

## **UC Irvine**

### **UC Irvine Electronic Theses and Dissertations**

#### **Title**

Expanding the bioluminescent tool box for multicomponent imaging

#### **Permalink**

<https://escholarship.org/uc/item/3f09r9dq>

#### **Author**

Zhang, Brendan Shin-kuei

#### **Publication Date**

2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,  
IRVINE

Expanding the bioluminescent tool box for multicomponent imaging

DISSERTATION

submitted in partial satisfaction of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

in Chemistry

by

Brendan S. Zhang

Dissertation Committee:  
Professor Jennifer Prescher, Chair  
Professor James Nowick  
Assistant Professor Chang Liu

2018

Chapter 1 © Elsevier

Select figures and material from Chapter 2 © John Wiley & Sons, Inc.

All other materials © 2018 Brendan S. Zhang

# DEDICATION

To

My parents, Victoria and Brian Zhang

For their life-long support and love

# TABLE OF CONTENTS

	Page
LIST OF FIGURES	v
LIST OF TABLES	vii
LIST OF SCHEMES	viii
ACKNOWLEDGMENTS	ix
CURRICULUM VITAE	x
ABSTRACT OF THE DISSERTATION	xii
CHAPTER 1: Advances in bioluminescence imaging: New probes from old recipes	
1.1 Introduction	1
1.2 Discovering new luciferases and luciferins	3
1.3 Generating a palette of bioluminescent probes	4
1.4 Engineering orthogonal luciferase-luciferin pairs	7
1.5 Monitoring new facets of biology	11
1.6 Conclusion	13
1.7 Objectives of this study	13
References	14
CHAPTER 2: Pyridone luciferins and mutant luciferases for bioluminescence imaging	
2.1 Introduction	20
2.2 Design and synthesis of luciferin analogs	22
2.3 Bioluminescence assays with luciferin analogs	24
2.4 In cellulo imaging	31
2.5 Improved imaging with mutant luciferases	33
2.6 Directed evolution of luciferase mutant 24	39
2.7 Conclusion	44
2.8 Materials and methods	45
References	61
CHAPTER 3: Pi-extended luciferin and mutant luciferases for multicomponent imaging	
3.1 Introduction	68
3.2 Design and synthesis of pi-extended analogs	70
3.3 Evolving luciferases with improved activity	74
3.4 Characterization of evolved luciferases	78
3.5 Orthogonal imaging with engineered bioluminescent pairs	80
3.6 Conclusion	81
3.7 Materials and methods	82

References	93
CHAPTER 4: Expression and purification of luciferases for X-ray crystallography	
4.1 Introduction	96
4.2 Expression and purification of firefly luciferase	98
4.3 Screening published crystallography conditions	100
4.4 Stabilizing firefly luciferase with ligands	101
4.5 Screening crystallography conditions with ligands	102
4.6 Conclusion	105
4.7 Materials and methods	105
References	109
APPENDIX A: NMR spectra	111

## LIST OF FIGURES

	Page	
Figure 1-1	Luciferase-luciferin pairs in nature	2
Figure 1-2	Expanded palette of bioluminescent probes	6
Figure 1-3	Substrate-selective luciferases for multi-component imaging	9
Figure 1-4	Visualizing cellular species with bioluminescent sensors	12
Figure 2-1	Firefly bioluminescence	20
Figure 2-2	Fluc catalyzes light emission with pyridone analogs	25
Figure 2-3	Bioluminescence spectra at various pH values	27
Figure 2-4	$^1\text{H}$ NMR spectra of tautomeric luciferins 2.3 and 2.12	29
Figure 2-5	$^1\text{H}$ NMR spectra of tautomeric luciferins 2.4 and 2.13	30
Figure 2-6	Absorbance spectra of luciferin analogs	31
Figure 2-7	Bioluminescence imaging in mammalian cells	32
Figure 2-8	Spectral resolution in mammalian cells	33
Figure 2-9	Enzyme kinetics of mutant luciferases	34
Figure 2-10	Improved mutant luciferases with luciferin analogs	35
Figure 2-11	Improved bioluminescence imaging with mutant luciferases	37
Figure 2-12	Luciferase expression was measured in cellular samples	38
Figure 2-13	SCA library screening identifies improved luciferases	39
Figure 2-14	Light output of purified SCA mutants and luciferin analogs	40
Figure 2-15	Thermal denaturation of mutant luciferases	41
Figure 2-16	Bioluminescence spectra of SCA mutant luciferases	43
Figure 3-1	Computational predictions of Pi-extended analogs	72

Figure 3-2	Fluorescence and bioluminescence of pi-extended luciferin analogs	74
Figure 3-3	Screening of site-directed libraries failed to yield an improved mutant	75
Figure 3-4	Computationally-guided library design	77
Figure 3-5	Directed evolution of mutant G1	78
Figure 3-6	Biochemical characterization of mutant G2	79
Figure 3-7	Pi-extended luciferin may adopt a twisted conformation	80
Figure 3-8	Bioluminescence imaging with an orthogonal quartet	81
Figure 4-1	Variances in H-bonding residues affect water network	97
Figure 4-2	Luciferase constructs for crystallography	98
Figure 4-3	SDS PAGE analysis and activity of recombinant luciferases	99
Figure 4-4	LC/MS analysis of recombinant luciferases	100
Figure 4-5	PEG optimization crystallography screens	101
Figure 4-6	Ligands solubilize and stabilize luciferase	102



## LIST OF TABLES

	Page
Table 2-1 Predicted $\Delta E_{\text{HOMO-LUMO}}$ and observed bioluminescent $\lambda_{\text{em}}$ values	23
Table 2-2 Sequencing analysis of “bright” mutants	36
Table 2-3 Sequencing analysis of mutant luciferases	41
Table 4-1 Crystal screens with ligands	104

## LIST OF SCHEMES

	Page
Scheme 2-1 Synthesis of pyridone luciferins	24
Scheme 2-2 Synthesis of tautomerically locked analogs	28
Scheme 3-1 Synthesis of pi-extended luciferins	73

## ACKNOWLEDGMENTS

I would like to thank Jenn for being a great mentor, always available and easy to approach with great scientific insights. My committee members for their advice over the years.

Members of the Prescher lab for their support throughout the years. I would specifically like to thank Dave McCutcheon for mentoring me when I started in the lab. Zi Yao for providing great scientific insights and for helping me with edits and experimental details. Anna Love for helping with the editing of this dissertation. And Sean Ngyuen for helping me edit defense slides.

Collaborators Prof. Jeremy Mills, Prof. Aaron Leconte and Prof. Celia Goulding and their lab personnel for their critical help with my project.

Friends at Calvary Chapel Costa Mesa and Bethel Bible Fellowship for their encouragement.

Financial support was provided by the University of California, Irvine, and the GAANN fellowship.

## CURRICULUM VITAE

### Brendan S. Zhang

Ph.D. Candidate, Department of Chemistry  
University of California, Irvine  
Irvine, CA 92617

Telephone: 732-887-5144  
E-mail: bszhang@uci.edu

### EDUCATION

---

<b>University of California, Irvine</b> , Irvine, CA <i>Ph.D. Chemistry</i>	2013-2018
<b>Willamette University</b> , Salem, OR <i>B.A. Chemistry (with honors)</i>	2009-2013

### RESEARCH EXPERIENCE

---

<b>University of California, Irvine</b> , Irvine, CA Doctoral candidate Department of Chemistry Advisor: Prof. Jennifer Prescher <b>Research:</b> Development of novel optical imaging tools to probe multicomponent systems	2013-2018
<b>Willamette University</b> , Salem, OR Undergraduate researcher Department of Chemistry Advisor: Professor Andrew Duncan <b>Research:</b> Synthesis of novel bis(oxazoline) ligands from sugar precursors for asymmetric catalysis.	2011-2013
<b>Bend Research</b> , Bend, OR Research Intern <b>Research:</b> Development of drug formulations with enhanced solubility via spray dried dispersion technology.	2012

### PUBLICATIONS

---

**Zhang, B. S.** ; Jones, K. A.; McCutcheon D. C.; Prescher, J. A. Pyridone luciferins and mutant luciferases for bioluminescent imaging. *ChemBioChem* **2018**, *19*, 470-477.

Liu, M. D.; Warner, E. A.; Morrissey, C. E.; Fick, C. W.; Wu, T. S.; Ornealas, M. Y.; Ochoa, G. V.; **Zhang, B. S.**; Rathbun, C. M.; Porterfield, W. B.; Prescher, J. A.; Leconte, A. M. Statistical Coupling Analysis-guided library design for discovery of mutant luciferases. *Biochemistry* **2018**, *57*, 663-671.

**Zhang, B. S.;** Yao, Z.; Prescher, J. A. Advances in bioluminescence imaging: New probes from old recipes. *Curr. Opin. Chem. Biol.* **2018**, *45*, 148-156.

## **PRESENTATIONS**

---

Zhang, B.S. *Mutant luciferases provide enhanced light emission with pyridone luciferins*. Presented at the 253<sup>rd</sup> American Chemical Society National Meeting & Exposition, San Francisco, CA, 2017. (poster)

Zhang, B. S. *Design and synthesis of pyridone luciferins for bioluminescence imaging*. Presented at the 251<sup>st</sup> American Chemical Society National Meeting & Exposition, San Diego, CA, 2016. (poster)

## **HONORS AND AWARDS**

---

GAANN Fellowship

2013-2015

# ABSTRACT OF THE DISSERTATION

Expanding the bioluminescent tool box for multicomponent imaging

By

Brendan S. Zhang

Doctor of Philosophy in Chemistry

University of California, Irvine, 2018

Professor Jennifer Prescher, Chair

Optical reporters have revolutionized our ability to visualize biological mechanisms in action. Among the most popular techniques for imaging in whole tissues and living organisms is bioluminescence. One of the most widely used bioluminescent systems comprises firefly luciferase (Fluc) and the small molecule D-luciferin. These components produce photons that can be detected with sensitive cameras. Since mammalian tissues produce little to no photons endogenously, bioluminescence is well suited for imaging in whole organisms. Consequently, Fluc and D-luciferin have been widely used to track cell movements, gene expression patterns and other activities in a variety of preclinical models. Bioluminescence has been largely limited to imaging one cell type at a time, though, due to a lack of distinguishable luciferase-luciferin pairs.

To address this void, my thesis work focused on developing spectrally resolved for multicomponent imaging. I generated both electronically modified luciferins and mutant luciferases. Isomeric luciferin analogs bearing a pyridone moiety were designed and synthesized. These analogs exhibited distinct bioluminescence spectra that could be distinguished using optical filters. However, the analogs were weak emitters

compared to the native substrate. Mutant luciferases with improved photon outputs were identified via library screening and directed evolution. Significant improvements in luciferase activity could be achieved in only 1-2 rounds of screening.

While spectrally resolved probes can be readily distinguished in transparent media, discriminating wavelengths through thick tissues is challenging. To address this issue, I developed a mutually orthogonal luciferase-luciferin pair. Pi-extended luciferin analogs were synthesized, and a complementary luciferase was engineered using computational enzyme design. This custom luciferase-luciferin pair was compatible with three existing bioluminescent tools, enabling selective imaging of four luciferases based on substrate preference.

Finally, I expressed and purified orthogonal luciferases for X-ray crystallography. Luciferase purity and fidelity were verified using a variety of analytical techniques. However, when the luciferases were subjected to published crystallization procedures, only protein precipitation was observed. Ligands were shown to stabilize the enzyme, and crystallization attempts with these molecules yielded several promising “hits”. Collectively, my thesis work expands the toolset for multicomponent imaging and builds towards an understanding of luciferase substrate specificity.

# Chapter 1: Advances in bioluminescence imaging: New probes from old recipes<sup>a</sup>

<sup>a</sup>Reproduced with permission from ref. [1]

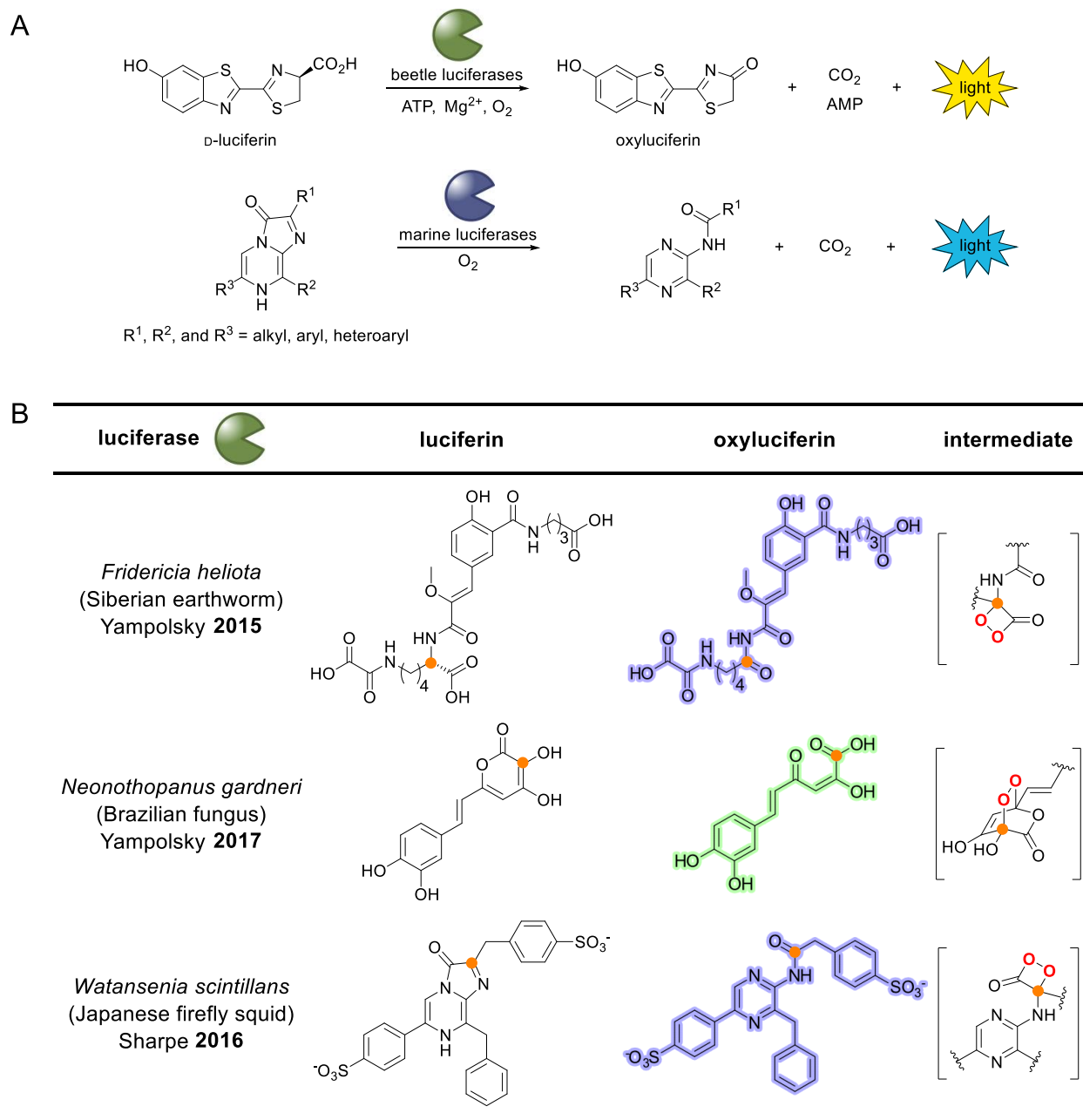
## 1.1 Introduction

Bioluminescent enzymes (luciferases) are among the most sensitive probes for imaging in thick tissues and whole organisms [1]. Luciferases catalyze light emission via the oxidation of small molecule substrates (luciferins). Since no external light is required, the background emission is virtually zero, enabling sensitive imaging *in vivo*. Bioluminescence has long been used to track cells, gene expression, and other biological features in tissues and whole organisms [2]. The emitted light is inherently weak, though, compared to conventional fluorescent tools. For this reason, luciferases are typically used in conjunction with fluorescent proteins. The bioluminescent enzymes survey processes on the macro scale and in heterogeneous environments. The fluorescent probes capture events at the micro scale or *ex vivo* – environments where excitation light is more efficiently delivered.

Historically, the most popular bioluminescent reporter for imaging *in vivo* has been firefly luciferase (Fluc). This enzyme emits the largest percentage of tissue-penetrant light with its cognate luciferin (D-luciferin, Figure 1-1) [3]. Other luciferases, including *Renilla* luciferase (Rluc) and *Gaussia* luciferase (Gluc) have also found broad utility in biological research [4]. These enzymes oxidize coelenterazine and emit blue light in the process. Rluc and Gluc require no additional cofactors (other than oxygen), making them well suited for extracellular work. Compared to their fluorescent protein counterparts, though, luciferases have been less frequently employed in bioimaging



studies. Fewer bioluminescent probes have been developed and even fewer have been optimized for application *in vivo*. There is a constant demand for more colors, improved



enzymes, and more biocompatible substrates.

**Figure 1-1. Luciferase-luciferin pairs in nature. A)** Beetle luciferases oxidize D-luciferin using ATP and  $\text{O}_2$ , generating primarily yellow-green light. Marine luciferases release blue photons via the oxidation of imidazopyrazinone analogs. **B)** Recently characterized luciferases exploit unique molecules and mechanisms to produce light. The relevant luciferins and oxyluciferin products are shown. Orange dots mark the sites of oxidation.

Advances in protein engineering and chemical syntheses are addressing these voids. The past few years, in particular, have seen an uptick in the number of sensitive and substrate-selective luciferases available for use. Much of the progress mirrors trends in fluorescent protein development, including identifying mechanistically distinct probes in nature and subsequently evolving for new function [5]. Systematic efforts to engineer fluorescent probes for altered colors of emission, photo-switching capabilities, and other features ultimately enabled new studies in biology. This iterative cycle of tool development and biological discovery is similarly driving the field of bioluminescence. Below we highlight recent efforts to discover and evolve new bioluminescent tools, and showcase their application to biological sensing.

## **1.2 Discovering new luciferases and luciferins**

Thousands of luminescent species exist in the natural world, but only a fraction of the associated luciferases and luciferins have been characterized in detail [4,6]. Even fewer have been coopted for use in heterologous systems [1]. Continued efforts to mine new luciferase and luciferin architectures from natural sources are expanding the number of available tools. For example, the luciferase gene from *Photinus scintillans* was recently cloned [7]. *P. scintillans* emits predominantly orange light, in contrast to the well-known North American firefly (which emits predominantly yellow-green light). The unique spectrum was traced to a single amino acid change (Y255F) in the luciferase structure. In 2016, Sharpe and colleagues reported the isolation of bioluminescent, crystalline protein assemblies in the Japanese firefly squid [8]. The

crystals comprise three different — but homologous — proteins that catalyze light emission with coelenterazine-disulfate and ATP (Figure 1-1B).

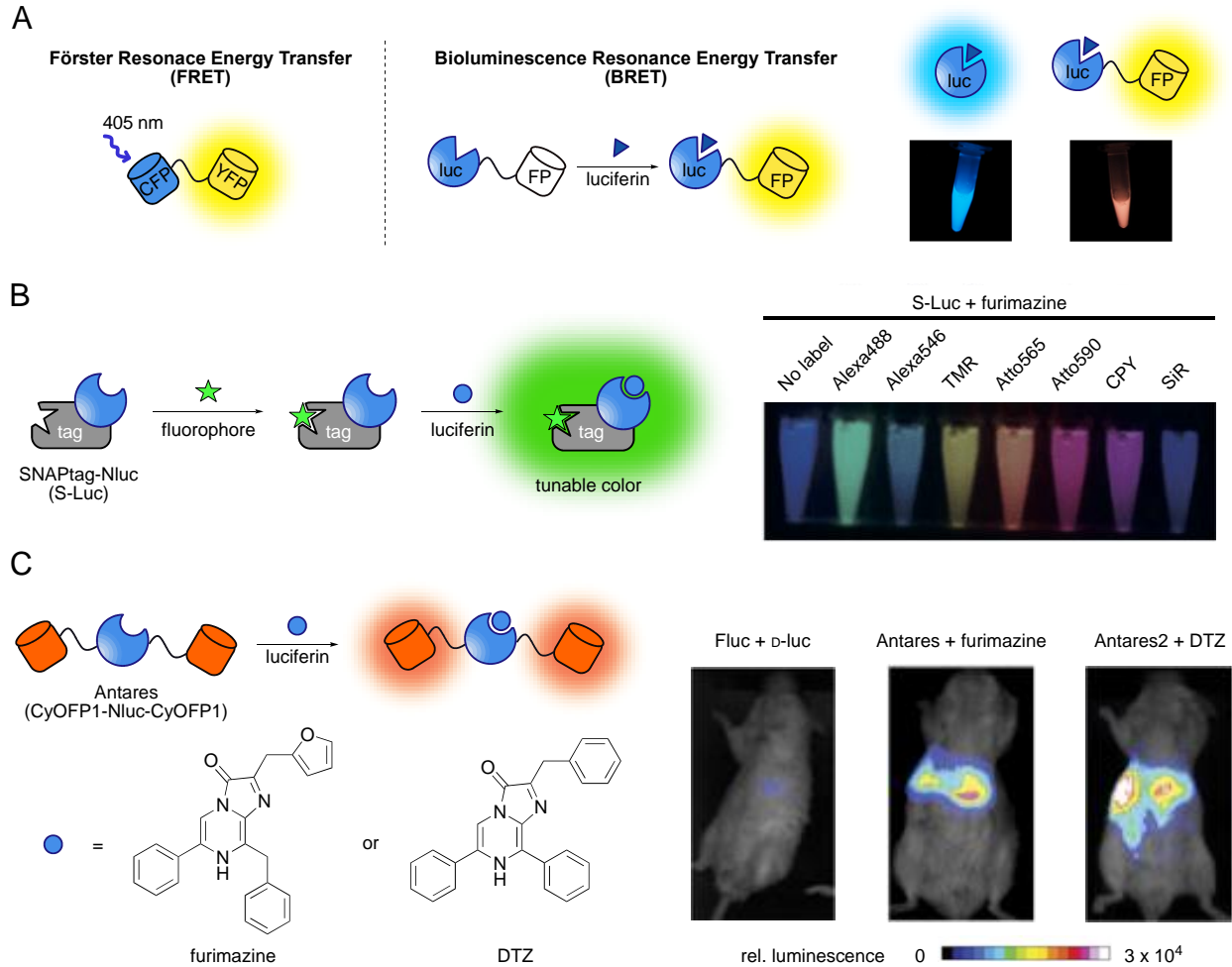
New luciferin scaffolds and light-emitting mechanisms have also been elucidated (Figure 1-1B). One recent example includes the peptide-like luciferin from *Fridericia heliota*. Yampolsky and colleagues speculated that this molecule undergoes oxidative decarboxylation in the light-emitting reaction (similar to D-luciferin), despite its highly divergent structure [9]. This same group also discovered a new bioluminescent mechanism operative within glowing fungi. Some species convert 3-hydroxyhispidin [10] to a putative endoperoxide intermediate en route to light emission. This scaffold is distinct from dioxetanones and other intermediates observed in classic bioluminescent reactions. Such unique luciferins and light-emitting mechanisms are potentially useful for multi-component imaging.

### **1.3 Generating a palette of bioluminescent probes**

The discovery and characterization of native bioluminescent systems, while important, has often not kept pace with the demand for user-friendly imaging tools. Thus, efforts to engineer bioluminescent probes with desirable properties have been critical to fill voids in the imaging toolbox. Many of the approaches have mirrored those in fluorescent protein development: mutagenesis and screening for desired properties such as thermostability, turnover, and color. Some of the most impactful luciferase engineering work in the past few years has centered around NanoLuciferase (Nluc) [11]. Nluc is a small (16 kDa) engineered variant of a luciferase found in sea shrimp *Oplophorus gracilirostris*. Nluc was evolved to process a more stable coelenterazine

analog (furimazine) in the light-emitting reaction. The Nluc-furimazine pair has been widely adopted for imaging studies in diverse fields, due to its brightness and stability. Split versions of Nluc have also been reported [12]. Like split GFP and other fluorescent proteins, these tools have been used for analyzing protein-protein interactions in cells [12] and screening inhibitors [13].

Nluc has also proven to be a versatile platform for broadening the palette of bioluminescent probes. Much like fluorescent proteins, distinct bioluminescent reporters are desirable for applications in multicellular imaging. An enhanced set of colors can be achieved via bioluminescence resonance energy transfer (BRET, Figure 1-2). BRET involves luminescent reactions that excite acceptor fluorophores, resulting in altered emission spectra. The process is analogous to Förster resonance energy transfer (FRET), where energy transfer processes between two fluorophores can tune emission spectra. Nagai and colleagues generated a set of Nluc-fluorescent protein conjugates for BRET imaging. These chimeras are similar to earlier generations of Rluc-fluorescent protein conjugates (i.e., the “Nano-lanterns”) [14-15]. The cyan, green, and red-emitting Nluc lanterns exhibit quantum yields on par or exceeding that of Nluc itself and can enable real-time colorimetric imaging [16]. Fluorescent dyes are also suitable BRET acceptors. Johnsson and coworkers pioneered a strategy to append different fluorophores to Nluc using SNAPtag and HaloTag technology (Figure 1-2). The suite of resulting probes provided a bioluminescent portrait reminiscent of the famous fluorescent protein collection [17]. The Nluc chimeras were also shown to be well suited for multi-component imaging in cells [18].



**Figure 1-2. Expanded palette of bioluminescent probes. A)** Resonance energy transfer processes can tune optical emission spectra. In FRET, donor fluorophores (e.g., CFP) are excited with external light. Emission from the acceptor fluorophore (e.g., YFP) is observed. In BRET, luminescent reactions can excite acceptor fluorophores, resulting in altered emission spectra. A sample BRET construct (ReNL, pairing Nluc with tdTomato) is shown. **B)** Nluc-fluorophore chimeras expand the palette of bioluminescent probes. Fluorescent molecules were appended to Nluc via SNAPtag ligation (left), generating a colorful array of S-Luc tags (right). S-Luc images were reproduced with permission from ref. 22. **C)** Far red-emitting BRET constructs enable sensitive imaging *in vivo*. Antares comprises Nluc and two copies of a fluorescent protein (CyOFP1). Red-shifted light is produced upon luciferin administration. Antares and a related construct (Antares2) were expressed in mice following hydrodynamic transfection. Light emission was observed upon furimazine or DTZ administration (3.3  $\mu\text{mol}$  i.p.). Mouse images were reproduced with permission from ref 24.

Pushing the frontiers in noninvasive imaging, the Lin lab reported a Nluc BRET construct (Antares) suitable for *in vivo* work [19]. Antares comprises Nluc flanked by 2

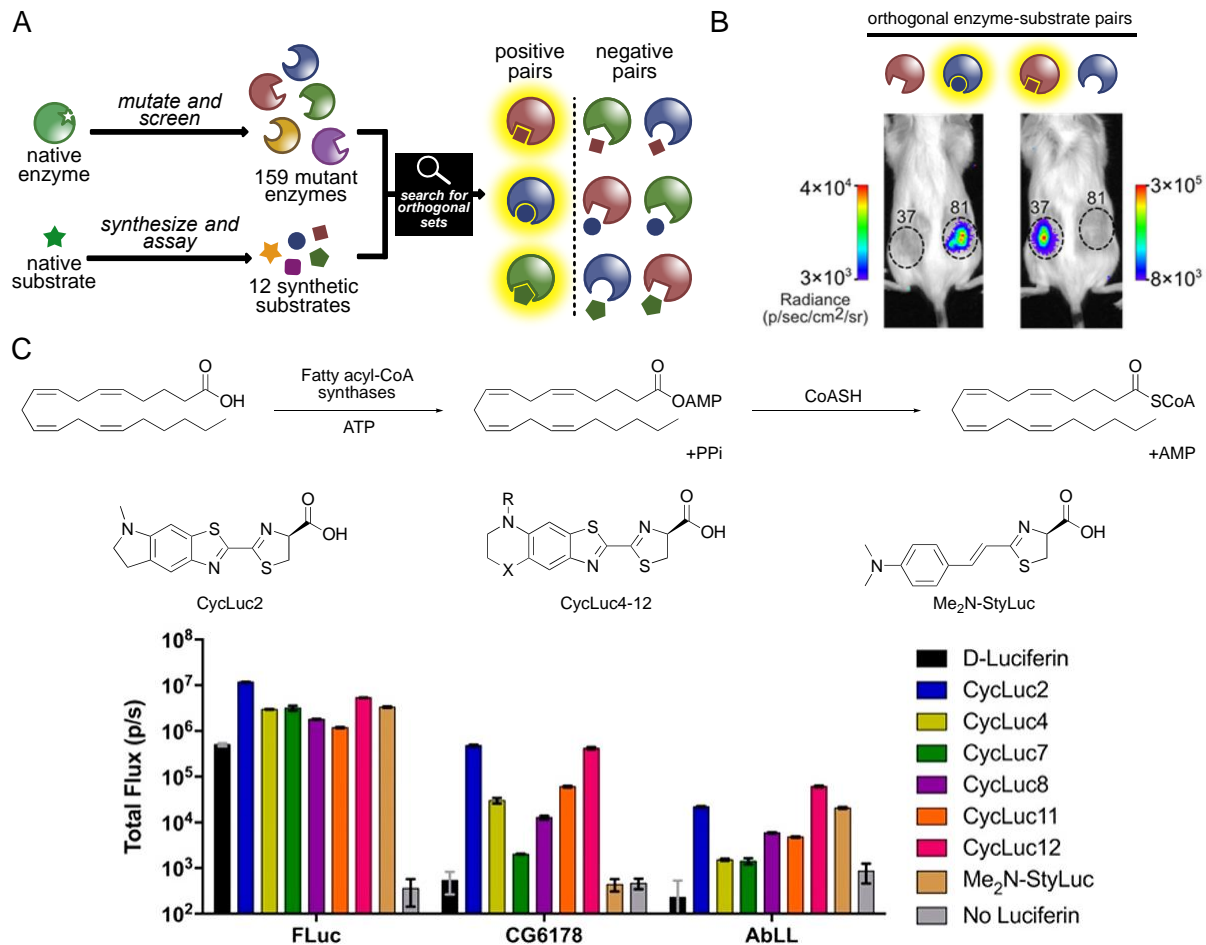
copies of an orange fluorescent protein (CYOFP1). This construct produces ~20-fold more tissue-penetrant photons (> 600 nm) compared to Nluc, enabling sensitive imaging in rodents. Further engineering and analog optimization yielded a second-generation reporter: Nluc (teLuc) that uses a modified furimazine analog (DTZ, Figure 1-2C). teLuc and DTZ provided improved spectral overlap with CYOFP1. The optimized construct (termed Antares2) exhibited enhanced red-shifted light emission and more robust bioluminescence in deep tissues [20]. Related BRET probes with Rluc have similarly provided bioluminescent probes that emit in the near-infrared regime [21].

Efforts to produce multi-spectral tools have historically focused on the luciferase enzyme, although modifying the luciferin architecture represents another viable route. Changes to the luciferin chromophore can directly impact the color of light released. For example, extending the conjugation of the luciferin pi system or altering heteroatom substituents can alter emission wavelengths [22-24]. Both blue- and red-shifted analogs have been produced, although most remain weak emitters with native luciferases [22-23,25]. Engineering enzymes to better process the modified analogs—and thus recover light intensity—has been successful in some cases [23,25-26].

#### **1.4 Engineering orthogonal luciferase-luciferin pairs**

Discriminating among wavelengths *in vivo* is challenging, as the perceived color changes with depth. Multi-component bioluminescence imaging has thus been most often achieved using *substrate*-resolved luciferases versus spectrally resolved pairs. For example, Fluc and Rluc oxidize completely different luciferins and can therefore be readily distinguished in two-component assays [27]. The Fluc/Rluc combination has

further inspired the expansion of orthogonal bioluminescent tools. Unique patterns of substrate use, rather than color, can serve as diagnostic fingerprints for collections of cells or other features (Figure 1-3A). In our own lab, we synthesized dozens of chemically distinct luciferins and screened them against a panel of Fluc mutants [28-29]. A computer algorithm was used to identify orthogonal enzyme-substrate pairs. Substrate selectivity was maintained in both mammalian cells and in mouse models, enabling multi-cellular imaging *in vivo* (Figure 1-3B). Additional screening analysis further revealed triplet sets and higher-order orthogonal combinations [29]. Simultaneous engineering of enzymes and substrates has also been applied to luciferases that use coelenterazine [30].



**Figure 1-3. Substrate-selective luciferases for multi-component imaging.** **A)** Orthogonal luciferases were identified via parallel screening of luciferase mutants and luciferin analogs. **B)** Dual imaging with engineered luciferase-luciferin pairs. DB7 cells expressing orthogonal mutants (37 and 81) were inoculated in opposing flanks. The populations were readily distinguished upon administration of the complementary luciferins. Bioluminescence images were reproduced with permission from ref. 33. **C)** Fatty acyl-CoA synthetases from non-luminous organisms (e.g., CG6178 from *D. melanogaster* and AbLL from *A. binodulus*) exhibit luciferase-like behavior with synthetic luciferin analogs. Bar graph was reproduced with permission from ref. 36.

Luciferase-like enzymes are further expanding the number of orthogonal probes.

Luciferases belong to the ANL (Acy-CoA synthetases, NRPS adenylation domains, and Luciferase enzymes) superfamily; these enzymes use a common mechanism to activate carboxylates as adenylates. These intermediates can be displaced with biological thiols (e.g. CoA) or, in the case of firefly luciferase, react with molecular oxygen [31]. Most ANL enzymes do not catalyze light-emitting reactions with their cognate substrates.



Excitingly, though, Miller showed that some fatty acyl-CoA synthetases exhibit “latent” luciferase activity when supplied with a luminogenic substrate [32]. For example, AbLL (a synthetase from the non-luminous beetle *Agrypnus binodulus*) catalyzed light emission with a panel of synthetic luciferin analogs (Figure 1-3C) [32]. Many of the latent luciferases exhibit unique patterns of substrate use, expanding the number of new and orthogonal tools.

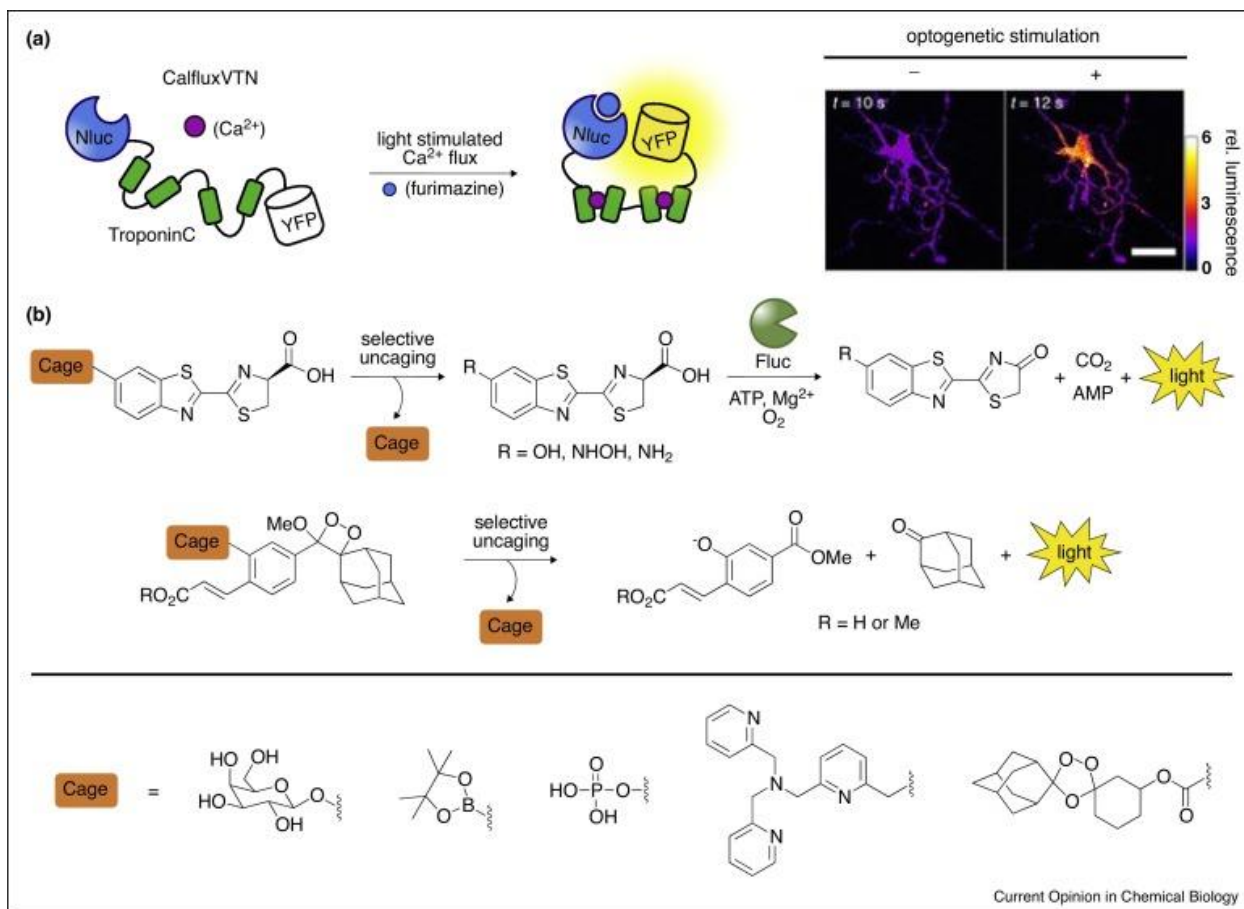
The functional and sequence similarity between luciferase and other ANL enzymes is further enabling luciferase engineering efforts. ANL enzymes are known to be promiscuous [31], and can potentially serve as starting points to identify new luciferase-luciferin pairs. The Leconte lab used the homologous enzymes in combination with a previously developed bioinformatic method, statistical coupling analysis (SCA) [33-34] to guide the design of mutant luciferase libraries [35]. SCA was used to analyze amino acid positions that were mutable and functionally important, along with networks of potentially synergistic interactions. In a single round of selection, mutants with desirable red-shifted emission spectra and improved thermostability were identified. Mutants with >50-fold changes in specificity for modified luciferins were also found.

Continued efforts to identify new enzyme-substrate pairs require rapid access to diverse collections of luciferins. Such molecules have historically been difficult to synthesize from common routes. Recent advances in luciferin chemistry, though, are beginning to address this issue. Modular coupling reactions to outfit D-luciferin with diverse steric modifications have been reported [28,36]. Ring-closing metathesis and carbene insertions have also been used to produce a series of conformationally

restricted and pi-extended coelenterazines [37-39]. Many these probes exhibited red-shifted emission or other desirable photophysical properties.

### **1.5 Monitoring new facets of biology**

Advances in luciferase engineering have ushered in a flurry of new sensors for metabolites and enzyme activities [40-43]. Many of the probes have parallels to classic fluorescent sensors, but are more tailored for *in vivo* work. A notable example is CalfluxVTN, a BRET-based calcium sensor comprising Nluc and Venus fluorescent protein (Figure 1-4A). In the absence of  $\text{Ca}^{2+}$ , Nluc emission is observed. Upon  $\text{Ca}^{2+}$  binding, the sensor undergoes a conformational change and BRET is observed. CalfluxVTN enabled sensitive imaging of calcium flux in response to stimulation of a rhodopsin photoreceptor [44]. Such measurements were refractory to FRET, as external light interfered with receptor activation. Johnsson and coworkers further developed a universal BRET sensor platform for analyte detection. They fused various antibody fragments to an Nluc-fluorophore pair; upon binding of a complementary analyte, a conformational change was induced, accompanied by a change in emission color [45]. The modularity of this system could enable point-of-care diagnosis for a variety of antigens.



**Figure 1-4. Visualizing cellular species with bioluminescent sensors. A)** CalfluxVTN comprises a calcium-binding protein (TroponinC) flanked by Nluc and YFP (left).  $\text{Ca}^{2+}$  binding induces a conformational change in the sensor, resulting in BRET. CalfluxVTN was expressed in neurons and used to monitor  $\text{Ca}^{2+}$  flux following photoreceptor firing (right, scale bar = 20  $\mu\text{m}$ ). Cellular images were reproduced with permission from ref. 48. **B)** Caged luciferins and dioxetanes can report on cellular activities. Selective removal of the caging groups provides an active luminophore. Light emission via luciferase oxidation (top) or direct chemiluminescence (bottom) thus provides a readout on the uncaging enzyme or analyte of interest.

Advances in luciferin synthesis have also enabled access to new probes of cellular function, including “caged” luciferins. These molecules typically contain a bulky group (i.e. “cage”) that renders the molecule non-emissive with luciferase. Upon removal of the cage (typically from an enzymatic reaction), an active luciferin is revealed and available for light emission. “Caged” luciferins have recently been used to detect biologically relevant metal ions and other species [46-49]. (Figure 1-4B). Some have

also been used to profile cell-cell interactions [50] and improve delivery [51]. The caged luminophore concept has recently been expanded to craft novel chemiluminescent sensors [52]. Some of these probes comprise embedded dioxetanes that are cleavable—and thus emit light—in response to a variety of triggers [53-54]. Unlike canonical caged luciferins, these reporters do not require a luciferase to produce light. Such probes further diversify the portfolio of tools for biological imaging.

## **1.6 Conclusion**

Many advances in bioluminescent probe technology have mirrored trends in fluorescent probe development. Dozens of luciferases have been evolved for new functions via iterative mutagenesis and screening. Collections of robust and structurally distinct luciferins have also been synthesized. A variety of unique bioluminescent mechanisms have further been uncovered in the natural world, providing platforms from which to craft new tools. The continued discovery and development of bioluminescence probes, like other optical imaging agents, promises to expand what researchers can 'see' in cells and tissues.

## **1.7 Objectives of this study**

Bioluminescent systems from nature are powerful tools for imaging in living organisms. Tracking the movements of cells and other biological features is critical for understanding the pathology of disease. However, bioluminescence tools have been largely restricted to studying one cell type or feature at a time. This limits our ability to interrogate cell-cell interactions and other multi-component processes in heterogeneous

environments. To solve this problem, my graduate work has involved developing bioluminescent tools for multicellular imaging.

I aimed to:

1. Expand the palette of bioluminescent tools by synthesizing luciferin analogs with modified heterocyclic cores.
2. Develop orthogonal bioluminescent tools by synthesizing structurally diverse luciferin analogs and evolving mutant luciferases that selectively utilize them.
3. Understand the origin of evolved luciferase selectivity (at the atomic level) by crystallizing luciferase for structural analysis.

## References

1. Yao, Z.; Zhang, B. S.; Prescher, J. A. Advances in bioluminescence imaging: new probes from old recipes. *Curr. Opin. Chem. Biol.* **2018**, *45*, 148-156.
2. Paley, M. A.; Prescher, J. A. Bioluminescence: A versatile technique for imaging cellular and molecular features. *MedChemComm* **2014**, *5*, 255-267.
3. Zhao, H.; Doyle, T. C.; Coquoz, O.; Kalish, F.; Rice, B. W.; Contag, C. H. Emission spectra of bioluminescent reporters and interaction with mammalian tissue determine the sensitivity of detection *in vivo*. *J. Biomed. Opt.* **2005**, *10*, 41210.
4. Kaskova, Z. M.; Tsarkova, A. S.; Yampolsky, I. V. 1001 lights: Luciferins, luciferases, their mechanisms of action and applications in chemical analysis, biology and medicine. *Chem. Soc. Rev.* **2016**, *45*, 6048-6077.
5. Rodriguez, E. A.; Campbell, R. E.; Lin, J. Y.; Lin, M. Z.; Miyawaki, A.; Palmer, A. E.; Shu, X. K.; Zhang, J.; Tsien, R. Y. The Growing and Glowing Toolbox of Fluorescent and Photoactive Proteins. *Trends Biochem. Sci.* **2017**, *42*, 111-129.
6. Martini, S.; Haddock, S. H. Quantification of bioluminescence from the surface to the deep sea demonstrates its predominance as an ecological trait. *Sci. Rep.* **2017**, *7*, 45750.
7. Branchini, B. R.; Southworth, T. L.; Fontaine, D. M.; Murtiashaw, M. H.; McGurk, A.; Talukder, M. H.; Qureshi, R.; Yetil, D.; Sundlov, J. A.; Gulick, A. M. Cloning of the

Orange Light-Producing Luciferase from *Photinus scintillans*-A New Proposal on how Bioluminescence Color is Determined. *Photochem. Photobiol.* **2017**, *93*, 479-485.

