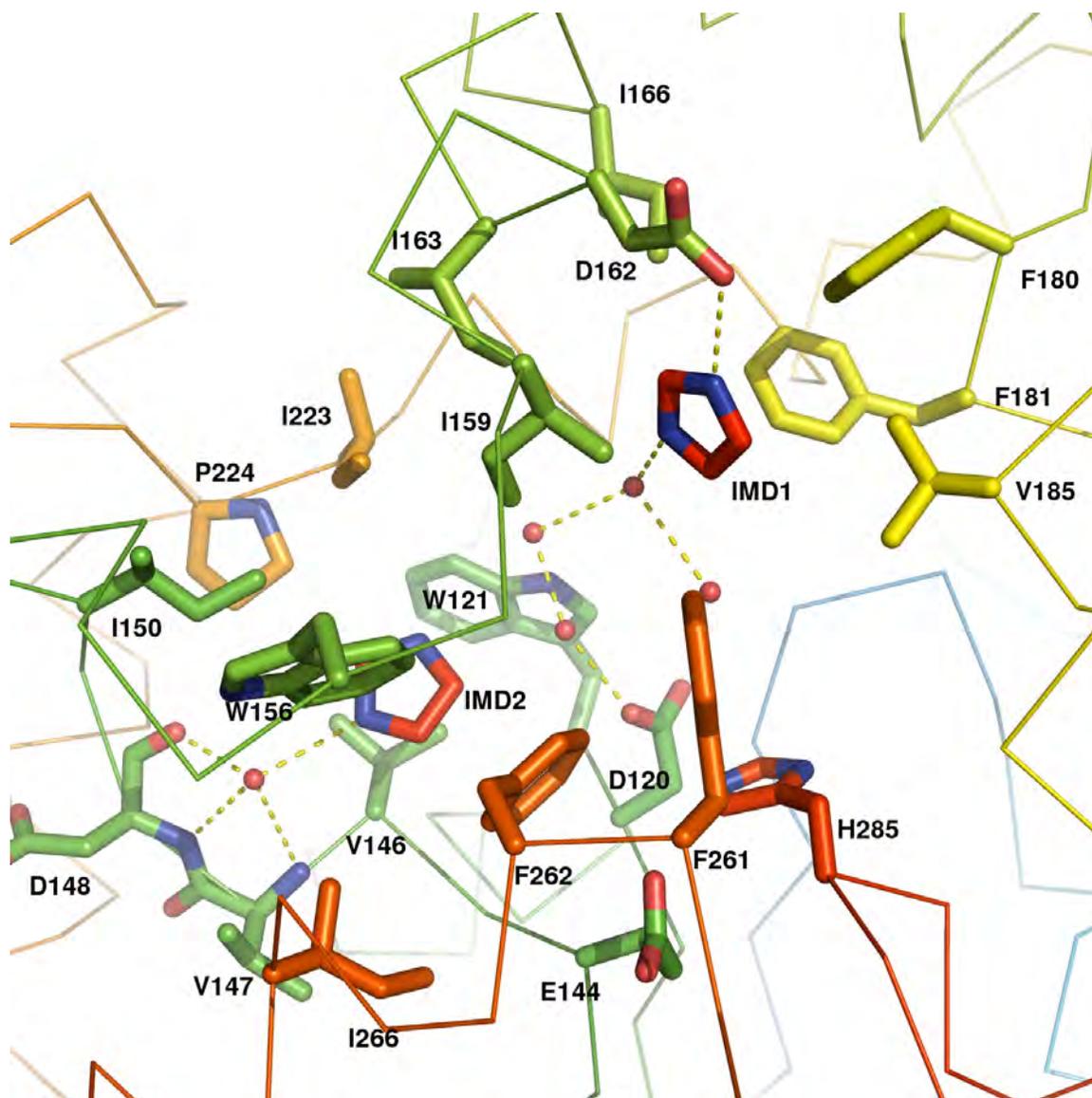
Figure 4.7: A cartoon representation of the imidazoles and water molecules within the active pocket of RLuc8 crystallized in the diammonium phosphate condition. The red balls represent water molecules, the yellow dashed lines represent predicted hydrogen bonds, and the two imidazoles are marked as in Figure 4.6.

(a) Topology Diagram
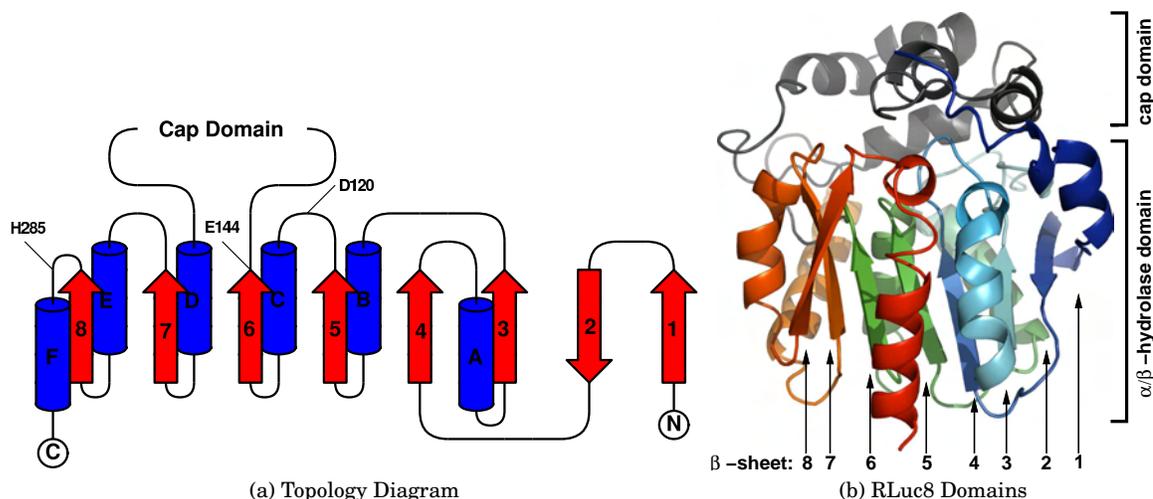
(b) RLuc8 Domains

Figure 4.8: The domains and α/β-hydrolase fold topology of RLuc8. Panel (a) shows the topology of RLuc8's α/β-hydrolase fold domain. α-helices are shown in blue, and β-sheets are shown in red. Numbering/lettering of the sheets/helices is done with respect to the standard for α/β-hydrolases [166], and the locations of the presumptive catalytic residues are marked. The cap domain is an excursion from the fold pattern from residues 146-230. Panel (b) shows the location of the cap domain (in gray) and α/β-hydrolase fold domain (blue to red) in the context of the crystal structure.

vious homology model (Figure 2.5), is the wrapping of the N-terminus around the front of the presumptive enzymatic pocket. As demonstrated in Table 4.2, clipping the N-terminal up to position I15 is tolerated, indicating that this N-terminal region is not required for enzymatic activity of the protein. It has been observed in our laboratory (unpublished data) that fusion proteins created by attachments to the N-terminus of RLuc invariably lead to very low luciferase activity of the resulting protein. Based on the structural data and the non-essentialness of the N-terminal region, it can be hypothesized that this drop in activity is simply a matter of steric hindrance to the active pocket of RLuc.

4.4 Structure of RLuc8 from the Potassium Thiocyanate Condition

The resultant structure from the potassium thiocyanate condition was identical to that already obtained from the diammonium condition shown in Figure 4.6 with the sole exception that only one of the imidazoles (IMD1 in Figure 4.6 was observable in the data. The corre-

sponding crystallography statistics are presented in Table 4.4. Interestingly, the mother liquor for this condition did not include imidazole. This would seem to indicate that the imidazole is retained with the protein during the nickel affinity purification, and is bound tightly enough to remain attach through at least one additional step of chromatography and a buffer exchange procedure.

4.5 Structure of S3RLuc8 from the n-Decyl-β-D-thiomaltoside condition

The structure for the condition containing the detergent n-Decyl-β-D-thiomaltoside is shown in Figure 4.9, and the corresponding crystallography statistics are presented in Table 4.4. This crystal was in a different space group than the previous diammonium phosphate and potassium thiocyanate conditions, containing two protomers in the asymmetric unit. Interestingly, the N-terminal (residues 4-13) of one of the two proteins in the unit cell (the right protein in Figure 4.9) interacts with its neighboring protein. This is a shift in placement for the N-terminus from the previous structures, and the finding was taken to imply that this N-terminal region of RLuc8 is structurally flexible.

4.6 Structures from the PEG 3350/Isopropanol Condition

The initial crystal structure obtained from the PEG 3350/isopropanol crystallization condition, shown in Figure 4.10, was for the protein RLuc8/K25A/E277A. The corresponding crystallography statistics are presented in Table 4.4. A relatively large region of the cap domain (residues 153-162) could not be identified in the electron density data, and this may indicate that this region of the protein can exist in a number of different conformational states. In addition to these missing residues, neither substrate nor product could be identified in the data.

The second structure obtained from the PEG 3350/isopropanol condition utilized cytoplasmically expressed RLuc8. This structure is presented in Figure 4.11, and the associated statistics can be found in Table 4.4. Notably, electron density corresponding to coelenteramide was present in the data.

The main conformation changes in this structure, compared to the previous diammonium phosphate condition (Figure 4.6), were a slight outward shift of the residues F261/F262/S263 and a larger outward movement of residues from W153 to A163. Residues 153-163 are within
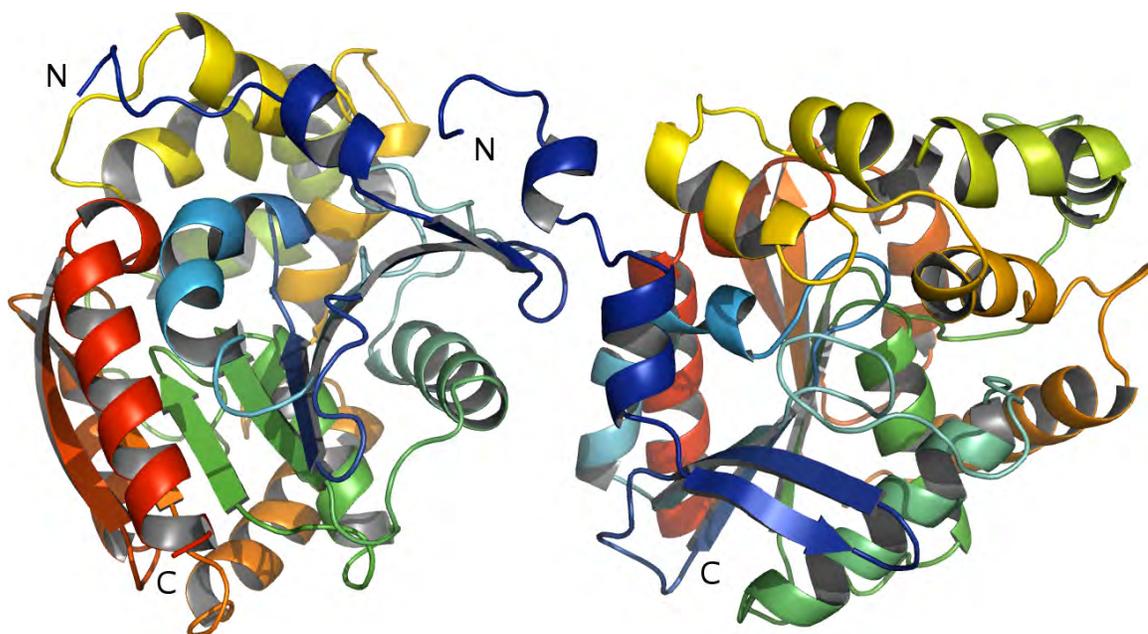
Figure 4.9: A cartoon representation of the structure of S3RLuc8 crystallized in the n-Decyl--β-D-thiomaltoside condition. There are two proteins in the unit cell. Residues 10 to 308 were visible for the protein on the left, and residues 4 to 309 were visible for the protein on the right. The N-termini (N) are in blue, and the C-termini (C) are in red. The condition was 0.4 M sodium acetate, 20% w/v PEG 3350, 1.8 mM n-Decyl-β-D-thiomaltoside, 20°C, using periplasmically expressed S3RLuc8 at 30 mg/ml. The corresponding photograph of the crystal can be seen in Figure 4.2a.
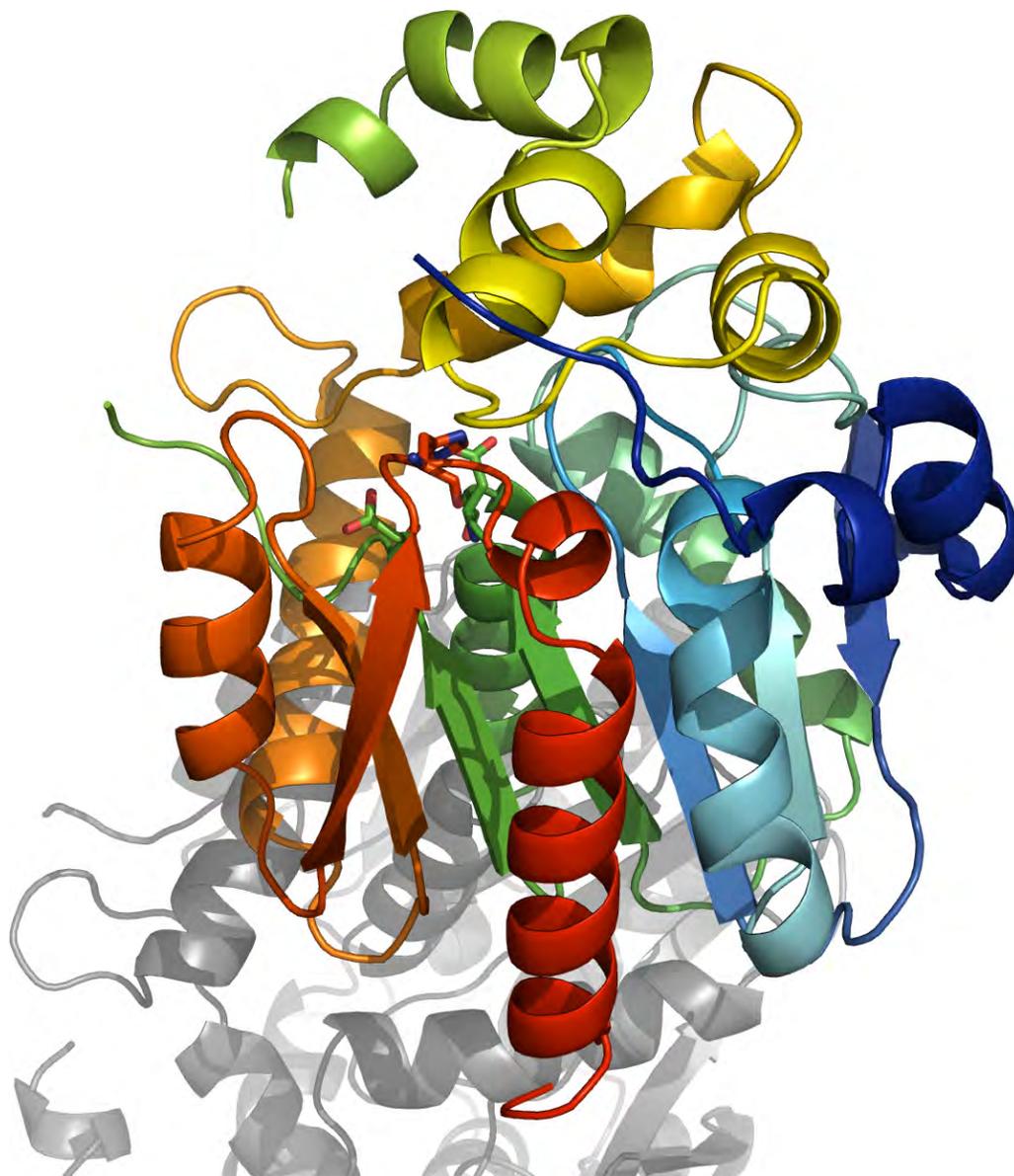
Figure 4.10: A cartoon representation of the structure of RLuc8/K25A/E277A crystallized in the PEG/isopropanol condition. Residues 1-152 and 163-310 are shown. The locations of residues 153-162 could not be determined in the electron density data and have not been modeled in. Electron density for neither benzyl-coelenterazine nor the reaction product (benzyl-coelenteramide) was seen in the diffraction data. The N-terminus is in blue, the C-terminus is in red, and the residues of the presumptive catalytic triad of D120, E144, and H285 are shown. The gray protein at bottom is the second molecule in the crystallographic unit. The condition was 0.1 M HEPES pH 7.5, 10% v/v isopropanol with 6 mg/ml benzyl-coelenterazine, 15% w/v PEG 3350, and used cytoplasmically expressed protein at 18 mg/ml. The corresponding photograph of the crystal can be seen in Figure 4.5b.

Figure 4.11: A cartoon representation of the structure of RLuc8 in complex with coelenter-amide crystallized in the PEG/isopropanol condition. Residues 3-152 and 155-310 are shown. The locations of residues 153-154 could not be determined in the electron density data and have not been modeled in. The N-terminus is in blue, the C-terminus is in red, and the molecule of coelenteramide is in orange. The portion of the enzyme corresponding to residues 153-163 has been marked. The residues of the presumptive catalytic triad of D120, E144, and H285 are shown. The gray protein at bottom is the second molecule of RLuc8 in the crystallographic unit. The condition was 0.1 M HEPES pH 7.5, 10% v/v isopropanol with 6 mg/ml coelenterazine, 15% w/v PEG 3350. The corresponding photograph of the crystal can be seen in Figure 4.5c.

the cap domain (Figure 4.8) of the enzyme, a domain that has been suggested to be flexible for the purposes of substrate binding in the haloalkane dehalogenases [183]. It can be expected that portions of the cap domain in *Renilla* luciferase, specifically residues 153-163, are similarly flexible for the same purpose. The finding of flexibility is further supported by the high B-factors found for the 153-163 residues (Figure 4.4), and the outward movement of this portion of the enzyme may indicate conformational changes in response to binding of the coelenteramide.

It was observed during the screening process that crystallization was more successful utilizing coelenterazine with RLuc8, and benzyl-coelenterazine with RLuc8/K25A/E277A. A possible explanation for this can be proposed based on the interactions of coelenteramide with the protein's residues as shown in Figure 4.12 (a corresponding electron density map is shown in Figure 4.13). This closeup shows that a hydroxyl group in coelenteramide is interacting with E277 through a predicted hydrogen bound. In RLuc8/K25A/E277A, this glutamate has been mutated to alanine and this predicted hydrogen bound cannot form. Benzyl-coelenterazine, however, lacks this particular hydroxyl group, and presumably there is a hydrophobic interaction between A277 and the hydroxyl-lacking benzene ring.

Several lines of evidence indicate that the observed location for coelenteramide in this structure is not its location during the enzymatic reaction and resultant emission of the bioluminescence photon. First, if the coelenteramide location shown in Figure 4.12 was the catalytic location, it would seem to indicate that two monomers of RLuc are involved in the enzymatic reaction. The reaction rate of RLuc, however, is first order with respect to enzyme concentration (Figure A.4) indicating that only one protein unit is involved in the reaction. Second, the residue N309 can be truncated out with little change in enzymatic activity (Table 4.2). Third, the K25A/E277A mutations resulted in only a 40% drop in activity (Table 4.3). One might expect a much larger drop in activity if E277 was actually involved in the enzymatic activity. Finally, many of the residues in the putative active pocket identified as being important for activity (Table 3.2) would be rather distant from the substrate/product if the observed location was correct.

It has been previously noted that the fluorescent emission spectrum of RLuc mixed with coelenteramide does not reconstitute the recorded bioluminescence emission spectrum [138]. Coelenteramide, however, is known to strongly inhibit the enzymatic reaction [139] so it must be able to bind to RLuc tightly. The explanation for this phenomenon has been that the chemical environment coelenteramide experiences changes immediately after emission of the bio-
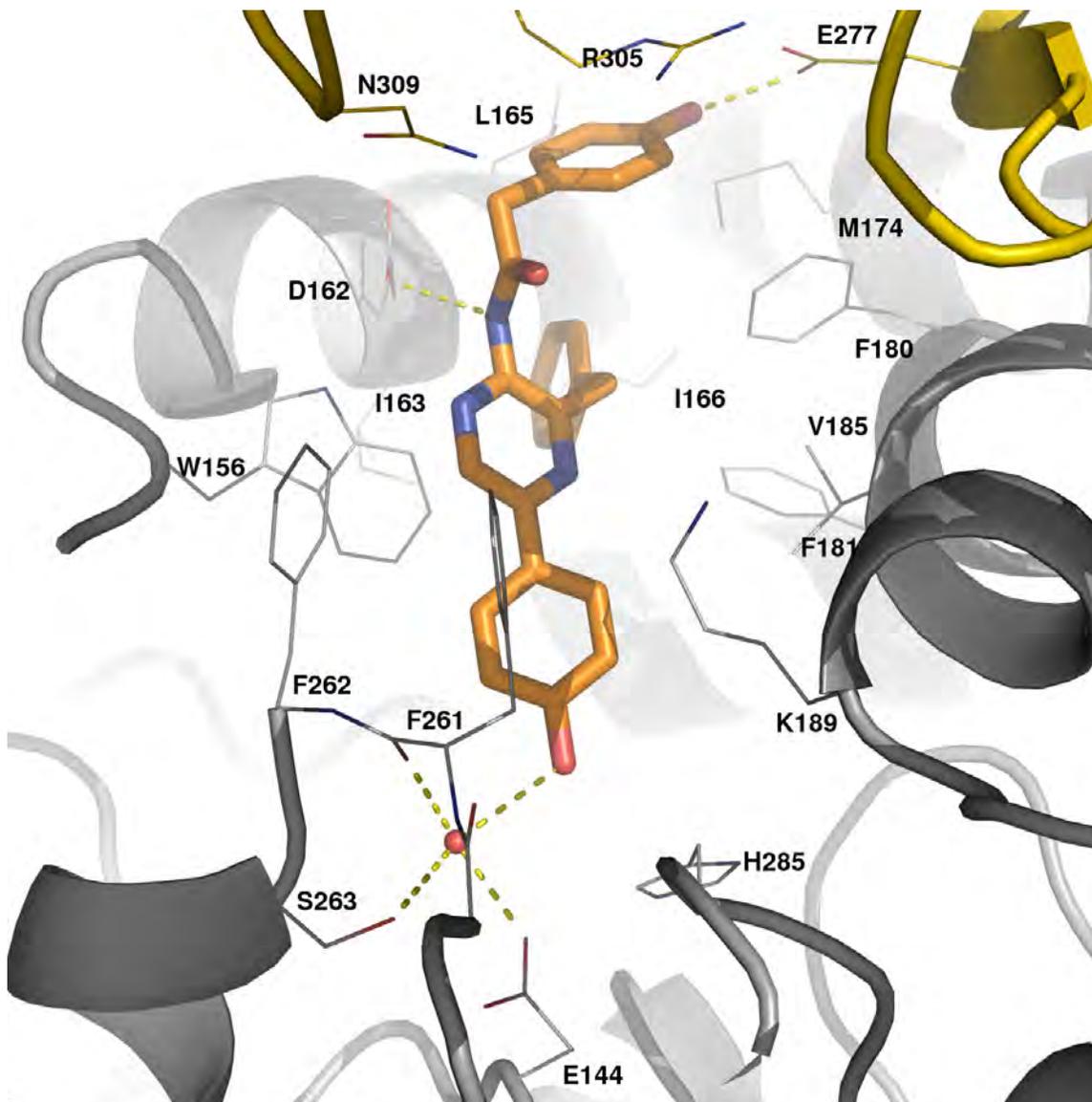
Figure 4.12: A close-up cartoon representation of the structure of RLuc8 in complex with coelenteramide crystallized in the PEG/isopropanol condition. Coelenteramide is shown in orange, the luciferase molecule binding the coelenteramide is shown in gray, and the neighboring luciferase is shown in yellow. The red balls represent water molecules, and the yellow dashed lines represent predicted hydrogen bonds.
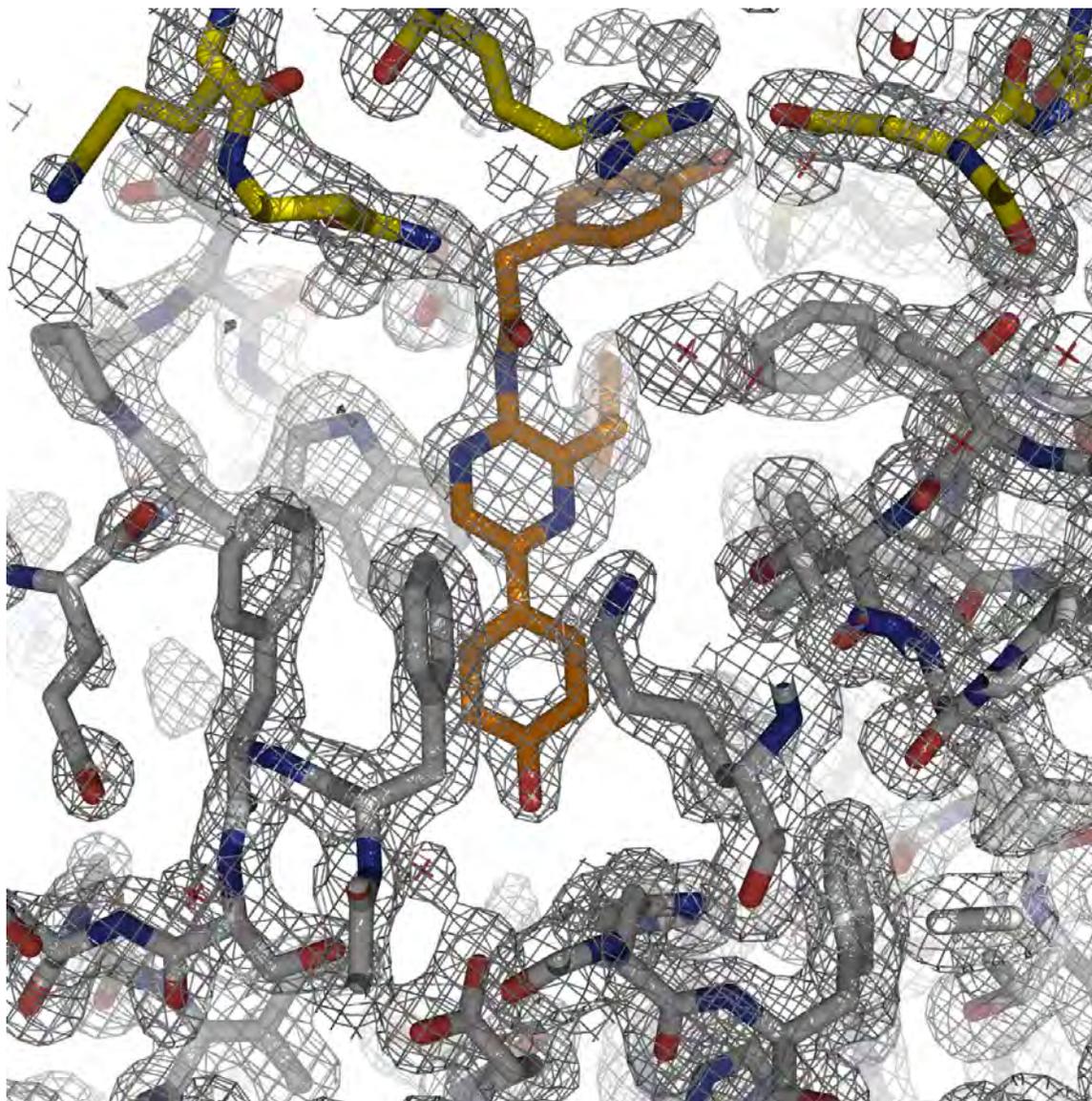
Figure 4.13: A view of the electron density associated with coelenteramide bound to RLuc8. Coelenteramide is shown in orange, the luciferase molecule binding the coelenteramide is shown in gray, and the neighboring luciferase is shown in yellow. The electron density information is shown as a mesh surrounding the atoms. The surrounding residues are identified in Figure 4.12.

luminescence photon [137]. This in turn leads us to the hypothesis that the coelenteramide location changes immediately after the enzymatic reaction and emission of the bioluminescence photon, with the coelenteramide sliding partially out of the presumptive active pocket due to a conformational change in the luciferase's structure. In this hypothesis, the location of coelenteramide in the crystal structure represents this "secondary" binding position and not the position of product/substrate during the enzymatic reaction.

4.7 Conclusions

At this point, the elucidated structures of RLuc8 beg as many questions as they offer answers. While the coelenteramide bound structure of RLuc8 is fascinating, it seems quite likely that the product is not bound in the location at which it undergoes enzymatic processing. To further elucidate the enzymatic reaction of *Renilla* luciferase, additional studies need to be done. A number of mechanism-based coelenterazine analog inhibitors have been previously synthesized [154] and may prove useful for both future crystallography work and for the study of the luciferase's enzymatic kinetics. Additionally, further work is ongoing in attempts to crystallize RLuc8 under anoxic conditions.

Chapter 5

AMIDE - Software for Multimodality Medical Image Analysis

In the molecular and medical imaging community today, there is a paucity of software tools available for volumetric image analysis that are freely available, modifiable, and relatively feature complete. While a number of packages exist, few of these packages encompass all the features that a researcher may desire, and only a subset of these can be freely modified by the researcher for her or his needs (see Table 5.1). For a researcher wishing to do multimodality image analysis, the choices are further constrained as the majority of packages are restricted to strictly orthogonal or planar processing of data sets. This particular limitation has become more pronounced as the role of multimodality imaging has increased in importance [135].

In light of this need, AMIDE [124] (Amide's a Medical Image Data Examiner) has been developed to provide the research community with a relatively full-featured, freely available, and open source solution for single and multimodality volumetric medical image analysis. AMIDE, licensed under the GNU General Public License GPL [61], is freely modifiable and redistributable, and is not dependent on any proprietary underlying packages.

In addition to being open source, AMIDE is unique in that it has been designed to avoid specific constraints of previous software packages. Data sets (e. g. PET, CT, MRI) and regions of interest (ROI's) are logically organized within a tree structure so that an unlimited number of

Package	Free	Source Code	Interface	Fusion	Platform Compatibility	URL
AMIDE	Yes	Open	Non-orthogonal	Yes	Windows, Mac OS X, Unix	amide.sf.net
CAPP	No	No	Orthogonal	No	Solaris	www.cti-pet.com/www/products.nsf/pages/ecat.htm
Hermes	No	No	Non-orthogonal	Yes	Solaris	www.nuclear-diagnostics.com/proc/processing.shtml
Mediman	Yes	No	Slice Based	No	Unix	www.topo.ucl.ac.be/iv_mediman.html
MIM	No	No	Non-orthogonal	Yes	Windows, Mac OS	www.zalen.com
MRIcro	Yes	No	Orthogonal	Overlap only	Windows, Linux	www.psychology.nottingham.ac.uk/staff/cr1/mricro.html
NucMed Image	Yes	No	Slice Based	No	Mac OS	nucmed.sluh.edu/NucMed_Image/NucMed_Image.html
OSIRIS	Yes	300€	Slice Based	Limited	Windows, Mac, Unix	www.expasy.org/www/UIN/html1/projects/osiris/osiris.html
3D-Doctor	No	No	Orthogonal	Yes	Windows	www.ablesw.com/3d-doctor/3ddoctor.html
Syngo	No	No	Non-orthogonal	Yes	Windows	www.syngo.com

Table 5.1: A list comparing AMIDE with several available molecular imaging software packages circa 2003. The "non-orthogonal" interface designation indicates that the program does not need to reslice the entire data set when showing a non-orthogonal viewing plane.

these items can be displayed, modified, and analyzed simultaneously. Furthermore, data sets in AMIDE are not restricted to processing along orthogonal directions. Instead, information is continuously interpolated as needed from the original data to allow for non-destructive and non-orthogonal reslicing of anisotropic data sets. This ability facilitates manual alignment and fused viewing of multiple medical images within an AMIDE session, and allows for seamless handling of data sets with differing voxel sizes and dimensions.

Another key design goal of AMIDE was to avoid encumbering it with an overly complex user interface. With the recent extension of medical imaging modalities into the realm of small animal research (e. g. microCAT, microPET), there has been a steady increase in the number of basic science researchers using these technologies who have not trained in medical imaging science. One of the major hurdles encountered by these researchers has been negative experiences with existing software packages. With this in mind, development has aimed at providing a consistent and intuitive interface for the casual research user. As one step in this process, AMIDE abstracts away the underlying digital representation of the medical data set whenever possible. For instance, the user is not presented with a fixed image plane and voxel based dimensions. Instead, slices of data are automatically extracted from the volumetric data sets at any user specified angle and thickness. Additionally, dimensions are handled in terms of real world units, and image units and statistics can be presented in terms of Percent Injected Dose per gram tissue (%ID/g) or Standardized Uptake Value (SUV) metrics.

AMIDE provides a variety of additional features useful to the molecular imaging researcher, including fully three dimensional ROI drawing and analysis for static and dynamic images, two and three way linked viewing (dual cursor mode), rigid body registration using fiducial markers, filtering and cropping of data sets, movie generation, series viewing, volume rendering, and the generation of line profiles.

5.1 Description

5.1.1 Underlying Concepts

The data hierarchy within AMIDE is built around a tree abstraction composed of a succession of objects such as data sets and ROI's (described below). Conceptually, any object type can be the child of any other object type, although not all pairings are necessarily logical. Generally, data set objects will be the children of the study object, and ROI objects will be the children of either the study object or a specific data set object. The tree based hierarchy allows operations performed on an object, such as shifts and rotations, to be successively mapped down to all of that object's children.

The following object types have been implemented in AMIDE:

Study The root object in AMIDE, this object is used for grouping a set of related medical images and ROI's into a logical unit, and keeps track of parameters that affect the whole study.

Data Set Used for encapsulating volumetric medical images, this object contains the raw image data along with information needed for interpreting that data (voxel sizes, color table, thresholds, patient weight, injected dose, calibration factors, etc.).

ROI Region of interest objects specify a volume of space over which statistics are to be calculated. Statistics include mean/median/max*etc.*, and can be calculated using all the voxels in the ROI or a defined subset [111]. Currently implemented ROI types are ellipsoid, box, cylinder, isocontour (2D or 3D), and freehand ROIs.

Fiducial Marker Fiducial reference markers encode only a location in space and are used for rigid body registration of data sets.

Each object in AMIDE is assigned its own Euclidean space, and the location of this local coordinate frame is defined with respect to the global coordinate frame. When information from one object is needed by another object, AMIDE automatically handles the requisite affine (linear plus translation) transformations between the spaces, as shown in Figure 5.1. This approach allows the rotation or movement of a data set object to be accomplished by a simple alteration of the parameters specifying the object's local coordinate frame, rather than the destructive reslicing of the image data.
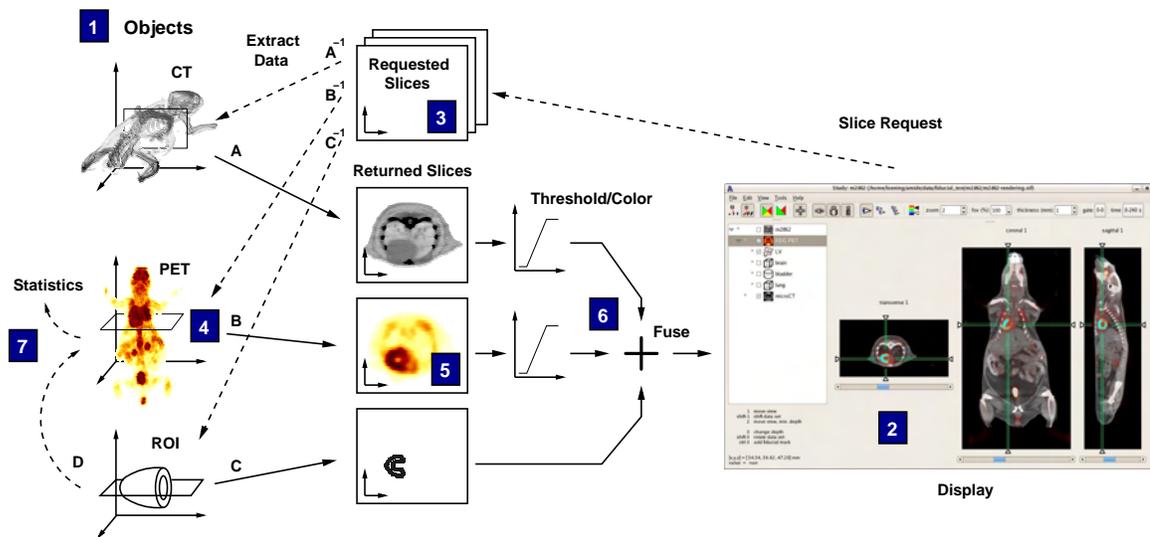
Figure 5.1: Diagram of coordinate transforms done by AMIDE. Each object in AMIDE (1) is defined with respect to its own Euclidean space, with transformations between different spaces performed as needed. When the user clicks on the display (2), a slice request is formed for each data set (3). The program then translates each slice request into the local coordinate frame of the appropriate data set (4), and fills the slice with interpolated data (5). The extracted data is thresholded, colored, and fused (6) into a single image for display to the user (2). Similarly, to calculate statistics for an ROI, the three dimensional ROI geometry is transformed into the Euclidean space of the data set (7), and the voxels enclosed in this volume are iterated over to tabulate the statistical values.

5.1.2 Implementation

The C programming language was chosen for the coding of AMIDE for several reasons, the three most important being the general familiarity of most researchers with this language, the ready availability of high performance C compilers on current operating systems, and the desire to avoid burdening the program with requirements on underlying proprietary packages such as Matlab or IDL. The specific compiler used in this work was GCC (GNU Compiler Collection, gcc.gnu.org) [198].

Version 2 of the GTK+/GNOME toolkit (www.gnome.org) was used for the user interface and object model. This toolkit was chosen for a combination of portability, a C language interface, and free licensing. Much of the core functionality of AMIDE has been written as an extension to this toolkit in order to provide a convenient interface for using AMIDE functionality in separate pieces of code.

Raw data in AMIDE is stored in the data format (8/16/32 bit integer, 32/64 bit float) and with the voxel size (isotropic or anisotropic) of the imported data file. The original data is never altered, rather, the program interpolates directly from the original data set as needed. This approach makes data set movements, scalings, and rotations computationally trivial as only the associated coordinate information is altered for these operations. The trade-off is that slice viewing is computationally more expensive compared to standard orthogonal data viewing. Zero order (nearest neighbor) and first order (trilinear) interpolation algorithms have been implemented for speed and image quality, respectively. Higher order interpolation methods [208] have not been employed since successive interpolations are never performed and because these interpolators become computationally prohibitive in three dimensions.

AMIDE saves studies in an XML (eXtensible Markup Language) based format using either a directory based or flat file structure. For the directory based version of the format, each object's parameters and data format information are saved as XML data in a text file, with the raw image data saved as a separate binary file. This approach allows data files to be easily manipulated externally to the program if the need ever arises. The flat file version of the format, which is the default, is similar to the directory based version except that the binary and text data are saved as different segments in a single file. This XML based format helps guard against endian incompatibilities (incompatibilities due to the inconsistent ordering of stored data between different processor architectures), and makes backward and forward file compatibility easy to maintain between different versions of the program.

File formats that can be used for importing of data include DICOM, ECAT 6.4/7.2, Acr/-

Nema 2.0, Analyze (SPM), InterFile 3.3, and Concorde. Support for these file formats is achieved primarily through the use of the (X)MedCon image conversion library [158]. Additionally, the DCMTK (DICOM Toolkit) libraries [104] are used to allow importing of clinical DICOM data files. Raw data importation in big, little, and PDP endian formats is handled natively for both integer and floating point data.

Volume rendering in AMIDE is performed using the VolPack [106] volume rendering library, which accelerates rendering using a shear-warp factorization algorithm. This software library based approach is portable and provides for true volume rendering capability, as opposed to the surface rendering approaches provided by many libraries and hardware accelerators. Series of rendered images, along with series of slices ("fly throughs"), can be encoded into MPEG-1 video files using the fame MPEG encoding library (fame.sf.net).

Rigid body registration is implemented inside of AMIDE through the use of fiducial reference markers and the Procrustes rigid body alignment algorithm without scaling [80]. Briefly, the transform needed to minimize the least squares error between a set of fixed and a set of movable fiducial marks is calculated. This transform is then applied to the coordinate space of the data set to be aligned.

Filtering is implemented using a "wizard" interface. Currently, Gaussian and median filters have been implemented, although any finite impulse response (FIR) filter would be a trivial extension. FIR filters are implemented using an overlap+add method with a 64^3 point fast Fourier transform. Median filters are of variable kernel size, and can be run as separable 1D or a single 3D filter. In the interest of algorithmic simplicity, spatial coherence is ignored and the median filter is implemented using a partial sort median finding algorithm [100].

