

UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Expanding the bioluminescent tool box for imaging macroscopic cell networks

Permalink

<https://escholarship.org/uc/item/8np7j8dz>

Author

Porterfield, William

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Expanding the bioluminescent tool box for imaging macroscopic cell networks

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Organic Chemistry

by

William Buchanan Porterfield

Dissertation Committee:
Associate Professor Jennifer Prescher, Chair
Professor Gregory Weiss
Professor David Van Vranken

2017

Chapter 1 © Elsevier

Select figures and material from Chapter 2 © American Chemical Society

Chapter 3 © American Chemical Society

All other materials © 2017 William B. Porterfield

DEDICATION

To

My parents, John and Andrea Porterfield

For their life-long support and love

TABLE OF CONTENTS

	Page
LIST OF FIGURES	v
LIST OF TABLES	ix
LIST OF SCHEMES	x
ACKNOWLEDGMENTS	xi
CURRICULUM VITAE	xii
ABSTRACT OF THE DISSERTATION	xv
CHAPTER 1: Tools for visualizing cell-cell 'interactomes'	
1.1 Introduction	1
1.2 Fluorescent probes for visualizing collections of cells in vivo	3
1.3 Engineered fluorescent tools for examining cell-cell contacts	7
1.4 Improved bioluminescent probes for noninvasive imaging	9
1.5 Next-generation bioluminescent probes for imaging cell-cell communication	12
1.6 Objectives of this study	13
References	15
CHAPTER 2: 'Caged' probes for imaging cell-cell contacts	
2.1 Introduction	23
2.2 Design and synthesis of 'caged' luciferins	26
2.3 In vitro characterization of caged luciferins	29
2.4 Examining substrate uncaging in live cells	34
2.5 Towards creating a permanent record of cell-cell contacts with a caged Cre-recombinase activator	42
2.6 Conclusion	45
2.7 Materials and methods	46
References	65
CHAPTER 3: Orthogonal luciferase-luciferin pairs for bioluminescence imaging	
3.1 Introduction	70
3.2 Designing and constructing sterically modified luciferins	73
3.3 Analyzing bioluminescent light emission with modified luciferins	77
3.4 Measuring the light-emitting potential of luciferin analogs	80
3.5 Evolving substrate-specific luciferases	82
3.6 Analyzing the origins of orthogonality	88
3.7 Cellular imaging with orthogonal pairs	95

3.8 Conclusions	104
3.9 Materials and methods	105
References	133
CHAPTER 4: A general method for identifying selective luciferase-luciferin pairs	
4.1 Introduction	139
4.2 Expanding the pool of candidate luciferins and luciferases	141
4.3 Screening for orthogonal luciferase-luciferin pairs	144
4.4 Analyzing trends in orthogonal substrate usage	148
4.5 Imaging with orthogonal pairs in vivo	154
4.6 Added diversity improves orthogonality	156
4.7 Conclusions	158
4.8 Methods	160
References	177
CHAPTER 5: Pursuing more orthogonal luciferase-luciferin pairs with improved selectivity	
5.1 Introduction	180
5.2 Expanding the compound base	181
5.3 Identification and characterization of a selective luciferase-luciferin pair	184
5.4 Directed evolution of a 7'-PyridoneLuc selective luciferase	187
5.5 Conclusions	192
5.6 Materials and Methods	193
References	203
APPENDIX A: NMR spectra	206

LIST OF FIGURES

	Page	
Figure 1-1	Multicellular imaging in live animals	5
Figure 1-2	New tools for imaging direct cell-cell interactions	8
Figure 1-3	Expanding the bioluminescent toolbox	11
Figure 1-4	Macroscale detection of microscale cellular interactions	14
Figure 2-1	General strategy for visualizing cell-cell interactions	23
Figure 2-2	Fluorescence spectra of caged luciferins	27
Figure 2-4	Luntr is stable in solution over time	29
Figure 2-5	Luntr reduction characterized by NMR	31
Figure 2-6	Stability of hydroxylamine luciferin	32
Figure 2-7	Reduced Luntr emits light with Fluc	33
Figure 2-8	Flow cytometry analysis of PLE expression	35
Figure 2-9	Cell culture assays with Lucy	36
Figure 2-10	Luntr can be selectively uncaged and locally consumed	37
Figure 2-11	Luntr shows limited fold induction upon media transfer	38
Figure 2-12	Comparison of Luntr activation in cell lines	39
Figure 2-13	Bioluminescence signal induction in Fluc-NTR cells	39
Figure 2-14	Imaging cell-cell interactions with DB7 cells	40
Figure 2-15	Luntr can report on cell-cell contacts	41
Figure 2-16	Bioluminescence microscopy of cell-cell interactions	42
Figure 2-17	Depiction of CreER based cell-cell contact reporter	43
Figure 3-1	Expanding the bioluminescence toolkit	71

Figure 3-2	Docking studies with sterically modified luciferins	74
Figure 3-3	Measuring luciferin light emission	78
Figure 3-4	Sterically modified luciferin light emission with Fluc	79
Figure 3-5	Effect of pH on bioluminescent light output	79
Figure 3-6	Bioluminescence emission spectra	81
Figure 3-7	Fluorescence emission spectra	82
Figure 3-8	Generating and screening mutant luciferases	83
Figure 3-9	Images of luciferin analogs screened on agar plates	84
Figure 3-10	Improved photon output with mutant luciferases	85
Figure 3-11	Sequencing of “hits” from site-directed libraries	85
Figure 3-12	Analyzing orthogonal enzyme-substrate pairs	87
Figure 3-13	Orthogonal substrate usage in mutant or Fluc lysate	89
Figure 3-14	Bioluminescent photon production with luciferase mutants	90
Figure 3-15	Comparative analyses of all analogs with mutants A-C and Fluc	91
Figure 3-16	Bioluminescent light emission with combinatorial enzymes	93
Figure 3-17	Orthogonality of combinatorial enzymes and Fluc	94
Figure 3-18	Flow cytometry analysis of mutant luciferase expression	95
Figure 3-19	Imaging cells with orthogonal luciferase-luciferin pairs	96
Figure 3-20	Luciferase expressing cells exhibit orthogonality	96
Figure 3-21	Cellular imaging with mutant luciferase and D-luciferin	97
Figure 3-22	Cellular imaging with mutant luciferase and luciferin analogs	98
Figure 3-23	Cellular bioluminescent photon production is sustained	99
Figure 3-24	Cellular imaging with orthogonal pairs	100

Figure 3-25	Cellular imaging with orthogonal pairs patterned with stencils	101
Figure 3-26	Orthogonal bioluminescent pairs can be spectrally resolved	102
Figure 3-27	Spectral resolution in vitro and in cells	103
Figure 4-1	General method for accessing luciferase-luciferin pairs	140
Figure 4-2	Brighter mutants may not lead to orthogonality	142
Figure 4-3	Luciferin analogue light emission with Fluc	143
Figure 4-4	Screening for diverse and functional luciferases	144
Figure 4-5	Percent unique “hits” per compound and library	145
Figure 4-6	Unique luciferase-luciferin pairings maintain substrate selectivity	146
Figure 4-7	Reproducibility of algorithm rankings	147
Figure 4-8	Origins of compound selectivity	148
Figure 4-9	Selectivity of luciferase enzymes of luciferin analogues	150
Figure 4-10	Frequency of luciferin analogues in top algorithm rankings	151
Figure 4-11	Enzyme-enzyme pairing frequency	152
Figure 4-12	Compound-residue pairing frequency	153
Figure 4-13	Orthogonality of luciferase-luciferin sets in mammalian cells	154
Figure 4-14	Animal model imaging of top orthogonal sets	155
Figure 4-15	Added diversity improves orthogonality and triple sets	157
Figure 4-16	Reproducibility of top unique triplet sets	158
Figure 5-1	Structures of electronically altered luciferin analogues	181
Figure 5-2	Light emission of PyrroLuc with Fluc	183
Figure 5-3	Light emission of analogues with mutant luciferases	184
Figure 5-4	Purified mutant enzyme light emission with luciferin analogues	186

Figure 5-5	Structure of luciferin analogues in luciferase active site	187
Figure 5-6	Selectivity and light emission over one round of evolution	188
Figure 5-7	Distribution of luciferase library light emission and selectivity	189
Figure 5-8	Compound-residue pairing frequency for analogues of interest	190
Figure 5-9	Location of third generation mutations in luciferase active site	191
Figure 5-10	Selectivity and light emission over rounds of evolution	192

LIST OF TABLES

	Page
Table 3-1 Enzymatic and optical parameters	80
Table 3-2 Biochemical analyses of orthogonal enzyme-substrate pairs	92

LIST OF SCHEMES

	Page
Scheme 2-1 Synthetic route to Lucy and Luntr	27
Scheme 2-2 Synthesis of hydroxylamine luciferin	32
Scheme 2-3 Synthetic route to cylco-NTF	44
Scheme 3-1 Initial synthetic route to C7'-modified luciferin analogues	75
Scheme 3-2 Synthesis of 7'-Methyl luciferin	75
Scheme 3-3 General procedure for synthesis of C7'-modified luciferins	76
Scheme 3-4 Synthesis of 4'-Methyl luciferin	77
Scheme 5-1 Synthetic route to PyrroLuc	182

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. The chemical biology book is one of my favorite reference books due to the detailed molecular discussion of biology.

My partner Anne, for always being there and supporting me through good times and bad.

Friends for making my time in Orange County very enjoyable.

Financial support was provided by the University of California, Irvine, the BEST IGERT program (funded by NSF award DGE-1144901), and the Allergan fellowship in synthetic organic chemistry.

CURRICULUM VITAE

William B. Porterfield

Ph.D. Candidate, Department of Chemistry
University of California, Irvine
Irvine, CA 92697

Telephone: (805) 766-2032
E-mail: wporterf@uci.edu

EDUCATION

University of California, Irvine, Irvine, CA <i>Ph.D. Chemistry</i>	2012-2017
University of San Diego, San Diego, CA <i>B.A. Biochemistry (with honors); Mathematics minor</i> GPA: 3.67	2005-2009

RESEARCH EXPERIENCE

University of California, Irvine, Irvine, CA Doctoral candidate Department of Chemistry Advisor: Prof. Jennifer Prescher Research: Development of novel optical imaging tools to probe multicomponent <i>in vivo</i> systems	2012-2017
Microtracers Inc., San Francisco, CA R&D and Technical Services Associate Research: Development of microengraved magnetic particles for use in animal feed	2012
Pfizer Global Research and Development, Oncology, La Jolla, CA Discovery Chemist Intern Advisor: Neal Sach Research: Construction of small molecule libraries using flow chemistry	2009
The Scripps Research Institute, La Jolla, CA Research Intern Advisors: Prof. Valery Fokin, Dr. Jason Hein Research: Generation of methods for measuring copper ion concentration using "click" chemistry	2009
University of San Diego, San Diego, CA Undergraduate Researcher Advisor: Prof. Debbie Tahmassebi Research: Synthesis of a novel fluorescent dideoxynucleoside	2007-2009

PUBLICATIONS

Jones, K.A.*; **Porterfield, W. B.***; Rathbun C. R.*; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. Orthogonal luciferase-luciferin pairs for bioluminescence imaging. *J. Am. Chem. Soc.* **2017**, *139*, 2351. [* denotes equal contribution]

Porterfield, W. B.; Jones, K. A.; McCutcheon, D.C.; Prescher, J.A. A 'Caged' Luciferin for Imaging Cell–Cell Contacts. *J. Am. Chem. Soc.* **2015**, *137*, 8656-8659.

Porterfield, W. B.; Prescher, J. A. Tools for visualizing cell-cell 'interactomes'. *Curr. Opin. Chem. Biol.* **2015**, *24*, 121-130.

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

Porterfield, W. B.; Tahmassebi, D. Synthesis of a fluorescent 2'3'-dideoxycytosine analog, tCdd. *Biorg. Med. Chem. Lett.* **2009**, *19*, 111-113.

PRESENTATIONS

Porterfield, W.B.; Jones, K. A.; McCutcheon, D.C.; Rathbun, C.M.; Prescher, J.A. "Development of novel luciferin-luciferase pairs for multicomponent imaging" 251st ACS National Meeting, San Diego, CA, March 2016. (presentation)

Porterfield, W.B.; Jones, K. A.; McCutcheon, D.C.; Prescher, J.A. "Visualizing cell-cell interactions with 'caged' luciferins" 248th ACS National Meeting, San Francisco, CA, August 2014. (poster)

Porterfield, W.B.; Tahmassebi, D. C. "Synthesis of a novel, fluorescent 2'3'-dideoxynucleoside" 236th ACS National meeting, Philadelphia, PA August 2008. (poster)

Porterfield, W.B.; Tahmassebi, D. C. "Synthesis of a dideoxy terminated nucleoside for the study of DNA polymerase" 41st Annual ACS Western Regional meeting, San Diego, CA, September 2007. (poster)

TEACHING AND SERVICE

Graduate Student Instructor , University of California, Irvine	
General Chemistry Lab	Spring 2016
Chemical Biology Graduate Course	Winter 2015
Chemical Biology Undergraduate Lab	Winter 2013
Organic Chemistry Lab	Spring 2013
General Chemistry Lecture	Fall 2012
Lab Safety Officer , Prescher lab, UC Irvine	2013-2017
Science Fair Judge and Advisor , Irvine Unified School District	2012-2014
High School Teacher , WorldTeach: Harvard Center for International Development Pohnpei, Micronesia	
High school-level chemistry, math, English	2010-2011
Snowboard Instructor , Squaw Valley Ski Corp. Olympic Valley, CA	2009-2010
Chemistry Club , University of San Diego	
President	2008-2009
Events & Programs Coordinator	2007-2008

HONORS AND AWARDS

Abbvie scholar	2016
Allergan graduate fellowship in synthetic organic chemistry	2015
Recipient of NSF IGERT biophotonics traineeship (UC Irvine)	2013-2015

NSF GRFP honorable mention	2013
Graduate with honors (University of San Diego)	2009
Phi Beta Kappa Member (University of San Diego)	2009-2017
Chemistry Service Award (University of San Diego)	2009
Outstanding Research Award (University of San Diego)	2008
Goldwater Scholar Nominee (University of San Diego)	2008
University of San Diego Trustee Scholarship	2005-2009

ABSTRACT OF THE DISSERTATION

Expanding the bioluminescent tool box for imaging macroscopic cell networks

By

William Buchanan Porterfield

Doctor of Philosophy in Organic Chemistry

University of California, Irvine, 2017

Associate Professor Jennifer Prescher, Chair

Optical imaging strategies have revolutionized our understanding of living systems. Among the most popular techniques for imaging in whole tissues and organisms is bioluminescence. The most widely used bioluminescent system comprises the luciferase enzyme from the firefly (Fluc) and its small molecule substrate, D-luciferin (D-luc). When introduced into cells, these components produce photons that can be captured by sensitive cameras. Since mammalian cells and tissues produce little to no endogenous light, bioluminescence imaging (BLI) is well suited for use in whole organisms. Despite its sensitivity and broad dynamic range, BLI has largely been limited to imaging one cell type at a time. Multicellular networks have been refractory to BLI owing to lack of distinct luciferase-luciferin probes. BLI also lacks adequate spatial resolution to directly visualize cell-to-cell contacts. To address these voids, my thesis work focused on: (1) developing “caged” luciferins to report on cell-cell interactions and (2) creating novel bioluminescent pairs via luciferin analog synthesis and luciferase enzyme engineering. Collectively, the tools developed from my thesis work will provide a better understanding of complex interactions *in vivo*.