8. Gimenez, G.; Metcalf, P.; Paterson, N. G.; Sharpe, M. L. Mass spectrometry analysis and transcriptome sequencing reveal glowing squid crystal proteins are in the same superfamily as firefly luciferase. *Sci. Rep.* **2016**, *6*.

9. Dubinnyi, M. A.; Kaskova, Z. M.; Rodionova, N. S.; Baranov, M. S.; Gorokhovatsky, A. Y.; Kotlobay, A.; Solntsev, K. M.; Tsarkova, A. S.; Petushkov, V. N.; Yampolsky, I. V. Novel mechanism of bioluminescence: oxidative decarboxylation of a moiety adjacent to the light emitter of *Fridericia* luciferin. *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 7065-7067.

10. Kaskova, Z. M.; Dorr, F. A.; Petushkov, V. N.; Purtov, K. V.; Tsarkova, A. S.; Rodionova, N. S.; Mineev, K. S.; Guglya, E. B.; Kotlobay, A.; Baleeva, N. S.; Baranov, M. S.; Arseniev, A. S.; Gitelson, J. I.; Lukyanov, S.; Suzuki, Y.; Kanie, S.; Pinto, E.; Di Mascio, P.; Waldenmaier, H. E.; Pereira, T. A.; Carvalho, R. P.; Oliveira, A. G.; Oba, Y.; Bastos, E. L.; Stevani, C. V.; Yampolsky, I. V. Mechanism and color modulation of fungal bioluminescence. *Sci. Adv.* **2017**, *3*, e1602847.

11. Hall, M. P.; Unch, J.; Binkowski, B. F.; Valley, M. P.; Butler, B. L.; Wood, M. G.; Otto, P.; Zimmerman, K.; Vidugiris, G.; Machleidt, T.; Robers, M. B.; Benink, H. A.; Eggers, C. T.; Slater, M. R.; Meisenheimer, P. L.; Klaubert, D. H.; Fan, F.; Encell, L. P.; Wood, K. V. Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem. Biol.* **2012**, *7*, 1848-1857.

12. Dixon, A. S.; Schwinn, M. K.; Hall, M. P.; Zimmerman, K.; Otto, P.; Lubben, T. H.; Butler, B. L.; Binkowski, B. F.; Machleidt, T.; Kirkland, T. A.; Wood, M. G.; Eggers, C. T.; Encell, L. P.; Wood, K. V. NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS Chem. Biol.* **2016**, *11*, 400-408.

13. Hayes, M. P.; Soto-Velasquez, M.; Fowler, C. A.; Watts, V. J.; Roman, D. L. Identification of FDA-Approved Small Molecules Capable of Disrupting the Calmodulin-Adenylyl Cyclase 8 Interaction through Direct Binding to Calmodulin. *ACS Chem. Neurosci.* **2018**, *9*, 346-357.

14. Saito, K.; Chang, Y. F.; Horikawa, K.; Hatsugai, N.; Higuchi, Y.; Hashida, M.; Yoshida, Y.; Matsuda, T.; Arai, Y.; Nagai, T. Luminescent proteins for high-speed single-cell and whole-body imaging. *Nat. Commun.* **2012**, *3*, 1262.

15. Takai, A.; Nakano, M.; Saito, K.; Haruno, R.; Watanabe, T. M.; Ohyanagi, T.; Jin, T.; Okada, Y.; Nagai, T. Expanded palette of Nano-lanterns for real-time multicolor luminescence imaging. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 4352-4356.

16. Suzuki, K.; Kimura, T.; Shinoda, H.; Bai, G.; Daniels, M. J.; Arai, Y.; Nakano, M.; Nagai, T. Five colour variants of bright luminescent protein for real-time multicolour bioimaging. *Nat. Commun.* **2016**, *7*, 13718.
17. Tsien, R. Tsien Lab Website. 2004 <http://www.tsienlab.ucsd.edu/>.
18. Hiblot, J.; Yu, Q.; Sabbadini, M. D. B.; Reymond, L.; Xue, L.; Schena, A.; Sallin, O.; Hill, N.; Griss, R.; Johnsson, K. Luciferases with Tunable Emission Wavelengths. *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 14556-14560.
19. Chu, J.; Oh, Y.; Sens, A.; Ataie, N.; Dana, H.; Macklin, J. J.; Laviv, T.; Welf, E. S.; Dean, K. M.; Zhang, F.; Kim, B. B.; Tang, C. T.; Hu, M.; Baird, M. A.; Davidson, M. W.; Kay, M. A.; Fiolka, R.; Yasuda, R.; Kim, D. S.; Ng, H. L.; Lin, M. Z. A bright cyan-excitable orange fluorescent protein facilitates dual-emission microscopy and enhances bioluminescence imaging *in vivo*. *Nat. Biotechnol.* **2016**, *34*, 760-767.
20. Yeh, H. W.; Karmach, O.; Ji, A.; Carter, D.; Martins-Green, M. M.; Ai, H. W. Red-shifted luciferase-luciferin pairs for enhanced bioluminescence imaging. *Nat. Methods* **2017**, *14*, 971-974.
21. Romyantsev, K. A.; Turoverov, K. K.; Verkhusha, V. V. Near-infrared bioluminescent proteins for two-color multimodal imaging. *Sci. Rep.* **2016**, *6*, 36588.
22. Anderson, J. C.; Grounds, H.; Jathoul, A. P.; Murray, J. A. H.; Pacman, S. J.; Tisi, L. Convergent synthesis and optical properties of near-infrared emitting bioluminescent infra-luciferins. *RSC Adv.* **2017**, *7*, 3975-3982.
23. Hall, M. P.; Woodroffe, C. C.; Wood, M. G.; Que, I.; Van't Root, M.; Ridwan, Y.; Shi, C.; Kirkland, T. A.; Encell, L. P.; Wood, K. V.; Lowik, C.; Mezzanotte, L. Click beetle luciferase mutant and near infrared naphthyl-luciferins for improved bioluminescence imaging. *Nat. Commun.* **2018**, *9*, 132.
24. Kuchimaru, T.; Iwano, S.; Kiyama, M.; Mitsumata, S.; Kadonosono, T.; Niwa, H.; Maki, S.; Kizaka-Kondoh, S. A luciferin analogue generating near-infrared bioluminescence achieves highly sensitive deep-tissue imaging. *Nat. Commun.* **2016**, *7*, 11856.
25. Zhang, B. S.; Jones, K. A.; McCutcheon, D. C.; Prescher, J. A. Pyridone Luciferins and Mutant Luciferases for Bioluminescence Imaging. *ChemBioChem* **2018**, *19*, 470-477.
26. Mofford, D. M.; Reddy, G. R.; Miller, S. C. Aminoluciferins extend firefly luciferase bioluminescence into the near-infrared and can be preferred substrates over D-luciferin. *J. Am. Chem. Soc.* **2014**, *136*, 13277-13282.

27. Chan, C. T.; Reeves, R. E.; Geller, R.; Yaghoubi, S. S.; Hoehne, A.; Solow-Cordero, D. E.; Chiosis, G.; Massoud, T. F.; Paulmurugan, R.; Gambhir, S. S. Discovery and validation of small-molecule heat-shock protein 90 inhibitors through multimodality molecular imaging in living subjects. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E2476-2485.
28. Jones, K. A.; Porterfield, W. B.; Rathbun, C. M.; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. Orthogonal Luciferase-Luciferin Pairs for Bioluminescence Imaging. *J. Am. Chem. Soc.* **2017**, *139*, 2351-2358.
29. Rathbun, C. M.; Porterfield, W. B.; Jones, K. A.; Sagoe, M. J.; Reyes, M. R.; Hua, C. T.; Prescher, J. A. Parallel Screening for Rapid Identification of Orthogonal Bioluminescent Tools. *ACS Cent. Sci.* **2017**, *3*, 1254-1261.
30. Nishihara, R.; Abe, M.; Nishiyama, S.; Citterio, D.; Suzuki, K.; Kim, S. B. Luciferase-Specific Coelenterazine Analogues for Optical Contamination-Free Bioassays. *Sci. Rep.* **2017**, *7*, 908.
31. Gulick, A. M. Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem. Biol.* **2009**, *4*, 811-827.
32. Mofford, D. M.; Liebmann, K. L.; Sankaran, G. S.; Reddy, G.; Reddy, G. R.; Miller, S. C. Luciferase Activity of Insect Fatty Acyl-CoA Synthetases with Synthetic Luciferins. *ACS Chem. Biol.* **2017**, *12*, 2946-2951.
33. Socolich, M.; Lockless, S. W.; Russ, W. P.; Lee, H.; Gardner, K. H.; Ranganathan, R. Evolutionary information for specifying a protein fold. *Nature* **2005**, *437*, 512-518.
34. Smock, R. G.; Rivoire, O.; Russ, W. P.; Swain, J. F.; Leibler, S.; Ranganathan, R.; Gierasch, L. M. An interdomain sector mediating allostery in Hsp70 molecular chaperones. *Mol. Syst. Biol.* **2010**, *6*, 414.
35. Liu, M. D.; Warner, E. A.; Morrissey, C. E.; Fick, C. W.; Wu, T. S.; Ornelas, M. Y.; Ochoa, G. V.; Zhang, B. S.; Rathbun, C. M.; Porterfield, W. B.; Prescher, J. A.; Leconte, A. M. Statistical Coupling Analysis-Guided Library Design for the Discovery of Mutant Luciferases. *Biochemistry* **2018**, *57*, 663-671.
36. Sharma, D. K.; Adams, S. T., Jr.; Liebmann, K. L.; Miller, S. C. Rapid Access to a Broad Range of 6'-Substituted Firefly Luciferin Analogues Reveals Surprising Emitters and Inhibitors. *Org. Lett.* **2017**, *19*, 5836-5839.
37. Nishihara, R.; Suzuki, H.; Hoshino, E.; Suganuma, S.; Sato, M.; Saitoh, T.; Nishiyama, S.; Iwasawa, N.; Citterio, D.; Suzuki, K. Bioluminescent coelenterazine



derivatives with imidazopyrazinone C-6 extended substitution. *Chem. Commun.* **2015**, *51*, 391-394.

38. Shakhmin, A.; Hall, M. P.; Machleidt, T.; Walker, J. R.; Wood, K. V.; Kirkland, T. A. Coelenterazine analogues emit red-shifted bioluminescence with NanoLuc. *Org. Biomol. Chem.* **2017**, *15*, 8559-8567.

39. Hosoya, T.; Iimori, R.; Yoshida, S.; Sumida, Y.; Sahara-Miura, Y.; Sato, J.; Inouye, S. Concise Synthesis of  $\nu$ -Coelenterazines. *Org. Lett.* **2015**, *17*, 3888-91.

40. Shigeto, H.; Ikeda, T.; Kuroda, A.; Funabashi, H. A BRET-based homogeneous insulin assay using interacting domains in the primary binding site of the insulin receptor. *Anal. Chem.* **2015**, *87*, 2764-2770.

41. den Hamer, A.; Dierickx, P.; Arts, R.; de Vries, J.; Brunsveld, L.; Merckx, M. Bright Bioluminescent BRET Sensor Proteins for Measuring Intracellular Caspase Activity. *ACS Sens.* **2017**, *2*, 729-734.

42. Takenouchi, O.; Kanno, A.; Takakura, H.; Hattori, M.; Ozawa, T. Bioluminescent Indicator for Highly Sensitive Analysis of Estrogenic Activity in a Cell-Based Format. *Bioconjug. Chem.* **2016**, *27*, 2689-2694.

43. Aper, S. J.; Dierickx, P.; Merckx, M. Dual Readout BRET/FRET Sensors for Measuring Intracellular Zinc. *ACS Chem. Biol.* **2016**, *11*, 2854-2864.

44. Yang, J.; Cumberbatch, D.; Centanni, S.; Shi, S. Q.; Winder, D.; Webb, D.; Johnson, C. H. Coupling optogenetic stimulation with NanoLuc-based luminescence (BRET)  $Ca^{++}$  sensing. *Nat. Commun.* **2016**, *7*, 13268.

45. Xue, L.; Yu, Q.; Griss, R.; Schena, A.; Johnsson, K. Bioluminescent Antibodies for Point-of-Care Diagnostics. *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 7112-7116.

46. Ke, B.; Wu, W.; Liu, W.; Liang, H.; Gong, D.; Hu, X.; Li, M. Bioluminescence Probe for Detecting Hydrogen Sulfide *in vivo*. *Anal. Chem.* **2016**, *88*, 592-595.

47. Kojima, R.; Takakura, H.; Kamiya, M.; Kobayashi, E.; Komatsu, T.; Ueno, T.; Terai, T.; Hanaoka, K.; Nagano, T.; Urano, Y. Development of a Sensitive Bioluminescent Probe for Imaging Highly Reactive Oxygen Species in Living Rats. *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 14768-14771.

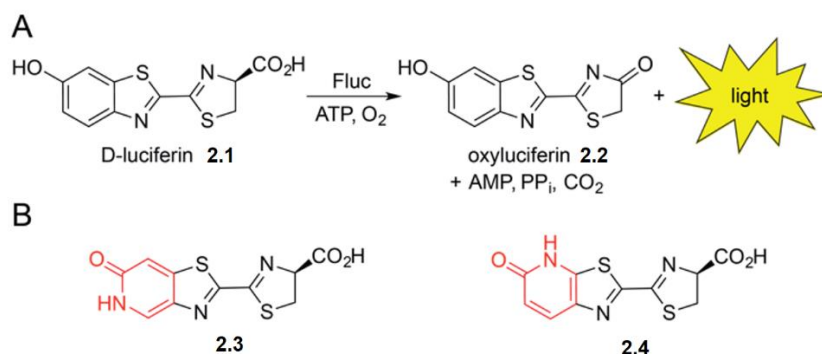
48. Heffern, M. C.; Park, H. M.; Au-Yeung, H. Y.; Van de Bittner, G. C.; Ackerman, C. M.; Stahl, A.; Chang, C. J. *In vivo* bioluminescence imaging reveals copper deficiency in a murine model of nonalcoholic fatty liver disease. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 14219-14224.

49. Aron, A. T.; Heffern, M. C.; Lonergan, Z. R.; Vander Wal, M. N.; Blank, B. R.; Spangler, B.; Zhang, Y.; Park, H. M.; Stahl, A.; Renslo, A. R.; Skaar, E. P.; Chang, C. J. *In vivo* bioluminescence imaging of labile iron accumulation in a murine model of *Acinetobacter baumannii* infection. *Proc Natl Acad Sci U S A* **2017**, *114*, 12669-12674.
50. Porterfield, W. B.; Prescher, J. A. Tools for visualizing cell-cell 'interactomes'. *Curr. Opin. Chem. Biol.* **2015**, *24*, 121-130.
51. Mofford, D. M.; Adams, S. T., Jr.; Reddy, G. S.; Reddy, G. R.; Miller, S. C. Luciferin Amides Enable *In Vivo* Bioluminescence Detection of Endogenous Fatty Acid Amide Hydrolase Activity. *J. Am. Chem. Soc.* **2015**, *137*, 8684-8687.
52. Hananya, N.; Shabat, D. A Glowing Trajectory between Bio- and Chemiluminescence: From Luciferin-Based Probes to Triggerable Dioxetanes. *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 16454-16463.
53. Green, O.; Gnaim, S.; Blau, R.; Eldar-Boock, A.; Satchi-Fainaro, R.; Shabat, D. Near-Infrared Dioxetane Luminophores with Direct Chemiluminescence Emission Mode. *J. Am. Chem. Soc.* **2017**, *139*, 13243-13248.
54. J. A. Cao, W. R., A. G.; Lippert, A. R. A Chemiluminescent Probe for Cellular Peroxynitrite Using a Self-Immolative Oxidative Decarbonylation Reaction. *Chem. Sci.* **2018**, *9*, 2552-2558.

## Chapter 2: Pyridone luciferins and mutant luciferases for bioluminescence imaging

### 2.1 Introduction

As noted in chapter 1, bioluminescence enables sensitive imaging of biological events *in vitro* and *in vivo* [1-2]. This method employs luciferase enzymes that catalyze the oxidation of small molecule luciferins, releasing visible light in the process [3-4] (Figure 2-1A). Several luciferase-luciferin pairs have been identified in nature [5], but only a handful have been optimized for routine use in cells and living organisms [6-10]. Among the most popular is firefly luciferase (Fluc) and D-luciferin (**2.1**). This bioluminescent pair emits the largest percentage of tissue-penetrant light among well characterized enzymes and substrates [11]. Consequently, Fluc and D-luciferin have been widely used to track cell movements [12-13], gene expression patterns [14-15] and other activities [16-18] in a variety of pre-clinical models.



**Figure 2-1. Firefly bioluminescence** **A**) Firefly luciferase catalyzes the oxidation of D-luciferin, producing a photon of light. **B**) Luciferin analogs **2.3** and **2.4** bearing pyridone moieties (red) were examined in this study.

While versatile, bioluminescence imaging with Fluc/D-luciferin has some notable shortcomings. D-Luciferin is not optimally cell permeant [19], requiring large doses to saturate Fluc and achieve robust photon production in cultured cell and tissue models [20]. Imaging with Fluc has also been historically limited to visualizing one feature at a time owing to a lack of distinguishable probes [2,21]. To craft improved bioluminescent tools, many groups have relied on Fluc mutagenesis to isolate “brighter” and spectrally resolved luciferases [22-24]. A complementary and potentially more general approach involves modulating the luciferin itself. The small molecule is the effective light emitter; thus, modifications to the scaffold can, in principle, impact both pharmacokinetic and spectral properties. There have been many notable successes in engineering unique luciferins in recent years [25-37]. However, many of the scaffolds exhibit weak photon outputs with luciferase [25,30,32-37], limiting their use in some cellular and tissue environments.

This chapter describes my efforts in developing spectrally resolved luciferin analogs. I apply these tools *in cellulo* and demonstrate groups of cells can be distinguished based on bioluminescent color. I then turned to developing brighter bioluminescent probes by engineering the enzyme, firefly luciferase. We screened the luciferin analogs with libraries of mutant luciferases and identified enzymes with improved photon output. To further improve the lead enzymes, we designed a second-generation library. Using phylogenetic data, we identified functionally important residues in homologous enzymes. Subsequent mutagenesis of these residues and screening led to improvements over the first-generation enzyme. Intriguingly, improvements in luciferase photon output were achieved with only 1-2 rounds of mutagenesis. This result

highlights the rapid identification of new tools by simultaneously engineering enzyme and substrate.

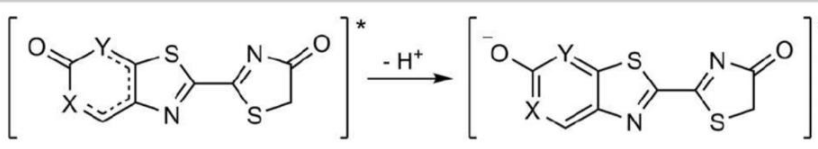
## 2.2 Design and synthesis of luciferin analogs

To further expand the scope of bioluminescence imaging, we aimed to develop a new class of spectrally resolved luciferins for cellular application. We initially focused on scaffolds comprising a pyridone moiety (**2.3** and **2.4**, Figure 2-1B). Pyridones are found in a number of drug molecules [38-40], suggesting that they would be sufficiently stable under physiological conditions. We further reasoned that **2.3** and **2.4** would be good substrates for Fluc or related mutants, given their structural similarity to the native substrate. Pyridones can also tautomerize in the excited state and thus potentially provide unique colors of light. Well-behaved luciferins with distinguishable spectra would be useful additions to the bioluminescence toolkit.

We reasoned that the Fluc-catalyzed oxidation of **2.3-2.4** would provide either a pyridone (neutral) or a deprotonated hydroxypyridine (from the corresponding tautomer, Table 2-1) as the functional light emitter. Luciferins and related chromophores are known to undergo proton transfers [41-43] in the excited state. In the case of the pyridones, deprotonation would provide fully aromatic architectures. Such species would likely emit visible light conducive to cellular imaging applications. To gain insight on the predicted wavelengths, we performed density functional theory (DFT) calculations. The relative HOMO and LUMO energies of the oxyluciferin emitters were used to approximate bioluminescent wavelengths [44]. The predicted emission for the pyridone structures were >100-nm blue-shifted compared to the corresponding

hydroxypyridines. Luciferin **oxy2.3** was also predicted to release ~50 nm more red-shifted light than **oxy2.4** (Table 2-1). A differential of this magnitude can easily be resolved with optical filters [45-46], further suggesting that **2.3** and **2.4** could be viable candidates for multispectral imaging.

**Table 2-1** Predicted  $\Delta E_{\text{HOMO-LUMO}}$  and observed bioluminescent  $\lambda_{\text{em}}$  values for the pyridone and hydroxypyridine luciferin tautomers.



pyridone  
**oxy 2.3** X = NH, Y = CH  
**oxy 2.4** X = CH, Y = NH

hydroxypyridine  
**oxy 2.3** X = N, Y = CH  
**oxy 2.4** X = CH, Y = N

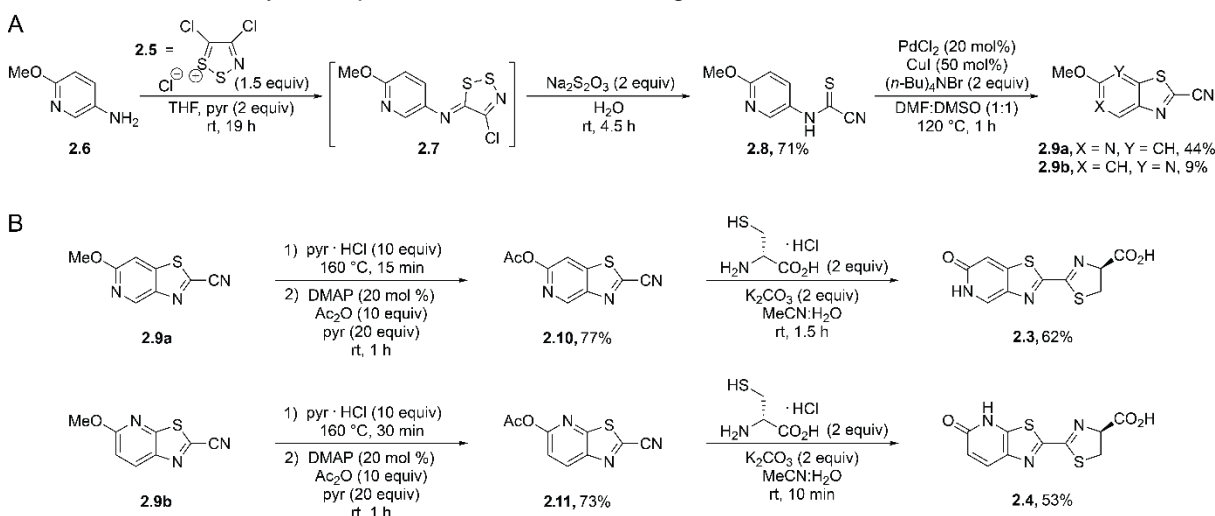
Compound	Predicted $\Delta E_{\text{HOMO-LUMO}}$ [a] [nm]		Bioluminescence $\lambda_{\text{em}}$ [nm]
	Pyridone	Hydroxypyridine	
<b>oxy 2.3</b>	403	582	570
<b>oxy 2.4</b>	349	510	530

[a] Wavelengths were calculated from predicted HOMO/LUMO energies from DFT calculations (B3LYP functional and 6-311 + G\*\* basis set).

We prepared the luciferins using a route previously developed in our group [25,47] (Scheme 2-11). This method relies on an auxiliary dithiazolium reagent (Appel's salt, **2.5**) [48] to access the annulated thiazole of the luciferin core. Appel's salt adducts can be readily formed and fragmented to provide cyanothioformamides [48] in a single pot; these latter intermediates are good candidates for cyclization via C-H activation [49]. To prepare the adduct en route to the pyridone luciferins, aniline **2.6** was first condensed with Appel's salt **2.5**. The resulting product (**2.7**) was subsequently fragmented [48] to access cyanothioformamide **2.8**. Oxidative cyclization of **2.8** with a Pd catalyst [49] provided intermediates **2.9a-b**. These regioisomers could be readily separated via flash column chromatography. The cyclized isomers were then subjected

to pyridine hydrochloride at elevated temperatures to remove the methyl protecting groups. Some undesired N-methylation was observed during this process. Fortunately, the unwanted byproducts could be readily separated from the desired cyanobenzothiazoles upon treatment with acetic anhydride and subsequent chromatography. Condensation of **2.10** and **2.11** with D-cysteine (under basic conditions) ultimately afforded luciferins **2.3** and **2.4**.

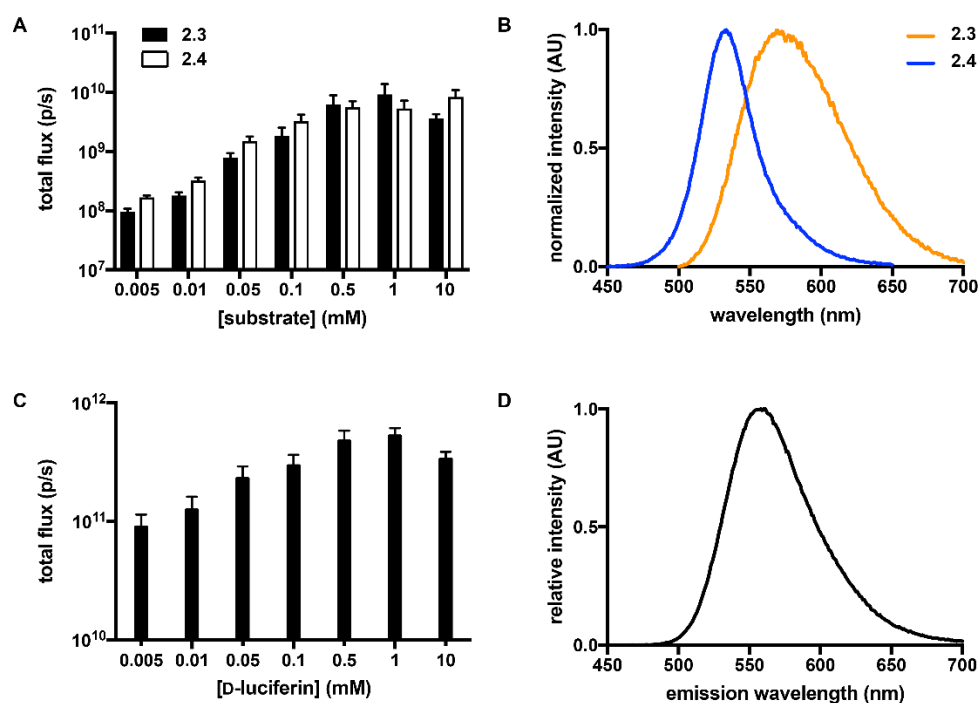
**Scheme 2-1. Synthesis of pyridone luciferins.** A) Formation of the thiazolo-pyridyl core. B) Condensations with D-cysteine provided the desired analogs.



## 2.3 Bioluminescence assays with luciferin analogs

With the desired pyridone luciferins in hand, we evaluated their light-emitting properties. Analogs **2.3** and **2.4** were first incubated with Fluc in the presence of ATP. Light emission was measured using a cooled CCD camera. Dose-dependent photon production was observed with both analogs, although the outputs were lower than that achieved with the native substrate, D-luciferin (Figure 2-2A,C). At saturating doses of

**2.3** or **2.4**, light emission reached ~1-3% of output achieved with the native pair. While weak, such outputs are on par with other luciferins comprising altered heterocyclic cores [25,30,34-35].

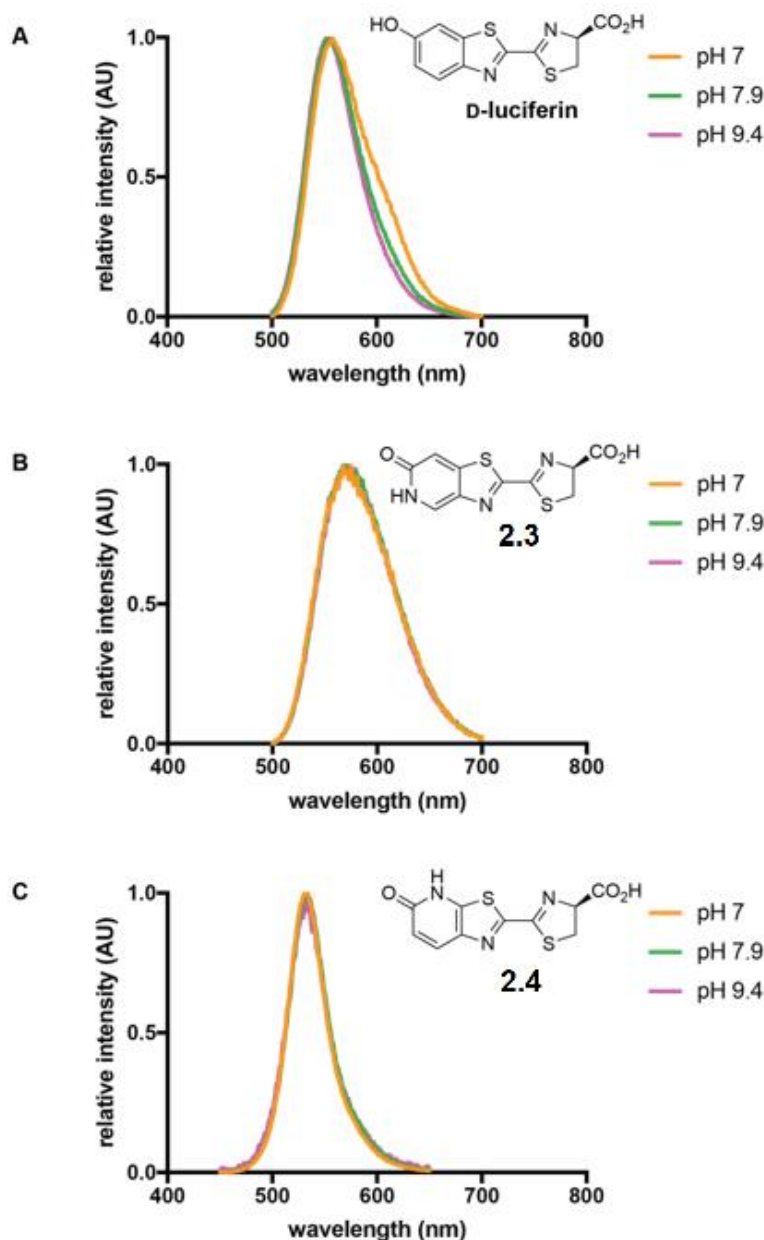


**Figure 2-2. Fluc catalyzes light emission with pyridone analogs** **A)** Pyridone analogs exhibit dose-dependent light emission. Analogs **2.3** (black) and **2.4** (white) (5  $\mu$ M – 10 mM) were incubated with Fluc (1  $\mu$ g) in imaging buffer. Images were acquired immediately post-Fluc addition. Error bars represent the standard error of the mean for  $n \geq 3$  experiments. **B)** Bioluminescence emission spectra for **2.3** ( $\lambda_{max} = 570$  nm) and **2.4** ( $\lambda_{max} = 530$  nm). Each analog was incubated with Fluc in imaging buffer and emission spectra were acquired over a range of wavelengths. **C)** D-luciferin (0-10 mM) was dissolved in imaging buffer, and photons were acquired one minute after Fluc (1  $\mu$ g) addition. Error bars represent the standard error for the mean of  $n = 3$  experiments. Light emission at saturating substrate conditions agrees with previous reports [50]. **D)** D-luciferin exhibits maximum bioluminescence emission at 566 nm (pH 7.6), consistent with previous reports [51].

As predicted, the luciferin analogs exhibited distinct emission profiles. The bioluminescence spectra for **2.3** and **2.4** are shown in Figure 2-2B. The peak emission for pyridone **2.3** ( $\lambda_{em} = 570$  nm) was ~40 nm red-shifted compared to analog **2.4** ( $\lambda_{em} =$



530 nm). These wavelengths flank the emission spectrum for D-luciferin [51] ( $\lambda_{em} = 566$  nm, Figure 2-2D), providing further evidence that modifications to the luciferin heterocycle can impact the color of bioluminescent light. The emission maxima also matched the values predicted for the hydroxypyridine (anionic) emitters shown in Table 2-1. The spectra did not vary from pH 7-9 (Figure 2-3), further suggesting that the light emitter is a single species over this physiological range. The emission wavelength of D-luciferin, by contrast, varies under these conditions [51].

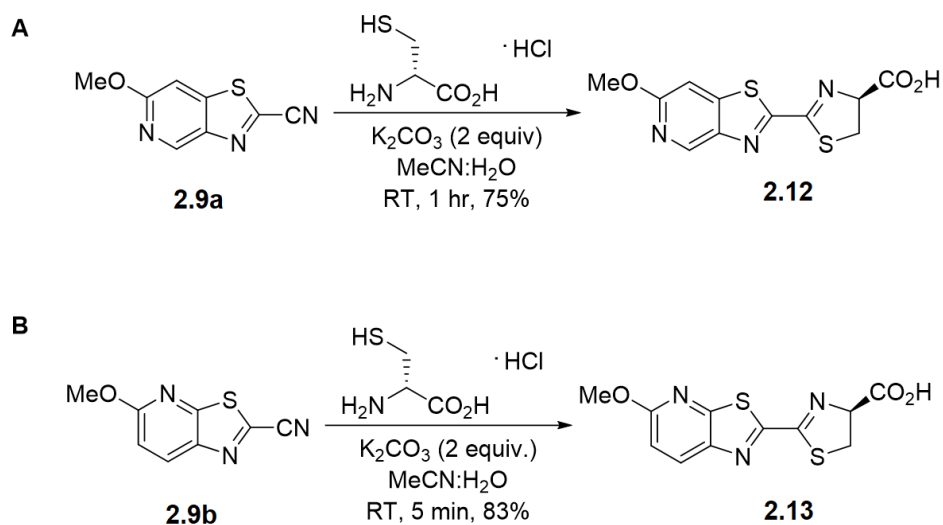


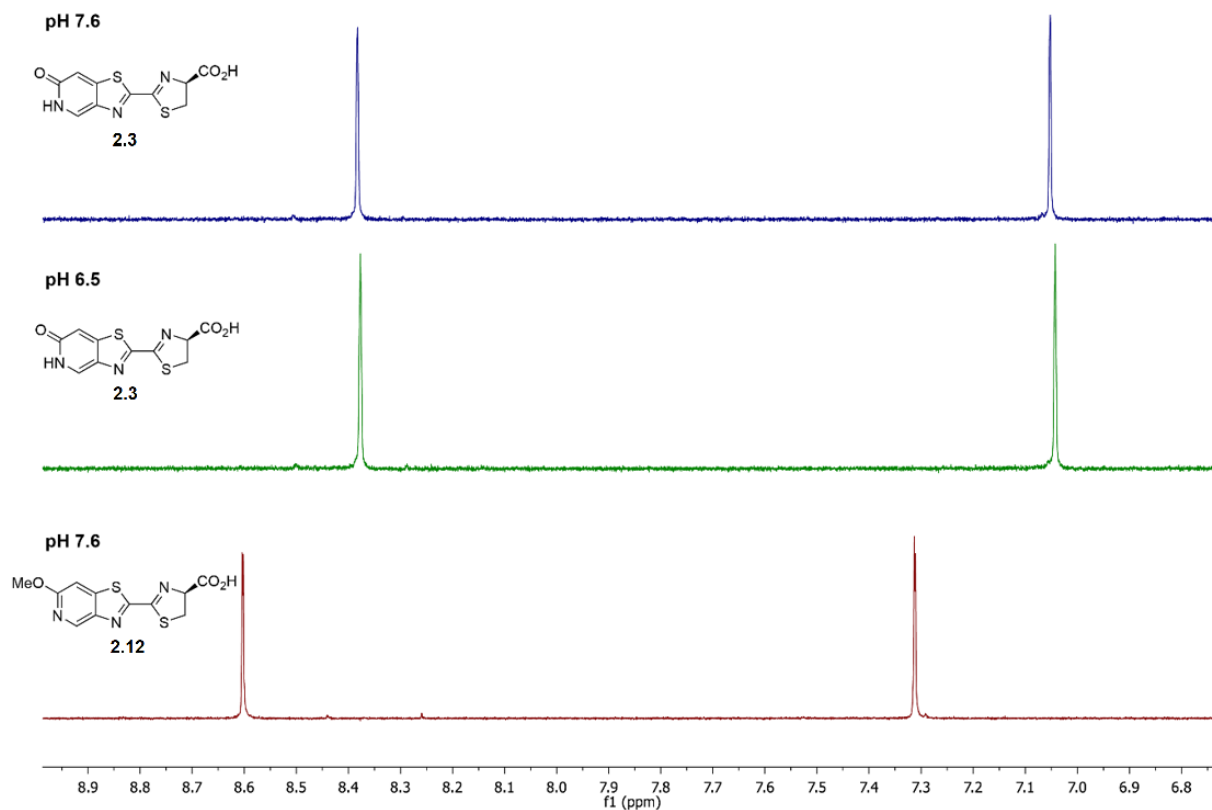
**Figure 2-3. Bioluminescence spectra at various pH values.** **A)** D-Luciferin, **B)** **2.3**, and **C)** **2.4** were dissolved in imaging buffer (50 mM Tris-HCl, 0.5 mg/ml BSA, 0.1 mM EDTA, 1 mM TCEP, 2 mM MgSO<sub>4</sub>, at pH 7-9.4) with ATP (1 mM) and LiCoA (0.5 mM). Bioluminescence was initiated by the addition of Fluc (0.01 mg/mL) and spectra were acquired.

While photophysical data suggested that the oxidized forms of **2.3** and **2.4** emit light as hydroxypyridines, the molecules appear to exist as pyridones in the ground state. NMR analyses were used to compare **2.3** and **2.4** to methylated and tautomerically locked

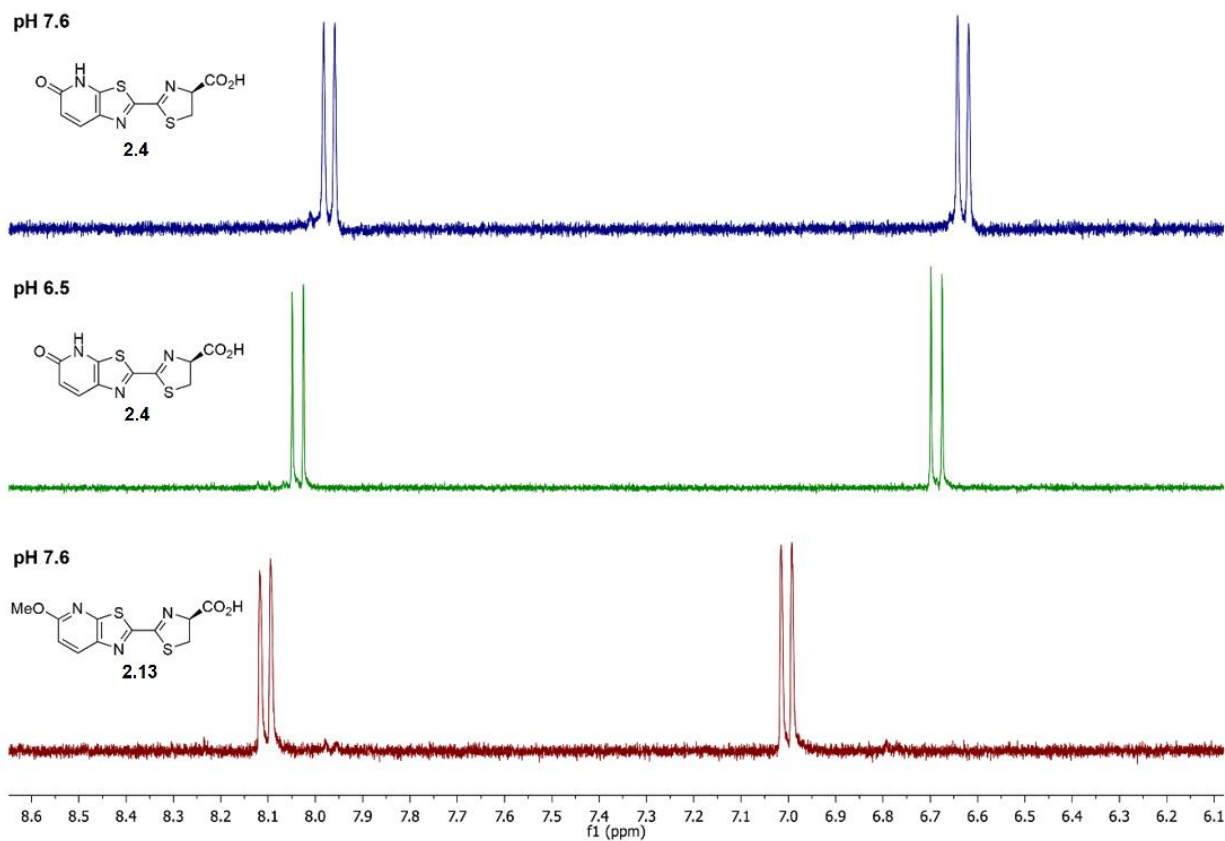
analog (2.12 and 2.13, Schemes 2-2A and B). The aryl proton resonances of 2.3 and 2.4 were shifted upfield relative to the locked analogs, consistent with the pyridone assignment [52] (Figures 2-4 and 2-5). Deprotonation of the analogs was observed at pH 9.5 (Figure 2-6), though, suggesting that N-H proton is susceptible to transfer in the excited state similar to the phenolic O-H in D-luciferin [41-42].