5.1.3 Validation

The validation animal data set, consisting of PET and CT scans, was acquired as follows: A nude mouse (Charles River Laboratories), anesthetized with pentobarbital, was injected with 200 μCi [^{18}F]-fluoro 2-deoxy-glucose ([^{18}F]FDG). One hour was allowed for tracer uptake and clearance. The mouse was then placed on a plastic bed and 4 fiducial reference markers were affixed. Fiducial markers consisted of 200 μl PCR tubes containing 1 μCi of [^{18}F]FDG and 10 μl Omnipaque (iohexal) nonionic iodinated contrast solution. The mouse was scanned on a microPET scanner built at UCLA [34] using 7 bed positions at 4 minutes/bed. Immediately after, a two bed position CT scan with 196 views/bed was acquired using an ImTek microCAT scanner [170] with the X-ray tube at 50 KVp, 300 mA, and 1.0 mm Aluminum filtration. All

	AMIDE	CAPP	Mediman	MRIcro	CRIIISP
cylinder	0.46 ± 0.093	0.46 ± 0.097	0.46 ± 0.10	0.47 ± 0.084	0.47 ± 0.096
heart	5.5 ± 0.73	5.9 ± 0.55	5.6 ± 0.83	5.2 ± 0.58	5.2 ± 1.0
brain	1.9 ± 0.19	2.0 ± 0.10	1.9 ± 0.18	1.8 ± 0.18	1.9 ± 0.13
bladder	48 ± 14	50 ± 11	45 ± 14	45 ± 11	47 ± 14

Table 5.2: ROI statistics generated for similarly placed and sized ROI's using five different image analysis programs. The first entry is from a calibration cylinder, with the remaining three entries from an [^{18}F]FDG PET Scan. Values expressed as 1000 * Mean Image Units$\pm$SD.

animal care and euthanasia was performed with the approval of the University of California Animal Research Committee. The validation cylinder consisted of a 37 mm diameter polysulfone cylinder filled with 262 μCi of [^{18}F]FDG in 70 ml water, and was scanned in a single bed position for 4 hours.

MicroPET scans were reconstructed using the MAP reconstruction algorithm [175] with a beta value of 0.5, and multiple beds were combined into a single image. MicroCAT scans were reconstructed using the company supplied 3D-filter back projection reconstruction software. ImTek's file format was converted to ECAT 6.4 format using the imtk_conv program supplied with (X)MedCon, and the two beds were combined into a single image. The resolutions of the data sets were 1.5 mm and 0.4 mm for the PET and CT, respectively.

After loading the data sets into AMIDE, the four fiducial reference markers were located for each of the two scans, and the data sets were aligned using a rigid body alignment. Results of this alignment are shown in Figures 5.3 and 5.4. The calculated fiducial reference error for this alignment was 0.2 mm/reference point.

For validation of the ROI statistics, similar ROI's were drawn in AMIDE, CTI's Clinical Applications Programming Package (CAPP), Mediman [44], MRIcro, and CRIIISP, an IDL based image package developed previously in our laboratory. The results are shown in Table 5.2. The values generated by AMIDE were not significantly different from any of the other packages when compared using a two tailed paired t-test at a significance level of p<0.2.

5.1.4 Availability/Requirements

All required source code and installation instructions can be found through the AMIDE web site (amide.sf.net), along with binaries for Linux, Macintosh OS X [Apple Computer, Inc., Cupertino, CA], and Microsoft Windows [Microsoft Corporation, Redmond, WA] operating systems. The Macintosh version relies on the freely available Fink (fink.sf.net) add-on distribu-

Function	Time (s)
Extract transverse slice of data	0.01
Extract coronal/sagittal slice of data	0.04
Calculate statistics for 7500 voxel ROI	0.3/frame
Initial setup for volume rendering	4.8
Volume rendering a data set	0.06

Table 5.3: Times needed for performing various functions in AMIDE. These functions were performed utilizing a 128x128x159 PET image on a 750 MHz i686 class computer with 384 MB RAM.

tion.

There are no strict hardware requirements for running AMIDE, but the computer will need to have at least as much memory available as the cumulative size of the loaded data sets. Additionally, since the continuous reslicing approach is computationally expensive, a modern processor ($\geq$500 MHz) is recommended when working with larger (512^3) data sets. Representative computational times needed for various functions are shown in Table 5.3.

5.1.5 Walk-through

Figure 5.2 points out the most important elements of the AMIDE user interface. A brief walk-through follows to explain how to use the program.

1. On program startup, the user is presented with a blank study, and can then import the requisite medical images into the program. Most medical image formats are correctly detected by the program (e.g. DICOM, Analyze, Concorde, ECAT 7, InterFile). If the format cannot be determined, the user is allowed to explicitly specify the import file type. After importing data, the entire study can be saved in AMIDE's native file format for direct loading in subsequent sessions.

2. The data sets will now appear as objects in the study's tree. Left clicking on an object will select it for appearance in the three orthogonal views, and a check will appear in the corresponding checkbox. Right clicking on any object in the tree will bring up a dialog for changing parameters relevant to that object, such as voxel sizes, scale factors, and thresholds. Note that one of the data sets in Figure 5.2 is highlighted, which indicates that this data set is the "active" object. When operations are performed that can logically apply to only one data set, the active object is the one chosen. For instance, pressing the
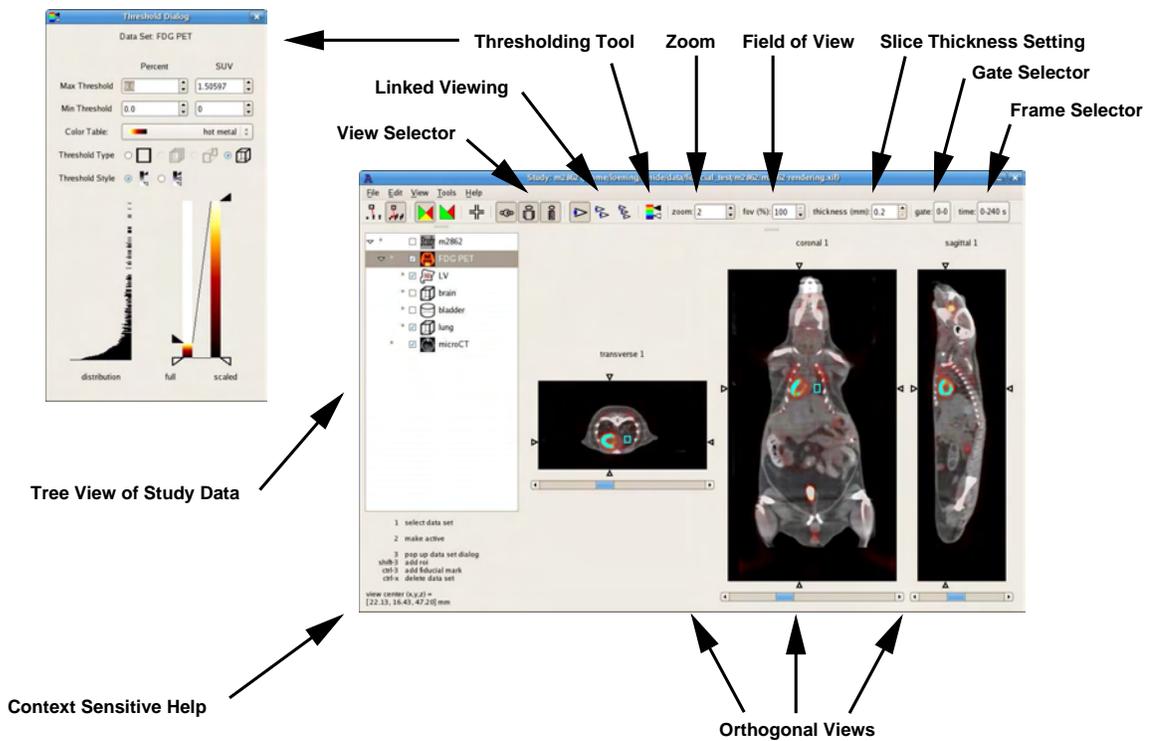
Figure 5.2: Salient user interface elements of AMIDE. A standard AMIDE session is shown, with the most important elements of the user interface labeled.

thresholding tool button will bring up the thresholding dialog for the active object. The middle mouse button can be used for switching the active object.

3. Context sensitive help is displayed in the lower left corner of the application to explain what the different mouse buttons and key strokes will do at any given point. Complete documentation is also available from the help menu.

4. Moving through the data set is accomplished by directly clicking on any of the orthogonal views. For instance, clicking on the transverse view will update the coronal and sagittal views to correspond to the chosen point. In Figure 5.2, all three orthogonal views are shown, although the viewer can select fewer views by toggling the view selector buttons on the toolbar. The active data set can be shifted or rotated with respect to the other objects in the study by using the shift key together with the left or middle mouse buttons, respectively.

5. The zoom and the thickness of the viewed slices can be altered by using the corresponding entries on the toolbar. For dynamic studies, clicking the "frame selector" button will pop-up a dialog for frame selection purposes. Also on the toolbar is a set of toggle buttons for switching between single and multiple (linked) cursor mode viewing. An example of three cursor mode is shown in Figure 5.3.

6. ROI's are added to the study either from the menu, or by directly clicking in the tree. After an ROI has been added to the tree, the next mouse click on any of the views will initiate the process of drawing an ROI, with the left button initiating edge-to-edge drawing and the middle button initiating center-out drawing. Subsequent modifications of the ROI can be done by clicking on the ROI in any of the views. Shifting, rotating, and scaling are accomplished by the left, middle, and right mouse buttons, respectively. Statistics for ROI's are generated by selecting "ROI Statistics" underneath the tools menu.

7. Underneath the view menu are options for generating series of slices and volume renderings of the currently selected data sets. Both of these options will pop up separate windows for the corresponding purpose. Series of slices can be displayed over space or time. From the volume rendering dialog, animated movies can be generated and saved as MPEG-1 files. An example of a stereoscopically rendered fusion data set generated inside AMIDE is shown in Figure 5.4.
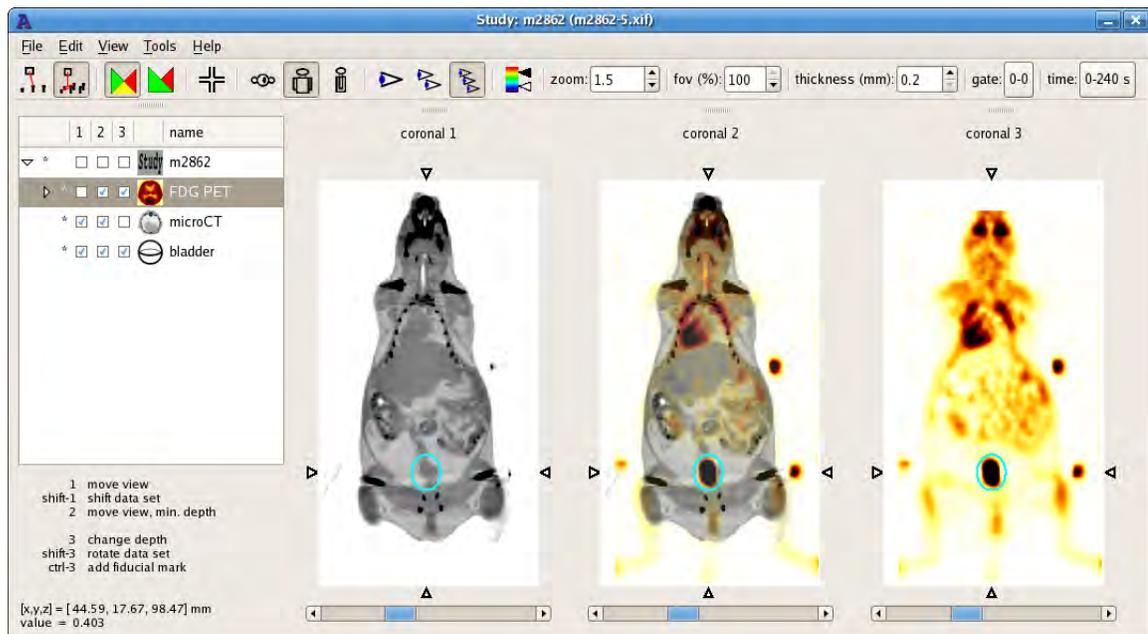
Figure 5.3: Main window of AMIDE shown in three cursor mode with two aligned data sets loaded and displayed on coronal views. The right view shows an [^{18}F]FDG image of a mouse centered around the left ventricle in an inverted hot metal color scale. The left view displays the corresponding CT scan in gray scale. The center view shows the fused images. The heart is on the left because the animal was scanned in the prone position. The spots of activity exterior to the mouse are fiducial reference markers. The PET data set has a voxel size of 0.4x0.4x0.7 mm, while the CT data set has a voxel size of 0.2x0.2x0.2 mm.
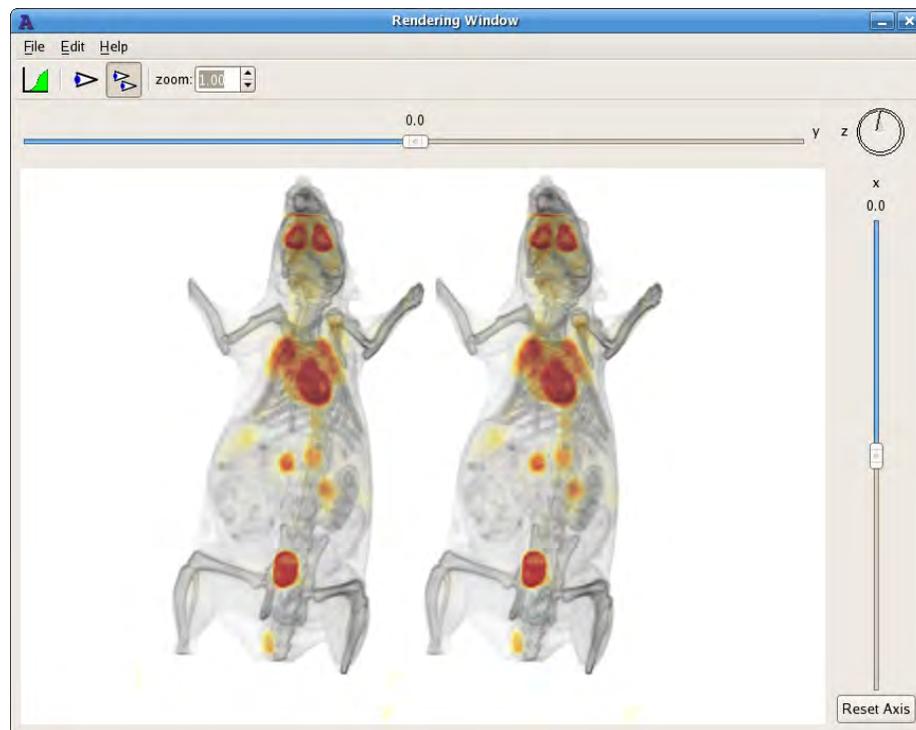
Figure 5.4: Example of fused data sets rendered stereoscopically by AMIDE using the VolPack volume rendering library. The CT scan is shown in gray scale with the [^{18}F]FDG PET scan shown in inverse hot metal.

8. Entries under the tools menu start wizards for various functions such as filtering, rigid body alignment, cropping of data sets, and fly through movie generation. Examples of both fly through and rendered movies can be found at the AMIDE web site (amide.sf.net/-output.html).

5.2 Discussion

With the increasing prevalence of multimodality imaging in the research community, a need has arisen for new, more sophisticated software tools that can handle and analyze these increasingly complex data sets. AMIDE provides such a tool, and AMIDE's capability for manipulating multiple, non-orthogonal data sets will become increasingly critical as multimodality image analysis becomes more common.

While a handful of proprietary tools exist that provide relatively comparable feature sets, to our knowledge, AMIDE is the only freely available and open source software package in its class. Since the source code is available, researchers are not only free to use the program but can also study and expand upon the program as fits their needs and interests. Furthermore, as the code is unencumbered by restrictive licensing users are free to redistribute both the code and any modifications made to it, although modified versions must be appropriately marked as such.

For the novice user, additional advantages of AMIDE are the simplified interface and unit handling. Continuous beta testing and feedback over the last two years from three basic science researchers with minimal imaging experience has been incorporated into the development of the user interface in order to make program interaction as intuitive as possible. Units in the program are, whenever practicable, specified in terms of real world values and the underlying digital representation of the data is, to a great extent, divorced from the user. For instance, in the slices viewed from the data set, the thickness is not restricted to integer multiples of the voxel size. As another example, given the correct conversion constants the program can present data and statistics to the user directly in terms of %ID/g or SUV's.

The continuous reslicing approach adapted by AMIDE has proven itself to be flexible from a development aspect and crucial for the arbitrary image fusion abilities of the package. It makes movement, scaling, and rotation of data sets essentially free from a computational standpoint while avoiding destructive interpolation of the original data set. The trade-off is that the computational expense of slice generation is greatly increased compared to an orthogonal slice

based approach. In practice, it has been found that modern processors ($\geq$500 MHz) are powerful enough that the added computational expense of this approach does not impact the user experience for standard sized data sets ($\leq 512^3$)

AMIDE now encompasses the core set of features needed for bringing multi-modality medical image analysis to the molecular imaging research community. Further work is shifting towards extending upon these core facilities, particularly in providing interactive "wizard" interfaces for making advanced medical imaging algorithms (e.g. factor analysis, cardiac polar maps) more accessible to the casual research user. It is hoped that not only will the package be a valuable addition to the molecular and medical imaging software toolkit, but that other research groups will seize upon the availability and extensibility of the package's source code, and choose AMIDE as a platform upon which their ideas and algorithms can be readily disseminated to the molecular and medical imaging research community as a whole.

Chapter 6

A Bioluminescent Imaging Probe for EGFR

The overlying element that remained for this dissertation, was an actual demonstration that a bioluminescently labeled imaging probe could work in an *in vivo* imaging context. The approach taken here was to create a genetic fusion of epidermal growth factor (EGF) and a luciferase, and to use the resulting protein as an imaging probe to study epidermal growth factor receptor (EGFR). The rationale behind the choice of the EGF/EGFR system is explained in Section 1.5. In hind sight, targeting a receptor using its native ligand is a poor choice for the reasons outlined in the conclusion section of this chapter. The choice of EGF/EGFR as the model system is particularly problematic, due to the lack of protein ligand antagonists for this receptor system.

For an example of a working bioluminescently tagged imaging probe, the interested reader is encouraged to skip ahead to Section 7.2. The work presented in this chapter does clearly illustrate one major difference between using a bioluminescent label versus a radioisotope or a small molecule fluorophore. Namely, that the bioluminescent label is a protein, and is therefore subject to degradative processes that do not effect fluorophores and radioisotopes.

Mature human EGF[1] (also known as β-Urogastrone), is a 6.2 kDa, 53 amino acid protein that is formed by proteolytic cleavage from a large membrane bound precursor [31]. It induces proliferative effects on epithelial cells through binding of its intended receptor, EGFR (also known as ErbB-1 or HER). EGFR, in turn, is known to be overexpressed in a large number of

127

cancers [150].

EGFR is a membrane bound tyrosine kinase and a member of the ErbB family of receptors. Binding of EGF to the extracellular domain of EGFR induces dimerization, with resultant autophosphorylation of the intracellular component of the receptor along with phosphorylation of other intracellular substrates [94]. The combination of dimerization and ligand binding also leads to an order of magnitude acceleration in clathrin mediated receptor internalization [234] as well as increasing the fraction of internalized activated receptor that traffics from the early endosome to the late endosome and subsequent degradation in the lysosome. The fraction of activated EGFR that is sorted to the lysosomes varies between 20-80% [62], with the remaining EGFR recycling to the cell membrane where it is again available to undergo internalization.

The exact nature of EGF/EGFR dimerization has been debated in the field. Many other protein ligands (e. g. vascular endothelial growth factor, interferon-γ) first dimerize, and it is the bivalent binding of the dimerized ligand that brings together two receptors leading to activation of the receptor's cytoplasmic component. Current evidence for EGF, however, conclusively indicates that the dimerization interface is on the receptor, and that binding of EGF to EGFR induces a conformational change so as to allow this interface to dimerize with a similarly EGF bound EGFR [130, 66, 162, 116]. EGFR can also heterodimerize with the other receptors in the ErbB family. Interestingly, ErbB-2 (also known as HER2, or neu) does not bind a ligand, and is in fact constitutively posed for dimerization [65].

The crystal structure of the EGF/EGFR complex was examined to ascertain the most appropriate orientation of an EGF-luciferase fusion. The complex structure (Figure 6.1) would seem to indicate that either the N or C terminus of the peptide would be amenable for fusing to another protein. This is consistent with a previous report [244] that found EGF tolerated both N and C-terminal fusions, with or without linkers, without any impact on binding to EGFR.

An important element to keep in mind when expressing EGF or an EGF containing fusion protein, is that the mature 53 amino acids have 3 disulfide bonds that need to be linked correctly in order to produce an active ligand. Recombinant EGF produced in *E. coli* is generally expressed in inclusion bodies within the reducing environment of the bacterial cytoplasm [56, 131], purified, and then refolded to obtain the active disulfide bond containing ligand. For the purposes of producing an RLuc-EGF fusion, however, cytoplasmic expression in inclusion bodies was not pursued as a successful RLuc refolding protocol has not been developed despite many attempts to do such (personal communication, Dr. Bruce Bryan).

---

[1]In this manuscript, "EGF" will always refer to the mature human protein, unless otherwise indicated.
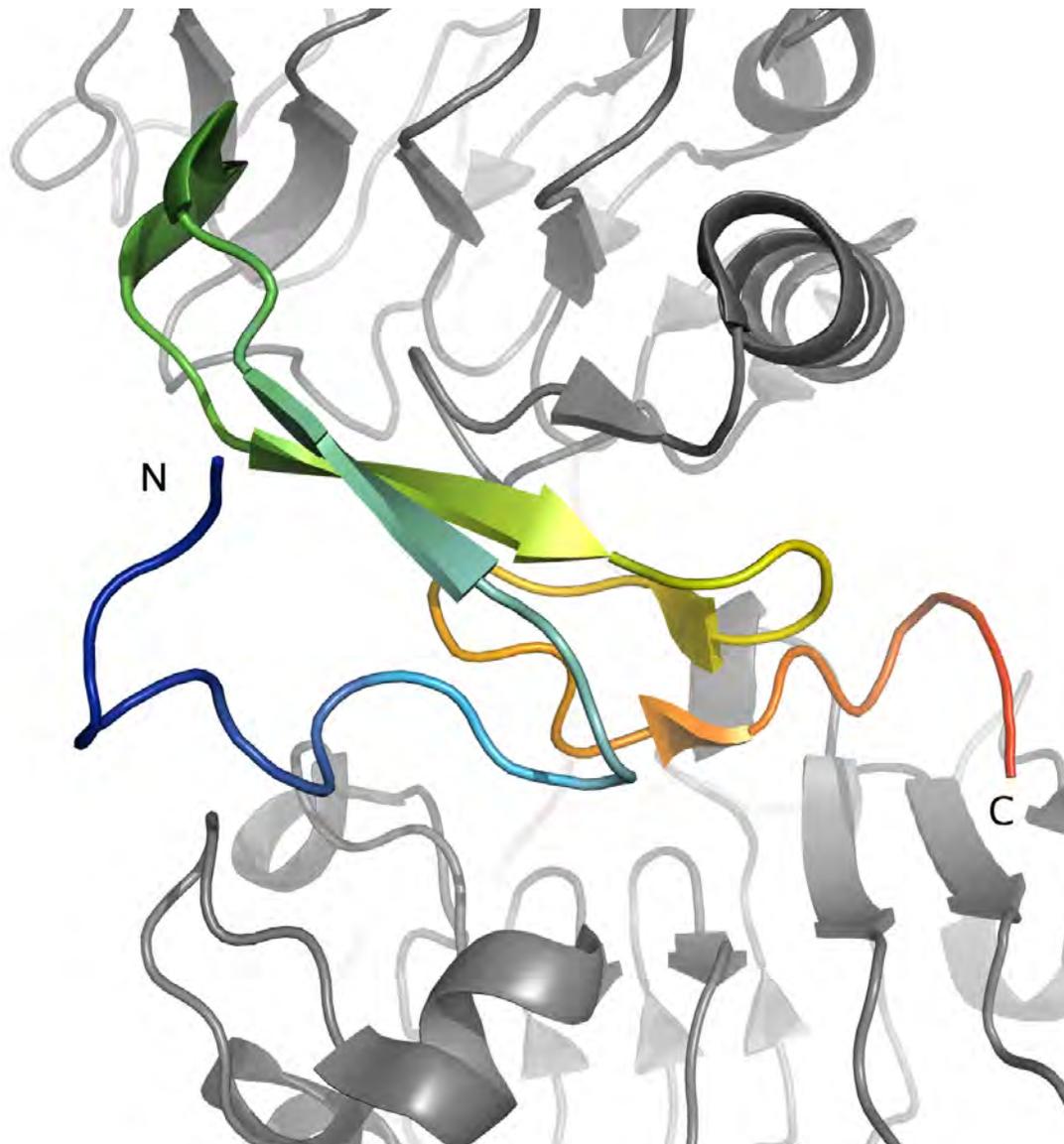
Figure 6.1: Crystal structure of mature human EGF and the extracellular domain of human EGFR. EGFR is shown in gray. EGF is in the center and shown in a rainbow color scheme with the N-terminus (N) in blue and the C-terminus (C) in red. The data set is PDB ID: 1ivo [162]. The first 4 residues on the N-terminal and the last 2 on the C-terminal of EGF were not resolved in this structure.

As an alternative production strategy, expression into the bacterial periplasm was pursued. The periplasm more readily allows for the correct formation of disulfide bonds due to its oxidative environment [12] and the presence of the DsB (Disulfide-Bond) machinery (Dsb is a family of proteins that catalyzes the formation of disulfide bonds as well as rearranges existing ones [39]).

Expression into the periplasm can be accomplished by making use of any one of a number of signal peptides that have been identified through their ability to make the protein to which they are attached to traffic to the bacterial periplasm. The most studied signal peptides (e. g. PelB, OmpA, PhoA) share many structural features and lead to export of the protein through the energy dependent Sec pathway [39]. These signal peptides are N-terminal sequences, generally 20-30 amino acids in length, that lead to the binding of SecB and other chaperone proteins. The chaperones are involved in directing the translated protein toward the translocation machinery, as well as maintaining the protein in an unfolded state prior to translocation across the inner cell membrane [47, 39]. During entry into the periplasm, the signal peptide is cleaved from the protein.

Note that the bacterial Sec system exports fully translated protein. This is in contrast to the generic eukaryotic secretion system, where binding of the translated signal peptide by the appropriate binding partner (Signal Recognition Particle) halts translation of the remaining protein. The complex is then situated such that translation can continue directly into the endoplasmic reticulum.

In addition to a more appropriate environment for disulfide bond formation, periplasmic expression has some additional advantages over cytoplasmic expression [13]. 1) Bacterial proteins begin with a formylmethionine, so periplasmic expression can be used to obtain protein with an authentic N-terminus due to the cleavage of the N-terminal signal peptide. 2) The periplasm has fewer proteases than the cytoplasm [39]. 3) Only ∼4% of total cellular protein is found in the periplasm [159], so the use of an initial periplasmic separation step leads to a substantially purer starting point than whole cell lysis.

Several disadvantages are also present with periplasmic expression. The use of a periplasmic fractionation step leads to a large and dilute volume to begin purification with, although this is not so detrimental when the protein in question contains an affinity tag. A more important limitation is that, besides the signal peptide, poorly understood structural features of the protein itself are involved in the translocation across the membrane [35, 131]. Many proteins will not express at satisfactory levels in the bacterial periplasm. For proteins that do

express well into the periplasm, only a couple mutations can drastically alter the efficiency of this secretion (as seen for the A2T mutations in Section 2.2.5, and for the surface mutations in Section 4.2.4).

Recombinant EGF has previously been expressed periplasmically in bacteria using the OmpA signal peptide [244, 213]. Although effectiveness of secretion into the periplasm can depend on the combination of signal peptide and protein [18], the PelB signal peptide was chosen for use here simply because we were more familiar with it.

There has been at least one previous report in the literature that tried to use EGF ligand as an imaging agent [180]. In that study, the authors used ^{111}In labeled EGF, and found that radiolabeled anti-EGFR monoclonal antibodies were a more effective approach for imaging EGFR. These results were probably due to the short 8 minute circulation half-life of EGF [32] that would have severely limited the amount of probe available to accumulate in the tumor of their animal model. In the approach here, the RLuc-EGF fusion imaging probe will be much larger due to the presence of the luciferase moiety, and circulation times should be long enough for adequate accumulation in the tumor to take place.

This chapter on generating an imaging probe for EGFR will progress in a roughly chronological order. It begins with the first generation luciferase/EGF fusion proteins based on RLuc, and switches to the use of variants more resistant to inactivation as these luciferases became available. The second generation consisted of RLuc/C124A fusions [119], and the third generation were fusions based on RLuc8 [126]. Within each generation, an iteration of protein production, *in vitro* validation, and the occasional *in vivo* study is presented.

6.1 Methods

Several of the methods used in this chapter, such as the luciferase assay and western blotting, are identical to what has already been described in the methods section of Chapter 2. These methods are not repeated here.

When values in terms of RLUs are used below, they refer to the RLUs of a Turner 20/20. Conversion factors between RLUs and absolute values are given in Appendix A.

6.1.1 Materials

Recombinant human EGF was from PreproTech (Rocky Hill, NJ). This EGF was reconstituted at 1 mg/ml in H_2O and stored in small aliquots at -20°C prior to use. Bafilomycin A_1 and

Phenylarsine Oxide (PAO) were from Sigma. Bafilomycin A$_1$ was dissolved as a 50 μM stock in DMSO and stored at -20°C. PAO was dissolved as a 25 mM stock in DMSO and stored at -20°C.

6.1.2 Constructs

The constructs shown in Figure 6.2 were assembled via PCR cloning. The plasmid CMV-hRL (Promega) was used as the initial template, and the DNA sequence encoding the 53 amino acids of mature human EGF was extended onto either the 5' or 3' end of the *hrluc* gene with the inclusion of single glycine spacers. The pelB leader sequence was then appended to the 5' end to provide a signal for protein export to the bacterial periplasm. The final fusion genes were cloned into the pBAD/Myc-His A plasmid (Invitrogen), using the NcoI/HindIII sites, such that a Myc epitope and a 6 histidine tag (6xHis) were attached to the 3' end of the fusion construct. Constructs with a C124A mutation in the luciferase sequence were created using a QuikChange mutagenesis kit (Stratagene).

To make RLuc8 based constructs, PCR was used to add an MfeI site to the 3' end of *rluc8*. PCR was also used to add an MfeI site to the 5' end and a SalI site to the 3' end of *egf*. The plasmid pBAD-pelB-RLuc8-L-EGF was then constructed by three-way ligation. The plasmid pBAD-pelB-RLuc8-L(G$_4$S)$_3$-EGF was made similarly, but with a primer to attach the (G$_4$S)$_3$ linker to the 5' end of EGF. Note that in referring to the proteins expressed from the various plasmids, the names of spacers have been dropped (e. g. RLuc8-EGF was produced from the pBAD-pelB-RLuc8-L-EGF plasmid).

The plasmid pET15b_DT_51E,148R, which contains the full length Diphtheria toxin gene with a disabled catalytic domain due to K51E and E148R mutations [64], was graciously provided by Dr. R. J. Collier (Harvard Medical School, Boston, MA). The translocation domain (TDT: residues 201 to 378) and the arginine-rich loop with the TDT (RL-TDT: residues 186 to 378) were obtained by PCR from this plasmid with appropriate primers to add 5' MfeI sites, and 3' XmaI or SalI sites. EGF had XmaI attached to its 5' end via PCR. Appropriate restriction digests and ligations were then done to make the plasmids pBAD-pelB-RLuc8-L-RL-TDT-G-EGF, pBAD-pelB-RLuc8-L-TDT-G-EGF, pBAD-pelB-RLuc8-L-RL-TDT, and pBAD-pelB-RLuc8-L-TDT.

A pET32 version of RLuc8-EGF was made by using PCR to replace the second and third codon of *rluc8* (arginine-serine) with a BamHI site (glycine-serine), and add a HindIII site 3' to the fusion gene's stop codon. The PCR product was then BamHI/HindIII digested, and ligated

	AA's	Size (kDa)
RLuc	335	38.8
EGF-RLuc	389	45.0
RLuc-EGF	390	45.1
RLuc8-EGF	372	43.3
RLuc8-TDT-EGF	550	62.6
RLuc8-RL-TDT-EGF	566	64.2
S3RLuc8-EGF	371	43.2

Table 6.1: Length and predicted sizes for several of the luciferase/EGF fusion constructs. The values are calculated assuming the signal peptides (or cleavable thioredoxin fusion domain, in the case of S3RLuc8-EGF) have been removed.

to the backbone fragment of a BamHI/HindIII digested pET32-trx-6xHis-s-thr-tGLuc plasmid (from Section C.1.3). The final plasmid was entitled pET32-trx-6xHis-s-thr-S3RLuc8-L-EGF-6xHis, where trx indicates a thioredoxin gene, s indicates an s affinity tag that was unused here, and thr indicates a thrombin protease site. Note that in addition to the C-terminal 6xHis tag, a redundant 6xHis sequence is located in the N-terminal portion of the construct. A KDEL version of this construct was made as above, by using a 3' primer in the PCR step that included codons coding for the amino acid sequence KDEL immediately preceding the stop codon.

All constructs were confirmed by sequencing. The predicted molecular weights for several of these constructs, after removal of the pelB leader sequence, are as given in Table 6.1.

6.1.3 Protein Production and Purification

Protein production and nickel affinity purification of the periplasmically expressed constructs were done almost identically to as described in Section 2.1.5. The differences for the RLuc and RLuc/C124A based constructs were that following induction culture times were 2 h, Luria-Bertani Broth (LB) was used instead of TB, and purified proteins were stored for later use in 20% glycerol at -80°C. The difference for the RLuc8 based constructs was that culture times following induction were 6 h.

Protein production and nickel affinity chromatography of the cytoplasmically expressed pET32 based constructs was performed similarly to what is described in Section C.1.3. The only difference was that instead of being digested off the nickel affinity chromatography medium with calf α-thrombin, the protein was first eluted and then digested. Anion exchange chromatography was then done using an 8 ml Source 15Q column to separate the S3RLuc8-EGF, the trx-6xHis-s-thr fragment, and the thrombin. The gradiant for anion exchange was from
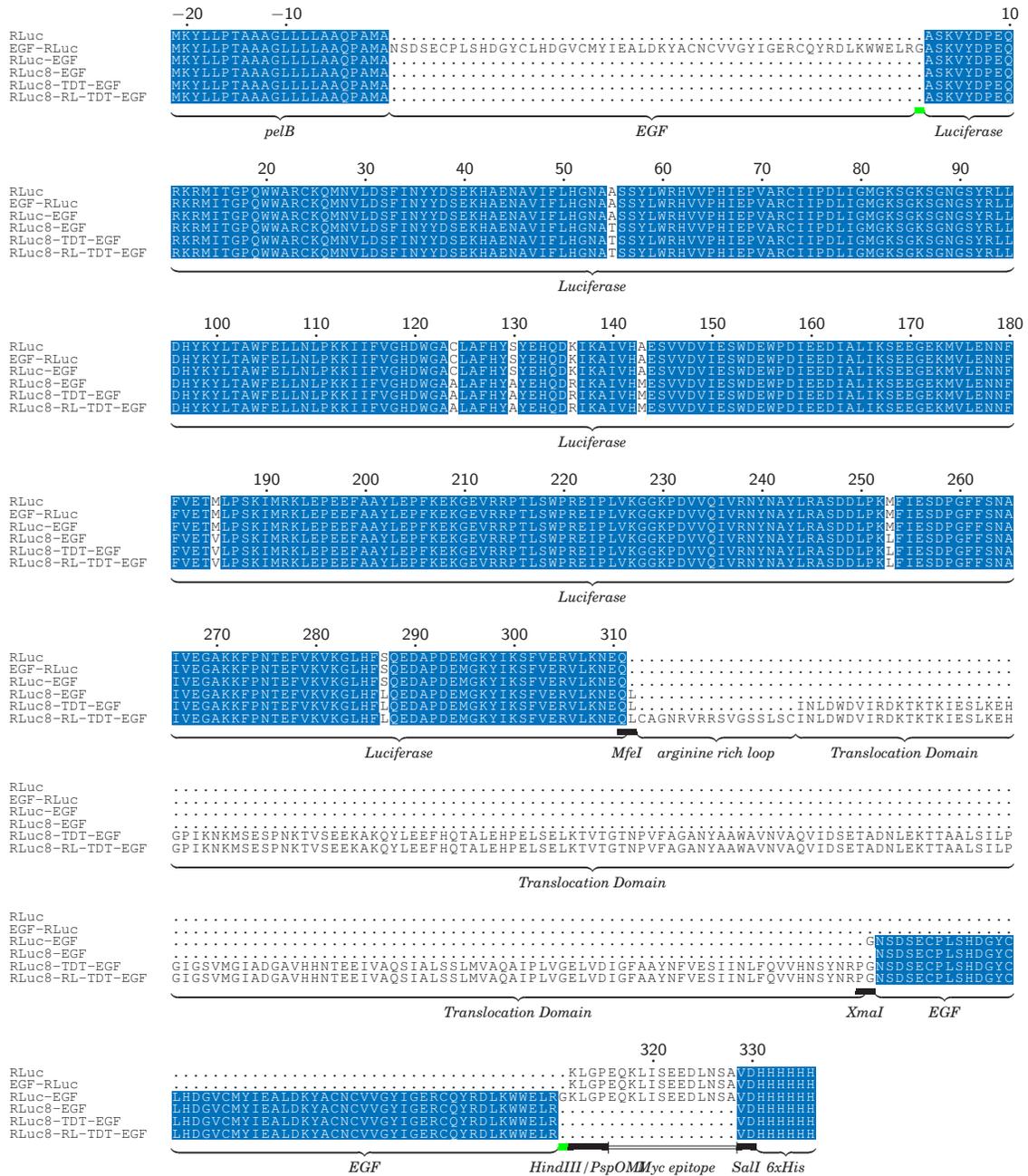
Figure 6.2: Protein sequences for several of the EGF/luciferase fusion constructs. The pelB leader sequence is cleaved during export into the periplasm and is absent in the mature protein. The RLuc/C124A based constructs discussed later are identical to the RLuc constructs shown here with the exception of the C124A mutation in the luciferase. TDT stands indicates the translocation domain of Diphtheria toxin. RL indicates the arginine rich loop of Diphtheria toxin. HindIII/PspOMI, MfeI, SalI, and XmaI are amino acids corresponding to restriction enzyme sites. Green boxes indicate additional spacers. Numbering is with respect to RLuc.

10 mM to 250 mM NaCl in buffer containing 30 mM Tris pH 8.0, with the S3RLuc8-EGF protein eluting around 200 mM NaCl.

Inactivation rates for the RLuc and RLuc/C124A based constructs were assessed by adding 5 μl of purified protein to 25 μl of freshly harvested mouse serum and incubating in a 37° C water bath. 1 μl samples were taken at intervals over the course of the experiment and measured for luciferase activity. Inactivation in serum for the RLuc8 constructs was assayed identically to the protocol already described in Section 2.1.6.

All protein intended for use in animals was desalted into phosphate buffered saline (PBS) using PD-10 columns (GE Healthcare Life Sciences).

6.1.4 Cell Culture

A431 human epidermoid carcinoma cells (ATCC Number CRL-1555) were obtained directly from ATCC (Manassas, VA) to use as high EGFR expressing target cells. These cells vastly overexpress EGFR at levels on the order of 2×10^6 receptors/cell [102]. NIH 3T3 murine fibroblasts were obtained from the laboratory of Dr. Harvey Herschman (UCLA) as low EGFR expressing controls. These cells have $\sim$3000 EGF receptors/cell [52, 167].

Both cell types were cultured in high glucose Dulbecco's modified Eagle medium (DMEM) with pyridoxine HCl, L-glutamine, 110 mg/L Na-pyruvate, and supplemented with 10% fetal bovine serum (FBS), 100 U/ml penicillin-G, and 100 μg/ml streptomycin. The cells were cultured in a 5% CO_2 environment at 37°C. For pulse-chase experiments, medium without phenol red was used.

For all cell culture experiments, cells were plated at 1×10^5/well in 24-well plates and allowed 24 h to attach before use. 24 h corresponds to approximately one doubling time for A431 cells [144].

When needed, conversions between luminometer units (RLUs) and the IVIS imaging systems units (photons/s/cm^2/steradian) were made using the calibration factors described in Section A.2.4 of Appendix A.

Cell Binding

Cell binding experiments for RLuc and RLuc/C124A based fusion proteins were performed on both A431 and NIH 3T3 cells. 8000 RLUs/well of either fusion or control protein was applied in 500 μl/well of medium without FBS. Following an incubation (0-320 minutes), the medium was aspirated, the cells washed twice with PBS, and the wells refilled with 500 μl/well room

temperature PBS. 0.5 μg/well coelenterazine was added and the plate was imaged for 60 s in an IVIS 100 imaging system (Xenogen, Alameda CA).

Competitive Binding

Competitive binding experiments for RLuc and RLuc/C124A based fusion proteins were performed on A431 cells using recombinant human epidermal growth factor as the competitor. For competitive binding, 8000 RLUs/well of fusion protein was applied in 500 μl/well of medium without FBS along with escalating amounts of recombinant human EGF. Following a 20 min (RLuc based fusions) or 40 min (RLuc/C124A based fusions) incubation at 37°C, the medium was aspirated and the cells washed and imaged as above.