Chapter 1: Tools for visualizing cell-cell 'interactomes'

1.1 Introduction

Humans rely on a multitude of cell types and communication networks to carry out specific tasks. Movements are controlled by nerve cell contacts with neighboring neurons and distant muscle tissue; pathogens are cleared by the coordinated actions of dozens of immune cell types. Many aspects of these and other cell networks are understood with exquisite *molecular* detail. Discrete cell adhesion molecules and secreted molecules have been identified for both neurons and immune cells both transmit information via direct cell adhesion and secreted molecules [1-3]. Our understanding of biological networks at the *cellular* level, by contrast, remains incomplete. The number of cell types involved in controlling movements, immune function, and other processes are unknown in most cases. The timing and location of cellular interactions, along with the long-term fates of the cells, are similarly poorly understood. Unraveling these networks of cellular interactions (i.e., "interactomes") is crucial to not only providing a complete picture of organismal biology, but to one day exploit cellular networks for therapeutic gain.

Understanding cellular "interactomes" requires tools that can function over long times and distances in heterogeneous tissues [4-5]. Optical imaging probes are uniquely suited for this purpose. These tools produce nontoxic visible light upon excitation, and when appended to whole cells, their signals can report on cell movements and other functions. Among the most popular optical agents for cell-based imaging are fluorescent proteins and small molecule fluorophores. These tools come in a wide assortment of colors, and combinations of fluorescent probes have been artfully

employed to visualize cells *in vitro* and *in vivo* [6]. Most applications to date, though, have been limited to *microscopic* observations of surface structures or exposed tissues where the required excitation light can be more effectively delivered.

A complementary class of optical agents, bioluminescent proteins (luciferases), are often better suited for imaging at the *macroscopic* scale. These tools do not require incident radiation to produce light. Rather, luciferases emit photons during the chemical oxidation of small molecule substrates (luciferins). The light produced is inherently weak, but since mammalian tissues do not produce large numbers of photons, bioluminescence can provide a sensitive readout on luciferase-labeled cells in intact organisms [5,7]. Indeed, tumor cell proliferation, immune cell homing, and other processes have been examined with this imaging modality [7].

While a popular choice for macroscopic imaging, bioluminescence has been difficult to employ for imaging microscopic features. The low levels of light produced are difficult to capture with conventional microscopes. Additionally, few unique luciferase-luciferin pairs exist, largely limiting the technology to visualizing one cell type or biological feature at a time. Thus, despite decades of optical imaging probe development and applications; there remains a void in our ability to visualize cells across long time and length scales. Recent advances in chemical biology, though, are beginning to address the need for reliable, user-friendly tools to observe collections of cells in intact organisms. These developments and their applications in profiling cellular interactomes are highlighted in the following sections.

1.2 Fluorescent probes for visualizing collections of cells *in vivo*

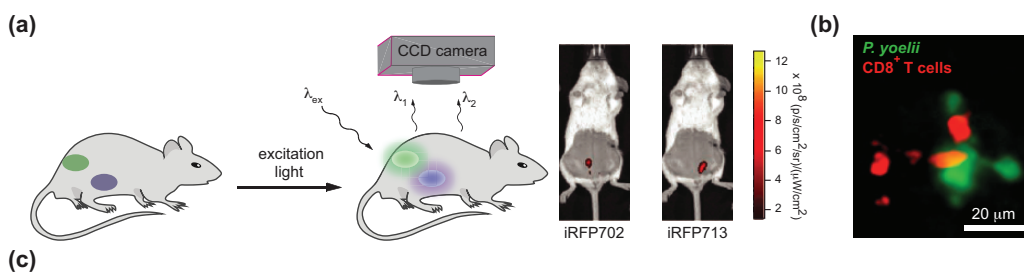
A necessary step in understanding cell-cell “interactomes” involves taking inventory of cells in their native habitats. Fluorescent proteins (FPs) and small molecule fluorophores are well suited for this task. These probes comprise chromophores that produce light upon irradiation. Several colors of FPs and fluorophores are now available, enabling multiple cell types to be simultaneously visualized. FPs can also be encoded in a variety of cell types for long-term, serial tracking [6]. Unfortunately, many of the most common fluorescent probes (e.g., GFP, fluorescein, etc.), are not ideal for macroscopic, deep-tissue imaging in rodent or other opaque models. The required excitation light induces high levels of tissue autofluorescence and is strongly absorbed by blood [8-9]. Thus, fluorescent probes have primarily been employed for imaging at superficial sites or surgically exposed regions [4]. Interference from blood and overlying tissue can be minimized using longer, more tissue-penetrant wavelengths in combination with multi-photon microscopy. This approach has been used to track a variety of cellular networks (labeled with FPs or other fluorophores) in live organisms [10-11]. Recent examples include cell-cell interactions relevant to cell homing [12], immune regulation [13] and pathogen clearance [14].

Optimized fluorescent proteins for cell tracking in vivo. Macroscopic cell imaging is possible using FPs with red-shifted excitation and emission spectra. Wavelengths > ~600 nm readily penetrate mammalian tissues, and identifying FPs with excitation and emission spectra in this range is thus an important goal [8]. The earliest red-emitting FPs were not sufficiently stable or bright for routine use *in vivo* [15-16]. Iterative improvements to these reporters have been made, though, and these tools are enabling

more sensitive imaging in live animals [17-19]. In a recent example, Lin and colleagues used structure-guided mutagenesis to engineer a far red-emitting, monomeric FP (mCardinal) with improved brightness and stability [20]. mCardinal enabled noninvasive, serial imaging of myoblast differentiation in deep tissues. Verkusha and colleagues also developed a suite of stable and spectrally distinct FPs that emit near-infrared light (iFPs) [21-22]. Some of these iFPs can be distinguished in a single animal model, enabling multicolor imaging of cellular networks (Figure 1-1a).

Combinatorial FP expression for color-coding cells. Additional “colors” for cell tracking are available using combinations of FPs. A stellar example is 'Brainbow', a genetic recombination technique that results in stochastic expression of four distinct FPs (orange, red, yellow, cyan) in individual cells. The combination of expressed FPs effectively marks each cell with a unique color for long-term visualization [23-24]. 'Brainbow' was originally applied to color code mouse neurons for imaging cell contacts in brain tissue (i.e., “connectomes”). More recently, the approach has been used to track non-neuronal cells in mice [25], along with cell migration patterns in other organisms [26-27]. Application of the 'Brainbow' technique to other cellular networks and in combination with new FPs will bolster efforts to track cellular interactomes.

Photoswitchable FPs for long-term tracking. A variety of FPs exhibit unique photophysical properties that can be exploited for imaging cell-cell interactions. One such FP, Kaede, irreversibly changes from green to red in color [28]. Mathis and colleagues exploited this photoconvertible protein to “mark” immune cells and monitor their movements to and from mouse intestines [29]. The intestinal tract is inhabited by diverse collections of microbes, and our understanding of their roles in



(c)

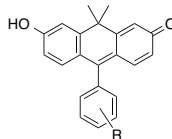
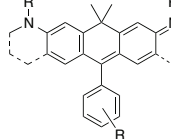
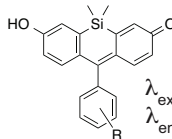
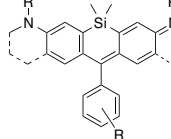
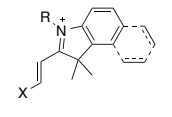
Recently reported far-red to infrared fluorescent proteins (FPs)							
Fluorescent protein	λ_{ex} (nm)	λ_{em} (nm)	ϵ ($M^{-1} cm^{-1}$)	ϕ	Photostability $t_{1/2}$ (s)	Quaternary structure	Sample refs
iRFP670	643	670	114,000	0.07	290	dimer	Filanov 2011 Scherbakova 2013
iRFP682	663	682	90,000	0.11	490		
iRFP702	673	702	93,000	0.08	630		
iRFP713	690	713	98,000	0.06	960		
iRFP720	702	720	96,000	0.06	490		
mCardinal	604	659	87,000	0.19	730	monomer	Chu 2014
TagRFP675	598	675	46,000	0.08	35	monomer	Piatkevich 2014
iFP2.0	690	711	86,125	0.08	—	monomer	Yu 2014
Recently reported red-shifted small molecule fluorophores							
Fluorophore type	Structure						Sample refs
Carbofluoresceins/ carborhodamines	 fluoresceins $\lambda_{ex} = 544$ nm $\lambda_{em} = 567$ nm			 rhodamines $\lambda_{ex} = 552-640$ nm $\lambda_{em} = 520-664$ nm			Kolmakov 2012 Grimm 2013
	Silafluoresceins/ silarhodamines	 fluoresceins $\lambda_{ex} = 472-597$ nm $\lambda_{em} = 593-609$ nm			 rhodamines $\lambda_{ex} = 646-721$ nm $\lambda_{em} = 660-740$ nm		
Cyanine derivatives					X = coumarin $\lambda_{ex} = 608-690$ nm $\lambda_{em} = 677-716$ nm X = fluorescein $\lambda_{ex} = 635$ nm $\lambda_{em} = 705$ nm X = rhodamine $\lambda_{ex} = 688-728$ nm $\lambda_{em} = 721-763$ nm		

Figure 1-1. Multicellular imaging in live animals. (a) Visualizing fluorescent proteins (FPs) in vivo. Upon excitation, FPs emit light that can be captured by sensitive cameras. In some cases, unique FPs (and, thus, the cells expressing them) can be spectrally resolved. **(b)** MTLn3 cells expressing infrared fluorescent proteins iRFP702 (left) or iRFP7 (right) were imaged in a single mouse model. Following irradiation, images were acquired in 19 distinct channels. Spectral unmixing provided the images shown. Images are reprinted with permission from ref. [22], copyright (2013) Nature Publishing Group. **(c)** Intravital imaging of cellular interactions in mouse liver. T cells (red) were observed to contact and destroy GFP-expressing *Plasmodium yoelii* parasites (green). This image is reprinted with permission from ref. [14], copyright (2013) National Academy of Sciences. U. S. A. **(d)** Overview of recently developed FPs and small molecule fluorophores for cellular imaging.

influencing immunological functions—and mammalian biology in general—is only in its infancy. The authors used an endoscope to deliver light and photoconvert Kaede-labeled immune cells in mouse models. They discovered large populations of leukocytes that travel to and from the gut. These results have important implications for the role of gut microbiota in influencing immune function.

Improved fluorophores for short-term tracking. While fluorescent proteins are the go-to choice for long-term visualization, small molecule fluorophores are attractive for short-term studies of cell-cell interactions [30-31]. These probes can be appended to cells in a variety of manners (intercalation, antibody-delivery, etc.), but they do not propagate with cell division. Small molecule fluorophores also come in an assortment of colors, and some (e.g., fluorescein and rhodamine) have been used extensively in combination with intravital imaging and multi-photon microscopy (Figure 1-1b) [4].

Similar to fluorescent protein technology, red- and infrared-emitting fluorophores are desirable for noninvasive cell tracking [32-35]. These probes have been notoriously difficult to access and handle, though, and remain the subject of intense research [34,36-37]. Recent efforts include generating more photostable cyanine dyes [38] and fluorophores with altered electronic properties [39-40]. In one example, Lavis and colleagues replaced the xanthene oxygens in fluorescein and rhodamine with quaternary carbons [39]. The resulting "carbofluorescein" and "carborhodamine" probes exhibited robust fluorescent properties and enabled sensitive imaging in cells (in the absence of anti-bleaching agents). The divergent syntheses of the carbofluorophores also enabled the facile construction of numerous fluorogenic probes for facile cell

labeling [39]. Similar strategies have been used to access more red- and infrared-emitting fluorophores for in vivo imaging [34,41-42].

Improved strategies for probe delivery. Concomitant with improved scaffold development; new methods to append small molecule probes to cellular targets are advancing studies of cell-cell interactions. Traditional approaches to target labeling have involved targets via lipid insertion, non-selective bioconjugation chemistries, or antibody labeling of ubiquitous cell targets. Many of these approaches result in non-uniform labeling or poor probe retention. More selective and modular methods for covalent probe attachment have been developed in recent years and are expanding the scope of cell visualization [6,43-44]. Many of these are also amenable to use with quantum dots and related optical nanostructures [45]. Among the most versatile technologies for attaching probes to cell surfaces are enzyme-assisted tagging methods. These strategies rely on enzyme-mediated recognition and modification of small, genetically encoded peptide tags [42,46]. In one example, Ting and colleagues engineered bacterial lipoic acid ligase (LplA) to append various probes to a short, 13-amino acid recognition sequence (LAP1) [47]. LplA can transfer a variety of probes to LAP1, including bioorthogonal azides [48]. Azides can be selectively ligated with fluorescent probes in a second step using “click” chemistry. This approach was used to selectively attach imaging probes to cell surface proteins.

1.3 Engineered fluorescent tools for examining cell-cell contacts

Imaging cells with discrete fluorescent probes is a classic method for visualizing networks of cells in live tissues and organisms. Multicolor imaging, though, does not

provide an unambiguous readout on physical cell contacts. Direct cell-cell contacts underlie memory formation, immune cell function, among other processes. Thus, methods that unequivocally report on physical cell contacts would enable critical aspects of cell communication to be more accurately mapped.

Contact-dependent “split” reporters. One technique for visualizing cell-cell interactions, GFP Reconstitution Across Synaptic Partners (or GRASP), is based on protein complementation. In this approach, nonfluorescent GFP fragments are appended to unique pairs of extracellular proteins. When brought into contact, the split

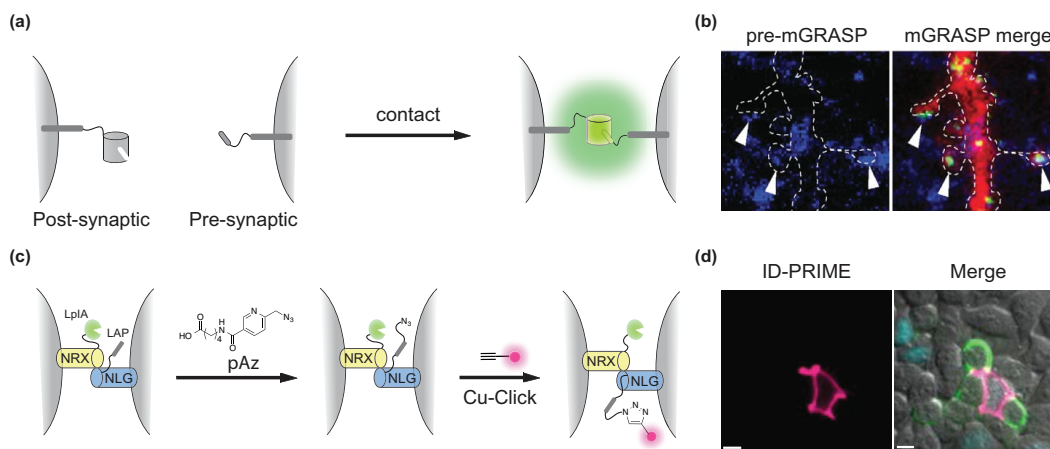


Figure 1-2. New tools for imaging direct cell-cell interactions. (a) Illuminating cell contacts with GRASP technology. In one example, split fragments of GFP were tethered to CD4 transmembrane proteins in adjacent neurons. Upon cell contact, the split fragments assembled to provide functional GFP. (b) Synaptic connections in mouse brain visualized with GRASP. Left: Pre-synaptic neurons expressing mCerulean (blue). Right: GFP reconstitution and fluorescence (green) is observed upon contact between pre- (blue) and post-synaptic (red) neurons. Sites of GFP reconstitution are denoted by white arrows. Reprinted with permission from ref. [49], copyright (2012) Nature Publishing Group. (c) ID-PRIME labeling of transcellular interactions. LplA and LAP were fused to neurexin (NRX) and neuroligin (NLG), respectively. Upon NRX-NLG binding, LplA can modify LAP with an exogenously supplied azido probe (pAz). The azides were detected in a second step using an alkyne fluorophore and copper-catalyzed “click” chemistry (Cu-Click). (d) ID-PRIME imaging of intercellular interactions. HEK cells expressing NRX-LplA were plated with cells expressing NLG-LAP. The cultures were incubated with pAz for 5 min, then “clicked” with an alkyne-AF647 conjugate. Left: AF647 (pink) labeling of cell contacts. Right: AF647 labeling merged with a marker of NLG-LAP-expressing cells (blue) and a marker for NRX-LplA (green). Reprinted with permission from ref. [50], copyright (2013) American Chemical Society.

fragments assemble to provide functional, light-emitting GFP (Figure 1-2a). GRASP was initially used to rapidly locate synapses in intact nematodes and fruit flies with high spatial resolution [51-52]. More recently, the technique has been adapted to imaging new synapse formation in mouse brains (Figure 1-2b) [49,53]. Other split FPs will likely be developed for *intercellular* imaging [54].