**Scheme 2-2. Synthesis of tautomerically locked analogs A) 2.12 and B) 2.13.**

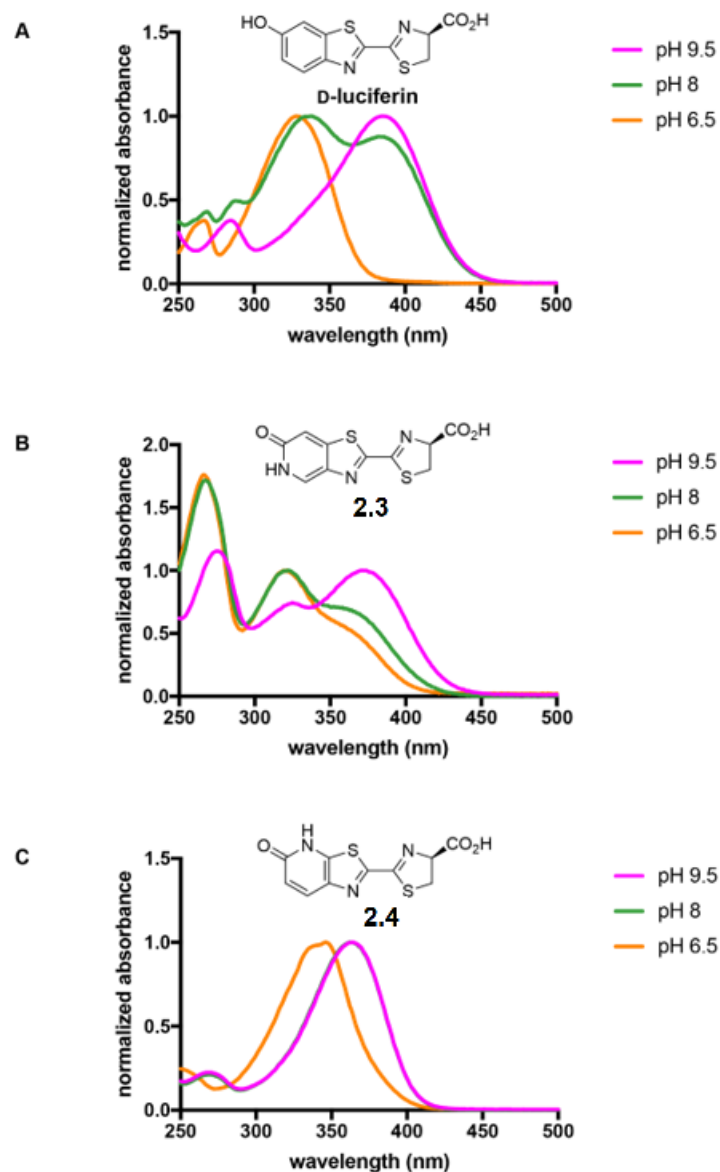




**Figure 2-4.** <sup>1</sup>H-NMR spectra of tautomeric luciferins **2.3** and **2.12**. Luciferin analogs **2.3** or **2.12** (10 mM) were dissolved in deuterated phosphate buffer (pH 6.5-7.6). NMR spectra were recorded immediately upon dissolution.



**Figure 2-5. <sup>1</sup>H-NMR spectra of tautomeric luciferins 2.4 and 2.12.** Luciferin analogs **2.4** or **2.13** (10 mM) were dissolved in deuterated phosphate buffer (pH 6.5 or 7.6). NMR spectra were recorded immediately upon dissolution.

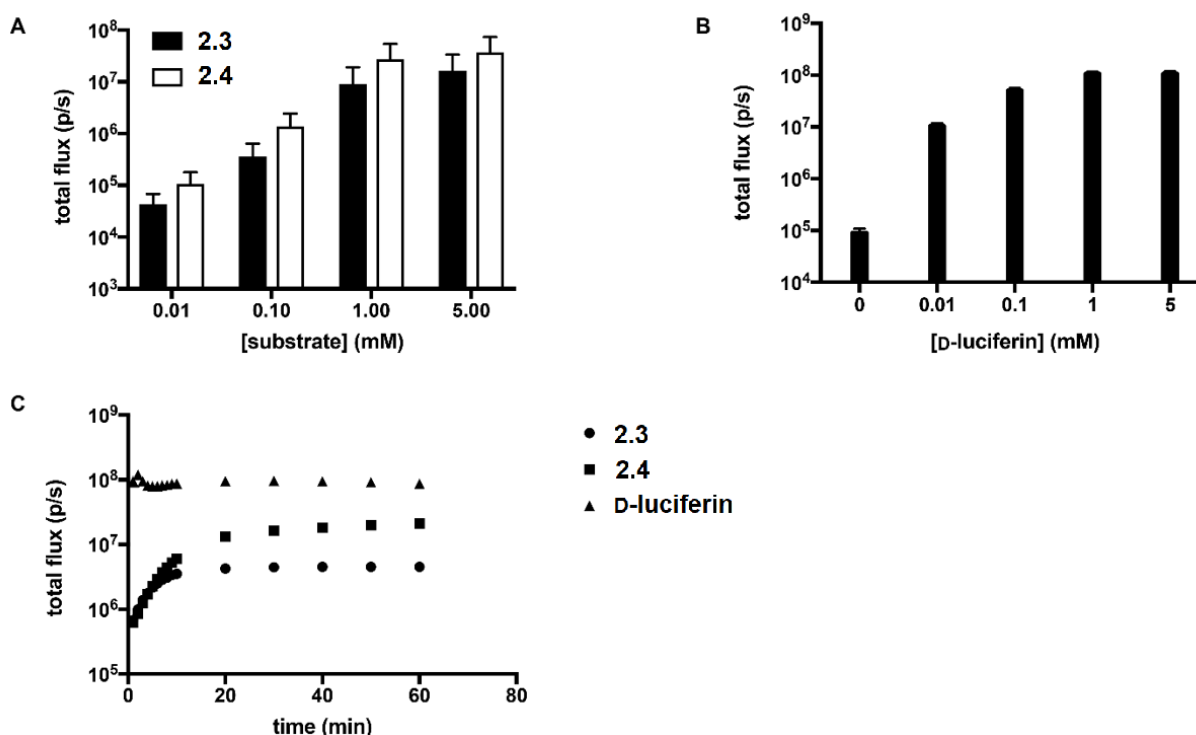


**Figure 2-6. Absorbance spectra of luciferin analogs.** **A)** D-Luciferin and luciferin analogs **B) 2.3** and **C) 2.4** were dissolved in buffer (bis-tris propane 20 mM, pH 6.5-9.5). Absorbance spectra were normalized.

## 2.4 *In cellulo* imaging

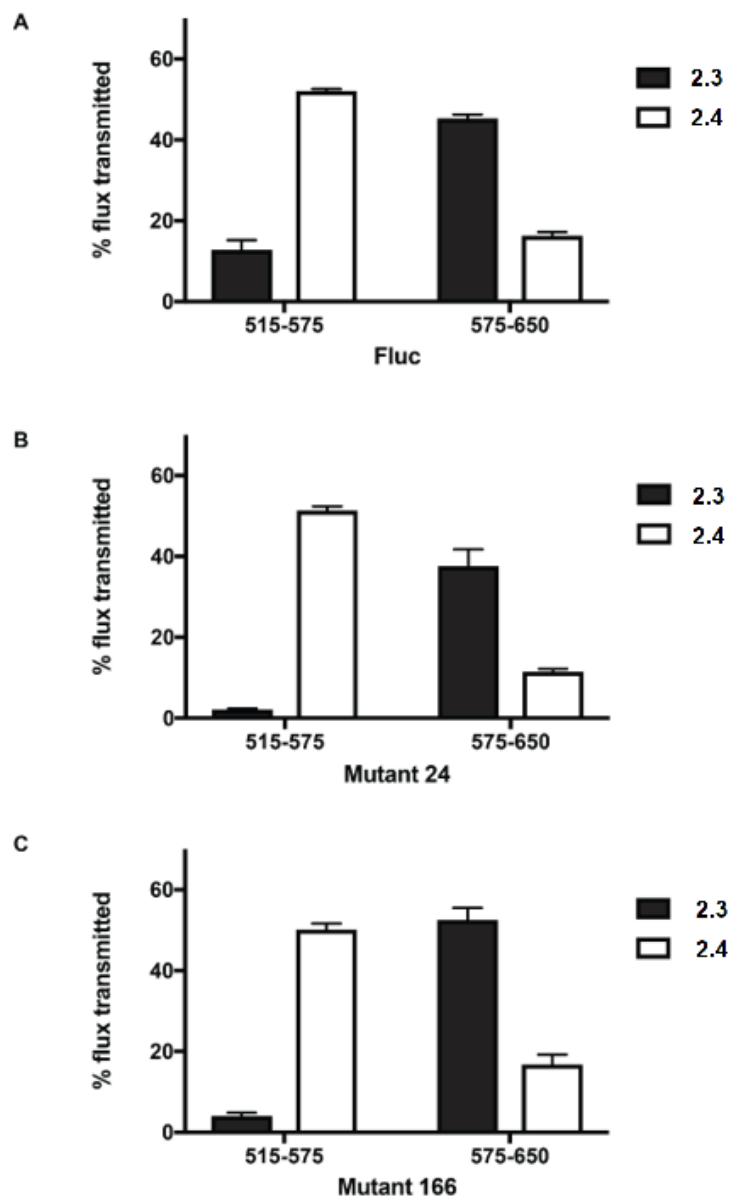
The pyridone luciferins were sufficiently cell permeable for bioluminescence imaging applications. Analogs **2.3** and **2.4** (0.01-5 mM) were incubated with Fluc-expressing HEK293 cells. Photon emission was measured over time. Stronger light outputs were observed with increasing concentrations of both analogs (Figure 2-7A). At

saturating doses (~ 5 mM), maximum light emission was achieved 20 minutes post-luciferin administration (Figure 2-7C). The intensities were ~3-20% of the total photon flux observed with D-luciferin in the same cells (Figure 2-7 A and B). Light emission was also sustained for 1 h, suggesting that the pyridone analogs are suitable for a variety of cellular assays (Figure 2-7C).



**Figure 2-7. Bioluminescence imaging in mammalian cells.** **A)** Analogs **2.3** and **2.4** (0.01-5 mM) were incubated with Fluc-expressing HEK293 cells in PBS (pH 7.6). Images were acquired 10 min post-luciferin addition. Error bars represent the standard error of the mean for  $n \geq 3$  experiments. Sample images are shown. **B)** HEK293 cells expressing Fluc were treated with D-luciferin (0-10 mM) in PBS and imaged after 10 minutes. Error bars represent the standard error of the mean for  $n = 3$  experiments. **C)** HEK293 cells expressing luciferase were treated with D-luciferin, **2.3** or **2.4** (5 mM) in PBS and photon flux was measured over 60 minutes. Error bars represent the standard error of the mean for  $n = 3$  experiments.

Analogs **2.3** and **2.4** could also be distinguished in cells based on their unique spectra (Figure 2-8). As noted earlier, the emission maxima for **2.3** and **2.4** are ~50 nm resolved, which is on par with some luciferases used for dual color imaging [45,53].



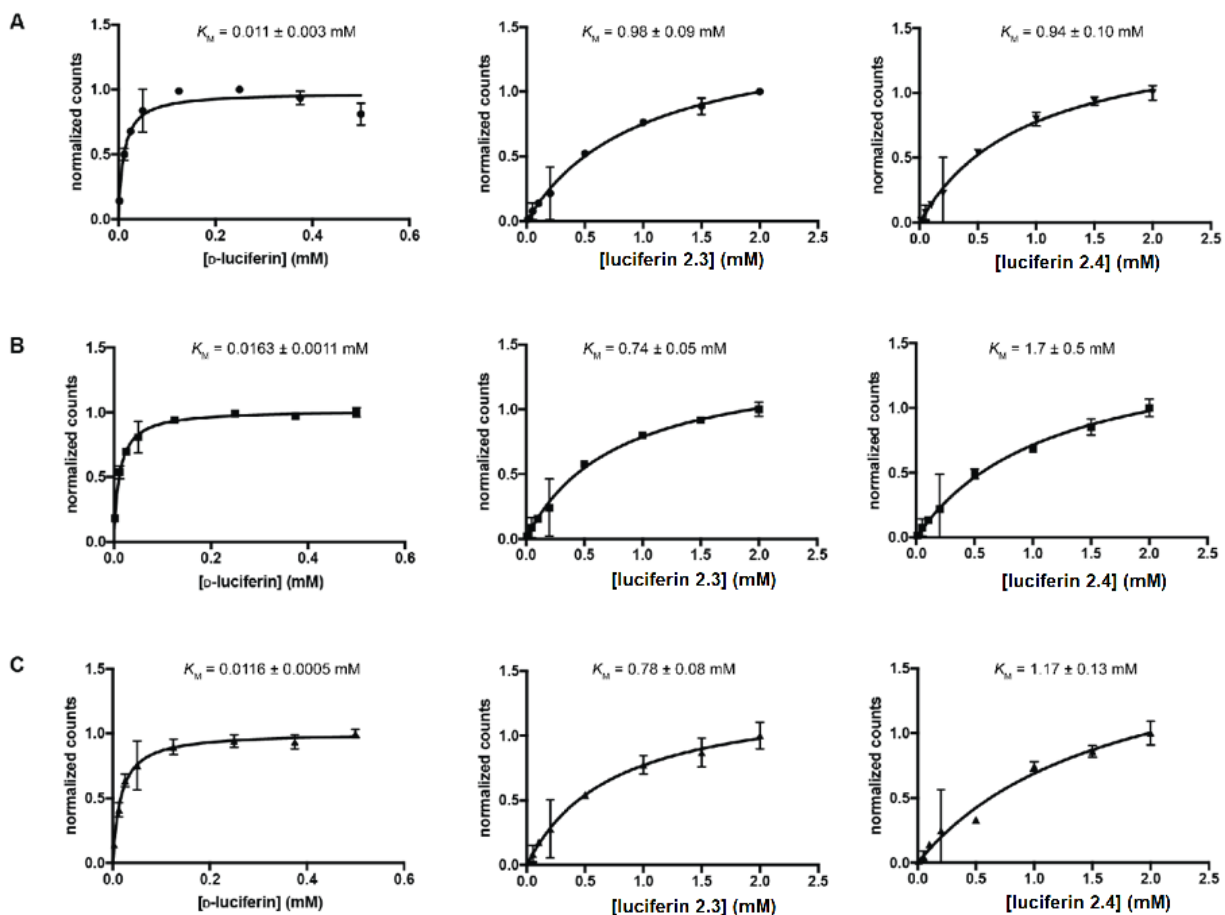
**Figure 2-8. Spectral resolution in mammalian cells.** HEK293 cells expressing **A)** Fluc, **B)** mutant 24, or **C)** mutant 166, were imaged upon incubation with **2.3** or **2.4** (1 mM). Optical filters (515-575 nm or 575-650 nm) were used during the acquisition. Photon outputs acquired with the filters were divided by the total luminescence (no filter) to afford “% flux transmitted”. Error bars represent the standard error of the mean for  $n \geq 3$  experiments. Bioluminescent signals were also normalized to luciferase expression levels (measured via flow cytometry).

## 2.5 Improved imaging with mutant luciferases

While the pyridone analogs were sufficiently robust for use in cellular imaging experiments, biochemical analyses suggested room for improvement. The apparent  $K_m$



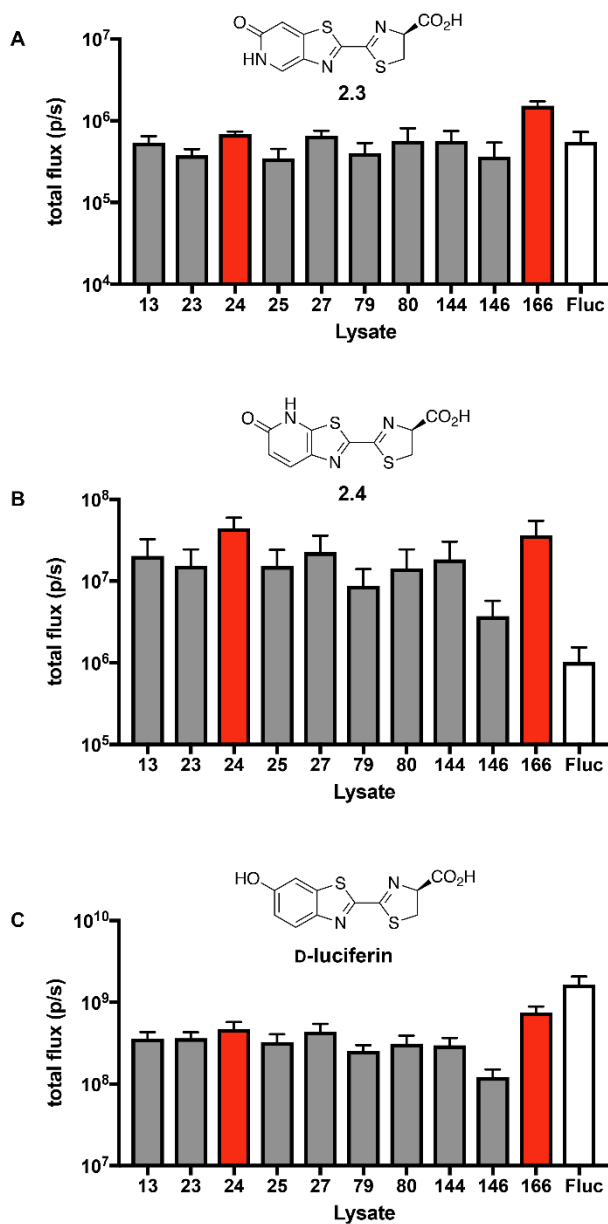
values measured for both luciferins were in the low millimolar range (Figure 2-9). These values are >100-fold greater than the  $K_m$  for D-luciferin [54], suggesting that the pyridones are poor binders of Fluc.



**Figure 2-9. Enzyme kinetics of mutant luciferases.** **A)** Fluc, **B)** mutant 24, and **C)** mutant 166 were incubated with D-luciferin or luciferin analogs **2.3** or **2.4** (0.01 – 2 mM) and light emission values were recorded. Data were then fit according to Michaelis-Menten kinetics and binding constants ( $K_m$ ) were obtained. Kinetic constants are apparent values, determined via measurements of initial rates of light emission. Apparent  $K_m$  values for D-luciferin and Fluc agree with reported values [54]. Error bars represent the standard error of the mean for  $n = 9$  experiments.

To identify enzymes that could better process the analogs, we incubated luciferin **2.4** with a panel of 166 mutant luciferases in bacterial lysate [50]. These enzymes were previously identified from screens with sterically modified luciferins [50] and were thus

known to be functionally active. Upon treatment with **2.4**, ten of the mutant enzymes exhibited 10-80-fold brighter outputs compared to Fluc (Figure 2-10) in lysate.

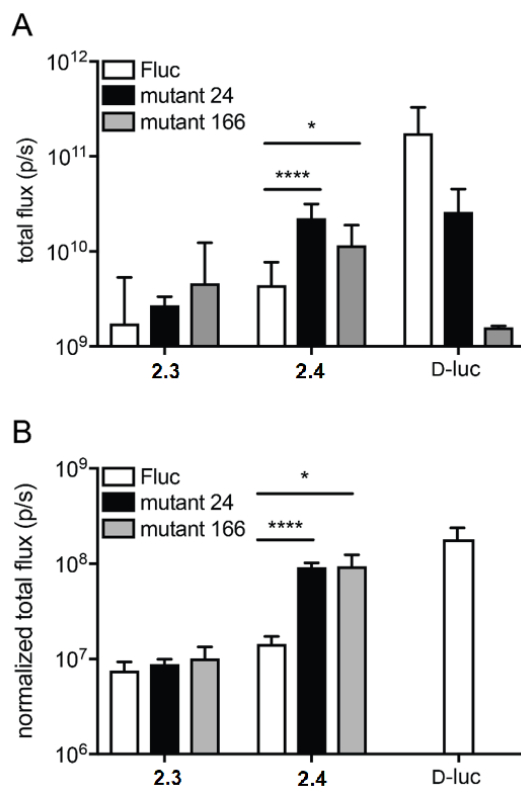


**Figure 2-10. Improved mutant luciferases with luciferin analogs.** Bacteria expressing mutant 24 or 166 (red bars), other mutant enzymes (grey bars) or Fluc (white bars) were selected from primary screens, expanded, and lysed. Lysates were imaged with luciferins **A**) **2.3**, **B**) **2.4** or **C**) D-luciferin. Error bars represent the standard error of the mean for n = 3 experiments.

**Table 2-2. Sequencing analysis of “bright” mutants**

<b>Mutant</b>	<b>Sequence</b>
<b>13</b>	M249F, T252S, F295L, S314T, G316T, A326V, P334S
<b>23</b>	V240A, M249L, L264F, S314T, G316T, K321R
<b>24</b>	<b>M249L, Q283R, S314T, G316T</b>
<b>25</b>	I226V, V240A, M249L, I282T, F295L, S314T, G316T
<b>27</b>	I232T, M249L, S314T, G316T
<b>79</b>	M249L, S314T, G316S
<b>80</b>	M249L, S314T, G316T
<b>144</b>	V241A, F247Y, S314T, G316T
<b>146</b>	V241A, F247L, S314T, G316T
<b>166</b>	<b>M249L, I257F, F295L, S314T, G316T, A326V, P334S</b>

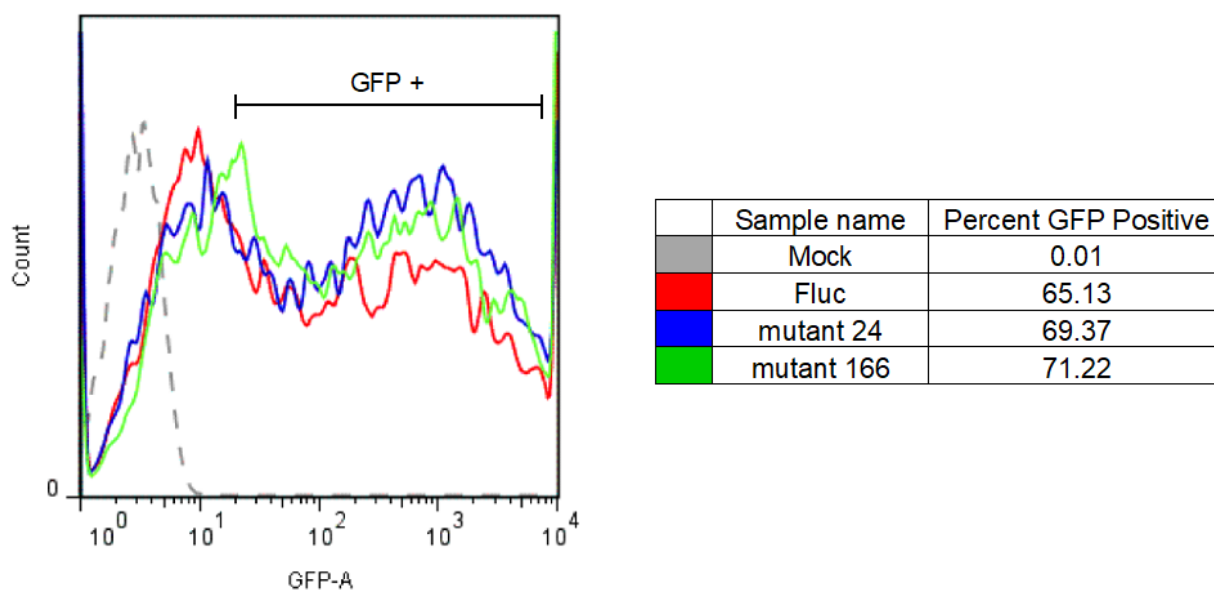
The two "brightest" mutants (24 and 166) were selected for isolation and further characterization (Table 2-2). Upon purification from bacterial expression systems, the luciferases maintained enhanced photon outputs with analog **2.4** (Figure 2-11). Interestingly, increased light emission was not observed with isomeric luciferin **2.3**.



**Figure 2-11.** Improved bioluminescence imaging with mutant luciferases. **A)** Analogs **2.3-2.4** (1000  $\mu$ M) or D-luciferin (D-luc) were incubated with mutant luciferases (24 or 166) or native Fluc and ATP. Light emission was measured and quantified. Error bars represent the standard error of the mean for  $n \geq 3$  experiments. **B)** HEK 293 cells expressing mutant luciferases (24 or 166) or native Fluc were incubated with analogs **2.3-2.4** (1000  $\mu$ M) or D-luciferin (D-luc). Error bars represent the standard error of the mean for  $n \geq 3$  experiments. For A) and B),  $p$ -values (from one-tailed unpaired t tests) are annotated by asterisks (\* $p < 0.05$ , \*\*\*\* $p < 0.0001$ ).

Enzymes 24 and 166 comprise M249L, S314T and G316T mutations. S314T and G316T have been previously shown to improve the catalytic efficiency of Fluc [50]. Furthermore, the G316A mutation is known to affect bioluminescent color [54], which may suggest a change in the polarity of the active site [55]. Mutant 166 further comprised an F295L mutation. This mutation is found in thermostable luciferases [23]. The  $K_m$  values measured for luciferin **2.4** and each mutant were similar to that of native Fluc, suggesting no improvement in substrate affinity (Figure 2-9). Thus, the mutated residues are likely contributing to a change in active site architecture that promotes more favorable substrate turnover.

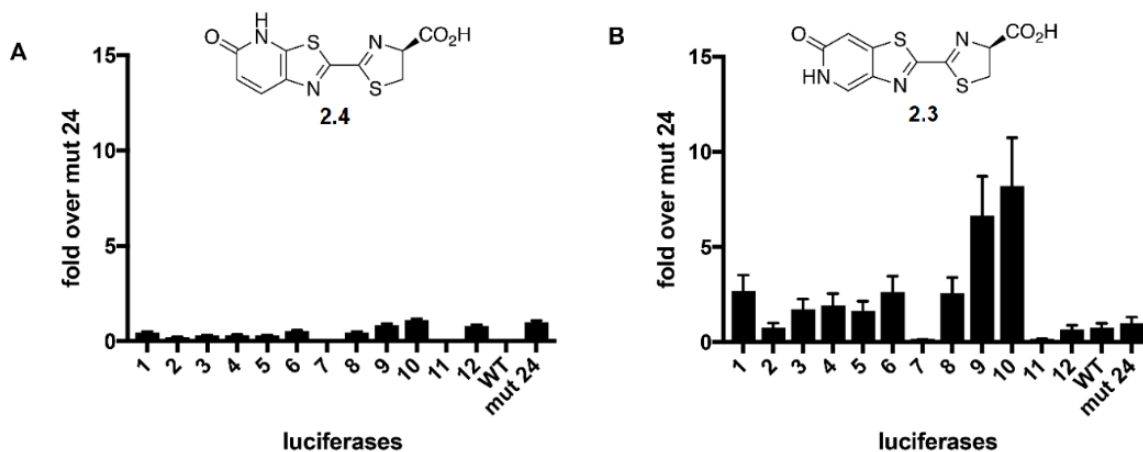
Encouraged by the *in vitro* findings, we analyzed mutants 24 and 166 in cellular studies. Plasmids encoding the luciferases were transfected into HEK293 cells. Luciferase expression was verified and normalized to a fluorescent reporter (GFP) in each case (Figure 2-12). The cells were incubated with either analog **2.3** or **2.4** and bioluminescent photon production was measured. The light output obtained from cells expressing mutant 24 or 166 (in the presence of analog **2.4**) was on par with bioluminescence levels obtained from conventional imaging tools (Fluc/D-luciferin, Figure 2-11). Like the *in vitro* studies, no improvement in the light output with analog **2.3** was obtained. The bioluminescence emission spectra from both pyridone analogs could still be resolved in this setting (Figure 2-8), though, suggesting that the engineered pairs will be useful for multicomponent imaging.



**Figure 2-12. Luciferase expression was measured in cellular samples.** HEK293 cells transiently expressing luciferase (as an IRES-GFP fusion) were analyzed by flow cytometry. Cells exhibiting  $\geq 10$  fold higher GFP fluorescence intensity (GFP-A) than mock cells were considered GFP positive (GFP +). GFP positive cells were counted and compared to the entire population to determine percent positive values for normalization. A representative plot is shown.

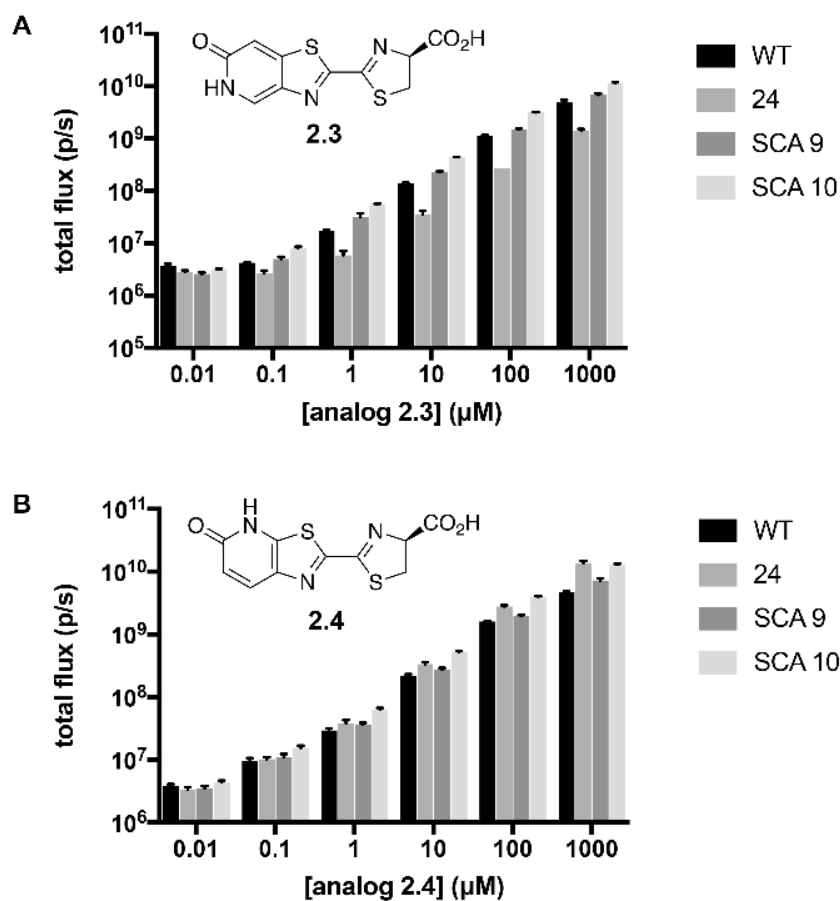
## 2.6 Directed evolution of luciferase mutant 24

To guide further mutagenesis efforts, statistical coupling analysis (SCA) was used to compare phylogenetic data from hundreds of homologous enzymes and identify functionally important mutations [56]. These mutations were systematically introduced to the mutant 24 gene using gene assembly. The resulting library was transformed into *E. coli* and screened for activity with analogs **2.3** and **2.4**. From this screen, two luciferases (mutants 9 and 10) were identified that exhibited a 10-fold increase in photon production with luciferin **2.3** (Figure 2-13). However, no mutants were identified with improved activity with luciferin **2.4**. This result suggested that mutant 24 represents a local maximum on the evolutionary landscape for utilization of luciferin **2.4**. In contrast, fitness of mutant 24 towards luciferin **2.3** could be improved when sampling the same evolutionary trajectory.



**Figure 2-13. SCA library screening identifies improved luciferases.** Bacteria expressing mutant luciferases 1-12, WT, or mutant 24 were expanded and lysed. Lysates were imaged with either **A**) analog **2.4** or **B**) analog **2.3**. Error bars represent the standard error of the mean of  $n = 3$  experiments.

Mutants 9 and 10 were expressed and purified. When assayed with the luciferin analogs, the mutant luciferases retained improved activity with analog **2.3** (Figure 2-14). Mutant 10 exhibited 8-fold improved activity over the first-generation lead (mutant 24) and 2-fold improved activity over native Fluc (Figure 2-14A). The SCA-designed mutants did not exhibit improved activity over native Fluc or mutant 24 with analog **2.4**, a result consistent with bacterial lysate data (Figure 2-14B).



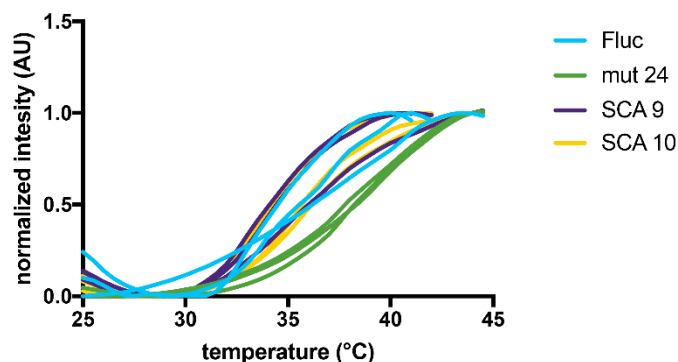
**Figure 2-14. Light output of purified SCA mutants and luciferin analogs.** Analogs **2.3** or **2.4** (0.01-1000  $\mu\text{M}$ ) were incubated with mutant luciferases (24, SCA 9, SCA 10) or native Fluc (0.01 mg/mL) and ATP (1 mM) in imaging buffer. Light emission was measured and quantified. Error bars represent the standard error of the mean for  $n \geq 3$  experiments.

We next sought to understand the origins of improved activity for mutants SCA 9 and 10. These luciferases share beneficial mutations N229T, S239T, G246A, and L264V (Table 2-3). S239T and G246A are known to impart thermostability in other luciferase mutants [56]. When subjected to differential scanning fluorimetry [57], though, no reproducible differences in thermal denaturation curves were observed between mutants 9 and 10 relative to mutant 24 (Figure 2-15).

**Table 2-3.** Sequencing analysis of mutant luciferase leads

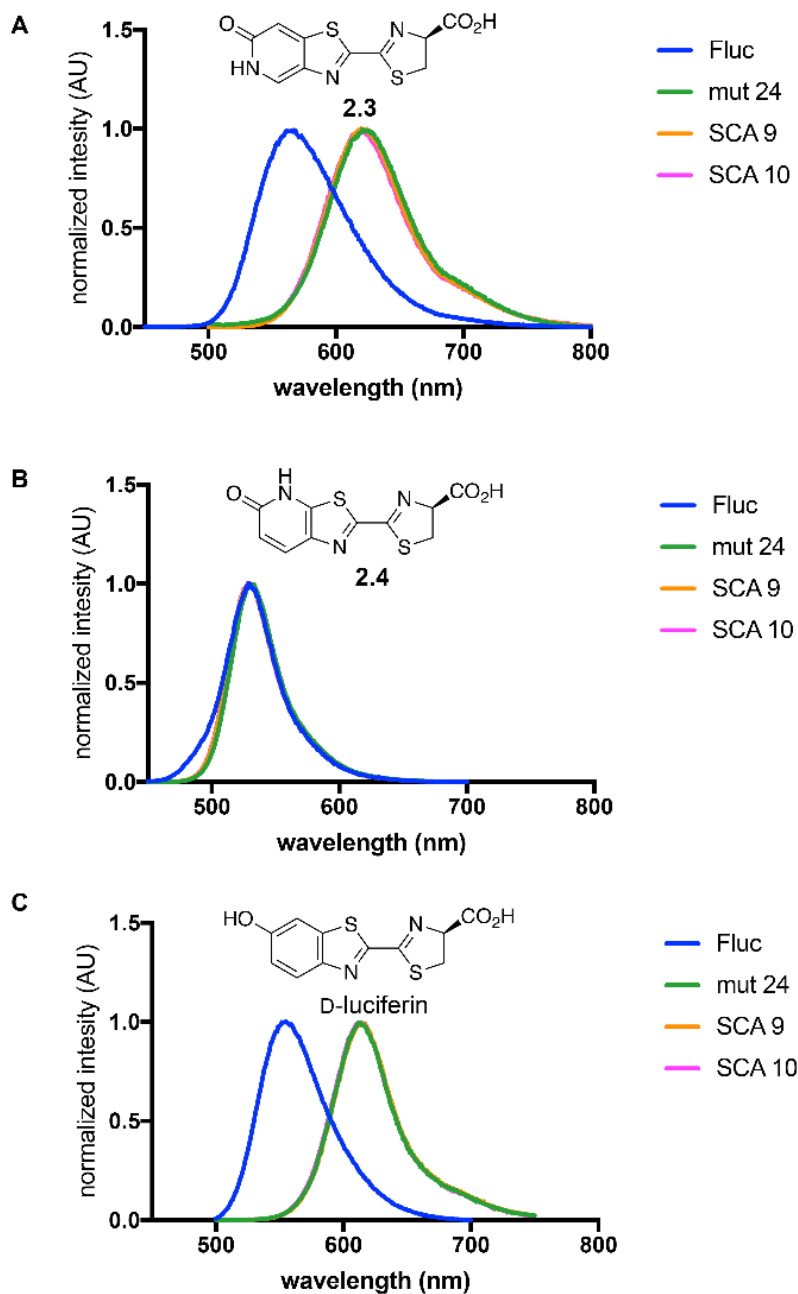
Mutant	P211	N229	I237	S239	G246	M249	F250	T251	L264	Q283	S314	G316
1	S	T	V	T	A	L	F	T	M	R	T	T
2	P	N	V	G	A	L	F	T	M	R	T	T
3	A	N	I	S	A	L	F	T	L	R	T	T
4	S	N	I	S	A	L	F	T	M	R	T	T
5	S	N	V	S	A	L	F	T	V	R	T	T
6	S	T	I	T	A	L	F	T	L	R	T	T
7	P	N	V	S	A	L	A	I	V	R	T	T
8	P	N	V	T	A	L	F	T	V	R	T	T
9	P	T	V	T	A	L	F	T	V	R	T	T
10	P	T	I	T	A	L	F	T	V	R	T	T
11	A	N	I	V	A	L	C	I	M	R	T	T
12	A	N	V	S	G	L	F	T	M	R	T	T
mutant 24	P	N	I	S	G	L	F	T	L	R	T	T





**Figure 2-15.** Thermal denaturation of mutant luciferases. Protein denaturation was monitored as a function of temperature via differential scanning fluorimetry. Luciferases (1 mg/mL) in storage buffer (25 mM Tris pH 7.8, 200 mM  $(\text{NH}_4)_2\text{SO}_4$ , 1 mM DTT, 1 mM EDTA) were incubated with 0.2% v/v SYPRO dye. Temperatures were sampled from 25-95°C and fluorescence was measured. Three replicate melt curves are plotted for each protein.

N229T is a natural mutation found in pH insensitive luciferases [58]. N229 is involved in an H-bond network in the binding pocket comprised of water molecules and other residues [59]. This complex is known to stabilize a closed active site conformation, promoting yellow-green bioluminescence. Disruption of the H-bond network causes an open binding pocket conformation, enabling red bioluminescence [59-61]. Mutation at N229T could possibly cause a similar change in binding pocket conformation. Indeed, the SCA mutants exhibit a red shift in emission with D-luciferin and analog **2.3** (Figure 2-16). Interestingly, the first generation mutant 24 also exhibits red-shifted emission, though it comprises mutations distal from the water network. In lieu of structural information for mutant 24 and the SCA mutants 9 and 10, we can only speculate on the binding pocket conformation of these luciferases and the mutational origins of improved activity.



**Figure 2-16. Bioluminescence spectra of SCA mutant luciferases.** Mutant luciferases (24, SCA 9, SCA 10) or native Fluc (0.1 mg/mL) were incubated with luciferins (100  $\mu$ M) **A**) **2.3**, **B**) **2.4** or **C**) D-luciferin and ATP (1mM) in imaging buffer and emission spectra were acquired over a range of wavelengths.

The lead mutants from this screen, SCA 9 and 10, contain improvements on two fronts. They exhibit photon output brighter than that of native Fluc with analog **2.3** and

produce red-shifted bioluminescence. Red-shifted emission is desirable for *in vivo* imaging applications. In addition, red-shifted probes are further resolved from blue emitting probes, facilitating multispectral imaging.

## 2.7 Conclusion

In summary, we identified two luciferin analogs with spectrally resolved emission wavelengths. These probes can be selectively detected using appropriate filter sets, suggesting they are useful as multispectral imaging probes. However, these probes were still weak emitters compared to the native substrate. To identify a luciferase variant with improved activity, a panel of luciferase mutants was screened with the luciferin analogs. From this screen, several luciferase mutants were identified with improved activity with luciferin **2.3**. We sought to further evolve the best performing mutant, targeting residues from phylogenetic analysis of homologous enzymes. Interestingly, luciferin **2.4** was not improved, suggesting a local maximum has been reached on the evolutionary landscape. In contrast, luciferin **2.3** was improved when screened against the same luciferase library. The improved luciferases also exhibit red-shifted light, useful for imaging applications. Importantly, these improvements could be achieved in only 1-2 rounds of mutagenesis, highlighting the utility of structurally diverse luciferins and evolved mutant luciferases. Collectively, the engineered luciferases and luciferins expand the collection of bioluminescent reporters available for multispectral imaging. Future work will involve validating luciferin 2.3 and engineered luciferases as imaging probes in living cells.

## **2.8 Materials and methods**

### **2.8a Computational methods**

Calculations were performed with Spartan Student Edition Version 6. Organic structures were modeled and energy minimized.  $\Delta E_{\text{HOMO-LUMO}}$  values were determined using the B3LYP level of theory and basis set 6-311+G\*\*.

### **2.8b Recombinant protein expression and purification**

Plasmids encoding Fluc or Fluc mutants (pET vectors) were transformed into chemically competent *E. coli* BL21 cells. Cultures (1 L) were grown in LB media (containing 40  $\mu\text{g/mL}$  kanamycin, LB-Kan) at 37 °C (with shaking). Cultures were grown to mid-log phase ( $\text{OD}_{600} \sim 0.8 - 1.0$ ), induced with 0.5 mM IPTG, and incubated at 22 °C for 16 – 18 h. Cells were harvested by centrifugation (4000  $\times g$ , 4 °C), then resuspended in 20 mL of 50 mM  $\text{NaPO}_4$ , 300 mM NaCl, 1 mM DTT, and 1 mM PMSF (pH 7.4). Lysozyme (1 mg) was added, and the mixtures were sonicated and centrifuged (10000  $\times g$ , 1 h at 4 °C). Luciferases were purified from the clarified supernatants using nickel-affinity chromatography (BioLogic Duo Flow Chromatography System, Bio-Rad). Protein isolates were dialyzed into 25 mM Tris-acetate buffer (pH 7.8) containing 1 mM EDTA and 0.2 mM ammonium sulfate (4 °C). DTT (1 mM) and 15% glycerol were added to the dialyzed samples prior to storage at -20 °C. Final protein concentrations were determined via UV spectroscopy (NanoDrop 2000c). Samples were also analyzed by SDS-PAGE and Coomassie staining.

### **2.8c General bioluminescence imaging**

Samples were imaged using an IVIS Lumina (Xenogen) system equipped with a cooled CCD camera. Images were acquired using small to medium binning. Exposure times varied from 1 s to 10 min. All acquisition parameters were controlled using Living Image software. Living Image was also used to measure photon flux values from defined regions of interest. These data were exported to Microsoft Excel for further analyses.

### **2.8d Screening collection of 166 previously characterized luciferase mutants**

*E. coli* BL21 cells (in 50% glycerol) expressing Fluc or mutant luciferases (see section 2.8x for list of mutants) were added to 5 mL of LB-Kan. The cultures were incubated at 37 °C for 2-5 h (with shaking) until OD<sub>600</sub> values reached ~0.8. A portion of each culture (0.5 mL) was reserved for sequencing analysis. IPTG (0.5 mM) was added to the remainder of the cultures and protein expression was induced for 2 h at 30 °C or 12 h at 25 °C. After induction, the cultures were centrifuged for 10 min at 3400 xg (4 °C). The supernatant was removed, and the pellet was resuspended in 0.4 – 0.6 mL of lysis buffer (50 mM Tris•HCl, 500 mM NaCl, 0.5% (v/v) Tween, 5 mM MgCl<sub>2</sub>, pH 7.4). Lysates were added to black 96-well plates (90 µL/well, in triplicate). Lysates were imaged using a stock solution of 10X luciferin and 10X ATP in PBS (1 mM luciferin, 10 mM ATP). The stock solution (10 µL) was added to each well, 6 wells at a time (final [luciferin] = 100 µM and [ATP] = 1 mM). Images were acquired as described above. Mutants with >10-fold light output compared to Fluc (in the same assay) were further analyzed. The reserved cultures for these mutants were added to 3.5 mL of LB-Kan and grown overnight in a shaking incubator (37 °C). Plasmid DNA was then harvested and sequenced.

### **2.8e Light emission assays with recombinant luciferase**

All bioluminescence assays were carried out in triplicate, using black, flat-bottom, 96-well plates (Greiner). Each well contained purified Fluc or a mutant luciferase (0-1  $\mu$ g), luciferin substrate (0-10 mM), ATP (Sigma, 1 mM), coenzyme-A (trilithium salt, LiCoA Calbiochem, 0.5 mM), and reaction buffer (50 mM Tris-HCl, 0.5 mg/ml BSA, 0.1 mM EDTA, 1 mM TCEP, 2 mM  $MgSO_4$ ), totaling 100  $\mu$ L. Additionally, all non-enzyme assay components were premixed in the wells prior to luciferase addition. Wells containing 0-10 mM substrate were prepared at pH 7.6 or 9. Images for all assays were acquired as described above.