Pulse-Chase

For pulse-chase experiments, cells were exposed to either fusion or control protein (100,000 RLUs/well) for 60 min in medium without phenol red and with either 10% FBS or 1% human serum albumin (HSA). Following the 60 min pulse, the medium was aspirated and replaced with fresh medium without protein for the chase. After the variable time chase interval (0 minutes to 2 days), the medium was again replaced with fresh, 37°C medium, 0.5 μg of coelenterazine/well was added, and the plate was imaged for 60 s using an IVIS 50 imaging system.

A control plate containing medium with protein but without cells was set up in parallel and stored at 37°C. Following the 1 hr pulse, 0.5 μg of coelenterazine/well was added and the plate was imaged for 10 s. The values from this control plate were then used to normalize the values from the experimental plates.

In experiments using bafilomycin A$_1$, the final concentration used was 200 nM. In experiments using PAO, the final concentration was 100 μM.

6.1.5 Animal Models

All animal work at UCLA was approved by UCLA's Animal Research Committee (ARC). All animal work at Stanford was approved by Stanford University's Administrative Panel on Laboratory Animal Care (APLAC).

Tumor models were constructed by injecting 1×10^6 A431 cells subcutaneously into the right shoulder of athymic (nu/nu) mice. The tumors were allowed to grow until palpable ($\sim$10 days)

prior to use. In some cases, 1 million NIH 3T3 cells were similarly injected into the contralateral shoulder with the intention of creating a control tumor. The NIH 3T3 cells however never grew into tumors, so no mention of them are made in the results.

For imaging, either fusion or control protein was injected *intravenously* (*i. v.*) using the tail vein. Scans were then performed with the mouse in the prone position at various time points following the injection of protein, with an *i. v.* injection of coelenterazine into the tail vein contralateral to the one used for protein injection immediately preceding each scan. The injected coelenterazine consisted of 10 μl of 1 μg/μl coelenterazine in ethanol mixed with 90 μl of PBS. The acquisition time of the scans was 60 s. In some instances, after acquisition of the prone image, the mouse was flipped over and a supine image was acquired. These supine images were acquired without the injection of an additional dose of coelenterazine. On the following day, the imaging protocol was repeated with a different protein.

6.2   Results and Discussion - EGF Fusions Using RLuc

6.2.1   Protein Production

Several periplasm fractionation protocols were tested, including protocols from Qiagen [176], Epicentre (PeriPreps Periplasting Kit), and Ames *et al.* [3]. In the end, a slightly modified osmotic shock protocol based on the publication from Neu and Heppel [156] (described in Section 2.1.5) gave the best results (data not shown) and was used for the work presented here. A comparison of protein production at 32°C and 37°C showed that the cooler temperature yielded 5-fold greater activity (data not shown). 32°C was tried as it is the temperature at which RLuc is maximally active [138]. A time course study of protein expression in bacteria showed that recovered activity was greatest 2-4 h following induction of the culture (data not shown).

For RLuc, 75% of the activity was found in the periplasmic fraction following osmotic shock, with the remainder in the cell pellet. For the RLuc fusions, only 10-20% of the total activity was found in the periplasmic fraction, with the majority of the activity remaining associated with the cell pellet.

6.2.2   Cell Binding Time Course

Specificity of the EGF-RLuc and RLuc-EGF fusion proteins was assessed by comparing binding of the different proteins on A431 and NIH 3T3 cells. In these studies, the fusions proteins were

applied to the cells for a specified time. The wells were then washed, and imaged immediately. As shown in Figure 6.3a, the RLuc-EGF fusion protein performed better with respect to binding on high EGFR expressing A431 cells than the EGF-RLuc construct. A further study of RLuc-EGF (Figure 6.3b) showed that RLuc-EGF was specific for A431 cells, showing no appreciably binding to NIH 3T3 cells that express very low levels of EGFR.

The poor binding result for EGF-RLuc, is a result that will be repeated often in this chapter for constructs containing N-terminal EGF fusions. This could potentially indicate that EGF disfavors large fusions on its C-terminus. More likely, it is an artifact of having the 6xHis tag on the C-terminus of the fusion proteins. If the EGF domain of EGF-RLuc is cleaved during expression or purification, the luciferase portion of the construct will still be purified in the nickel affinity chromatography step. This will result in the "EGF-RLuc" condition actually being a mixture of intact EGF-RLuc and RLuc, with a resulting drop in the percentage of the condition that is retained on the cells in binding experiments. The RLuc-EGF construct does not suffer from this same effect, as if proteolysis occurs between its two domains, the RLuc will be lost in the Ni-NTA step and will not lead to aberrant activity in the *in vitro* assays.

6.2.3   Competitive Binding

A competitive binding experiment, as shown in Figure 6.4, further demonstrated the specificity of the RLuc-EGF fusion protein for EGFR. Some instructive calculation can be made to verify that this data meets expectations. The dissociation constants for the EGF ligand (K_D) and the RLuc-EGF ($K_{D'}$) probe can be stated as:

$$K_D = \frac{[L][R]}{[LR]} \tag{6.1}$$

$$K_{D'} = \frac{[L'][R]}{[L'R]} \tag{6.2}$$

Where $[L]$ and $[LR]$ represent the concentrations of free ligand and ligand bound to receptor, and $[L']$ and $[L'R]$ represent the concentrations of free probe and probe bound to receptor. Let $[T_L]$, $[T_{L'}]$, and $[T_R]$ be defined as the total EGF, the total probe, and the total receptor concentrations, respectively. The value of $[T_R]$ is given as the abscissa of the graph in Figure 6.4. $[T_{L'}]$ can be estimated to be on the order of 0.1 nM based on the amount of activity. Given that 100,000 cells were plated, one cell doubling time has occurred, and an A431 cell has $\sim2\times10^6$ EGF receptors [102], $[T_R]$ can be estimated as ~1 nM. In the limit of low concentrations of EGF
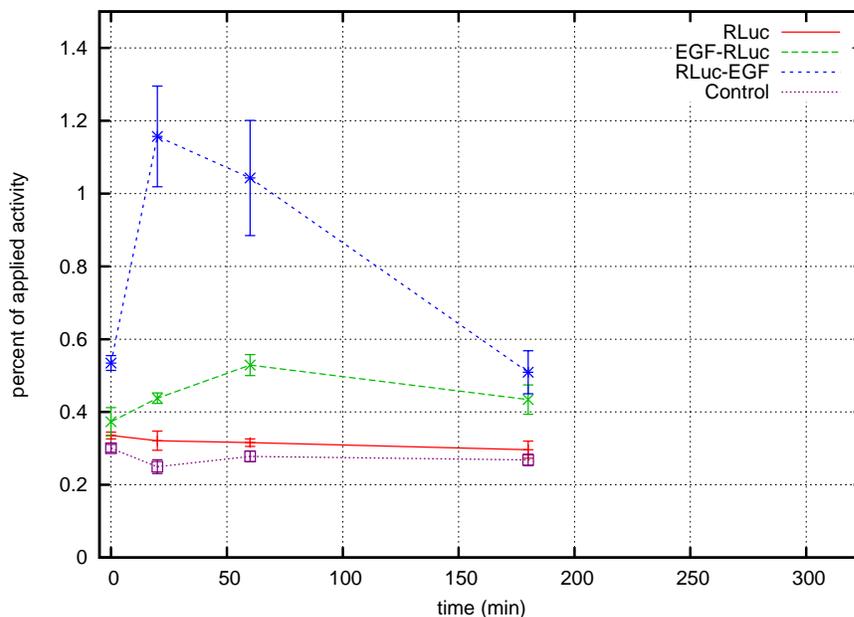
(a) EGF-RLuc and RLuc-EGF on A431 Cells
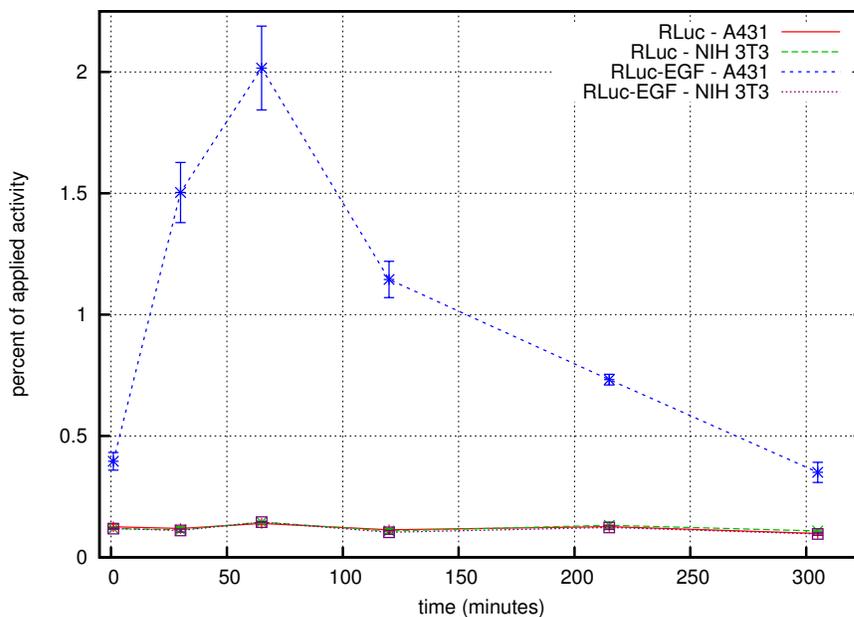
(b) RLuc-EGF on A431 and NIH 3T3 Cells

Figure 6.3: Time course studies of EGF-RLuc and RLuc-EGF fusion protein binding to A431 and NIH 3T3 cells. In Panel (a), A431 cells in 24-well plates were exposed to protein (RLuc, EGF-RLuc, RLuc-EGF) or nothing (control) for the given periods of time, washed, and imaged. In Panel (b), A431 or NIH 3T3 cells in 24-well plates were exposed to protein (RLuc or RLuc-EGF) for the given periods of time, washed, and imaged. Error bars represent standard error.

ligand, the following assumption can be made.

$$[T_R] = [R] + [LR] + [L'R] \approx [R] \tag{6.3}$$

This assumption is valid as $[LR] < [T_L] \ll [T_R]$ and $[L'R] < [T_{L'}] \ll [T_R]$. With this simplification, algebra quickly yields:

$$Fraction\ Probe\ Bound = \frac{[T_R]}{K_{D'} + [T_R]} \tag{6.4}$$

From Figure 6.4, it can be seen that the maximum percent of probe bound is ∼3%. Solving the above equation gives a value for the dissociation constant of the RLuc-EGF fusion of $K_{D'} \leq$ 35 *nM*. This value is expressed as an upper limit, because probe degradation has not been taken into account (see Section 6.2.4) and the affinity may in fact be tighter (lower value).

For the case of high EGF concentrations, the following approximations can be made:

$$[T_R] = [R] + [LR] + [L'R] \approx [R] + [LR] \tag{6.5}$$

$$[T_L] = [L] + [LR] \approx [L] \tag{6.6}$$

The first assumption is sound, as $[L'R] < [T_{L'}] \ll [T_R]$. The second assumption is weaker and can be expected to be invalid at some of the concentrations of EGF used, but the calculations here are only meant to be illustrative. The expected fraction of bound probe can be derived from Equation 6.2 and stated as:

$$Fraction\ Probe\ Bound = \frac{[T_R]}{K_{D'} + [T_R] + (K_D/K_{D'})[T_L]} \tag{6.7}$$

Assuming a previously published binding affinity of $K_D = 5$ *nM* for EGF on A431 cells [95], this equation gives an inflection point at 5 nM. This matches closely with the inflection point of ∼4 nM that can be estimated from the data in Figure 6.4. The inflection point of Equation 6.7 is fairly insensitive to the value of $K_{D'}$ if it is greater than K_D, so consideration of the curve shape yields the inequality $K_{D'} \geq 5$ *nM*. Combining this calculation with the previous limit gives 5 *nM* $\leq K_{D'} \leq$ 35 *nM*.
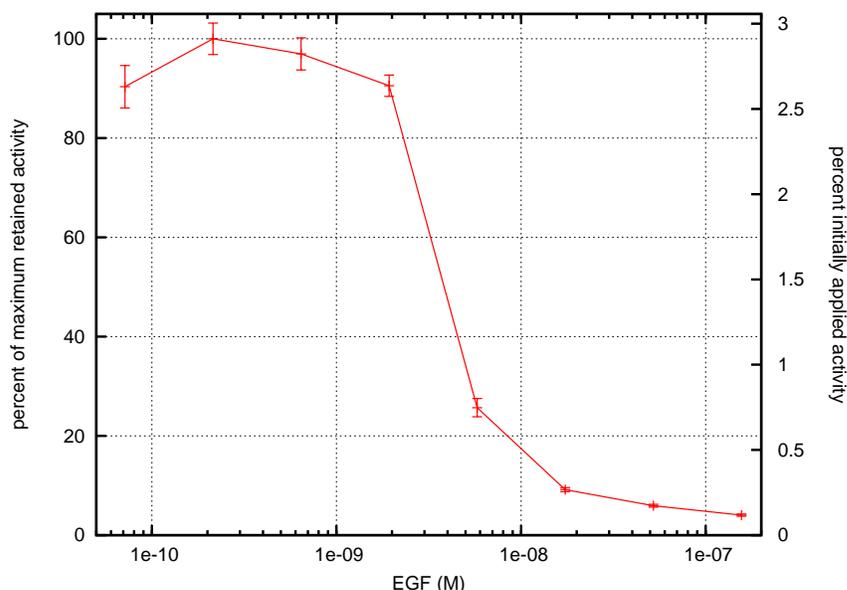
Figure 6.4: Competitive of RLuc-EGF versus EGF on A431 cells. A431 in 24-well plates were exposed to a fixed amount of RLuc-EGF and varying levels of recombinant human EGF for 20 min. The wells were then washed and imaged. Error bars represent standard error.

6.2.4 Stability under Culture Conditions

An experiment was done to evaluate how quickly RLuc-EGF degrades under the conditions used in the *in vitro* assays. The results shown in Figure 6.5a indicate that inactivation of the luciferase moiety in this probe is occurring at a very fast rate ($\tau_{1/2} \sim 20$ min), and that this inactivation is independent of whether cells are present or not.

One thought was that some of this activity loss could be due to adsorption of the luciferase onto the polystyrene of the 24-well cell culture plates with subsequent denaturation of the enzyme. To assess this, an experiment was performed to compare RLuc-EGF incubated in polystyrene and polypropylene containers, both in the presence and absence of a carrier protein (1% HSA). The results shown in Figure 6.5b indicate that some of the activity loss can be attributed to the polystyrene of the 24-well plates, and this loss can be prevented by the presence of a carrier protein. However, the majority of the activity loss could not be prevented by the use of polypropylene or the presence of a carrier protein, pointing to known issues with the stability of RLuc [120] as the primary cause of inactivation.

Interestingly, the HSA containing conditions initially showed less activity. The reason for this is not entirely clear, but it may have to do with interactions of the coelenterazine and HSA.
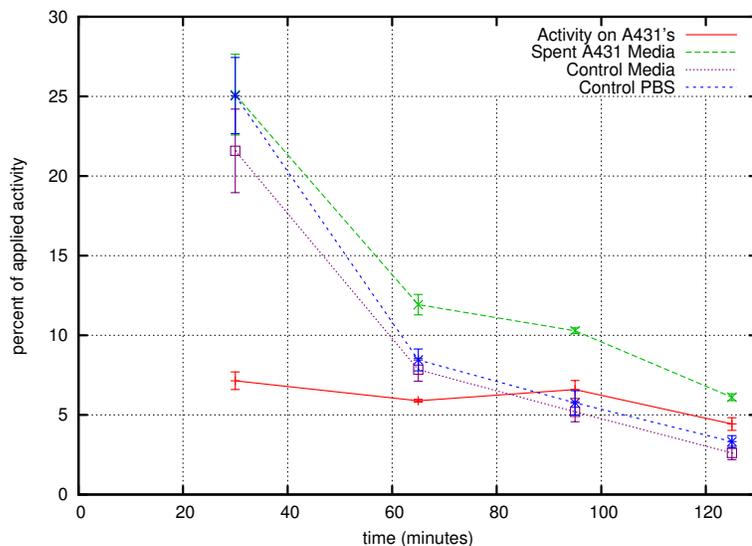
(a) Inactivation of RLuc-EGF under cell culture conditions
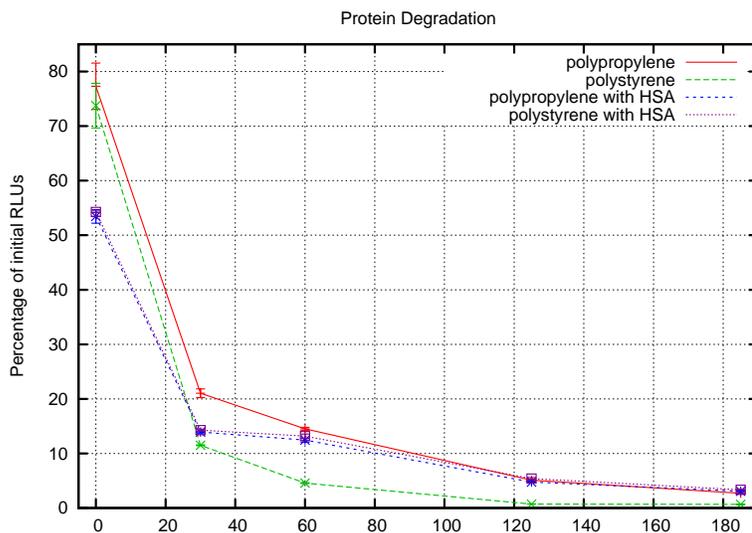
(b) Inactivation of RLuc-EGF exposed to plastic containers

Figure 6.5: Inactivation of the RLuc-EGF fusion protein under cell culture conditions and exposure to plastic containers. In Panel (a), A431 cells were exposed to RLuc-EGF in medium without FBS for the given amounts of time. The cells were then washed and imaged, with spent medium samples taken and measured in the luminometer. Additionally, an empty plate was set up with RLuc-EGF in medium or PBS, and incubated identically to the cell containing plate. In Panel (b), RLuc-EGF was diluted in PBS with our without 1% HSA and exposed to polystyrene or polypropylene containers at 37°C. At the indicated time points, aliquots were removed and assayed. Error bars represent standard error.

	Specific Activity	μg protein per ml culture
RLuc/C124A	1.2	7.5
EGF-RLuc/C124A	0.05	1.1
RLuc/C124A-EGF	0.07	0.7

Table 6.2: Protein production values for EGF-RLuc/C124A and RLuc/C124A-EGF. Specific activity was measured using coelenterazine and is expressed as relative to that of RLuc (See Table 2.1 for absolute values).

In assaying the medium for this experiment, 10 μl of medium was added directly to the sodium phosphate buffer rather than the usual 1 μl of protein diluted in EB with 1% HSA. The end effect is that medium conditions with the carrier protein had 10-fold more HSA in the assay than usual, and the medium conditions without a carrier protein had no HSA in the assay.

6.3 Results and Discussion - EGF Fusions Using RLuc/-C124A

The results with the RLuc based constructs have shown that inactivation of the probe's luciferase moiety at 37°C is a major issue. This thermolability of *Renilla* luciferase's activity has been encountered by other authors [120], and had lead to the identification of a C124A point mutation [119] (listed as C152A in that publication) that greatly improves the resistance of RLuc to inactivation. A more detailed discussion of this mutation is given in Chapter 2. Based on the previously published success with the C124A mutation, it was decided to incorporate this mutation into the EGF fusion constructs.

6.3.1 Protein Production and Stability

Production values for RLuc/C124A, EGF-RLuc/C124A, and RLuc/C124A-EGF are given in Table 6.2. The expression conditions used were 2 hr at 32°C following induction of the culture. For EGF-RLuc/C124A and RLuc/C124A-EGF, ~25% of the activity was present in the periplasmic fraction after osmotic shock, with the remainder in the pellet. In total, for equal volumes of culture ~50-fold more activity was recovered from the RLuc/C124A fusions than the previous RLuc based constructs.

Both silver stained gels and westerns gave bands of appropriate sizes for nickel affinity purified RLuc/C124A, EGF-RLuc/C124A, and RLuc/C124A-EGF, with the western shown in
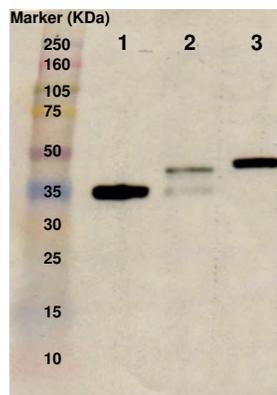
Figure 6.6: A western performed on purified RLuc/C124A, EGF-RLuc/C124A, and RLuc/-C124A-EGF. The purified proteins were run on a Tris-HCl 10-20% gradient gel under reducing conditions, transferred to a PVDF membrane, and blotted using a monoclonal antibody to RLuc. Lanes are 1: RLuc/C124A, 2: EGF-RLuc/C124A, 3: RLuc/C124A-EGF.

Figure 6.6. The secondary band present in the EGF-RLuc/C124A lane of the western was present in the silver stain gel as well (not shown), and indicates that some amount of clipping is occurring between the EGF and RLuc/C124A domains of this fusion. No band corresponding to EGF-6xHis was observed in the lane of RLuc/C124A-EGF on the silver stained gel (not shown). As these constructs had not yet been buffer exchanged on a PD-10 gel filtration column, low molecular weight fragments (<10 kDa) should have appeared if present and the absence of an EGF-6xHis band indicates that interdomain clipping was not occurring with the RLuc/C124A-EGF construct.

Serum stability for the RLuc and RLuc/C124A based fusion proteins was measured to assess the effect of the C124A point mutation on the constructs. The results in Figure 6.7 show that this point mutation does enhance the protein's resistance to inactivation, but also indicates that the EGF fusion constructs are remarkably more labile than their correspond non-fused luciferases.

6.3.2 Cell Binding

Specificity of the RLuc/C124A based fusion proteins was assessed by comparing binding of the different constructs on A431 and NIH 3T3 cells. The results shown in Figure 6.8 indicate that both the EGF-RLuc/C124A and the RLuc/C124A-EGF fusion proteins specifically bound to the high EGFR expressing A431 cells. The EGF-RLuc/C124A fusion protein did not perform as well as the RLuc/C124A-EGF construct, a result consistent with the known contamination of
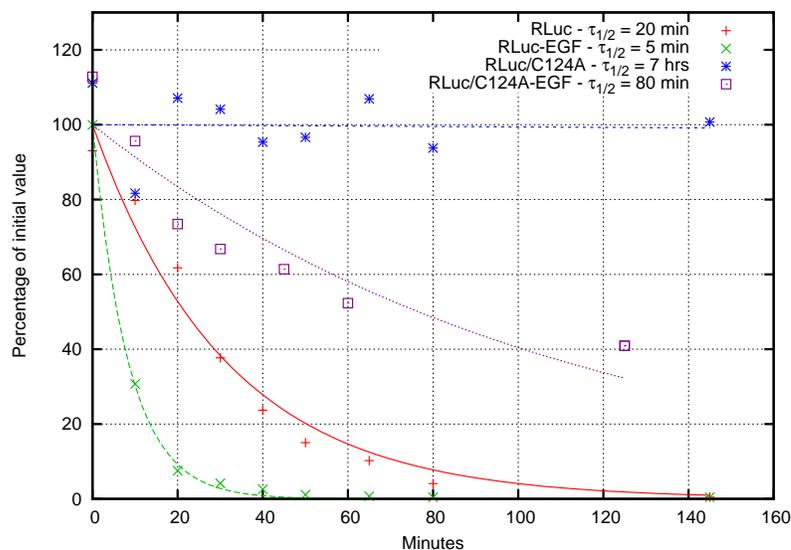
Figure 6.7: Resistance of RLuc, RLuc-EGF, RLuc/C124A, and RLuc/C124A-EGF to inactivation in mouse serum at 37°C. The raw data (points) are shown superimposed on the results of monoexponential non-linear curve fits (lines). The half-life for RLuc/C124A is from the data in Table 2.1, as insufficient data is present here to fit the exponential.

the EGF-RLuc/C124A sample with the cleaved RLuc/C124A fragment.

6.3.3 Stability under Culture Conditions

An experiment was done to test the benefit of the C124A point mutation constructs under cell culture conditions. The results, shown in Figure 6.9, indicate that both RLuc/C124A and RLuc/C124A-EGF are far more resistant to inactivation under these *in vitro* conditions than the corresponding RLuc based constructs (Figure 6.5a).

From these results, another estimate of the probe binding affinity can be made. Rewriting Equation 6.2 with respect to the total concentrations of probe and receptor yields:

$$K_{D'} = \frac{(T_{L'} - [L'R])(T_R - [L'R])}{[L'R]} \tag{6.8}$$

Based on the data in Figure 6.9, $[L'R] = 0.065[T_{L'}]$. Based on the RLUs of activity applied to the cells and the specific activity of RLuc/C124A-EGF, $[T_{L'}]$ can be estimated as roughly 1 nM. As before, the receptor concentration is ~1 nM. Substituting these into Equation 6.8 yields a value of $K_{D'} \leq 13\ nM$. This value is most likely more accurate than the estimation in Section 6.2.3 as the C124A based probe is more stable, but it is again an overestimation of the
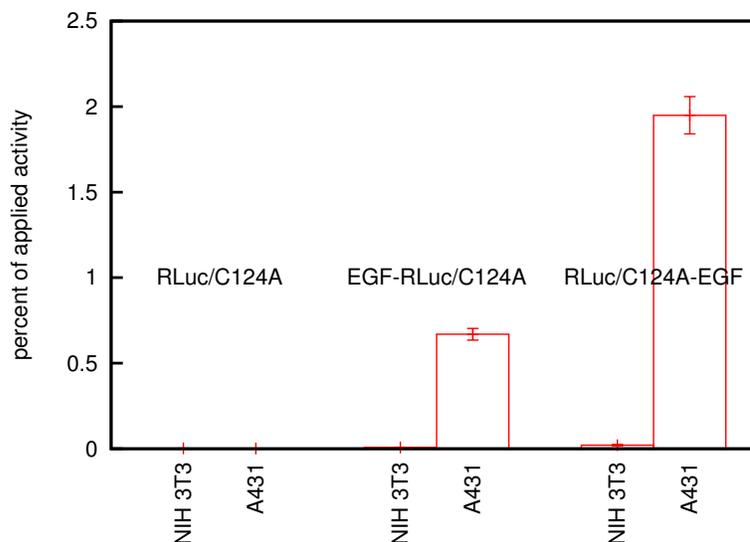
Figure 6.8: Study of EGF-RLuc/C124A and RLuc/C124A-EGF fusion protein binding to A431 and NIH 3T3 cells. Cells in 24-well plates were exposed to protein (RLuc/C124A, EGF-RLuc/-C124A, or RLuc/C124A-EGF) for 40 min, washed, and imaged. Error bars represent standard error.

dissociation constant as will be explained in Section 6.4.7.

6.3.4 Competitive Binding

A competitive binding experiment was done using RLuc/C124A-EGF to further demonstrate the specificity of RLuc/C124A-EGF for EGFR. The results shown in Figure 6.10 are almost identical to the previous results for RLuc-EGF discussed in Section 6.2.3.

6.3.5 Animal Experiments

Distribution of RLuc/C124A

The distribution and persistence of RLuc/C124A *in vivo* was assessed by injecting the protein into non-tumor bearing athymic (nu/nu) mice. RLuc/C124A ($\sim$5 μg) was injected *i.v.* at $t = 0$, with additional *i.v.* injections of 10 μg coelenterazine using the contralateral tail vein immediately preceded each imaging time point. A representative mouse study is shown in Figure 6.11. The results demonstrated that imaging of *i.v.* injected C124A mutated RLuc is feasible over a time scale of hours. Foci of activity along the flanks of the mice are presumably from luciferase passing through the kidney, as the relatively small size of this probe (39 kDa) would lend it to
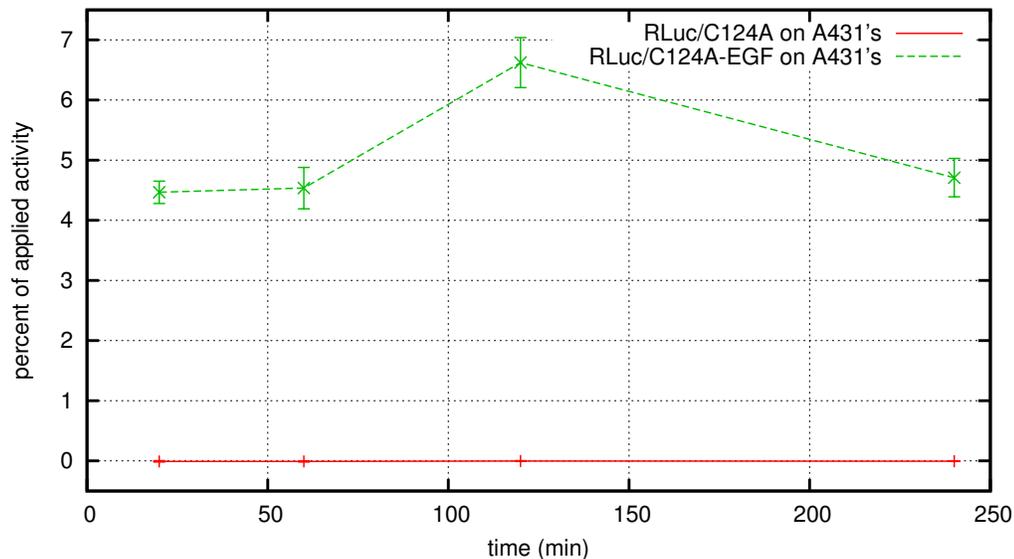
(a) Luciferase Activity Affixed to Cells
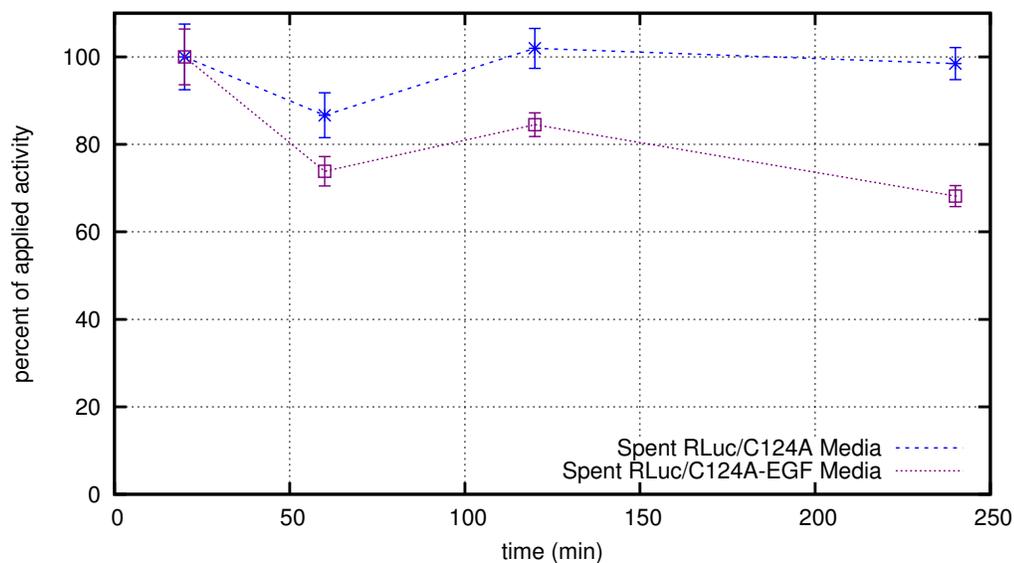
(b) Luciferase Activity in the Medium

Figure 6.9: Binding of RLuc/C124A-EGF to A431 cells and stability of RLuc/C124A and RLuc/-C124A-EGF to culture conditions. In Panel (a), A431 cells were exposed to RLuc/C124A-EGF (50,000 RLU/well ∼1 nM) or RLuc/C124A in medium without FBS for the given amounts of time, the cells were then washed and imaged. Panel (b) shows the results from spent medium samples that were taken and measured in the luminometer. Error bars represent standard error.
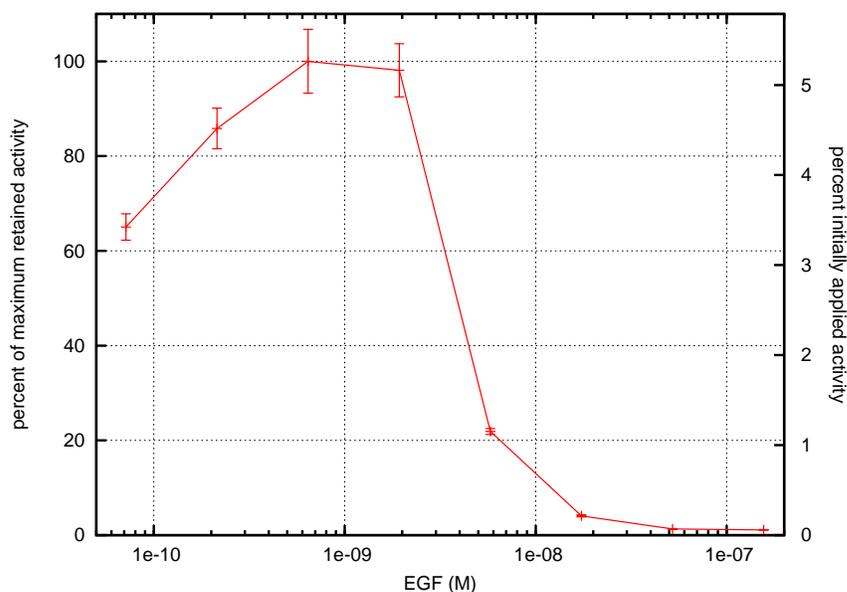
Figure 6.10: Competitive of RLuc/C124A-EGF versus EGF on A431 cells. A431 in 24-well plates were exposed to a fixed amount of RLuc-EGF (∼0.2 nM) and varying levels of recombinant human EGF for 40 min. The wells were then washed and imaged. Error bars represent standard error.

a low rate of glomerular filtration [129]. This renal clearance of luciferase was confirmed in an [124]I labeled RLuc8 study (described in Section 6.4.8). An interesting observation apparent in a subset of the scans (not shown), is that bite marks present on the backs of some of the animals show up very brightly in these studies.

Rluc/C124A-EGF in a Tumor Model

Following the distribution study, A431 xenograft models were set up in athymic mice. When the tumors became palpable, an *i. v.* injection of either 45 μg of RLuc/C124A-EGF or 5 μg of RLuc/C124A was made. Scans were then performed at approximately 5 min, 1 h, and 2.5 h by *i. v.* injection of coelenterazine (10 μg) and imaging in the prone position. On the following day, the study was repeated but with the alternate protein.

A representative study is shown in Figure 6.12. The results show that at the early time points examined, any amount of specific binding to the A431 tumor that may have occurred with the fusion protein (RLuc/C124A-EGF) could not be distinguished from non-specific retention as observed with the control (RLuc/C124A). The non-specific retention processes could involve either protein extravasation through leaky tumor endothelium [142] and/or tumor hy-
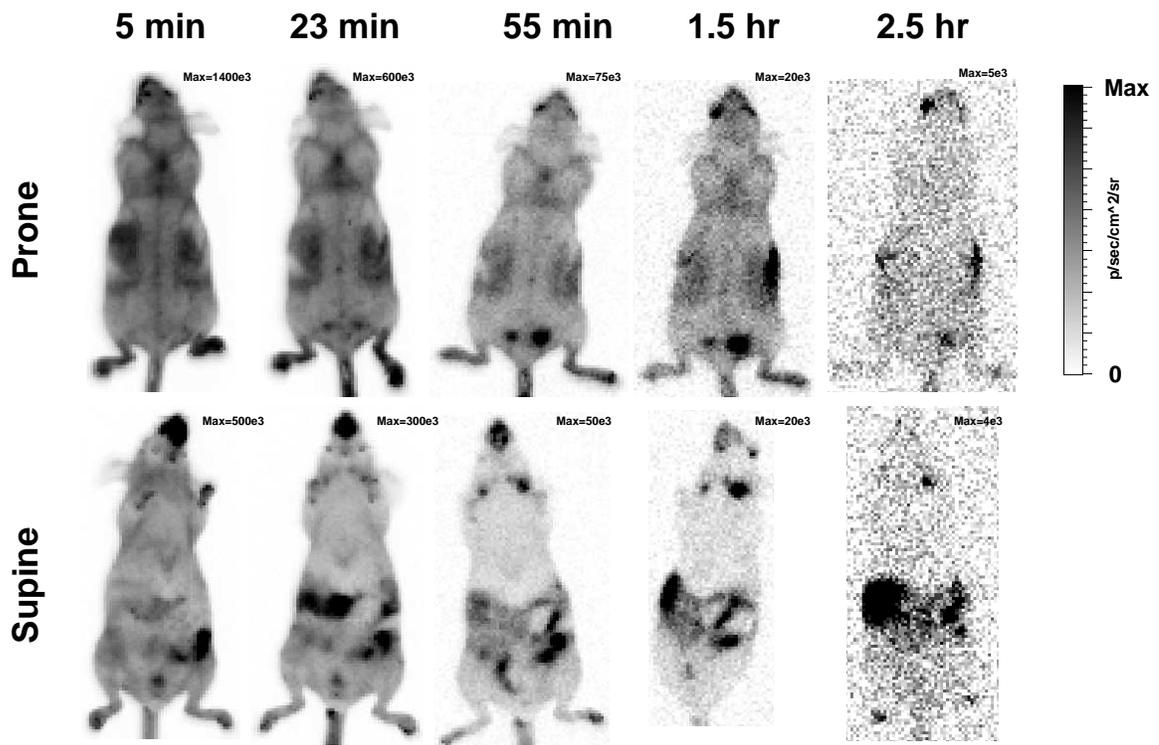
Figure 6.11: A representative mouse from a study of RLuc/C124A distribution in non-tumor bearing mice. A 10 μg injection of coelenterazine was made immediately before each set of prone/supine images. Image acquisitions were 60 s. As the supine images were taken immediately after the prone images, the somewhat lower activity levels seen in these images are expected.

pervascularization and resultant increased blood pool signal. In either case, due the quick *in vivo* kinetics with which the tumor associated activity fades, later time points at which non-specifically retained probe might clear could not be imaged.

Of note for the *in vivo* work performed in this chapter using mouse xenograft models, although there is ∼70% homology between human and mouse EGF, they are known not to activate the receptor of the opposite species [31]. As activation of EGFR is coupled to binding [184], their inability to cross-react presumably means they also do not bind the other species' receptor.

6.4 Results and Discussion - EGF Fusions Using RLuc8

At this point in the project, the stabilized variant of *Renilla* luciferase (RLuc8) described in Chapter 2 had been developed. For the third time, EGF fusion constructs were made, and the corresponding proteins expressed. One alteration in these constructs versus earlier ones was that the RLuc8 based fusions did not contain the Myc epitope present in the previous probes. Another alteration, is that some of the constructs were made with a $(G_4S)_3$ linker between the domains. These linker constructs were pursued based on a previous report that a linker in an EGF based fusion protein facilitated binding of the fusion to EGFR [226].

6.4.1 Protein Production and Properties

Production values for the various RLuc8 based fusion proteins, along with some additional constructs incorporating elements described in Sections 6.4.3 and 6.4.7, are summarized in Table 6.3. The half-lives reported were derived from the somewhat noisy data in Figure 6.13. As the data only went out 50 h in time, the half-lives are not fit with a high level of accuracy.

A silver stain gel on several of these proteins (Figure 6.14) showed that after a single step of purification, the proteins were essentially pure. Note that the gel was vastly over-developed during the silver stain in an attempt to bring out the RLuc8-EGF band and this led to the artifactual "tails" seen in the other lanes.