Enzymatic tagging of cell-cell interactions. In addition to marking cells with unique probes, enzyme-assisted tagging technologies can be co-opted for imaging cell-cell contacts. In recent work, Ting and Zegelbone utilized cell-surface LpIA to selectively label proteins on interacting cells (Figure 1-2c-d) [47,50]. When the ligase and acceptor peptide were fused to surface proteins neurexin (NRX) and neuroligin (NLG), LpIA transferred picolyl azide only when the proteins were in direct contact. The azide was tagged with a visual probe in a second step. This approach—interaction-dependent probe incorporation mediated by enzymes (ID-PRIME)—enabled selective imaging of *trans*-cellular interactions in multiple cell types (Figure 1-2d). ID-PRIME is also versatile technology and should be amenable to tagging cell-cell interactions with a variety of probes for *in vivo* translation is expected.

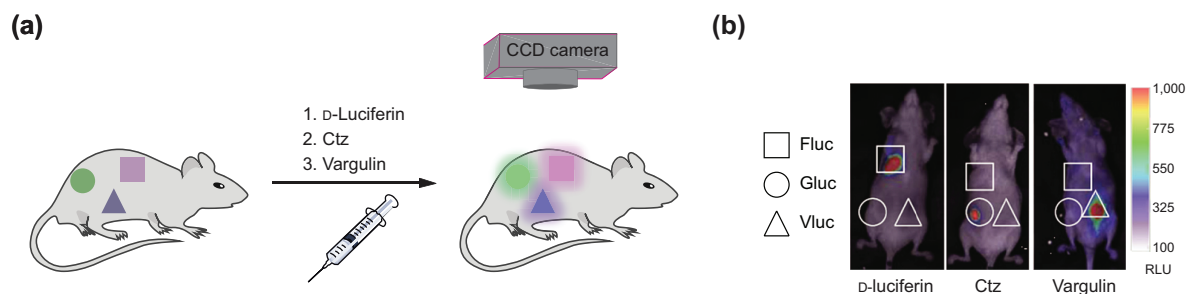
1.4 Improved bioluminescent probes for noninvasive imaging

While the expansive palette of fluorescent probes enables cell-cell contacts to be readily discerned over microscopic distances, these tools have been historically difficult to transition to larger length scales (i.e., bigger organisms). This is largely due to the requirement for photon excitation: one must know “when and where” to shine the light. Bioluminescence imaging, by contrast, does not require excitation light and is often

more suitable for global cell tracking experiments *in vivo*. The genes encoding luciferases can be expressed in numerous cell types, and when luciferin is administered, light is produced that can be captured by sensitive cameras [7,9]. Luciferase-tagged cells can also be imaged repeatedly and noninvasively, making bioluminescence particularly well suited for longitudinal studies in a variety of preclinical models [55-56].

While versatile, bioluminescence has been largely limited to imaging one cell type or biological feature at a time owing to a lack of unique probes. Dozens of unique luciferases and luciferins have been identified in nature, but most are difficult to resolve *in vivo* owing to overlapping substrate usage or emission spectra [5,7]. Multicellular imaging is possible with biochemically distinct luciferase-luciferin pairs; however, the luciferins must be administered sequentially (often days apart) for unique cell populations to be discerned [57-58] (Figure 1-3a-b). In a recent example, Tannous and colleagues visualized three distinct tumor cell populations in mice [58]. The cells were tagged with three unique luciferases (*Vargula* luciferase, *Renilla* luciferase, and firefly luciferase) and imaged via sequential administration of the cognate luciferins (vargulin, coelenterazine, and D-luciferin, respectively).

Engineered luciferases and luciferins for noninvasive imaging. Several groups are attempting to address the need for distinct luciferase and luciferins probes. Mutant versions of luciferase have been crafted and found to possess altered spectral properties [59]. More recent efforts to engineer luciferins have similarly provided luciferins capable of emitting different colors of light (Figure 1-3c) [60-63]. Other new



(c)

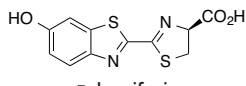
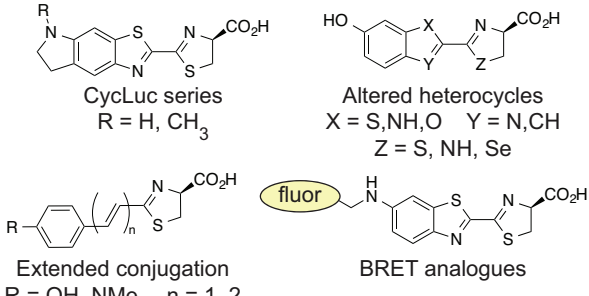
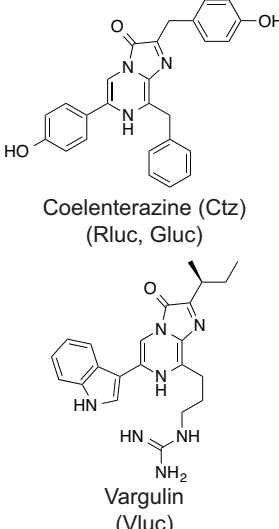
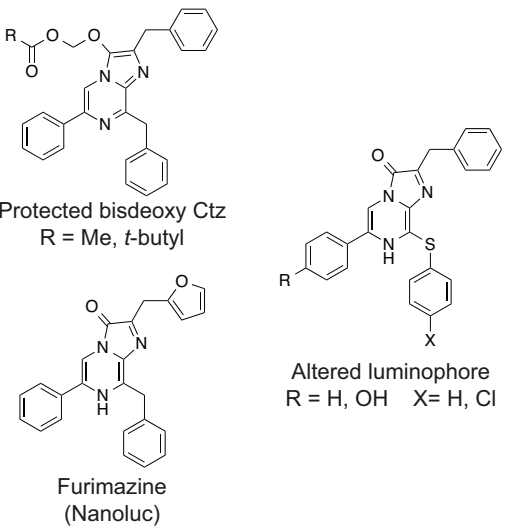
Luciferase substrates and recent derivatives			
Luciferase	Native substrate(s)	Recently developed analogs	Sample refs
Firefly luciferase (Fluc)	 D-Luciferin	 CycLuc series $R = H, CH_3$ Altered heterocycles $X = S, NH, O$ $Y = N, CH$ $Z = S, NH, Se$ Extended conjugation $R = OH, NMe_2$ $n = 1, 2$ BRET analogues	Reddy 2010 Evans 2014 McCutcheon 2012 Woodroffe 2012 Conley 2012 Iwano 2013 Kojima 2013
Marine luciferases	 Coelenterazine (Ctz) (Rluc, Gluc) Vargulin (Vluc)	 Protected bisdeoxy Ctz $R = Me, t\text{-butyl}$ Furimazine (Nanoluc) Altered luminophore $R = H, OH$ $X = H, Cl$	Macquire 2013 Levi 2007 Hall 2012 Giuliani 2012

Figure 1-3. Expanding the bioluminescent toolbox. (a) Multicomponent bioluminescence imaging. Cells expressing unique luciferases: firefly luciferase (Fluc) Gaussia luciferase (Gluc), and Vargula luciferase (Vluc) were implanted in mice (locations denoted by shapes). Sequential administration of the cognate luciferins (D-luciferin, coelenterazine (Ctz), and vargulin) enabled detection of the three populations. Each luciferin was administered one day apart. Reprinted with permission from [58]. Copyright (2013) Nature Publishing Group **(b)** Luciferase-luciferin pairs for *in vivo* imaging.

substrates offer improved cell and tissue permeabilities and shelf stability [64-68]. Collectively, these tools remain difficult to spectrally resolve *in vivo*.

Perhaps a more fruitful approach involves simultaneous modification of luciferase enzymes and luciferin substrates. In recent work, Wood and colleagues designed a cell permeable, stable analog of coelenterazine (furimazine, Figure 1-3c). The authors then evolved a luciferase from a deep-sea shrimp specific for this designer luciferin [66]. The resulting luciferase (Nanoluc) provided brighter and more stable light emission than related bioluminescent systems, and has found application in point-of-care drug monitoring [69] and pathogen detection [70]. Miller and co-workers used a similar approach to identify mutant versions of firefly luciferase that more efficiently catalyze light production with aminoluciferin variants [71-72]. These tools will facilitate the simultaneous tracking of cells *in vivo* and thus enhance our understanding of cellular interactomes.

1.5 Next-generation bioluminescent probes for imaging cell-cell communication

Bioluminescence is sufficiently sensitive to monitor contacts between small numbers of cells in live animals, but can only approximate cell location owing to its low spatial resolution. By contrast, cell-cell interactions can be readily detected via intravital microscopy. However, these methods require invasive surgical procedures to expose or collect tissue. Developing general methods for studying cellular interactions *in vivo* that blend the macroscopic and noninvasive features of bioluminescence with the spatial resolution and rich information content of microscopy remain important goals.

The Prescher group is crafting novel bioluminescent tools that produce light only when two cells come into close proximity or direct physical contact. In one approach, we synthesized “caged” probes—molecules comprising a luciferin core outfitted with appendages (i.e., “cages”) that preclude binding to luciferase. In the presence of cells capable of excising the cage (“activator” cells), luciferin is liberated and available for use by luciferase-expressing (“reporter”) cells. Reporter cells nearest the activator cells consume the most substrate; thus, light intensity correlates with the proximity of the two populations (Figure 1-4a,b). We initially synthesized a galactose-caged luciferin (Lugal) to report on the proximity between β -gal-expressing activator cells and luciferase-expressing reporter cells [73]. When activator cells were localized to sites of metastases (e.g., lymph nodes), Lugal administration signaled the invasion of luciferase-expressing tumor cells and thus metastases in live mice (Figure 1-4c). Additional caged luciferins will enhance this more robust readout on cell-cell interactions, along with alternative technologies.

1.6 Objectives of this study

Cellular communication drives diverse aspects of organismal biology ranging from immune function to memory formation. The mechanisms by which cells transmit information *in vivo*, though, are only partially understood. For example, how immune cells interact with tumor cells or pathogens has been difficult to study in their native habitat. These cellular interactions have been studied at the microscopic level [14,74], but not at the macroscopic, whole animal level. Bioluminescence is a promising technique for improving the utility of optical imaging in whole animal models. It is more

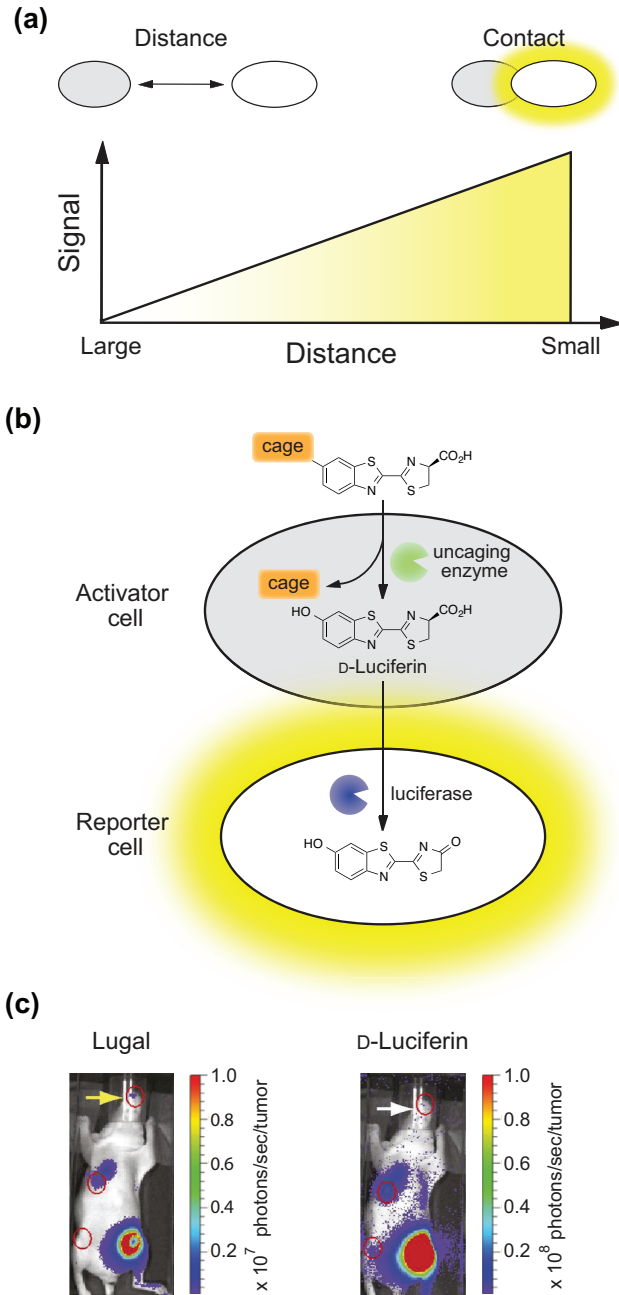


Figure 1-4. Macroscale detection of microscale cellular interactions. (a) Strategy for imaging distance-dependent interactions. (b) “Caged” luciferins can report on cell proximity. The “cage” can be selectively released in activator cells, and the liberated luciferin can be utilized by nearby luciferase-expressing reporter cells. Robust photon production is observed only when the activator and reporter cells are in close proximity. Right: Tumor-immune cell interactions visualized in live mice. Mice were inoculated with bone marrow from beta-galactosidase transgenic mice. Upon implantation of luciferase-expressing metastatic tumor cells (4T1) in these mice, tumor-immune cell interactions could be selectively visualized using Lugal (left) or D-luciferin (right). Lugal imaging was able to provide more sensitive imaging of cell interactions at metastatic sites (red circles). Subsequent *ex vivo* analysis confirmed both beta-gal expressing immune cells and luciferase expressing tumor cells were present at sites. Reprinted with permission from ref. [73]. Copyright (2013) National Academy of Sciences U.S.A.

sensitive than fluorescence due to the decreased background of luminescence. In addition, the simplicity and relative inexpensiveness of BLI combine to make this technique more approachable than PET, MRI or other similar techniques. Yet, the current state of BLI is lacking in two major ways. (1) The spatial resolution of BLI is too poor to report on microscopic, or cellular level events. (2) The small number (2 practically used pairs) of unique and selective bioluminescent systems precludes the imaging of multiple cell types, genes, etc. where two or more targets are to be imaged simultaneously. The tools and methods developed throughout my thesis work improves the utility of bioluminescence for use in complex *in vivo* environments.