### **2.8f Cellular bioluminescence assays**

HEK293 cells were grown in DMEM media supplemented with fetal bovine serum (FBS, 10%), penicillin (10 U/mL), and streptomycin (10  $\mu$ g/mL). The cells were cultured in a 5%  $CO_2$  humidified incubator at 37  $^{\circ}C$ . Fluc-expressing HEK cells were provided by the Contag lab (Stanford). Transient transfections (with pcDNA-IRES-GFP vectors) were performed using cationic lipid formulations (Lipofectamine 2000, Invitrogen). For light emission assays, approximately 100,000 cells were plated in black 96 well plates in PBS (pH 7.6) containing 1% BSA. The cells were then incubated with luciferin analogs (0.01-5 mM) in PBS (pH 7.6) containing 1% BSA. Bioluminescence images were acquired as above.

### **2.8g Flow cytometry**

GFP expression was verified in transiently transfected cells via flow cytometry. Cells were trypsinized and washed with FACS buffer (PBS with 1% BSA) prior to analysis. For each sample, 10,000 cells were collected and data were analyzed using FloJo software (Tree Star, Inc.).

### **2.8h Bioluminescence emission spectra**

Emission spectra for **2.3**, **2.4** and D-luc (**1**) were recorded on a Horiba Jobin-Yvon FluoroMax®-4 spectrometer. Each luciferin (1 mM) was incubated in a cuvette (10-mm path length) with ATP (1 mM), LiCoA (0.5 mM) and reaction buffer totaling 900  $\mu$ L. Purified Fluc (10  $\mu$ g) was added immediately prior to data acquisition. The excitation and emission slits on the instrument were adjusted to 0 and 5 nm, respectively. Emission data were collected at 1 nm intervals from 400-700 nm at ambient temperature. The acquisition times of 1 s per wavelength were used. Light emission values were measured in relative luminescence units (RLU), and the intensities were normalized to the peak emission intensity for each analog.

### **2.8i Bioluminescence kinetic measurements**

Measurements were acquired on a Tecan F200 PRO luminometer with a neutral density filter. Reactions were performed in black 96-well, flat-bottom plates (Greiner). Solutions of luciferin (20  $\mu$ M-4 mM analog) in bioluminescence buffer (20 mM Tris-HCl pH 7.6, 2 mM MgSO<sub>4</sub>, 2 mM ATP, 0.1 mM EDTA, 1 mM TCEP, 0.5 mg/mL BSA) were prepared, and 50  $\mu$ L was added to each well. The luminescence from each well was measured for 1.5 s prior to the addition of Fluc in bioluminescence buffer (50  $\mu$ L). For wells containing

D-luc, a 0.002 mg/mL solution of enzyme was used. For all other compounds, a 0.02 mg/mL solution of enzyme was used. Luminescence values were recorded every 0.1 s over a 13.5 s period. Samples were analyzed in triplicate and three runs of each compound-enzyme pair were performed. Emission maxima were determined by averaging the five maximum photon outputs per run. Photon outputs were then normalized and plotted as a function of concentration.  $K_m$  values were determined using nonlinear regression analyses and robust fit outlier removal in GraphPad Prism (version 6.0f for Macintosh, GraphPad Software).

### **2.8j Absorbance and fluorescence emission spectra**

Absorbance spectra were acquired using a Cary 50 absorption spectrometer. Fluorescence spectra were acquired using a Cary Eclipse spectrometer (with 1 nm excitation and emission slit widths). Solutions were prepared via a 1:100 dilution of luciferin stock solution (10 mM in DMSO). To analyze the pH dependence of absorbance and fluorescence, compound stocks were diluted into BIS-TRIS propane buffer (20 mM, pH 6.5-9.5).

### **2.8k NMR analyses**

Luciferin analogs (1 mM) were dissolved in deuterated phosphate buffer saline (20 mM phosphate, pH 7.4, containing 4.5 mM KCl and 228 mM NaCl).  $^1\text{H-NMR}$  spectra were acquired on Bruker Advanced spectrometers at ambient temperature.

### **2.9I SCA library generation**



A 330 bp fragment of the luciferase gene was targeted for gene assembly, as described previously [56]. Primers bearing SCA mutations of interest were commercially synthesized (see primer list in section 2.9m). Mutations corresponding to the mutant 24 gene were left constant. Briefly, primers were commercially synthesized, diluted to 10  $\mu$ M in water, and mixed in equimolar portions. The primer mix (4  $\mu$ L) was assembled in a PCR in Q5 buffer containing 1 mM dNTPs and Q5 polymerase (NEB) using the following conditions 94 °C 2 min, 25 cycles of 98 °C 15 s, 60 °C 15 s, 72 °C 15 s, followed by a 72 °C extension for 10 min). Melting, annealing and elongation steps were repeated for 30 cycles. The PCR mixture (1  $\mu$ L) was then directly added to a second PCR using Q5 buffer containing 1 mM dNTPs, Q5 polymerase, 1  $\mu$ M AL158 and AL161 primers (see primer list) using the same thermal cycling conditions as the assembly reaction. PCR product was purified using PCR Purification kit (Qiagen). The purified assembled gene product was mixed with purified PCR-generated vector (generated using vector amplification primers, see primer table) in a 5:1 insert:vector ratio in a Gibson Assembly mix (New England BioLabs) using manufacturer's protocols. The post-assembly mixture was subject to DPN1 digest overnight and purified using PCR purification kit (Qiagen). The resultant purified library was transformed into DH5 $\alpha$  cells (New England BioLabs), and plated on large LB/agar plates containing 1 mM kanamycin. Following overnight incubation at 37°C, library was recovered in LB, and plasmid was purified using Qiagen plasmid purification kit.

Linear vector was constructed from template pET28a-luc2, described previously [62]. Mutation at S314T G316T was introduced using a Quikchange™ protocol. Quikchange™ mutagenesis was performed in Q5 buffer containing 1 mM dNTPs, Q5

polymerase, Quikchange™ primers (BZ\_forquickchange\_314 and BZ\_revquickchange314) at 0.3 mM and pET28a-luc2 template (0.3 ng/μL) using the following conditions (94 °C 2 min, 25 cycles of 98 °C 15 s, 73.2 °C 15 s, 72 °C 210 s, followed by a 72°C extension for 10 min). Melting, annealing and elongation steps were repeated for 30 cycles. The resulting plasmid was subject to DpnI digest overnight and purified using Qiagen plasmid purification kit. This purified vector was transformed into DH5α cells and plasmid DNA was extracted using a Zymo™ kit. The recovered plasmid will be referred to as pET28a-luc2-S314T\_G316T.

pET28a-luc2-S314T\_G316T was linearized in Q5 buffer containing 1 mM dNTPs, Q5 polymerase, 0.3 mM vector amplification primers (BZ\_SCA\_vector\_amp\_for and BZ\_SCA\_vector\_amp\_rev) using the following conditions (94 °C 2 min, 25 cycles of 98 °C 15 s, 72 °C 15 s, 72°C 180 s, followed by a 72 °C extension for 10 min). Melting, annealing and elongation steps were repeated for 30 cycles. The resulting linearized vector was subjected to DpnI digest overnight and purified using a PCR purification kit (Qiagen).

## 2.9m Primer list

### Gene assembly primers\*

\*codons encoding parent mutant 24 are underlined, SCA mutagenic sites are described previously.

**AL148\_F1:** GAACAGTAGTGGATCCACCGGATTGCCCAAGGGCGTAGCCCTANCTCATCGCACCGCTT

**AL149\_F2:** GTGTCCGATTCAGTCATGCCCGCGACCCCATCTTCGGCAMTCAGATCATCCCCGACACC

**AL150\_F3\_M249L:**

GCTRTTCTCRBTGTGGTGCCATTTACCACGSTTTTCGGGCCTGKBTAYTACGCTGGGCTA

**AL151\_F4:** CTTGATCTGCGGCTTTTCGGGTCGTGVTGATGTACCGCTTCGAGGAGGAGCTGTTCTTGC

**AL152\_R1:** GGTCCGCGGCATGACTGAATCGGACACAAGCGGTGCGATGAGNTAGGGCTACGCCCTT

**AL153\_R2:**

GTGAAATGGCACCAVYGAGAAAYAGCGGTGTCGGGGATGATCTGAKTGCCGAAGATGG

**AL154\_R3\_M249L:**

ABCACGACCCGAAAGCCGCAGATCAAGTAGCCCAGCGTARTAVM**CAG**GCCGAAASC GTG

**AL155\_R4\_Q283R:**

**AGCG**AATCTTATAGTCTTGCAAGCTTCGCAAGAACAGCTCCTCCTCGAAGCGGTACATC

### **Insert amplification primers**

**BZ\_SCA\_insert\_amp\_for3:**

AAAACCATCGCCCTGATCATGAACAGTAGTGGATCCACCG

**BZ\_SCA\_insert\_amp\_rev:**

TGTGGGCACCAGCAGGGCAGAG**GCG**AATCTTATAGTCTTGCAAGC

### **Vector amplification primers**

**BZ\_SCA\_vector\_amp\_for**

GCAAGACTATAAGATT**CGC**TCTGCCCTGCTGGTGCCCA

**BZ\_SCA\_vector\_amp\_rev**

CGGTGGATCCACTACTGTTTCATGATCAGGGCGATGGTTTT

### **Quick Change primers**

**BZ\_forquickchange\_314**

**ACCGGCACG**GCGCCGCTCAGCAA

**BZ\_revquickchange\_314**

**CGTGCCGGT**GGCGATCTCGTGCAAGT

## **2.9n Agar-plate screening of SCA library with luciferin analogs**

Library DNA was transformed into BL21 *E. coli* (New England BioLabs) and plated on agar plates containing 1 mM kanamycin. Plates were incubated at 37 °C for 12 h. Luciferin analog (1 mM) in 100 mM phosphate buffer pH 7.6 was lightly sprayed on the colonies. Plates were subsequently imaged and light emitting colonies were selected for further analysis.

## **2.9o Secondary screen of candidate SCA mutants in lysate using autoinduction**

Colonies emitting detectable light from the agar-plate screen were used to inoculate new cultures. Each colony was added to an individual well of a 96 deep well plate (Greiner Bio-One) containing 500  $\mu$ L of autoinduction media, which comprised 6 g/L  $\text{Na}_2\text{HPO}_4$  (Sigma Aldrich), 3 g/L  $\text{KH}_2\text{PO}_4$  (Sigma Aldrich), 20 g/L tryptone (BD Biosciences), 5 g/L yeast extract (Difco), 5 g/L NaCl (Sigma Aldrich), 0.6% v/v sterile filtered glycerol (Sigma Aldrich), 0.05% w/v sterile filtered glucose (Sigma Aldrich), 0.2% w/v sterile filtered lactose (Fisher), and 20  $\mu$ g/mL kanamycin (Fisher). For each 96-well plate, one culture expressing native Fluc and one culture expressing mutant 24 were included for screening controls. Cultures were grown with shaking at 30  $^\circ\text{C}$  for 24 h. A portion of each culture (50  $\mu$ L) was reserved for plasmid recovery and sequencing analysis. The remainder of the culture was centrifuged at 3400 rpm for 10 min at 4  $^\circ\text{C}$ . Following centrifugation, the supernatant was removed and the cells were used in screens within 24 h. Pelleted cultures were resuspended in 600  $\mu$ L lysis buffer (500 mM NaCl, 50 mM Tris (pH 7.4), 5 mM  $\text{MgCl}_2$ , 0.5% v/v Tween20). Lysed bacterial suspensions (90  $\mu$ L/well) were added to a 96-well black plate (in duplicate). The lysed cultures were imaged using a stock solution of 10X luciferin and 10X ATP in PBS (10 mM luciferin, 10 mM ATP). The appropriate luciferin stock solution (10  $\mu$ L) was added to each well (final [luciferin] = 1 mM and [ATP] = 1 mM). Images were acquired over 5 min. Cultures exhibiting  $\geq$  2-fold improved activity over mutant 24 were selected for sequencing analysis. The 50  $\mu$ L aliquot from the mutants of interest on each plate were used to inoculate a new culture (3 mL LB broth, 20  $\mu$ g/mL kanamycin), grown overnight at 37  $^\circ\text{C}$  with shaking, and plasmid purified using a plasmid purification kit (Qiagen).

## **2.9p Reproducing candidate SCA mutants in lysate using IPTG induction**

*E. coli* cells expressing candidate luciferases from the autoinduction screen (section 2.9o), native Fluc, or mutant 24 were added to 1 mL of LB broth (containing 1 mM Kan) and grown at 37 °C for 16 h. A 35  $\mu$ L aliquot of this culture was added to 5 mL of LB broth (containing 1 mM Kan). All culture tubes were shaken and incubated at 37 °C for 2-5 h until OD<sub>600</sub> ~0.8. Rapidly growing cultures were stored at 4 °C until all cultures reached the desired OD<sub>600</sub>. IPTG (0.5 mM) was added to the remainder of the cultures to induce protein expression. The cultures were incubated with shaking for 2 h at 30 °C or for 12 h at 25 °C. After induction, the cultures were centrifuged for 10 min at 3400  $\times$ g (4 °C). The supernatant was removed, and the pellet was resuspended in 0.6 mL of Fluc lysis buffer (50 mM Tris•HCl, 500 mM NaCl, 0.5% (v/v) Tween 20, 5 mM MgCl<sub>2</sub>, pH = 7.4). Lysed bacterial suspensions (90  $\mu$ L/well) were added to a 96-well black plate (in duplicate). The lysed cultures were imaged using a stock solution of 10X luciferin and 10X ATP in PBS (10 mM luciferin, 10 mM ATP). The appropriate luciferin stock solution (10  $\mu$ L) was added to each well (final [luciferin] = 1 mM and [ATP] = 1 mM). Images were acquired over 5 min.

### **2.9q Differential scanning fluorimetry (DSF)**

DSF was performed as described previously [57]. Purified luciferases (1 mg/mL) were incubated with 0.2% SYPRO dye (Sigma Aldrich) in storage buffer (25 mM Tris-HCl (pH 7.8) containing 1 mM EDTA and 0.2 mM ammonium sulfate) at 4 °C. Samples were added in triplicate to a 96-well PCR plate. The temperature was incrementally increased starting from 25 °C and ending at 95 °C at an interval of 0.5 °C. Fluorescence of the

SYPRO dye was measured at each temperature interval using a qPCR instrument (Agilent Technologies, Stratagene Mx3005P).

## **2.9r General synthetic methods**

All reagents purchased from commercial suppliers were of analytical grade and used without further purification. Anhydrous solvents were dried by passage over neutral alumina. Reaction vessels were either flame or oven dried prior to use. Appel's salt [48] **2.5** (4,5-dichloro-1,2,3-dithiazolium chloride), and 5-amino-6-methoxypyridine [63] **2.6** were prepared according to literature procedure. Reaction progress was monitored by thin-layer chromatography on EMD 60 F254 plates and visualized with UV light. Compounds were purified via flash column chromatography using Sorbent Technologies 60 Å, 230-400 mesh silica gel, unless otherwise stated. HPLC purifications were performed on a Varian ProStar equipped with 325 Dual Wavelength UV-Vis Detector. Semi-preparative runs were performed using an Agilent Prep-C18 Scalar column (9.4 x 150 mm, 5 µm) with a 4.2 mL/min flow rate, eluting with a gradient of 5-95% MeOH in ammonium acetate buffer (25 mM, pH 8). NMR spectra were acquired with Bruker Advanced spectrometers at 298 K. <sup>1</sup>H-NMR spectra were acquired at either 500 or 400 MHz, and <sup>13</sup>C-NMR spectra were acquired at 126 or 101 MHz. Chemical shifts are reported in ppm relative to residual non-deuterated NMR solvent, and coupling constants (*J*) are provided in Hz. Low and high-resolution electrospray ionization (ESI) mass spectra were collected at the University of California-Irvine Mass Spectrometry Facility. Infrared spectra were collected using a Nicolet iS5 FT-IR Spectrometer (Thermo Scientific). The abbreviations used can be found in the

## 2.9s Synthetic procedures

***N*-(6-Methoxy-pyridin-3-yl)cyanothioformamide (2.8):** Appel's salt (**2.5**, 31.3 g, 150 mmol) was added to a suspension of amino pyridine **2.6** (12.4 g, 100 mmol) in THF (300 mL). The mixture was stirred under N<sub>2</sub>. The mixture turned grey/green within 10 min. After 1.5 h, pyridine (16.6 mL, 205 mmol) was added. The mixture immediately turned dark red, and a yellow precipitate formed. After 19 h, a solution of sodium thiosulfate (31.6 g, 200 mmol) in 150 mL of H<sub>2</sub>O was added. The reaction was stirred for an additional 4.5 h. The mixture was then filtered, and the organic phase was removed *in vacuo*. The remaining aqueous solution was extracted with EtOAc (3 x 100 mL), and the combined organic fractions were washed with water (3 x 50 mL), and dried with MgSO<sub>4</sub>. The organics were then filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography (eluting with a gradient of 10:1 hexanes:EtOAc to 4:1 hexanes:EtOAc), affording **2.8** as a light orange solid (13.7 g, 71%). <sup>1</sup>H NMR (400 MHz, acetone-*d*<sub>6</sub>, mixture of tautomers) δ 8.66 (d, *J* = 4.0, 0.8H), 8.41(d, *J* = 4.0, 0.2H), 8.23 (dd, *J* = 8.0, 4.0, 0.8H), 7.92 (dd, *J* = 8.0, 4.0, 0.2H), 6.94 (d, *J* = 8.0, 0.2H), 6.90 (d, *J* = 8.0, 0.8H), 3.95 (s, 0.6H), 3.93 (s, 2.4H). <sup>13</sup>C NMR (101 MHz, DMSO-*d*<sub>6</sub>) (mixture of tautomers, not all tautomeric carbon signals were distinct) δ 162.0, 161.7, 141.7, 134.4, 128.9, 113.8, 110.6, 53.6. HRMS (ESI<sup>+</sup>) calcd for C<sub>8</sub>H<sub>8</sub>N<sub>3</sub>OS [M + H]<sup>+</sup> 194.0338, found 194.0394.

**2-Cyano-6-methoxy-thiazolo[4,5-c]pyridine (2.9a) and 2-Cyano-5-methoxy-thiazolo[5,4-b]pyridine (2.9b):** Palladium chloride (0.354 g, 2.00 mmol), CuI (0.956 mg, 0.502 mmol), TBAB (6.45 g, 20.0 mmol), and **2.8** (1.94 g, 10.0 mmol) were suspended in anhydrous 1:1 DMF:DMSO (160 mL). The resultant red-brown mixture was placed under N<sub>2</sub> and stirred at 120 °C for 1 h. The reaction was then diluted with H<sub>2</sub>O (100 mL) and extracted with EtOAc (7 x 100 mL). The organic layer was dried over MgSO<sub>4</sub>, filtered, and concentrated *in vacuo*. The crude product was purified by flash column chromatography (eluting with 4:1 hexanes:EtOAc). Fractions containing **2.9b** were combined and concentrated, affording **2.9b** (0.17 g, 8.6%) as a white solid. Fractions containing regioisomer **2.9a** were combined and concentrated *in vacuo*. The resulting solid was triturated with cold hexanes, and the precipitate was collected to afford regioisomer **2.9a** (0.85 g, 44%). **2.9a:** <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 9.21 (s, 1H), 7.73 (s, 1H), 3.98 (s, 3H). <sup>13</sup>C NMR (126 MHz, acetone-*d*<sub>6</sub>) δ 163.91, 147.54, 145.78, 145.42, 136.76, 113.54, 102.58, 54.99. HRMS (ESI<sup>+</sup>) calcd for C<sub>8</sub>H<sub>6</sub>N<sub>3</sub>OS [M + H]<sup>+</sup> 192.0232, found 192.0224. **2.9b:** <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 8.55 (d, *J* = 9.0, 1H), 7.22 (d, *J* = 9.0, 1H), 4.01 (s, 3H). <sup>13</sup>C NMR (101 MHz, DMSO-*d*<sub>6</sub>) δ 164.3, 155.4, 140.7, 135.0, 132.4, 113.3, 113.1, 54.6. HRMS (ESI<sup>+</sup>) calcd for C<sub>8</sub>H<sub>6</sub>N<sub>3</sub>OS [M + H]<sup>+</sup> 192.0232, found 192.0233.

**2-Cyano-6-acetoxy-thiazolo[4,5-c]pyridine (2.10):** Pyridine hydrochloride (1.8 g, 16 mmol) and **2.9a** (300 mg, 1.6 mmol) were combined in a septum-sealed flask and stirred at 160 °C for 15 min. The resulting red-brown liquid was left to cool to room temperature. Acetic anhydride (1.5 mL, 16 mmol), pyridine (2.6 mL, 32 mmol) and



DMAP (0.037 g, 0.32 mmol) were added to the flask and stirred for 1 h. The beige-colored heterogeneous mixture was then extracted with EtOAc (20 mL). The organics were washed with sat. NaHSO<sub>4</sub> (3 x 5 mL), dried with MgSO<sub>4</sub>, filtered, and then concentrated *in vacuo*. The crude mixture was purified by flash column chromatography (3:1 hexanes:EtOAc to 1:1 hexanes:EtOAc) affording **2.10** as a white solid (270 mg, 77%). <sup>1</sup>H NMR (500 MHz, acetone-*d*<sub>6</sub>) δ 9.31 (s, 1H), 8.10 (s, 1H), 2.36 (s, 3H). <sup>13</sup>C NMR (126 MHz, acetone-*d*<sub>6</sub>) δ 169.4, 157.0, 148.6, 147.2, 146.4, 140.3, 113.3, 110.6, 21.0. HRMS (ESI<sup>-</sup>) calcd for C<sub>7</sub>H<sub>2</sub>N<sub>3</sub>SO [M – CH<sub>3</sub>CO]<sup>-</sup> 175.9919, found 175.9927. IR (dry film)  $\nu_{\max}$  (cm<sup>-1</sup>): 3106 (CH, aromatic), 2946 (CH, aromatic), 2236 (CN, nitrile), 1744 (CO, ester).

**2-Cyano-5-acetoxy-thiazolo[5,4-*b*]pyridine (2.11):** Pyridine hydrochloride (1.3 g, 1.2 mmol) and **2.9b** (190 mg, 1.0 mmol) were combined in a sealed tube and stirred at 160 °C for 30 min. The resulting red-brown liquid was left to cool. Acetic anhydride (1.0 mL, 10 mmol), pyridine (1.6 mL, 20 mmol) and DMAP (0.026 g, 0.20 mmol) were added to the flask and stirred for 1 h. The beige-colored heterogeneous mixture was extracted with EtOAc (20 mL). The organics were washed with sat. NaHSO<sub>4</sub> (3 x 5 mL), dried with MgSO<sub>4</sub>, filtered, then concentrated *in vacuo*. The crude mixture was purified by flash column chromatography (3:1 hexanes:EtOAc to 1:1 hexanes:EtOAc) affording **2.11** as a white solid (0.16 g, 73%) <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.55 (d, *J* = 8.0 Hz, 1H), 7.40 (d, *J* = 8.0, 1H), 2.41 (s, 3H). <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>) δ 168.6, 158.1, 155.9, 144.4, 137.6, 135.7, 117.1, 112.6, 21.4. HRMS (ESI<sup>-</sup>) calcd for C<sub>7</sub>H<sub>2</sub>N<sub>3</sub>SO [M – CH<sub>3</sub>CO]<sup>-</sup>

175.9919, found 175.9916. IR (dry film)  $\nu_{\max}$  (cm<sup>-1</sup>): 3073 (CH, aromatic), 2233 (CN, nitrile), 1758 (CO, ester).

**(S)-2-(Thiazolo[4,5-c]pyridin-6(5H)-one-2-yl)-4,5-dihydrothiazole-4-carboxylic acid**

**(2.3):** D-Cysteine hydrochloride monohydrate (0.211 g, 1.20 mmol) and K<sub>2</sub>CO<sub>3</sub> (0.166 g, 1.20 mmol) were dissolved in H<sub>2</sub>O (1.6 mL). The mixture was added to a suspension of **2.10** (0.142 g, 0.648 mmol) in MeCN (8.8 mL) and stirred. Yellow precipitate began to form within minutes. After 1.5 h, TLC (1:1 hexanes:EtOAc) revealed complete consumption of **2.10**, and the reaction mixture was then filtered. The precipitate was purified by HPLC as described in the general synthetic methods. The desired fractions were combined and concentrated to provide **2.3** as a yellow solid (0.11 g, 62%). <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  8.35 (d, *J* = 0.8, 1H), 7.03 (d, *J* = 0.8, 1H), 5.31 (dd, *J* = 10.0, 8.0, 1H), 3.88 (dd, *J* = 11.2, 10.0, 1H), 3.69 (dd, *J* = 11.2, 8.1, 1H). <sup>13</sup>C NMR (126 MHz, D<sub>2</sub>O)  $\delta$  179.8, 167.6, 164.3, 164.2, 155.9, 141.8, 134.5, 110.9, 83.0, 39.4. HRMS (ESI<sup>-</sup>) calcd for C<sub>9</sub>H<sub>6</sub>N<sub>3</sub>OS<sub>2</sub> [M - COOH]<sup>-</sup> 235.9952, found 235.9949.

**(S)-2-(Thiazolo[5,4-b]pyridin-5(4H)-one-2-yl)-4,5-dihydrothiazole-4-carboxylic acid**

**(2.4):** D-Cysteine hydrochloride monohydrate (0.462 g, 2.63 mmol) and K<sub>2</sub>CO<sub>3</sub> (0.363 g, 2.63 mmol) were dissolved in H<sub>2</sub>O (2.6 mL). The mixture was added to a suspension of **2.11** (0.291 g, 1.33 mmol) in MeCN (5.3 mL) and stirred. Yellow precipitate began to form within minutes. After 10 min, TLC (3:2 hexanes:EtOAc) revealed complete consumption of **2.11** and the reaction mixture was filtered. The precipitate was purified by HPLC as described in the general synthetic methods. The desired fractions were

combined and concentrated to provide **2.4** as a yellow solid (0.235 g, 53%).  $^1\text{H}$  NMR (500 MHz,  $\text{D}_2\text{O}$ )  $\delta$  7.99 (d,  $J = 9.2$ , 1H), 6.63 (d,  $J = 9.2$ , 1H), 5.24 (dd,  $J = 9.8$ , 7.9, 1H), 3.84 (dd,  $J = 11.2$ , 9.8, 1H), 3.63 (dd,  $J = 11.2$ , 7.9, 1H).  $^{13}\text{C}$  NMR (126 MHz,  $\text{D}_2\text{O}$ )  $\delta$  180.6, 174.0, 168.9, 159.6, 154.4, 139.9, 135.6, 118.6, 82.5, 38.9. HRMS (ESI $^-$ ) calcd for  $\text{C}_{10}\text{H}_6\text{N}_3\text{O}_3\text{S}_2$  [ $\text{M} - \text{H}$ ] $^-$  279.9850, found 279.9846.

**(S)-2-(6-Methoxy-thiazolo[4,5-c]pyridin-2-yl)-4,5-dihydrothiazole-4-carboxylic acid (2.12):** D-Cysteine hydrochloride monohydrate (34.7 mg, 0.219 mmol) and **2.9a** (39.9 mg, 0.209 mmol) were suspended in 30:1 MeCN:DMF (1.8 mL) in a 20 mL vial. A solution of  $\text{K}_2\text{CO}_3$  (29.2 mg, 0.211 mmol in 0.42 mL  $\text{H}_2\text{O}$ ) was added, and the mixture was stirred for 3 h. The reaction mixture was then filtered, providing **2.12** as a white solid (50 mg, 75%).  $^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  8.80 (d,  $J = 0.8$ , 1H), 7.46 (d,  $J = 0.8$ , 1H), 5.32 (dd,  $J = 10.0$ , 8.4, 1H), 4.01 (s, 3H), 3.91 (dd,  $J = 11.2$ , 10.0, 1H), 3.71 (dd,  $J = 11.6$ , 8.4, 1H).  $^{13}\text{C}$  NMR (126 MHz,  $\text{D}_2\text{O}$ )  $\delta$  177.3, 165.2, 161.7, 161.3, 147.5, 144.8, 142.3, 100.6, 80.3, 55.2, 36.7. HRMS (ESI $^-$ ) calcd for  $\text{C}_{11}\text{H}_9\text{N}_3\text{O}_3\text{S}_2$  [ $\text{M} - \text{COOH}$ ] $^-$  250.0109, found 250.0119.

**(S)-2-(6-Methoxy-thiazolo[5,4-b]pyridin-2-yl)-4,5-dihydrothiazole-4-carboxylic acid (2.13):** D-Cysteine hydrochloride monohydrate (0.0181 g, 0.103 mmol) and  $\text{K}_2\text{CO}_3$  (0.0145 g, 0.105 mmol) were dissolved in  $\text{H}_2\text{O}$  (0.4 mL). The mixture was added to a suspension of **2.9b** (0.0198 g, 0.104 mmol) in MeCN (0.5 mL) and DMF (0.03 mL). The resulting mixture was stirred. After 10 min, TLC revealed complete consumption of **2.9b**, and the reaction mixture was filtered. The filtrate was collected and concentrated *in*

*vacuo*. The crude mixture was washed with MeOH (3 x 5 mL), affording **2.13** as a white solid (29 mg, 83%). <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ 8.11 (d, *J* = 12.0, 1H), 7.00 (d, *J* = 8.0, 1H), 5.29 (t, *J* = 12.0, 1H), 4.00 (s, 3H), 3.88 (t, *J* = 12.0, 1H), 3.68 (t, *J* = 8.0, 1H). <sup>13</sup>C NMR (126 MHz, D<sub>2</sub>O) δ 177.4, 165.7, 163.5, 157.6, 154.5, 141.1, 133.9, 112.0, 80.2, 54.7, 36.4. HRMS (ESI<sup>-</sup>) calcd for C<sub>11</sub>H<sub>9</sub>N<sub>3</sub>O<sub>3</sub>S<sub>2</sub> [M - COOH]<sup>-</sup> 250.0109, found 250.0115.

## References

1. Massoud, T. F.; Gambhir, S. S. Molecular imaging in living subjects: seeing fundamental biological processes in a new light. *Genes Dev.* **2003**, *17*, 545-580.
2. Paley, M. A.; Prescher, J. A. Bioluminescence: a versatile technique for imaging cellular and molecular features. *MedChemComm* **2014**, *5*, 255-267.
3. Branchini, B. R.; Behney, C. E.; Southworth, T. L.; Fontaine, D. M.; Gulick, A. M.; Vinyard, D. J.; Brudvig, G. W. Experimental Support for a Single Electron-Transfer Oxidation Mechanism in Firefly Bioluminescence. *J. Am. Chem. Soc.* **2015**, *137*, 7592-7595.
4. Fraga, H.; Fernandes, D.; Novotny, J.; Fontes, R.; da Silva, J. C. G. Firefly luciferase produces hydrogen peroxide as a coproduct in dehydroluciferyl adenylate formation. *ChemBioChem* **2006**, *7*, 929-935.
5. Kaskova, Z. M.; Tsarkova, A. S.; Yampolsky, I. V. 1001 lights: Luciferins, luciferases, their mechanisms of action and applications in chemical analysis, biology and medicine. *Chem. Soc. Rev.* **2016**, *45*, 6048-6077.
6. Contag, C. H.; Spilman, S. D.; Contag, P. R.; Oshiro, M.; Eames, B.; Dennery, P.; Stevenson, D. K.; Benaron, D. A. Visualizing gene expression in living mammals using a bioluminescent reporter. *Photochem. Photobiol.* **1997**, *66*, 523-531.
7. Bhaumik, S.; Gambhir, S. S. Optical imaging of *Renilla* luciferase reporter gene expression in living mice. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 377-382.
8. Tannous, B. A.; Kim, D. E.; Fernandez, J. L.; Weissleder, R.; Breakefield, X. O. Codon-optimized *Gaussia* luciferase cDNA for mammalian gene expression in culture and *in vivo*. *Mol. Ther.* **2005**, *11*, 435-443.

9. Hall, M. P.; Unch, J.; Binkowski, B. F.; Valley, M. P.; Butler, B. L.; Wood, M. G.; Otto, P.; Zimmerman, K.; Vidugiris, G.; Machleidt, T.; Robers, M. B.; Benink, H. A.; Eggers, C. T.; Slater, M. R.; Meisenheimer, P. L.; Klaubert, D. H.; Fan, F.; Encell, L. P.; Wood, K. V. Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. *ACS Chem. Biol.* **2012**, *7*, 1848-1857.
10. Chu, J.; Oh, Y.; Sens, A.; Ataie, N.; Dana, H.; Macklin, J. J.; Laviv, T.; Welf, E. S.; Dean, K. M.; Zhang, F. J.; Kim, B. B.; Tang, C. T.; Hu, M.; Baird, M. A.; Davidson, M. W.; Kay, M. A.; Fiolka, R.; Yasuda, R.; Kim, D. S.; Ng, H. L.; Lin, M. Z. A bright cyan-excitable orange fluorescent protein facilitates dual-emission microscopy and enhances bioluminescence imaging *in vivo*. *Nature Biotechnology* **2016**, *34*, 760.
11. Zhao, H.; Doyle, T. C.; Coquoz, O.; Kalish, F.; Rice, B. W.; Contag, C. H. Emission spectra of bioluminescent reporters and interaction with mammalian tissue determine the sensitivity of detection *in vivo*. *J. Biomed. Opt.* **2005**, *10*, 41210.
12. Liu, H.; Patel, M. R.; Prescher, J. A.; Patsialou, A.; Qian, D.; Lin, J.; Wen, S.; Chang, Y. F.; Bachmann, M. H.; Shimono, Y.; Dalerba, P.; Adorno, M.; Lobo, N.; Bueno, J.; Dirbas, F. M.; Goswami, S.; Somlo, G.; Condeelis, J.; Contag, C. H.; Gambhir, S. S.; Clarke, M. F. Cancer stem cells from human breast tumors are involved in spontaneous metastases in orthotopic mouse models. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 18115-18120.
13. Sweeney, T. J.; Mailander, V.; Tucker, A. A.; Olomu, A. B.; Zhang, W.; Cao, Y.; Negrin, R. S.; Contag, C. H. Visualizing the kinetics of tumor-cell clearance in living animals. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 12044-12049.
14. Sanada, Y.; Yamamoto, T.; Satake, R.; Yamashita, A.; Kanai, S.; Kato, N.; van de Loo, F. A. J.; Nishimura, F.; Scherer, P. E.; Yanaka, N. Serum Amyloid A3 Gene Expression in Adipocytes is an Indicator of the Interaction with Macrophages. *Sci. Rep.* **2016**, *6*, 38697.
15. Thorne, S. H.; Contag, C. H. Using *in vivo* bioluminescence imaging to shed light on cancer biology. *Proc. IEEE* **2005**, *93*, 750-762.
16. Fan, F.; Binkowski, B. F.; Butler, B. L.; Stecha, P. F.; Lewis, M. K.; Wood, K. V. Novel genetically encoded biosensors using firefly luciferase. *ACS Chem. Biol.* **2008**, *3*, 346-351.
17. Luker, K. E.; Smith, M. C.; Luker, G. D.; Gammon, S. T.; Piwnica-Worms, H.; Piwnica-Worms, D. Kinetics of regulated protein-protein interactions revealed with firefly luciferase complementation imaging in cells and living animals. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 12288-12293.

18. Sellmyer, M. A.; Bronsart, L.; Imoto, H.; Contag, C. H.; Wandless, T. J.; Prescher, J. A. Visualizing cellular interactions with a generalized proximity reporter. *P. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 8567-8572.
19. Evans, M. S.; Chaurette, J. P.; Adams Jr, S. T.; Reddy, G. R.; Paley, M. A.; Aronin, N.; Prescher, J. A.; Miller, S. C. A synthetic luciferin improves bioluminescence imaging in live mice. *Nat. Meth.* **2014**, *11*, 393-395.
20. Keyaerts, M.; Verschueren, J.; Bos, T. J.; Tchouate-Gaikam, L. O.; Peleman, C.; Breckpot, K.; Vanhove, C.; Caveliers, V.; Bossuyt, A.; Lahoutte, T. Dynamic bioluminescence imaging for quantitative tumour burden assessment using IV or IP administration of D-luciferin: effect on intensity, time kinetics and repeatability of photon emission. *Eur. J. Nucl. Med. Mol. Imaging* **2008**, *35*, 999-1007.
21. Rathbun, C. M.; Prescher, J. A. Bioluminescent Probes for Imaging Biology beyond the Culture Dish. *Biochemistry* **2017**, *56*, 5178-5184.
22. Branchini, B. R.; Ablamsky, D. M.; Rosenman, J. M.; Uzasci, L.; Southworth, T. L.; Zimmer, M. Synergistic mutations produce blue-shifted bioluminescence in firefly luciferase. *Biochemistry* **2007**, *46*, 13847-13855.
23. Branchini, B. R.; Ablamsky, D. M.; Murtiashaw, M. H.; Uzasci, L.; Fraga, H.; Southworth, T. L. Thermostable red and green light-producing firefly luciferase mutants for bioluminescent reporter applications. *Anal. Biochem.* **2007**, *361*, 253-262.
24. Li, X.; Nakajima, Y.; Niwa, K.; Viviani, V. R.; Ohmiya, Y. Enhanced red-emitting railroad worm luciferase for bioassays and bioimaging. *Protein Sci.* **2010**, *19*, 26-33.
25. McCutcheon, D. C.; Paley, M. A.; Steinhardt, R. C.; Prescher, J. A. Expedient Synthesis of Electronically Modified Luciferins for Bioluminescence Imaging. *J. Am. Chem. Soc.* **2012**, *134*, 7604-7607.
26. Steinhardt, R. C.; O'Neill, J. M.; Rathbun, C. M.; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. Design and Synthesis of an Alkynyl Luciferin Analogue for Bioluminescence Imaging. *Chem. Eur. J.* **2016**, *22*, 3671-3675.
27. Steinhardt, R. C.; Rathbun, C. M.; Krull, B. T.; Yu, J. M.; Yang, Y.; Nguyen, B. D.; Kwon, J.; McCutcheon, D. C.; Jones, K. A.; Furche, F.; Prescher, J. A. Brominated Luciferins Are Versatile Bioluminescent Probes. *ChemBioChem* **2017**, *18*, 96-100.
28. Reddy, G. R.; Thompson, W. C.; Miller, S. C. Robust light emission from cyclic alkylaminoluciferin substrates for firefly luciferase. *J. Am. Chem. Soc.* **2010**, *132*, 13586-13587.

29. Mofford, D. M.; Reddy, G. R.; Miller, S. C. Aminoluciferins extend firefly luciferase bioluminescence into the near-infrared and can be preferred substrates over D-luciferin. *J. Am. Chem. Soc.* **2014**, *136*, 13277-13282.
30. Woodroffe, C. C.; Meisenheimer, P. L.; Klaubert, D. H.; Kovic, Y.; Rosenberg, J. C.; Behney, C. E.; Southworth, T. L.; Branchini, B. R. Novel heterocyclic analogues of firefly luciferin. *Biochemistry* **2012**, *51*, 9807-9813.
31. Conley, N. R.; Dragulescu-Andrasi, A.; Rao, J. H.; Moerner, W. E. A Selenium Analogue of Firefly D-Luciferin with Red-Shifted Bioluminescence Emission. *Angew. Chem. Int. Ed.* **2012**, *51*, 3350-3353.
32. Takakura, H.; Sasakura, K.; Ueno, T.; Urano, Y.; Terai, T.; Hanaoka, K.; Tsuboi, T.; Nagano, T. Development of luciferin analogues bearing an amino group and their application as BRET donors. *Chem. Asian J.* **2010**, *5*, 2053-2061.
33. Kojima, R.; Takakura, H.; Ozawa, T.; Tada, Y.; Nagano, T.; Urano, Y. Rational Design and Development of Near-Infrared-Emitting Firefly Luciferins Available *In Vivo*. *Angew. Chem. Int. Ed.* **2013**, *52*, 1175-1179.
34. Iwano, S.; Obata, R.; Miura, C.; Kiyama, M.; Hama, K.; Nakamura, M.; Amano, Y.; Kojima, S.; Hirano, T.; Maki, S.; Niwa, H. Development of simple firefly luciferin analogs emitting blue, green, red, and near-infrared biological window light. *Tetrahedron* **2013**, *69*, 3847-3856.
35. Miura, C.; Kiyama, M.; Iwano, S.; Ito, K.; Obata, R.; Hirano, T.; Maki, S.; Niwa, H. Synthesis and luminescence properties of biphenyl-type firefly luciferin analogs with a new, near-infrared light-emitting bioluminophore. *Tetrahedron* **2013**, *69*, 9726-9734.
36. Jathoul, A. P.; Grounds, H.; Anderson, J. C.; Pule, M. A. A Dual-Color Far-Red to Near-Infrared Firefly Luciferin Analogue Designed for Multiparametric Bioluminescence Imaging. *Angew. Chem. Int. Ed.* **2014**, *53*, 13059-13063.
37. Anderson, J. C.; Grounds, H.; Jathoul, A. P.; Murray, J. A. H.; Pacman, S. J.; Tisi, L. Convergent synthesis and optical properties of near-infrared emitting bioluminescent infra-luciferins. *Rsc Adv.* **2017**, *7*, 3975-3982.
38. Bisai, V.; Sarpong, R. Methoxypyridines in the Synthesis of Lycopodium Alkaloids: Total Synthesis of (+/-)-Lycoposerramine R. *Org. Lett.* **2010**, *12*, 2551-2553.
39. de Candia, M.; Fossa, P.; Cellamare, A.; Mosti, L.; Carotti, A.; Altomare, C. Insights into structure-activity relationships from lipophilicity profiles of pyridin-2(1H)-one analogs of the cardiotoxic agent milrinone. *Eur. J. Pharm. Sci.* **2005**, *26*, 78-86.
40. Dragovich, P. S.; Prins, T. J.; Zhou, R.; Brown, E. L.; Maldonado, F. C.; Fuhrman, S. A.; Zalman, L. S.; Tuntland, T.; Lee, C. A.; Patick, A. K.; Matthews, D. A.;

Hendrickson, T. F.; Kosa, M. B.; Liu, B.; Batugo, M. R.; Gleeson, J. P. R.; Sakata, S. K.; Chen, L. J.; Guzman, M. C.; Meador, J. W.; Ferre, R. A.; Worland, S. T. Structure-based design, synthesis, and biological evaluation of irreversible human rhinovirus 3C protease inhibitors. 6. Structure-activity studies of orally bioavailable, 2-pyridone-containing peptidomimetics. *J. Med. Chem.* **2002**, *45*, 1607-1623.

41. Erez, Y.; Huppert, D. Excited-state intermolecular proton transfer of the firefly's chromophore D-luciferin. *J. Phys. Chem. A* **2010**, *114*, 8075-8082.

42. Erez, Y.; Presiado, I.; Gepshtein, R.; Pinto da Silva, L.; Esteves da Silva, J. C.; Huppert, D. Comparative study of the photoprotolytic reactions of D-luciferin and oxyluciferin. *J. Phys. Chem. A* **2012**, *116*, 7452-7461.

43. Rini, M.; Magnes, B. Z.; Pines, E.; Nibbering, E. T. Real-time observation of bimodal proton transfer in acid-base pairs in water. *Science* **2003**, *301*, 349-352.

44. Mcelroy, W. D.; Seliger, H. H.; White, E. H. Mechanism of Bioluminescence, Chemiluminescence and Enzyme Function in Oxidation of Firefly Luciferin. *Photochem. Photobiol.* **1969**, *10*, 153-170.