6.4.2 Pulse-Chase with RLuc8 Constructs

At this point in the project, the decision was made to begin using pulse-chase *in vitro* assays as a more physiologically accurate way to assess the binding of the different fusion proteins. In
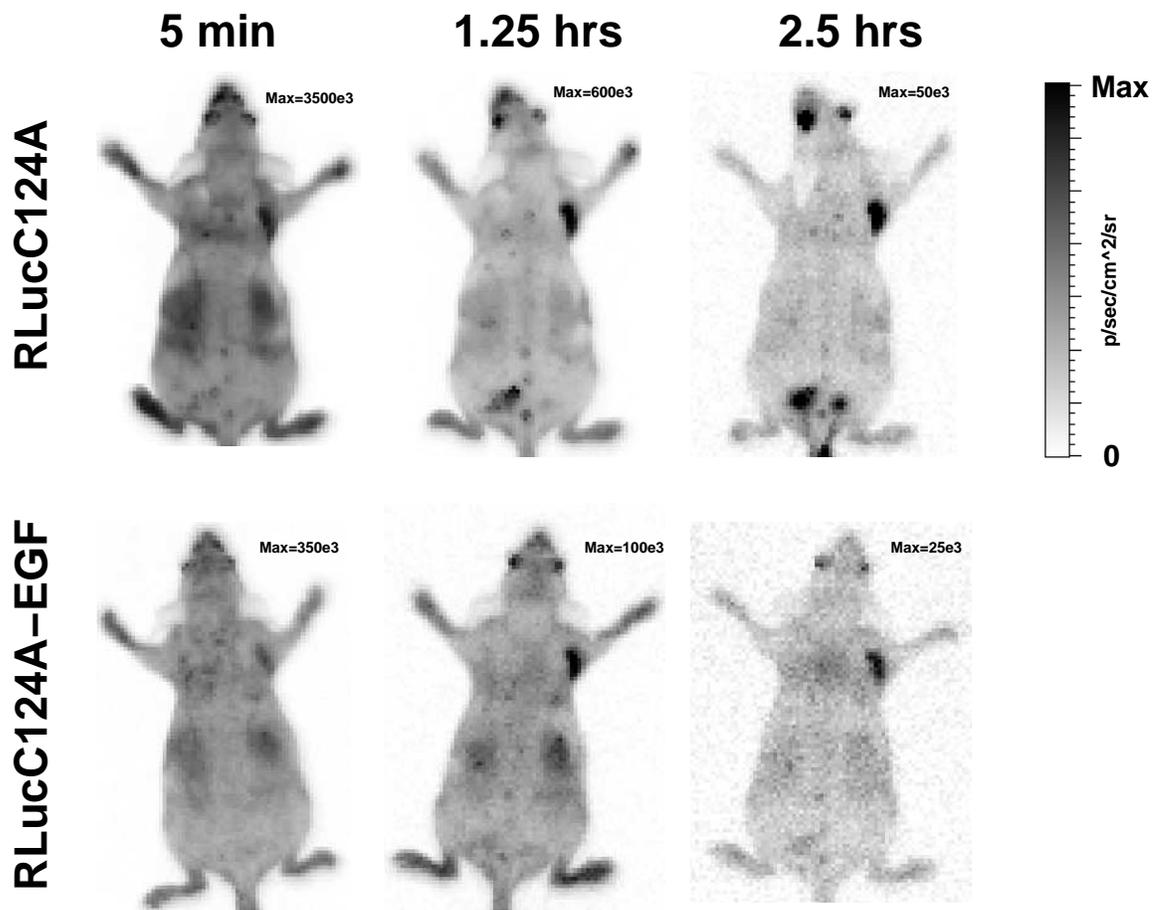
Figure 6.12: A representative mouse from a study of RLuc/C124A versus RLuc/C124A-EGF imaging in A431 tumor bearing mice. The tumor is located on the right shoulder. A 10 μg injection of coelenterazine was made immediately before each image, and acquisition times were 60 s.

| | Specific Activity | μg protein per ml culture | Serum Inact. $\tau_{1/2}$ (h) | Wavelength (nm) | | | %>600 nm |
				peak	mean	fwhm	
RLuc8	4.3±0.2	82	281±49	486	503	94	4
EGF-(G$_4$S)$_3$GS-RLuc8	0.66	4.4					
RLuc8-EGF	0.33	1.1	42	484	502	92	4
RLuc8-(G$_4$S)$_3$-EGF	0.20	1.9					
RLuc8-TDT	4.2	33	190				
RLuc8-TDT-EGF	1.3	4.8	36				
RLuc8-RL-TDT	1.7	16	187				
RLuc8-RL-TDT-EGF	1.1	6.9	89	482	499	98	4
RLuc8-RL-TDT-(G$_4$S)$_3$-EGF	1.0	4.6					
S3RLuc8-EGF	2.2	0.6					
S3RLuc8-EGF-KDEL	2.1	1.4					

Table 6.3: Protein production values for various RLuc8 fusion proteins. Coelenterazine was used for measuring the spectra and specific activity. Specific activity values are expressed as relative to that of RLuc (See Table 2.1 for absolute values). The TDT (Translocation domain of Diphtheria Toxin) constructs are described further in Section: 6.4.3. Except for S3RLuc8 and S3RLuc8-EGF, all proteins were periplasmically expressed. Blank entries were not determined.
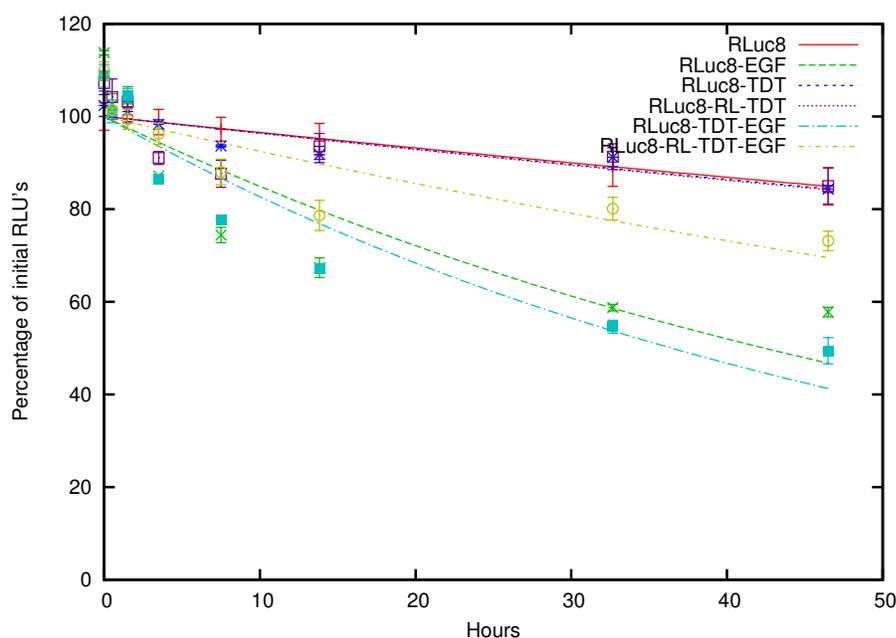
Figure 6.13: Stability of the RLuc8 based constructs in mouse serum at 37°C. The points show the raw data, and the results of monoexponential non-linear curve fits are given by the lines. The fitted half-life values are given in Table 6.2.
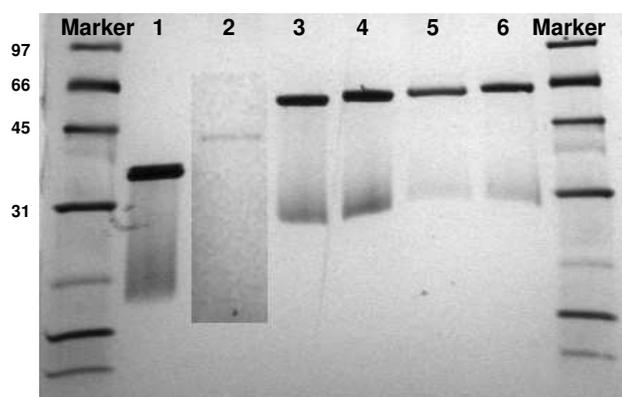
Figure 6.14: A silver stained gel of the RLuc8 fusion constructs. The purified proteins were run on a Tris-HCl 10-20% gradient gel under reducing conditions. The silver stain was over developed to bring out the band in lane 2, and the contrast in this lane was additionally increased in software. Lanes are 1: RLuc8, 2: RLuc8-EGF, 3: RLuc8-TDT, 4: RLuc8-RL-TDT, 5: RLuc8-TDT-EGF, 6: RLuc8-RL-TDT-EGF.

the pulse-chase experiments, the protein was first allowed to bind for 1 hr to the A431 cells (the "pulse"). The medium was then removed, and the cells were allowed a variable time "chase", before being washed again with medium and imaged. Essentially, the pulse-chase experiment tries to emulate the bolus of probe that is given to an animal and the quick wash out of activity that occurs thereafter.

As an additional change to better reflect *in vivo* conditions, either 10% FBS or 1% HSA was included in the cell culture medium. This had the beneficial effect of reducing inter-experiment variability as the albumin present acted as a carrier protein, but had the negative effect of increasing background levels of luminescence. For all the pulse-chase results, a value of ≤ 0.5-1% of the initially applied activity can be considered background.

A pulse-chase experiment with several of the RLuc8 fusion proteins is shown in Figure 6.15. The first result that can be drawn from this experiment, is that the presence or absence of a linker domain did not seem to alter the functionality of the C-terminal EGF fusion proteins. This is in agreement with a previous report in the literature on an Angiogenin/EGF fusion protein [244], in which a single glycine, a G_4S linker, and an absence of linker in both C and N-terminal EGF fusion constructs gave binding results equivalent to recombinant EGF. As an alternative hypothesis based on the structures presented in Chapter 4, it may simply be because the C-terminus of RLuc extends away from the protein and a further flexible linker is not required to make a successful fusion protein.

Another result from this pulse-chase experiment is that the N-terminal EGF fusion protein
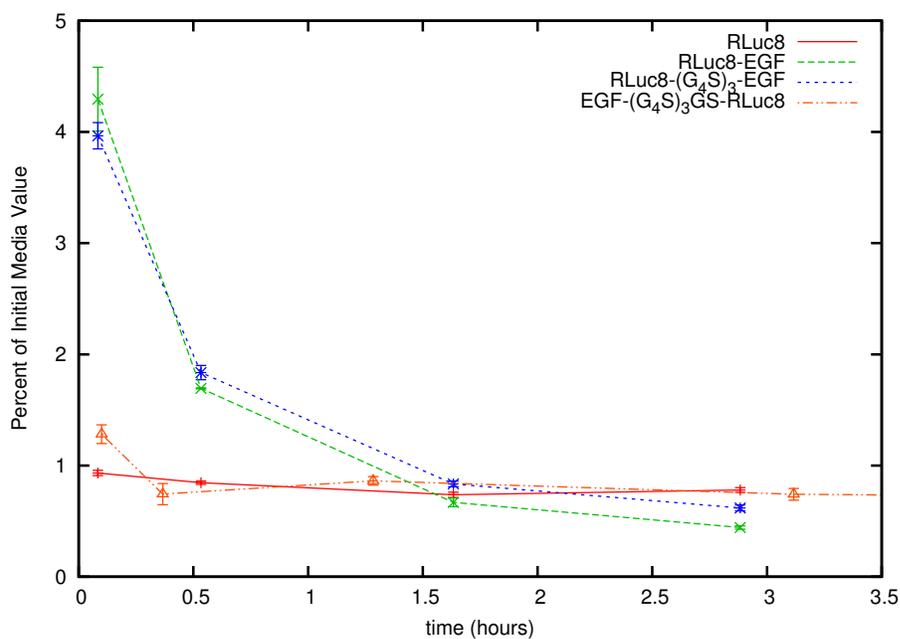
Figure 6.15: Pulse-chase experiment of the RLuc8 based fusion proteins on A431 cells. A431 cells were exposed to the indicated fusion proteins for 1 hr. For the EGF fusions, ∼0.3 nM of probe was used for each condition, while for the RLuc8 control condition only 0.03 nM was used in order to keep the imaging system from saturating. The medium in the wells was then replaced with fresh unlabeled medium (t=0). Following a variable incubation time, the medium was again replaced, coelenterazine was added, and the plate was imaged. Error bars represent standard error of the mean.

performed poorly. As previously mentioned for the other N-terminal EGF fusions, this is most likely due to contamination of the purified protein with protein that has had the EGF domain clipped off. The final and most important result from this pulse-chase experiment, is that following binding to EGFR the activity of the fusion protein is rapidly lost.

Pulse-chase experiments were also done with RLuc8, RLuc8-EGF, and RLuc8-$(G_4S)_3$-EGF on NIH 3T3 cells (data not shown). At all time points, the cell associated activity was less than 1% of the initial activity in the medium.

6.4.3 Pulse-Chase with Translocation Domain Constructs

The hypothesis generated from the pulse-chase experiments, was that binding of the EGF fusion proteins to EGFR was leading to activation of the receptor, internalization, and subsequent degradation of the fusion protein/EGFR complex. As this is an internalization process, it was hoped that the use of bacterial toxin translocation domain could potentially circumvent this process.

Bacterial toxins generally consist of a binding, a translocation, and a catalytic domain. Following binding to a cellular receptor and subsequent internalization, the acidification of the endosome triggers the translocation domain to go through a conformational change such that the catalytic (toxin) domain is translocated into the cytoplasm. Fusion proteins based on bacterial toxin translocation domains have seen some amount of success in directing payloads into the cytoplasm, with examples in the literature including a *Pseudomonas* exotoxin A translocation domain/TGF-α fusion [59], a Diphtheria toxin translocation domain/EGF fusion [186, 121], and a Diphtheria toxin translocation domain/antibody fusion specific for ErbB2 [215]. An added benefit of using a translocation domain based approach would be that it could enable the use of reporter proteins that function optimally only in the cytoplasm. Firefly luciferase would be particularly interesting due to its suitable emission spectrum for *in vivo* imaging, and its requirement for ATP would ensure that it would only be able to emit light following translocation into the cytoplasm.

The translocation domain of Diphtheria toxin, a protein produced by *Corynebacterium diphtheriae*, was chosen for two reasons. First, it is the best understood of the various bacterial toxins containing translocation domains [58]. Second, its primary structure, with an enzymatic domain on the N-terminal side of the translocation domain and the binding domain on the C-terminal end, matched the orientation of the RLuc8-EGF construct.

Only the translocation domain of Diphtheria toxin (TDT) is needed to form the channel

through which the catalytic domain is translocated [163]. However, in the full length toxin the arginine rich loop (RL) between the catalytic and translocation domains is cleaved following endocytosis by the mammalian protease furin. This cleavage leaves the N-terminal catalytic domain tethered to the rest of the toxin through a disulfide bond formed from the two cysteine residues flanking the arginine rich loop. This arrangement allows release of the catalytic domain following its translocation by the reduction of the disulfide bond in the reducing environment of the cytoplasm. As it was unclear whether the arginine rich loop would be absolutely required for the purposes here, constructs were made both with and without this loop.

The protein production values of the luciferase/TDT/EGF fusion proteins are given in Table 6.3, and the results of a pulse-chase experiment performed with these constructs are shown in Figure 6.16. One result from this experiment was that the translocation domain by itself, as in the RLuc8-TDT and RLuc8-RL-TDT constructs, did not increase background binding of the fusion proteins. A less fortunate result was that the presence of the translocation domain led to a decrease in the amount of probe initially bound to EGFR, and did not lead to a sustained increase in luciferase activity associated with the A431 cells. A pulse-chase experiment was also done with RLuc8-RL-TDT-$(G_4S)_3$-EGF on NIH 3T3 cells. At all time points, cell associated activity on these cells was less than 1% of the initial activity in the medium (data not shown).

The conclusion from this experiment, is that the Diphtheria translocation domain could not successfully translocate the luciferase portion of these constructs into the cytoplasm of the A431 cells. For Diphtheria toxin, the domain that is translocated can only pass through the pore created by the translocation domain in an unfolded manner [57]. Unfolding of the translocated domain may be a general requirement for all bacterial translocation toxins, and is known to be required in at least Anthrax toxin [232] and ricin [9]. The lack of success in this experiment may simply be that RLuc8 is unable to unfold in the acidified endosome, or that RLuc8 is unable to refold correctly after it has entered the cytoplasm. In essence, the catalytic domain of the Diphtheria toxin has evolved so that it can be translocated, while reporter proteins have not. It would be interesting to see whether a translocation domain from another toxin and/or use of another reporter protein (e. g. firefly luciferase) could lead to successful translocation.

6.4.4 Pulse-Chase with Inhibition of Lysosomal Acidification

An experiment using bafilomycin A_1 was done in order to demonstrate that lysosomal mediated degradation was leading to the loss of cell associated activity in the pulse-chase experiments.
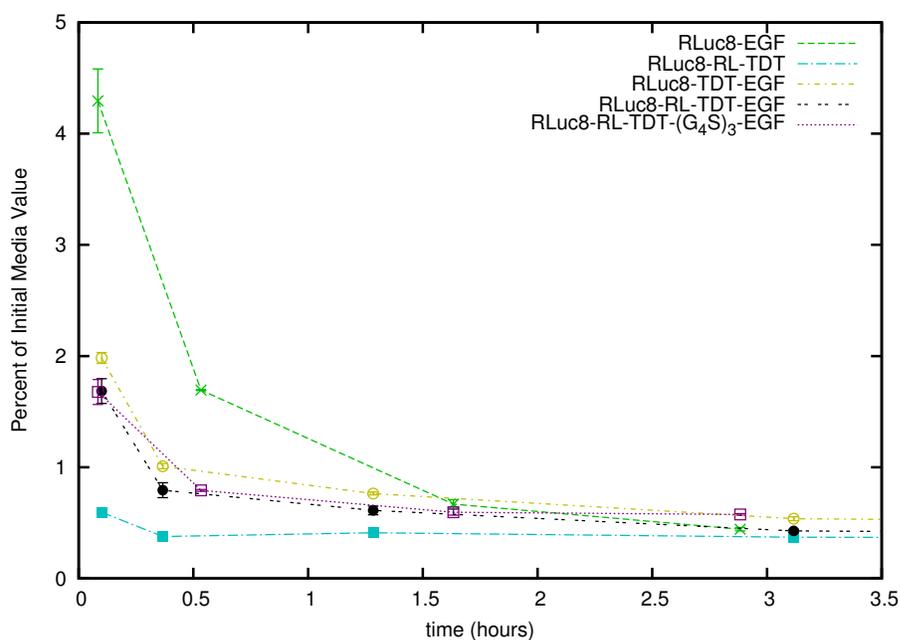
Figure 6.16: A pulse-chase experiment on A431 cells using RLuc8 based fusion proteins incorporating the Diphtheria translocation domain. A431 cells were exposed to the indicated fusion proteins (at 0.1-0.3 nM) for 1 hr, and the chase begins at t=0. The data for RLuc8-EGF is duplicated from Figure 6.15 for the purpose of comparison. Error bars represent standard error of the mean.
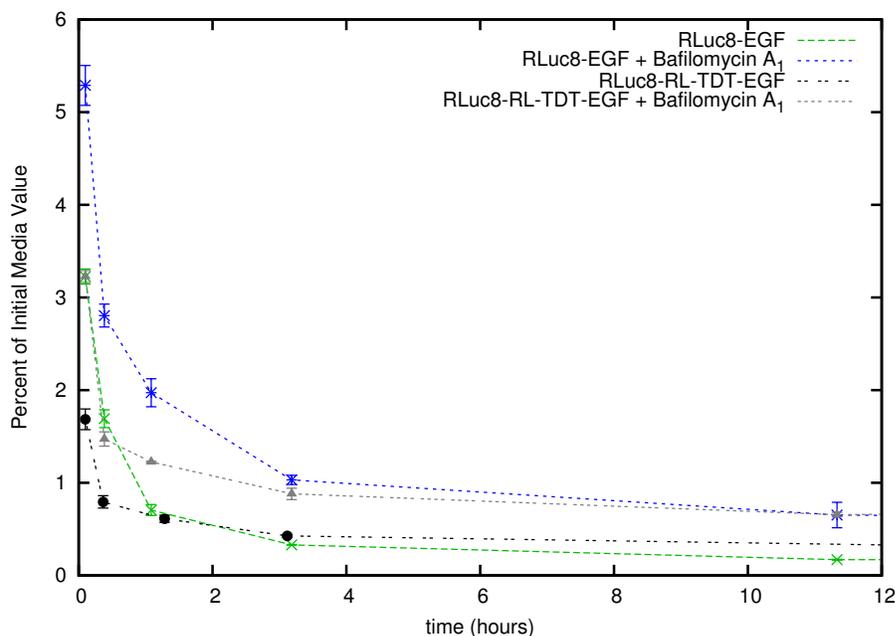
Figure 6.17: Pulse-chase experiments in the presence of bafilomycin A_1 to inhibit lysosomal acidification. Bafilomycin A_1 was used at a final concentration of 200 nM. RLuc8-EGF was used at 0.3 nM, and RLuc8-RL-TDT-EGF was used at 0.1 nM. The chase begins at t=0. Error bars represent standard error of the mean.

Bafilomycin A_1, a macrolide antibiotic isolated from *Streptomyces* sp., is a specific inhibitor of the endosomal proton pumps (vacuolar type H^+-ATPases) and is active at nanomolar concentrations [23]. Bafilomycin-A_1 has previously been shown to inhibit EGF degradation following internalization of the EGF/EGFR complex in A431 cells [245], presumably because the increase in pH inhibits activation of many lysosomal proteases.

The results shown in Figure 6.17 help confirm that it is receptor mediated endocytosis followed by lysosomal degradation that is leading to the loss of activity in the pulse-chase experiments. The results also confirm that the Diphtheria toxin translocation domain constructs are not working as had been hoped. If the translocation domains were functional, inhibition of endosomal acidification would prevent activation of the translocation domain and reduce the amount of cell associated activity with these constructs. This is contrary to what is seen.

6.4.5 Pulse-Chase using Coelenterazine Analogs

There has recently been interest in the coelenterazine analogs coelenterazine-*cp* and coelenterazine-*n* as, in contrast to the native substrate, they are not substrates for MDR1 P-glycoprotein
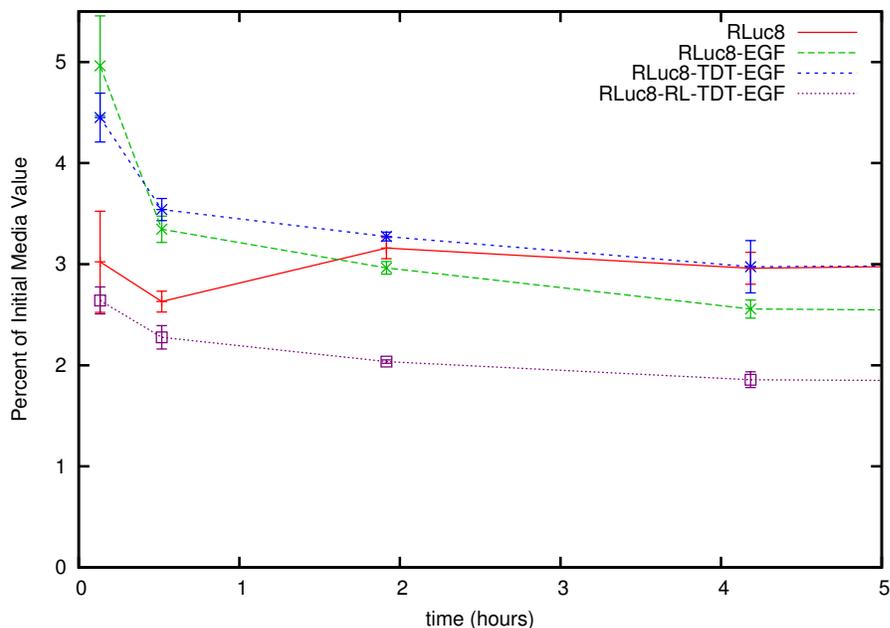
Figure 6.18: Pulse-chase experiment on RLuc8 based fusion proteins using coelenterazine-*cp*. The chase begins at t=0. Error bars represent standard error of the mean.

(Pgp) [173]. Although A4431 cells have Verapamil-sensitive efflux pumps and are not known to overexpress Pgp [117], these coelenterazine analogs were tested anyway.

Pulse-chase experiments using coelenterazine-*n* had major issues with background signal, with an initial background luminescence rate due to this substrate that was 15% of the luciferase containing medium. This background signal actually increased as the chase continued, with the rate of increase being roughly the same as the doubling time for A431 cells. Pulse-chase experiments performed with coelenterazine-*cp* were more successful (Figure 6.18), but still showed a background level of signal that was at least a 3-fold increase over comparable experiments using the native substrate. The comparable trends between the results here (Figure 6.18) and the previous pulse-chase experiments confirm that the use of the native substrate is not leading to complications in these experiments.

6.4.6   Pulse-Chase with EGF Mutants

Mutations of EGF were examined to determine whether they could lead to increased initial binding on the A431 cells or enhance the retention of activity. Initially, a series of mutations identified by Mullenbach *et al.* were examined [151]. The four mutations were G12Q, Y13W,
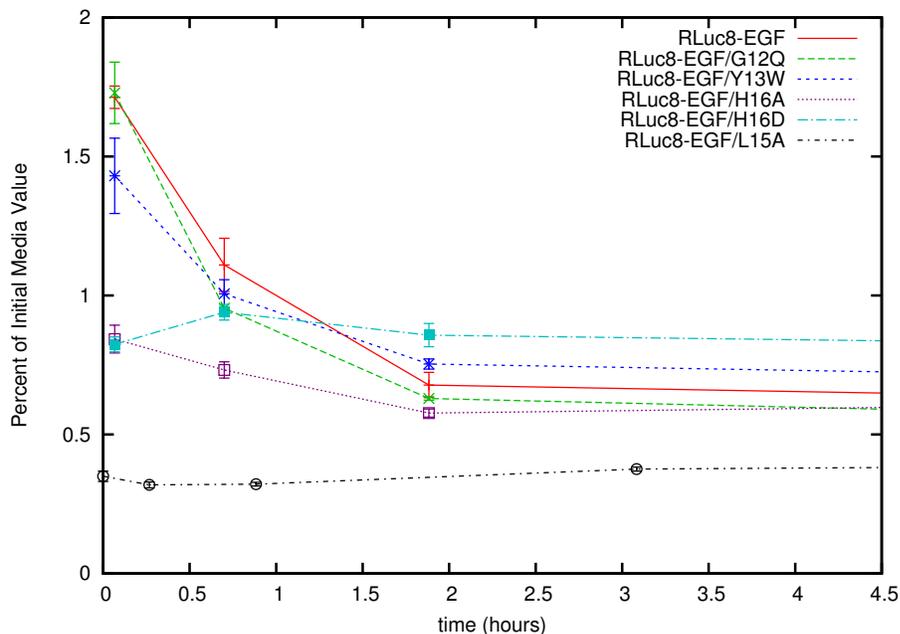
Figure 6.19: A pulse-chase experiment on A431 cells using mutants of EGF identified by Mullenbach *et al.* or Nandagopal *et al.* The chase begins at t=0, and uses each protein at a concentration of 0.3 nM. Error bars represent standard error of the mean.

H16A, and H16D. G12Q, Y13W, and H16D had been reported to increase the affinity of EGF to its receptor by a factor of 2-4, while H16A had been reported to maintain affinity while making the binding more stable with respect to pH. Additionally, an L15A mutation that had been reported to be a partial agonist of EGFR by Nandagopal *et al.* was examined [155]. The authors had speculated that tyrosine kinase activation and EGFR dimerization are not necessarily coupled, and it was hoped that this mutant might slow internalization of the complexed receptor.

The results of applying these mutants to A431 cells in a pulse-chase experiment are shown in Figure 6.19. All 5 of the mutants actually performed worse than the native ligand in the context of this assay. L15A in particular showed poor binding to the cells, a result that has been reported previously in the literature [151].

An additional set of EGF mutants were made based on variants of EGF identified by Dr. Jennifer Cochran [42].[2] These mutants were created by a combination of random mutagenesis and DNA shuffling, with yeast display technology employed as the screening process for selecting the higher binding affinity variants. Compared to wild-type EGF, these mutants exhibited higher binding affinities, increased EGFR down regulation, and led to increased chemotactic
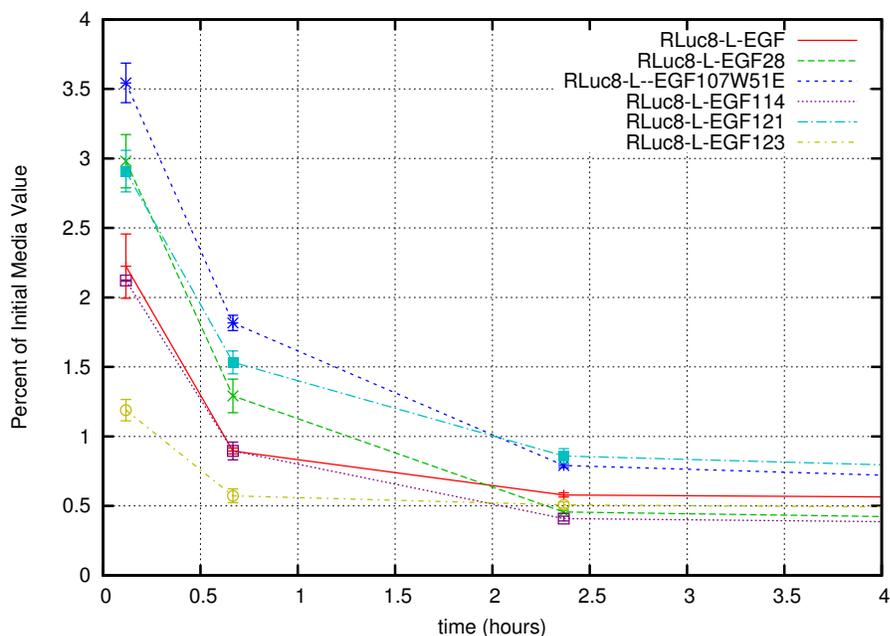
Figure 6.20: A pulse-chase experiment using mutants of EGF identified by the Cochran Laboratory on A431 cells. The chase begins at t=0, and uses each protein at a concentration of 0.3 nM. Error bars represent standard error of the mean.

responses in various cell migration assays. Several of the mutations identified through this selection process are actually found in the EGF homologs present in other species, this serves as an interesting footnote to the consensus guided mutagenesis approach that had been used in Chapter 2.

In any case, it was hoped that the increases seen in the functional assays (e. g. chemotactic response) indicated that the mutants might exhibit prolonged activation of the EGFR tyrosine kinase. As the kinase domain of EGF bound EGFR is active while the complex is on the cell surface and in the early endosome [234], prolonged retention of kinase activation would indicate a longer period between receptor binding and eventual degradation of the complex in the lysosome.

The results of a pulse-chase experiment with these mutants (Figure 6.20) showed that several of these mutants increased the initial binding of the fusion construct to EGFR. Nevertheless, they all showed no long term gains compared to the native ligand, reaching background levels of activity within 2 h of the chase period.

---

[2]The mutant 107W51E used here is mutant 107 in the reference with a W51E mutation to enhance its solubility.

6.4.7   Pulse-Chase with Cytoplasmically Expressed Fusion Proteins, a KDEL Tagged Protein, and an Inhibitor of Internalization

Following the success in cytoplasmically expressing *Gaussia* luciferase (GLuc) using Origami strain *E. coli* cells and a pET32 based plasmid (Section C.5), it seemed logical to follow a similar expression approach with RLuc8-EGF and see if this leads to any changes in the functionality of the fusion protein. GLuc is thought to require 5 disulfide bonds to fold correctly, and the hope was that this same expression system would correctly handle the 3 disulfide bonds of EGF. The resultant fusion protein was entitled S3RLuc8-EGF as the introduced thrombin site necessitated an A2G mutation in RLuc8, resulting in the parental sequence beginning at S3.

As an additional strategy, a KDEL sequence was added to the C-terminus of S3RLuc8-EGF. The KDEL sequence is recognized by the KDEL receptor in the intracellular sorting network, and it is normally used as a signal that a transcribed protein should be retained in the endoplasmic reticulum (ER) [115]. But the KDEL receptor can also function as a scavenger, moving KDEL tagged proteins in a retrograde fashion from later compartments in the sorting network (e. g. the Golgi apparatus) back to the ER [58]. The hope was that an internalized, KDEL tagged EGF/Luciferase fusion would be rescued from the lysosome directed pathway and end up being retained in the ER.

Expression results for S3RLuc8-EGF and S3RLuc8-EGF-KDEL are given in Table 6.3, and demonstrated very high specific activities compared to the previous luciferase/EGF fusions. Using denaturing gel electrophoresis, both of these fusion proteins were shown to run as single bands of the correct size (Figure 6.21).

Finally, following a suggestion from Jennifer Cochran, incubation with phenylarsine oxide (PAO) was tried as an additional condition. PAO is an irreversible inhibitor of clathrin mediated endocytosis [235, 77], and as a result it should inhibit the bioluminescently labeled EGF/EGFR complex from trafficking to the lysosome.

The results shown in Figure 6.22 indicate that cytoplasmically expressed S3RLuc8-EGF performed no differently than the periplasmically expressed RLuc8-EGF used in previous experiments (e. g. Figure 6.15). Furthermore, the KDEL sequence made no difference with respect to retention of luciferase activity on the cells. The results did show that with inhibition of clathrin mediated endocytosis (PAO condition), the activity associated with the cell increased several fold at the initial time point and remained several fold over background levels for at least 16 h.

Using the same mathematics as described in Section 6.3.3, the dissociation constant for
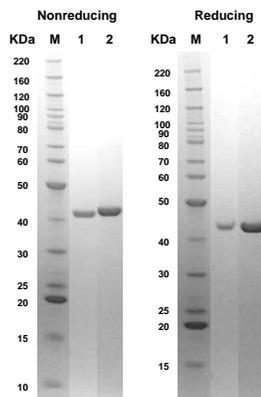
Figure 6.21: Coomassie blue stained gels of S3RLuc8-EGF and S3RLuc8-EGF-KDEL. Lanes are M: Marker, 1: S3RLuc8-EGF, and 2: S3RLuc8-EGF-KDEL. Both gels were 4-12% gradient Bis/Tris gels.
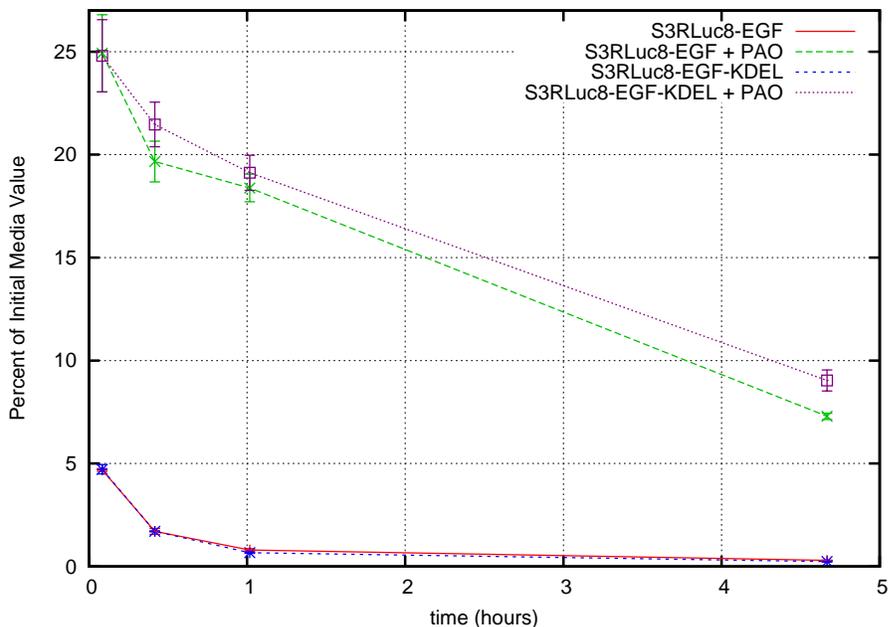
Figure 6.22: A pulse-chase experiment using cytoplasmically expressed RLuc8 based EGFR probes and phenylarsine oxide (PAO) on A431 cells. PAO was used at a concentration of 100 μM, while the fusion proteins were used at a concentration of 0.1 nM. The chase begins at t=0. Error bars represent standard error of the mean.

S3RLuc8-EGF can be calculated as $K_{D'} = 3$ *nM* for the PAO treated case. This value of 3 nM is very close to the reported $K_D \sim 5$ nM for EGF binding EGFR on A431 cells [95].

The apparent dissociation constant for S3RLuc-EGF in the non-PAO treated case can be calculated as 20 nM. Clearly this is incorrect, as the dissociation constant calculation is not taking into account destruction of the probe. For this reason, 20 nM only serves as an upper limit for the $K_{D'}$ value, and this explains the reason that previously calculated $K_{D'}$ values (e. g. Section 6.2.3, 6.3.3) were given as inequalities.

6.4.8 Animal Experiments

Distribution of RLuc8

The *in vivo* distribution of RLuc8 was studied using ^{124}I labeled RLuc8 and positron emission tomography (PET). For this work, purified protein was sent to Dr. Anna Wu's lab at UCLA where it was iodinated and used for imaging in mice. The results shown in Figure 6.23 indicate that the majority of the activity enters the renal system within 30 min of injection and clears into the bladder within 1.5 hrs, consistent with the results in Section 6.3.5. Further results along with methodology are available in other publications [216, 219].

RLuc8 Based EGF Fusions in a Tumor Model

A431 xenograft models were set up in nude mice to study the effectiveness of the various RLuc8 based EGF fusion proteins for *in vivo* EGFR imaging. When the tumors became palpable, the mice were imaged on subsequent days using the various proteins. Representative images from these studies are shown in Figure 6.24.

As with the previous *in vivo* tumor imaging studies (Section 6.3.5), the control luciferase (RLuc8) was retained in the tumor for at least a matter of hours, reflecting non-specific retention processes due to leaky vascular endothelium in the tumor blood vessels. At later time points (>5 h), the output luminescence from the EGF fusion proteins became very weak, presumably from a combination of non-specific binding clearing from the tumor, and specifically retained probe being degraded through a receptor mediated endocytosis process. The translocation domain constructs did visibly "better" with respect to the length of time that they were retained in the tumor. This result is expected, as the translocation domain constructs are significantly larger than the RLuc8 and RLuc8-EGF proteins, and can be expected to have prolonged circulation times along with slower clearance rates from tumors.

(a) 0-15 s   (b) 2-3 min   (c) 7-8 min   (d) 20-25 min   (e) 90-100 min

(f) 0-15 s   (g) 2-3 min   (h) 7-8 min   (i) 20-25 min   (j) 90-100 min
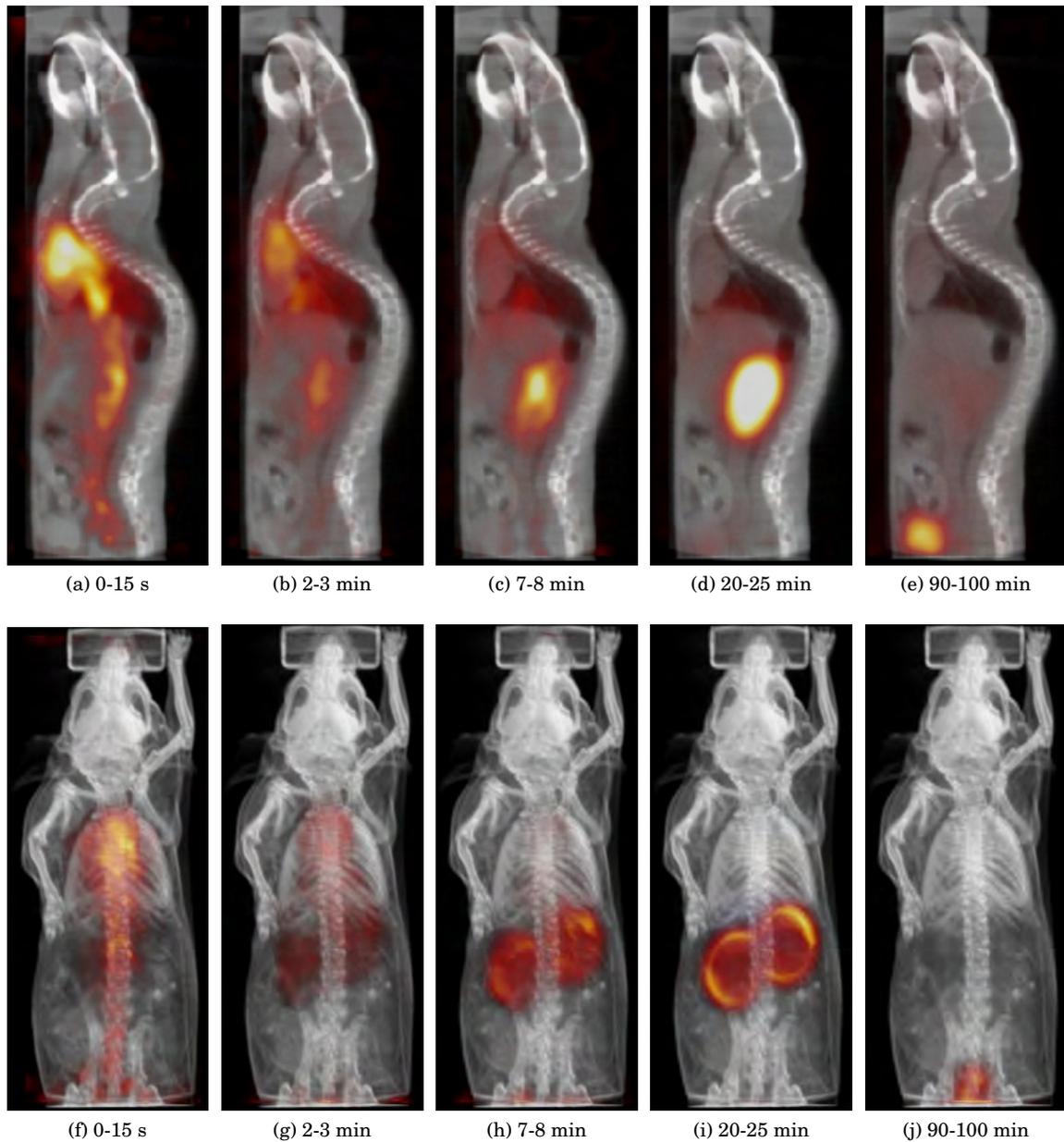
Figure 6.23: PET images from a mouse injected with ^{124}I labeled RLuc8. 141 μCi of radiolabeled RLuc8 was injected via tail vein into the mouse (t=0), with scanning occurring during the injection so as to acquire data for the initial blood curve. A CT image was acquired following the scan and registered to the PET data to allow anatomical visualization [40]. Panels (a)-(e) show sagittal slices, with a slick thickness of 4 mm utilized so that multiple organs can be visualized in a single slice. Panels (f)-(j) show the same time points as volume rendered images to allow visualization of the entire animal. Time intervals over which the data in a given image were acquired is shown in the caption below each figure. All images were created using AMIDE (Chapter 5).