I aimed to:

1. Synthesize and characterize novel 'caged' luciferins for their use as a reporter on cell-cell contact and increase the spatial resolution of BLI.
2. Develop a screening platform to engineer orthogonal luciferase-luciferin pairs for multicomponent imaging.
3. Synthesize novel luciferins and evolve selective luciferase enzymes for improved *in vivo*, multicomponent imaging.

References

1. Mittelbrunn, M.; Sanchez-Madrid, F., Intercellular communication: Diverse structures for exchange of genetic information. *Nat. Rev. Mol. Cell. Biol.* **2012**, *13*, 328.
2. Mittelbrunn, M.; Gutierrez-Vazquez, C.; Villarroya-Beltri, C.; Gonzalez, S.; Sanchez-Cabo, F.; Gonzalez, M. A.; Bernad, A.; Sanchez-Madrid, F., Unidirectional transfer of microrna-loaded exosomes from t cells to antigen-presenting cells. *Nat. Commun.* **2011**, *2*, 282. doi: 10.1038/ncomms1285.

3. Lai, C. P.; Tannous, B. A.; Breakefield, X. O., Noninvasive in vivo monitoring of extracellular vesicles. *Bioluminescent Imaging: Methods and Protocols* **2014**, 1098, 249.
4. Germain, R. N.; Robey, E. A.; Cahalan, M. D., A decade of imaging cellular motility and interaction dynamics in the immune system. *Science* **2012**, 336, 1676.
5. Prescher, J. A.; Contag, C. H., Guided by the light: Visualizing biomolecular processes in living animals with bioluminescence. *Curr. Opin. Chem. Biol.* **2010**, 14, 80.
6. Dean, K. M.; Palmer, A. E., Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nat. Chem. Biol.* **2014**, 10, 512.
7. Paley, M. A.; Prescher, J. A., Bioluminescence: A versatile technique for imaging cellular and molecular features. *Medchemcomm* **2014**, 5, 255.
8. Rice, B. W.; Contag, C. H., The importance of being red. *Nat. Biotechnol.* **2009**, 27, 624.
9. Adams, S. T.; Miller, S. C., Beyond d-luciferin : Expanding the scope of bioluminescence imaging in vivo. *Curr. Opin. Chem. Biol.* **2014**, 21, 112.
10. Drobizhev, M.; Makarov, N. S.; Tillo, S. E.; Hughes, T. E.; Rebane, A., Two-photon absorption properties of fluorescent proteins. *Nature Methods* **2011**, 8, 393.
11. Li, J. L.; Goh, C. C.; Keeble, J. L.; Qin, J. S.; Roediger, B.; Jain, R.; Wang, Y.; Chew, W. K.; Weninger, W.; Ng, L. G., Intravital multiphoton imaging of immune responses in the mouse ear skin. *Nat. Protoc.* **2012**, 7, 221.
12. Matheu, M. P.; Teijaro, J. R.; Walsh, K. B.; Greenberg, M. L.; Marsolais, D.; Parker, I.; Rosen, H.; Oldstone, M. B. A.; Cahalan, M. D., Three phases of cd8 t cell response in the lung following H1N1 influenza infection and sphingosine 1 phosphate agonist therapy. *PLoS One* **2013**, 8, e58033.
13. Matheu, M. P.; Su, Y.; Greenberg, M. L.; Blanc, C. A.; Parker, I.; Scott, D. W.; Cahalan, M. D., Toll-like receptor 4-activated B cells out-compete toll-like receptor 9-activated B cells to establish peripheral immunological tolerance. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, 109, E1258.
14. Cockburn, I. A.; Amino, R.; Kelemen, R. K.; Kuo, S. C.; Tse, S. W.; Radtke, A.; Mac-Daniel, L.; Ganusov, V. V.; Zavala, F.; Ménard, R., In vivo imaging of CD8⁺ T cell-mediated elimination of malaria liver stages. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, 110, 9090.

15. Shu, X.; Royant, A.; Lin, M. Z.; Aguilera, T. A.; Lev-Ram, V.; Steinbach, P. A.; Tsien, R. Y., Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science* **2009**, *324*, 804.
16. Strack, R. L.; Hein, B.; Bhattacharyya, D.; Hell, S. W.; Keenan, R. J.; Glick, B. S., A rapidly maturing far-red derivative of DsRed-express2 for whole-cell labeling. *Biochemistry* **2009**, *48*, 8279.
17. Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y., A guide to choosing fluorescent proteins. *Nature Methods* **2005**, *2*, 905.
18. Piatkevich, K. D.; Subach, F. V.; Verkhusha, V. V., Engineering of bacterial phytochromes for near-infrared imaging, sensing, and light-control in mammals. *Chem. Soc. Rev.* **2013**, *42*, 3441.
19. Yu, D.; Gustafson, W. C.; Han, C.; Lafaye, C.; Noirclerc-Savoie, M.; Ge, W.-P.; Thayer, D. A.; Huang, H.; Kornberg, T. B.; Royant, A.; Jan, L. Y.; Jan, Y. N.; Weiss, W. A.; Shu, X., An improved monomeric infrared fluorescent protein for neuronal and tumour brain imaging. *Nat. Commun.* **2014**, *5*, 3626. DOI: 10.1038/ncomms4626
20. Chu, J.; Haynes, R. D.; Corbel, S. Y.; Li, P.; González-González, E.; Burg, J. S.; Ataie, N. J.; Lam, A. J.; Cranfill, P. J.; Baird, M. A.; Davidson, M. W.; Ng, H.-L.; Garcia, K. C.; Contag, C. H.; Shen, K.; Blau, H. M.; Lin, M. Z., Non-invasive intravital imaging of cellular differentiation with a bright red-excitable fluorescent protein. *Nature Methods* **2014**, *11*, 572.
21. Filonov, G. S.; Piatkevich, K. D.; Ting, L.-M.; Zhang, J.; Kim, K.; Verkhusha, V. V., Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat. Biotechnol.* **2011**, *29*, 757.
22. Shcherbakova, D. M.; Verkhusha, V. V., Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nature Methods* **2013**, *10*, 751.
23. Livet, J.; Weissman, T. A.; Kang, H. N.; Draft, R. W.; Lu, J.; Bennis, R. A.; Sanes, J. R.; Lichtman, J. W., Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **2007**, *450*, 56.
24. Cai, D.; Cohen, K. B.; Luo, T.; Lichtman, J. W.; Sanes, J. R., Improved tools for the brainbow toolbox. *Nature Methods* **2013**, *10*, 540.
25. Schepers, A. G.; Snippert, H. J.; Stange, D. E.; van den Born, M.; van Es, J. H.; van de Wetering, M.; Clevers, H., Lineage tracing reveals Igr5⁺ stem cell activity in mouse intestinal adenomas. *Science* **2012**, *337*, 730.
26. Gupta, V.; Poss, K. D., Clonally dominant cardiomyocytes direct heart morphogenesis. *Nature* **2012**, *484*, 479.

27. Hadjieconomou, D.; Rotkopf, S.; Alexandre, C.; Bell, D. M.; Dickson, B. J.; Salecker, I., Flybow: Genetic multicolor cell labeling for neural circuit analysis in *Drosophila melanogaster*. *Nature Methods* **2011**, *8*, 260.
28. Tomura, M.; Yoshida, N.; Tanaka, J.; Karasawa, S.; Miwa, Y.; Miyawaki, A.; Kanagawa, O., Monitoring cellular movement in vivo with photoconvertible fluorescence protein "kaede" transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 10871.
29. Morton, A. M.; Sefik, E.; Upadhyay, R.; Weissleder, R.; Benoist, C.; Mathis, D., Endoscopic photoconversion reveals unexpectedly broad leukocyte trafficking to and from the gut. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 6696.
30. Lavis, L. D.; Raines, R. T., Bright building blocks for chemical biology. *ACS Chem. Biol.* **2014**, *9*, 855.
31. Guo, Z.; Park, S.; Yoon, J.; Shin, I., Recent progress in the development of near-infrared fluorescent probes for bioimaging applications. *Chem. Soc. Rev.* **2014**, *43*, 16.
32. Koide, Y.; Urano, Y.; Hanaoka, K.; Piao, W.; Kusakabe, M.; Saito, N.; Terai, T.; Okabe, T.; Nagano, T., Development of NIR fluorescent dyes based on si-rhodamine for in vivo imaging. *J. Am. Chem. Soc.* **2012**, *134*, 5029.
33. Yuan, L.; Lin, W.; Yang, Y.; Chen, H., A unique class of near-infrared functional fluorescent dyes with carboxylic-acid-modulated fluorescence on/off switching: Rational design, synthesis, optical properties, theoretical calculations, and applications for fluorescence imaging in living animals. *J. Am. Chem. Soc.* **2012**, *134*, 1200.
34. Yuan, L.; Lin, W.; Zhao, S.; Gao, W.; Chen, B.; He, L.; Zhu, S., A unique approach to development of near-infrared fluorescent sensors for in vivo imaging. *J. Am. Chem. Soc.* **2012**, *134*, 13510.
35. McCormack, E.; Silden, E.; West, R. M.; Pavlin, T.; Micklem, D. R.; Lorens, J. B.; Haug, B. E.; Cooper, M. E.; Gjertsen, B. T., Nitroreductase, a near-infrared reporter platform for in vivo time-domain optical imaging of metastatic cancer. *Cancer Res.* **2013**, *73*, 1276.
36. Xiong, X.; Song, F.; Chen, G.; Sun, W.; Wang, J.; Gao, P.; Zhang, Y.; Qiao, B.; Li, W.; Sun, S.; Fan, J.; Peng, X., Construction of long-wavelength fluorescein analogues and their application as fluorescent probes. *Chem. Eur. J.* **2013**, *19*, 6538.
37. Bochkov, A. Y.; Akchurin, I. O.; Dyachenko, O. A.; Traven, V. F., NIR-fluorescent coumarin-fused bodipy dyes with large stokes shifts. *Chem. Commun.* **2013**, *49*, 11653.

38. Altman, R. B.; Terry, D. S.; Zhou, Z.; Zheng, Q.; Geggier, P.; Kolster, R. A.; Zhao, Y.; Javitch, J. A.; Warren, J. D.; Blanchard, S. C., Cyanine fluorophore derivatives with enhanced photostability. *Nature Methods* **2012**, *9*, 68.
39. Grimm, J. B.; Sung, A. J.; Legant, W. R.; Hulamm, P.; Matlosz, S. M.; Betzig, E.; Lavis, L. D., Carbofluoresceins and carborhodamines as scaffolds for high-contrast fluorogenic probes. *ACS Chem. Biol.* **2013**, *8*, 1303.
40. Kolmakov, K.; Belov, V. N.; Wurm, C. A.; Harke, B.; Leutenegger, M.; Eggeling, C.; Hell, S. W., A versatile route to red-emitting carbopyronine dyes for optical microscopy and nanoscopy. *Eur. J. Org. Chem.* **2010**, Doi 10.1002/Ejoc.201000343, 3593.
41. Koide, Y.; Urano, Y.; Hanaoka, K.; Terai, T.; Nagano, T., Evolution of group 14 rhodamines as platforms for near-infrared fluorescence probes utilizing photoinduced electron transfer. *ACS Chem. Biol.* **2011**, *6*, 600.
42. Lukinavičius, G.; Umezawa, K.; Olivier, N.; Honigmann, A.; Yang, G.; Plass, T.; Mueller, V.; Reymond, L.; Corrêa, I. R.; Luo, Z.-G.; Schultz, C.; Lemke, E. A.; Heppenstall, P.; Eggeling, C.; Manley, S.; Johnsson, K., A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. *Nat. Chem.* **2013**, *5*, 132.
43. Mali, P.; Aach, J.; Lee, J. H.; Levner, D.; Nip, L.; Church, G. M., Barcoding cells using cell-surface programmable DNA-binding domains. *Nature Methods* **2013**, *10*, 403.
44. Patterson, D. M.; Nazarova, L. A.; Prescher, J. A., Finding the right (bioorthogonal) chemistry. *ACS Chem. Biol.* **2014**, *9*, 592.
45. Bao, G.; Mitragotri, S.; Tong, S., Multifunctional nanoparticles for drug delivery and molecular imaging. *Annu. Rev. Biomed. Eng.* **2013**, *15*, 253.
46. Liu, T. K.; Hsieh, P. Y.; Zhuang, Y. D.; Hsia, C. Y.; Huang, C. L.; Lai, H. P.; Lin, H. S.; Chen, I. C.; Hsu, H. Y.; Tan, K. T., A rapid snap-tag fluorogenic probe based on an environment-sensitive fluorophore for no-wash live cell imaging. *ACS Chem. Biol.* **2014**, 10.1021/cb500502n, dx.doi.org/10.1021/cb500502n.
47. Liu, D. S.; Loh, K. H.; Lam, S. S.; White, K. a.; Ting, A. Y., Imaging trans-cellular neurexin-neuroligin interactions by enzymatic probe ligation. *PLoS One* **2013**, *8*, e52823.
48. Uttamapinant, C.; Sanchez, M. I.; Liu, D. S.; Yao, J. Z.; Ting, A. Y., Site-specific protein labeling using prime and chelation-assisted click chemistry. *Nat. Protoc.* **2013**, *8*, 1620.

49. Kim, J.; Zhao, T.; Petralia, R. S.; Yu, Y.; Peng, H.; Myers, E.; Magee, J. C., Mgrasp enables mapping mammalian synaptic connectivity with light microscopy. *Nature Methods* **2012**, *9*, 96.
50. White, K. A.; Zegelbone, P. M., Directed evolution of a probe ligase with activity in the secretory pathway and application to imaging intercellular protein-protein interactions. *Biochemistry* **2013**, *52*, 3728.
51. Feinberg, E. H.; VanHoven, M. K.; Bendesky, A.; Wang, G.; Fetter, R. D.; Shen, K.; Bargmann, C. I., GFP reconstitution across synaptic partners (grasp) defines cell contacts and synapses in living nervous systems. *Neuron* **2008**, *57*, 353.
52. Gordon, M. D.; Scott, K., Motor control in a drosophila taste circuit. *Neuron* **2009**, *61*, 373.
53. Cavanaugh, D. J.; Geratowski, J. D.; Wooltorton, J. R.; Spaethling, J. M.; Hector, C. E.; Zheng, X.; Johnson, E. C.; Eberwine, J. H.; Sehgal, A., Identification of a circadian output circuit for rest:activity rhythms in drosophila. *Cell* **2014**, *157*, 689.
54. Filonov, G. S.; Verkhusha, V. V., A near-infrared *BifC* reporter for in vivo imaging of protein-protein interactions. *Chem. Biol.* **2013**, *20*, 1078.
55. Gupta, G. P.; Nguyen, D. X.; Chiang, A. C.; Bos, P. D.; Kim, J. Y.; Nadal, C.; Gomis, R. R.; Manova-Todorova, K.; Massague, J., Mediators of vascular remodelling co-opted for sequential steps in lung metastasis. *Nature* **2007**, *446*, 765.
56. Acharyya, S.; Oskarsson, T.; Vanharanta, S.; Malladi, S.; Kim, J.; Morris, P. G.; Manova-Todorova, K.; Leversha, M.; Hogg, N.; Seshan, V. E.; Norton, L.; Brogi, E.; Massague, J., A CXCL1 paracrine network links cancer chemoresistance and metastasis. *Cell* **2012**, *150*, 165.
57. Wang, H.; Cao, F.; De, A.; Cao, Y.; Contag, C.; Gambhir, S. S.; Wu, J. C.; Chen, X. Y., Trafficking mesenchymal stem cell engraftment and differentiation in tumor-bearing mice by bioluminescence imaging. *Stem Cells* **2009**, *27*, 1548.
58. Maguire, C. A.; Bovenberg, M. S.; Crommentuijn, M. H.; 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. Nucleic Acids* **2013**, *2*, e99.
59. 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.
60. 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.