45. Mezzanotte, L.; Que, I.; Kaijzel, E.; Branchini, B.; Roda, A.; Lowik, C. Sensitive Dual Color *In Vivo* Bioluminescence Imaging Using a New Red Codon Optimized Firefly Luciferase and a Green Click Beetle Luciferase. *PLoS One* **2011**, *6*, e19277.

46. Krishnamoorthy, A.; Robertson, J. B. Dual-Color Monitoring Overcomes the Limitations of Single Bioluminescent Reporters in Fast-Growing Microbes and Reveals Phase-Dependent Protein Productivity during the Metabolic Rhythms of *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **2015**, *81*, 6484-6495.

47. McCutcheon, D. C.; Porterfield, W. B.; Prescher, J. A. Rapid and scalable assembly of firefly luciferase substrates. *Org. Biomol. Chem.* **2015**, *13*, 2117-2121.

48. Michaelidou, S. S.; Koutentis, P. A. The Synthesis of 2-Cyano-cyanothioformanilides from 2-(4-Chloro-5H-1,2,3-dithiazol-5-ylideneamino)benzonitriles Using DBU. *Synthesis* **2009**, 4167-4174.

49. Inamoto, K.; Hasegawa, C.; Hiroya, K.; Doi, T. Palladium-Catalyzed Synthesis of 2-Substituted Benzothiazoles via a C-H Functionalization/Intramolecular C-S Bond Formation Process. *Org. Lett.* **2008**, *10*, 5147-5150.

50. Jones, K. A.; Porterfield, W. B.; Rathbun, C. M.; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. Orthogonal Luciferase–Luciferin Pairs for Bioluminescence Imaging. *J. Am. Chem. Soc.* **2017**, *139*, 2351-2358.



51. Branchini, B. R.; Murtiashaw, M. H.; Magyar, R. A.; Portier, N. C.; Ruggiero, M. C.; Stroh, J. G. Yellow-green and red firefly bioluminescence from 5,5-dimethyloxyluciferin. *J. Am. Chem. Soc.* **2002**, *124*, 2112-2113.
52. Forlani, L.; Cristoni, G.; Boga, C.; Todesco, P. E.; Del Vecchio, E.; Selva, S.; Monari, M. Reinvestigation of the tautomerism of some substituted 2-hydroxypyridines. *ARKIVOC* **2002**, 198-215.
53. Gammon, S. T.; Leevy, W. M.; Gross, S.; Gokel, G. W.; Piwnica-Worms, D. Spectral unmixing of multicolored bioluminescence emitted from heterogeneous biological sources. *Anal. Chem.* **2006**, *78*, 1520-1527.
54. Branchini, B. R.; Southworth, T. L.; Murtiashaw, M. H.; Boije, H.; Fleet, S. E. A mutagenesis study of the putative luciferin binding site residues of firefly luciferase. *Biochemistry* **2003**, *42*, 10429-10436.
55. Viviani, V. R.; Neves, D. R.; Amaral, D. T.; Prado, R. A.; Matsushashi, T.; Hirano, T. Bioluminescence of beetle luciferases with 6'-amino-D-luciferin analogues reveals excited keto-oxyluciferin as the emitter and phenolate/luciferin binding site interactions modulate bioluminescence colors. *Biochemistry* **2014**, *53*, 5208-5220.
56. Liu, M. D.; Warner, E. A.; Morrissey, C. E.; Fick, C. W.; Wu, T. S.; Ornelas, M. Y.; Ochoa, G. V.; Zhang, B. S.; Rathbun, C. M.; Porterfield, W. B.; Prescher, J. A.; Leconte, A. M. Statistical Coupling Analysis-Guided Library Design for the Discovery of Mutant Luciferases. *Biochemistry* **2017**, *57*, 663-671.
57. Rosa, N.; Ristic, M.; Seabrook, S. A.; Lovell, D.; Lucent, D.; Newman, J. Meltdown: A Tool to Help in the Interpretation of Thermal Melt Curves Acquired by Differential Scanning Fluorimetry. *J. Biomol. Screen* **2015**, *20*, 898-905.
58. Viviani, V. R.; Arnoldi, F. G. C.; Silva-Neto, A. J.; Barbosa, J. A.; Ohmiya, Y. The structural origin of pH-sensitivity in beetle luciferases: Influence of the loop between residues 223-235. *Luminescence* **2006**, *21*, 295-295.
59. Branchini, B. R.; Southworth, T. L.; Fontaine, D. M.; Murtiashaw, M. H.; McGurk, A.; Talukder, M. H.; Qureshi, R.; Yetil, D.; Sundlov, J. A.; Gulick, A. M. Cloning of the Orange Light-Producing Luciferase from *Photinus scintillans*-A New Proposal on how Bioluminescence Color is Determined. *Photochem. Photobiol.* **2017**, *93*, 479-485.
60. Viviani, V. R.; Simoes, A.; Bevilaqua, V. R.; Gabriel, G. V. M.; Arnoldi, F. G. C.; Hirano, T. Glu311 and Arg337 Stabilize a Closed Active-site Conformation and Provide a Critical Catalytic Base and Countercation for Green Bioluminescence in Beetle Luciferases. *Biochemistry* **2016**, *55*, 4764-4776.

61. Nakatsu, T.; Ichiyama, S.; Hiratake, J.; Saldanha, A.; Kobashi, N.; Sakata, K.; Kato, H. Structural basis for the spectral difference in luciferase bioluminescence. *Nature* **2006**, *440*, 372-376.
62. Jones, K. A.; Porterfield, W. B.; Rathbun, C. M.; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. Orthogonal Luciferase-Luciferin Pairs for Bioluminescence Imaging. *J. Am. Chem. Soc.* **2017**, *139*, 2351-2358.
63. Ahmad, Y.; Hey, D. H. Acylarylnitrosamines .7. Reactions of 3-Aminomethoxypyridines and 3-Amino-1-Methylpyridones - the Preparation of 2-Methoxy-3-Phenyl-Pyridine, 4-Methoxy-3-Phenyl-Pyridine, and 2-Methoxy-5-Phenylpyridine. *J. Chem. Soc.* **1954**, 4516-4523.

## Chapter 3: Pi-extended luciferins and mutant luciferases for multicomponent imaging<sup>a</sup>

<sup>a</sup>Zi Yao contributed to work in this chapter

### 3.1 Introduction

As demonstrated in chapter 2, spectrally resolved probes can be readily discriminated in transparent samples, enabling multicomponent imaging *in vitro* [1-2]. However, discriminating wavelengths through thick tissue is challenging. Tissues readily absorb wavelengths in the visible spectrum (< 600 nm), attenuating the signal from conventional bioluminescent probes. Historically, multicomponent bioluminescence imaging has been achieved using *substrate*-resolved luciferases instead of spectrally resolved pairs [3-4]. Substrate-resolved bioluminescence is well preceded in nature, as structurally distinct luciferase-luciferin pairs have been identified across diverse phyla [5-6]. Select orthogonal luciferase-luciferin pairs from nature, such as those from the firefly and the sea pansy *Renilla reniformis*, find routine use in multicomponent imaging assays [3,7-8]. These tools enable multicomponent imaging in heterogeneous environments, complementing conventional imaging tools such as fluorescent proteins. However, the development of orthogonal bioluminescent systems from nature as imaging tools has not kept pace with the diverse palette of fluorescent proteins [9].

To circumvent this bottleneck, we and others have turned to engineering orthogonal bioluminescent tools with luciferin analogs and mutant luciferases. Mutant luciferases have been shown to exhibit unique substrate selectivity. Work by the Miller group demonstrated that mutating residues critical for binding the native substrate (R218K, L286M, and S347A) oblates activity with D-luciferin, but retains activity with a luciferin analog [10]. Complementary studies in the Prescher lab have involved

screening large collections of mutants and substrates for orthogonal luciferase-luciferin pairs [4,11]. We designed libraries of luciferase mutants, also targeting residues important for substrate binding. Parallel screening of these libraries with luciferin analogs generated novel orthogonal luciferin/luciferase pairs. In total, we screened 159 mutant luciferases with 12 synthetic substrates, generating more than 800,000 possible pairs. Interestingly, the top orthogonal pairs collectively included seven unique luciferin substrates and 18 mutational sites. This result suggests there are a variety of paths to achieve orthogonality. Further exploration of luciferin structure and luciferase mutational sites could further expand the orthogonal bioluminescent toolset.

This chapter describes my efforts to engineer orthogonal luciferase-luciferin pairs. We were inspired by previous examples of luciferin structures with extended conjugation [12-16]. Luciferins with extended conjugation are generally red-shifted, which is desirable for *in vivo* imaging. We thought they could also be valuable as orthogonal tools. Niwa and co-workers showed that vinylogous luciferins exhibit different reactivity with luciferase based on the chromophore length [12]. Not surprisingly, the luciferin analogs closer in length to the native substrate (D-luciferin) exhibited the brightest emission with Fluc, though all analogs had some activity with the enzyme. This result suggests that the binding pocket could be optimized to complement elongated luciferin substrates. In a recent example, a luciferase derived from the click beetle was engineered to process extended naphthyl luciferins. The engineered enzyme-substrate was shown to be orthogonal to other probes [15].

Here, we developed pi-extended, rotameric luciferins (Figure 3-1) as orthogonal probes. These luciferins were poor substrates for the wild-type luciferase, making them

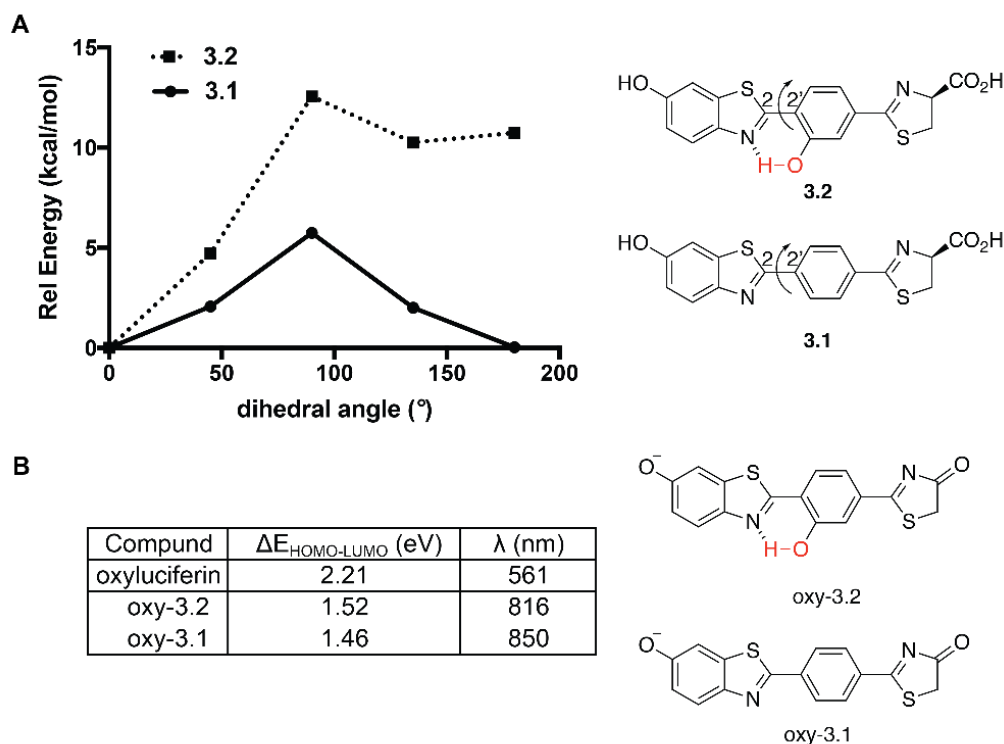
ideal candidates for orthogonal tool development. Initial attempts at mutating residues involved with substrate binding failed to produce a light-emitting luciferase. We then turned to computation to design a binding pocket to complement the luciferin analogs. Using a computationally guided approach, we generated mutant enzymes predicted to improve binding with the luciferin analogs. Subsequent screening of this library with the luciferin analogs produced a light emitting enzyme. Additional rounds of mutagenesis further improved the photon output of the luciferase lead and luciferin analog. Gratifyingly, this luciferase-luciferin pair was orthogonal to existing bioluminescent tools, enabling three- and four-component imaging. This result highlights the utility of structurally diverse luciferins and mutant luciferases for expanding orthogonal bioluminescent tools beyond pairs to higher order multiplets.

### **3.2 Design and synthesis of pi-extended analogs**

We sought to expand the diversity of orthogonal luciferins by designing pi-extended analogs **3.1** and **3.2** (Figure 3-1). Probes of this type generally produce tissue-penetrant, red-shifted luminescence [17]. Extended conjugation also lengthens structure, distinguishing them from less conjugated analogs [12]. Ideally, an orthogonal luciferin would be selectively processed by a mutant luciferase, while being minimally active with other luciferases. Specificity in enzyme/substrate interactions can be engineered by designing structurally modified analogs and subsequent screening for a complementary protein binding pocket.

Pi-extended luciferins **3.1** and **3.2** can adopt multiple rotameric states. Based on DFT calculations, the rotational barrier of the 2-2' C-C bond is 5-10 kcal/mol (Figure 3-

1A). Thus, we envisioned the luciferin could adopt multiple conformations in the enzyme binding pocket. A complementary protein binding pocket could be designed to prefer the fully planar chromophore, enabling optimal luminescence efficiency. We also designed planarized luciferin **3.2**, containing an intramolecular H-bond, which restricts bond rotation about the 2-2' C-C bond. In addition to being orthogonal luciferins, these pi-extended luciferins could produce red-shifted bioluminescence when fully planarized. Based on DFT calculations, luciferins **3.1** and **3.2** were predicted to emit infrared light (Figure 3-1B).



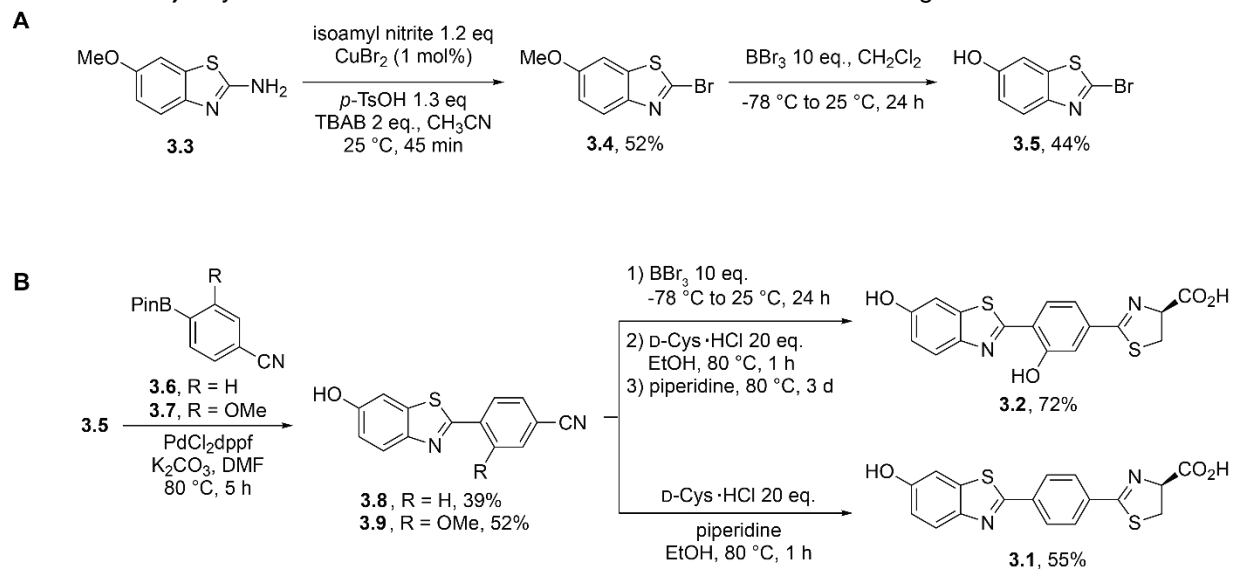
**Figure 3-1. Computational predictions of pi-extended analogs 3.1 and 3.2.** **A)** DFT calculations<sup>[a]</sup> were performed to determine torsion energy as a function of dihedral angle (N-C2-C2'-C) for analogs **3.1** and **3.2**. **B)** DFT calculations<sup>[a]</sup> were performed to determine  $\Delta E_{\text{HOMO-LUMO}}$  values for oxyluciferin chromophores of **3.1** and **3.2**. Wavelength of emission ( $\lambda$ ) values were obtained by converting  $\Delta E_{\text{HOMO-LUMO}}$  values in eV to wavelength in nm.

[a] Bond torsion energies were calculated using B3LYP 6-311G\* and HOMO-LUMO energies were calculated using B3LYP 6-311++G(2df,2pd).

Both luciferins could be efficiently synthesized from common intermediates (Scheme 3-1). The synthesis began with a bromination of commercially available aniline **3.3** via a Sandmeyer reaction. The reaction was mediated by substoichiometric copper (II) bromide [18], affording 2-bromo benzothiazole **3.4**. This intermediate was then demethylated by boron tribromide to afford **3.5**. Suzuki coupling of aryl-bromide **3.5** with boronic esters **3.6** and **3.7** afforded bi-aryl intermediates **3.8** and **3.9**. Condensation of **3.8** according to literature procedures [19] afforded luciferin analog **3.1**. Intermediate

**3.9** was first demethylated with boron tribromide, then condensed with D-cysteine to afford analog **3.2**.

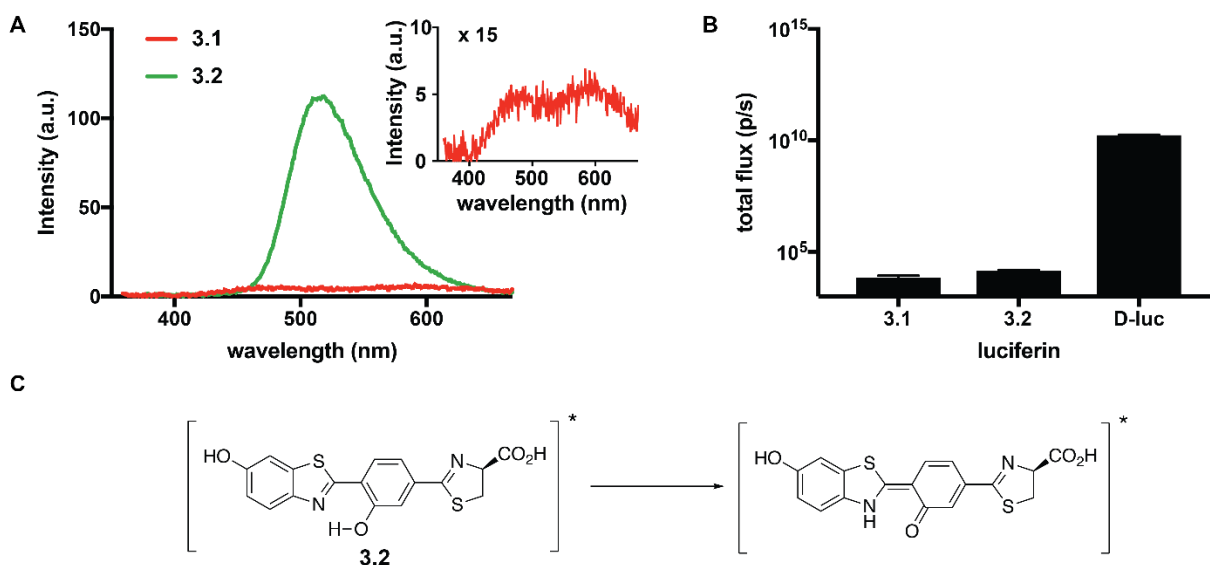
**Scheme 3-1. Synthesis of pi-extended luciferins.** **A)** Preparation of 2-bromo benzothiazole intermediate **B)** Key Suzuki reaction and condensation to afford luciferin analogs **3.1** and **3.2**.



With the analogs in hand, we examined the conformational flexibility of these chromophores. Conformational flexibility in a chromophore is known to hinder fluorescence efficiency [20]. Indeed, luciferin **3.1** exhibited weak fluorescence emission in water (Figure 3-2). Two emission bands were observed at 480 and 600 nm. Emission at 480 nm corresponds to the benzothiazole chromophore, suggesting a twisted excited-state conformation. However, luciferin **3.2** exhibited a single fluorescence band of stronger intensity, likely corresponding to intramolecular proton transfer in the excited state (Figure 3-2C) [21]. Chromophore **3.2** is more conformationally locked about the 2'-2' C-C bond. Thus, we expected luciferin **3.2** to be a more robust bioluminescent emitter than luciferin **3.1**. However, to our surprise, both analogs exhibited very weak light output, six orders of magnitude less intense than D-luciferin (Figure 2B). This result



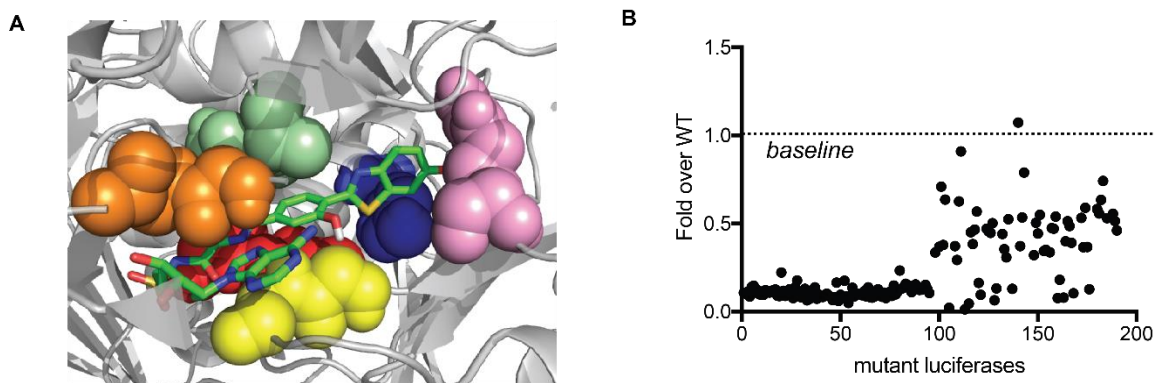
suggested that luciferins **3.1** and **3.2** are poorly processed by the wildtype enzyme. Such analogs could potentially be orthogonal to firefly luciferase. However, an engineered luciferase with improved activity toward the extended analogs would have to be identified.



**Figure 3-2. Fluorescence and bioluminescence of pi-extended luciferin analogs.** **A)** Fluorescence spectra of luciferin analogs **3.1** and **3.2** (100  $\mu$ M) in water. Luciferin **3.2** emits maximally at 518 nm. Luciferin **3.1** exhibits two weaker emission bands at 488 and 617 nm. **B)** Luciferin analogs are poor bioluminescent substrates with firefly luciferase (Fluc). Luciferins **3.1**, **3.2**, or D-luciferin (100  $\mu$ M) were incubated with Fluc (0.01 mg/mL) in imaging buffer and photon output was measured. Error bars represent the standard error of the mean for  $n = 3$  experiments. **C)** Tautomerization of luciferin **3.2** in the excited state mediated by intramolecular proton transfer.

### 3.3 Evolving luciferases with improved activity

We initially targeted residues in the luciferase active site important for binding the native substrate (Figure 3-3A) [22]. These residues were systematically mutated to a panel of amino acids and libraries of mutant luciferases were screened for activity. However, we were unable to find a light-emitting starting point (Figure 3-3B).



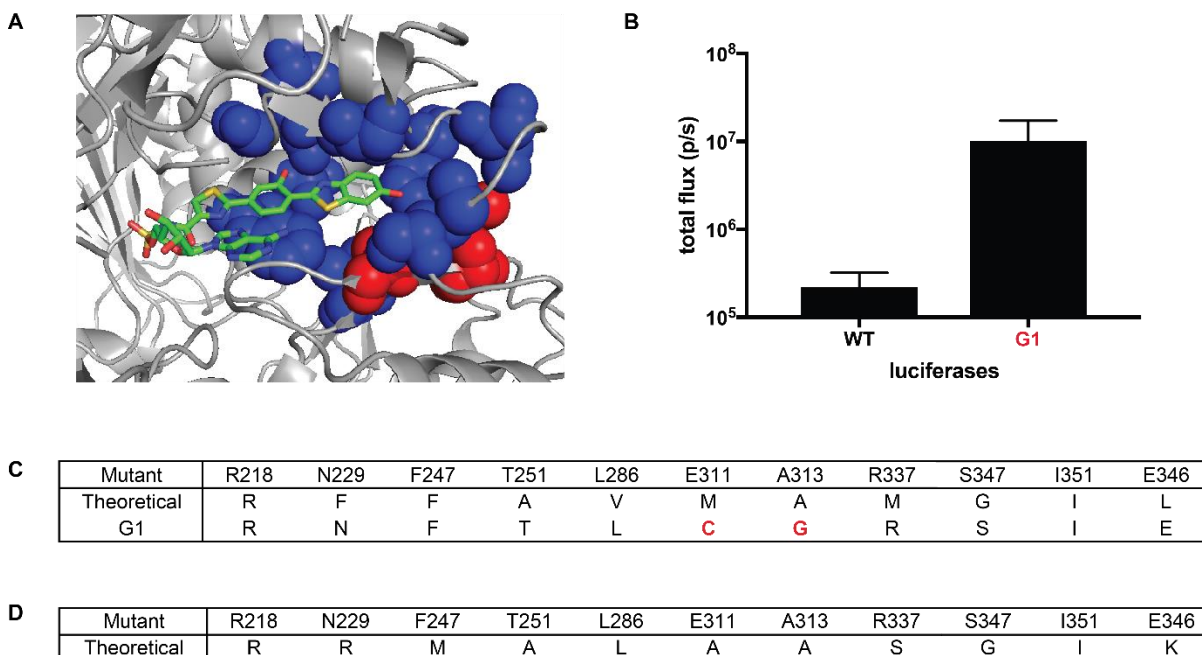
**Figure 3-3. Screening of site-directed libraries failed to yield and improved mutant.** **A)** The following residues were targeted for mutagenesis: G246 and F247 (red), Y255 (blue), G315 and G316 (yellow), R337 (pink), G341 and L342 (orange), S347 and A348 (green). **B)** Library members did not show improvement over Fluc (WT). Representative screening data of 200 mutant luciferases is shown.

To remodel the luciferase active site (in collaboration with Prof. Jeremy Mills, ASU), we employed Rosetta modeling to design a luciferase binding pocket with improved packing with the luciferin analogs (Figure 3-4A). From the computationally optimized protein sequence, we selected 11 residues for mutagenesis (Figure 3-4C,D). Combinatorial codon mutagenesis (CCM) [23] was used to target this large number of sites in a single library. On average, 2-3 mutations were combined on a single clone. While beneficial epistasis resulting from >3 mutations may be missed by this approach, such a high degree of mutations would likely result in unfolded protein.

Upon screening the library for activity with analogs **3.1** and **3.2**, we identified a double mutant E311C, A313G (**G1**) with nearly 50-fold improved activity with analog **3.2** (Figure 3-4B,C). Mutation at E311 was predicted in the Rosetta-optimized sequence, suggesting computation was useful in identifying functionally important residues. However, mutation at all the predicted sites was not required to restore light emission. We hypothesize mutations E311C and A313G induce a change in binding pocket shape

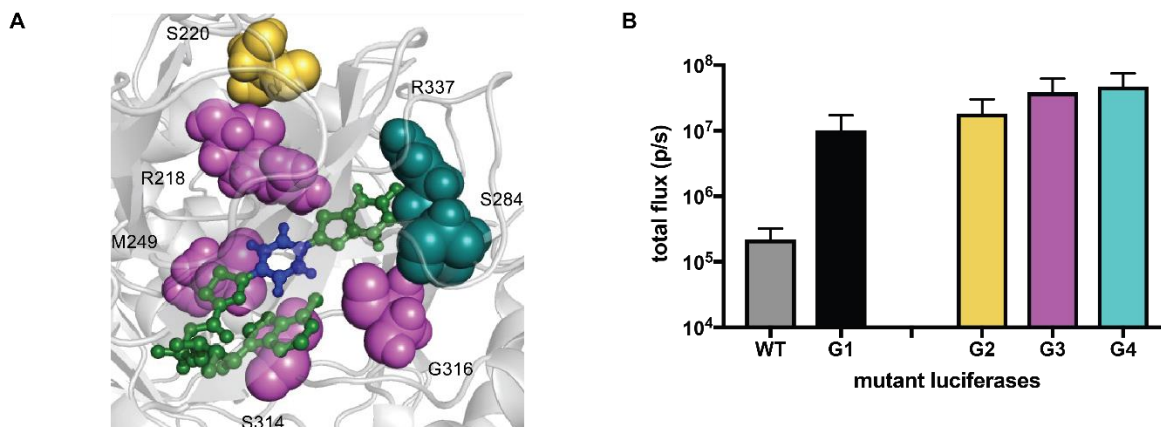
amenable for light emission with analog **3.2**. We explore changes in luciferase binding pocket shape and luciferin conformation in section 3.4.

Unfortunately, a luciferase “hit” was not identified for analog **3.1**. This result suggests that the binding pocket could be further optimized to promote binding with analog **3.1**. However, remodeling of the active site may require several (i.e. >2) synergistic mutations, as suggested by the Rosetta modeling (Figure 3-4D). Such a mutant may be difficult to find when screening the large libraries required to target multiple mutations simultaneously. Alternatively, additional screening of the CCM library may be necessary. This library contains  $20^{11}$  unique luciferases, though we only screened a small fraction of the total population (see section 3.7k).



**Figure 3-4. Computationally-guided library design. A)** Rosetta modeling was used to optimize residues in the binding pocket for binding with pi-extended luciferin analogs. 11 residues were selected for mutagenesis (blue). Beneficial mutations (red) were identified upon screening the mutant luciferase library. **B)** A mutant luciferase with 50-fold improved activity was identified. Luciferases (mutant G1 or Fluc) were expressed in E-coli and lysate was incubated with luciferin analog **3.2** (100  $\mu$ M) and ATP (1 mM). Photon output was then measured. **C)** Sequence analysis of improved mutant G1 and Rosetta-optimized mutant for luciferin **3.2** (theoretical). Beneficial mutations are highlighted in red. **D)** Sequence analysis of Rosetta-optimized mutant for luciferin analog **3.1**.

With a light emitting luciferase (mutant G1) in hand, we sought to evolve for further improvement in light output. Like our initial screening attempts, we again targeted binding pocket residues for mutagenesis (Figure 3-5). Three subsequent generations of site-directed mutagenesis and screening identified lead mutants G2-G4 with 2-5 fold further improvement in activity (Figure 3-5B). However, improvement in activity seemed to plateau, suggesting that we reached a local maximum on the evolutionary landscape [24].



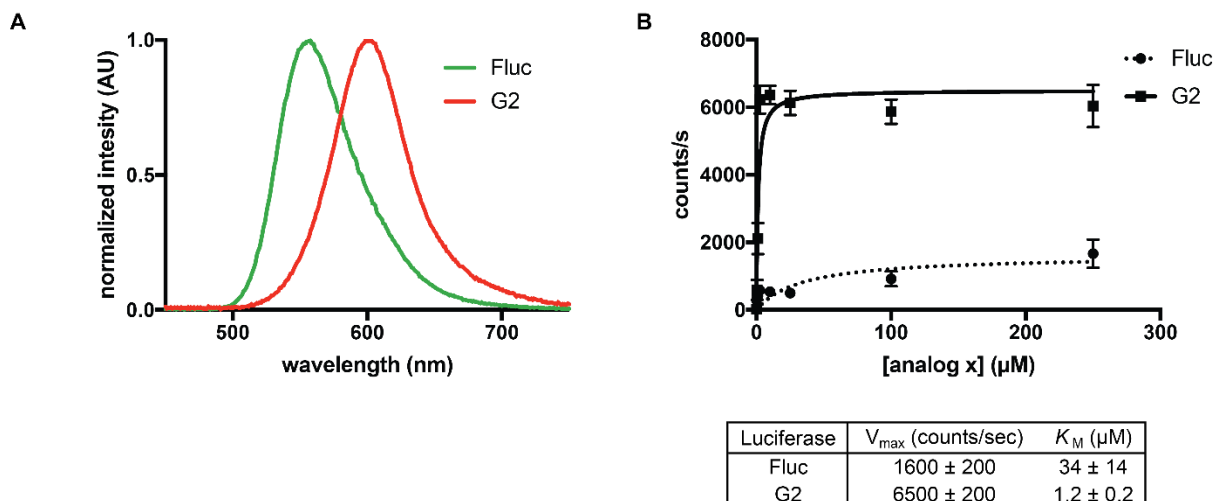
**Figure 3-5. Directed evolution of mutant G1.** Activity of first generation mutant (G1) was improved through three additional rounds of mutagenesis and screening. **A)** Residues targeted for mutagenesis in second generation (yellow), third generation (purple) and fourth generation (teal) libraries. **B)** The top hit from each generation exhibited improved activity (2-5 fold) over the first-generation lead.

### 3.4 Characterization of evolved luciferases

We wished to understand the origin of improved activity. The most significant improvement in bioluminescence was observed with mutant G1. Mutant G1 is comprised of mutations E311C, A313G. E311 is known to participate in H-bonding interactions with a water network located in the binding pocket [25]. Mutation at E311 is known to red-shift bioluminescent color, which corresponds to an opening of the binding pocket [26]. Indeed, mutant G1 does exhibit red-shifted emission with D-luciferin (Figure 3-6A), suggesting a similar change in active site conformation is occurring.

We next sought to understand how these mutations affect the kinetics of the luciferase enzyme. Mutant G2 was chosen to be expressed and purified based on its enhanced thermostability (data not shown) relative to the other mutants. The purified enzyme was incubated with a range of substrate concentrations. Initial rates of photon output were then measured and apparent  $K_m$  and  $k_{cat}$  values were determined (Figure 3-6B). From this data, we determined that mutant G2 exhibited improved catalytic

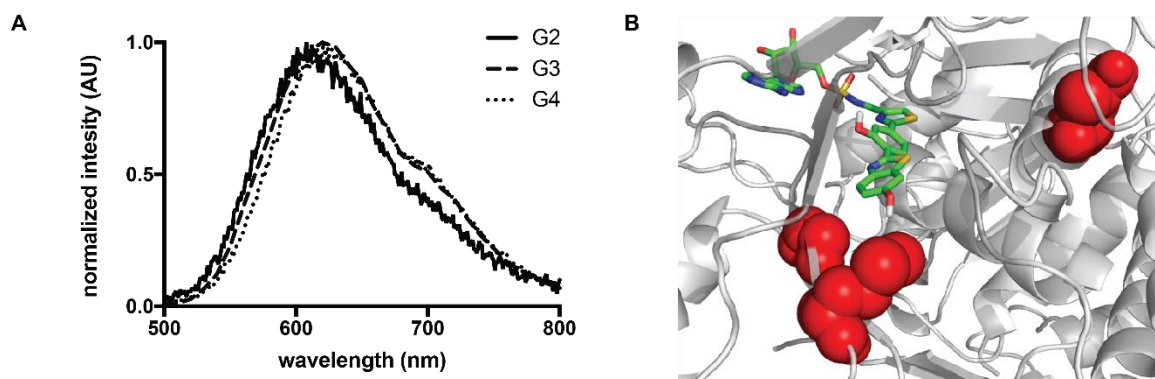
efficiency over the wildtype enzyme. We hypothesize that mutations at E311C and A313G enable an improvement in substrate turnover and binding because of a reshaping of the binding pocket.



**Figure 3-6. Biochemical characterization of mutant G2.** **A)** Mutant G2 exhibits red-shifted bioluminescence with D-luciferin, suggesting an open active site conformation. Luciferases (0.1 mg/mL) were incubated with D-luciferin (100  $\mu\text{M}$ ) and ATP (1 mM) in imaging buffer. Emission spectra were acquired over a range of wavelengths. **B)** Initial rates of light emission for mutant G2 and Fluc with luciferin analog **3.2**. Fluc and mutant G2 were incubated with analog **3.2** (1 – 250  $\mu\text{M}$ ) and light emission values were recorded. Data were then fit according to Michaelis-Menten kinetics and binding constants ( $K_M$ ) and rate constants ( $k_{\text{cat}}$ ) were obtained. Error bars represent the standard error of the mean for  $n = 9$  experiments.

We wished to explore how the engineered luciferases affect substrate conformation in the binding pocket. To do this, we obtained bioluminescence spectra for mutants G2-G4 with luciferin **3.2** (Figure 3-7A). In all cases, bioluminescent emission was constant, with maximum emission observed at  $\sim 600$  nm. This wavelength is inconsistent with the planarized chromophore observed in fluorescence (Figure 3-2A), which emits maximally at 520 nm. Furthermore, the observed spectrum doesn't match the computational prediction for the planarized chromophore (Figure 3-1B). Thus, we hypothesize that luciferin **3.2** adopts a twisted conformation in the luciferase binding

pocket. This hypothesis is corroborated by docking simulations, which predicted a non-planar ligand conformation (Figure 3-7B). We reasoned that regardless of ligand conformation, analog **3.2** and the engineered luciferases could be orthogonal to existing bioluminescent probes.

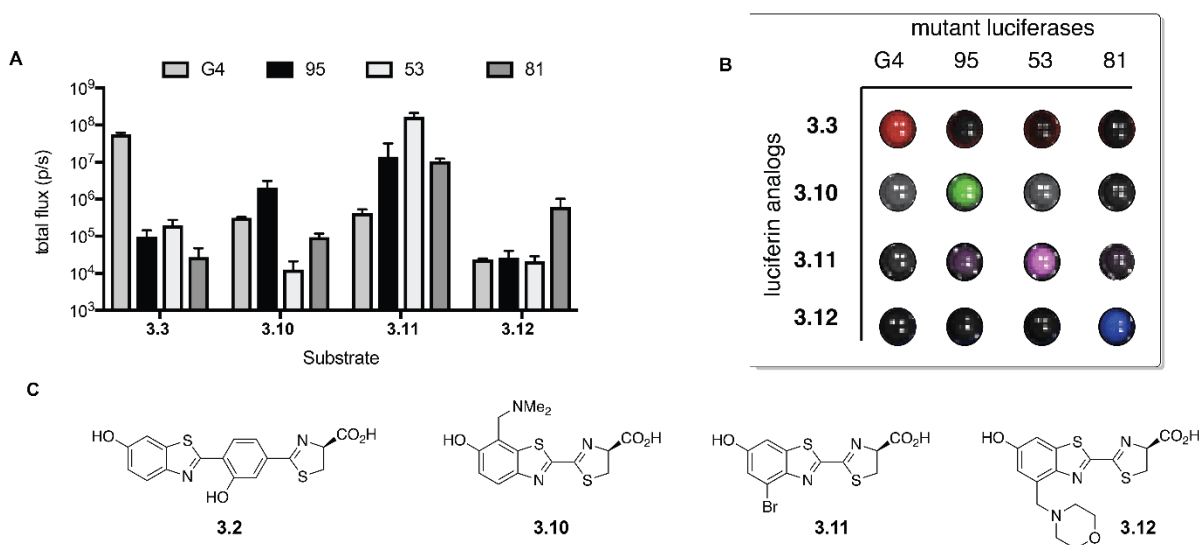


**Figure 3-7. Pi-extended luciferin may adopt a twisted conformation** **A)** Bioluminescence spectra of evolved luciferases G2-G4 with luciferin analog **3.2**. Luciferases (0.1 mg/mL) were incubated with analog **3.2** (100  $\mu$ M) and ATP (1 mM) in imaging buffer and light emission was detected over a range of wavelengths. **B)** Computationally docked structure of luciferin **3.2** and mutant G2. Analog **3.2** adopts a twisted conformation. Simulations were performed using Autodock Vina [27].

### 3.5 Orthogonal imaging with engineered bioluminescent pairs

Orthogonal luciferase-luciferin pairs can be rapidly identified through parallel screening of luciferin analogs and mutant luciferases. Such methods generate large amounts of data. The Prescher lab recently developed a computer script to evaluate  $>10^5$  bioluminescent output pairs. However, computational cost increases greatly when searching for higher order orthogonal sets, such as triplets and quartets [4]. We reasoned that luciferin **3.2** and an engineered luciferase would be orthogonal to existing orthogonal pairs and triplets. This would rapidly provide expanded sets of orthogonal probes. Luciferin **3.2** and luciferase G4 were orthogonal to an existing triplet pair [4],

providing a novel quartet set of luciferases (Figures 3-8A-C). Across the set, orthogonality is resolved by at least an order of magnitude in light output. Such signals can be unmixed from background emission using standard algorithms [1]. Indeed, each matching luciferase-luciferin pair could be preferentially detected (Figure 3-8B), suggesting these probes could be used to monitor four components simultaneously.



**Figure 3-8. Bioluminescence imaging with an orthogonal quartet. A)** Orthogonal quartet set. Bacteria expressing luciferases were lysed and incubated with their corresponding luciferin (100  $\mu$ M) and ATP (1 mM). Photon outputs were quantified and error bars represent the standard error of the mean for  $n = 3$  experiments. **B)** Sample images are shown. Light emission from matching luciferin/luciferase pairs were unmixed from background emission using an unmixing algorithm. Light emission was false colored for clarity. **C)** Structures of orthogonal luciferin analogs.

### 3.6 Conclusion

In summary, we designed pi-extended, conformationally flexible luciferins as orthogonal bioluminescence substrates. These probes were good candidates as orthogonal probes based on their unique structures relative to conventional luciferin substrates. Indeed, these luciferin analogs were completely inactive with the wildtype enzyme, highlighting its divergence in structure from the natural substrate, D-luciferin.



We then turned to protein engineering to identify a luciferase that could process the luciferin analogs. We initially constructed site-directed libraries targeting residues in the binding pocket. However, when we screened these libraries for activity with the analogs, we were unable to find a light-emitting enzyme. We reasoned that rational remodeling of the active site was needed to complement the luciferin analogs. Rosetta modeling was employed to design a complementary binding pocket. The resulting optimized protein sequence was used to identify residues of interest for mutagenesis. Mutational sites were then combined in combinatorial fashion to create a single library. Upon screening this diverse library of luciferases, we identified a double mutant with >10-fold improved activity with analog **3.2**. Subsequent generations of site directed mutagenesis produced mutants with 2-5 fold further improvement in activity. This engineered luciferin/luciferase pair was broadly orthogonal to existing luciferase-luciferin pairs, expanding the toolset of orthogonal probes. We demonstrated an orthogonal quartet, selectively identifying four luciferase mutants based on substrate preference. These tools expand the orthogonal bioluminescent toolkit for multicomponent imaging.

### **3.7 Materials and methods**

#### **3.7a Computational methods**

Calculations were performed with Spartan Student Edition Version 6. Organic structures were modeled and energy minimized. Bond torsion energies were calculated using B3LYP 6-311G\* and HOMO-LUMO energies were calculated using B3LYP 6-311++G(2df,2pd).

### **3.7b Fluorescence emission spectra**

Fluorescence spectra were acquired using a Cary Eclipse spectrometer (with 1 nm excitation and emission slit widths). Luciferin analogs (0.1 mM) in water were placed in a quartz cuvette (10 mm path length), excited at the appropriate wavelength and fluorescence emission was scanned over a range of wavelengths.

### **3.7c Recombinant protein expression and purification**

Native Fluc and mutant luciferases were expressed and purified as previously described in chapter 2 section 2.8b.

### **3.7d Light emission assays with recombinant luciferase**

All bioluminescence assays were carried out as described in sections 2.8f,h and i.

### **3.7e Multicomponent imaging with orthogonal quartet**

Orthogonal luciferases were expressed and assayed as described in section 2.9p. Bacterial lysates (90  $\mu$ L) expressing each orthogonal luciferases (G4, 95, 53, 81) were incubated with luciferin analog **3.2** (100  $\mu$ M) and ATP (1 mM). Bacterial lysate containing the positive pairing (G4 and analog **3.2**) was separately prepared as a positive control. Bioluminescence signal from the positive pairing could be unmixed from signal from the negative pairings using a spectral unmixing algorithm [1]. This experiment was repeated on separate imaging plates for luciferin analogs **3.10**, **3.11** and **3.12** and their respective selective luciferases.