(a) Mouse 7239-3

(b) Mouse 7239-4
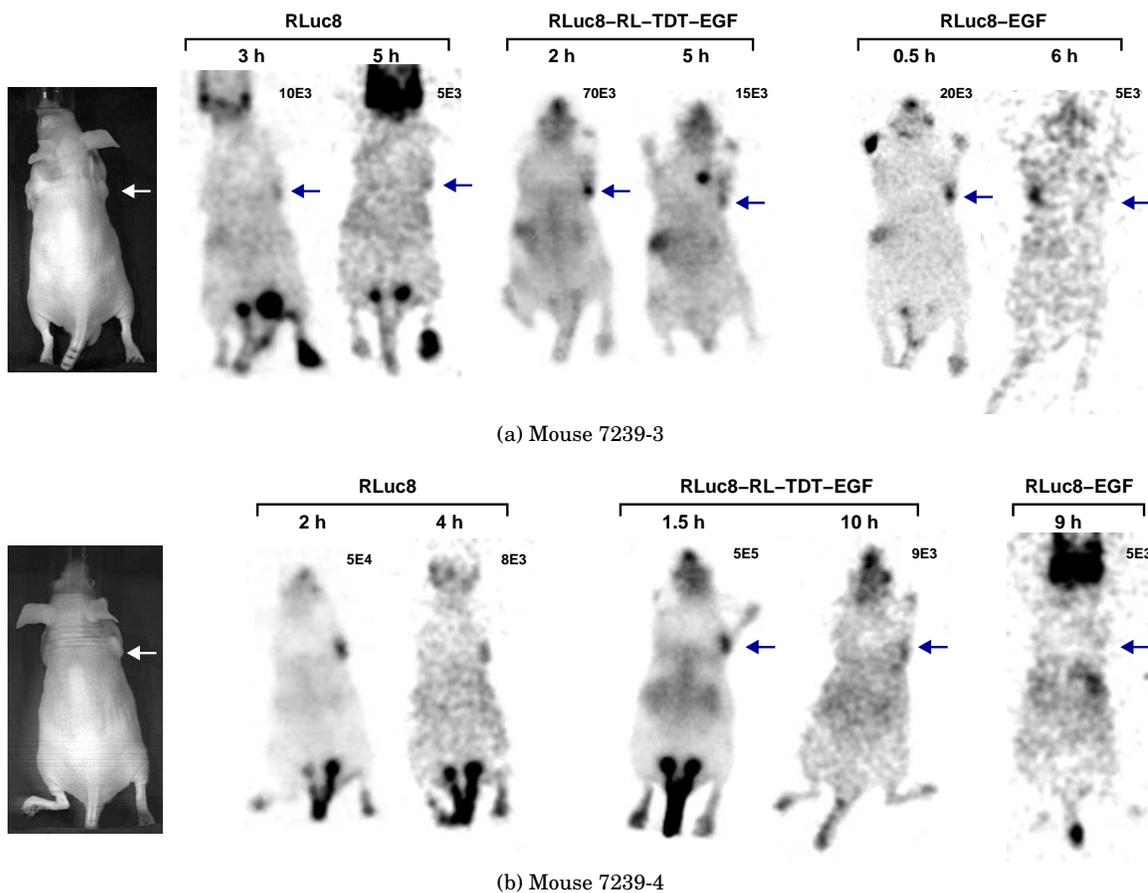
Figure 6.24: Example scans of mice injected with various RLuc8 based imaging probes intended to target EGFR. Both mice carry an A431 tumor on the right shoulder, indicated by an arrow in the figures. Injections of the various proteins were done on subsequent days, and the time listed indicates the number of hours following administration of the given protein. For each image, the mouse was injected with 10 μg of coelenterazine immediately prior to acquisition. The actual order of protein injections for the mouse in Panel (a) was RLuc8-EGF, RLuc8-RL-TDT-EGF, and RLuc8. The actual order of protein injections for the mouse in Panel (b) was RLuc8-RL-TDT-EGF, RLuc8, and RLuc8-EGF. The amount of protein injected was 5 μg, 7 μg, and 1.5 μg for RLuc8, RLuc8-RL-TDT-EGF, and RLuc8-EGF, respectively. The number next to each luminescence image indicates the thresholding maximum level, in photons/s/cm^2/steradian. In Panel (a), the bright spot on the top left of the 0.5 h RLuc8-EGF luminescence image is the animal's paw. What are likely lymphatic channels along the tail are visible in several of the scans, and may indicate extravasated protein at the injection site.

6.5 Conclusion

The end conclusion of the above experiments is that functional EGF/luciferase proteins have been made, but due to rapid loss of the luciferase moiety's activity following binding to EGFR the fusion proteins are not useful as imaging probes. The presumption is that this activity loss is due to receptor mediated endocytosis of the imaging probe followed by subsequent degradation in the lysosome, a hypothesis that is strongly supported by the results of the bafilomycin A_1 and phenylarsine oxide pulse-chase experiments.

A native ligand based imaging probe is a difficult challenge in many imaging modalities and is especially problematic when the probe's imaging moiety is a protein. A further complication here was the choice of the EGF/EGFR system. Without activation, EGFR internalizes with a $\tau_{1/2}$ ~30 min [234], but is rapidly returned from the early endosome to the cell membrane. Upon binding of ligand and receptor activation, the rate of internalization increases by an order of magnitude. A rough estimate for the time between activated EGFR internalization and arrival in the lysosome is 20 min [77], a number that does not bode well for imaging studies that are likely to require clearance times on the order of hours before specific retention can be discriminated from non-specific processes.

So what about using a variant of EGF that binds the receptor but does not lead to activation? With respect to the prevention of internalization, the real bugbear of the EGF/EGFR system is that ligand binding is closely coupled to a conformational change in the receptor and it is the receptor that undergoes the dimerization. To date, there have been no ligand based antagonists of EGFR reported in the literature. In many other ligand systems, it is the ligand that is dimerized, with the resultant bivalent binding of two receptors leading to activation of the receptor system. Such ligand systems would have been more appropriate for the demonstration of a bioluminescently labeled imaging probe, as they allow the generation of protein level receptor antagonists. For instance, protein ligand antagonists of vascular endothelial growth factor (VEGF) have been constructed by creating heterodimers of the ligand in which one of the monomers has an inactivated binding interface [193, 20]. Such a heterodimer can bind to one receptor, but is unable to bind a second receptor and lead to activation.

In truth, a more tractable and generalizable approach would be to use a non-activating antibody specific for the receptor of interest rather than trying to make use of the receptor's native ligand. The main advantage of using a ligand is that its smaller size should allow for faster clearance of background signal due to non-specific retention processes. However, in the case

of bioluminescent labeling where a 36 kDa luciferase protein is being attached to the ligand, it's debatable how much one gains from the smaller size of a ligand based imaging probe, especially when engineered antibody variants such as single chain Fv fragments can be as small as 28 kDa. Additionally, antibodies (engineered or otherwise) have well understood structural properties, established pharmacokinetics, and known methods for producing them recombinantly. An example of a functional RLuc8/engineered antibody fusion that is appropriate for use as an imaging probe is presented in Section 7.2.

Chapter 7

Conclusion

As mentioned in Chapter 1, in the context of small animal imaging bioluminescently labeled imaging probes can be expected to exhibit better sensitivity than probes utilizing either radioactive or fluorescent labels. However, bioluminescent labels have the important limitation that since the imaging moiety is a protein, the environments to which the imaging probe is exposed must be tolerated by that protein. As an example of this limitation, Chapter 6 demonstrated a bioluminescently labeled imaging probe targeted at EGFR that was ultimately unsuccessful for *in vivo* imaging purposes, presumably due to degradation of the bioluminescent moiety following internalization of the target receptor. As this particular probe was ultimately ineffective, validation of the bioluminescent labeling concept has been pursued through three different collaborations.

This chapter begins by discussing the first of these collaborations, an as of yet unsuccessful attempt to create a bioluminescent imaging probe targeted at the vascular endothelial growth factor receptor. The discussion then continues with two examples in which bioluminescent labeling has been successfully utilized for small animal imaging. The first example involves a luciferase/antibody fusion targeted at carcinoembryonic antigen, and the second demonstrates the creation of bioluminescently labeled quantum dots optimized for *in vivo* imaging. Finally, a few pieces of additional work are discussed that may be worth pursuing in order to advance the technological state of bioluminescent labeling.

7.1 Bioluminescent Labeling of VEGF

The vascular endothelial growth factor (VEGF) receptor, along with its related receptors, is involved in tumor angiogenesis and has been more recently implicated as a tumor autocrine survival factor [30]. Unsurprisingly, it makes for an interesting imaging target for the purposes of following experimental animal cancer models [179]. Before the work on the EGF ligand based imaging probe discussed in Chapter 6 had been completed, an additional project was started focused on imaging VEGF receptor through the creation of a VEGF ligand/RLuc8 fusion protein. This work was originally undertaken by Jinha Park, with additional work performed by Olivier Gheysens and myself.

In many ways, the results of the bioluminescently labeled VEGF project have essentially duplicated the conclusions reached with the bioluminescent probe for EGFR. While initial *in vitro* experiments were successful [169], the animal imaging results were equivocal and later *in vitro* pulse-chase experiments pointed to serious issues with probe internalization and degradation. The lack of success for a probe based on native VEGF should have been expected, as the half-life of internalization for activated VEGF receptor is 20-30 min [225].

Further work on this project, however, may prove to be fruitful. As mentioned in Section 6.5, it is possible to create a heterodimeric version of the VEGF ligand that is able to bind its receptor without dimerizing and therefore activating the receptor. A bioluminescent imaging probe based on such a VEGF heterodimer should be able to avoid internalization and may eventually prove to be successful.

7.2 Bioluminescent Labeling of an Engineered Antibody

At the same time that the EGF/luciferase fusion proteins were being developed, a parallel project was being worked on in the laboratory of Dr. Anna M. Wu (University of California, Los Angeles). In this parallel project, the goal was to create a bioluminescently labeled antibody specific for carcinoembryonic antigen (CEA). CEA, which is actually a number of related glycoproteins involved in cell adhesion, is upregulated in neoplastic tissues and as a consequence has been the target of antibody mediated therapy and imaging [48, 239].

Using an antibody or an engineered antibody as the targeting moiety in a bioluminescently labeled imaging probe is in many ways a much better choice than a native ligand. For one thing, antibodies are extremely generalizable. The details of fusing and expressing an anti-

body/luciferase fusion need to be worked out once. After this, altering the probe's target receptor can be as simple as swapping the particular antibody that is being utilized in the imaging probe. In contrast, ligands show a great amount of sequence and structural variability. As such, details of creating the fusions and successfully expressing the fusion proteins would have to be worked out for each imaging probe. More importantly, unlike native ligand based imaging probes, most antibodies do not activate their target receptors. Internalization may still present an issue when targeting some receptors with fast non-activated receptor recycling rates (e. g. EGFR [234]), but it should not be of the same magnitude a problem as internalized non-activated receptors are often returned to the cell surface. As an example, although antibody bound CEA internalizes rapidly ($\tau_{1/2} \sim 0.5\ h$), most of it is quickly returned to the cell surface and the effective degradation half-life of the antibody/receptor complex is thought to be more on the order of 24 h [238].

The Wu laboratory has generated a genetic fusion between a CEA binding engineered antibody (diabody) developed in their laboratory and the stabilized *Renilla* luciferase variant (RLuc8) developed in this dissertation [217, 219, 216]. This construct was successfully expressed in mammalian cells, purified, and shown to target CEA in a variety of *in vitro* studies. Excitingly, a series of well controlled experiments utilizing mouse xenograft models showed that this imaging probe could be used for imaging CEA levels *in vivo*. Representative images from these *in vivo* experiments, demonstrating specific retention of the engineered antibody/RLuc8 fusion in CEA positive tumors, are shown in Figure 7.1.

7.3   Self-illuminating Quantum Dot/RLuc8 Conjugates

An interesting application of the stabilized *Renilla* luciferase variant (RLuc8) developed in Chapter 2, has been the bioluminescent labeling of quantum dots in order to make these quantum dots self-illuminating and therefore more useful for *in vivo* imaging [194].

Quantum dots have generated considerable excitement in the molecular imaging community, as several of their properties make them favorable for *in vivo* imaging purposes. Namely, they have a high resistance to photobleaching and exhibit tunable, size dependent emission spectra [145]. Detractions to the *in vivo* use of quantum dots include biodistribution issues due to their size, as well as sensitivity limitations due to the autofluorescence background inherent in all fluorescence imaging modalities. While not much can be done in ameliorating the size issue, the autofluorescence background problem could potentially be eliminated by
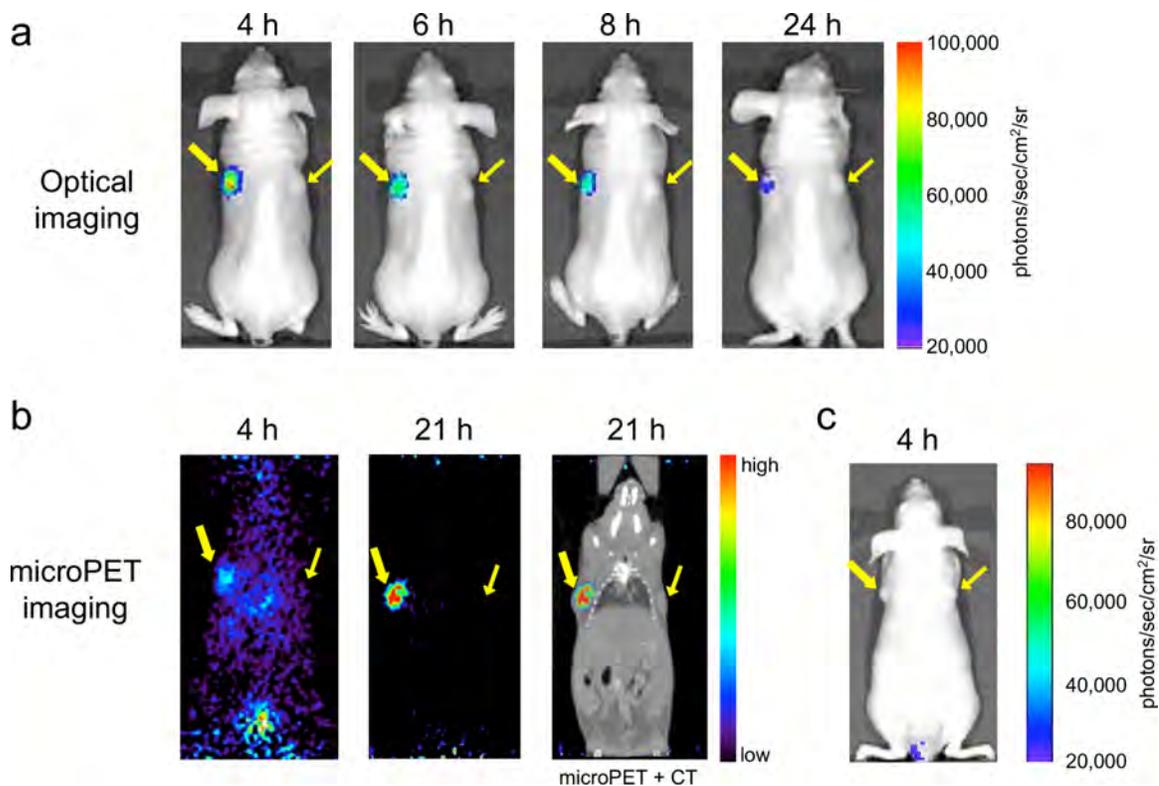
Figure 7.1: Bioluminescence and PET imaging of CEA antigen using an engineered antibody/RLuc8 fusion protein. The athymic mouse bears a CEA-positive LS174T tumor (thick arrow) on its left shoulder, and a CEA-negative C6 tumor on its right (thin arrow). Panel (a) shows bioluminescence images acquired 4, 6, 8, and 24 hours after $i.\,v.$ injection of the fusion protein. Coelenterazine injections were done $i.\,v.$ immediately prior to each image. An ROI analysis of the CEA-positive tumor gave values of 9.7×10^4, 7.0×10^4, 6.4×10^4, and 3.5×10^4 photons/s/cm^2/steradian for the 4, 6, 8, and 24 h images, respectively. Panel (b) shows a PET image of a similar tumor bearing mouse that has been injected with the engineered antibody-RLuc8 fusion protein labeled with ^{124}I. Panel (c) shows a third tumor bearing mouse that has been injected with RLuc8 as a control. The bioluminescence image was acquired at 4 h. This data is from Venisnik $et\ al.$ [219]
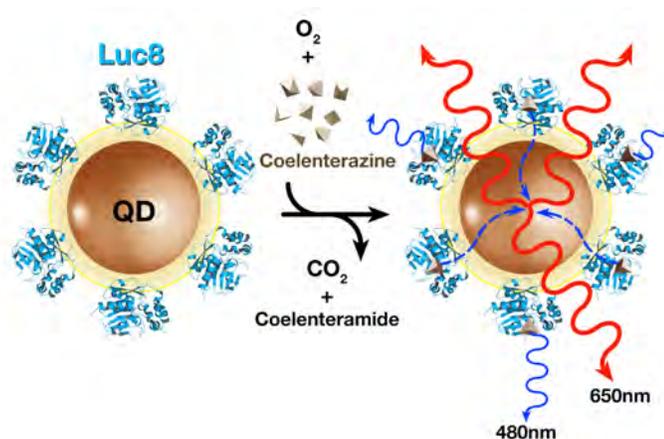
Figure 7.2: Schematic diagram of a quantum dot/RLuc8 conjugate. The RLuc8 is covalently bound to the quantum dot through the coupling of amino groups on the protein to carboxylate groups on the quantum dot's polymer coating. On average, 6 Rluc8 molecules have been estimated to attach to each quantum dot. When coelenterazine is added, the luciferase catalyzes the oxidation of the substrate. The energy yielded from this reaction can be directly released from the enzyme as a photon of blue light or passed via a resonance energy transfer mechanism to the quantum dot and emitted as a red photon of light. Graphic by Jim Strommer.

removing the need for an excitation light source.

In order to improve on the capabilities of quantum dots for *in vivo* imaging, Dr. JiangHong Rao's laboratory has worked on making them self-illuminating by coupling RLuc8 directly to the surface of polymer coated CdSe/ZnS core-shell quantum dots. A schematic diagram of how these quantum dot/RLuc8 conjugates work is shown in Figure 7.2. For the conjugation, a couple of different strategies have been attempted, with the most successful so far being a covalent bond between amino groups on RLuc8 and carboxylates present on the quantum dot surface.

The ability of these quantum dots to self-illuminate through the conjugated RLuc8 is demonstrated in Figure 7.3. When coelenterazine is added to a conjugate, a large portion of the energy released by the enzymatic reaction is transferred to the quantum dot and emitted with the dot's emission spectrum. The spectrum in turn can be altered simply by changing the size of the quantum dot used in the conjugate.

Figure 7.4 shows an application of these self-illuminating quantum dots in a mouse. The important thing to note from this figure, is that when the quantum dots are used for fluorescence imaging the sensitivity is insufficient to detect those that have been injected intramuscularly. When the self-illuminating quantum dots are imaged, however, the intramuscularly injected
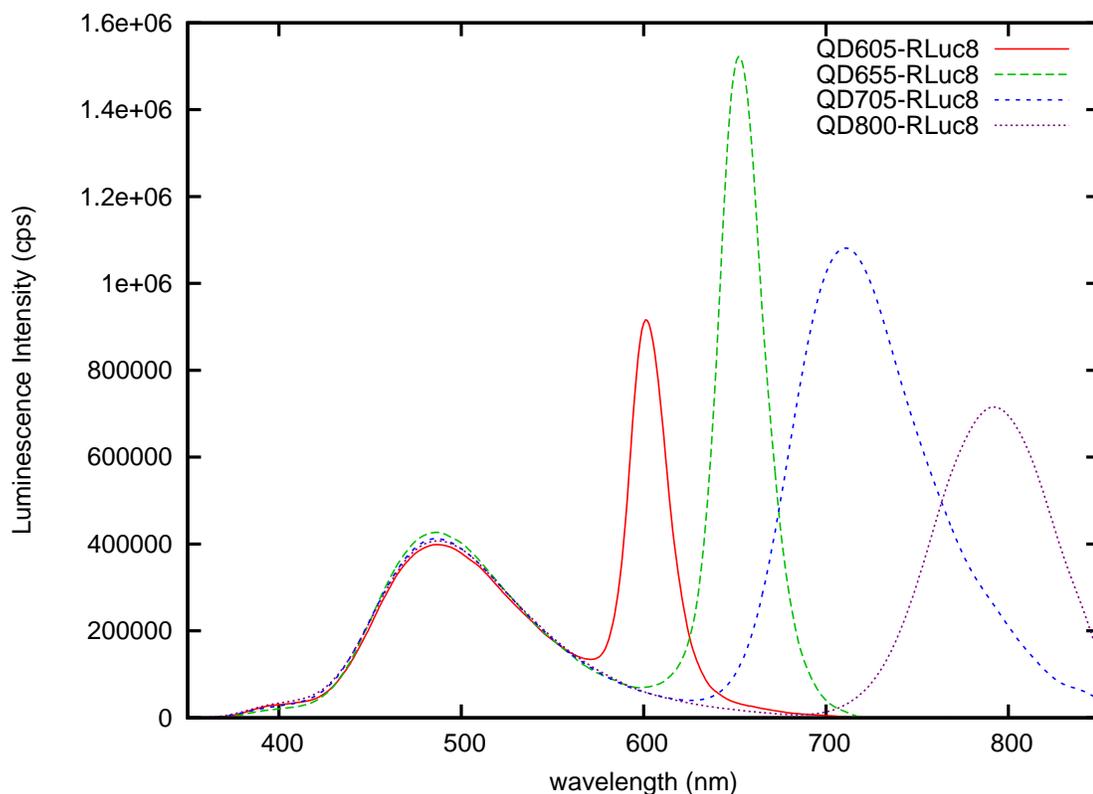
Figure 7.3: Bioluminescence emission spectra for various quantum dot/RLuc8 conjugates. The peak at 480 nm is due to photon emission directly from RLuc8, while the red-shifted emission peaks are due to resonance energy transfer to the quantum dot with associated quantum dot luminescence. The quantum dots are named according to their fluorescence emission peaks, 605, 655, 705, and 800 nm. The BRET efficiencies for the 605, 655, 705, and 800 nm quantum dots were 41, 55, 79, and 57% respectively. Coelenterazine was used as the substrate, the spectrophotometer's wavelength dependent sensitivity has been compensated for, and the curves have been smoothed. This figure is based on data published in So *et al.* [194]
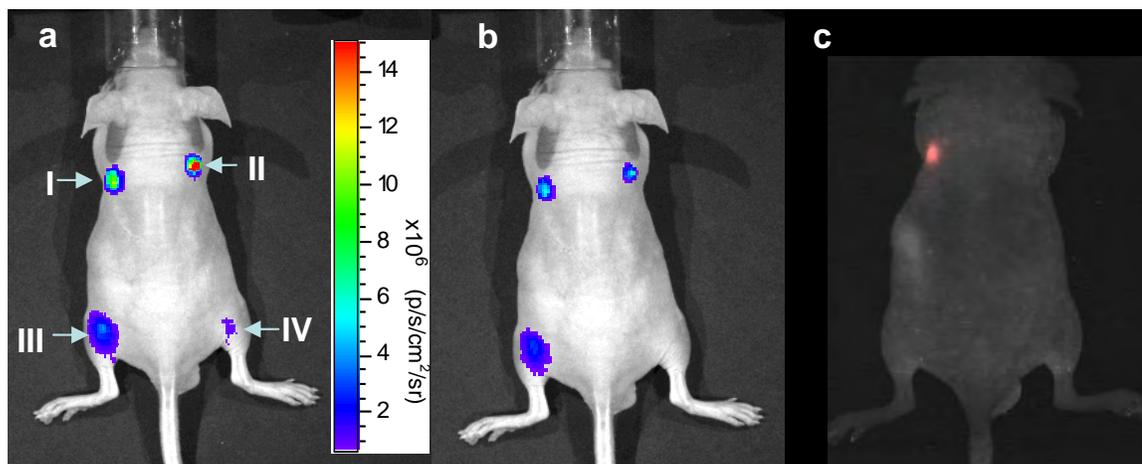
Figure 7.4: *In vivo* bioluminescence and fluorescence imaging of RLuc8 and a 655 nm emission quantum dot conjugated to RLuc8. 30 pmoles of RLuc8 was injected subcutaneously (I) or intramuscularly (III) into a mouse. Similarly, 5 pmoles of the quantum dot/RLuc8 conjugate were injected subcutaneously (II) or intramuscularly (IV). Following an *i. v.* injection of coelenterazine, images were acquired without the use of a filter, shown in Panel (a), and with the use of a 575-650 nm band-pass filter, shown in Panel (b). Panel (c) shows a fluorescence image of the same mouse acquired using a Maestro Imaging System (Cambridge Research and Instrumentation), in which spectral unmixing has been applied to reduce background autofluorescence. The excitation filter for the fluorescence was 503-550 nm. This figure is from So *et al.* [194]

conjugates are readily detectable. Additionally, the conjugates perform better than RLuc8 on its own, as the emission spectrum of the conjugate is less attenuated by biological tissue and therefore more conducive to *in vivo* imaging.

The potential this technology has for multiplex imaging was tested by injecting conjugates made of RLuc8 and 655, 705, or 800 nm emission quantum dots. The different quantum dot conjugates could be readily discriminated after imaging with emission filters (Figure 7.5), although the spectra were not completely resolved due to the filters not being optimized for these particular quantum dots.

7.4 Possible Human Interactions

An application of bioluminescently tagged imaging probes, proposed independently several times but probably first by Dr. Bruce Bryan [28, 29], is their intraoperative use to aid in the surgical resection of tumors. One can imagine a scenario where a bioluminescently labeled antibody specific for the patient's tumor is injected prior to surgery. The surgeon would then

(a) Open Filter   (b) 575-650 nm   (c) 680-720 nm   (d) 810-875 nm
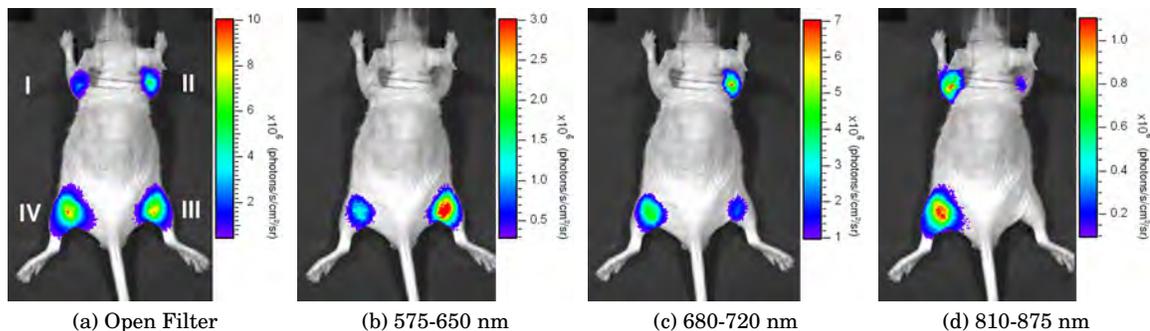
Figure 7.5: A demonstration of multiplexed bioluminescence imaging using four different quantum dot/RLuc8 conjugates. The following were injected subcutaneously, with the injection sites marked in Panel (a): (I) 15 pmoles 800 nm emission quantum dot/RLuc8 conjugate, (II) 15 pmoles 705 nm emission quantum dot/RLuc8 conjugate, (III) 5 pmoles 655 nm emission quantum dot/RLuc8 conjugate, (IV) a mixture of the 3 conjugates. The animal was then injected *i. v.* with coelenterazine and imaged using an IVIS 200 bioluminescence imaging system. This figure is adapted from So *et al.* [194]

be able to visualize the extent of tissue to be resected during the operation through the use of image intensifying goggles. Many hurdles remain before such an idea can become reality. Two potential barriers are the unknown toxicity of coelenterazine and the possible immunogenicity of the probe's luciferase moiety.

Toxicity of coelenterazine has not been studied in any formal manner. At least in mice, reports of toxicity concomitant to injection of coelenterazine have in hindsight always been due to the solvent (e. g. methanol, ethanol) in which the coelenterazine is dissolved rather than the substrate itself. Coelenterazine is present in the marine food chain [69], and at least one source of this compound appears to be crustaceans [212]. At least in small amounts, coelenterazine is not overtly toxic as it is present in the sea food we eat. Whether the amounts used in bioluminescence imaging are toxic is another issue that needs to be addressed, although finding a non-toxic solvent capable of dissolving mass amounts of coelenterazine will be a challenge. For future researchers undertaking such toxicity studies, hydroxypropyl-β-cyclodextrin may be of use, as this compound at a concentration of 50 mM can increase the solubility of coelenterazine in aqueous solution by 280-fold [207].

As to the potential immunogenicity of the bioluminescent probe's luciferase moiety, this issue has never been assessed. The animal studies in this dissertation as well as the imaging studies mentioned previously in this chapter have all been done using immunodeficient mice. At some point, the potential for immunogenicity with a bioluminescent label will need to be

addressed.

If immunogenicity is indeed an issue, a possible route around this problem could be the transformation of a human enzyme into a functional coelenterazine using luciferase. As mentioned previously (Section 2.2.1), *Renilla* luciferase exhibits a classic α/β-hydrolase fold, and this fold pattern is found in all kingdoms of life. A quick BLAST search will reveal that the most similar (41%) *homo sapiens* protein is the α/β-hydrolase fold domain of the soluble epoxide hydrolase EPHX2 (GenBank ID GI:1359739). To demonstrate the degree of structural similarity, the tertiary folds of these two enzymes are shown in figure 7.6. Transformation of this human enzyme into a functional luciferase could plausibly be done by a combination of rational design (e. g. loop grafting) and random mutagenesis.

7.5 Improved Coelenterazine Analogs

Chapter 3 described the creation of *Renilla* luciferase variants with green peaked emission spectra. For the purposes of *in vivo* imaging, even further red-shifts of the emission spectrum would be desirable, but it is unclear at this point whether continued mutagenesis of *Renilla* luciferase could accomplish this. The coelenteramide pyrazine anion that is the presumptive green emitter may represent the longest wavelength emitting state achievable within the context of the luciferase.

An alternative and perhaps more accessible path to achieving further red-shifts for this luciferin/luciferase system would be to alter the luciferin, with the idea that the red-shift from the luciferin and the red-shifts from the mutants developed in Chapter 3 could potentially be additive. An analog of coelenterazine called coelenterazine-*v* [191, 88] (Figure 7.7) has already been synthesized, shown to exhibit a high level of bioluminescence with *Renilla* luciferase, and shown to yield a ~35 nm red shifted emission compared to the native substrate. Although this analog has been reported to be relatively unstable and difficult to synthesize (personal communication with Dr. Bruce Bryan), its potential advantages for *in vivo* imaging indicate that it should be examined further. Also worth investigation are several additional analogs reported in the literature with equivalent red shifts to coelenterazine-*v* [241]. Although these analogs suffer from low levels of light output when used with the native luciferase, they may exhibit acceptable levels of luminescence when used with the various mutants developed in this dissertation.

(a) RLuc8
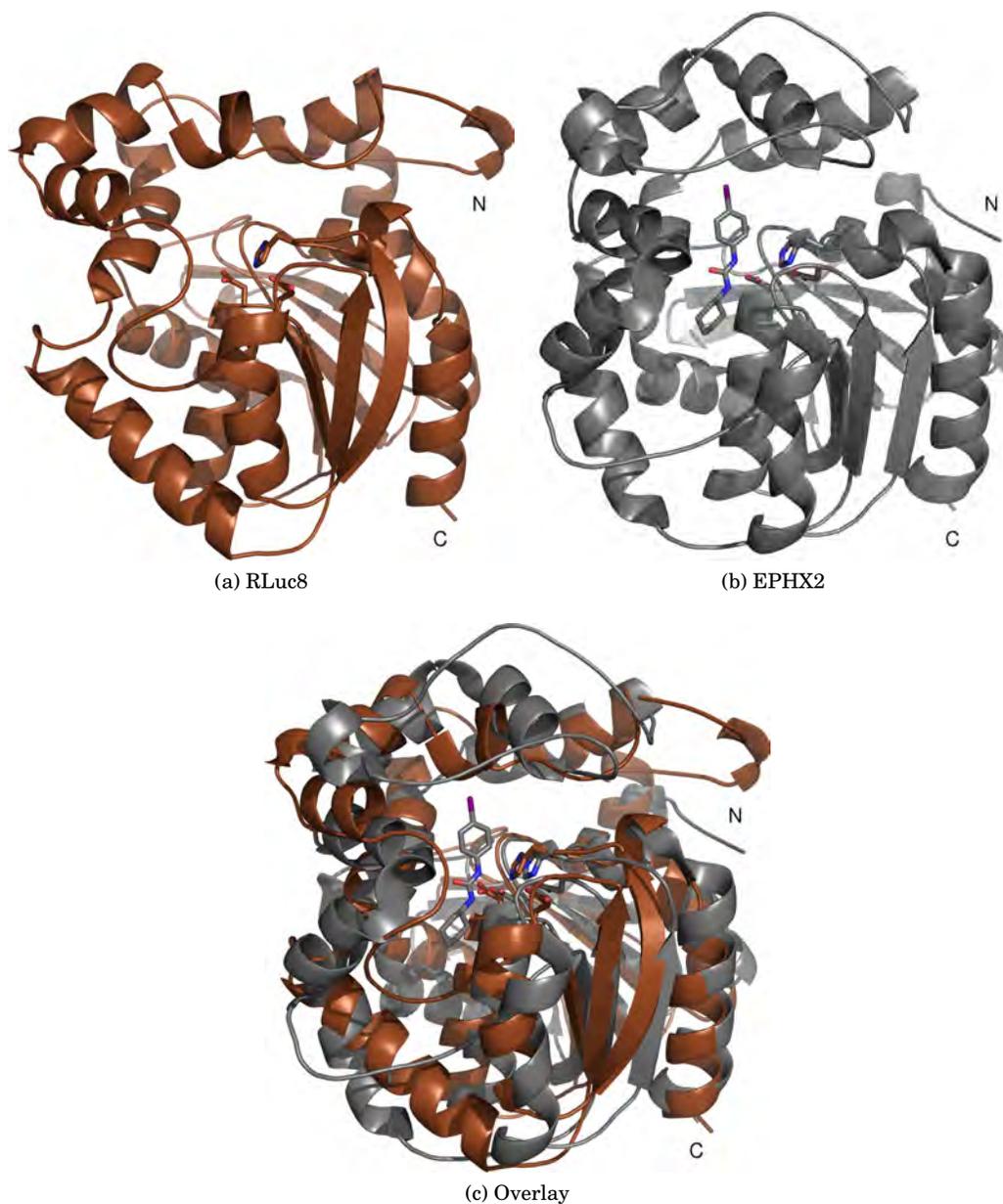
(b) EPHX2

(c) Overlay

Figure 7.6: The cartoon structure of RLuc8 overlayed with that of the human soluble epoxide hydrolase EPHX2. Only the α/β-hydrolase fold domain of EPHX2 is shown. The N-termini of the proteins are at the top right, and the C-termini are at the bottom right. The catalytic triad of D120/E144/H285 is shown for RLuc8, with the corresponding residues of D333/D495/H523 shown for EPHX2 [8]. The chemical complexed to the EPHX2 structure is *N*-cyclohexyl-*N'*-(iodophenyl)urea (CIU). The PDB ID for EPHX2 is 1VJ5 [67].
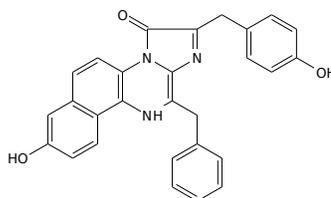
Figure 7.7: Chemical structure of coelenterazine-*v*.

7.6 Kinetics

Some kinetic work was done in Chapter 2 in an attempt to understand the effects that different mutations had on the catalytic process of *Renilla* luciferase. A model was proposed that might explain the observed kinetic data, but it was in no way rigorously tested and could quite possibly be deficient.

Future researchers may be interested in returning to the study of *Renilla* luciferase's enzymatic process, a full elucidation of which would be extremely interesting from a basic science perspective. Additionally, the interpretation of the structural information from Chapter 4 in the context of a full kinetic model would be quite illuminating. Some work on relating structure and enzymatics for coelenterazine based photoproteins has been done [122], but no such work has ever been performed for the coelenterazine utilizing luciferases.

7.7 Outro

"I told him, sir, that fruit baskets is like life - until you've got the pineapple off of the top you never know what's underneath." - From "Going Postal" by Terry Pratchet

In many ways, this dissertation has been as much about the journey as the end goal. Over the course of this work many side projects have been undertaken that, while perhaps peripheral to the goal of bioluminescent labeling, are arguably of more relevance to the general researcher. Such "dual use" technologies include: variants of *Renilla* luciferase optimized for use as transient or constitutive reporter genes in mammalian cells (Chapter 2), red-shifted variants with improved emission spectra for *in vivo* imaging (Chapter 3), a better understanding of *Renilla* luciferase at the structural level (Chapter 4), and software for viewing and analyzing multi-modality imaging studies (Chapter 5).

The end goal of this dissertation was to develop the underlying technologies needed to make bioluminescent labeling of molecular imaging probes a reality. In terms of this goal the disser-

tation has been a success, as demonstrated by the engineered antibody/RLuc8 fusion as well as the RLuc8 based self-illumination quantum dots discussed in this chapter. The hope is that bioluminescent labeling proves to be a generalizable strategy for the construction of imaging probes, and that it becomes an invaluable tool in the molecular imaging toolbox.

Appendix A

Calibration of the Luciferase Assay

Whenever possible, the results in this dissertation have been presented in terms of absolute units (e. g. photons/s/mole) or as relative to an absolute value (e. g. relative to the photons/s/mole of *Renilla* luciferase). This is in contrast to most contemporary publications making use of bioluminescence data, in which values are simply presented in terms of relative light units (RLUs). While the use of RLUs makes comparisons between different publications close to impossible, there are very pragmatic reasons why most authors do not use absolute bioluminescence values - calibration of equipment is difficult, and rarely do absolute units aid in supporting the hypothesis at hand.

Over the course of this dissertation, seminal work done during the 1970's in the laboratory of Dr. Milton Cormier was often referred to [86, 137, 138, 139, 228, 229, 71]. One of the main reasons this body of work has been so useful, is that much of their data was recorded on equipment that had been calibrated in absolute units. While equipment has changed in the past 30 years, the values they published remain accurate and useful for the purposes of comparison.

The main focus of this appendix is to explain how the Turner 20/20 and 20/20n luminometers (Turner Biosystems, Sunnyvale, CA) used in this work were calibrated from RLUs to absolute units (photons/s). Both these luminometers make use of an R1924P photomultiplier tube (PMT) as their light detectors. From looking at the quantum efficiency of this PMT with respect to wavelength (Figure A.1), it's readily apparent that these luminometers have a strong
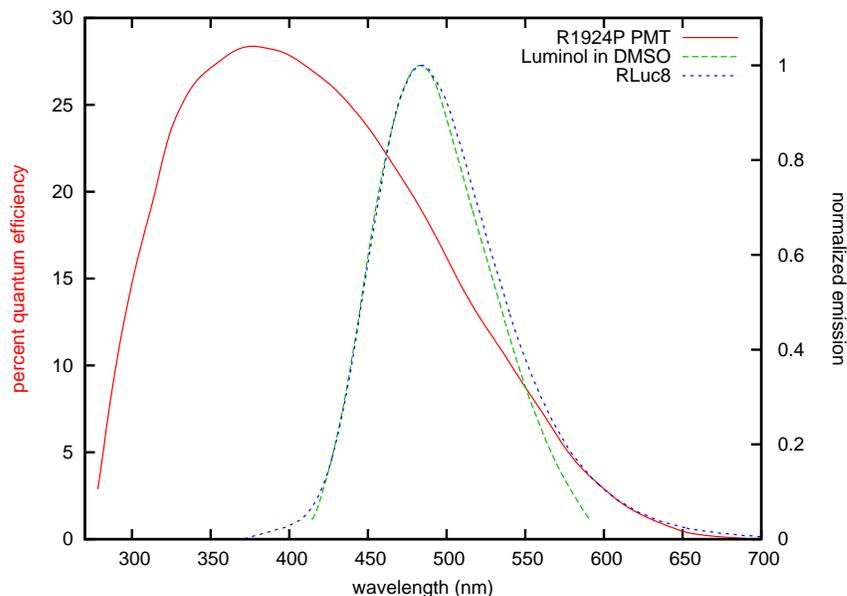
Figure A.1: Quantum efficiency of the R1924P photomultiplier tube, along with the normalized emission spectra of luminol in DMSO and RLuc8. The R1924P quantum efficiency is derived from the Hamamatsu catalog [70]. The luminol chemiluminescence spectrum data is taken from Lee and Seliger [109], and is for luminol dissolved in DMSO and activated using *tert*-butoxide. The RLuc8 emission spectrum is repeated from Figure 3.8.

spectral dependence to their sensitivity. For this reason, the decision was made to calibrate using the luminol light standard performed in dimethyl sulfoxide (DMSO) [109, 110, 165].