61. 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.
62. 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.
63. 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.
64. Evans, M. S.; Chaurette, J. P.; Adams, 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. *Nature Methods* **2014**, *11*, 393.
65. 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.
66. 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.
67. Giuliani, G.; Molinari, P.; Ferretti, G.; Cappelli, A.; Anzini, M.; Vomero, S.; Costa, T., New red-shifted coelenterazine analogues with an extended electronic conjugation. *Tetrahedron Lett.* **2012**, *53*, 5114.
68. Levi, J.; De, A.; Cheng, Z.; Gambhir, S. S., Bisdeoxycoelenterazine derivatives for improvement of bioluminescence resonance energy transfer assays. *J. Am. Chem. Soc.* **2007**, *129*, 11900.
69. Griss, R.; Schena, A.; Reymond, L.; Patiny, L.; Werner, D.; Tinberg, C. E.; Baker, D.; Johnsson, K., Bioluminescent sensor proteins for point-of-care therapeutic drug monitoring. *Nat. Chem. Biol.* **2014**, *10*, 598.
70. Sun, C.; Gardner, C. L.; Watson, A. M.; Ryman, K. D.; Klimstra, W. B., Stable, high-level expression of reporter proteins from improved alphavirus expression vectors to track replication and dissemination during encephalitic and arthritogenic disease. *J. Virol.* **2014**, *88*, 2035.
71. Harwood, K. R.; Mofford, D. M.; Reddy, G. R.; Miller, S. C., Identification of mutant firefly luciferases that efficiently utilize aminoluciferins. *Chem. Biol.* **2011**, *18*, 1649.

72. 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.
73. Sellmyer, M. A.; Bronsart, L.; Imotoa, H.; Contag, C. H.; Wandless, T. J.; Prescher, J. A., Visualizing cellular interactions with a generalized proximity reporter. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 8567.
74. Germain, R. N.; Robey, E. A.; Cahalan, M. D., A decade of imaging cellular motility and interaction dynamics in the immune system. *Science* **2012**, *336*, 1676.

Chapter 2: ‘Caged’ probes for imaging cell-cell contact

2.1 Introduction

Cell-cell contacts govern numerous biological processes, including cell growth, motility, and immune function [1]. Our understanding of these interactions is critically dependent on our ability to “see” them, and several fluorescence imaging techniques have been developed for this purpose [1-2]. While powerful, these strategies require intense excitation sources and are thus largely limited to visualizing interactions on a microscopic scale [2-4]. As introduced in Chapter 1, bioluminescence imaging (BLI), a complementary optical technique, is often more suitable for capturing macroscopic cell movements in whole tissues and organisms as it does not require excitation light.

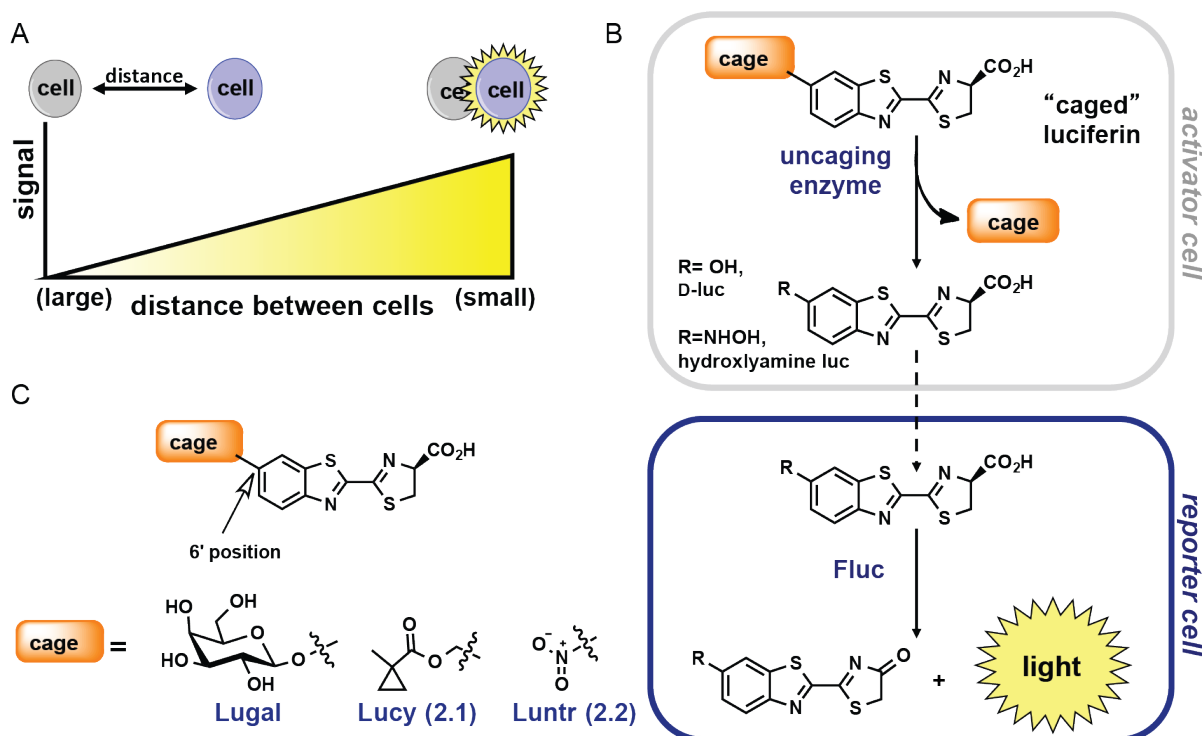


Figure 2-1. General strategy for visualizing cell-cell interactions. (A) Light is produced only when cells are in close contact. (B) “Caged” luciferins enter activator cells, where an uncaging enzyme (e.g., β -gal, PLE, or NTR) liberates an active luciferin. This molecule diffuses into nearby reporter cells (expressing Fluc), and light is produced. (C) Structures of “caged” luciferins discussed in this work.

However, in its current form, BLI is incapable of reporting on direct cell-cell interactions owing to its low spatial resolution [3,5-6]. This leaves a gap in our ability to investigate dynamic cellular interactions across large length and time scales without knowing when and where to look.

To address this void, we aimed to engineer bioluminescent probes that could report on cell-cell contacts. Bioluminescence exploits enzymes (luciferases) that catalyze light production via the oxidation of small molecule substrates (luciferins). The most widely used enzyme-substrate pair comprises the luciferase from the firefly (Fluc) and its small molecule substrate, D-luciferin [3,7]. When introduced into non-luminescent cells, these components produce photons that can be captured by sensitive cameras. The Fluc/D-luciferin pair has been widely used for tracking cells and gene expression patterns in mouse models of human biology [3,7-9]. Synthetic analogs of D-luciferin are also gaining traction in imaging studies, owing to their unique light emission profiles and cell permeabilities [10-15].

In previous work, the Prescher lab exploited a modified luciferin—known as a “caged” probe—to visualize cell-cell proximity in rodent models [5]. “Caged” luciferins typically comprise appendages (i.e. “cages”) at the 6′-position of the scaffold, rendering the molecules incapable of interacting with luciferase to produce light [16]. However, upon removal of the cage, a functional luciferin (with an electron-donating group at the 6′-position) is generated and available for the light-emitting reaction. “Caged” probes have been previously employed to measure enzyme activities [17-18], track bioactive small molecules [19-20], and alter gene expression via Cre-recombinase [21-22] in vitro and in vivo.

The initial probe explored was Lugal, a galactose-caged luciferin that is selectively cleaved by the uncaging enzyme, β -galactosidase (β -gal) [5]. If β -gal is expressed in one cell (i.e., an “activator” cell) and Fluc is expressed in another (i.e., a “reporter” cell), administration of Lugal can report on the proximity of the two cells (Fig. 1) [23]. In this scenario, luciferin released by activator cells enters neighboring reporter cells and is used by Fluc to produce light (Fig. 1B) [5]. While Lugal is able to report on relative distances between cell populations, the molecule is prone to non-specific uncaging in biological media [5]. Premature uncaging results in luciferin release and bioluminescent light emission in regions devoid of activator cells. Thus, sensitive imaging of cellular interactions with Lugal remains challenging. Additionally, Lugal enables only transient imaging. In some cases, long-term visualization of cell-cell contacts is desired. For example, studies focused on long-term tracking of small subsets of immune cells or implanted stem cells would benefit from a sensitive and permanent record of cell-cell contacts.

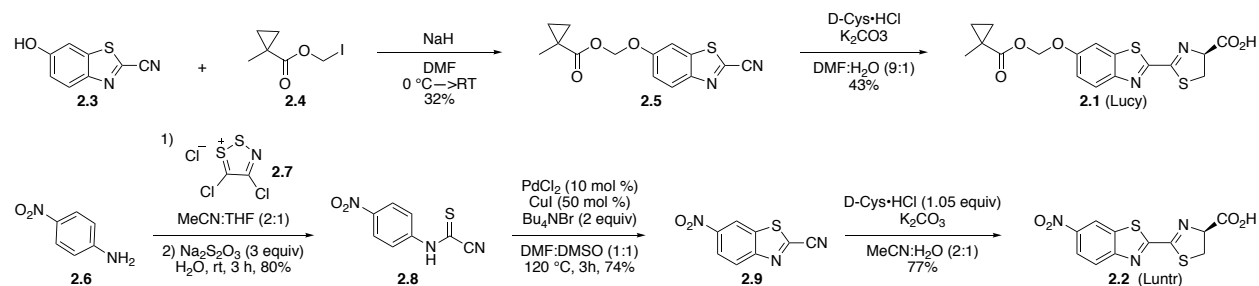
This chapter describes my efforts to develop improved caged molecules. I initially focused on more stable luciferin cages with increased selectivity for their activating enzymes. The luciferins pursued were a 6'-nitro cage (Luntr) and a methylcyclopropane carboxymethyl cage (Lucy) with activating enzymes nitroreductase (NTR) and porcine liver esterase (PLE), respectively. A complementary caged reporter system based on an activatable Cre-recombinase was also pursued. This caged-activator Cre-recombinase system would permanently turn on luciferase expression in cells contacted by a cell type of interest allowing for long-term studies and a recording of all cell-cell contacts.

2.2 Design and synthesis of ‘caged’ luciferins

I initially aimed to develop alternative caged luciferins with improved robustness and specificity. The ideal molecules would be readily accessible from common synthetic procedures, produce little to no light with Fluc, and, importantly, be resistant to non-specific uncaging. Based on these considerations, we were drawn to two luciferins, Lucy (**2.1**) and Luntr (**2.2**) (Fig. 1C). Lucy comprises a methylcyclopropane carboxymethyl ester as a cage. This appendage has been previously used as a cage that can be selectively cleaved by Porcine Liver Esterase (PLE) for fluorescein and other fluorophores. In these cases, the masked fluorophores were stable and remained “dark” in cells and in the presence of other esterases [24]. The increased stability of the cyclopropyl methyl ester, as compared to other esters, was attributed to more favorable overlap between the sigma orbitals of the cyclopropane and the π -system of the carbonyl [25-26].

The second caged probe, Luntr (**2.2**), comprises a nitro group at the 6-position and is a candidate substrate for bacterial nitroreductase (NTR). These enzymes have been widely used to activate prodrugs [27-31] and fluorophores for imaging [32-33]. Most recently, NTR was used to uncage near infrared fluorophores for imaging cancer metastases [32]. Neither NTR nor PLE are expressed in human or rodent cells, suggesting these enzymes would be suitable for selective uncaging in our approach. Furthermore, we reasoned that both Lucy and Luntr were likely to be poor light emitters in their caged forms. Robust luciferin emitters require an electron-donating group at the 6-position [34]. PLE and NTR, activate their respective caged luciferins by restoring electron density to the six position (Fig. 1B-C): PLE by cleaving the ester linkage from

Scheme 2-1: Synthetic route to Lucy (2.1) and Luntr (2.2).



Lucy to reveal a phenol, and NTR by reducing the nitro group of Luntr to provide a hydroxyl group.

Lucy and Luntr were readily synthesized via a short and modular synthetic route developed in the Prescher lab [11]. In brief, the cyanobenzothiazole structure **2.3** was prepared from the condensation of dithiazolium chloride reagent **2.7** (Appel's salt) and *p*-anisidine, followed by a cyclization and deprotection. Lucy was derived from intermediate **2.3** and known compound **2.4** (Scheme 2.1A) [11]. Luntr (**2.2**) was prepared analogously (Scheme 2.1B) from commercially available 4-nitroaniline (**2.6**).

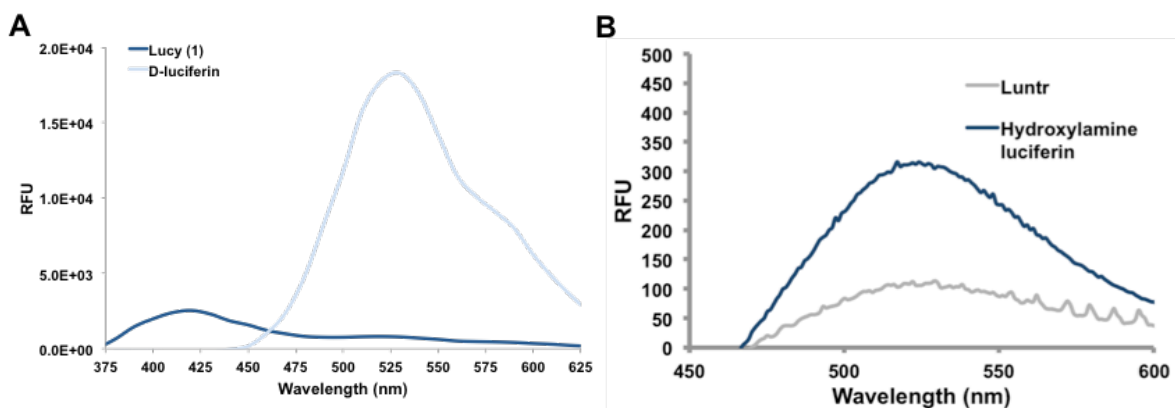


Figure 2-2. (A) Fluorescence spectra of D-luciferin (light blue) and Lucy (dark blue). Each compound was prepared at 50 μ M in PBS. An excitation wavelength of 350 nm was used. (B) Fluorescence spectra of hydroxylamine **2.10** (blue) and Luntr (gray). Each compound was prepared at 250 μ M in PBS, and an excitation wavelength of 350 nm was used. RFU is Relative Fluorescence Units.

The aniline was condensed with **2.7** and fragmented to form cyanothioformamide **2.8**. This intermediate was then cyclized [35] to provide cyanobenzothiazole **2.9** and ultimately condensed with D-cysteine to afford the caged probe **2.2**.