### 3.7f Mutant luciferase library generation

Two sections of the luciferase gene (*pgl4-luc2*), denoted R1 and R2, were targeted for gene assembly. Libraries and primers are denoted by “SD” (site directed), followed by the number of the first residue mutated in the library. The R1 region comprises amino acids 199-257 and was assembled according to Jones, *et al.* [11] with site directed mutagenesis primers replacing the respective wild type primers. The R2 region comprises amino acids 285-346 and was assembled according to Jones *et al.* [11] with site directed mutagenesis primers replacing the respective wild-type primers at the desired sites (see primer list). For generational libraries (Figure 3-5), mutations from the previous generation were included in the assembly primers and kept constant. Libraries SD246, 341, and 347 (Figure 3-3) were provided by K. Jones. All library assembly PCR reactions were run using Q5 HotStart DNA polymerase (New England BioLabs).

### 3.7g Construction of mutant G4 construct

Mutants generated from generation 3 screening were combined to yield mutant G4. R1 and R2 regions were amplified and ligated according to Jones, *et. al.*[11] The R1 region was amplified from template containing mutant 213 (R218K, S220N, M249L, E311C, A313G). The R2 region was amplified from template containing mutant 212 (S220N, E311C, A313G, S314C, G316S). The R1 and R2 regions were ligated, amplified and introduced to pET28a vector via Gibson assembly as described previously [11].

### 3.7h Gene assembly primer list

#### **SD337 primers**

**SD337-F1**

GCATC**NDT**CAGGGCTACG

**SD337-R1**

CGTAGCCCTGAHNGATGC

**SD218 249 generation 3 library primers**

**R1\_F35\_NDT218\_S220N**  
GCACCGCACCGCTTGTGTCNDT T TCAA  
**R1\_F35\_VHG218\_S220N**  
GCACCGCACCGCTTGTGTCVHG T TCAA  
**R1\_F35\_R218W\_S220N**  
GCACCGCACCGCTTGTGTC TGG T TCAA  
**R1\_R51\_NDT218\_S220N**  
GCCGAAGATGGGGTTCGCGGGCATGGTTGAAAHN GAC  
**R1\_R51\_VHG218\_S220N**  
GCCGAAGATGGGGTTCGCGGGCATGGTTGAA CDB GAC  
**R1\_R51\_R218W\_S220N**  
GCCGAAGATGGGGTTCGCGGGCATGGTTGAA CCA GAC  
**R1\_F141\_NDT249**  
TTCGGCNDT T TCCACCACGCTGGGCTACTTGATCTGCG  
**R1\_F141\_VHG249**  
TTCGGCVHG T TCCACCACGCTGGGCTACTTGATCTGCG  
**R1\_F141\_M249W**  
TTCGGCTGG T TCCACCACGCTGGGCTACTTGATCTGCG  
**R1\_R124\_NDT249**  
CGTGGTGAAAHN GCCGAAGCCGTGGTGAAATGGCA  
**R1\_R124\_VHG249**  
CGTGGTGAA CDB GCCGAAGCCGTGGTGAAATGGCA  
**R1\_R124\_M249W**  
CGTGGTGAA CCA GCCGAAGCCGTGGTGAAATGGCA

**SD314 316 generation 3 library primers**

**Millsg2\_F117\_1**  
CTAAGCAACTTGCAC TGCATCGGCNDTNDTNDT  
**Millsg2\_F117\_2**  
CTAAGCAACTTGCAC TGCATCGGCNDTNDTVHG  
**Millsg2\_F117\_3**  
CTAAGCAACTTGCAC TGCATCGGCNDTNDTTGG  
**Millsg2\_F117\_4**  
CTAAGCAACTTGCAC TGCATCGGCVHGNDTNDT  
**Millsg2\_F117\_5**  
CTAAGCAACTTGCAC TGCATCGGCVHGNDTVHG  
**Millsg2\_F117\_6**  
CTAAGCAACTTGCAC TGCATCGGCVHGNDTTGG  
**Millsg2\_F117\_7**  
CTAAGCAACTTGCAC TGCATCGGCTGGNDTNDT  
**Millsg2\_F117\_8**  
CTAAGCAACTTGCAC TGCATCGGCTGGNDTVHG  
**Millsg2\_F117\_9**  
CTAAGCAACTTGCAC TGCATCGGCTGGNDTTGG  
**Millsg2\_R135\_1**  
TTGCTGAGCGGCGCAHNAHNAHNGCCGATCGG  
**Millsg2\_R135\_2**  
TTGCTGAGCGGCGCCDBAHNAHNGCCGATCGG  
**Millsg2\_R135\_3**  
TTGCTGAGCGGCGCCAAHNAHNGCCGATCGG  
**Millsg2\_R135\_4**

TTGCTGAGCGGCGCAHNAHNCDBGCCGATCGG  
Millsg2\_R135\_5  
TTGCTGAGCGGCGCCDBAHNCDBGCCGATCGG  
Millsg2\_R135\_6  
TTGCTGAGCGGCGCCAAHNCDBGCCGATCGG  
Millsg2\_R135\_7  
TTGCTGAGCGGCGCAHNAHNCAGCCGATCGG  
Millsg2\_R135\_8  
TTGCTGAGCGGCGCCDBAHNCAGCCGATCGG  
Millsg2\_R135\_9  
TTGCTGAGCGGCGCCAAHNCAGCCGATCGG  
Millsg2\_R97\_1  
GTGCAAGTTGCTTAGGTCGTA CTGTGCGATGAGAGT

### **SD284 337 generation 4 library primers**

R2-F117-gen4a  
CTAAGCAACTTGCACTGCATCGGCACCGGCGGG  
R2-R97-gen4  
CAGTGCAAGTTGCTTAGGTCGTA CTGTGCGATGAGAGT  
R2-R135-gen4a  
TTGCTGAGCGGCGCCCCGCCGGTGCCGATG  
R2-F117-gen4b  
CTAAGCAACTTGCACTGCATCGGCTGCGGCTCT  
R2-R97-gen4  
CAGTGCAAGTTGCTTAGGTCGTA CTGTGCGATGAGAGT  
R2-R135-gen4b  
TTGCTGAGCGGCGCAGAGCCGCAGCCGATG  
R2-F36-gen4-1  
AGACTATAAGATTCAANDTGCCCTGCTGGTGCCACAC  
R2-F36-gen4-2  
AGACTATAAGATTCAAVHGGCCCTGCTGGTGCCACAC  
R2-F36-gen4-3  
AGACTATAAGATTCAATGGGCCCTGCTGGTGCCACAC  
R2-R19-gen4-1  
GGGCAHNTTGAATCTTATAGTCTTGCAAGCTTCGCAAGAA  
R2-R19-gen4-2  
GGGCCDBTTGAATCTTATAGTCTTGCAAGCTTCGCAAGAA  
R2-R19-gen4-3  
GGGCCCAATTGAATCTTATAGTCTTGCAAGCTTCGCAAGAA  
R2-F181-gen4  
GTGGCCAAACGCTTCCACCTACCAGGCATC  
R2-F213-gen4-1  
NDTCAGGGCTACGGCCTGACAGAAACA ACTAGTGCCA  
R2-F213-gen4-2  
VHGCAGGGCTACGGCCTGACAGAAACA ACTAGTGCCA  
R2-F213-gen4-3  
TGGCAGGGCTACGGCCTGACAGAAACA ACTAGTGCCA  
R2-R197-gen4-1  
AGCCCGTAGCCCTGAHNGATGCCTGGTAGGT  
R2-R197-gen4-2  
AGCCCGTAGCCCTGCDBGATGCCTGGTAGGT  
R2-R197-gen4-3  
AGCCCGTAGCCCTGCCAGATGCCTGGTAGGT

### **Vector amplification primers**

**R2-vector-Fwd**  
TAGTGCCATTCTGATCACCCCC  
**R1-vector-Rev**  
CGCAAGAATAGCTCCTCCTCG

### **3.7h Library plasmid construction**

The libraries generated above were ultimately inserted into a vector lacking the R1 and R2 regions. This deletion vector (pET28-R1-R2del) was provided by W. Porterfield and generated according to Jones, et al.[11] The pET28-R2del was linearized by amplification using primers R2-vector-Fwd and R1-vector-Rev (Table S3). Q5 Hot start polymerase (New England BioLabs) was used in the amplifications. The DNA was then digested with DpnI for 1 h at 37 °C, and the products were purified via a PCR cleanup kit (Qiagen). The resulting linear vector (50 ng) was combined in a 1:2 (vector:insert) ratio with library insert DNA described above. The insert and vector were added to a Gibson master mix (New England BioLabs) and assembled at 50 °C for 20 min. A portion of the assembled product (2.5 µL) was utilized directly for transformation into Top10 chemically competent *E. coli* cells. Transformants were plated on agar plates containing kanamycin antibiotic (40 µg/mL) and saved for further analysis.

### **3.7i CCM library construction**

From a computationally optimized protein sequence (Rosetta modelling) 11 mutational sites were selected for mutagenesis. Forward and reverse primers encoding the NDT degenerate codon at these mutational sites were commercially synthesized. These primers were used in tandem to introduce 11 mutational sites to luciferase template

pET28a-luc2 (generation 1 library) or mutant G1 (generation 2 library) in combinatorial fashion as described previously.

### 3.7j CCM primer list

#### Primers for generation 1 CCM library

**CCM1-F-218**  
CGCTTGTGTC**NDT**TTTCAGTCATGCC  
**CCM1-R-218**  
GGGCATGACTGAA**AHN**GACACAAGCG  
**CCM1-F-222**  
GATTCAGTCAT**NDT**CGCGACCCCATC  
**CCM1-R-222**  
GATGGGGTCGCG**AHN**ATGACTGAATC  
**CCM1-F-240**  
GCTATCCTCAGC**NDT**GTGCCATTTAC  
**CCM1-R-240**  
GTGAAATGGCAC**AHN**GCTGAGGATAGC  
**CCM1-F-247**  
CACCACGGC**NDT**GGCATGTTC  
**CCM1-R-247**  
GAACATGCC**AHN**GCCGTGGTG  
**CCM1-F-250**  
CTTCGGCATG**NDT**ACCACGCTG  
**CCM1-R-250**  
CAGCGTGGT**AHN**CATGCCGAAG  
**CCM1-F-311**  
CAACTTGAC**NDT**ATCGCCAGCG  
**CCM1-R-311**  
CGCTGGCGAT**AHN**GTGCAAGTTG  
**CCM1-F-313**  
ACGAGATC**NDT**AGCGGCGGG  
**CCM1-R-313**  
CCCGCCGCT**AHN**GATCTCGT  
**CCM1-F-340**  
CCAGGGC**NDT**GGCCTGAC  
**CCM1-R-340**  
GTCAGGCC**AHN**GCCCTGG  
**CCM1-F-347**  
CTGACAGAAACA**NDT**GCCATTCTGATCACC  
**CCM1-R-347**  
GGTGATCAGAATGGC**AHN**AGTTGTTTCTGTCAG  
**CCM1-F-354**  
CACCCCC**NDT**GGGGACGA  
**CCM1-R-354**  
TCGTCCCC**AHN**GGGGGTG  
**Mills\_CCM1-F-229**  
GACCCCATCTTCGGC**NDT**CAGATCATC  
**Mills\_CCM1-R-229**  
GATGATCTG**AHN**GCCGAAGATGGGGTC  
**Mills\_CCM1-F-251**  
CTTCGGCATGTT**NDT**ACGCTGG

**Mills\_CCM1-R-251**  
CCAGCGT**AHN**GAACATGCCGAAG  
**Mills\_CCM1-F-286**  
ATTCAATCTGCC**NDT**CTGGTGCCC  
**Mills\_CCM1-R-286**  
GGGCACCAG**AHN**GGCAGATTGAAT  
**Mills\_CCM1-F-337**  
CCAGGCATC**NDT**CAGGGCTACG  
**Mills\_CCM1-R-337**  
CGTAGCCCTG**AHN**GATGCCTGG  
**Mills\_CCM-F-351**  
TGCCATTCTG**NDT**ACCCCCGAAG  
**Mills\_CCM-R-351**  
CTTCGGGGGT**AHN**CAGAATGGCA  
**Mills\_CCM-F-351-354**  
TGCCATTCTG**NDT**ACCCCC**NDT**GGGGACGACAAGCCTGGC  
**Mills\_CCM-R-351-354**  
GCCAGGCTTGTCGTCCCC**AHN**GGGGGT**AHN**CAGAATGGCA  
**Mills\_CCM-F-311-313**  
GCAACTTGCAC**NDT**ATC**NDT**AGCGGCGGGGCGC  
**Mills\_CCM-R-311-313**  
GCGCCCCGCCGCT**AHN**GAT**AHN**GTGCAAGTTGC

### 3.7k Screening protocol

Screening of luciferase libraries in *E. coli* were performed as described in section 2.9 n and o. Site directed (SD) libraries were screened to completion (3X of the library size) on agar plates. Only a small fraction of the CCM libraries (~5000 colonies) were screened on agar plates.

### 3.7l General synthetic methods

Synthetic methods were performed as described in section 2.9r. Luciferin analog **3.2** and all synthetic intermediates leading to analog **3.2** were synthesized and by Z. Yao.

### 3.7m Synthetic procedures

**2-Bromo-6-methoxybenzo[d]thiazole (3.4):** The title compound was synthesized using a procedure from Yoshida, *et al.*, with some modifications [28]. To an oven-dried round bottom flask was added 6-methoxybenzo[d]thiazol-2-amine (**3.3**) (1.80 g, 10.0



mmol), isopentyl nitrite (1.6 mL, 12 mmol), *p*-toluenesulfonic acid (0.230 g, 13.0 mmol), CuBr<sub>2</sub> (23.0 mg, 0.10 mmol) and tetrabutylammonium bromide (6.50 g, 20.0 mmol). The flask was then flushed with nitrogen, and anhydrous MeCN (100 mL) was added. The resulting mixture was stirred at 25 °C for 45 min. The reaction was then diluted with ethyl acetate (100 mL), and washed with water (2 x 100 mL), and brine (1 x 100 mL). The organic layer was then dried with MgSO<sub>4</sub>, filtered and concentrated *in vacuo*. The resulting residue was purified via flash-column chromatography (eluting with a gradient of 5:95 to 10:90 EtOAc:hexanes) to yield **8** as a pale yellow solid (1.3 g, 5.2 mmol, 52%). The spectra were in accordance with reported characterization [28].

**2-Bromobenzo[*d*]thiazol-6-ol (3.5):** To a flame dried round bottom flask was added 2-bromo-6-methoxybenzo[*d*]thiazole (**3.4**) (1.22 g, 5.00 mmol). The flask was then flushed with nitrogen, and anhydrous CH<sub>2</sub>Cl<sub>2</sub> (150 mL) was added. The flask was cooled to -78 °C, and a 1.0 M solution of BBr<sub>3</sub> in CH<sub>2</sub>Cl<sub>2</sub> (50.0 mmol, 50.0 mL) was added dropwise. The mixture was allowed to slowly warm to room temperature and stirred for 24 h. The reaction was then quenched with a saturated solution of ammonium chloride (100 mL) and extracted into EtOAc (200 mL). The organic layer was then washed with saturated ammonium chloride (2 x 100 mL) and brine (1 x 100 mL). The organic layer was dried with MgSO<sub>4</sub>, filtered and concentrated *in vacuo*. The residue was purified via flash-column chromatography (eluting with a gradient of 10:90 to 30:70 EtOAc:hexanes) to yield **8** as a white solid (0.844 g, 3.67 mmol, 73%). <sup>1</sup>H NMR (400 MHz, acetone-*d*<sub>6</sub>) δ 7.77 (d, *J* = 8.9 Hz, 1H), 7.41 (d, *J* = 2.4 Hz, 1H), 7.05 (dd, *J* = 8.8, 2.5 Hz, 1 H); <sup>13</sup>C

(100 MHz, acetone- $d_6$ )  $\delta$  157.0, 147.2, 139.7, 135.0, 124.1, 117.1, 107.3; HRMS (ESI-)  $m/z$  calcd for  $C_7H_5BrNOS$   $[M+H]^+$  229.9275, found 229.9283.

**4-(4,4,5,5-Tetramethyl-1,3,2-dioxaborolan-2-yl)benzotrile (3.6):** To an oven dried round bottom flask was added 4-bromobenzotrile (1.00 g, 5.50 mmol), 1,1'-bis(diphenylphosphino)ferrocene]palladium(II) dichloride dichloromethane complex (140 mg, 0.17 mmol), bis(pinacolato)diboron (1.54 g, 6.16 mmol), and potassium acetate (1.6 g, 17 mmol). The flask was then flushed with nitrogen, and anhydrous DMSO (33 mL) was added. The flask was sealed and heated at 80 °C for 2 h. The flask was then diluted with ethyl acetate (5 mL), and washed with water (2 x 5mL), and brine (1 x 5 mL). The organic layer was dried with  $MgSO_4$ , filtered and concentrated *in vacuo*. The resulting residue was purified via flash-column chromatography (eluting with 1:20 EtOAc:hexanes) to yield **x** as a white solid (760 mg, 3.3 mmol, 60%).  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  7.89 (d,  $J$  = 8.4 Hz, 2H), 7.65 (d,  $J$  = 8.0 Hz, 2H), 1.36 (s, 12 H);  $^{13}C$  (125 MHz,  $CDCl_3$ )  $\delta$  135.3, 131.3, 119.0, 114.7, 84.7, 25.0 HRMS (ESI+)  $m/z$  calcd for  $C_{13}H_{16}BNNaO_2$   $[M+Na]^+$  252.1174, found 252.1174.

**4-(6-Hydroxybenzo[*d*]thiazol-2-yl)benzotrile (3.8):**

To an oven-dried round bottom flask was added 2-bromobenzo[*d*]thiazol-6-ol (**3.5**) (0.10 g, 0.43 mmol), 4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzotrile (**3.6**) (0.11 g, 0.52 mmol), 1,1'-bis(diphenylphosphino)ferrocene]palladium(II) dichloride dichloromethane complex (19 mg, 0.022 mmol), and dipotassium carbonate (0.36 g, 2.6 mmol). The flask was then flushed with nitrogen and DMF (15 mL) was added. The flask was sealed and heated to 80 °C for 5 h. The flask was then diluted with ethyl acetate

(50 mL), and washed with water (2 x 40 mL), and brine (1 x 40 mL). The organic layer was then dried with MgSO<sub>4</sub>, filtered and concentrated *in vacuo*. The residue was purified via flash-column chromatography (eluting with a gradient of 1:4 to 1:1 EtOAc:hexanes) followed by recrystallization in EtOAc to yield **10** as a white solid (42 mg, 0.17 mmol, 38%). <sup>1</sup>H NMR (500 MHz, acetone-*d*<sub>6</sub>) δ 8.26 (d, *J* = 8.5 Hz, 2H), 7.95 (d, *J* = 8.5 Hz, 1H), 7.51 (d, *J* = 2.5 Hz, 1H), 7.13 (dd, *J* = 9, 2.5 Hz, 1H); <sup>13</sup>C (100 MHz, CDCl<sub>3</sub>) δ 161.3, 156.4, 147.1, 137.0, 136.6, 133.2, 127.3, 124.1, 118.4, 116.8, 112.6, 106.8; HRMS (ESI-) *m/z* calcd for C<sub>14</sub>H<sub>7</sub>N<sub>2</sub>OS [M-H]<sup>-</sup> 251.0279, found 251.0287.

**(S)-2-(4-(6-hydroxybenzo[*d*]thiazol-2-yl)phenyl)-4,5-dihydrothiazole-4-carboxylic acid (3.1):** To a round bottom flask was added 4-(6-hydroxybenzo[*d*]thiazol-2-yl)benzotrile (**3.8**) (0.044g, 0.18 mmol), D-cysteine (0.31 g, 1.8 mmol), and sodium bicarbonate (0.15 g, 1.8 mmol). The flask was then flushed with nitrogen and absolute ethanol (8.8 mL) was added. Piperidine was added dropwise until the solution was pH 9. The flask was sealed and heated at 80 °C for 20 h. The mixture was then acidified to pH ~4 with 1 M sodium hydrogen sulfate, diluted with EtOAc (20 mL), and washed with 1 M sodium hydrogen sulfate (2 x 10 mL). The organic layer was basified to pH ~8 with saturated sodium bicarbonate, washed with EtOAc, then re-acidified to pH ~4 with 1 M sodium hydrogen sulfate. The aqueous layer was extracted with EtOAc (3 x 20 mL). The organic layers were combined and washed with water (2 x 20 mL) and brine (1 x 20 mL). The organic layer was then dried with MgSO<sub>4</sub>, filtered and concentrated *in vacuo* to provide **2** as a yellow solid (34 mg, 0.10 mmol, 55%). <sup>1</sup>H NMR (500 MHz, (CD<sub>3</sub>)<sub>2</sub>SO) δ 8.13 (d, *J* = 10 Hz, 2H), 7.95 (d, *J* = 10 Hz, 2H), 7.9 (d, *J* = 5 Hz, 1H), 7.45 (s, 1H),

7.03 (dd,  $J = 10, 5$  Hz, 1H), 5.35 (dd,  $J = 10, 5$  Hz, 1H), 3.76 (dd,  $J = 10, 10$  Hz, 1H), 3.67 (dd,  $J = 10, 10$  Hz, 1H);  $^{13}\text{C}$  (150 MHz,  $(\text{CD}_3)_2\text{SO}$ )  $\delta$  171.6, 167.5, 162.2, 156.1, 147.1, 136.3, 135.8, 133.8, 127.1, 123.8, 116.5, 106.8, 78.5, 35.1; HRMS (ESI+)  $m/z$  calcd for  $\text{C}_{17}\text{H}_{12}\text{N}_2\text{O}_3\text{S}_2$   $[\text{M}+\text{H}]^+$  357.0368, found 357.0367.

## References

1. Gammon, S. T.; Leevy, W. M.; Gross, S.; Gokel, G. W.; Piwnica-Worms, D. Spectral unmixing of multicolored bioluminescence emitted from heterogeneous biological sources. *Anal. Chem.* **2006**, *78*, 1520-1527.
2. Branchini, B. R.; Ablamsky, D. M.; Murtiashaw, M. H.; Uzasci, L.; Fraga, H.; Southworth, T. L. Thermostable red and green light-producing firefly luciferase mutants for bioluminescent reporter applications. *Anal. Biochem.* **2007**, *361*, 253-262.
3. Maguire, C. A.; Bovenberg, M. S.; Crommentuijn, M. H. W.; Niers, J. M.; Kerami, M.; Teng, J.; Sena-Esteves, M.; Badr, C. E.; Tannous, B. A. Triple Bioluminescence Imaging for *In Vivo* Monitoring of Cellular Processes. *Mol. Ther. Nucl. Acids* **2013**, *2*.
4. Rathbun, C. M.; Porterfield, W. B.; Jones, K. A.; Sagoe, M. J.; Reyes, M. R.; Hua, C. T.; Prescher, J. A. Parallel Screening for Rapid Identification of Orthogonal Bioluminescent Tools. *ACS Cent. Sci.* **2017**, *3*, 1254-1261.
5. Kaskova, Z. M.; Tsarkova, A. S.; Yampolsky, I. V. 1001 lights: Luciferins, luciferases, their mechanisms of action and applications in chemical analysis, biology and medicine. *Chem. Soc. Rev.* **2016**, *45*, 6048-6077.
6. Martini, S.; Haddock, S. H. Quantification of bioluminescence from the surface to the deep sea demonstrates its predominance as an ecological trait. *Sci. Rep.* **2017**, *7*, 45750.
7. Korpai, M.; Yan, J.; Lu, X.; Xu, S. W.; Lerit, D. A.; Kang, Y. B. Imaging transforming growth factor-beta signaling dynamics and therapeutic response in breast cancer bone metastasis. *Nat. Med.* **2009**, *15*, 960.
8. Stacer, A. C.; Nyati, S.; Moudgil, P.; Iyengar, R.; Luker, K. E.; Rehemtulla, A.; Luker, G. D. NanoLuc Reporter for Dual Luciferase Imaging in Living Animals. *Mol. Imaging* **2013**, *12*, 1-13.

9. Yao, Z.; Zhang, B. S.; Prescher, J. A. Advances in bioluminescence imaging: new probes from old recipes. *Curr. Opin. Chem. Biol.* **2018**, *45*, 148-156.
10. Mofford, D. M.; Reddy, G. R.; Miller, S. C. Aminoluciferins extend firefly luciferase bioluminescence into the near-infrared and can be preferred substrates over D-luciferin. *J. Am. Chem. Soc.* **2014**, *136*, 13277-13282.
11. Jones, K. A.; Porterfield, W. B.; Rathbun, C. M.; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. Orthogonal Luciferase-Luciferin Pairs for Bioluminescence Imaging. *J. Am. Chem. Soc.* **2017**, *139*, 2351-2358.
12. Iwano, S.; Obata, R.; Miura, C.; Kiyama, M.; Hama, K.; Nakamura, M.; Amano, Y.; Kojima, S.; Hirano, T.; Maki, S.; Niwa, H. Development of simple firefly luciferin analogs emitting blue, green, red, and near-infrared biological window light. *Tetrahedron* **2013**, *69*, 3847-3856.
13. Kuchimaru, T.; Iwano, S.; Kiyama, M.; Mitsumata, S.; Kadonosono, T.; Niwa, H.; Maki, S.; Kizaka-Kondoh, S. A luciferin analogue generating near-infrared bioluminescence achieves highly sensitive deep-tissue imaging. *Nat. Commun.* **2016**, *7*, 11856.
14. Jathoul, A. P.; Grounds, H.; Anderson, J. C.; Pule, M. A. A Dual-Color Far-Red to Near-Infrared Firefly Luciferin Analogue Designed for Multiparametric Bioluminescence Imaging. *Angew. Chem. Int. Ed.* **2014**, *53*, 13059-13063.
15. Hall, M. P.; Woodroffe, C. C.; Wood, M. G.; Que, I.; Van't Root, M.; Ridwan, Y.; Shi, C.; Kirkland, T. A.; Encell, L. P.; Wood, K. V.; Lowik, C.; Mezzanotte, L. Click beetle luciferase mutant and near infrared naphthyl-luciferins for improved bioluminescence imaging. *Nat. Commun.* **2018**, *9*, 132.
16. Anderson, J. C.; Grounds, H.; Jathoul, A. P.; Murray, J. A. H.; Pacman, S. J.; Tisi, L. Convergent synthesis and optical properties of near-infrared emitting bioluminescent infra-luciferins. *RSC Adv.* **2017**, *7*, 3975-3982.
17. Iwano, S.; Sugiyama, M.; Hama, H.; Watakabe, A.; Hasegawa, N.; Kuchimaru, T.; Tanaka, K. Z.; Takahashi, M.; Ishida, Y.; Hata, J.; Shimozono, S.; Namiki, K.; Fukano, T.; Kiyama, M.; Okano, H.; Kizaka-Kondoh, S.; McHugh, T. J.; Yamamori, T.; Hioki, H.; Maki, S.; Miyawaki, A. Single-cell bioluminescence imaging of deep tissue in freely moving animals. *Science* **2018**, *359*, 935-939.
18. Yoshida, S.; Yano, T.; Nishiyama, Y.; Misawa, Y.; Kondo, M.; Matsushita, T.; Igawa, K.; Tomooka, K.; Hosoya, T. Thiazolobenzynes: a versatile intermediate for multisubstituted benzothiazoles. *Chem. Commun.* **2016**, *52*, 11199-11202.

19. Loughlin, W. A.; Knevitt, S. A.; Hosking, R. E.; Marshall, R. L. Approaches to the high-throughput synthesis of analogues of dihydroaeruginosic acid. *Aust. J. Chem.* **2000**, *53*, 457-462.
20. Sasaki, S.; Drummen, G. P. C.; Konishi, G. Recent advances in twisted intramolecular charge transfer (TICT) fluorescence and related phenomena in materials chemistry. *J. Mater. Chem. C* **2016**, *4*, 2731-2743.
21. Das, K.; Sarkar, N.; Ghosh, A. K.; Majumdar, D.; Nath, D. N.; Bhattacharyya, K. Excited-State Intramolecular Proton-Transfer in 2-(2'-Hydroxyphenyl)Benzimidazole and 2-(2'-Hydroxyphenyl)-Benzoxazole - Effect of Rotamerism and Hydrogen-Bonding. *J. Phys. Chem.* **1994**, *98*, 9126-9132.
22. Branchini, B. R.; Southworth, T. L.; Murtiashaw, M. H.; Boije, H.; Fleet, S. E. A mutagenesis study of the putative luciferin binding site residues of firefly luciferase. *Biochemistry* **2003**, *42*, 10429-10436.
23. Belsare, K. D.; Andorfer, M. C.; Cardenas, F. S.; Chael, J. R.; Park, H. J.; Lewis, J. C. A Simple Combinatorial Codon Mutagenesis Method for Targeted Protein Engineering. *ACS Synth. Biol.* **2017**, *6*, 416-420.
24. Romero, P. A.; Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Bio.* **2009**, *10*, 866-876.
25. Nakatsu, T.; Ichiyama, S.; Hiratake, J.; Saldanha, A.; Kobashi, N.; Sakata, K.; Kato, H. Structural basis for the spectral difference in luciferase bioluminescence. *Nature* **2006**, *440*, 372-376.
26. Viviani, V. R.; Simoes, A.; Bevilaqua, V. R.; Gabriel, G. V. M.; Arnoldi, F. G. C.; Hirano, T. Glu311 and Arg337 Stabilize a Closed Active-site Conformation and Provide a Critical Catalytic Base and Counteraction for Green Bioluminescence in Beetle Luciferases. *Biochemistry* **2016**, *55*, 4764-4776.
27. Trott, O.; Olson, A. J. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31*, 455-461.
28. Yoshia, S.; Yano, T.; Nishiyama, Y.; Misawa, Y.; Kondo, M.; Matsushita, T.; Igawa, K.; Tomooka, K.; Hosoya, T. Thiazolobenzynes: a versatile intermediate for multisubstituted benzothiazoles. *Chem. Comm.* **2016**, *52*, 11199.

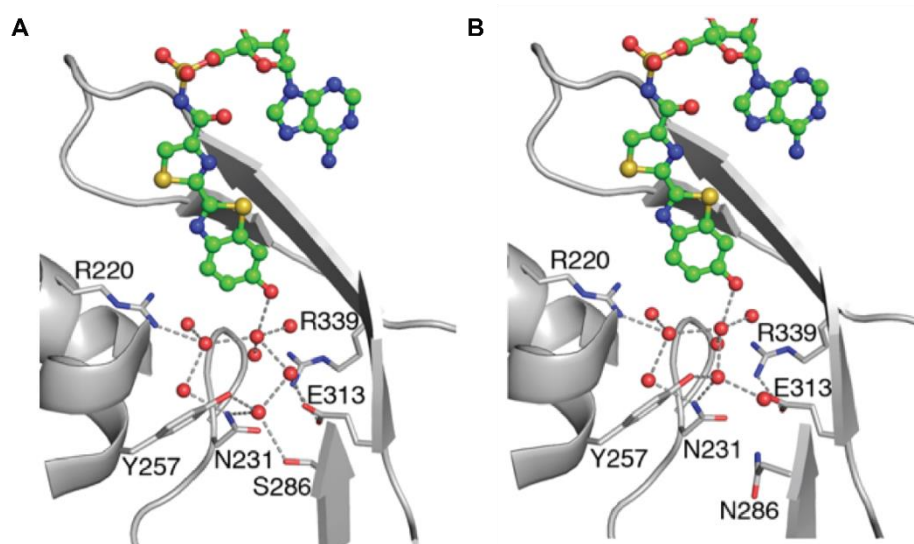
## Chapter 4: Expression and purification of luciferases for X-ray crystallography

### 4.1 Introduction

As exhibited in the evolved luciferases detailed in chapters 2 and 3, improvement in catalytic performance and shift in bioluminescence wavelength can be substrate specific. For example, in chapter 2, mutant 24 exhibited improved apparent catalytic efficiency with analog **2.4**. Mutant 24 also exhibited red-shifted emission and reduced catalytic efficiency with the analog **2.3**. Subsequent targeting of residues indicated by SCA analysis and a second round of screening with the pyridone analogs produced a lead mutant that restored catalytic efficiency with **2.3**. These results suggest that the luciferase-binding pocket can be reshaped to complement analogs with subtle differences in structure, such as the regioisomeric luciferin analogs. In chapter 3, we identified mutant luciferases with improved activity for extended analog **3.2**. The engineered luciferase-luciferin pair was orthogonal to other bioluminescent probes, again suggesting the luciferase binding pocket can be engineered to complement structurally distinct substrates. As described in the introduction, this paradigm is most obviously evident in the diversity of luciferin and luciferase structures in nature, though engineering orthogonal luciferase-luciferin pairs has been explored only recently.

While we collected indirect evidence of interactions between engineered luciferases and luciferin analogs through biochemical assays, we lacked concrete evidence of the structural features underlying selectivity. Understanding protein-substrate interactions at the molecular level could help inform further attempts to rationally design improved luciferases. Toward this end, we pursued crystallographic

structures of luciferase. Variances in luciferase structure have been previously studied through crystallography. For example, the structure of a red-shifted mutant luciferase revealed variances in residues that affect the structure of ordered water molecules in the binding pocket. (Figure 4-1) [1-2]. Subtle changes in interactions between the water network and the luciferin chromophore were evident in the crystal structures, suggesting the importance of this interaction to chromophore binding and color modulation.



**Figure 4-1. Variances in H-bonding residues affect water network.** Crystal structures of **A)** Native *Luciola cruciata* and **B)** *Luciola cruciata* bearing an S286N mutation. The native luciferase emits maximally at 560 nm and the mutant emits maximally at 605 nm. Figure reproduced with permission from ref. [1].

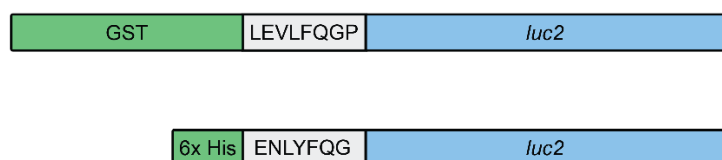
This chapter describes my efforts to obtain structural information of luciferase mutants through X-ray crystallography. Luciferase expression and purification were optimized in the process. Initial attempts to crystallize the apo protein using published methods resulted in protein precipitation and lack of protein crystals. Supplying known ligands to the crystallization conditions resulted in reduced protein precipitation. Thermal denaturation experiments suggest ligands stabilize the luciferase structure. We



hypothesize the protein/ligand complex accesses a more ordered and thermodynamically stable structure relative to the apo protein. Screening results with ligands have thus far revealed promising hits and will be the subject of continued investigation.

## 4.2 Expression and purification of firefly luciferase

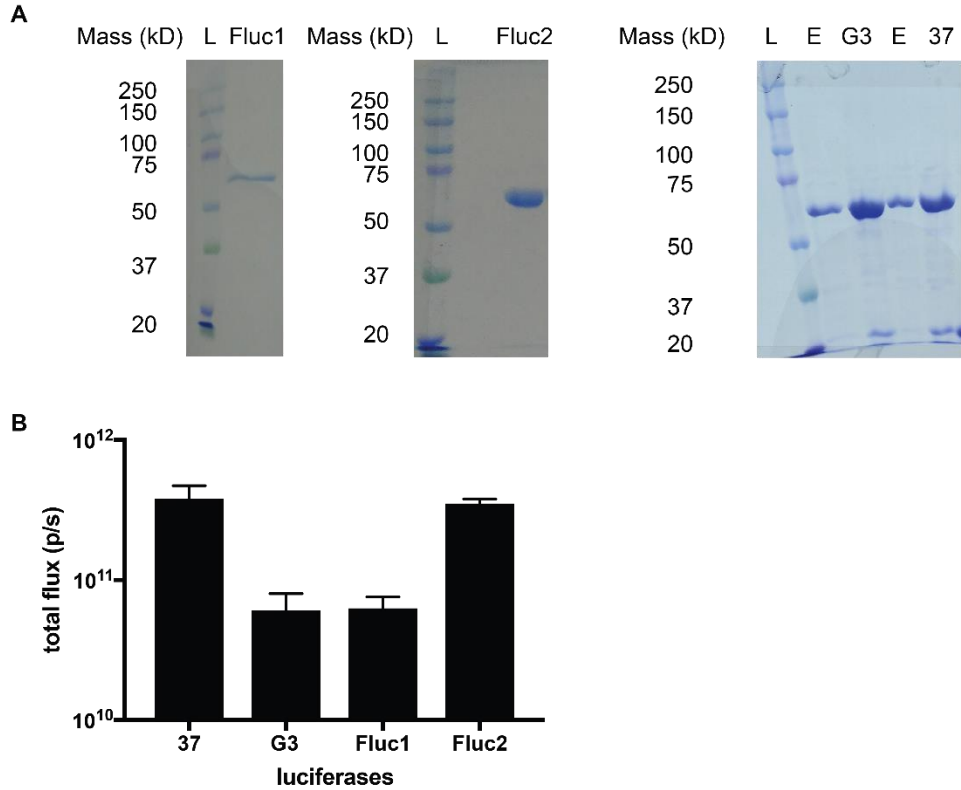
Structural studies of firefly luciferase have involved recombinantly expressed protein. In general, firefly luciferase has been expressed in *E. coli* and purified via affinity chromatography. We expressed and purified wild-type and mutant luciferases bearing either an N-terminal histidine tag [3] or glutathione S-transferase (GST) tag [4], linked by a protease recognition sequence (Figure 4-2). Bacterial lysates containing the luciferase constructs were bound to the appropriate affinity resin, washed, and eluted according to manufacturer specifications. Prior to crystallization experiments, the affinity tags were removed by the appropriate protease.



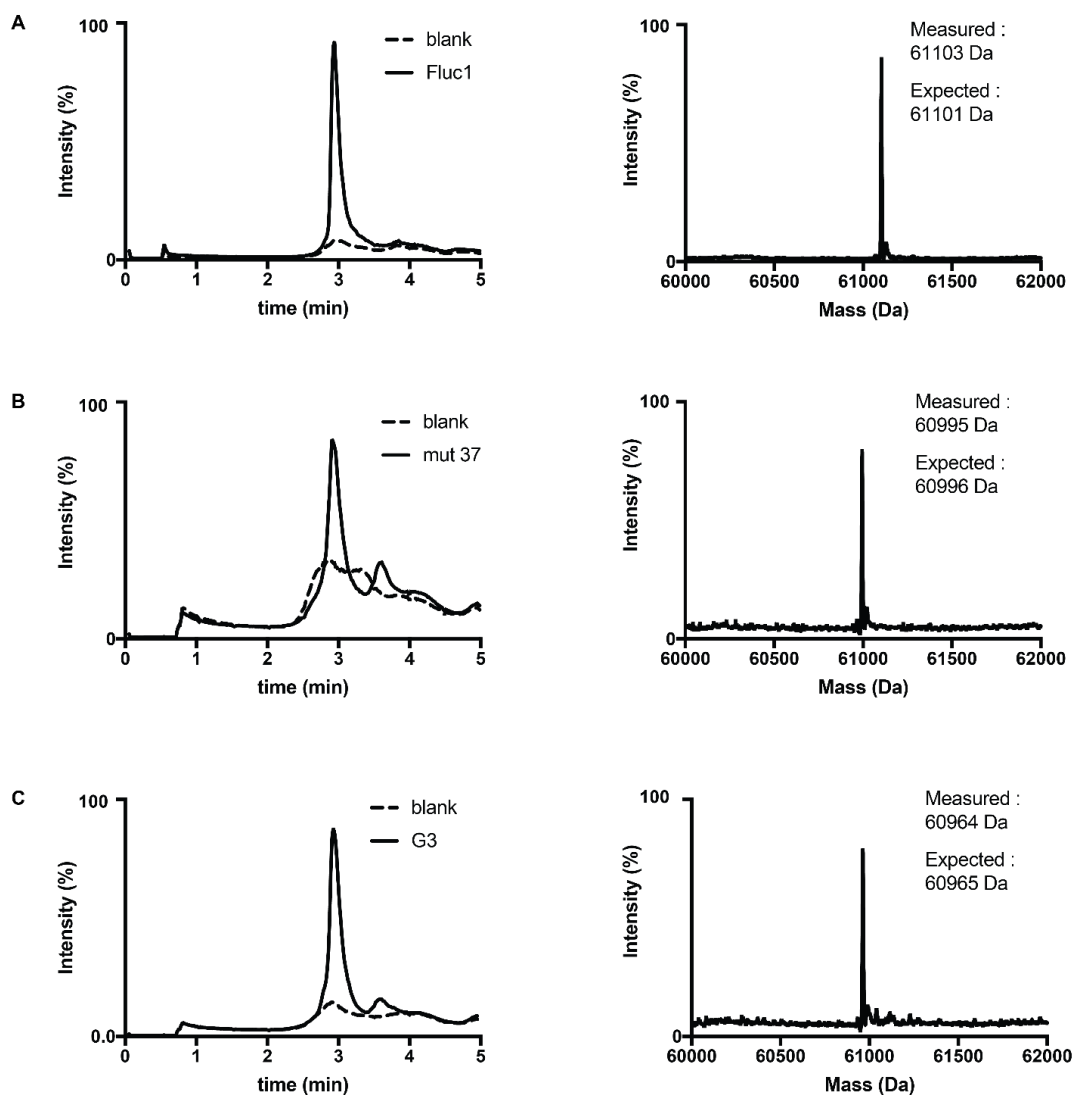
**Figure 4-2. Luciferase constructs for crystallography** **A)** N-terminal GST tag linked to luciferase via PreScission protease recognition sequence or **B)** N-terminal 6x His tag linked to luciferase via Tobacco Etch Virus (TEV) protease recognition sequence.

After subsequent removal of the protease, the luciferases were used in crystallization experiments. The purity and fidelity of the luciferases were validated by SDS-PAGE and bioluminescence assays (Figure 4-3). The chemical identity of the

luciferases was verified by LC/MS (Figure 4-4). Only the expected masses of the luciferases were observed.



**Figure 4-3. SDS PAGE analysis and activity of recombinant luciferases. A)** Prior to loading on 10% SDS PAGE gel, purified luciferase samples (2  $\mu$ g) were denatured at 95% for 20 min with 1x loading dye and 1% BME. L = ladder, Fluc1 = GST-Fluc construct post protease treatment, Fluc2 = 6xHis-Fluc construct post protease treatment, E = mutant luciferase pre-protease treatment, G3 = mutant G3 post protease, 37 = mutant 37 post protease. **B)** Bioluminescence of mutant luciferases and native Fluc with D-luciferin. D-luciferin (100 $\mu$ M) was incubated with Fluc (1  $\mu$ g) and ATP (1 mM) in imaging buffer and light emission was measured. Photon output was consistent with previous reports. Error bars represent the standard error of the mean for  $n \geq 3$  experiments.