This particular luminol based light standard was developed because in DMSO, the chemiluminescence emission spectrum of luminol (Figure A.1) is very similar to the bioluminescence emission spectra of many marine luciferases (e. g. *Renilla* luciferase). Because of this similarity in emission spectra, a correction for the PMT's wavelength dependent sensitivity is not absolutely required in calculating an accurate calibration factor. An additional advantage of this calibration method, is that the geometry of the standard luciferase assay (100 μl in an Eppendorf tube) can be matched exactly when performing the calibration. Although the refraction index of the DMSO/butanol solution is different than the aqueous solution used for the luciferase assay, the proximate geometry employed in the Turner luminometers means that the correction for refractive index is close to unity [165].

This appendix also includes some additional information on the luciferase assay employed through most of the text, as a guide for those who would like to replicate these results. Finally, some additional calibration factors (e. g. conversions between a Turner 20/20 luminometer and

an IVIS 100) that were made use of during the dissertation are also included, although these are probably of little value beyond the current work.

A.1 Methods

A.1.1 Materials

Potassium *tert*-butoxide was ordered alternatively from Sigma, Aldrich, or Alfa Aegar. Luminol (3-aminophthalhydrazide) and all other chemicals were from Sigma. In cases where a luciferase was used for calibrations, the 8 mutation stabilized variant of *Renilla* luciferase (RLuc8) described in Chapter 2 was used.

A.1.2 Equipment

Both a Turner 20/20, and later a Turner 20/20n luminometer (Turner Designs, Sunnyvale, CA) were used in this dissertation. Both these units make use of a Hamamatsu R1924P PMT operating in photon counting mode. The R1924P has a bialkali photocathode with a spectral response range from 300-650 nm (Figure A.1). Its peak quantum efficiency of 26% is reached at 390 nm [70].

A.1.3 Luminol Standard

The luminol standard performed here is closely based off of a previously published protocol [109, 110, 165]. All solutions were made immediately before use, with great care taken to ensure that they were maintained anhydrous. Please note that DMSO is especially hygroscopic.

To ensure dryness, anhydrous DMSO was incubated under argon with molecular sieves (Type 4A) for 24 hrs before use. A small quantity of luminol was dissolved in this DMSO, the $OD_{359.5}$ was measured, and the solution was diluted to a final $OD_{359.5}$ of 0.0001. A saturated solution of potassium *tert*-butoxide in warm, dry *tert*-butanol was made and kept under argon.

The injectors of the luminometer were well flushed with anhydrous DMSO immediately prior to use to ensure that any traces of water in the injector lines were cleared. 2 μl of the saturated potassium *tert*-butoxide/*tert*-butanol solution was pre-mixed with 50 μl of anhydrous DMSO and placed in a tube inside the luminometer. An additional 50 μl of DMSO containing the luminol was then injected into the tube, and the output of the luminometer was recorded

for 17 min. A background reaction was also recorded by using DMSO without luminol for the injection.

The calibration constant was derived by integrating the output of the luminometer (average RLUs multiplied by acquisition time), subtracting the equivalent value from the background recording, and making use of the knowledge that 1 ml of $OD_{359.5}$=1 luminol should emit 9.75×10^{14} quanta [109].

A.1.4 Standard Luciferase Assay

Before assaying, all luciferases were diluted in nickel affinity chromatography elution buffer (EB: 300 mM NaCl, 250 mM imidazole, 20 mM HEPES, pH 8) containing 1% human serum albumin (HSA) in order to bring them into range of the luminometer. As the amounts of protein being handled are extraordinary tiny, the presence of a carrier protein (e. g. 1% HSA) is absolutely critical to prevent loss of luciferase from adsorption onto the surface of the dilution containers.

Luciferase activity was measured by adding 1 μl of sample (diluted as necessary) to 100 μl room temperature 100 mM sodium phosphate buffer (pH 7) [195, 68], adding 1 μl of 0.5 μg/μl coelenterazine or analog, manually mixing, and reading for 10 s in the luminometer. Manual mixing is more convenient than using the auto-injectors, and it has been found to give more consistent results both here as well as for other authors [137]. The time between the addition of the luciferin and the start of measurement was approximately 4 s. Samples were always diluted in duplicate, with both dilutions measured and averaged. Obtained values were corrected for background levels of luminescence (primarily due to albumin [33]) unless otherwise indicated. The use of sodium phosphate buffer for this assay is due to unknown historical reasons in our laboratory. Note that after mixing the protein and sodium phosphate buffer the final solution has 2.5 mM imidazole and 0.01% HSA, both of which are concentrations at which *Renilla* luciferase shows maximum activity [137].

Imidazolopyrazine compounds, such as coelenterazine and its analogs, are photolabile and readily oxidize [33, 178]. As such, care should be taken in their handling and storage. Coelenterazine was dissolved at 0.5 mg/ml in propylene glycol and stored at -80°C in small, single use aliquots. Propylene glycol proved much more repeatable than ethanol or methanol for *in vitro* assay use, as due to the small volumes handled evaporation of the alcohols during pipetting was a major problem.
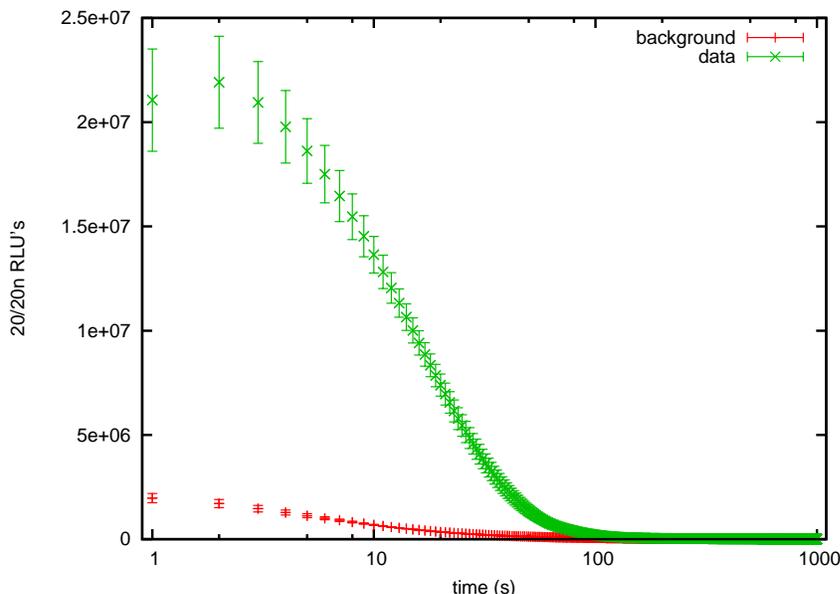
Figure A.2: An example of luminol calibration data. Preparation and injection of the solutions was as described in the methods section. Background data was measured in triplicate, luminol data was measured in quadruplicate, and error bars are standard error of the mean (SEM).

A.2 Results and Discussion

A.2.1 Calibration Factor for the Luminometers

The calibration factor determined for the Turner 20/20n was 12.9$\pm$0.9 (photons/s/Turner 20/-20n RLU). By making use of the conversion factor described below in Section A.2.2, the corresponding calibration factor for the Turner 20/20 was calculated as 18000$\pm$1200 (photons/s/-Turner 20/20 RLU). An example of the acquired data used to calculate this calibration factor is shown in Figure A.2.

Since the chemiluminescence spectrum of luminol and the bioluminescence spectrum of RLuc8 are so similar, no corrections were applied for the spectral sensitivity in determining the calibration factor. Using the quantum efficiency data from Figure A.1, it can be estimated that the difference from not including this information in the calibration factor calculation is at most 3%.

As a side note, quantum efficiency in the context of a PMT actually refers to the ratio of electrons released from the photocathode to photons that are incident on the detector. More useful for luminometers is the concept of counting efficiency (also called detection efficiency), which in a photon counting device refers to the ratio of recorded photons to incident photons.

	Mean Emission Peak (nm)	Compensation Factor
RLuc8 with bisdeoxycoelenterazine	416	0.6
RLuc/A54P	482	0.9
RLuc	497	1.0
RLuc8	503	1.0
RLuc8/A123S/D162L/I163V	515	1.2
RLuc8/F261W	524	1.3
RLuc8/A123S/D162E/I163L	538	1.5
RLuc8/A123S/D162E/I163L/V185L	545	1.6
RLuc8/A123S/D154M/E155G/D162E/I163L/V185L	550	1.7
RLuc8/A123S/D154K/E155N/D162E/I163L/F261W	560	2.0
RLuc8/A123S/D154A/E155G/D162E/I163V/F262W	564	2.0

Table A.1: Compensation factors for the spectral sensitivity of the luminometer. These compensation factors were estimated by assuming that the counting efficiency of the Turner luminometers is directly proportional to the quantum efficiency of their photomultiplier tubes. The factors are expressed as multipliers to the reported specific activities.

Counting efficiency depends in part on the PMT characteristics (like quantum efficiency) as well as the associated electronics. Sadly, Turner to date has not made counting efficiency data available for their devices. In lieu of this information, the quantum efficiency has been used instead as it is the primary factor in the device's wavelength dependent sensitivity.

Many of the luciferases described in Chapter 3 have emission spectra significantly shifted from that of RLuc. Additionally, some measurements in Chapters 2 and 3 were made using bisdeoxycoelenterazine, a coelenterazine analog with an emission peak typically ∼80 nm blue shifted from that of the native substrate. Rough estimates of the compensations that would be required in these cases are given in Table A.1, and are based on the PMT quantum efficiency values in the absence of the more appropriate counting efficiency data. The reported bisdeoxycoelenterazine specific activity and quantum yield values in Chapters 2 and 3 make use of the 0.6 compensation factor for this substrate. The reported specific activity values for the various color shifted mutants, however, do not.

Previous protocols [109, 110, 165] injected the DMSO/luminol solution directly onto the potassium *tert*-butoxide/*tert*-butanol, as opposed to premixing the *tert*-butoxide/*tert*-butanol with DMSO before the DMSO/luminol injection as was done here. The reason for this change in protocol was that implementing the previously published methods yielded a brief but extraordinarily high background of signal that overwhelmed the Turner 20/20n's PMT. It was suggested by Dr. Dennis O'Kane, that the effect may be due to a singlet oxygen produced by the potassium *tert*-butoxide giving rise to a red light emission ∼630 nm (personal communication). He has encountered a similar problem in his work, but was able to overcome it by
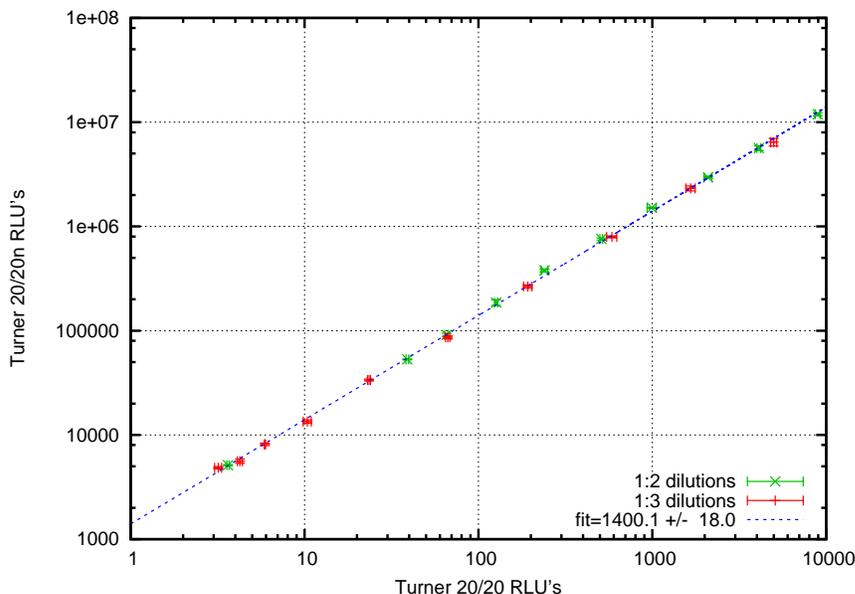
Figure A.3: Conversion between the Turner 20/20 and 20/20n Luminometers. Serial dilutions of RLuc8 were assayed in triplicate on both luminometers, and the data was fit using a linear function. The lowest data point contained no luciferase, and represents background levels of activity primarily from albumin. Error bars are SEM.

ordering several lots of potassium *tert*-butoxide from various manufacturers and identifying a batch that did not give rise to this effect. All the lots ordered for the work here, however, did give off this background signal, necessitating the premixing step as well as subtracting the background chemiluminescence from the luminol chemiluminescence signal.

A.2.2 Conversion between the Turner 20/20 and 20/20n Luminometers

During the course of the dissertation, the Turner 20/20 luminometer in the laboratory was upgraded to a Turner 20/20n. This switch necessitated a calibration between the two machines. Serial 1:2 and 1:1 dilutions of RLuc8 were made and measured using the luciferase assay on both the Turner 20/20 and the Turner 20/20n. This data was not corrected for background levels of luminescence. By fitting the data, as shown in Figure A.3, the scaling factor was determined to be 1400±18 Turner 20/20n RLUs/Turner 20/20 RLUs.

Note that in this dissertation, at times when values are expressed in RLUs, it is always the Turner 20/20 RLUs that are being used. In cases where the measurements were actually made on the Turner 20/20n, the units were first converted to those of the Turner 20/20 using the above scale factor.

A.2.3   Linearity of the Luciferase Assay on the Turner 20/20n

The linearity of both the luciferase assay and the Turner 20/20n was assessed by measuring serial 1:2 dilutions of RLuc8, with the results shown in Figure A.4. The luciferase assay was linear over the first 5 logarithmic decades, and only deviated from the expected value in the highest decade of the luminometer's range. In turn, linearity over the Turner 20/20's smaller range is implied by the data presented in Figure A.3.

The reason for the positive deviation in the Turner 20/20n's highest decade is not understood. As the luminometer is reporting more light than expected, it is not likely that this effect is due to a problem with the luciferase assay at high concentrations of RLuc8. The deviation could possibly arise if the Turner 20/20n electronics were inappropriately overcompensating for the PMT dead time at this high flux level. In any case, all assays done for this dissertation were diluted so as to avoid the last logarithmic decade of this luminometer's range. Even without this upper decade of range, the Turner 20/20n has almost two logarithmic decades greater range than the Turner 20/20.

A.2.4   Conversion between the IVIS 100 and the Turner 20/20

Conversion factors, determined for translating between Turner 20/20 RLUs and the IVIS 100 units, are shown in Figure A.5 along with the data from which they were calculated. These conversion factors were utilized for normalizing the results of the cell culture experiments performed at UCLA (measured in the IVIS 100) to the initial amount of RLuc or RLuc/C124A fusion protein present in the conditions (measured with the Turner 20/20). As the methods used for calibrating the IVIS machines vary between institutions, and in fact at UCLA varied between when the RLuc and the RLuc/C124A based studies were done, these results are probably not applicable beyond the experiments in Chapter 6 for which they were used.

In the later experiments that utilized RLuc8 based probes, a separate control plate containing the different experimental media conditions was used. This plate would be imaged in the IVIS system at the same time as the cell culture plates, and the resulting values were used to control for the amount of protein applied to the cell culture plates. This method proved to be much more reproducible than normalizing based on the luminometer measurements.
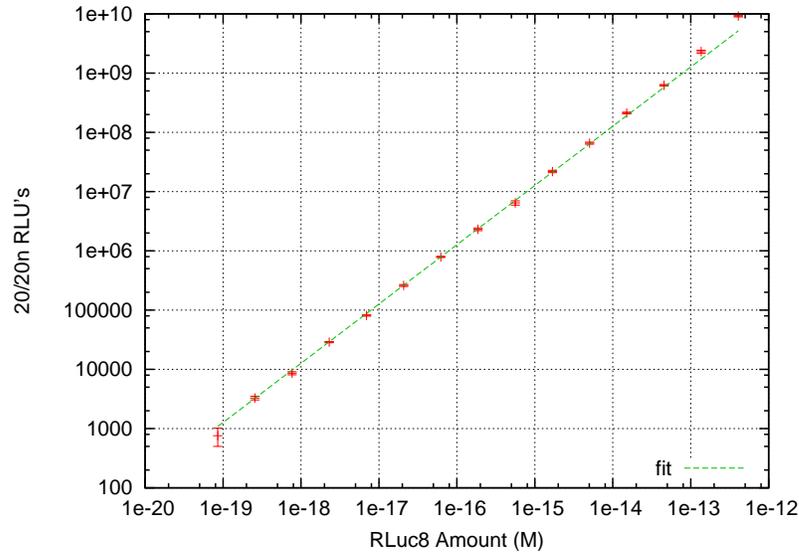
(a) Assay Linearity
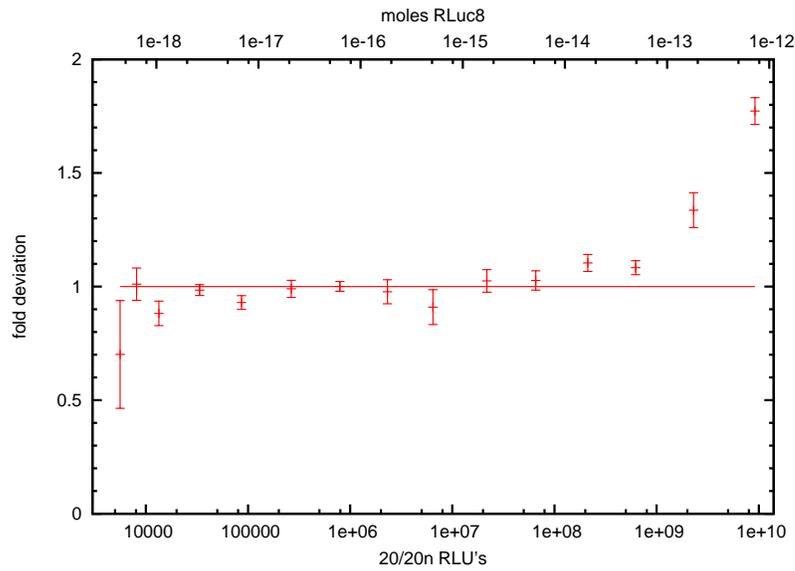
(b) Fold Deviation

Figure A.4: Testing the linearity of the luciferase assay on the Turner 20/20n luminometer. Serial 1:2 dilutions of RLuc8 were assayed on the Turner 20/20n and measured in triplicate. Background luminescence was subtracted from the values. After excluding the two highest values, a fit was made between the amount of protein and the measurement. Panel (a) shows the value obtained (in Turner 20/20n RLUs) versus the total amount of luciferase in the assay. Panel (b) shows the same information graphed as the fold deviation from the expected value based on the fit. The range of protein assayed was from 85 zeptomoles (51,000 molecules) to 410 femtomoles. At 85 zeptomoles, the value was significantly above background (p=0.03) using a one-tailed t-test. Error bars are SEM.
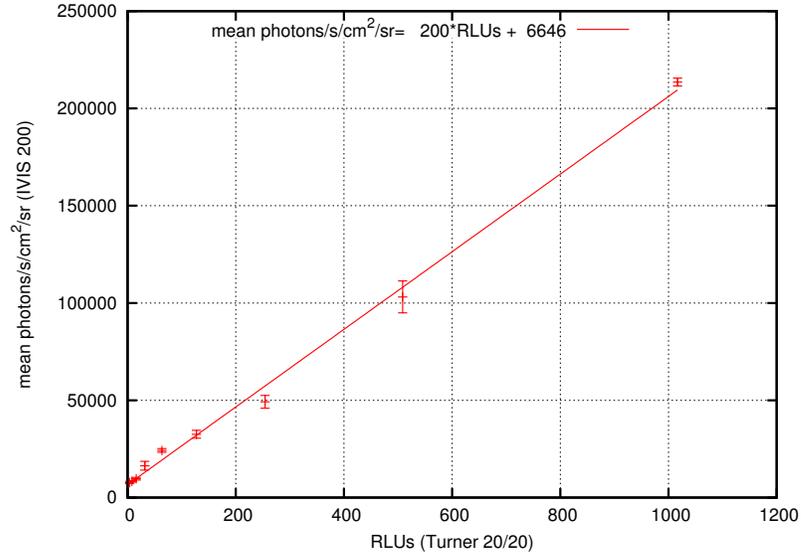
(a) Conversion for RLuc based experiments
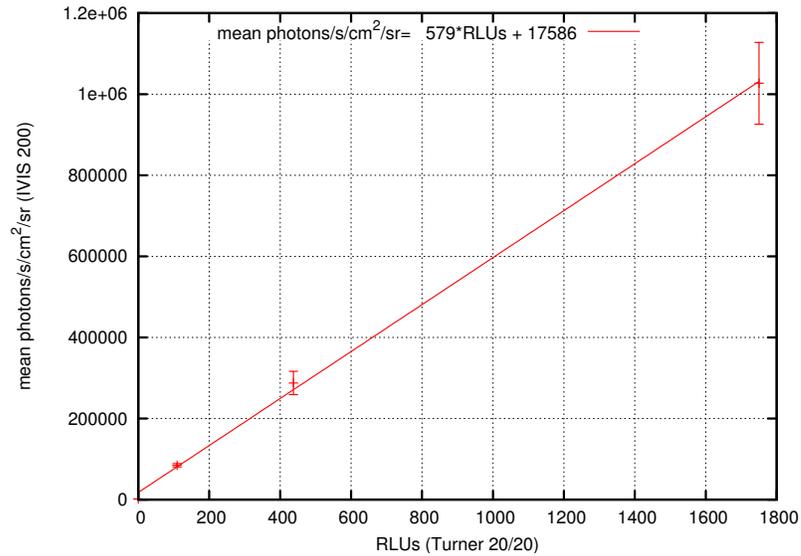
(b) Conversion for RLuc/C124A based experiments

Figure A.5: Conversions between the Turner 20/20 luminometer and the IVIS 100 imaging system. RLuc (Panel (a)) or RLuc/C124A (Panel (b)) was assayed in triplicate on the Turner 20/20, and serial dilutions were made. These dilutions were placed in triplicate on a 24 well plate with 500 μl of PBS/well. 0.5 μg coelenterazine/well was added and the plate was imaged for 60 s in the IVIS 100. A fit to the data was made using a linear function. Differences in the calculated conversion factors between Panel (a) and (b) are due to changes in the IVIS 100 setup between when the RLuc experiments were done and when the RLuc/C124A experiments were performed. Error bars are SEM.

A.3 Concluding Remarks

At this point, a couple comments are in order as to the quantitativeness of the values reported in this dissertation. First, a systemic variation in the measurements of luciferase activity that has not been taken into account is temperature variations encountered in the laboratory. The laboratory in the James Clark Building is essentially a barn, with both seasonal and diurnal variations in temperature present. However, based on previously published data [137], the temperature fluctuations would not be expected to add a variability of greater than 15%.

The greater limitation is that the luminometers were only calibrated once, and these calibration factors were used throughout the work presented in this dissertation. As the Turner luminometers operate their PMTs in photon counting mode, the recorded values are relatively more resistant to voltage fluctuations and changes in PMT sensitivity than if the PMTs were operated in current mode [70]. However, changes in sensitivity of the PMT with age (called "life characteristics") and long term voltage fluctuations will still lead to some level of change in the luminometer's counting efficiency. Turner Biosystems in fact gives two reasons for not calibrating their luminometers; individual PMTs exhibit significant variability that would necessitate a separate calibration of every machine produced, and the poor long term stability of the PMT counting efficiency would make the calibration ineffective in the long run (Personal communication from Emily Matheu, Turner Biosystems).

For future researchers, a better method that has been suggested several times before would be to standardize the luminometer once with luminol, and then use a radioactive standard for day to day calibration [223]. An appropriate radioactive standard could be made in the lab [164], or alternatively purchased from a commercial supplier (e.g. Glowell Microplate Standards from LUX Biotech, Edinburgh, UK).

Appendix B

Expression, Purification, and Structure of *Renilla Reniformis* GFP

The work presented in this appendix is the result of a collaboration with Dr. Timothy D. Fenn and the laboratory of Dr. Axel Brunger. Dr. Fenn assisted with the crystallization screens, acquired the X-ray diffraction data, and calculated the structures. Data interpretation was done jointly.

Early studies on partially purified *Renilla* Luciferase (RLuc) showed that the bioluminescence emission spectrum measured *in vitro* ($\sim$480 nm peak) did not match what was observed from the intact animal ($\sim$505 nm peak). A contaminant was found in some of these initial RLuc purifications that could account for the observed *in vivo* bioluminescence emission, and this contaminant was originally thought to be a luciferase bound chromophore [224]. Later studies showed that this "chromophore" was in fact a separate, chromophore containing protein. This protein was named *Renilla reniformis* Green Fluorescent Protein (RrGFP), and it was first fully purified and characterized in the late 1970's [228, 229].

Although RrGFP has not been studied as thoroughly as the GFP from *Aequorea victoria* (AvGFP), it is safe to assume from the level of primary sequence similarity (50%) that they function in essentially the same manner. The chromophore in these fluorescent proteins, *p*-hydroxybenzylidene-imidazolidone, is produced by a post-translational modification

of a serine-tyrosine-glycine sequence through an auto-catalyzed cyclization, dehydration, and oxidation. For AvGFP at 22°C, this maturation process occurs with a time constant of ~4 h [76].

In *Renilla reniformis*, RLuc is always found in close cellular association with RrGFP [5, 196] and it is the RrGFP that accounts for the observed ~505 nm peaked *in vivo* emission spectrum. This alteration of the RLuc emission spectrum is achieved through a radiationless process called bioluminescence resonance energy transfer (BRET) [228], in which the energy from the luciferase catalyzed oxidation of coelenterazine is transferred to the RrGFP fluorophore. BRET, a direct analog of the Förster resonance energy transfer [60] that can occur between fluorophores, was originally identified in studies of the photoprotein from *Aequorea victoria* and AvGFP [148].

Unlike a trivial transfer mechanism, where a photon of light released from the luciferase would lead to excitation of the fluorescent protein, BRET involves a direct, radiationless transfer of energy from the excited state coelenteramide to the fluorescent protein's chromophore. Resonance energy transfer is far more efficient than trivial transfer and can have efficiencies of close to 100%, but requires close physical proximity (generally <50 Å [202]) between the energy donor (RLuc) and energy acceptor (RrGFP). In *Renilla*, there is evidence that this physical proximity is achieved via protein-protein interactions between a monomer of RLuc and a dimer of RrGFP [228]. These observed interactions, however, are only strong in presumably non-physiological low ionic strength (<100 mM NaCl) environments. It is debatable whether the observed *in vitro* protein-protein interactions are the true mechanism whereby close proximity between these two proteins is maintained in *Renilla reniformis*.

In any case, RrGFP was expressed and purified with the hope that it might aid in the crystallization of RLuc. Use of RrGFP in attempts to cocrystallize RrGFP and RLuc are discussed in Chapter 4. The work discussed in this appendix covers the bacterial expression, purification, crystallization, and structural determination of RrGFP.

B.1   Methods

B.1.1   Construct

The gene for RrGFP was obtained from the plasmid pUC18-RrGFP (gift of Dr. Bruce Bryan, NanoLight Technology). PCR was used to create an NcoI site straddling the initial methionine, and to add a SalI site 3' to the final codon in the gene. The PCR product was then NcoI/SalI digested and ligated into a correspondingly digested pBAD/Myc-His A plasmid (Invitrogen) to

create the plasmid pBAD-RrGFP, in the process adding a C-terminal 6xHis tag and a stop codon to the construct. The RrGFP protein expressed from this plasmid contains the amino acid sequence VDHHHHHH (arising from the SalI site and 6xHis tag) C-terminal to the last codon in the RrGFP sequence, and has an expected size of 27.1 kDa for the monomer.

B.1.2 Protein Production and Purification

Cytoplasmic expression and purification of RrGFP was done identically to the method used for cytoplasmic expression of RLuc8 described in Section 4.1.3, with the exception that the culture was grown at $24°C$ and the time after induction was increased to 24 h.

B.1.3 Structural Determination

A homology model of RrGFP was created using SWISS-MODEL [185] and the crystal structures of similar fluorescent proteins from the corals *Montipora efflorescens*, *Favia favus*, and a *Discosoma* species (PDB IDs 1MOV, 1MOU, 1XSS, and 1GGX). The loop regions of this model were then removed, and the resulting structure was used to bootstrap the phasing via molecular replacement. Based on the expected molecular weight of RrGFP and the volume of the crystal cell, Matthews calculations suggested 4 protomers in the asymmetric unit [136]. All four were found during the molecular replacement search using Phaser [141].

All other computational methods used were identical to what has already been described in Section 4.1.6, although 4-fold NCS restraints were used in the case of RrGFP given the presence of 4 protomers in the asymmetric unit

To appropriately model the chromophore in RrGFP, idealized coordinates for *p*-hydroxy-benzylidene-imidazolidone were derived from HIC-Up database entry CRO [97, 98].

B.2 Results and Discussion

B.2.1 Protein Production

One nicety of RrGFP purification, is that the level of purity attained can easily be monitored by the ratio of the protein's OD_{498} to OD_{280}. Higher ratio values imply greater purity and/or a greater percentage of RrGFP in which the fluorophore has matured. The previously reported ratio for pure RrGFP is $OD_{498}/OD_{280}=5.6$ [229].

An initial expression culture was attempted by incubating at 32°C for 16 h following induction. Nickel affinity purification of the lysed cell pellet under native conditions yielded 22 μg protein/ml culture with an OD_{498}/OD_{280} of 2.2. Later cultures utilized a 24 h incubation at 24°C post-induction, and following nickel affinity purification yielded 100 μg protein/ml culture with an OD_{498}/OD_{280} of 2.3. This protein was further purified by anion exchange chromatography, with roughly 60% of the protein eluting as a single peak at a NaCl concentration of 100 mM. This peak had an OD_{498}/OD_{280} of 5.6. Most likely, the increase in the OD_{498}/OD_{280} ratio was due primarily to maturation of the protein's fluorophore, as only minor contaminant peaks were noted in the chromatography trace. The protein was further purified by gel filtration chromatography, and eluted at a position appropriate for the dimer species. 80% of the protein was recovered from this gel filtration step, with an OD_{498}/OD_{280} of 5.8.

Bioluminescence resonance energy transfer with RrGFP was checked by mixing the protein with purified luciferase in 10 mM NaCl, 10 mM HEPES pH 7.4, and observing the wavelength of light emitted following the addition of coelenterazine (Figure B.1). At the 4 μM concentration of RrGFP dimer used, the dimers are on average separated by 750 Å. As BRET requires short distances (generally <50 Å), the emission spectra observed could only occur if the luciferase and RrGFP were interacting. Note that the RLuc8+RrGFP spectrum contains a larger shoulder in the blue wavelengths than the S3RLuc+RrGFP spectrum. This is presumable due to a weaker interaction between RLuc8 and RrGFP, and may be due to the mutations introduced into RLuc8. Alternatively, it may simply be the 6xHis tag present on this construct (absent on S3RLuc) interfering with the interaction.

B.2.2 Crystallography

RrGFP was screened for crystallization conditions, and found to crystallize easily in a number of different mother liquors. The results of these crystallization trials are shown in Figure B.2. Interestingly, in many of these conditions the crystals appeared extremely rapidly. As an extreme example, the 0.1 M MES pH 6.5, 12% w/v PEG 20,000 condition formed crystals before the setting up of the crystallization tray had been completed.

RrGFP will in fact slowly crystallize on its own when stored in 10 mM NaCl, 10 mM HEPES pH 7.4. After a matter of months, a precipitate that forms in the storage tube can be shown to consist of small, poorly formed crystals capable of polarizing light. Because of this tendency to crystallize, freshly prepared protein is best for crystallography. If older protein is used, too many nucleation sites are present and the crystals grown will be of small size. The 0.1 M MES
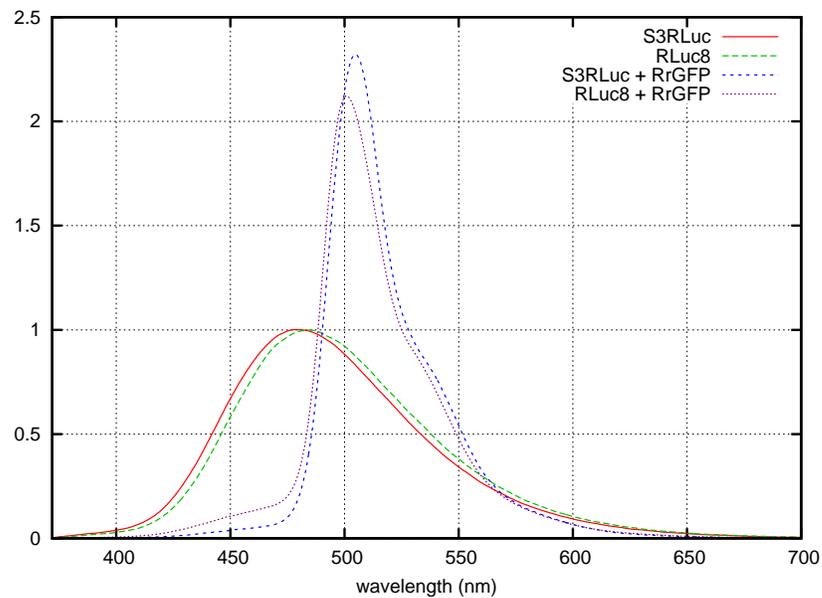
Figure B.1: Bioluminescence emission spectra of luciferase and luciferase mixed with RrGFP. A mixture of 2 μM luciferase (cytoplasmically expressed RLuc8 or S3RLuc) and 4 μM of RrGFP dimer (a 2 fold excess) in 100 μl 10 mM NaCl, 10 mM HEPES, pH 7.4 was mixed with 0.5 μg coelenterazine. The emission spectrum was recorded as described in Section 2.1.6. Both luciferases are described in Chapter 4. The peak for S3RLuc+RrGFP is at 505 nm. The RLuc8+RrGFP peak is at 501 nm.
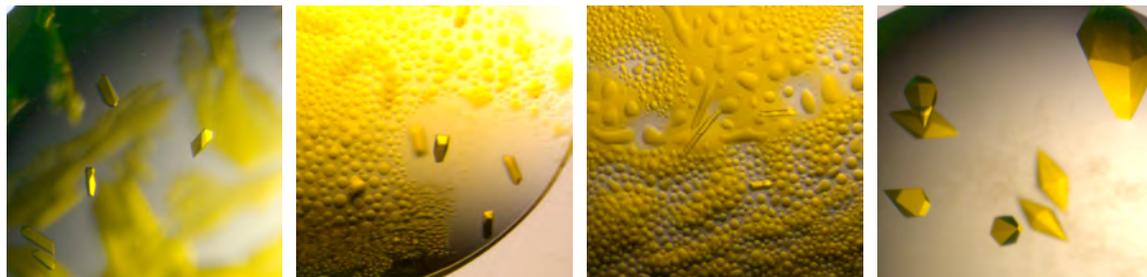
(a) 0 days, 0.1 M MES pH 6.5, 12% w/v PEG 20,000

(b) 1 day, 0.1 M HEPES pH 7.5, 10% w/v PEG 8000, 8% v/v ethylene glycol

(c) 1 day, 0.1 M HEPES pH 7.5, 20% w/v PEG 10,000

(d) 2 days, 2.0 M ammonium sulfate, 5% v/v isopropanol
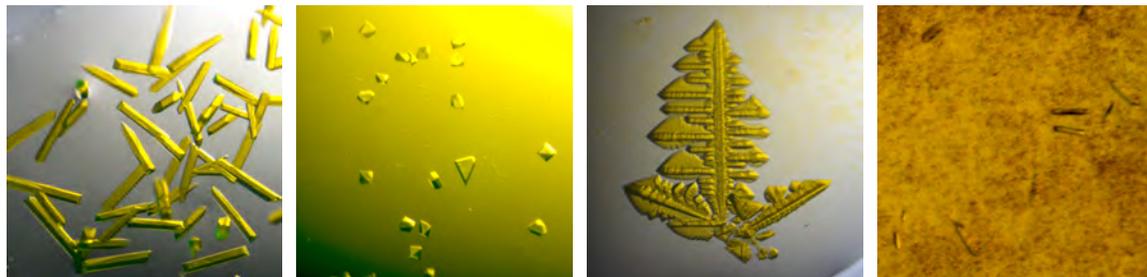
(e) 4 days, 0.1 M HEPES pH 7.5, 10% w/v PEG 6000, 5% v/v MPD

(f) 4 days, 25% v/v ethylene glycol

(g) 1 week, 0.1 M bicine pH 9.0, 2% v/v dioxane, 10% w/v PEG 20,000

(h) 2 weeks, 0.1 M HEPES pH 7.5, 70% v/v MPD
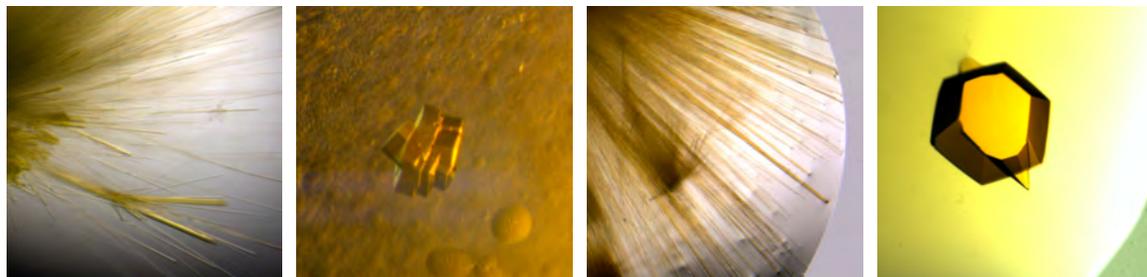
(i) 2 weeks, 0.1 M sodium acetate trihydrate pH 4.6, 2.0 M ammonium sulfate

(j) 1.5 months, 0.1 M tri-sodium citrate dihydrate pH 5.6, 20% v/v isopropanol, 20% w/v PEG 4000

(k) 4 months, 2.0 M ammonium sulfate

(l) 6 months, 1.5 M ammonium sulfate, 0.1 M Tris pH 8.5, 12% v/v glycerol anhydrous

Figure B.2: Results of screening RrGFP for crystallization conditions. Freshly purified RrGFP (48 mg/ml) was used for screening of crystallization conditions using Crystal Screen and Crystal Screen II (Hampton Research) at 20°C. The mother liquor condition and the length of time between set up of the tray and taking of the image are shown below each photograph.

.

Cell Parameters	
space group	P 2_1 2_1 2
a, b, c dimensions (Å)	73.744, 85.416, 158.433
α, β, γ angles	90.00, 90.00, 90.00
Data Collection Statistics	
number of reflections	135,330
possible reflections	145,458
completeness	93%
Model Statistics	
resolution range (Å)	50-1.55
R$_{free}$ (%)	35.2
R$_{working}$ (%)	30.7

Table B.1: Statistics for the crystallographic structure of RrGFP. The crystallographic condition was 0.1 M HEPES pH 7.5, 10% w/v PEG 6000, 5% v/v MPD, and the corresponding structure is shown in Figure B.3. The cross-validation statistic R$_{free}$ was computed from a randomly chosen subset (5%) of the diffraction data that had been excluded from the model refinement process [26]. R$_{working}$ was calculated as $\sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$.

pH 6.5, 12% w/v PEG 20,000 condition was particularly sensitive to this effect.

Crystals from several of the conditions were tried for X-ray crystallography. Although the 2.0 M ammonium sulfate, 5% v/v isopropanol condition gave excellent appearing crystals (Figure B.2d), the particularly long C axis of the crystal cell ($\sim$250 Å) required a significant loss of diffraction resolution in order to keep the diffraction spots properly separated for the purposes of peak integration. In the end, the best diffraction data was obtained with the 0.1 M HEPES pH 7.5, 10% w/v PEG 6000, 5% v/v MPD condition (Figure B.2e) using an additional v/v 35% MDP as the cryoprotectant. The resulting structure from this data is presented in Figures B.3, with the corresponding statistics given in Table B.1. A comparison of RrGFP and AvGFP is shown in Figure B.4.

In the obtained structures of RrGFP, the initial 6 and last 7 amino acids of the primary sequence could not be located in the electron density. The tertiary fold pattern seen is the classic β-barrel characteristic of the fluorescent proteins (e. g. AvGFP).