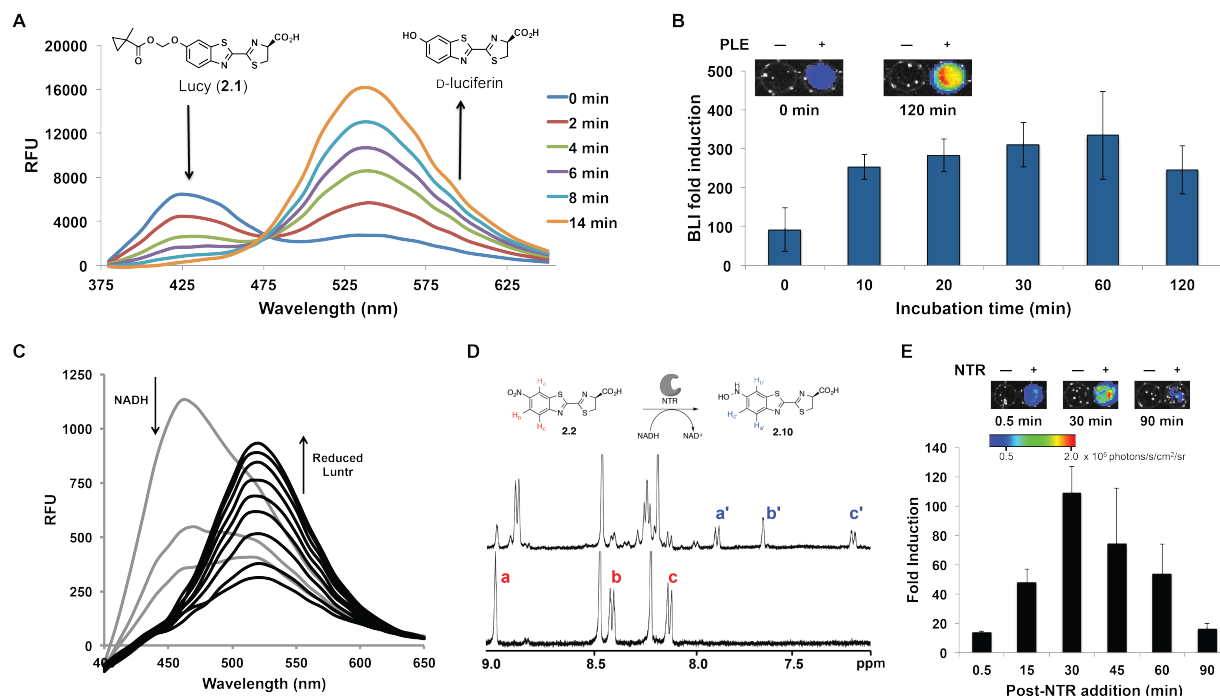


Figure 2-3. Uncaging of Lucy and Luntr in vitro. (A) Fluorescence spectra of Lucy (50 μM) incubated with PLE (10 $\mu\text{g}/\text{mL}$) and monitored over time (RFU is Relative Fluorescence Units). An excitation wavelength of 350 nm was used. (B) Lucy (500 μM) was incubated in the presence (+) or absence (-) of PLE (5 $\mu\text{g}/\text{mL}$) at 37 $^{\circ}\text{C}$ for 0-120 min. Aliquots were then added to a solution of Fluc (10 $\mu\text{g}/\text{mL}$) in BLI buffer and bioluminescent images were recorded. (C) Fluorescence emission spectra of Luntr (100 μM) incubated with NTR (2 $\mu\text{g}/\text{mL}$). The reaction was monitored over 170 min using an excitation wavelength of 350 nm. (D) Luntr (1, 5 mM) was combined with NTR (2 $\mu\text{g}/\text{mL}$) and NADH (7.5 mM) in deuterated PBS, and the reduction was monitored by ^1H -NMR spectroscopy. Sample spectra (2 h post-NTR addition) are shown, and diagnostic resonances are labeled. (E) Reduced Luntr is cell permeant. Luntr (2.75 mM) was incubated with NADH in the presence (+) or absence (-) of NTR for 0-90 min. Aliquots were collected over 90 min, transferred to wells containing Fluc+ HEK (reporter) or control cells, and imaged. Sample images are shown, and the fold induction in bioluminescent signal (from reporter vs. control cells) is plotted. Error bars represent the standard deviation of the mean for $n=3$ experiments.

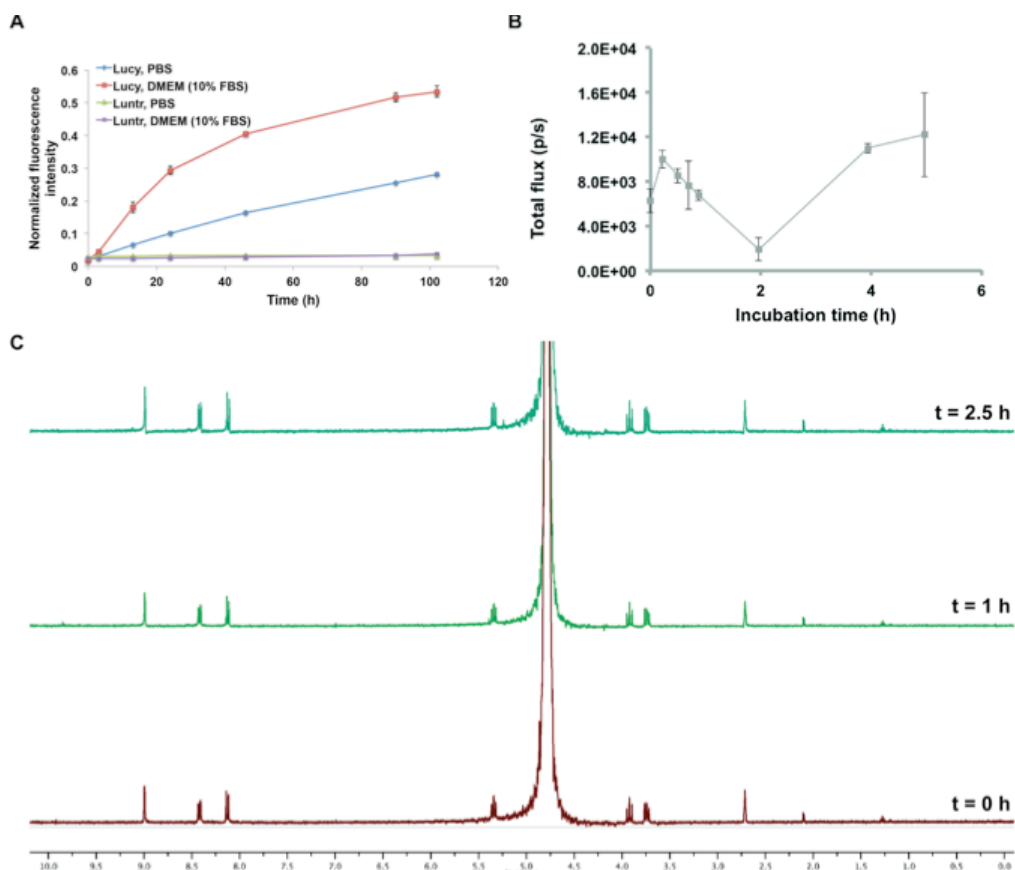


Figure 2-4. Luntr is stable in solution over time. (A) The stability of Lucy and Luntr monitored over time. The caged probes (50 μ M) was incubated at 37 $^{\circ}$ C in PBS or DMEM (10% FBS) and monitored by fluorescence (λ_{ex} 350 nm, λ_{em} 534 nm) over time. RFU is relative fluorescence units. (B) Luntr was incubated with 50,000 NTR⁻ DB7 cells for 0-5 h. Media aliquots (50 μ L) were transferred to wells containing 50,000 Fluc⁺ DB7 cells (final Luntr concentration = 250 μ M) and bioluminescence images were acquired. (C) Luntr was monitored for degradation over 2.5 h *via* ¹H NMR. The compound was dissolved in deuterated PBS (containing 10 % DMSO-*d*₆) at 25 $^{\circ}$ C.

2.3 In vitro characterization of caged luciferins

With the caged luciferins in hand, we evaluated their responsiveness to the respective uncaging enzymes, along with their overall solution stabilities. For the Lucy-PLE pair, we exploited the difference in fluorescence emission between the caged and uncaged luciferins (Fig. 2-2) [36]. When Lucy was incubated with PLE, uncaging was

observed over 14 min (Fig. 2-3A). The uncaged probe (i.e., D-luciferin) could be readily detected using its complementary luciferase, Fluc (Fig. 2-3B). At later time points, a slight decrease in bioluminescence induction was observed, possibly due to non-specific hydrolysis of the probe. However, when the hydrolytic stability of Lucy was measured in PBS or cell media (DMEM + 10% serum), a half-life of greater than 3 days in PBS and 20 h in cell media was observed (Fig. 2-4C). While not as stable as other methycyclopropyl esters [24], Lucy represents an improvement over Lugal. Chemically robust cages, as noted earlier, are critical for our approach to imaging cell contacts.

The uncaging capability and hydrolytic stability of Luntr were similarly assessed. The probe was incubated with purified NTR and NADH [27,37], and substrate uncaging was monitored via fluorescence spectroscopy. As shown in Figures 2-3C and 2-3E, uncaging was both selective and rapid, with product being detected after only 15 min of incubation. Luntr was stable in both aqueous solution and media (Fig. 2-4) for greater than 2 days.

We hypothesized that the product of Luntr reduction was hydroxylamine **2.10**, as this molecule is capable of robust light emission with Fluc (Fig. 2-5). However, we were unable to directly isolate and characterize **2.10** from the enzymatic reaction. Thus, we turned to ¹H-NMR to monitor the enzymatic reduction in situ. Luntr was incubated with NTR and NADH in deuterated PBS, and aliquots of the reaction mixture were compared to synthetic standards of hydroxylamine **2.10** and amine **2.13** (Fig. 2-5 and Scheme 2-2). NTR is known to reduce aryl nitro groups to hydroxylamines [27-30,38-39], although nitroso and amine products have been detected in a few cases [31-32,40].

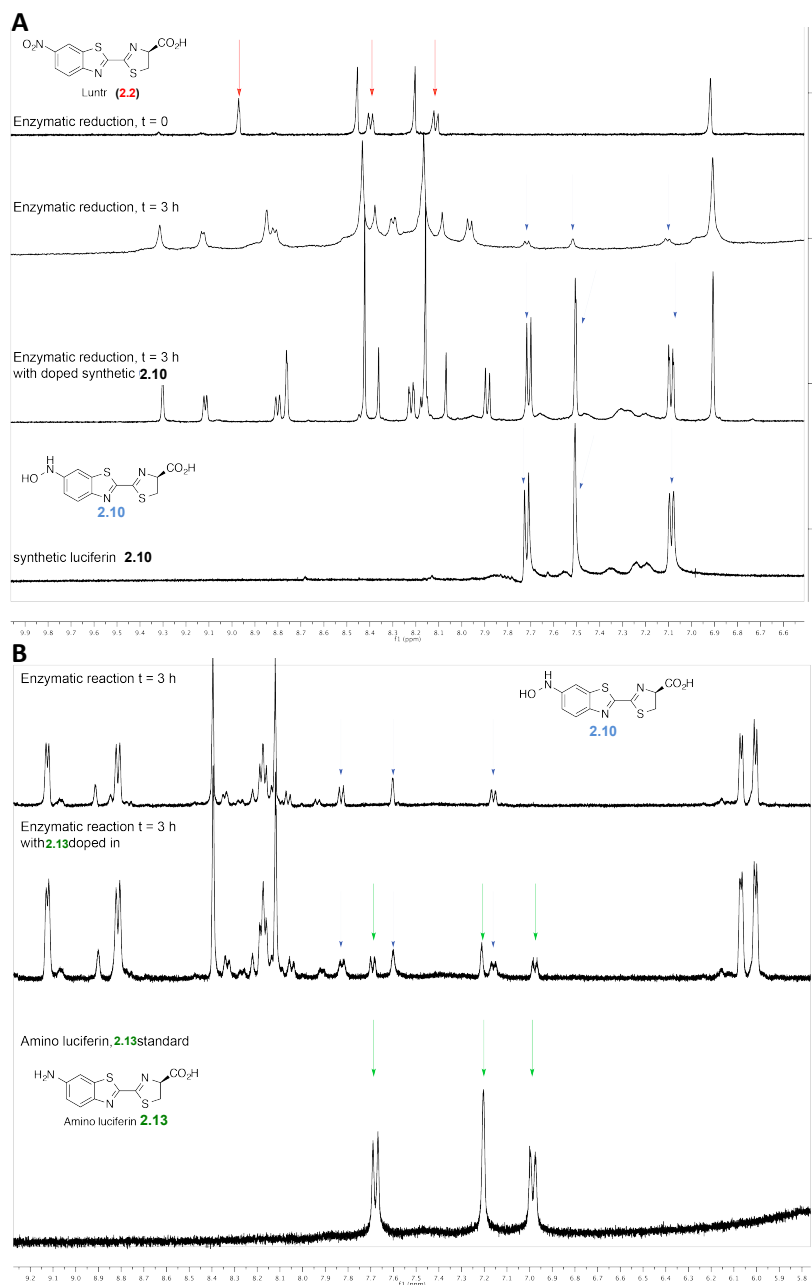


Figure 2-5. (A) NTR-mediated reduction of Luntr (**2.2**). The bottom panel shows a synthetic standard of hydroxylamine luciferin **2.10** (diagnostic resonances are labeled in blue). The top panel shows the reaction mixture before addition of NTR (diagnostic resonance for Luntr (**2.2**) are labeled in red). The second panel shows the reaction 3 h post NTR addition. The third panel shows the reaction 3 h post NTR addition with synthetic luciferin **2.10** doped into the reaction. (B) The top panel shows reduced product **2.10** after 3 h of incubation with NTR and NADH (resonances are labeled in blue). The middle panel shows amino luciferin **2.13** doped into the NTR reaction. The bottom panel shows a synthetic standard of **2.13** (diagnostic resonances labeled in green). All spectra were acquired in deuterated PBS (pH 7.4) and referenced to residual solvent. Unlabeled resonances are NADH or NAD⁺.

Scheme 2-2. Synthesis of hydroxylamine luciferin.