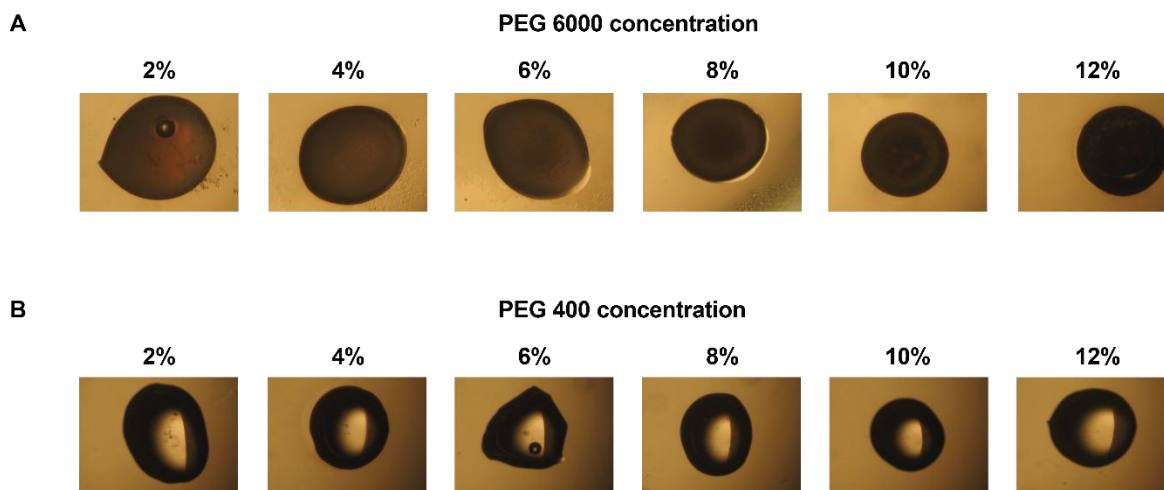


**Figure 4-4. LC/MS analysis of recombinant luciferases A) Fluc, B) mutant 37 and C) mutant G3.** Luciferases (100 ng) or a blank sample was passed through a C4 HPLC column and ion abundance was measured as a function of retention time. Mass spectra were summed over the area of the most predominant peak in the chromatograms. Deconvolution software was used to determine the average isotopic mass of each protein, which matched well with expected values.

### 4.3 Screening published crystallography conditions

We first attempted to reproduce conditions for crystallizing wild-type firefly luciferase [4], using a concentration of 10-20 mg/mL. However, heavy precipitation was observed. Precipitation was dependent on polyethylene glycol (PEG) molecular weight

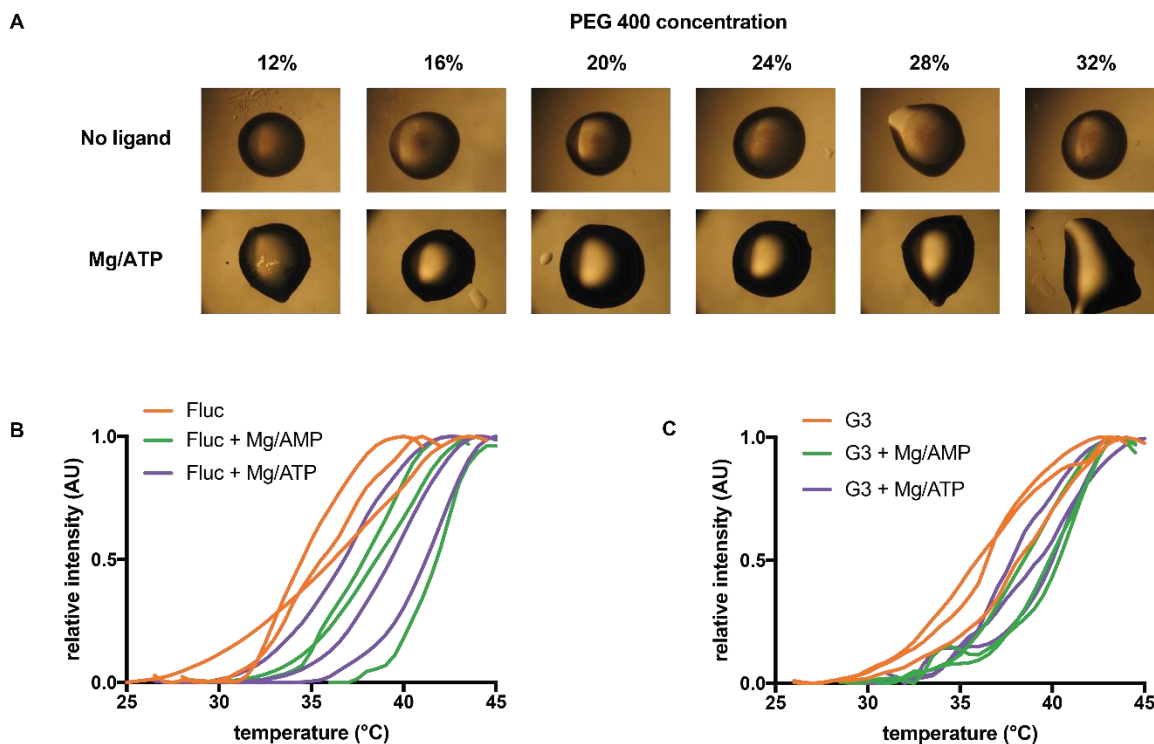
and concentration (Figure 4-5). Precipitation is known to be caused by disordered protein aggregation, a process that can be exacerbated by the presence of PEG [5].



**Figure 4-5. PEG optimization crystallography screens.** Hanging drop vapor diffusion experiments were prepared by mixing 1  $\mu$ L Fluc (15 mg/mL) in storage buffer (25 mM Tris pH 7.8, 200 mM  $(\text{NH}_4)_2\text{SO}_4$ , 1 mM DTT, 1 mM EDTA) with 1  $\mu$ L well condition (100 mM Tris pH 7.5, 300 mM Na/K tartrate, and 2-12% **A)** PEG 6000 or **B)** PEG 400 at 4  $^\circ\text{C}$ . Drops were imaged after 24 h.

#### 4.4 Stabilizing firefly luciferase with ligands

Firefly luciferase has been routinely crystallized in the presence of ligands [2,4]. Unfortunately, attempts to reproduce published conditions were not successful. However, when we supplied ATP and  $\text{Mg}^{2+}$ , known ligands for Fluc, we observed less protein precipitation than in the absence of ligands (Figure 4-6A). We hypothesized that the presence of ligands stabilizes the folded state of the protein. To test this theory, we measured the melt curves of the protein in the presence and absence of ligands using differential scanning fluorimetry. In general, the presence of ligands shifted the melt curves to higher temperatures (Figure 4-6B). Similar effects to Fluc stability have been observed with inhibitors and ATP [6].



**Figure 4-6. Ligands solubilize and stabilize luciferase.** **A)** Hanging drop vapor diffusion experiments were prepared by mixing 1  $\mu\text{L}$  protein (15 mg/mL) in storage buffer (50 mM  $\text{Na}_2\text{HPO}_4$ , 50 mM NaCl) containing 0-27 mM  $\text{MgCl}_2$  and 0-13 mM ATP with 1  $\mu\text{L}$  well condition (50 mM  $\text{Na}_2\text{HPO}_4$ , 50 mM NaCl, and 12-32% PEG 400) at 4°C. Drops were imaged after 24 hours. **B)** and **C)** Protein denaturation was monitored as a function of temperature via differential scanning fluorimetry. Luciferases (1 mg/mL) in storage buffer (25 mM Tris pH 7.8, 200 mM  $(\text{NH}_4)_2\text{SO}_4$ , 1 mM DTT, 1 mM EDTA) were incubated with AMP or ATP (1 mM) and  $\text{MgCl}$  (1 mM) and 0.2% v/v SYPRO dye. Temperatures were sampled from 25-95 °C and fluorescence was measured. Three replicate melt curves are plotted for each unique condition.

#### 4.5 Screening crystallography conditions with ligands

When bound to ATP and  $\text{Mg}^{2+}$ , Fluc adopts a catalytic conformation relevant to the adenylate forming step of the bioluminescence reaction [2]. This function is shared among adenylate-forming enzymes, including acyl-CoA synthases (ACS) [7-10] and the adenylation domain of non-ribosomal peptide synthases [11]. Catalytic conformations of Fluc [2] and homologous enzymes [9,12-13] in the presence of ligands have been extensively studied through crystallography. We reasoned that the luciferase mutants of interest may be crystallized in the adenylating conformation since orthogonal luciferases

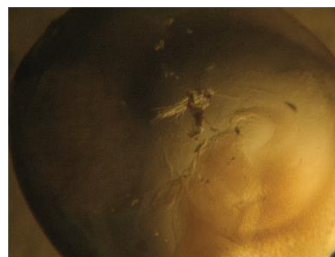
conserve residues involved in adenylation catalysis [14]. We chose to screen crystallization conditions for mutants 37 and G3 based on their utility as orthogonal luciferases.

Several factors can influence protein crystallization such as precipitant identity and concentration [5]. Thus, we screened a wide variety of precipitants, ionic strengths and pH using commercial screens. We were pleased to find several hits for mutant luciferases 37 and 213 (Table 4-1). Many of these hits reflect published conditions. We are currently in the process of reproducing these conditions to produce crystals for X-ray diffraction experiments.

**Table 4-1. Crystal screens with ligands. A)** Hanging drop vapor diffusion experiments were prepared by mixing enzyme and screening condition at a ratio of 1:2, 1:1 and 2:1 by volume. 50-100 nL Fluc (15 mg/mL) in storage buffer (25 mM Tris pH 7.8, 1 mM DTT, 1 mM EDTA, 4.8 mM ATP, 10 mM MgCl<sub>2</sub>, 1 mM D-luciferin) was mixed with 50-100 nL screening condition to a total volume of 150 nL. Crystals appeared within one week. **B)** (Left picture) Representative crystals from coordinate B3 from screen “Crystal” in the presence of mutant 37. (Right picture) Representative crystals from coordinate C10 from screen “PEG/Ion” (not shown in table 1A). Hanging drop vapor diffusion experiments were prepared by mixing enzyme and screening condition at a ratio of 1:1 by volume. 75 nL mutant 37 (15 mg/mL) in storage buffer (25 mM Tris pH 7.8, 200 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1 mM DTT, 1 mM EDTA, 4.8 mM ATP, 10 mM MgCl<sub>2</sub>) was mixed with 75 nL screening condition. Crystals appeared within one week.

A	Screen	Coordinate	Condition
	ProComplex	E5	0.1 M Sodium Chloride, 0.1 M Tris pH 8, 8% (w/v) PEG 20000
	ProComplex	H5	0.1 M Tris pH 8, 1.2 M Sodium, Potassium Tartrate
	ProComplex	H8	0.05 M Calcium acetate, 0.1 M Sodium cacodylate pH 6, 25% (v/v) MPD
	Index	B4	0.1M Tris pH 8.5, 0.3 M Magnesium formate dihydrate
	Index	B7	pH 8.2, 0.056 M Sodium phosphate monobasic monohydrate, 1.344 M Potassium phosphate dibasic
	Index	C6	0.1 M Sodium chloride, 0.1 M BIS-Tris pH 6.5, 1.5 M Ammonium sulfate
	Index	F2	0.2 M Trimethylamine N-oxide dihydrate, 0.1 M Tris pH 8.5, 20% w/v Polyethylene glycol monomethyl ether 2,000
	PACT	A3	0.1 M SPG buffer pH 6, 25% (w/v) PEG 1500
	PACT	A4	0.1 M SPG buffer pH 7, 25% (w/v) PEG 1500
	PACT	A5	0.1 M SPG buffer pH 8, 25% (w/v) PEG 1500
	PACT	A6	0.1 M SPG buffer pH 9, 25% (w/v) PEG 1500
	PACT	A10	0.2 M Magnesium chloride, 0.1 M Sodium acetate pH 5, 20% (w/v) PEG 6000
	JCSG+	A7	0.1 M CHES pH 9.5, 20% (w/v) PEG 8000
	JCSG+	B3	0.1 M Bicine pH 8.5, 20% (w/v) PEG 6000
	JCSG+	B4	0.1 M HEPES pH 7.5 10% (w/v) PEG 8000, 8% (v/v) Ethylene glycol
	JCSG+	D6	0.2 M Magnesium chloride, 0.1 M Tris pH 8.5, 20% (w/v) PEG 8000
	JCSG+	D7	0.2 M Lithium sulfate, 0.1 M Tris pH 8.5, 40% (v/v) PEG 400
	JCSG+	E8	1 M di-Ammonium phosphate, 0.1 M Sodium acetate pH 4.5
	Crystal	A6	0.2 M Magnesium chloride hexahydrate, 0.1 M Tris hydrochloride pH 8.5, 30% (w/v) Polyethylene glycol 4,000
	Crystal	B3	0.2 M Ammonium sulfate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 30% (w/v) polyethylene glycol 8000
	Crystal	B11	0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES sodium pH 7.5, 30% (w/v) Polyethylene glycol 400
	Crystal	E2	0.5 M Sodium chloride, 0.01 M Magnesium chloride hexahydrate, 0.01 M Hexadecyltrimethylammonium bromide

**B**



## 4.6 Conclusion

In conclusion, crystallization and characterization of luciferases will provide an increased structural understanding of engineered luciferase-luciferin pairs. The fidelity of purified luciferases was verified by SDS PAGE analysis, bioluminescence and LC/MS, suggesting that they were good candidates for crystallography experiments. However, when the luciferases were subjected to published crystallization procedures, only protein precipitation was observed. By incubating natural ligands with the luciferase, fewer precipitates were observed, suggesting disordered protein aggregation is inhibited by the presence of ligands. We screened crystallization conditions for mutant luciferases in the presence of ligands, resulting in several promising “hits”. Optimization of these conditions are ongoing.

## 4.7 Materials and methods

### 4.7a Cloning *luc2* gene into pGEX-6p vector

The *luc2* gene was amplified in Q5 buffer containing 1 mM dNTPs, Q5 polymerase, 0.3 mM primers (BZ\_Luc2pGEX\_insertfor and BZ\_Luc2pGEX\_insertfor (primer table)) and pET28a-luc2 template (0.3 ng/μL) using the following conditions (94 °C 2 min, 25 cycles of 98 °C 15s, 68 °C 15 s, 72 °C 60 s, followed by a 72°C extension for 10 min). Melting, annealing and elongation steps were repeated for 34 cycles.

The pGEX-6p vector was linearized in Q5 buffer containing 1 mM dNTPs, Q5 polymerase, 0.3 mM primers (pGEX\_vector\_rev and pGEX\_vector\_for (primer table)) and pGEX-6p template (0.3 ng/μL) using the following conditions (94 °C 2 min, 25 cycles of 98 °C 15 s, 72.8 °C 15 s, 72 °C 140 s, followed by a 72°C extension for 10



min). Melting, annealing and elongation steps were repeated for 22 cycles. The linear vector was then digested with DpnI for 1 h at 37 °C, and the products were purified via a PCR cleanup kit (Qiagen). The resulting linear vector (50 ng) was combined in a 1:2 (vector:insert) ratio with library insert DNA described above. The insert and vector were added to a Gibson master mix (New England BioLabs) and assembled at 50 °C for 20 min. A portion of the assembled product (2.5 µL) was utilized directly for transformation into Top10 chemically competent *E. coli* cells. Transformants were plated on agar plates containing ampicillin antibiotic (40 µg/mL) and saved for further analysis.

#### **4.7b Primer list**

**Luc\_pGEX\_vectorrev\_ext**

GCCCTTCTTAATGTTTTTGGCATCTTCCATGGATCCCAGGGGCCCTG

**Luc\_pGEX\_vectorfor\_ext**

GCCAAGAGGGCGGCAAGTCTAAATTATAATAAGTCGACTCGAGCGGCCGC

**pGEX\_vector\_rev**

GGATCCCAGGGGCCCTG

**pGEX\_vector\_for**

GTCGACTCGAGCGGCCGC

#### **4.7c Recombinant protein expression and purification**

Luciferase constructs bearing an N-terminal His tag were expressed and purified as described in section 2.8b. Following nickel affinity chromatography, purified luciferases were treated with TEV protease bearing an N-terminal His tag (Waugh lab ref. [15]) (10 mol%) and simultaneously dialyzed into 25 mM Tris-HCl (pH 7.8) overnight. TEV protease was removed from cleaved luciferase by binding to nickel resin and eluting with dialysis buffer. The resulting eluent containing cleaved luciferase was dialyzed into 25 mM Tris-HCl buffer (pH 7.8) containing 1 mM EDTA and 1 mM DTT. The resulting solution was concentrated to 15 mg/mL protein and used in crystallization experiments.

Luciferase constructs bearing an N-terminal GST tag were expressed as described in section 2.8b. Following expression in *E. coli*, the luciferase constructs were purified using Glutathione Sepharose 4B GST-tagged protein purification resin (GE Healthcare). GST tags were cleaved by PreScission Protease (GE) on resin according to the manufacturer's protocol. Cleaved luciferase was eluted in PreScission cleavage buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT), pH 7.5). The eluent was dialyzed into 25 mM Tris-HCl buffer (pH 7.8) containing 1 mM EDTA, 200 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 1 mM DTT and concentrated to 15 mg/mL protein concentration. The resulting solution was used in crystallization experiments.

#### **4.7d LC/MS analysis of purified luciferases**

Purified luciferases (100 ng) were passed through a C4 HPLC column eluting with 0.1% TFA/MeCN and elutions were analyzed using a Waters Xevo G2-XS QToF mass spectrometer.

#### **4.7e SDS PAGE analysis of purified luciferases**

Purified luciferases (2 µg) and fractions from affinity purification were denatured at 95 °C for 20 min in the presence of 1X loading dye and 1% BME. Denatured samples were loaded on pre-cast 10% SDS PAGE gels comprising 10% acrylamide/bis-acrylamide (37.5:1), 0.4 M Tris-HCl pH 8.8, 1% SDS, 0.001% TEMED, 0.003% ammonium persulfate in running buffer (25 mM Tris-HCl pH 8.3, 192 mM glycine, 0.1% SDS). Gel electrophoresis was performed at 210 V.

#### **4.7f Light emission assays with recombinant luciferase**

Bioluminescence assays with recombinant luciferases were performed as described in section 2.8e.

#### **4.7g PEG optimization screens**

Hanging drop vapor diffusion experiments [4] were prepared by mixing 1  $\mu$ L Fluc (15 mg/mL) in storage buffer (25 mM Tris pH 7.8, 200 mM  $(\text{NH}_4)_2\text{SO}_4$ , 1 mM DTT, 1 mM EDTA) with 1  $\mu$ L well condition (100 mM Tris pH 7.5, 300 mM sodium/potassium tartrate, and 2-12% PEG 6000 or PEG 400) at 4°C. Drops were imaged after 24 h.

#### **4.7h Mg/ATP optimization screens**

Hanging drop vapor diffusion experiments were prepared by mixing 1  $\mu$ L protein (15 mg/mL) in storage buffer (50 mM  $\text{Na}_2\text{HPO}_4$ , 50 mM NaCl) containing 0-27 mM  $\text{MgCl}_2$  and 0-13 mM ATP with 1  $\mu$ L well condition (50 mM  $\text{Na}_2\text{HPO}_4$ , 50 mM NaCl, and 12-32% PEG 400) at 4°C. Drops were imaged after 24 h.

#### **4.7i Nanoliter reagent screening**

Reagent mixing was performed by mosquito® crystal liquid handler (ttplabtech). Hanging drop vapor diffusion experiments were prepared by mixing enzyme and screening condition at a ratio of 1:2, 1:1 and 2:1 by volume. Mutant 37 or G3 (15 mg/mL) in storage buffer (25 mM Tris pH 7.8, 1 mM DTT, 1 mM EDTA, 4.8 mM ATP, 10 mM  $\text{MgCl}_2$ , 1 mM D-luciferin or analog 3.2) was mixed with 50-100 nL screening condition at ratios 1:2, 1:1, and 2:1 to a total volume of 150 nL. Crystals appeared within one week.

## 2.7j Differential scanning fluorimetry

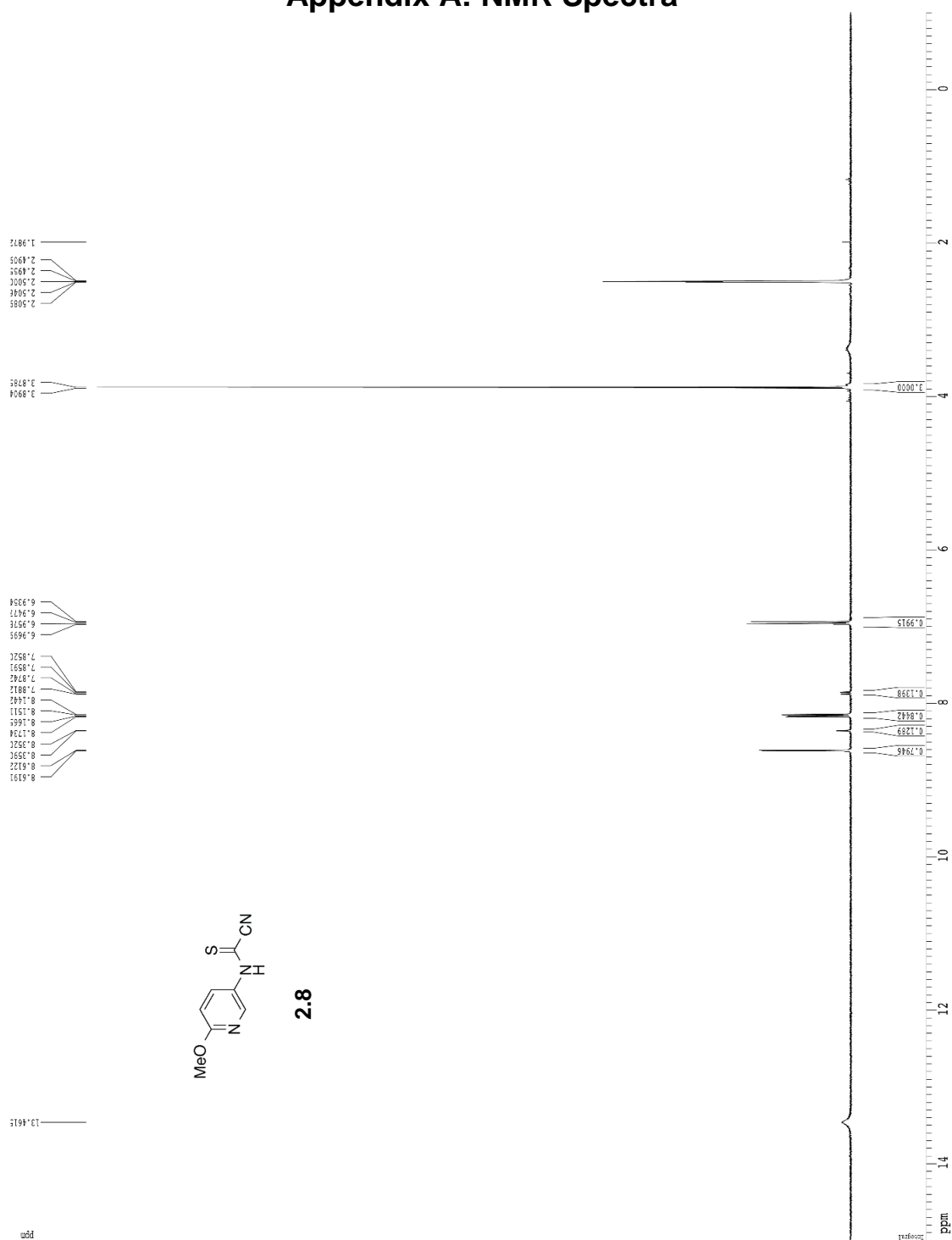
Differential scanning fluorimetry experiments were performed as described in section 2.9q.

## References

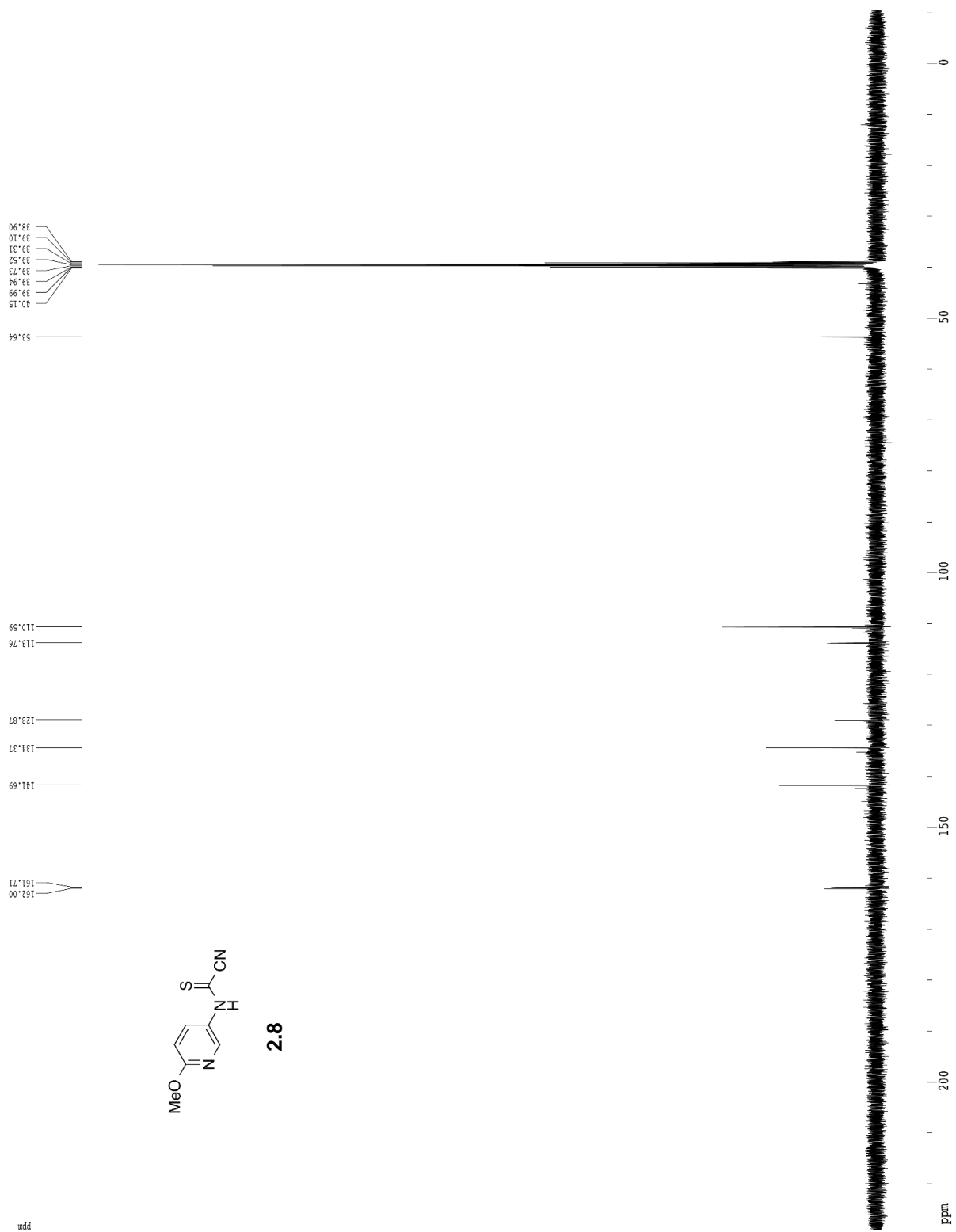
1. Branchini, B. R.; Southworth, T. L.; Fontaine, D. M.; Murtiashaw, M. H.; McGurk, A.; Talukder, M. H.; Qureshi, R.; Yetil, D.; Sundlov, J. A.; Gulick, A. M. Cloning of the Orange Light-Producing Luciferase from *Photinus scintillans*-A New Proposal on how Bioluminescence Color is Determined. *Photochem. Photobiol.* **2017**, *93*, 479-485.
2. Nakatsu, T.; Ichiyama, S.; Hiratake, J.; Saldanha, A.; Kobashi, N.; Sakata, K.; Kato, H. Structural basis for the spectral difference in luciferase bioluminescence. *Nature* **2006**, *440*, 372-376.
3. Zhang, B. S.; Jones, K. A.; McCutcheon, D. C.; Prescher, J. A. Pyridone Luciferins and Mutant Luciferases for Bioluminescence Imaging. *Chembiochem* **2018**.
4. Sundlov, J. A.; Fontaine, D. M.; Southworth, T. L.; Branchini, B. R.; Gulick, A. M. Crystal Structure of Firefly Luciferase in a Second Catalytic Conformation Supports a Domain Alternation Mechanism. *Biochemistry* **2012**, *51*, 6493-6495.
5. Van Driessche, A. E. S.; Van Gerven, N.; Bomans, P. H. H.; Joosten, R. R. M.; Friedrich, H.; Gil-Carton, D.; Sommerdijk, N. A. J. M.; Sleutel, M. Molecular nucleation mechanisms and control strategies for crystal polymorph selection. *Nature* **2018**, *556*, 89-94.
6. Thorne, N.; Shen, M.; Lea, W. A.; Simeonov, A.; Lovell, S.; Auld, D. S.; Inglese, J. Firefly Luciferase in Chemical Biology: A Compendium of Inhibitors, Mechanistic Evaluation of Chemotypes, and Suggested Use As a Reporter. *Chem. Biol.* **2012**, *19*, 1060-1072.
7. Mofford, D. M.; Reddy, G. R.; Miller, S. C. Latent luciferase activity in the fruit fly revealed by a synthetic luciferin. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 4443-4448.
8. Mofford, D. M.; Liebmann, K. L.; Sankaran, G. S.; Reddy, G.; Reddy, G. R.; Miller, S. C. Luciferase Activity of Insect Fatty Acyl-CoA Synthetases with Synthetic Luciferins. *ACS Chem. Biol.* **2017**, *12*, 2946-2951.

9. Kochan, G.; Pilka, E. S.; von Delft, F.; Oppermann, U.; Yue, W. W. Structural Snapshots for the Conformation-dependent Catalysis by Human Medium-chain Acyl-coenzyme A Synthetase ACSM2A. *J. Mol. Biol.* **2009**, *388*, 997-1008.
10. Babbitt, P. C.; Kenyon, G. L.; Martin, B. M.; Charest, H.; Slyvestre, M.; Scholten, J. D.; Chang, K. H.; Liang, P. H.; Dunawaymariano, D. Ancestry of the 4-Chlorobenzoate Dehalogenase - Analysis of Amino-Acid-Sequence Identities among Families of Acyl-Adenyl Ligases, Enoyl-CoA Hydratases Isomerases, and Acyl-CoA Thioesterases. *Biochemistry* **1992**, *31*, 5594-5604.
11. Marahiel, M. A.; Stachelhaus, T.; Mootz, H. D. Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* **1997**, *97*, 2651-2673.
12. Reger, A. S.; Wu, R.; Dunaway-Mariano, D.; Gulick, A. M. Structural characterization of a 140 degrees domain movement in the two-step reaction catalyzed by 4-Chlorobenzoate : CoA ligase. *Biochemistry* **2008**, *47*, 8016-8025.
13. Conti, E.; Stachelhaus, T.; Marahiel, M. A.; Brick, P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* **1997**, *16*, 4174-83.
14. Rathbun, C. M.; Porterfield, W. B.; Jones, K. A.; Sagoe, M. J.; Reyes, M. R.; Hua, C. T.; Prescher, J. A. Parallel Screening for Rapid Identification of Orthogonal Bioluminescent Tools. *ACS Cent. Sci.* **2017**, *3*, 1254-1261.
15. Kapust, R. B.; Tozser, J.; Fox, J. D.; Anderson, D. E.; Cherry, S.; Copeland, T. D.; Waugh, D. S. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* **2001**, *14*, 993-1000.

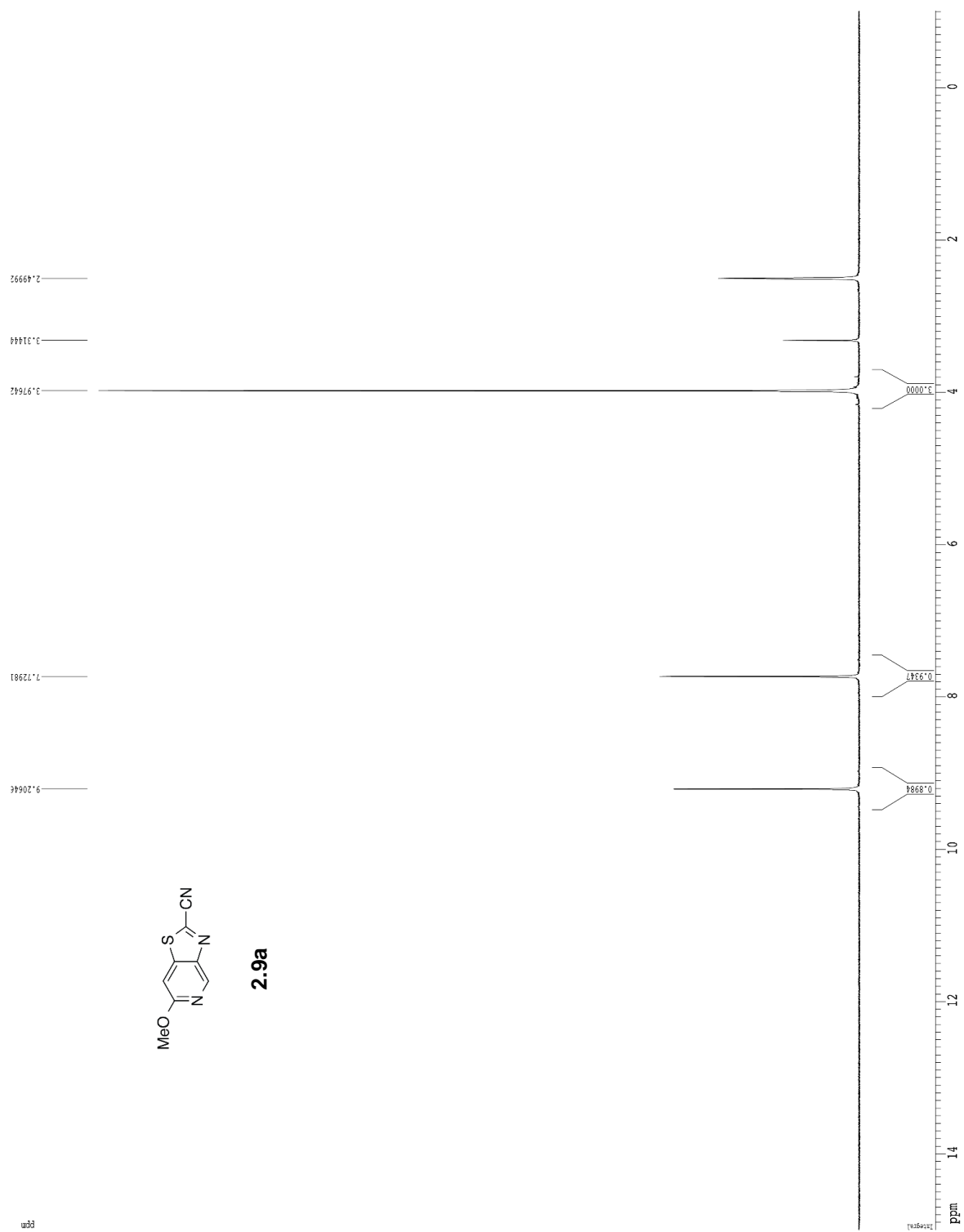
# Appendix A: NMR Spectra



2.8

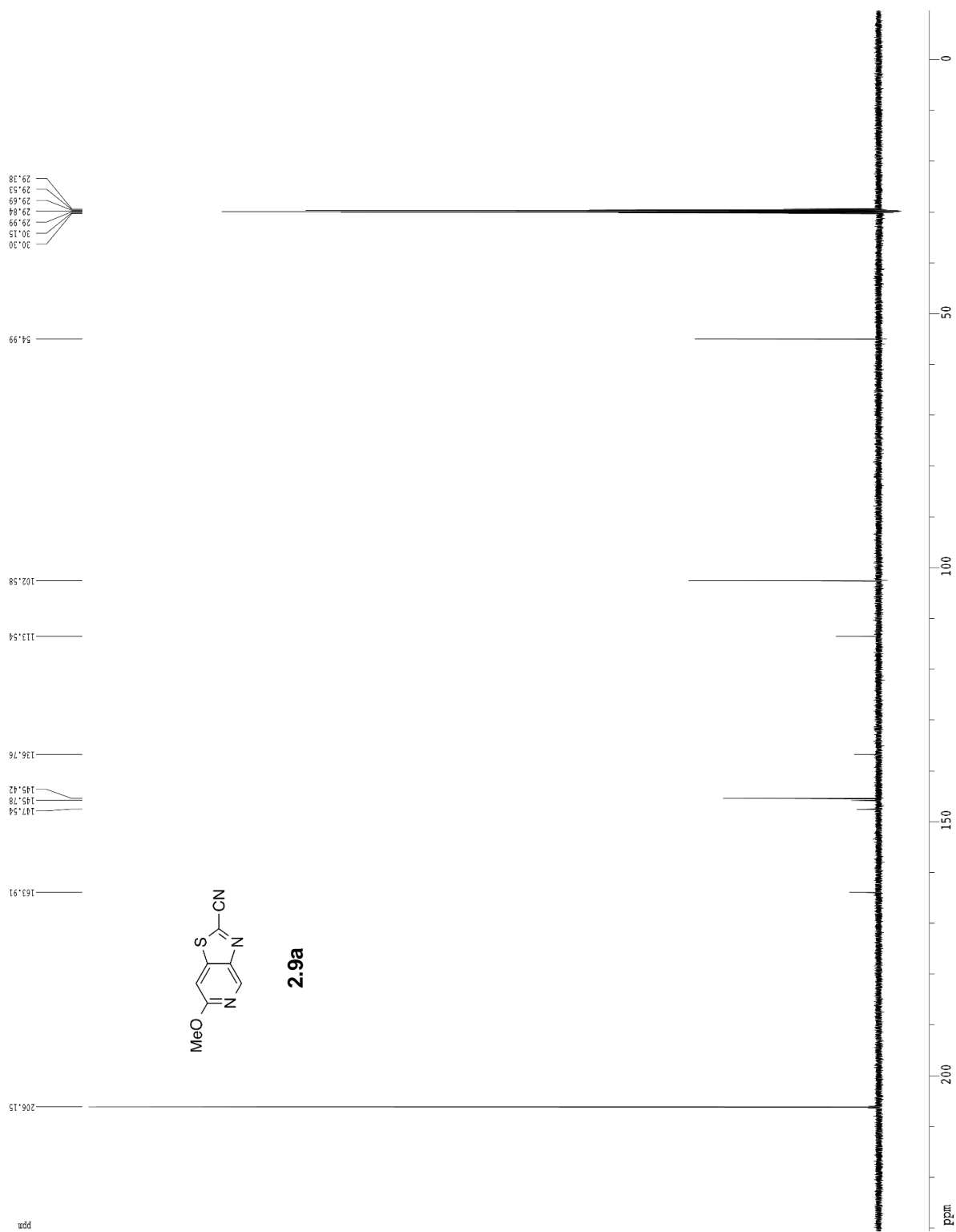


2.9a

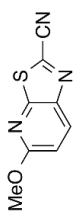
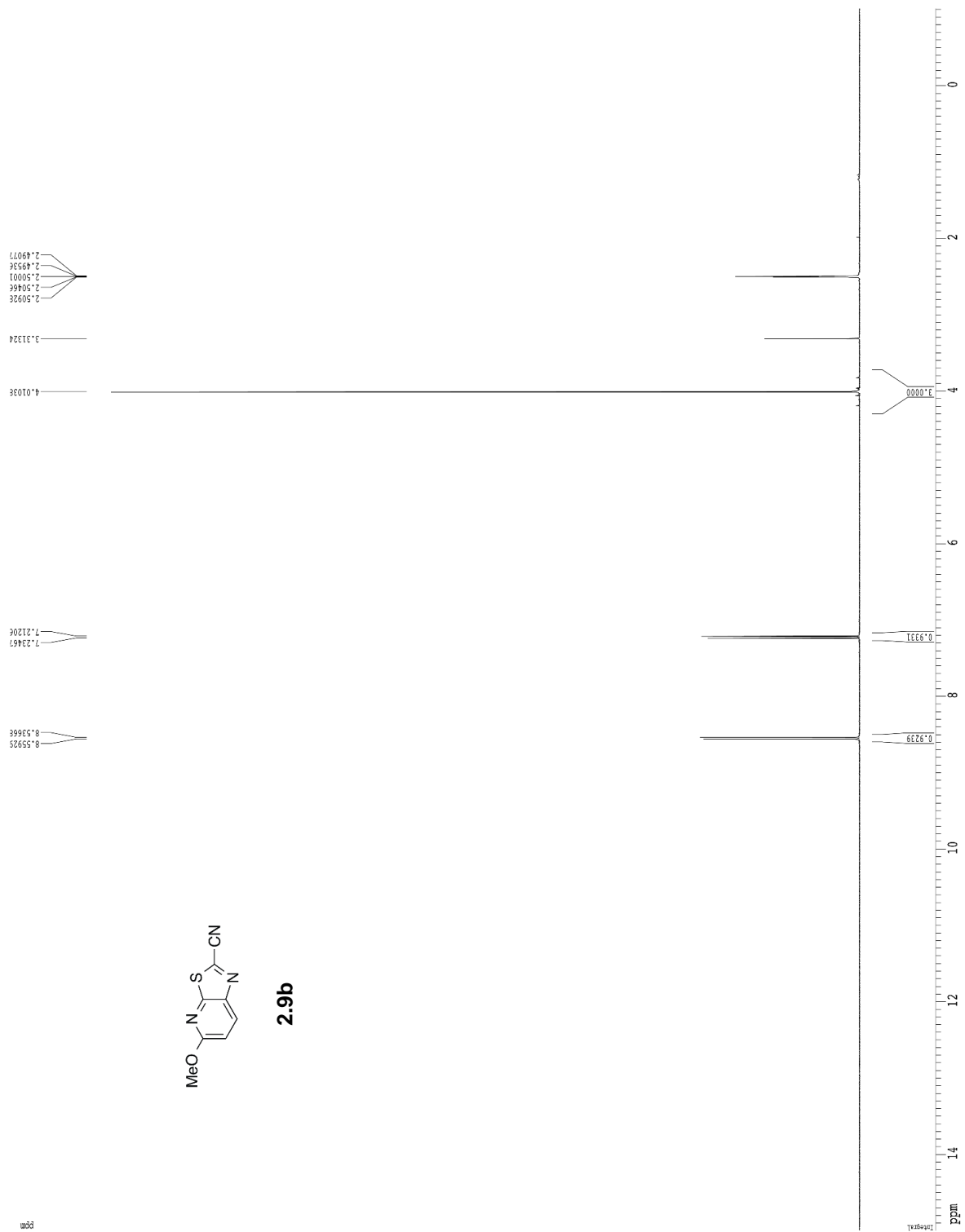