RrGFP is known to exist as a dimer [229], and through accessible surface area (ASA) calculations this dimerization interface can be demonstrated between the protomers A and B (equivalently C and D) in Figure B.3. This calculation was made using the Protein-Protein Interaction Server [93, 92], which utilizes an implementation of the algorithm of Lee *et al.*
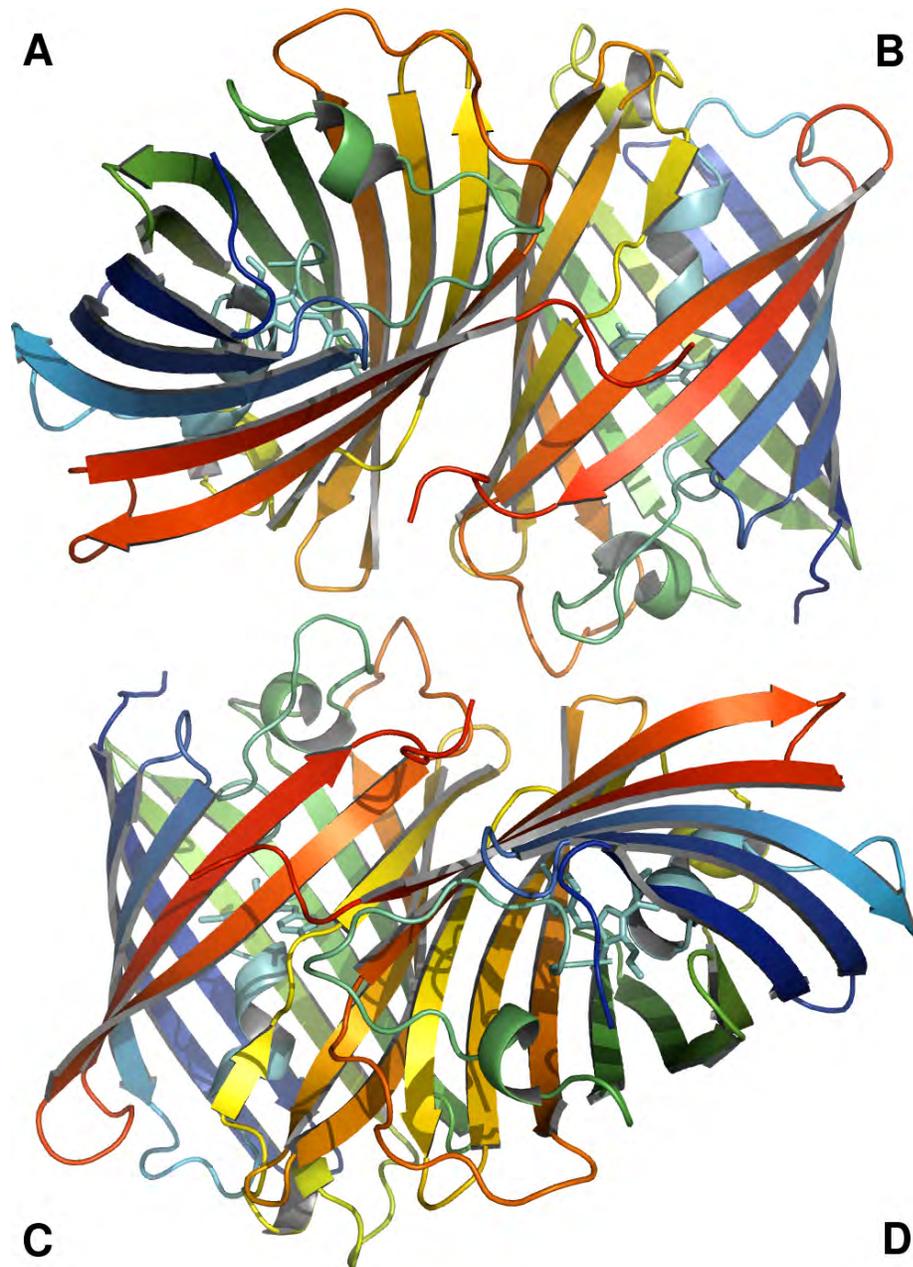
Figure B.3: Cartoon representation of a single unit cell of the RrGFP crystal. The four protomers in each unit cell are labeled A-D. For each protomer, its N-terminus (N) is shown in blue and its C-terminus (C) is shown in red. The mother liquor for this condition was 0.1 M HEPES pH 7.5, 10% w/v PEG 6000, 5% v/v MPD (Figure B.2e). Residues from 7-226 (of 233 total) were identified in the data.

(a) AvGFP

(b) RrGFP
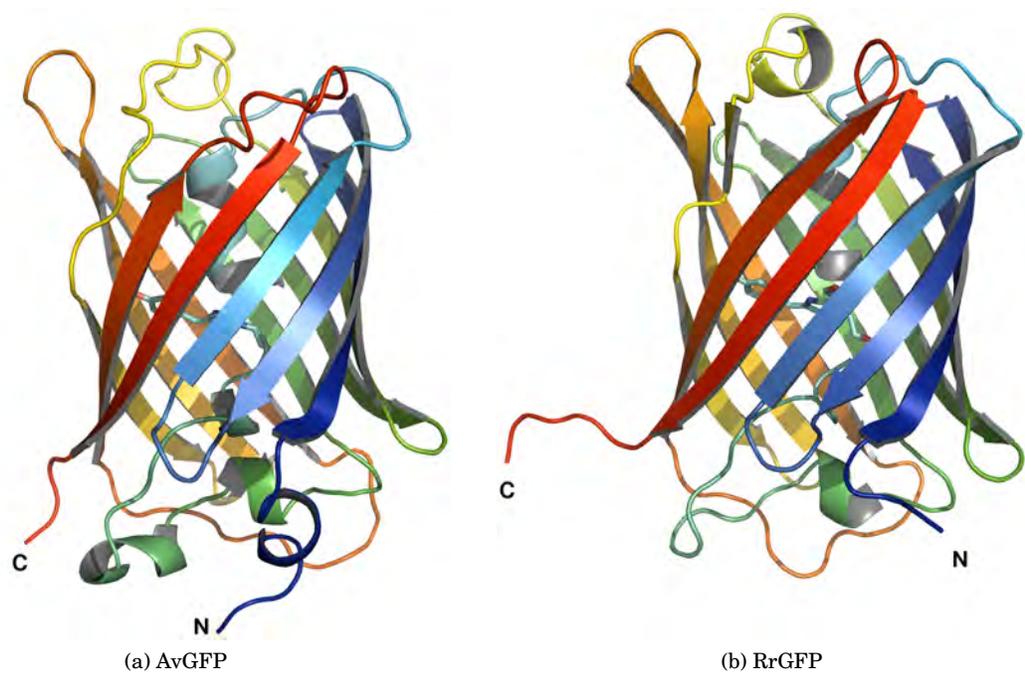
Figure B.4: Cartoon representations of the GFPs from *Aequorea victoria* and *Renilla reniformis*. The N-termini (N) are shown in blue, with the C-termini (C) shown in red. The molecule at the center of the β-barrel is the fluorophore. The primary sequences of the two GFPs are 28% identical and 50% similar.

to calculate accessible surface area [108] and the program HBPLUS to calculate hydrogen bonds [143]. The results were that the A-B (or C-D) interface contained an interface ASA of 1316 Å^2 and 14 hydrogen bonds, appropriate values for a dimerization interface [93]. In contrast, the other interface in the crystallographic unit (A-C, or B-D) had an interface ASA of 151 Å^2 and no hydrogen bonds.

B.3   Conclusion

In conclusion, RrGFP was readily expressed, purified, and crystallized. Unsurprisingly, due to the close primary sequence similarity between RrGFP and AvGFP, the resultant structure for RrGFP is very similar to the previously known structure for AvGFP.

Appendix C

Bacterial Expression and Analysis of *Gaussia* Luciferase

Gaussia princeps, shown in Figure C.1, is a marine copepod (a shrimp-like crustacean). It resides at a depth of ∼800 m in the daytime, and migrates vertically in the water column to ∼300 m at night. When disturbed, the animal ejects a bolus of brilliant blue bioluminescence that is most likely used as a glowing decoy when fleeing predators. The luciferase from *Gaussia* responsible for this bioluminescence was first cloned in 1999 [205], and is now commercially available from NanoLight Technologies as well as in a mammalian codon optimized form from New England Biolabs.

Recently, a closely related luciferase (68% similar) has been described from the copepod *Metridia longa* [134]. This is not surprising, as both *Metridia longa* and *Gaussia princeps* are from the *Metridinidae* family. Interestingly, a luciferase from a species of *Pleuromamma* (GenBank ID GI:12621056), also in the *Metridinidae* family, does not appear to be related.

Gaussia luciferase (GLuc) has attracted interest as a reporter gene in mammalian cell studies as it may be a more sensitive reporter than firefly or *Renilla* luciferase [206]. Its use as a reporter gene for *in vivo* imaging purposes, however, may be limited as GLuc is a secreted protein.

For the purposes of a bioluminescent label, GLuc would appear to have many favorable properties. At 20 kDa (18 kDa for the mature protein) it is the smallest currently known luciferase [205], and it is resistant to inactivation from heat as well as acidic and basic conditions [11]. This stability may be due to the number of cysteines in the molecule, which likely

Figure C.1: A photograph of *Gaussia princeps*. Actually size of the animal is ∼1 cm. This photograph is courtesy of Dr. Bruce Bryan.

form disulfide bonds [205].

As things turn out, GLuc is extremely challenging to produce in bacterial expression systems. This recalcitrance to bacterial expression is the primary reason it was not utilized for the bioluminescent labeling experiments done in this dissertation (Chapter 6). As one can imagine, with 10 cysteine residues in the mature protein there are many possible combinations of disulfide bonds that are incorrect. Another annoyance with GLuc, as will be encountered below, is that it is much more sensitive to the composition of the reaction buffer than *Renilla* luciferase.

This appendix contains some analysis of the GLuc primary sequence, and details the various techniques that were attempted for expressing the protein in bacteria. As a review of bacterial protein expression techniques is not covered here, the interested reader is referred to review articles by Makrides [131] and Baneyx [12], as well as the product manuals from Qiagen's nickel affinity resin [176] and Novagen's pET expression system [160].

C.1 Methods

C.1.1 Materials

M9 minimal medium was made using the standard protocol [182] with the carbon source consisting either of glucose or arabinose.

C.1.2   Periplasmic Constructs

The gene for *Gaussia* luciferase was a generous gift from Dr. Bruce Bryan. This construct had been partially codon optimized for the purposes of mammalian expression. The construct also had the predicted N-terminal signal peptide domain (first 17 amino acids) replaced with a single methionine, and as a consequence is referred to as truncated GLuc (tGLuc) in the text.

PCR was used to replace the methionine codon of *tgluc* with two codons (met/ala) containing an NcoI site, and to add a HindIII site to the 3' end of the gene. The NcoI/HindIII digested PCR product was then inserted into the backbone from an NcoI/HindIII digested pBAD-pelB-RLuc (Section 2.1.4), to create pBAD-pelB-tGLuc-Myc. The Myc epitope, along with a C-terminal 6xHis tag and stop codon, were added by the ligation into the expression plasmid. The pBAD-pelB-tGLuc construct, which does not contain the Myc epitope, was similarly made by using the pBAD SalI site instead of the HindIII site. The construct pBAD-pelB-GLuc was made by using PCR to reattach a DNA sequence corresponding to the missing signal peptide.

Using PCR, an N-terminal 6xHis tagged, thrombin cleavable version was made by replacing the initial methionine codon of *tgluc* with a sequence containing an NcoI site, an alanine, a 6xHis tag, and a thrombin site (LVPR/GS). A PmeI site and stop codon were added 3' to the gene as well. The PCR product was NcoI/PmeI digested, and ligated into a correspondingly digested pBAD-pelB backbone to create pBAD-pelB-A-6H-thr-tGLuc.

In the hopes of pursuing NMR studies of GLuc, additional constructs based on the pET42 vector (Novagen, Darmstadt, Germany) were made, as the pBAD based system is not practical to use for stable isotope labeling experiments. This is due to a combination of two reasons. First, ^{13}C labeled L-(+)-arabinose is not available commercially so the minimal media experiments would need to be done with a different carbon source and a *E. coli* strain deficient in arabinose utilization. Second, all *E. coli* strains that are deficient for arabinose metabolism are unable to grow on minimal media.

PCR was used on pBAD-pelB-A-6H-thr-tGLuc to add an NdeI site 5' to the pelB, and switch the PmeI site C-terminal to the stop codon to an AvrII site. The product was then NdeI/AvrII digested and ligated into a similarly digested pET42 backbone to make pET42-pelB-A-6H-thr-tGLuc. The constructs pET42-pelB-A-6H-3C-P2tGLuc, pET42-pelB-A-tGLuc-thr-6H, and pET42-pelB-A-thr-tGLuc-thr-6H were made in similar fashion by appropriate use of PCR. The "3C" indicates a protease site (LEVLFQ/GP) for the HRV-3C protease. P2tGLuc indicates that the initial residue of tGLuc (K1) has been replaced, in this case by a glycine, and the native sequence begins with P2.

C.1.3   Cytoplasmic Constructs

An initial cytoplasmic expression plasmid, pBAD-tGLuc, was made by using PCR to replace the pelB sequence in pBAD-pelB-tGLuc with a methionine codon.

Additional plasmids based on the pET32 vector (Novagen) were constructed in order to express GLuc cytoplasmically. The pET32 plasmids are designed to allow fusing the protein of interest to a thioredoxin (Trx) protein. The purpose of the Trx is two fold. First, Trx is highly soluble, and this can lead to enhanced solubility of the fusion partner. Second, Trx appears to catalyze disulfide bond formation when used in *E. coli* strains that have an oxidizing cytoplasmic environment due to mutations in thioredoxin reductase (trxB) and glutathione reductase (gor) [174, 200, 160]. The ability to catalyze disulfide bonds is presumably of benefit for the expression of GLuc.

Using pET42-pelB-A-6H-thr-tGLuc and pET42-pelB-A-6H-3c-P2tGLuc as templates, PCR was used to add a BglII site 5' to the protease restriction site and a HindIII site 3' to the stop codon in these constructs. Following BglII/HindIII digestion of the PCR product, the constructs were ligated into a correspondingly digested pET32b(+) plasmid. The resulting plasmids were titled pET32-trx-6H-s-thr-P2tGLuc and pET32-trx-6H-s-3c-P2tGLuc, where "s" designates an s affinity tag that was not used.

C.1.4   Protein Production

Protein production was done in either LMG194, BL-21, or Rosetta-gami B (DE3) cells (Novagen). The Rosetta-gami cells are K-12 derivatives that have oxidizing cytoplasmic environments due to mutations in both trxB and gor [174]. Negative mutations of the trxB/gor genes generally reduce the growth rate of the cells considerably. Rosetta-gami cells, however, are derivatives of the FA113 strain that was developed to incorporate the trxB/gor mutations and still maintain a considerable growth rate [19]. In addition to the oxidizing environment in their cytoplasm, the Rosetta-gami cells also have supplemental tRNAs for codons that are rarely used in *E. coli*.

The medium used for protein expression was always Terrific Broth (TB) unless otherwise indicated. The medium contained the appropriate antibiotics for the given plasmid/cell line combination (e. g. ampicillin for pBAD/LMG194, ampicillin/kanamycin/chloramphenicol/tetracycline for pET32/Rosetta-gami B). The cells were grown in the medium at 37°C until the culture obtained an OD_{600} of 0.7. At this point, the cells were induced (0.2% arabinose for

pBAD based vectors, 1 mM IPTG for pET based vectors) and generally grown for 24 hr at 32°C unless stated otherwise.

For periplasmic expression of the pelB containing plasmids, the cells were centrifuged and submitted to the osmotic shock/nickel affinity chromatography steps that have already been described in Section 2.1.5. In some cases in which minimal media was used for the culture, protein was purified directly from the culture supernatant. In these cases, following centrifugation to remove the cells, the supernatants were brought to 20 mM imidazole from a 400 mM stock, 0.2 μm filtered, and then purified by nickel affinity chromatography.

For cytoplasmic expression plasmids (pET32 constructs), the cells were centrifuged and frozen, with the pellets submitted to lysis and nickel-affinity chromatography as has already been described in Section 4.1.3. Following nickel-affinity chromatography, the protein was cleaved using the appropriate protease (1 μg calf α-thrombin per mg protein, or 10 U HRV-3C per mg protein) overnight at 4°C, and further purified by gel filtration chromatography on a Sephacryl S-100 column (GE Healthcare) using 10 mM NaCl, 10 mM HEPES pH 7.4 as the running buffer. Anion exchange chromatography was attempted, but GLuc has insufficient surface charge to bind the Source 15Q resin (GE Healthcare) that was used.

MALS on purified protein was done as explained in Section 4.1.4.

C.1.5 Assay

The luciferase assay was performed as described in Section A.1.4, with the exception that a different assay buffer was used. GLuc activity was found to be highly buffer dependent. The buffer originally used for assays was Buffer D (10 g/L Roche blocking reagent in 0.1 M maleic acid, 0.15 M NaCl, 1 mM EDTA, pH7.5) [220]. The assay was later switched over to using phosphate buffered saline (PBS) with 0.01% Tween-20, as this yields ∼7-fold more light than the use of Buffer D.

Bioluminescence spectrophotometry was measured as described in Section 2.1.6.

C.2 Results

C.3 Primary Sequence Analysis

A sequence alignment between GLuc and *Metridia longa* luciferase (MLuc) is shown in Figure C.2. A signal peptide sequence along with its corresponding cleavage site were predicted

at a high level of confidence for both luciferases through the use of the SignalP 3.0 server [17]. As an aside, the luciferase from *Pleuromamma* was not predicted to have a signal peptide.

In addition to the signal peptide prediction, the sequences were also checked for internal duplications using the Radar server [74]. The results, as indicated in Figure C.2, pointed to a likely region of duplication in both GLuc and MLuc. Each of these duplicated domains contains 5 cysteines, 4 of which line up exactly to cysteines in the corresponding domain. An alignment of the duplicated domain from GLuc is shown in Figure C.3.

C.4   Periplasmic Constructs

For the periplasmic expression constructs, the thought was that replacing the eukaryotic signal peptide sequence with a prokaryotic version would allow efficient expression of the protein into the oxidative environment of the bacterial periplasm. However, recovered activity from the periplasmic expression of the pBAD-pelB-tGLuc construct was low. The emission spectrum of this protein was recorded, and is shown in Figure C.4.

To test the effects of the missing eukaryotic signal peptide, a construct was created containing the full length GLuc sequence and the pelB sequence preceding the eukaryotic signal peptide. Expression of this plasmid (pBAD-pelB-GLuc) resulted in recovered activity values an order of magnitude worse than from the corresponding tGLuc construct.

With the idea that inappropriate or incorrect disulfide bond formation may be limiting the activity of the periplasmically expressed tGLuc, a series of 10 tGLuc point mutants were created with each containing a single cysteine to alanine mutation. All these variants led to large drops in the recovered activity from the bacterial periplasm, with the exception of the C76A mutation that only led to a ∼50% reduction. Interestingly, C76 is the only cysteine in its domain that does not have a cysteine at the corresponding location in the duplicated domain (Figure C.3).

It was noticed that a majority of the activity from the pBAD-pelB-tGLuc expressions was appearing in the culture supernatant, probably indicating that expression of this plasmid was to some extent toxic to the cells [197]. As proteins cannot easily be purified by nickel affinity chromatography from rich media (e. g. TB), due to chelators present in these media stripping the nickel from the resin, the decision was made to try expression in a minimal medium. The LMG194 cells that had been used previously cannot, however, grow in minimal media. For this reason, BL21 *E. coli* cells were used for the minimal medium (M9) protein expression. As
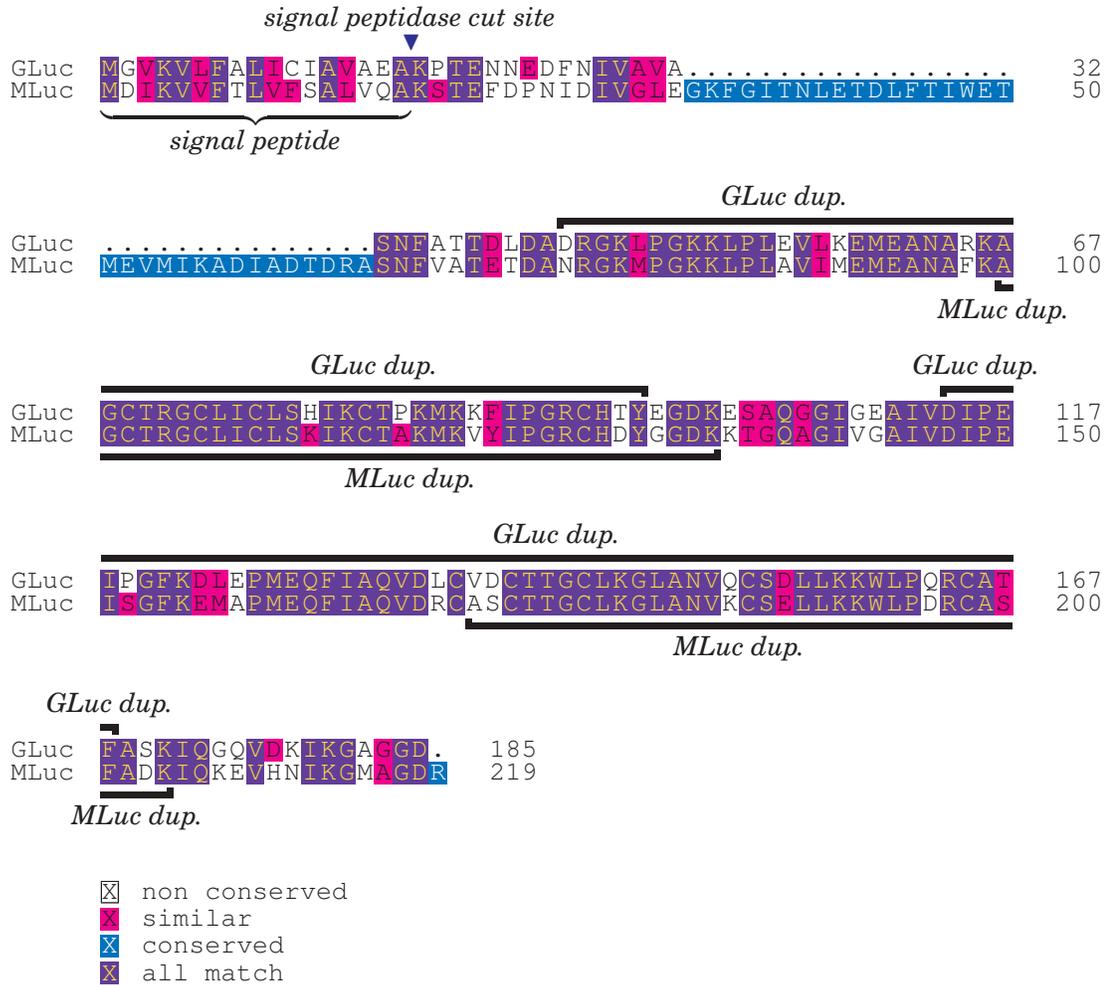
Figure C.2: Aligned sequences for the luciferases from *Gaussia princeps* (GLuc) and *Metridia longa* (MLuc). The signal peptide region and duplicated domains (dup.) were predicted as described in the text. GenBank identifiers for GLuc and MLuc are GI:12621054 and GI:38326874, respectively.

```
GLuc-1    43  DRGKLPGKKLPLEVLKEMEANAR.KAGCTRGCLIC    76
GLuc-2   114  DIPEIPGFK.DLEPMEQFIAQVDLCVDCITGCLKG   147


GLuc-1    77  LSHIKCTPKMKKFIPGRCHTY    97
GLuc-2   148  LANVQCSDLLKKWLPQRCATF   168
```

X non conserved
X similar
X conserved
X all match

Figure C.3: Putative duplicated domain in *Gaussia princeps* luciferase. The duplicated domain was identified using the Radar server as described in the text. GLuc-1 is residues 43-97. GLuc-2 is residues 114-168.
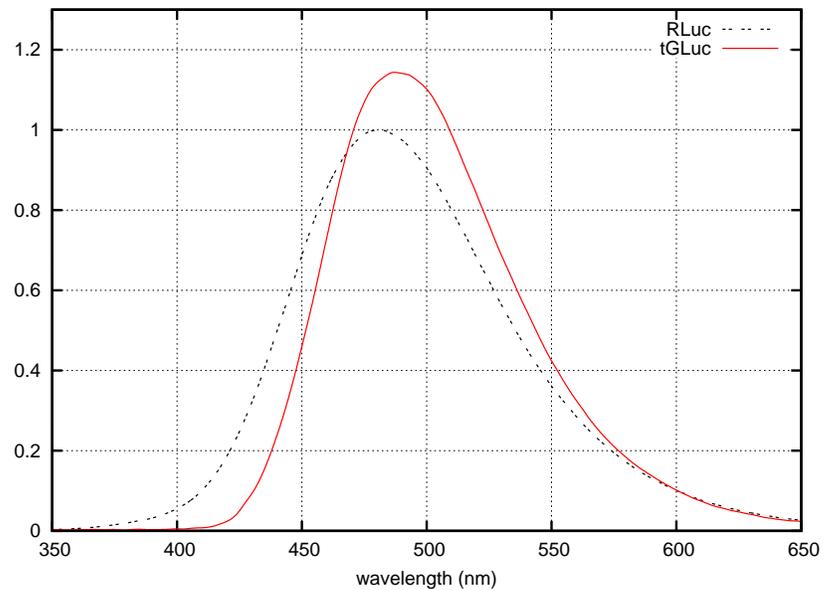
Figure C.4: Normalized bioluminescence emission spectrum for *Gaussia* luciferase. The emission spectrum for *Renilla* luciferase (RLuc) is included for comparison. *Gaussia* luciferase (tGLuc) had a peak emission at 488 nm, and its mean emission wavelength was 506 nm. Coelenterazine was used for the emission spectra, and the curves are normalized such that the total area under the curve is equal.

these cells can utilize arabinose as an energy source, arabinose was used as the carbon source in these experiments instead of glucose. These cultures generated 17 μg luciferase/ml culture, with a specific activity (averaged over 10 s) in Buffer D of 4.1×10^{22} photons/s/mole enzyme. Denaturing gel electrophoresis under reducing conditions of the protein purified from these cultures showed two bands, presumably corresponding to protein with and without proper removal of the pelB signal peptide.

The above experiment was retried with the pBAD-pelB-A-6H-thr-tGLuc construct, in which the pelB and 6xHis tag are separated from tGLuc by a thrombin cleavage site. Expression of this construct was worse (1 μg luciferase/ml culture) and the specific activity was reduced (1.1×10^{22} photons/s/mole enzyme in Buffer D). Additionally, the thrombin site did not appear to be available for cleavage as assessed by gel electrophoresis.

With respect to generating protein labeled with stable isotopes for NMR, the pBAD based protein expression methods are inappropriate, as ^{13}C labeled L-(+)-arabinose does not appear to be commercially available. Even if an isotopically labeled version of this sugar was available, its price would likely be prohibitively expensive. For this reason, it was decided to switch to an IPTG inducible expression plasmid (e. g. pET42) so that ^{13}C labeled glucose could be used as the carbon source.

A total of four pET42 based vectors were created. These constructs were grown in BL-21 cells using M9 minimal medium with glucose as a carbon source at 24°C. Following induction with IPTG, the cultures were grown for 3 days, at which point the activity in the culture supernatants plateaued and the cultures were ended. The results following purification from the culture supernatants are shown in Table C.1. For both the N-terminal 6xHis tagged constructs, only ∼10% of the total activity bound onto the nickel affinity resin, with the remainder coming through on the flow through fraction. For the C-terminal 6xHis tagged constructs, two isoforms were present on denaturing gel electrophoresis under reducing conditions, most likely indicating the presence of protein both with and without the pelB signal peptide.

C.5 Cytoplasmic Constructs

The soluble protein fraction from the pBAD based cytoplasmic construct (pBAD-tGLuc) yielded only ∼0.1 μg of tGLuc/ml culture. This protein had a specific activity of 3.3×10^{21} photons/s/-mole as measured in Buffer D. Performing the cell lysis step in the presence of 8 M urea improved yield to 6.5 μg tGLuc/ml culture, with a specific activity of 5.2×10^{21} photons/s/mole

Plasmid	Yield (μg/ml)	Specific Activity (photons/s/mole)
pET42-pelB-A-6H-tGLuc	1.3	1.3×10^{23}
pET42-pelB-A-6H-3c-P2tGLuc	0.1	2.0×10^{23}
pET42-pelB-A-tGLuc-thr-6H	8.4	2.2×10^{23}
pET42-pelB-A-thr-tGLuc-thr-6H	1.0	1.3×10^{23}

Table C.1: Yield and specific activity of the pET42 based *Gaussia* constructs. Yield represents μg of recovered protein per ml of culture following nickel affinity chromatography. Specific activities were measured in Buffer D.

Plasmid	Step	Yield (μg/ml)	Specific Activity (photons/s/mole)
pET32-trx-6H-s-thr-tGLuc	nickel affinity	28	2.4×10^{23}
	gel filtration	5	3.8×10^{23}
pET32-trx-6H-s-3c-P2tGLuc	nickel affinity	36	1.8×10^{23}
	gel filtration	5	4.6×10^{23}
			(3.2×10^{24} †)

Table C.2: Yield and specific activity of the pET32 based *Gaussia* constructs. Yield represents μg of recovered protein per ml of culture following the given step of purification. Protease digestion of the luciferase from the trx fusion was done between the nickel affinity and gel filtration chromatography steps. Specific activities were measured in Buffer D, except for the entry marked with † that is a remeasure of the gel filtration purified P2tGLuc using PBS with 0.01% Tween-20 as the assay buffer.

in Buffer D. Using 2-mercaptoethanol in place of the 8 M urea led to a further increase in yield to 16 μg tGLuc/ml culture, with a specific activity of 1.4×10^{22} photons/s/mole in Buffer D.

At this point, protein production was switched over to the pET32 based expression plasmids. As explained in the methods, these experiments made use of Rosetta-gami B (DE3) cells allowing cytoplasmic expression in an oxidizing environment. Results for these constructs using TB medium and purifying from the soluble protein fraction are given in Table C.2. The overnight protease digestion at 4°C, as assessed by gel electrophoresis, gave complete cleavage between the luciferase and the trx fusion.

Gel filtration chromatography was performed following protease digestion of the constructs and showed a large amount of protein coming through the column in the void fraction. This protein showed very little luciferase activity, but ran at the correct size for tGLuc on denaturing gel electrophoresis under reducing conditions and most likely represents incorrectly disulfide bonded oligomers of the luciferase. The main peak eluted from the Sephacryl S-100 column at a point earlier than would be expected for a monomer of the luciferase, but was confirmed to

be the monomer for the P2tGLuc condition by light scattering (see below). In addition, a small peak eluted, at a point appropriate for the dimer species, intermediate to the void fraction and the major peak. A bioluminescence emission spectrum was recorded for P2tGLuc, and showed a peak wavelength of 488 nm, a mean wavelength of 503 nm.

Protein size and monodispersity of the P2tGLuc monomeric fraction were confirmed using a Shodex 802.5 gel-filtration column followed by in-line multi-angle light scattering (MALS) and refractive index detectors (DAWN EOS and Optilab DSP, Wyatt Technologies, Santa Barbara, CA). A dn/dc value of 0.185 ml/g was assumed in all calculations, and all processing was performed using the ASTRA software package (Wyatt Technologies). P2tGLuc eluted from the Shodex column at a point appropriate in size for the monomer species. The calculated mass based on the MALS data was 19.6 kDa (actual expected mass 18.1 kDa).

Crystallography screens at 20°C were set up for P2tGLuc at a protein concentration of 13 mg/ml in the presence and absence of 0.1% Tween-20. The screens used were the Index Screen HT and Crystal Screen HT from Hampton Research. An additional Index Screen HT was set up for P2tGLuc at a protein concentration of 32 mg/ml with and without 0.1% Tween-20. After 6 months of observation, none of the conditions produced a protein crystal.

When stored for 3 months at 4°C in 10 mM NaCl, 10 mM HEPES pH 7.4 and at a protein concentration of 13 mg/ml, P2tGLuc was observed to lose 90% of its activity.

C.6 Discussion

Previous attempts at expressing the luciferase from *Gaussia princeps* in both bacterial and yeast expression systems have met with disappointing results (personal communication, Dr. Bruce Bryan). Unsurprisingly, given the sequence similarity, the related luciferase from *Metridia longa* also expresses poorly in bacteria cells [134]. In an attempt to successfully produce *Gaussia* luciferase from bacteria, two expression systems were tried where the protein would be exposed to an oxidizing environment.

The first method tried was expression of GLuc into the bacterial periplasm through the use of the pelB leader sequence. The conclusion from this work, is that the protein is in some way toxic to the cells with the result that the expressed protein ends up being located in the culture supernatant [197]. By purifying from the culture supernatant, small amounts of active protein were recovered, but this protein was heterogeneous due to inconsistent cleavage of the pelB signal peptide.

The second method attempted was intracellular expression in a host strain that presents an oxidizing environment in the cytoplasm. This technique was successful in producing small amounts of homogeneous, highly active GLuc. Work is currently ongoing to isotopically label the protein and obtain an NMR structure of the luciferase.

Bibliography

[1] Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ. Basic local alignment search tool. *J. Mol. Biol.* 215:403–410, 1990.

[2] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, and Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* 25:3389–3402, 1997.

[3] Ames GF, Prody C, and Kustu S. Simple, rapid, and quantitative release of periplasmic proteins by chloroform. *J. Bacteriol.* 160:1181–1183, 1984.

[4] Amin N, Liu AD, Ramer S, Aehle W, Meijer D, Metin M, Wong S, Gualfetti P, and Schellenberger V. Construction of stabilized proteins by combinatorial consensus mutagenesis. *Protein Eng. Des. Sel.* 17(11):787–793, 2004.

[5] Anderson JM and Cormier MJ. lumisomes, the cellular site of bioluminescence in coelenterates. *J. Biol. Chem.* 248:2937–2943, 1973.

[6] Anderson JM and Cormier MJ. Transductive coupling in bioluminescence: Effects of monovalent cations and ionophores on the calcium-triggered luminescence of *Renilla* lumisomes. *Biochem. Biophys. Res. Commun.* 68:1234–1241, 1976.

[7] Araki F, Nakamura H, Nojima N, Tsukumo K, and Sakamoto S. Stability of recombinant human epidermal growth factor in various solutions. *Chem. Pharm. Bull.* 37:404–406, 1989.

[8] Arand M, Wagner H, and Oesch F. Asp333, Asp495, and His523 form the catalytic triad of rat soluble epoxide hydrolase. *J. Biol. Chem.* 271:4223–4229, 1996.

[9] Argent RH, Roberts LM, Wales R, Robertus JD, and Lord JM. Introduction of a disulfide bond into ricin A chain decreases the cytotoxicity of the ricin holotoxin. *J. Biol. Chem.* 269:26705–26710, 1994.

[10] Aristidou AA, Yu P, and San KY. Effects of glycine supplement on protein production and release in recombinant *Escherichia coli*. *Biotechnol. Lett.* 15:331–336, 1993.

[11] Ballou BT, Szent-Gyorgyi C, and Finley G. Properties of a new luciferase from the copepod *Gaussia princeps*. In *Abstracts of the 11th International Symposium on Bioluminescence and Chemiluminescence*. 2000.

[12] Baneyx F. Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* 10:411–421, 1999.

[13] Baneyx F and Mujacic M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol.* 22:1399–1408, 2004.

[14] Baud F and Karlin S. Measures of residue density in protein structures. *Proc. Natl. Acad. Sci.* 96:12494–12499, 1999.

[15] Beilhack A, Schulz S, Baker J, Beilhack GF, Wieland CB, Herman EI, Baker EM, Cao YA, Contag CH, and Negrin RS. *In vivo* analyses of early events in acute graft-versus-host disease reveal sequential infiltration of T cell subsets. *Blood* 206:1113–1122, 2005.

[16] Beitz E. TEXshade: shading and labeling multiple sequence alignments using LATEX2 epsilon. *Bioinformatics* 16:135–139, 2000.

[17] Bendtsen JD, Nielsen H, von Heijne G, and Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 340:783–795, 2004. `http://www.cbs.dtu.dk/services/SignalP`.

[18] Bergès H Joseph-Liauzun E FO. Combined effects of the signal sequence and the major chaperone proteins on the export of human cytokines in *Escherichia coli*. *Appl. Environ. Microbiol.* 62:55–60, 1996.

[19] Bessette PH, Åslund F, Beckwith J, and Georgiou G. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc. Natl. Acad. Sci.* 96:13703–13708, 1999.

[20] Boesen TP, Soni B, Schwartz TW, and Halkier T. Single-chain vascular endothelial growth factor variant with antagonist activity. *J. Biol. Chem.* 277:40335–40341, 2002.

[21] Bond CS. TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics* 19:311–312, 2003.

[22] Bordo D and Argos P. Suggestions for "safe" residue substitutions in site-directed mutagenesis. *J. Mol. Biol.* 217:721–729, 1991.

[23] Bowman EJ, Siebers A, and Altendorf K. Bafilomycins: a class of inhibitors of membrane atpases from microorganisms, animal cells, and plant cells. *Proc. Natl. Acad. Sci.* 85:7972–7976, 1988.

[24] Bradford M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248–254, 1976.

[25] Bréfort J. GChemPaint: a 2D chemical structures editor. on the World Wide Web at `http://www.nongnu.org/gchempaint`.

[26] Brünger AT. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355:472–475, 1992.

[27] Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, and Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallog.* ˚ 54:905–921, 1998.

[28] Bryan B. Detection and visualization of neoplastic tissues and other tissues. United States Patent 6,416,960. Filed August 8, 1997, Awarded July 9, 2002.

[29] Bryan B and Szent-Gyorgyi C. Luciferases, fluorescent proteins, nucleic acids encoding the luciferases and fluorescent proteins and the use thereof in diagnostics, high throughput screening and novelty items. United States Patent 6,232,107. Filed March 26, 1999, Awarded May 15, 2001.

[30] Byrne AM, Bouchier-Hayes DJ, and Harmey JH. Angiogenic and cell survival functions of vascular endothelial growth factor (VEGF). *J. Cell. Mol. Med.* 9:777–794, 2005.

[31] Callard RE and Gearing AJH. *The Cytokine FactsBook*. Academic Press, London, UK, 1994.

[32] Calnan DP, Fagbemi A, Berlanga-Acosta J, Marchbank T, Sizer T, Lakhoo K, Edwards AD, and Playford RJ. Potency and stability of C terminal truncated human epidermal growth factor. *Gut* 47:622–627, 2000.

[33] Campbell AK and Herring PJ. Imidazolopyrazine bioluminescence in copepods and other marine organisms. *Marine Biol.* 104:219–225, 1990.

[34] Chatziioannou AF, Cherry SR, Shao Y, Silverman RW, Meadors K, Farquhar TH, Pedarsani M, and Phelps ME. Performance evaluation of microPET: a high-resolution lutetium oxyorthosilicate PET scanner for animal imaging. *J. Nucl. Med.* 40:1164–1175, 1999.

[35] Cheah KC, Harrison S, King R, Crocker L, Wells JR, and Robins A. Secretion of eukaryotic growth hormones in *Escherichia coli* is influenced by the sequence of the mature proteins. *Gene* 138:9–15, 1994.

[36] Chen J and Stites WE. Packing is a key selection factor in the evolution of protein hydrophobic cores. *Biochemistry* 40:15280–15289, 2001.

[37] Cheong WF, Prahl SA, and Welch AJ. A review of the optical properties of biological tissues. *IEEE J. Quantum Electron.* 26:2166–2185, 1990.

[38] Cheong WF, Prahl SA, Welch AJ, Wang L, and Jacques SL. Optical properties of biological tissues. on the World Wide Web at `http://omlc.ogi.edu/pubs/abs/cheong90a.html`, 1993.

[39] Choi JH and Lee SY. Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 62:625–635, 2004.

[40] Chow PL, Stout DB, Komisopoulou E, and Chatziioannou AF. A method of image registration for small animal, multi-modality imaging. *Phys. Med. Biol.* 51:379–390, 2006.

[41] Clamp M, Cuff J, Searle SM, and Barton GJ. The jalview java alignment editor. *Bioinformatics* 12:426–427, 2004.

[42] Cochran JR, Kim YS, Lippow SM, Rao B, and Wittrup KD. Improved mutants from directed evolution are biased to orthologous substitutions. *Protein Eng. Des. Sel.* 19:245–253, 2006.

[43] Contag CH, Spilman SD, Contag PR, Oshiro M, Eames B, Dennery P, Stevenson DK, and Benaron DA. Visualizing gene expression in living mammals using a bioluminescent reporter. *Photochem. Photobiol.* 66:523–531, 1997.

[44] Coppens A, Sibomana M, Bol A, and Michel C. Mediman: an object oriented programming approach for medical image analysis. *IEEE Trans. Nucl. Sci.* 40:950–955, 1993. URL http://www.topo.ucl.ac.be/iv_mediman.html.

[45] Cormier MJ, Lee J, and Wampler JE. Bioluminescence: recent advances. *Annu. Rev. Biochem.* 44:255–272, 1975.

[46] De A, Loening AM, and Gambhir SS. Novel donor mutation enhances sensitivity for non-invasive imaging of bioluminescence resonance energy transfer signal in living subjects. *Mol. Imaging Biol.* 8:62, 2006.

[47] de Keyzer J, van der Does C, and Driessen AJ. The bacterial translocase: a dynamic protein channel complex. *Cell. Mol. Life Sci.* 60:2034–2052, 2003.

[48] de Kleijn EM and Punt CJ. Biological therapy of colorectal cancer. *Eur. J. Cancer.* 38:1016–1022, 2002.

[49] DeLano WL. The PyMOL molecular graphics system. on the World Wide Web at http://www.pymol.org, 2002.

[50] Deng L, Vysotski ES, Liu Z, Markova SV, Malikova NP, Lee J, Rose J, and Wang B. Structural basis for the emission of violet bioluminescence from a W92F obelin mutant. *FEBS Lett.* 506:281–285, 2001.

[51] Derewenda ZS. Rational protein crystallization by mutational surface engineering. *Structure* 12:529–535, 2004.

[52] Di Fiore PP, Pierce JH, Fleming TP, Hazan R, Ullrich A, King CR, Schlessinger J, and Aaronson SA. Overexpression of the human EGF receptor confers an EGF-dependent transformed phenotype to NIH 3T3 cells. *Cell* 51:1063–1070, 1987.