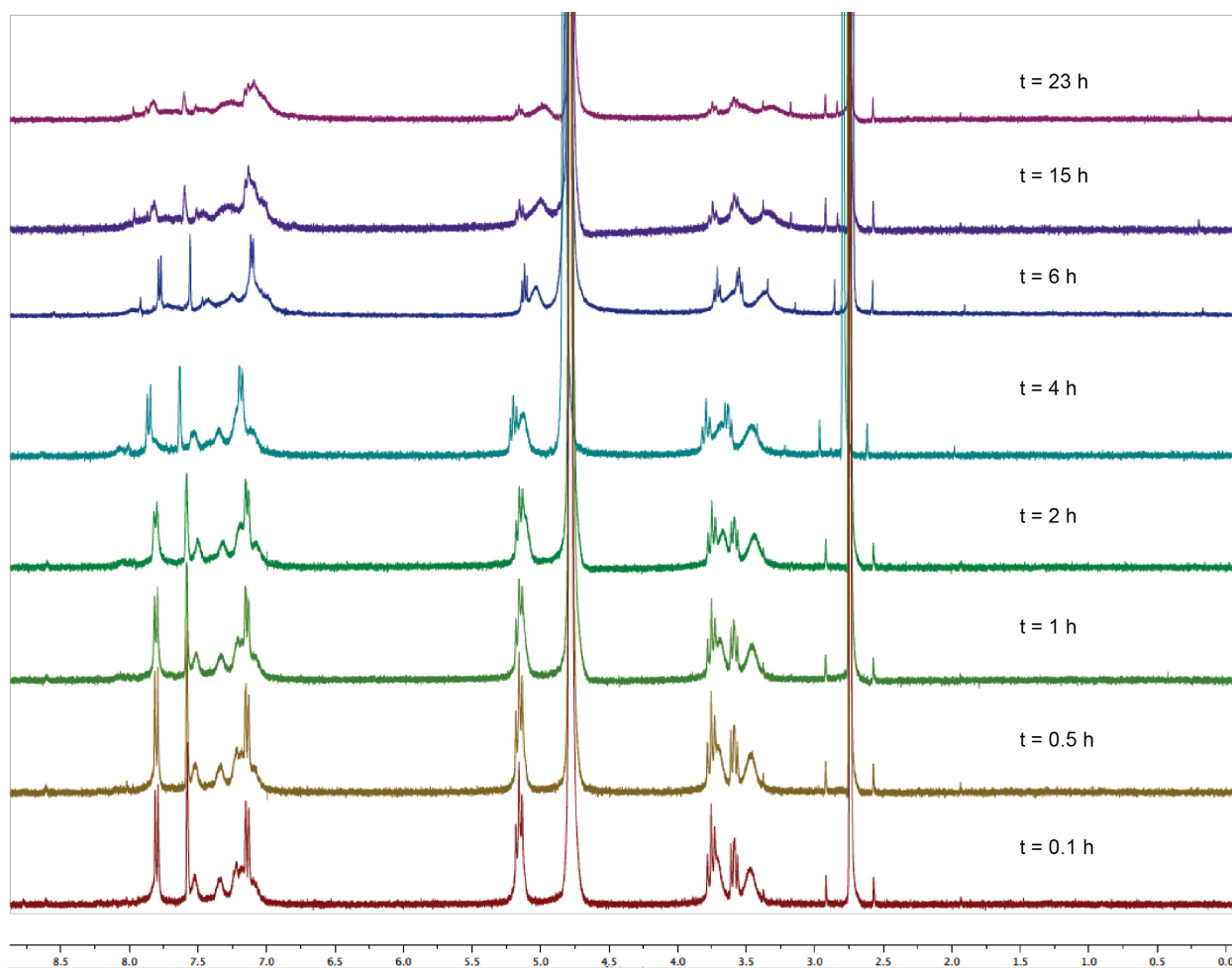
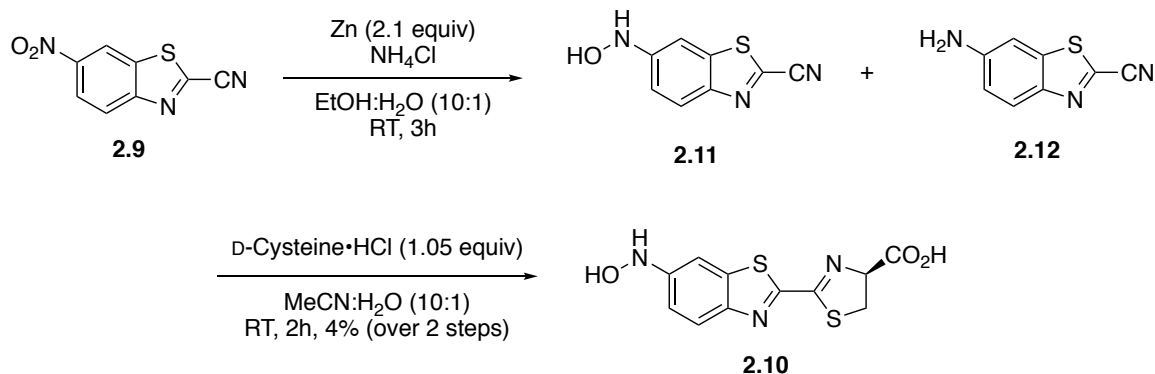


Figure 2-6. Stability of luciferin **2.10** monitored over 23 h. A 10 mM solution of **2.10** in deuterated PBS (pH 7.4) was incubated at 37 °C and monitored by ^1H NMR. DMSO was added as an internal standard.

Diagnostic proton resonances for **2.10** were observed during the first few hours of incubation (Fig. 2B), prior to compound degradation (Fig. 2-6). Formation of the fully reduced 6-amino luciferin (**2.13**) was not observed (Fig. 2-5B) during the course of the reaction. In the absence of NTR, no changes in the NMR spectra for Luntr or NADH were noted. Importantly, when the reaction mixtures (containing **2.10**) were added to purified Fluc, bioluminescent light was produced (Fig. 2-7). These data indicate that hydroxylamine **2.10** is the uncaged form of Luntr.

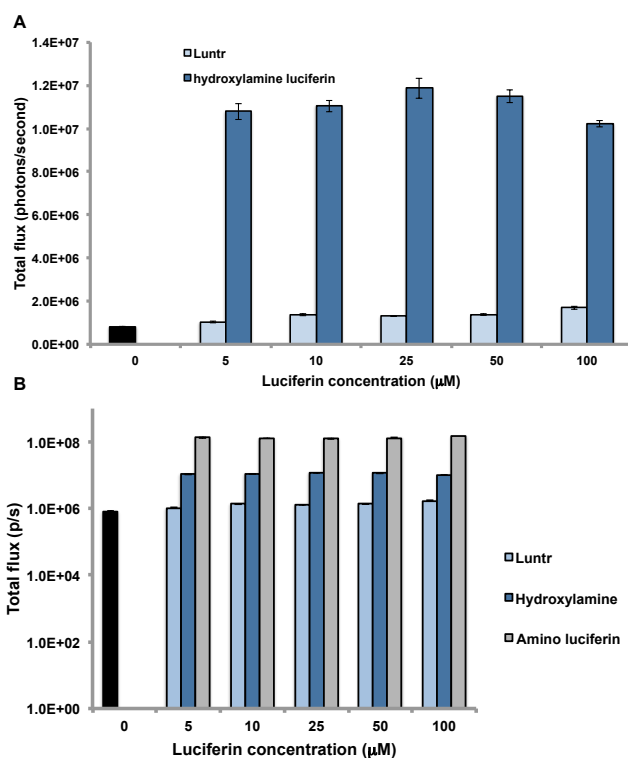


Figure 2-7. Reduced Luntr emits light with Fluc. (A) Luntr (**2.2**) and hydroxylamine (**2.10**) were incubated with Fluc and light emission was acquired. (B) Aminoluciferin (**2.13**) is a more robust light emitter than **2.10**. Luciferin samples were incubated with Fluc as above and emission data were acquired. For (A)-(B), error bars represent the standard deviation of the mean for $n = 4$ replicates. Data are representative of three independent experiments.

Uncaged Luntr was also sufficiently cell permeant to induce bioluminescent light production in luciferase-expressing (Fluc⁺) reporter cells. In these experiments, Luntr (**2.2**) was incubated in the presence or absence of recombinant NTR and NADH for 0-90 min. Aliquots of these reaction mixtures were added to Fluc⁺ reporter cells, and bioluminescent light production was measured over time. As shown in Figure 2.2E, a ~120-fold increase in bioluminescent signal was observed when Luntr was incubated with NTR and NADH for 30 min. Reduced light outputs were observed at longer incubation times, likely due to product inhibition or degradation of hydroxylamine **2.10** (Fig. 2-2E).

2.4 Examining substrate uncaging in live cells

The biochemical data prompted us to examine substrate uncaging in “activator” cells expressing the requisite enzymes. Toward this end, we cloned the PLE [24] and NTR [27] genes into HEK293 cells. The PLE vector contained an mCherry reporter and successful expression of PLE was confirmed using flow cytometry. (Fig. 2-8). To examine Lucy uncaging *in vivo*, we incubated the caged substrate with 500-25,000 activator (PLE⁺) cells or control (PLE⁻) cells for 30 min. Aliquots of the media containing the probe were then transferred to plates containing 50,000 luciferase-expressing (i.e., “reporter” cells) and imaged. Unfortunately, no increase in bioluminescent signal was observed in the presence of activator cells compared to control cells (Fig. 2-9A). Similar results were observed when Lucy was incubated with activator cells for extended time periods (Fig. 2-9B). This was likely due to non-specific uncaging of the probe by cellular esterases. PLE is also known to localize to the ER [24,41], potentially limiting Lucy’s

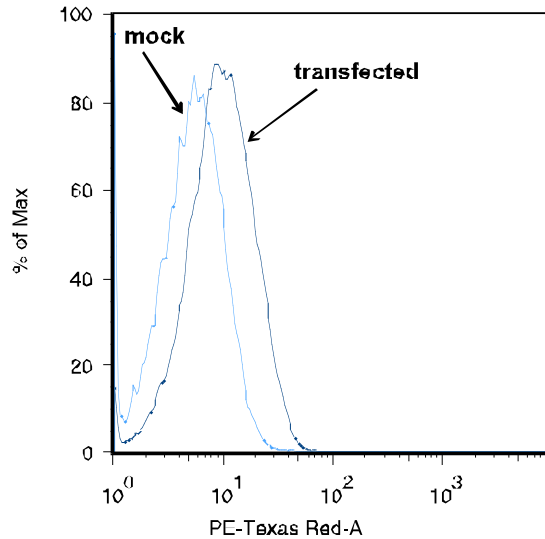


Figure 2-8. Flow cytometry analysis of PLE expression in HEK293 cells. HEK cells were transfected with a PLE-mCherry construct.

access to the uncaging enzyme and subsequent diffusion out of the cell. Removing the ER localization tag or modifying the methyl cyclopropyl ester attachment could potentially improve the performance of this probe. However, the poor stability of Lucy in cells led us to focus on Luntr in further cell-cell contact experiments.

We hypothesized that the limited release and half-life of uncaged Luntr would be beneficial for imaging cell contacts, as functional luciferin would remain localized near the NTR source. To examine this possibility, we first monitored Luntr uncaging in NTR-expressing (NTR⁺) activator cells [27]. Activator cells were seeded together with Fluc⁺ reporter cells in a 96-well plate. When Luntr (250 μ M) was added, only a modest increase in bioluminescent signal was observed in the mixed cultures compared to reporter cells incubated with control (NTR⁻) cells (Fig. 2.10A). Similarly weak signals were observed when Luntr was first incubated with NTR⁺ cells, followed by media transfer to reporter cells (Fig. 2-11A). Previous experiments indicated that reduced Luntr

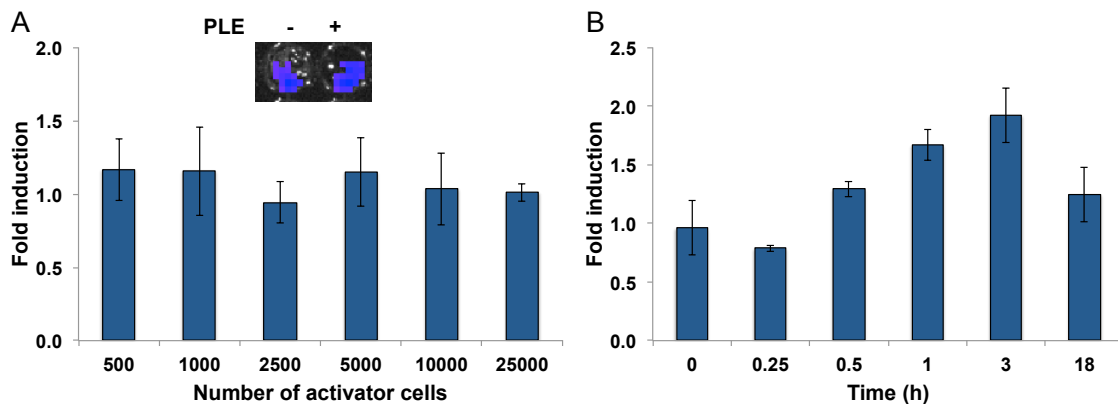


Figure 2-9. Cell culture assays with Lucy (A) No selective uncaging was observed when Lucy was incubated with activator or control cells. Activator (HEK PLE⁺) or control (HEK PLE⁻) cells (500-25,000 cells/well) were plated and incubated with Lucy (500 μ M) for 30 min and transferred to wells containing 50,000 HEK Fluc(reporter) cells. The inset is a representative bioluminescent image of Lucy incubated with 5,000 activator or control cells. (B) Minimal fold induction was observed over 3 h upon incubating Lucy with HEK PLE expressing cells versus control cells. Lucy was incubated with 50,000 HEK PLE⁺ or PLE⁻ cells, aliquots were transferred (final Lucy concentration, 250 μ M) to wells containing 50,000 HEK luciferase cells and bioluminescent images were acquired. Experiments were performed in triplicate and are representative of multiple experiments. Error bars indicate the standard deviation of the mean (n=3).

2.10 was sufficiently cell permeant to enter Fluc⁺ cells (Fig. 2-3E), and NTR enzyme assays suggested that uncaging activity was present in the activator cells (Fig. 2-11B). Thus, the low signals were likely due to limited diffusion of **2.10** out of the NTR⁺ activator cells. To test this hypothesis, we generated cells that stably expressed Fluc directly fused to NTR (Fluc-NTR). In these cells, reduced Luntr can be immediately processed by Fluc, without having to diffuse out of one cell and into another. Fluc-NTR⁺ cells were plated and incubated with Luntr (250 μ M) for 0-60 min prior to imaging.

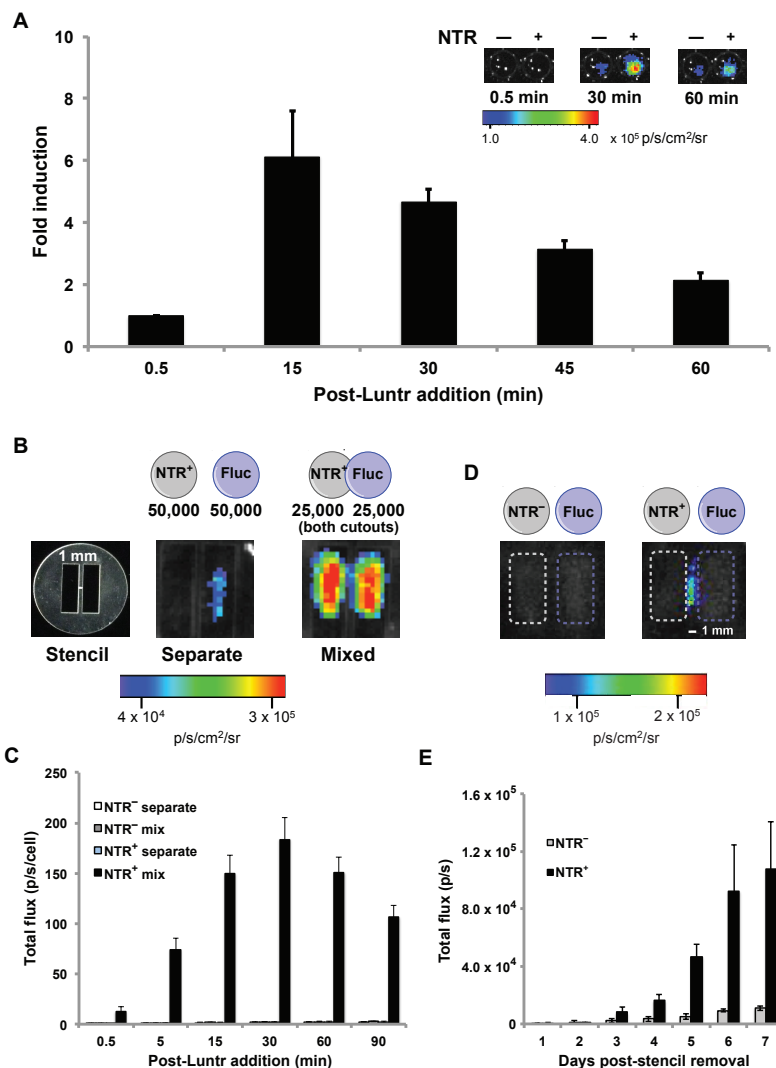


Figure 2-10. Luntr can be selectively uncaged and locally consumed. (A) Activator (NTR⁺) or control (NTR⁻) (50,000) cells were plated with 50,000 reporter (Fluc⁺) cells. Luntr (250 μ M) was added, and bioluminescence images were recorded over 60 min. Sample images are provided, and the fold induction in bioluminescent signal from NTR⁺ vs NTR⁻ cells is plotted. (B and C) Close proximity is necessary for signal induction. Fabricated stencils (left) were used to separate (blue bars) or mix (black bars) activator and reporter cells. Luntr (250 μ M) was added, and images were recorded from 0 to 90 min. NTR⁻ cells were also mixed (gray bars) or separated (white bars) with reporter cells and imaged. For (A) and (C), error bars represent the standard deviation of the mean for $n = 3$ experiments. (D and E) NTR⁻ or NTR⁺ cells were plated 1 mm apart from reporter cells (dotted lines indicate initial plating areas). The cells were allowed to grow into contact over 7 d. Images were acquired daily (25 min post-Luntr addition). Those shown are from day 6. Error bars in (E) represent the standard deviation of the mean for $n = 3$ samples. Data are representative of three replicate experiments.