2.9a

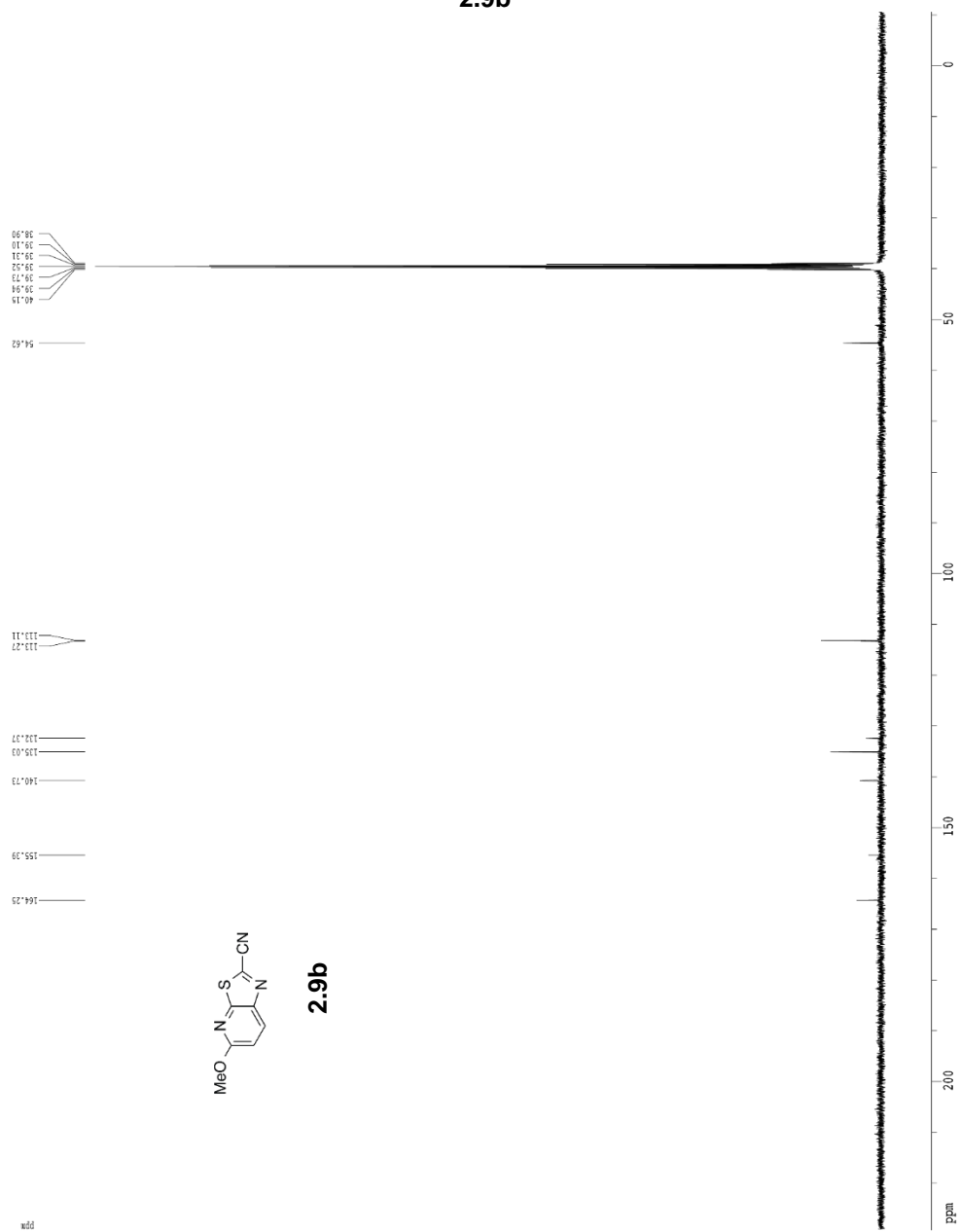


2.9b

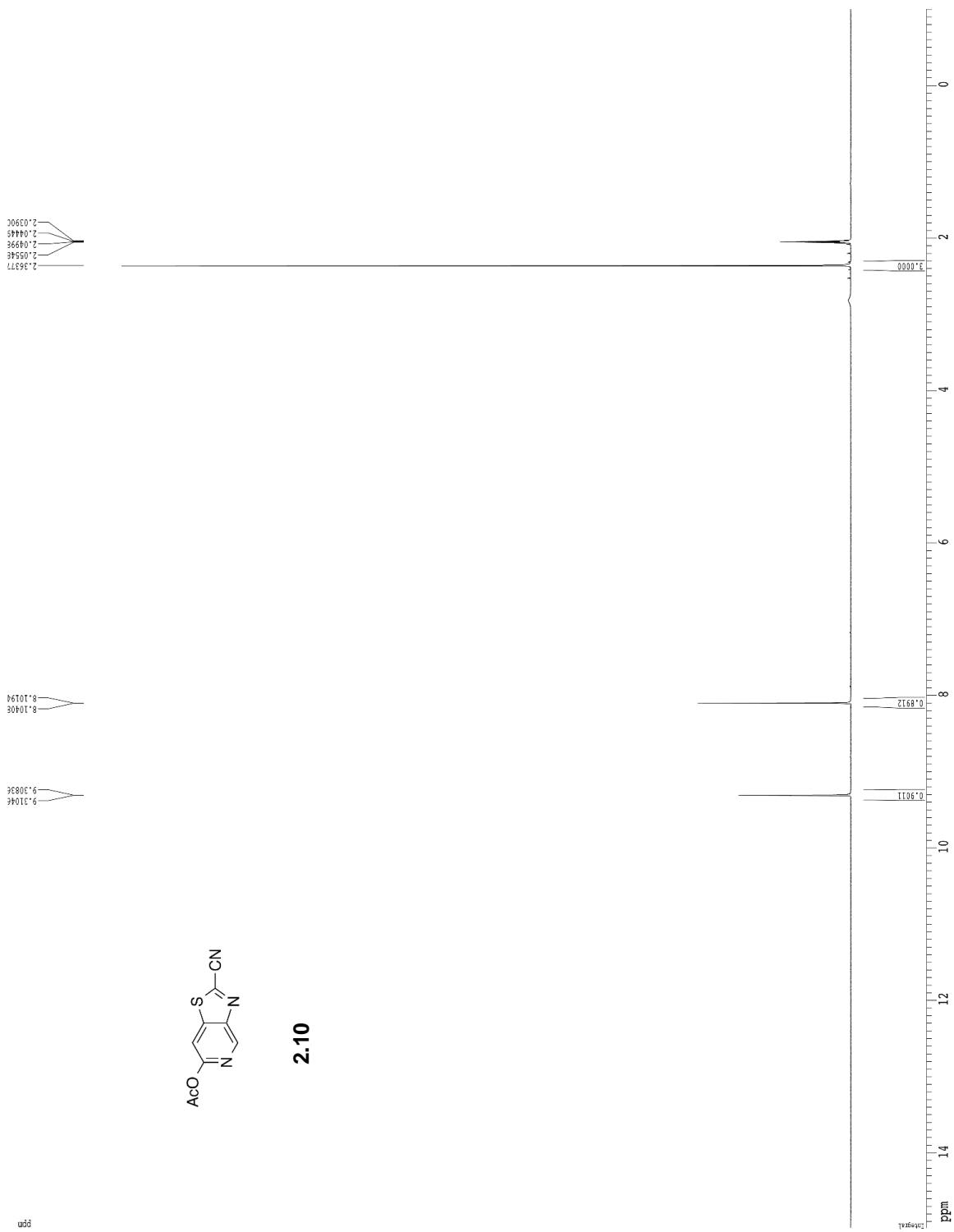


2.9b

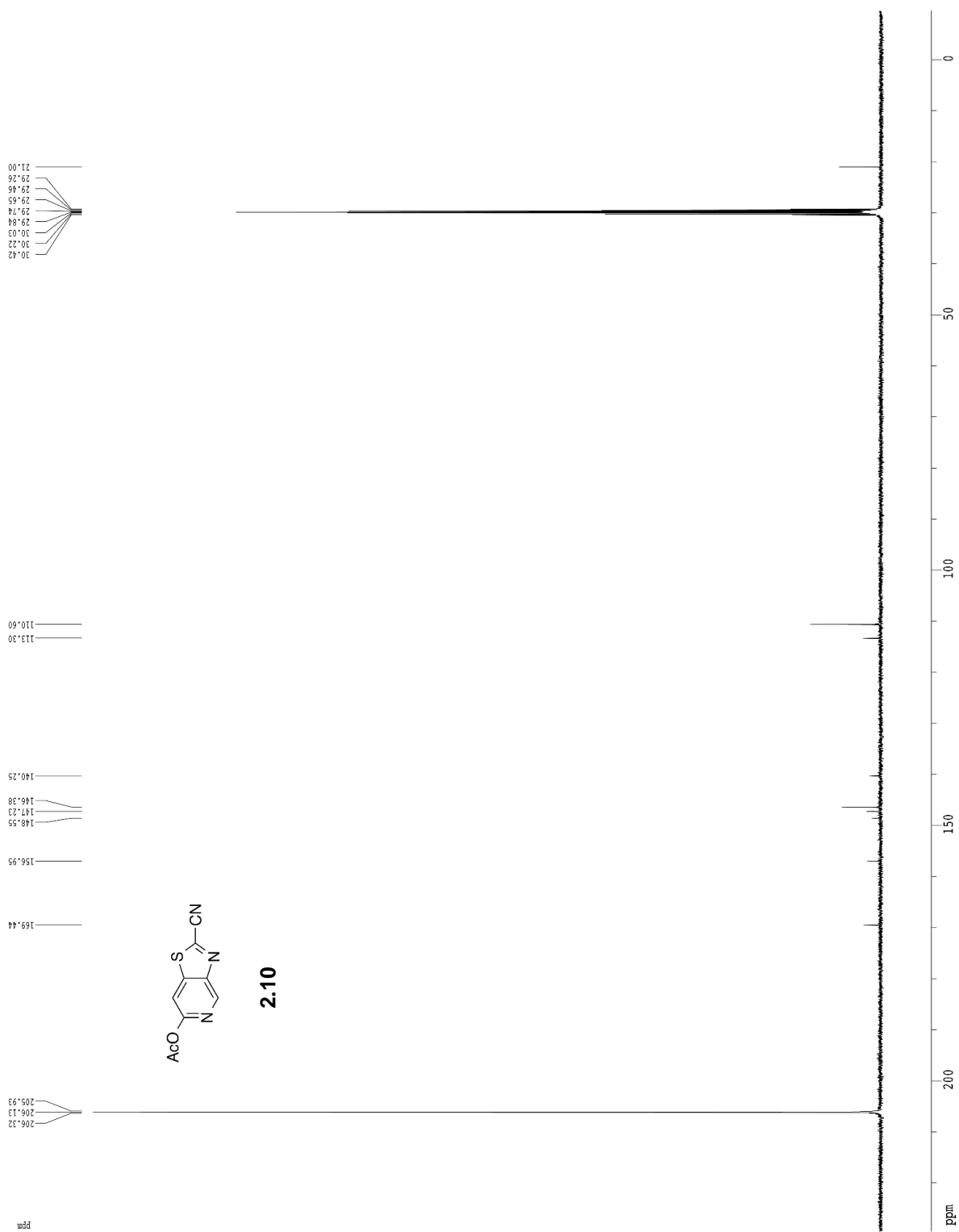
2.9b



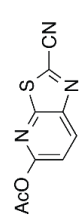
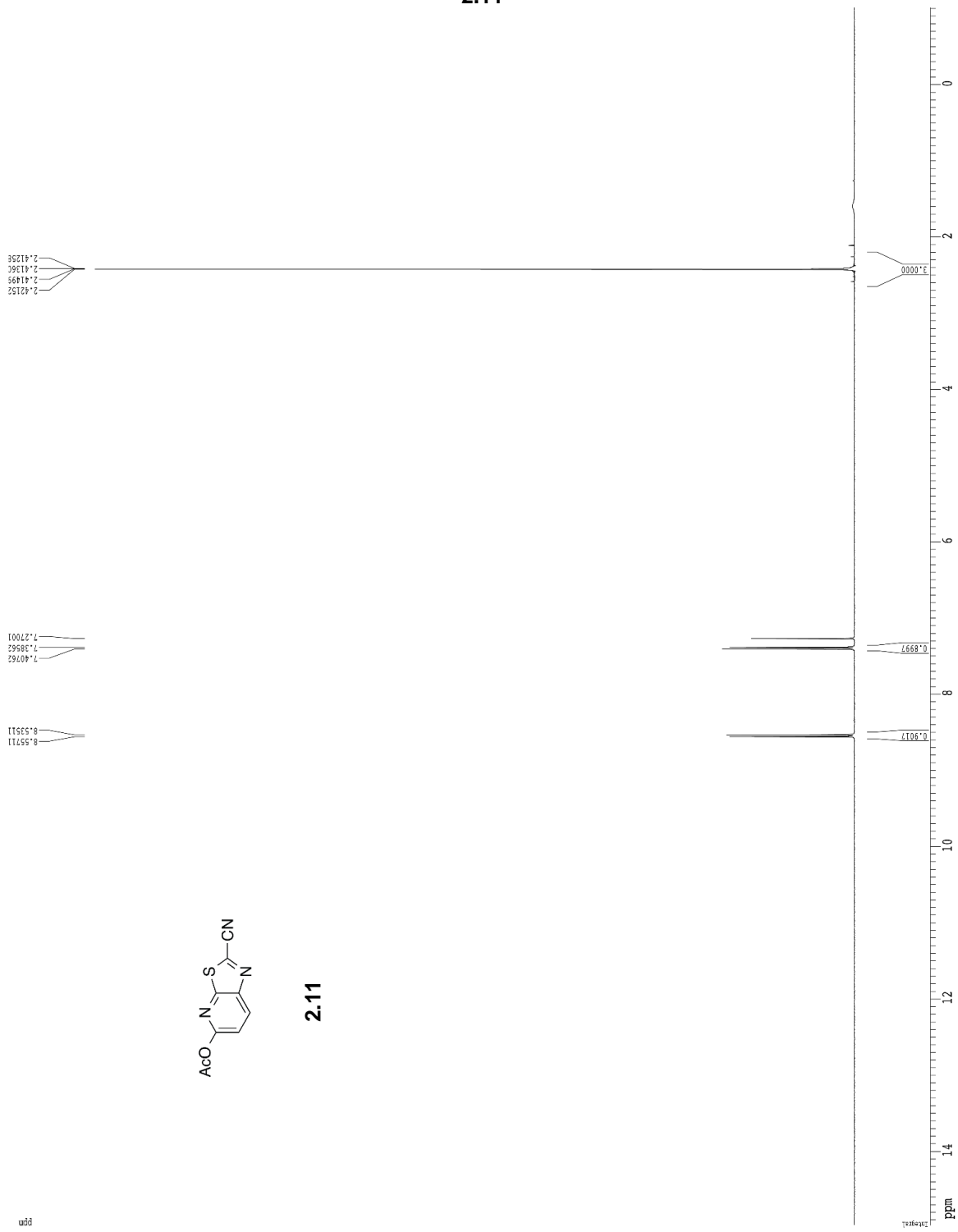
2.10



2.10

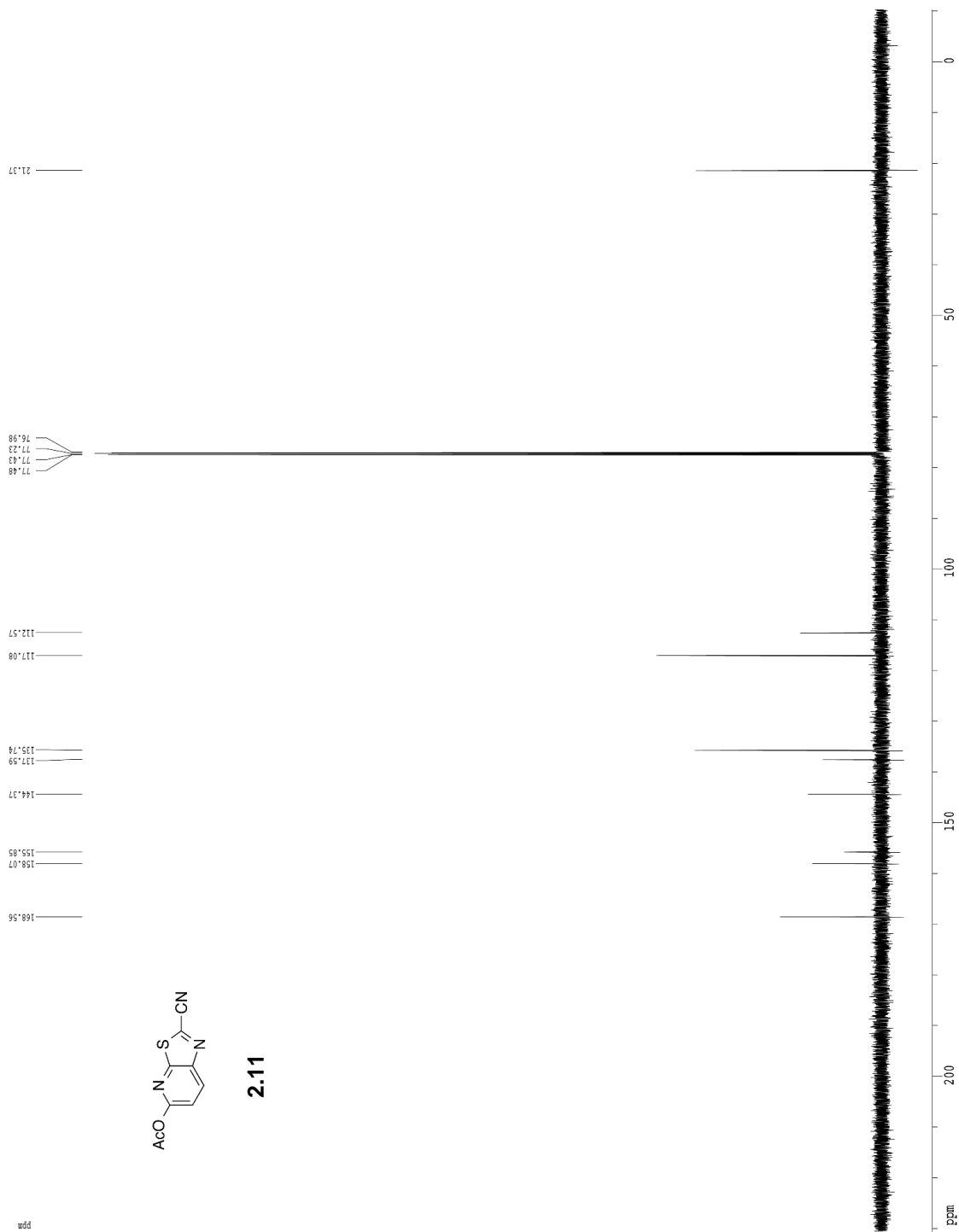


2.11

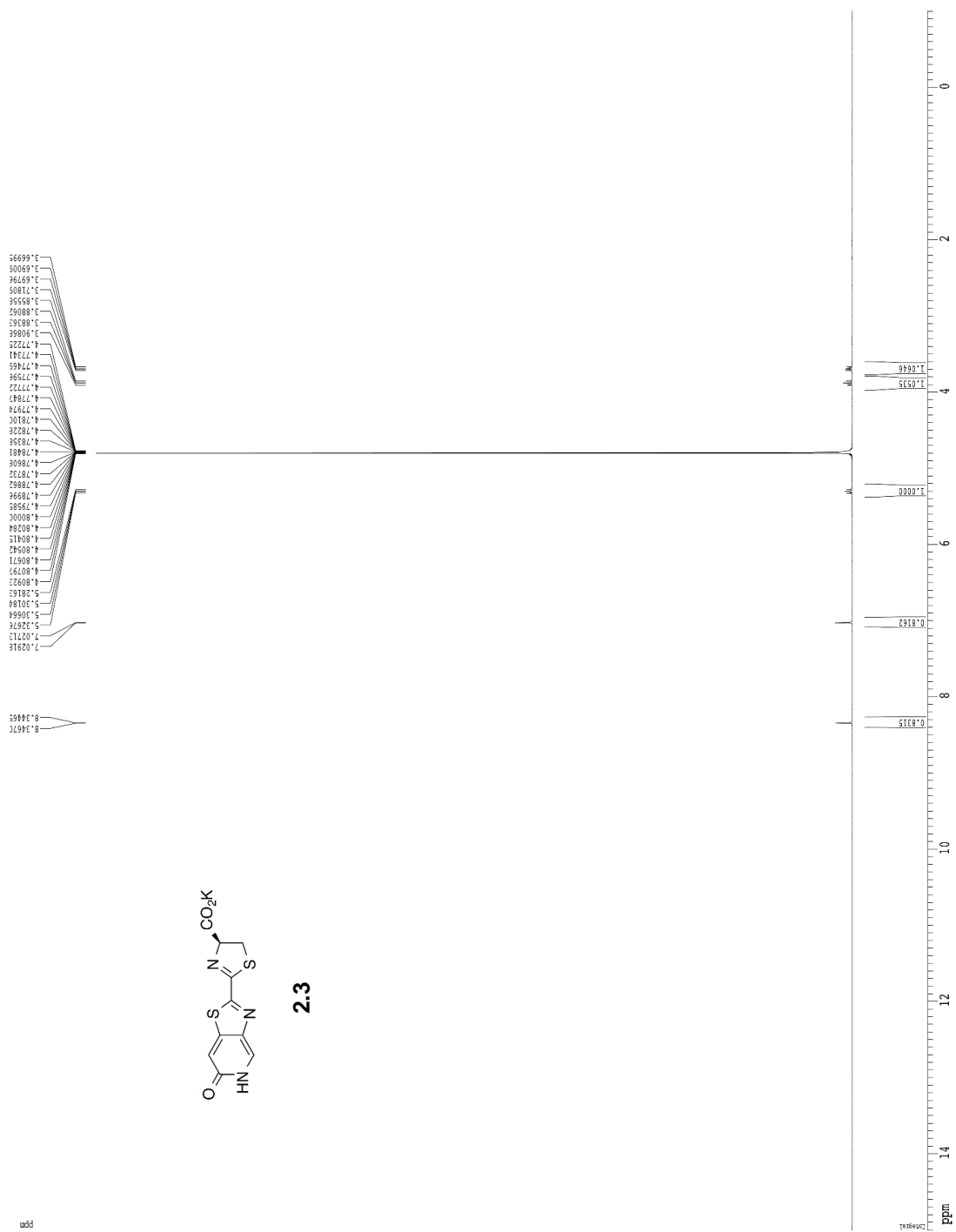


2.11

2.11



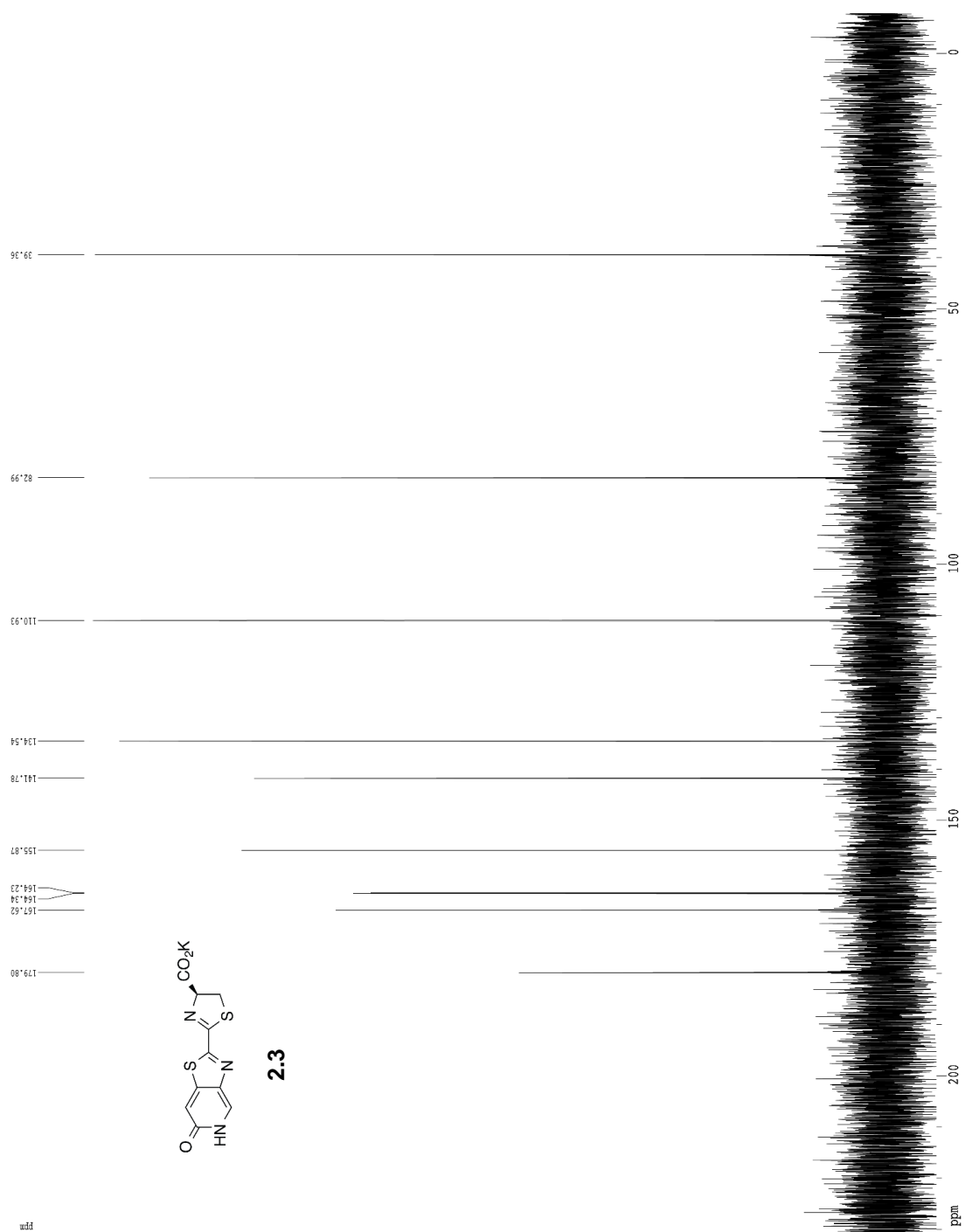
2.3



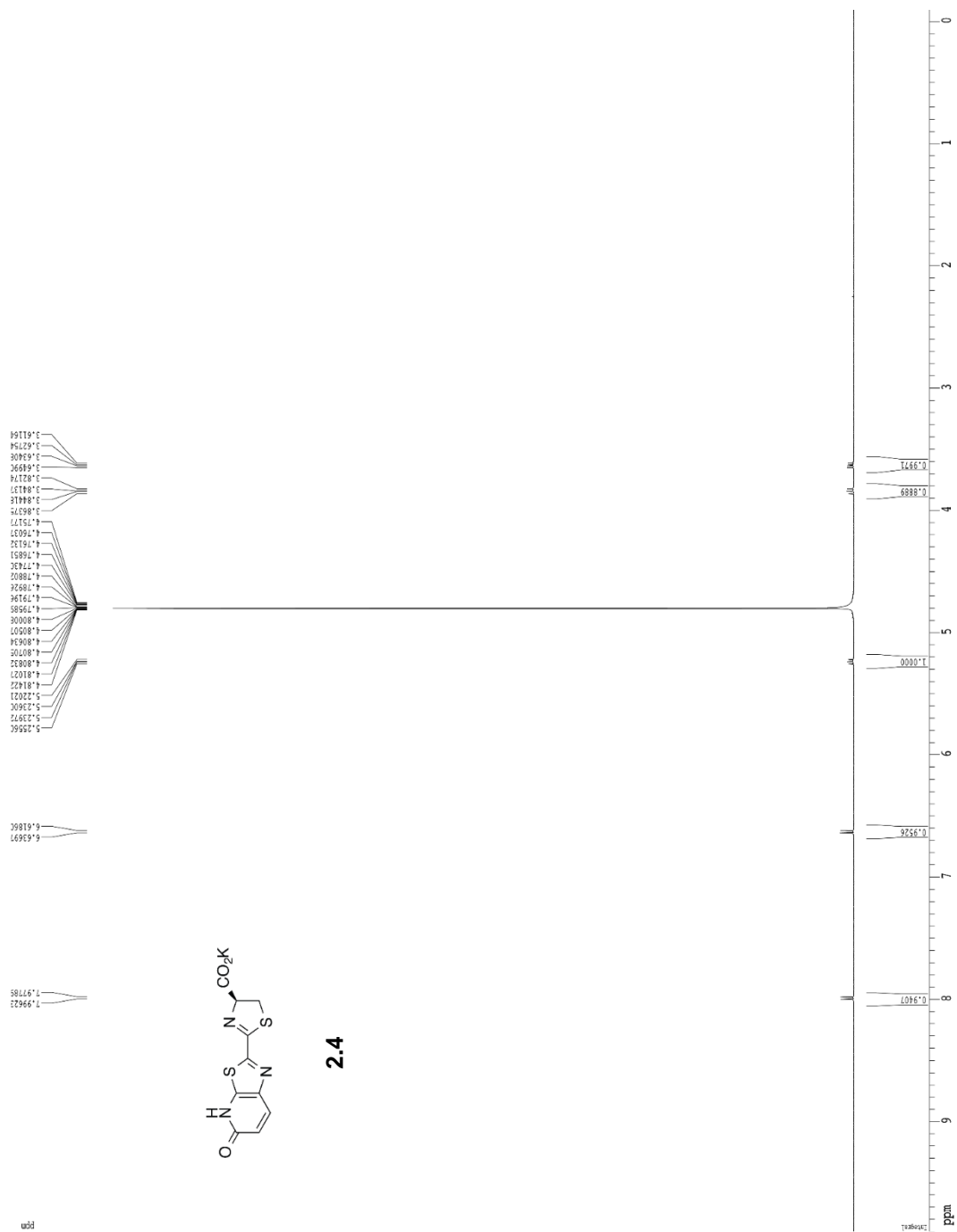
121



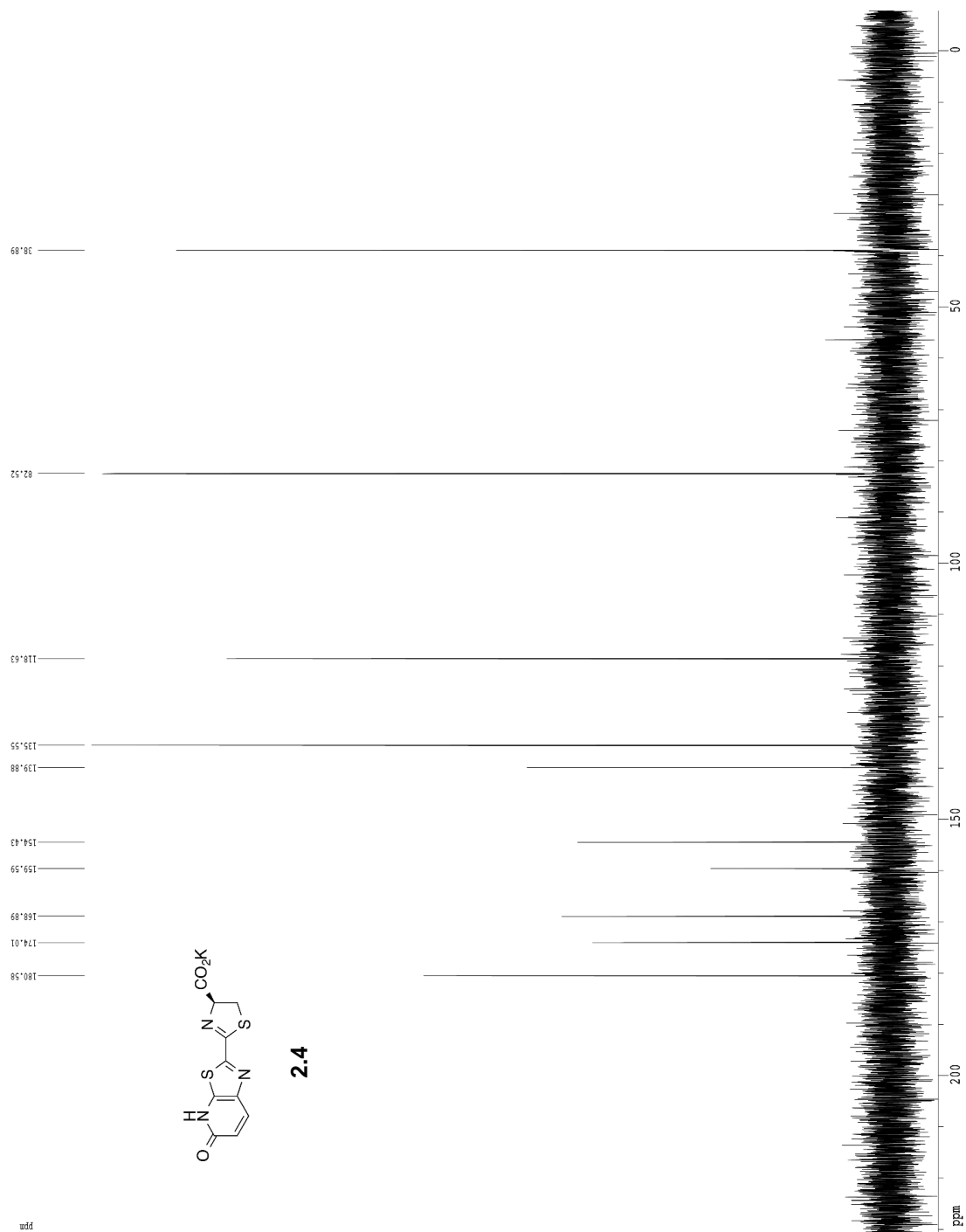
### 2.3



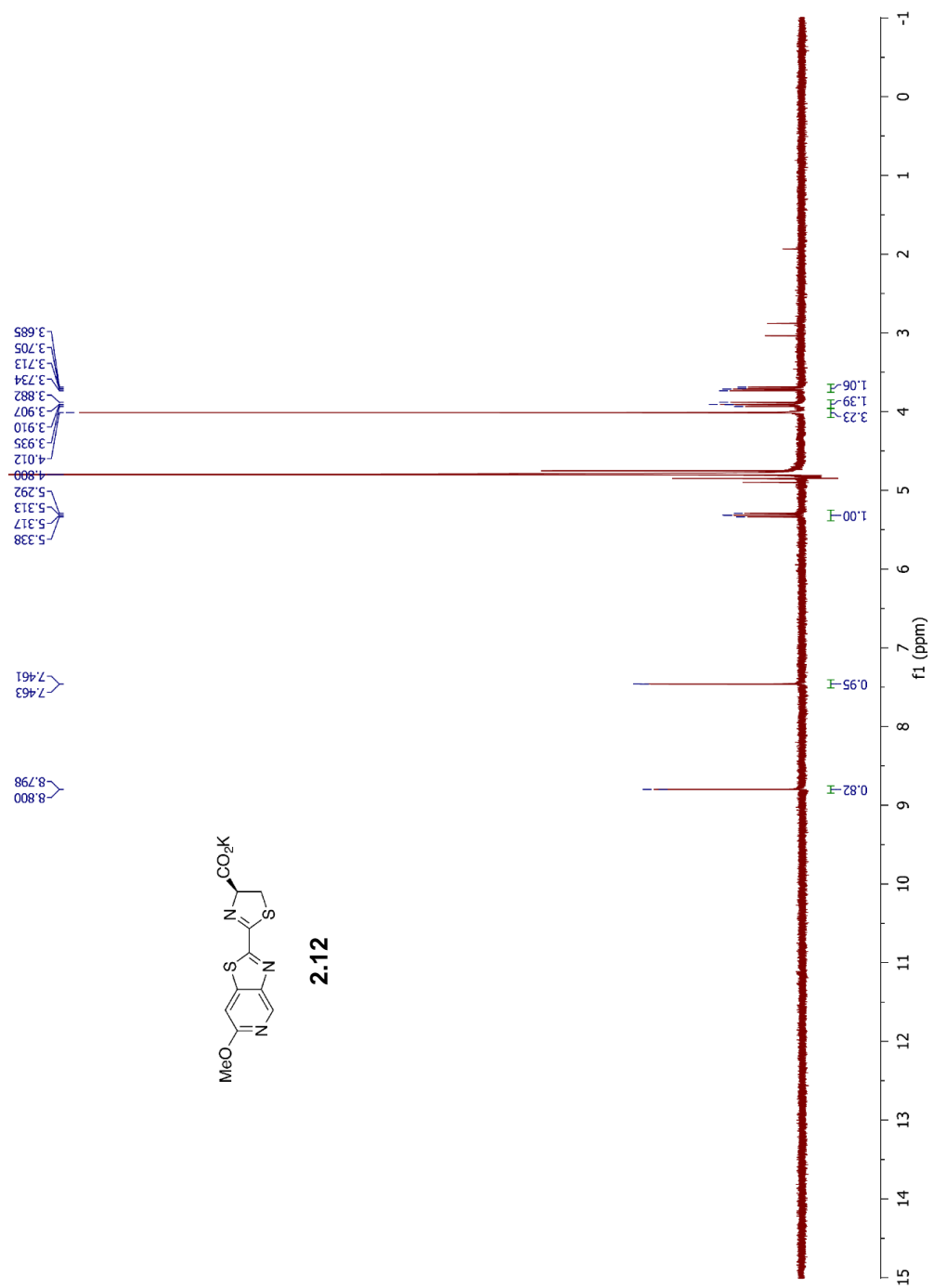
2.4



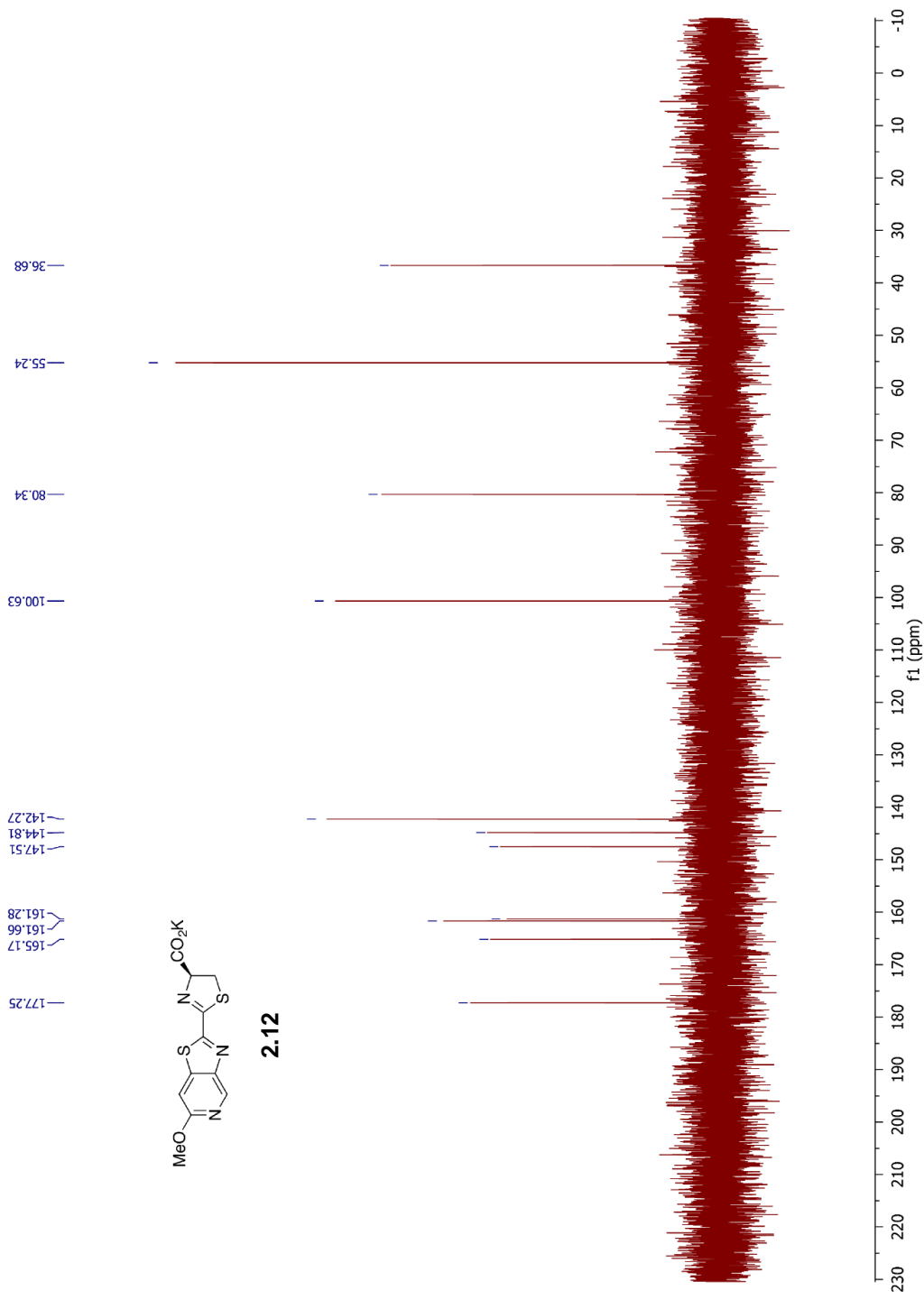
2.4



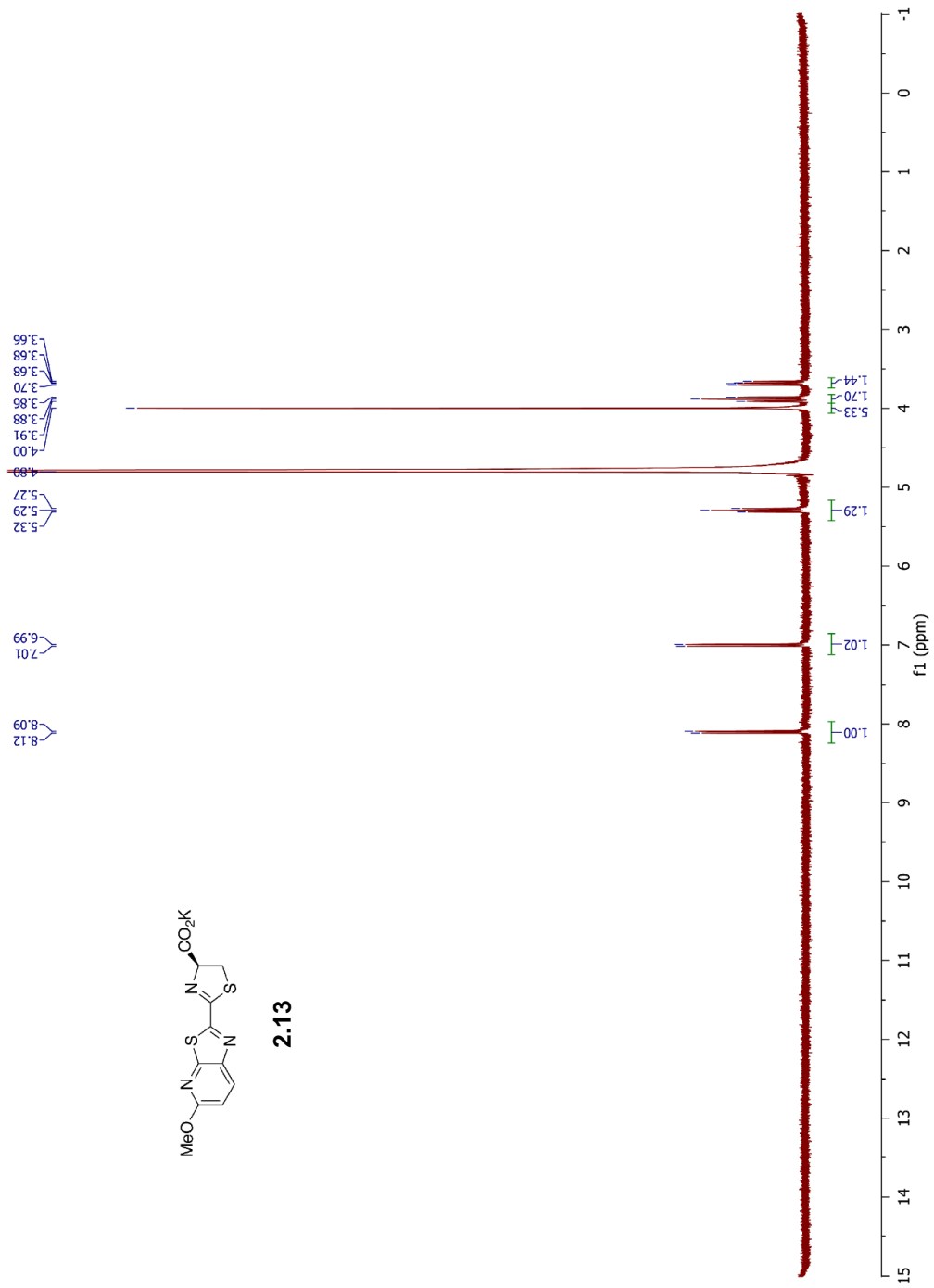
2.12



2.12



2.13



2.13