[53] DuBridge RB, Tang P, Hsia HC, Leong PM, Miller JH, and Calos MP. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol. Cell. Biol.* 7:379–387, 1987.

[54] Eaton JW. *GNU Octave: A High-Level Interactive Language for Numerical Computations*. Third edition. Network Theory, Clifton, UK, 1997.

[55] Emsley P and Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallog.* ° 60:2126–2132, 2004.

[56] Fahey RC, Hunt JS, and Windham GC. On the cysteine and cystine content of proteins. Differences between intracellular and extracellular proteins. *J. Mol. Evol.* 10:155–160, 1977.

[57] Falnes PO, Choe S, Madshus IH, Wilson BA, and Olsnes S. Inhibition of membrane translocation of diphtheria toxin A-fragment by internal disulfide bridges. *J. Biol. Chem.* 269:8402–8407, 1994.

[58] Falnes PO and Sandvig K. Penetration of protein toxins into cells. *Curr. Opin. Cell Biol.* 12:407–413, 2000.

[59] Fominaya J, Uherek C, and Wels W. A chimeric fusion protein containing transforming growth factor-alpha mediates gene transfer via binding to the EGF receptor. *Gene Ther.* 5:521–530, 1998.

[60] Förster VT. Experimentelle und theoretische untersuchung des zwischenmolekularen uebergangs von elektronenanregungsenergie. *Z. Naturforsch.* 4:321–327, 1949.

[61] Free Software Foundation, Inc. GNU general public license, Version 2, 1991. URL `http://www.gnu.org/copyleft/gpl.html`.

[62] French AR, Sudlow GP, Wiley HS, and Lauffenburger DA. Postendocytic trafficking of epidermal growth factor-receptor complexes is mediated through saturable and specific endosomal interactions. *J. Biol. Chem.* 269:15749–15755, 1994.

[63] Fu AY, Chou HP, Spence C, Arnold FH, and Quake SR. An integrated microfabricated cell sorter. *Anal. Chem.* 74:2451–2457, 2002.

[64] Fu H, Blanke SR, Mattheakis LC, and Collier RJ. Selection of diphtheria toxin active-site mutants in yeast. Rediscovery of glutamic acid-148 as a key residue. *Adv. Exp. Med. Biol.* 419:45–52, 1997.

[65] Garrett TP, McKern NM, Lou M, Elleman TC, Adams TE, Lovrecz GO, Kofler M, Jorissen RN, Nice EC, Burgess AW, and CW W. The crystal structure of a truncated ErbB2 ectodomain reveals an active conformation, poised to interact with other ErbB receptors. *Mol. Cell* 11:495–505, 2003.

[66] Garrett TP, McKern NM, Lou M, Elleman TC, Adams TE, Lovrecz GO, Zhu HJ, Walker F, Frenkel MJ, Hoyne PA, Jorissen RN, Nice EC, Burgess AW, and Ward CW. Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell* 110:763–773, 2002.

[67] Gomez GA, Morisseau C, Hammock BD, and Christianson DW. Structure of human epoxide hydrolase reveals mechanistic inferences on bifunctional catalysis in epoxide and phosphate ester hydrolysis. *Biochemistry* 43:4716–4723, 2004.

[68] Gomori G. Preparation of buffers for use in enzyme studies. *Methods Enzymol.* 1:138–146, 1955.

[69] Haddock SH, Rivers TJ, and Robison BH. Can coelenterates make coelenterazine? Dietary requirement for luciferin in cnidarian bioluminescence. *Proc. Natl. Acad. Sci.* 98:11148–11151, 2001.

[70] Hamamatsu Photonics K.K., Hamamatsu City, Japan. *Photomultiplier Tubes, Basics and Applications*, second edition, 1999.

[71] Hart RC, Matthews JC, Hori K, and Cormier MJ. *Renilla reniformis* bioluminescence: luciferase-catalyzed production of nonradiating excited states from luciferin analogues and elucidation of the excited state species involved in energy transfer to *Renilla* green fluorescent protein. *Biochemistry* 18:2204–2210, 1979.

[72] Hart RC, Stempel KE, Boyer PD, and Cormier MJ. Mechanism of the enzyme-catalyzed bioluminescent oxidation of coelenterate-type luciferin. *Biochem. Biophys. Res. Commun.* 81:980–986, 1978.

[73] Head JF, Inouye S, Teranishi K, and Shimomura O. The crystal structure of the photoprotein aequorin at 2.3 Å resolution. *Nature* 405:372–376, 2000.

[74] Heger A and Holm L. Rapid automatic detection and alignment of repeats in protein sequences. *Proteins* 41:224–237, 2000. `http://www.ebi.ac.uk/Radar`.

[75] Heikinheimo P, Goldman A, Jeffries C, and Ollis DL. Of barn owls and bankers: a lush variety of alpha/beta hydrolases. *Structure* 7:R141–146, 1999.

[76] Heim R, Prasher DC, and Tsien RY. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci.* 91:12501–12504, 1994.

[77] Hertel C, Coulter SJ, and Perkins JP. A comparison of catecholamine-induced internalization of beta-adrenergic receptors and receptor-mediated endocytosis of epidermal growth factor in human astrocytoma cells. Inhibition by phenylarsine oxide. *J. Biol. Chem.* 260:12547–12553, 1985.

[78] Hewes CS. *Sea Urchin Egg Bioluminescence: An Intermediate Event of Fertilization and Its Possible Relationship in Mitosis*. Ph.D. thesis, Sonoma State University, 1977.

[79] Hibbert EG and Dalby PA. Directed evolution strategies for improved enzymatic performance. *Microb. Cell Fact.* 4, 2005.

[80] Hill DLG, Batchelor PG, Holden M, and Hawkes DJ. Medical image registration. *Phys. Med. Biol.* 46:R1–R45, 2001.

[81] Holm S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6:65–70, 1979.

[82] Holmquist M. Alpha/beta hydrolase fold enzymes: structure, functions and mechanisms. *Cur. Protein Pept. Sci.* 1:209–235, 2000.

[83] Holyoak T, Fenn TD, Wilson MA, Moulin AG, Ringe D, and Petsko GA. Malonate: a versatile cryoprotectant and stabilizing solution for salt-grown macromolecular crystals. *Acta Crystallogr. D Biol. Crystallog.* ˚ 59:2356–2358, 2003.

[84] Hori K, Nakano Y, and Cormier MJ. Studies on the bioluminescence of *Renilla reniformis*. XI. Location of the sulfate group in luciferyl sulfate. *Biochim. Biophys. Acta* 256:638–644, 1972.

[85] Hori K, Wampler JE, and Cormier MJ. Chemiluminescence of *Renilla* (sea pansy) luciferin and its analogues. *J. Chem. Soc. Chem. Commun.* pages 492–493, 1973.

[86] Hori K, Wampler JE, Matthews JC, and Cormier MJ. Identification of the product excited states during the chemiluminescent and bioluminescent oxidation of *Renilla* (sea pansy) luciferin and certain of its analogs. *Biochemistry* 12:4463–4468, 1973.

[87] Imai Y, Shibata T, Maki S, Niwa H, Ohashi M, and Hirano T. Fluorescence properties of phenolate anions of coelenteramide analogues: the light-emitter structure in aequorin bioluminescence. *J. Photochem. Photobiol. A: Chem.* 146:95–107, 2001.

[88] Inouye S and Shimomura O. The use of *Renilla* luciferase, *Oplophorus* luciferase, and apoaequorin as bioluminescent reporter protein in the presence of coelenterazine analogues as substrate. *Biochem. Biophys. Res. Commun.* 233:349–353, 1997.

[89] Inouye S, Watanabe K, Nakamura H, and Shimomura O. Secretional luciferase of the luminous shrimp *Oplophorus gracilirostris*: cDNA cloning of a novel imidazopyrazinone luciferase. *FEBS Lett.* 481:19–25, 2000.

[90] Jancarik J and Kim SH. sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Cryst.* 24(4):409–411, 1991. doi:10.1107/S0021889891004430. URL `http://dx.doi.org/10.1107/S0021889891004430`.

[91] Jensen AA, Hansen JL, Sheikh SP, and Bräuner-Osborne H. Probing intermolecular protein-protein interactions in the calcium-sensing receptor homodimer using bioluminescence resonance energy transfer (BRET). *Eur. J. Biochem.* 269(20):5076–5087, 2002.

[92] Jones S. Protein-protein interaction server. on the World Wide Web at `http://www.biochem.ucl.ac.uk/bsm/PP/server`, 2003.

[93] Jones S and Thornton JM. Principles of protein-protein interactions derived from structural studies. *Proc. Natl. Acad. Sci.* 93:13–20, 1996.

[94] Jorissen RN, Walker F, Pouliot N, Garrett TP, Ward CW, and Burgess AW. Epidermal growth factor receptor: mechanisms of activation and signalling. *Exp. Cell Res.* 284:31–53, 2003.

[95] Kawamoto T, Sato JD, Le A, Polikoff J, Sato GH, and Mendelsohn J. Growth stimulation of A431 cells by epidermal growth factor: identification of high-affinity receptors for epidermal growth factor by an anti-receptor monoclonal antibody. *Proc. Natl. Acad. Sci.* 80:1337–1341, 1983.

[96] Kimura M and Ota T. On some principles governing molecular evolution. *Proc. Natl. Acad. Sci.* 71:2848–2852, 1974.

[97] Kleywegt G. Hetero-compound information centre HIC-Up. on the World Wide Web at `http://xray.bmc.uu.se/hicup`.

[98] Kleywegt GJ and Jones TA. Databases in protein crystallography. *Acta Crystallogr. D Biol. Crystallog.* ˚ 54:1119–1131, 1998.

[99] Klijn JGM, Berns PMJJ, Schmitz PIM, and Foekens JA. The clinical significance of epidermal growth factor receptor (EGF-R) in human breast cancer: a review of 5232 patients. *Endocr. Rev.* 13:3–17, 1992.

[100] Knuth DE. *Sorting and Searching*, volume 3 of *The Art of Computer Programming*. Addison-Wesley, Reading, MA, 1973.

[101] Ko JK and Ma J. A rapid and efficient PCR-based mutagenesis method applicable to cell physiology study. *Am. J. Physiol. Cell Physiol.* 288(6):C1273–C1278, 2005.

[102] Krupp MN, Connolly DT, and Lane MD. Synthesis, turnover, and down-regulation of epidermal growth factor receptors in human A431 epidermoid carcinoma cells and skin fibroblasts. *J. Biol. Chem.* 257:11489–11496, 1982.

[103] Kumar S, Harrylock M, Walsh KA, Cormier MJ, and Charbonneau H. Amino acid sequence of the Ca2(+)-triggered luciferin binding protein of *Renilla reniformis*. *FEBS Lett.* 268:287–290, 1990.

[104] Kuratorium OFFIS eV. DCMTK - DICOM toolkit. on the Word Wide Web at `http://dicom.offis.de/dcmtk.php.en`.

[105] Kuzmic P. Program DYNAFIT for the analysis of enzyme kinetic data: Application to HIV proteinase. *Anal. Biochem.* pages 260–273, 1996.

[106] Lacroute P and Levoy M. Shear-warp factorization of the viewing transformation. In *Computer Graphics Proceedings, Annual Conference Series*, pages 451–458. ACM SIGGRAPH, Orlando, Florida, 1994. URL `http://graphics.stanford.edu/software/volpack`.

[107] Lamport L. *LaTeX: A Document Preparation System*. Second edition. Addison-Wesley, Reading, MA, 1985.

[108] Lee B and Richards FM. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* 55:379–400, 1971.

[109] Lee J and Seliger HH. Absolute spectral sensitivity of phototubes and the application to the measurement of the absolute quantum yields of chemiluminescence and bioluminescence. *Photochem. Photobiol.* 4:1015–1048, 1965.

[110] Lee J, Wesley AS, Ferguson III JF, and Seliger HH. The use of luminol as a standard of photon emission. In *Bioluminescence in Progress*, editors Johnson FH and Haneda Y, pages 35–43. Princeton University Press, Princeton, New Jersey, 1966.

[111] Lee JR, Madsen MT, Bushnel D, and MENDA Y. A threshold method to improve standardized uptake value reproducibility. *Nucl. Med. Comm.* 21:685–690, 2000.

[112] Lee JY, Yoon CS, Chung IY, Lee YS, and Lee EK. Scale-up process for expression and renaturation of recombinant human epidermal growth factor from *Escherichia coli* inclusion bodies. *Biotechnol. Appl. Biochem.* 31:245–248, 2000.

[113] Lehmann M, Loch C, Middendorf A, Studer D, Lassen SF, Pasamontes L, van Loon APGM, and Wyss M. The consensus concept for thermostability engineering of proteins: further proof of concept. *Protein Eng.* pages 403–411, 2002.

[114] Lei SP, Lin HC, Wang SS, Callaway J, and Wilcox G. Characterization of the *Erwinia carotovora* pelB gene and its product pectate lyase. *J. Bacteriol.* 169:4379–4383, 1987.

[115] Lewis MJ and Pelham HR. Ligand-induced redistribution of a human KDEL receptor from the Golgi complex to the endoplasmic reticulum. *Cell* 68:353–364, 1992.

[116] Li S, Schmitz KR, Jeffrey PD, Wiltzius JJ, Kussie P, and Ferguson KM. Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. *Cancer Cell* 7:301–311, 2005.

[117] Lichtner RB, Rotgeri A, Bunte T, Buchmann B, Hoffmann J, Schwede W, Skuballa W, and Klar U. Subcellular distribution of epothilones in human tumor cells. *Proc. Natl. Acad. Sci.* 98:11743–11748, 2001.

[118] Lim WA, Hodel A, Sauer RT, and Richards FM. The crystal structure of a mutant protein with altered but improved hydrophobic core packing. *Proc. Natl. Acad. Sci.* 91:423–427, 1994.

[119] Liu J and Escher A. improved assay sensitivity of an engineered secreted *Renilla* luciferase. *Gene* 237:153–159, 1999.

[120] Liu J, O'Kane DJ, and Escher A. secretion of functional *Renilla reniformis* luciferase by mammalian cells. *Gene* 203:141–148, 1997.

[121] Liu TF, Cohen KA, Ramage JG, Willingham MC, Thorburn AM, and Frankel AE. A diphtheria toxin-epidermal growth factor fusion protein is cytotoxic to human glioblastoma multiforme cells. *Cancer Res.* 63:1834–1837, 2003.

[122] Liu Z, Vysotski E, Chen C, Rose J, Lee J, and Wang B. Structure of the Ca^{2+}-regulated photoprotein obelin at 1.7 Å resolution determined directly from its sulfur substructure. *Protein Sci.* 9:2085–2093, 2000.

[123] Liu ZJ, Stepanyuk GA, Vysotski ES, Lee J, Markova SV, Malikova NP, and Wang BC. Crystal structure of obelin after Ca^{2+}-triggered bioluminescence suggests neutral coelenteramide as the primary excited state. *Proc. Natl. Acad. Sci.* 103:2570–2575, 2006.

[124] Loening AM and Gambhir SS. AMIDE: a free software tool for multimodality medical image analysis. *Mol. Imaging* 2, 2003.

[125] Loening AM, Paulmurugan R, Wu AM, and Gambhir SS. A novel renilla luciferase/-epidermal growth factor fusion protein as an optical molecular probe for cancer imaging. *Mol. Imaging* 2:132P, 2003.

[126] Loening AM, Wu AM, and Gambhir SS. Alignment guided mutagenesis of renilla luciferase increases stability and light output. *Mol. Imaging* 3, 2004.

[127] Lorenz WW, Gray JP, Cormier MJ, Gibson BG, and O'Kane DJ. Overexpression and purification of recombinant *Renilla reniformis* luciferase. In *Bioluminescence and Chemiluminescence: Status Report*, pages 191–195. 7th International Symposium on Bioluminescence and Chemiluminescence, Banff, Canada, 1993.

[128] Lorenz WW, McCann RO, Longiaru M, and Cormier MJ. Isolation and expression of a cDNA encoding *Renilla reniformis* luciferase. *Proc. Natl. Acad. Sci.* 88:4438–4442, 1991. UniProtKB Entry: P27652.

[129] Lote CJ. *Principles of Renal Physiology*, pages 36–39. Chapman and Hall, London, 1992.

[130] Lu HS, Chai JJ, Li M, Huang BR, He CH, and Bi RC. Crystal structure of human epidermal growth factor and its dimerization. *J. Biol. Chem.* 276:34913–34917, 2001.

[131] Makrides SC. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol. Rev.* 60:512–538, 1996.

[132] Malakauskas SM and Mayo SL. Design, structure and stability of a hyperthermophilic protein variant. *Nat. Struct. Biol.* 5:470–475, 1998.

[133] Marchler-Bauer A, Anderson JB, DeWeese-Scott C, Fedorova ND, Geer LY, He S, Hurwitz DI, Jackson JD, Jacobs AR, Lanczycki CJ, Liebert CA, Liu C, Madej T, Marchler GH, Mazumder R, Nikolskaya AN, Panchenko A, Rao BS, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Vasudevan S, Wang Y, Yamashita RA, Yin JJ, and Bryant SH. CDD: a curated Entrez database of conserved domain alignments. *Nucl. Acids Res.* 31:383–387, 2003.

[134] Markova SV, Golz S, Frank LA, Kalthof B, and Vysotski ES. Cloning and expression of cDNA for a luciferase from the marine copepod *Metridia longa*. *J. Biol. Chem.* 279:3212–3217, 2004.

[135] Massoud TF and Gambhir SS. Molecular imaging in living subjects: seeing fundamental biological processes in a new light. *Genes Dev.* 17:545–580, 2003.

[136] Matthews BW. Solvent content of protein crystals. *J. Mol. Biol.* 33:491–497, 1968.

[137] Matthews JC. *Purification and Characteristics of* Renilla reniformis *Luciferase*. Ph.D. thesis, University of Georgia, 1976.

[138] Matthews JC, Hori K, and Cormier MJ. Purification and properties of *Renilla reniformis* luciferase. *Biochemistry* 16:85–91, 1977.

[139] Matthews JC, Hori K, and Cormier MJ. Substrate and substrate analogue binding properties of *Renilla* luciferase. *Biochemistry* 16:5217–5220, 1977.

[140] Mayer A and S N. Luminescent labels - more than just an alternative to radioisotopes? *Ang. Chemie, Int. Ed. Engl.* 33:1044–1072, 1994.

[141] McCoy AJ, Grosse-Kunstleve RW, Storoni LC, and Read RJ. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D Biol. Crystallog.* ˚ 61:458–464, 2005.

[142] McDonald DM and Baluk P. Significance of blood vessel leakiness in cancer. *Cancer Res.* 62:5381–5385, 2002.

[143] McDonald IK and Thornton JM. Satisfying hydrogen-bonding potential in proteins. *J. Mol. Biol.* 238:777–793, 1994.

[144] Mese H, Sasaki A, Alcalde RE, Nakayama S, and Matsumura T. Establishment and characterization of cisplatin-resistant human epidermoid carcinoma cell line, A431 cell. *Chemotherapy* 44:414–420, 1998.

[145] Michalet X, Pinaud FF, Bentolila LA, Tsay JM, Doose S, Li JJ, Sundaresan G, Wu AM, Gambhir SS, and Weiss S. Quantum dots for live cells, *in vivo* imaging, and diagnostics. *Science* 307:538–544, 2005.

[146] Mitchell M. Engauge digitizer. on the Word Wide Web at `http://digitizer.sf.net`.

[147] Morgan-Kiss RM, Wadler C, and Cronan Jr JE. Long-term and homogeneous regulation of the *Escherichia coli* araBAD promoter by use of a lactose transporter of relaxed specificity. *Proc. Natl. Acad. Sci.* 99:7373–7377, 2002.

[148] Morise H, Shimomura O, Johnson FH, and Winant J. Intermolecular energy transfer in the bioluminescent system of *Aequorea*. *Biochemistry* 13:2656–2662, 1974.

[149] Morishita H, Ohashi S, Oku T, Nakajima Y, Kojima S, Ryufuku M, Nakamura H, and Ohmiya Y. Cloning and characterization of an active fragment of luciferase from a luminescent marine alga, *Pyrocystis lunula*. *Photochem. Photobiol.* 75:311–315, 2002.

[150] Mosesson Y and Yarden Y. Oncogenic growth factor receptors: implications for signal transduction therapy. *Semin. Cancer. Biol.* 14:262–270, 2004.

[151] Mullenbach GT, Chiu CY, Gyenes A, Blaney J, Rosenberg S, Marlowe CK, Brown S, Stratton-Thomas J, Montelione GT, George-Nascimento C, and Stauber G. Modification of a receptor-binding surface of epidermal growth factor (EGF): analogs with enhanced receptor affinity at low pH or at neutrality. *Protein Eng.* 11:473–480, 1998.

[152] Murshudov GN, Vagin AA, and Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallog.* ˚ 53:240–255, 1997.

[153] Nakajima Y, Kobayashi K, Yamagishi K, Enomoto T, and Ohmiya Y. cDNA cloning and characterization of a secreted luciferase from the luminous Japanese ostracod, *Cypridina noctiluca. Biosci. Biotechnol. Biochem.* 68:565–570, 2004.

[154] Nakamura H, Wu C, Kato M, and Kimura M. Syntheses of the mechanism-based inhibitors of coelenterazine bioluminescence. In *11th International Symposium on Bioluminescence and Chemiluminescence*, pages 127–130. International Society for Bioluminescence and Chemiluminescence, Monterey, CA, 2000.

[155] Nandagopal K Popp DM and Niyogi SK. Utilization of a receptor reserve for effective amplification of mitogenic signaling by an epidermal growth factor mutant deficient in receptor activation. *J. Cell. Biochem.* 83:326–341, 2001.

[156] Neu HC and Heppel LA. The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts. *J. Biol. Chem.* 240:3685–3692, 1965.

[157] Newman J, Peat TS, Richard R, Kan L, Swanson PE, Affholter JA, Holmes IH, Schindler JF, Unkefer CJ, and Terwilliger TC. Haloalkane dehalogenases: structure of a *Rhodococcus* enzyme. *Biochemistry* 38:16105–16114, 1999.

[158] Nolf E, Voet T, Jacobs F, Dierckx R, and Lemahieu I. (X)MedCon - an open-source medical image conversion toolkit. *Eur. J. Nucl. Med.* 30:S246, 2003. URL `http://xmedcon.sf.net`.

[159] Nossal NG and Heppel LA. The release of enzymes by osmotic shock from *Escherichia coli* in exponential phase. *J. Biol. Chem.* 241:3055–3062, 1966.

[160] Novagen, Inc., Darmstadt, Germany. *pET System Manual*, tenth edition, 2002.

[161] Oakley AJ, Klvana M, Otyepka M, Nagata Y, Wilce MC, and Damborsky J. Crystal structure of haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* ut26 at 0.95 Å resolution: dynamics of catalytic residues. *Biochemistry* 43:870–878, 2004.

[162] Ogiso H, Ishitani R, Nureki O, Fukai S, Yamanaka M, Kim JH, Saito K, Sakamoto A, Inoue M, Shirouzu M, and Yokoyama S. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* 110:775–787, 2002.

[163] Oh KJ, Senzeldagger L, Colier RJ, and Finkelstein A. Translocation of the catalytic domain of diphtheria toxin across planar phospholipid bilayers by its own T domain. *Biophysics* 96:8467–8470, 1999.

[164] O'Kane DJ and Lee J. Encapsulated radiophosphorescent standards for day-to-day photometer calibration. *Photochem. Photobiol.* 52:723–734, 1990.

[165] O'Kane DJ and Lee J. Absolute calibration of luminometers with low-level light standards. *Methods Enzymol.* 305:87–96, 2000.

[166] Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J, Sussman JL, Verschueren KH, and Goldman A. The alpha/beta hydrolase fold. *Protein Eng.* 5:197–211, 1992.

[167] Pandiella A, Beguinot L, Velu TJ, and Meldolesi J. Transmembrane signalling at epidermal growth factor receptors overexpressed in NIH 3T3 cells. Phosphoinositide hydrolysis, cytosolic Ca^{2+} increase and alkalinization correlate with epidermal-growth-factor-induced cell proliferation. *Biochem. J.* 254:223–228, 1988.

[168] Park J and Gambhir S. Multimodality radionuclide, fluorescence, and bioluminescence small-animal imaging. *Proc. IEEE* 93:771–783, 2005.

[169] Park JM, Loening AM, Wu AM, and Gambhir SS. Bioluminescence imaging of angiogenesis with a novel *Renilla* luciferase-VEGF121 fusion reporter protein. *Mol. Imaging Biol.* 7:149, 2005.

[170] Paulus MJ, Sari-Sarraf H, Gleason SS, Bobrek M, Hicks JS, Johnson DK, Behel JK, Thompson LH, and Allen WC. A new X-ray computed tomography system for laboratory mouse imaging. *IEEE Trans. Nucl. Sci.* 46:558–564, 1999.

[171] Perrakis A, Harkiolaki M, Wilson KS, and Lamzin VS. ARP/wARP and molecular replacement. *Acta Crystallogr. D Biol. Crystallog.* ˚ 57:1445–1450, 2001.

[172] Perrakis A, Morris R, and Lamzin VS. Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* 6:458–463, 1999.

[173] Pichler A, Prior JL, and Piwnica-Worms D. Imaging reversal of multidrug resistance in living mice with bioluminescence: MDR1 P-glycoprotein transports coelenterazine. *Proc. Natl. Acad. Sci.* 101:1702–1707, 2004.

[174] Prinz WA, Åslund F, Holmgren A, and Beckwith J. The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *J. Biol. Chem.* 272:15661–15667, 1997.

[175] Qi J, Leahy R, Cherry S, Chatziioannou A, and Farquhar T. High-resolution 3D bayesian image reconstruction using the microPET small-animal scanner. *Phys. Med. Biol.* 43:1001–1013, 1998.

[176] Qiagen, Inc., Valencia, CA. *The QIAexpressionist: a handbook for high-level expression and purification of 6xHis-tagged proteins*, fifth edition, 2003.

[177] Rees JF, de Wergifosse B, Noiset O, Dubuisson M, Janssens B, and Thompson E. The origins of marine bioluminescence: turning oxygen defense mechanisms into deep-sea communication tools. *J. Exp. Biol.* 201:1211–1221, 1998.

[178] Rees JF and Thompson EM. Photophores: the analysis of bioluminescent systems. In *Biochemistry and molecular biology of fishes, vol. 3*, editors Hochachka PW and Mommsen TP, pages 215–229. Elsevier Science B.V., Amsterdam, the Netherlands, 1994.

[179] Reichardt W, Hu-Lowe D, Torres D, Weissleder R, and Bogdanov Jr A. Imaging of VEGF receptor kinase inhibitor-induced antiangiogenic effects in drug-resistant human adenocarcinoma model. *Neoplasia* 7:847–853, 2005.

[180] Reilly RM, Kiarash R, Sandhu J, Lee YW, Cameron RG, Hendler A, Vallis K, and Gariepy J. A comparison of EGF and MAb 528 labeled with 111In for imaging human breast cancer. *J. Nucl. Med.* 41:903–911, 2000.

[181] Rost B and Liu J. The PredictProtein server. *Nucl. Acids Res.* 31:3300–3304, 2003.

[182] Sambrook J and Russell DW. *Molecular Cloning: A Laboratory Manual*, volume 3, pages A2.2–A2.4. Third edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001.

[183] Schanstra JP and Janssen DB. Kinetics of halide release of haloalkane dehalogenase: evidence for a slow conformational change. *Biochemistry* 35:5624–5632, 1996.

[184] Schlessinger J. Ligand-induced, receptor-mediated dimerization and activation of EGF receptor. *Cell* 110:669–672, 2002.

[185] Schwede T, Kopp J, Guex N, and Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. *Nucl. Acids Res.* 31:3381–3385, 2003.

[186] Shaw JP, Akiyoshi DE, Arrigo DA, Rhoad AE, Sullivan B, Thomas J, Genbauffe FS, Bacha P, and Nichols JC. Cytotoxic properties of DAB486EGF and DAB389EGF, epidermal growth factor (EGF) receptor-targeted fusion toxins. *J. Biol. Chem.* 266:21118–21124, 1991.

[187] Sherf B A Navarro SL, Hannah RR, and Wood KV. Dual-luciferase reporter assay: An advanced co-reporter technology integrating firefly and *Renilla* luciferase assays. Promega Notes Magazine Number 57, p.02, 1996.

[188] Shimomura O, Goto T, and Johnson FH. Source of oxygen in the CO2 produced in the bioluminescent oxidation of firefly luciferin. *Proc. Natl. Acad. Sci.* 74:2799–2802, 1977.

[189] Shimomura O and Johnson FH. Mechanism of the luminescent oxidation of *Cypridina* luciferin. *Biochem. Biophys. Res. Commun.* 44:340–346, 1971.

[190] Shimomura O and Johnson FH. Peroxidized coelenterazine, the active group in the photoprotein aequorin. *Proc. Natl. Acad. Sci.* 75:2611–2615, 1978.

[191] Shimomura O, Musicki B, and Kishi Y. Semi-synthetic aequorin. an improved tool for the measurement of calcium ion concentration. *Biochem. J.* 251:405–410, 1988.

[192] Shimomura O and Teranishi K. Light-emitters involved in the luminescence of coelenterazine. *Luminescence* 15:51–58, 2000.

[193] Siemeister G, Schirner M, Reusch P, Barleon B, Marme D, and Martiny-Baron G. An antagonistic vascular endothelial growth factor (VEGF) variant inhibits VEGF-stimulated receptor autophosphorylation and proliferation of human endothelial cells. *Proc. Natl. Acad. Sci.* 95:4625–4629, 1998.

[194] So MK, Xu C, Loening AM, Gambhir SS, and Rao J. Self-illuminating quantum dot conjugates for in vivo imaging. *Nat. Biotechnol.* 24:339–343, 2006.

[195] Sørensen SPL. Ergänzung zu der abhandlung: Enzymstudien II: Über die messung und die bedeutung der wasserstoffionenkonzentration bei enzymatischen prozessen. *Biochem. Z.* 22:352–356, 1909.

[196] Spurlock BO and Cormier MJ. A fine structure of the anthocodium in *Renilla mülleri*. *J. Cell Biol.* 64:15–28, 1975.

[197] Stader JA and Silhavy TJ. Engineering *Escherichia coli* to secrete heterologous gene products. *Methods Enzymol.* 185:166–187, 1990.

[198] Stallman RM. *Using and Porting GNU CC*. Free Software Foundation, Boston, MA, 1998.

[199] Steipe E, Schiller B, Plückthun A, and Steinbacher S. Sequence statistics reliably predict stabilizing mutations in a protein domain. *J. Mol. Biol.* 240:188–192, 1994.

[200] Stewart EJ, Åslund F, and Beckwith J. Disulfide bond formation in the *Escherichia coli* cytoplasm: an *in vivo* role reversal for the thioredoxins. *EMBO J.* 17:5543–5550, 1998.

[201] Streltsov VA, Prokop Z, Damborsky J, Nagata Y, Oakley A, and Wilce MCJ. Haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* ut26: X-ray crystallographic studies of dehalogenation of brominated substrates. *Biochemistry* 42:10104–10112, 2003.

[202] Stryer L and Haugland RP. Energy transfer: a spectroscopic ruler. *Proc. Natl. Acad. Sci.* 58:719–726, 1967.

[203] Sutanthavibul S, Yap K, Smith BV, King P, Boyter B, and Sato T. Xfig drawing program for the X windows system. on the Word Wide Web at `http://www.xfig.org`.

[204] Suzuki C, Nakajima Y, Akimoto H, Wu C, and Ohmiya Y. A new additional reporter enzyme, dinoflagellate luciferase, for monitoring of gene expression in mammalian cells. *Gene* 344:61–66, 2005.

[205] Szent-Gyorgyi C, Ballou BT, Dagnal E, and Bryan B. Cloning and characterization of new bioluminescent proteins. In *Biomedical imaging: reporters, dyes, and instrumentation*, editors Bornhop DJ, Contag CH, and Sevick-Muraca EM, pages 4–11. Society of Photo-optical Instrumentation Engineers (SPIE), Bellingham, WA, 1999. Proceedings of the SPIE - Vol. 3600.

[206] Tannous BA, Kim DE, Fernandez JL, Weissleder R, and Breakefield XO. Codon-optimized *Gaussia* luciferase cDNA for mammalian gene expression in culture and in vivo. *Mol. Ther.* 11:435–443, 2005.

[207] Teranishi K and Shimomura O. Solubilizing coelenterazine in water with hydroxypropyl-beta-cyclodextrin. *Biosci. Biotechnol. Biochem.* 61, 1997.

[208] Thévenaz P, Blu T, and Unser M. Interpolation revisited. *IEEE Trans. Med. Imaging* 19:739–758, 2000.

[209] Thompson EM, Nagata S, and Tsuji FI. Cloning and expression of cDNA for the luciferase from the marine ostracod *Vargula hilgendorfii*. *Proc. Natl. Acad. Sci.* 86:6567–6571, 1989.

[210] Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, and Higgins DG. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.* 25:4876–4882, 1997.

[211] Thompson JD, Higgins DG, and Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 22:4673–4680, 1994.

[212] Thomson CM, Herring PJ, and Campbell AK. Coelenterazine distribution and luciferase characteristics in oceanic decapod crustaceans. *Marine Biol.* 124:197–207, 1995.

[213] Tong WY, Yao SJ, Zhu ZQ, and Yu J. An improved procedure for production of human epidermal growth factor from recombinant *E. coli*. *Appl. Microbiol. Biotechnol.* 57:674–679, 2001.

[214] Troy T, Jekic-McMullen D, Sambucetti L, and Rice B. Quantitative comparison of the sensitivity of detection of fluorescent and bioluminescent reporters in animal models. *Mol. Imaging* 3:9–23, 2004.

[215] Uherek C, Fominaya J, and Wels W. A modular DNA carrier protein based on the structure of diphtheria toxin mediates target cell-specific gene delivery. *J. Biol. Chem.* 273:8835–8841, 1998.

[216] Venisnik KM. In vivo *optical imaging of tumors expressing carcinoembryonic antigen (CEA) using engineered antibody-luciferase fusion proteins*. Ph.D. thesis, University of California, Los Angeles, 2006.

[217] Venisnik KM, Olafsen T, Kenanova V, Loening AM, Gambhir SS, and Wu AM. *In vivo* optical imaging of tumors expressing carcinoembryonic antigen (cea) using a diabody-*Renilla* luciferase fusion protein. *Mol. Imaging Biol.* 7:160, 2005.

[218] Venisnik KM, Olafsen T, Loening AM, Gambhir SS, and Wu AM. Optimization of expression and purification of a diabody-renilla luciferase fusion protein for localization of CEA expressing tumors. *Mol. Imaging* 3:P282, 2004.

[219] Venisnik KM, Olafsen T, Loening AM, Iyer M, Gambhir SS, and Wu AM. Bifunctional antibody-*Renilla* luciferase fusion protein for *in vivo* optical detection of tumors. *Protein Eng. Des. Sel.* page in press, 2006.

[220] Verhaegen M and Christopoulos TK. Recombinant *Gaussia* luciferase. overexpression, purification, and analytical application of a bioluminescent reporter for DNA hybridization. *Anal. Chem.* 74:4378–4385, 2002.

[221] Viviani VR, Oehlmeyer TL, Arnoldi FG, and Brochetto-Braga MR. A new firefly luciferase with bimodal spectrum: identification of structural determinants of spectral ph-sensitivity in firefly luciferases. *Photochem. Photobiol.* 81:843–848, 2005.

[222] Vysotski ES and Lee J. Ca^{2+}-regulated photoproteins: Structural insight into the bioluminescence mechanism. *Acc. Chem. Res.* 37:405–415, 2004.

[223] Wampler JE. Measurements and physical characteristics of luminescence. In *Bioluminescence in Action*, editor Herring PJ, pages 1–48. Academic Press, New York, 1978.

[224] Wampler JE, Hori K, Lee JW, and Cormier MJ. Structured bioluminescence. Two emitters during both the in vitro and the in vivo bioluminescence of the sea pansy, *Renilla*. *Biochemistry* 10:2903–2909, 1971.

[225] Wang D, Lehman RE, Donner DB, Matli MR, Warren RS, and Welton ML. Expression and endocytosis of VEGF and its receptors in human colonic vascular endothelial cells. *Am. J. Physiol. Gastrointest. Liver Physiol.* 282:G1088–1096, 2002.

[226] Wang J, Reilly RM, Chen P, Yang S, Bray MR, Gariépy J, Chan C, and Sandhu J. Fusion of the CH1 domain of IgG1 to epidermal growth factor (EGF) prolongs its retention in the blood but does not increase tumor uptake. *Cancer Biother. Radiopharm.* 17:665–671, 2002.

[227] Wang L, Jackson WC, Steinbach PA, and Tsien RY. Evolution of new nonantibody proteins via iterative somatic hypermutation. *Proc. Natl. Acad. Sci.* 101:16745–16749, 2004.

[228] Ward WW and Cormier MJ. Energy transfer via protein-protein interaction in *Renilla* bioluminescence. *Photochem. Photobiol.* 27:389–396, 1978.

[229] Ward WW and Cormier MJ. An energy transfer protein in coelenterate bioluminescence. characterization of the *Renilla* green-fluorescent protein. *J. Biol. Chem.* 254:781–788, 1979.

[230] Weiss RF. The solubility of nitrogen, oxygen and argon in water and seawater. *Deep Sea Res.* 17:7221–735, 1970.

[231] Weissleder R. A clearer vision for *in vivo* imaging. *Nat. Biotechnol.* 19:316–317, 2001.

[232] Wesche J, Elliott JL, Falnes PO, Olsnes S, and Collier RJ. Characterization of membrane translocation by anthrax protective antigen. *Biochemistry* 37:15737–15746, 1998.

[233] Wet JRD, Wood KV, Helinski DR, and DeLuca M. Cloning of firefly luciferase cDNA and the expression of active luciferase in *Escherichia coli*. *Proc. Natl. Acad. Sci.* 82(23):7870–7873, 1985.

[234] Wiley HS. Trafficking of the ErbB receptors and its influence on signaling. *Exp. Cell Res.* 284:78–88, 2003.

[235] Wiley HS and Cunningham DD. The endocytotic rate constant. a cellular parameter for quantitating receptor-mediated endocytosis. *J. Biol. Chem.* 257:4222–42229, 1982.

[236] Wilson T and Hastings JW. Bioluminescence. *Annu. Rev. Cell Dev. Biol.* 14:197–230, 1998.

[237] Wirtz P and Steipe B. Intrabody construction and expression III: engineering hyperstable V(H) domains. *Protein Sci.* 8:2245–2250, 1999.

[238] Wittrup KD. Experimental & theoretical analyses of tumor targeting. on the World Wide Web at `http://web.mit.edu/kdw-lab/pdf/IBC_Ab_Engr_05.pdf`, 2005.

[239] Wong JY, Chu DZ, Williams LE, Yamauchi DM, Ikle DN, Kwok CS, Liu A, Wilczynski S, Colcher D, Yazaki PJ, Shively JE, Wu AM, and Raubitschek AA. Pilot trial evaluating an 123I-labeled 80-kilodalton engineered anticarcinoembryonic antigen antibody fragment (cT84.66 minibody) in patients with colorectal cancer. *Clin. Cancer Res.* 10:5014–5021, 2004.

[240] Wood KV, Lam YA, Seliger HH, and McElroy WD. Complementary DNA coding click beetle luciferases can elicit bioluminescence of different colors. *Science* 244:700–702, 1989.

[241] Wu C, Nakamura H, Murai A, and Shimomura O. Chemi- and bioluminescence of coelenterazine analogues with a conjugated group at the c-8 position. *Tetr. Lett.* 42:2997–3000, 2001.

[242] Wu JC, Sundaresan G, Iyer M, and Gambhir SS. Noninvasive optical imaging of firefly luciferase reporter gene expression in skeletal muscles of living mice. *Mol. Ther.* 4:297–306, 2001.

[243] Yegutkin GG, Samburski SS, and Jalkanen S. Soluble purine-converting enzymes circulate in human blood and regulate extracellular ATP level via counteracting pyrophosphatase and phosphotransfer reactions. *FASEB J.* 17(10):1328–1330, 2003.

[244] Yoon JM, Han SH, Kown OB, Kim SH, Park MH, and Kim BK. Cloning and cytotoxicity of fusion proteins of EGF and angiogenin. *Life Sci.* 64:1435–1445, 1999.

[245] Yoshimori T, Yamamoto A, Moriyama Y, Futai M, and Tashiro Y. Bafilomycin A1, a specific inhibitor of vacuolar-type H(+)-ATPase, inhibits acidification and protein degradation in lysosomes of cultured cells. *J. Biol. Chem.* 266:17707–17712, 1991.

[246] Zhao H, Doyle TC, Coquoz O, Kalish F, Rice BW, and Contag CH. Emission spectra of bioluminescent reporters and interaction with mammalian tissue determine the sensitivity of detection *in vivo*. *J. Biomed. Opt.* 10:41210, 2005.

[247] Zhao H, Doyle TC, Wong RJ, Cao Y, Stevenson DK, Piwnica-Worms D, and Contag CH. Characterization of coelenterazine analogs for measurements of *Renilla* luciferase activity in live cells and living animals. *Mol. Imaging* 3:43–54, 2004.

[248] Zheng L, Baumann U, and Reymond J. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucl. Acids Res.* 32:e115, 2004.