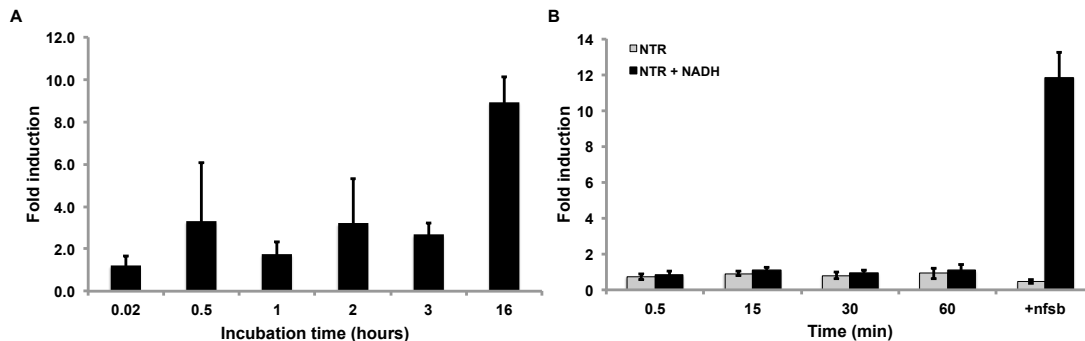


Figure 2-11. Luntr shows limited fold induction upon media transfer. (A) Luntr was incubated with 500,000 NTR⁺ or NTR⁻ HEK cells for 0-16 h. Media aliquots were transferred to wells containing 50,000 Fluc⁺ HEK cells and bioluminescent images were acquired. The fold induction in bioluminescent signal for NTR⁺ versus NTR⁻ cells is plotted. (B) NTR⁺ or NTR⁻ HEK cells were lysed and the lysate was plated with or without NADH. Luntr was then added and incubated with the cell lysates for 0-60 min. Lysate aliquots were subsequently transferred to a 96-well black microplate containing Fluc⁺ HEK reporter cells. The samples were then imaged as described previously. After 60 min, 10 µg/mL recombinant NTR enzyme was added to each well and the samples were incubated for 15 min prior to image acquisition. The fold induction in bioluminescent signal for NTR⁺ versus NTR⁻ cells is plotted. For (A) and (B), error bars represent the standard deviation of the mean for experiments performed in triplicate. The data are representative of multiple experiments.

Control experiments included mixtures of NTR⁺ and Fluc⁺ cells, along with Fluc⁺ cells plated with NTR⁻ cells. Comparable levels of NTR activity were observed in the Fluc⁺ NTR⁺ and NTR⁺ cells (Fig. 2.12). However, the Fluc-NTR⁺ cells provided a nearly 40-fold increase in bioluminescent signal compared to Fluc⁺ cells mixed with NTR⁺ cells (Fig. 2.13). These data suggest that once Luntr is uncaged, the molecule does not readily escape activator cells for use by neighboring reporter cells.

Limited release of uncaged Luntr is advantageous for imaging direct cell-cell interactions, as active luciferin should be completely consumed by the most proximal reporter cells. Fluc⁺ cells located further away should remain dark, owing to insufficient quantities of luciferin reaching distant areas. Indeed, when Fluc⁺ and NTR⁺ cells were

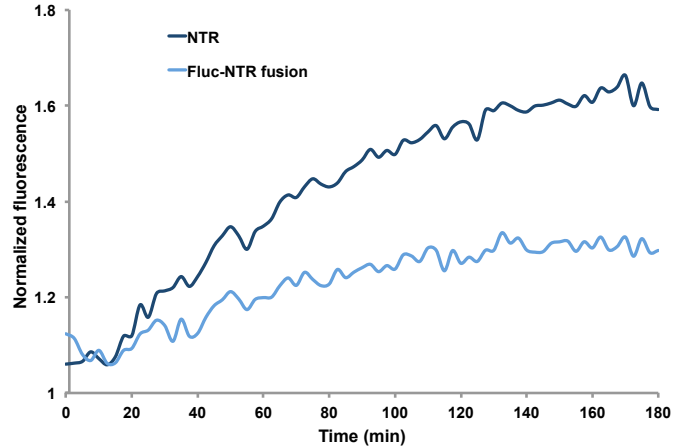


Figure 2-12. Comparable Luntr activation is observed with NTR⁺ (dark blue) and Fluc-NTR⁺ cells (light blue). Cells were suspended in PBS containing 5% FBS and plated in clear, flat-bottom 96-well plates (100,000 cells/well). Luntr (500 μ M) was added to each well, and fluorescence emission spectra (λ_{ex} = 350 nm, λ_{em} 525 nm) were recorded every 5 min for 180 min. Assays were performed in triplicate and are representative of two independent experiments. The data are normalized to Luntr incubated with HEK293 (NTR⁻) control cells only.

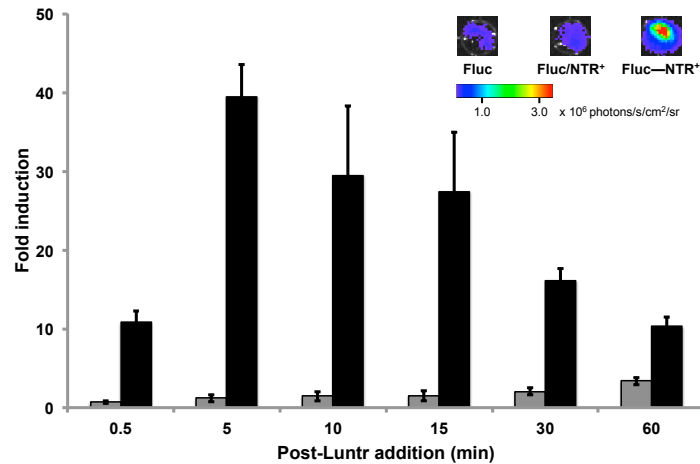


Figure 2-13. Larger bioluminescence signal inductions observed when cells are engineered to express a Fluc-NTR fusion enzyme. Fluc⁺ cells (100,000 cells/well) were plated together with NTR⁺ (100,000) cells (gray bars) as a comparison and Fluc-NTR cells (100,000 cells/well) were plated by themselves (black bars). Fold induction is compared to Fluc⁺ (100,000 cells/well) plated with NTR⁻ control cells (100,000 cells/well). In all cases, cells were incubated with Luntr (250 μ M) for 0-60 min and bioluminescent images were acquired. Experiments were performed in triplicate and are representative of at least 2 biological replicates. Error bars represent the standard deviation about the mean.

plated in direct contact, light emission was readily observed following Luntr addition (Figs. 2-10A, 2-14A-C). When separated by just 1 mm, though, Fluc⁺ and NTR⁺ cells produced little light in the presence of Luntr (Fig. 2-10B-C). Similar trends were observed using another cell type (Fig. 2-14A), when cells were separated by different

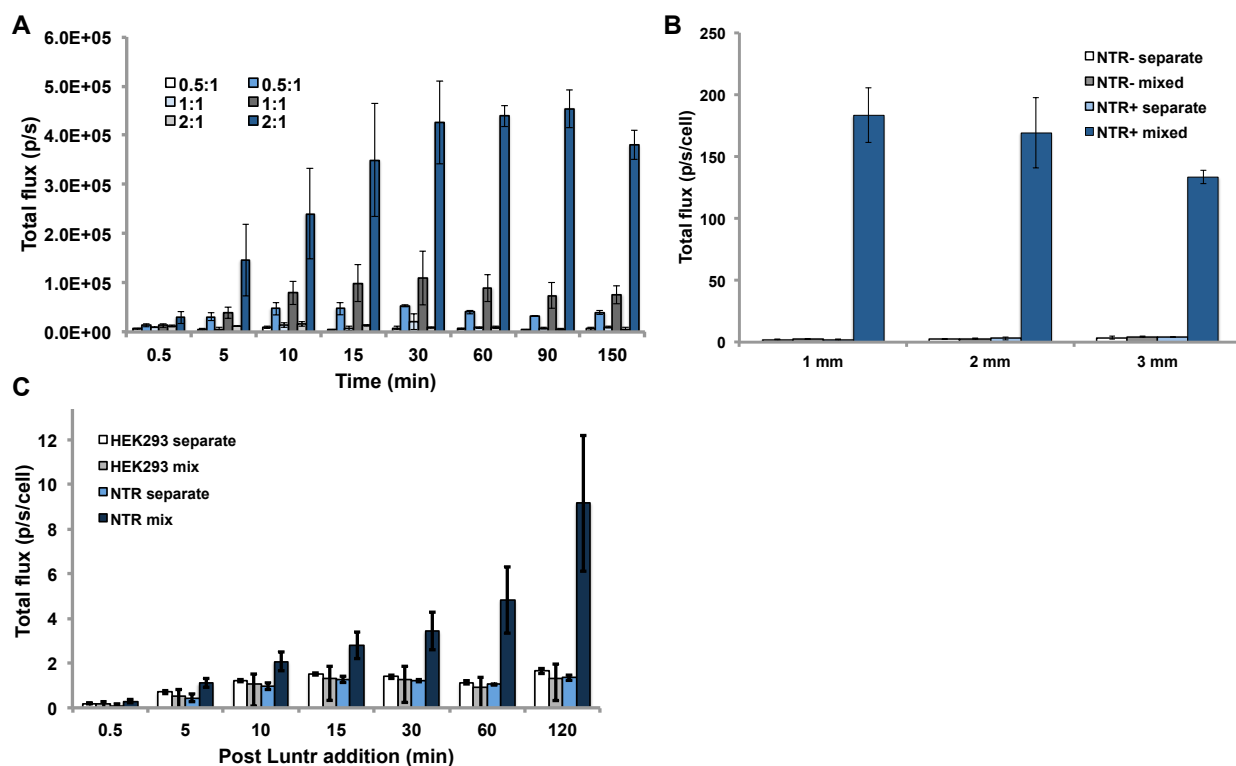


Figure 2-14. (A) Imaging DB7 cell-cell interactions. DB7 NTR⁺ or NTR⁻ cells were plated with DB7 Fluc⁺ cells in 0.5:1, 1:1, or 2:1 ratios. The cells were allowed to adhere for 12 h. Luntr was then added to each well (final concentration of 250 μ M) and images were acquired over 150 min. The experiment was carried out in triplicate and error bars represent standard deviation of the mean. The data are representative of two biological replicates. (B) HEK293 cells in close contact are visualized. Fabricated stencils were used to separate (1-3 mm) or mix cells as described in figures 2-10 and 2-15. Luntr (250 μ M in media) was flowed over the cells and bioluminescent images were acquired after 30 min. (C) Increased light emission is observed when NTR and Fluc cells are imaged immediately post plating. Fabricated stencils (see Figure 2-10B-C) were used to separate activator and control cells (light blue) by 3 mm or contain mixed populations (dark blue) of cells. Immediately post-plating Luntr was added to a final concentration of 250 μ M. Bioluminescence images were then recorded from 0-120 min. HEK NTR⁻ control cells were also plated with reporter cells (gray) and separate (white) in the same manner. The experiment was carried out in quadruplet and error bars represent the standard deviation of the mean. Data are representative of five replicate experiments.

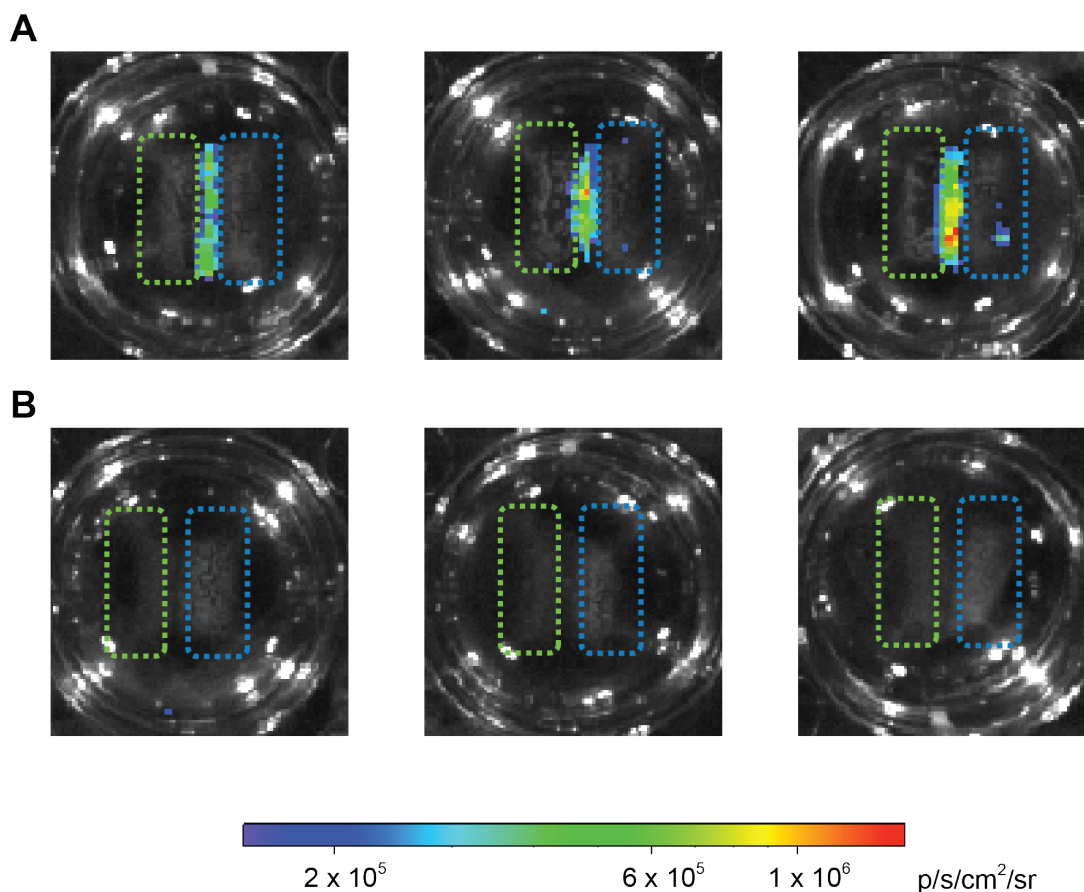


Figure 2-15. Luntr can report on cell-cell contacts. Repeat images for the experiment described in Figure 2-10D-E. Cells were plated 1 mm apart from Fluc⁺ cells and allowed to grow into contact over 7 d. Dotted lines indicate the initial plating areas (green = NTR⁺ (A) or NTR⁻ (B) cells, blue = Fluc⁺ cells). Images were acquired daily following Luntr addition. The images shown above were acquired 7 d post-stencil removal.

distances (Fig. 2-14B), or when cells were imaged immediately post-plating versus, adhering overnight (Fig. 2-14C). Robust light emission was also observed when Fluc⁺ and NTR⁺ cells were allowed to grow into contact (Figs. 2-10D-E, 2-15, 2-16). In all cases, no cell death was observed upon trypan blue staining. The stringent dependence of light emission on cellular distance is ideal for use in tissues and other complex environments, as only cells in close contact should produce light.

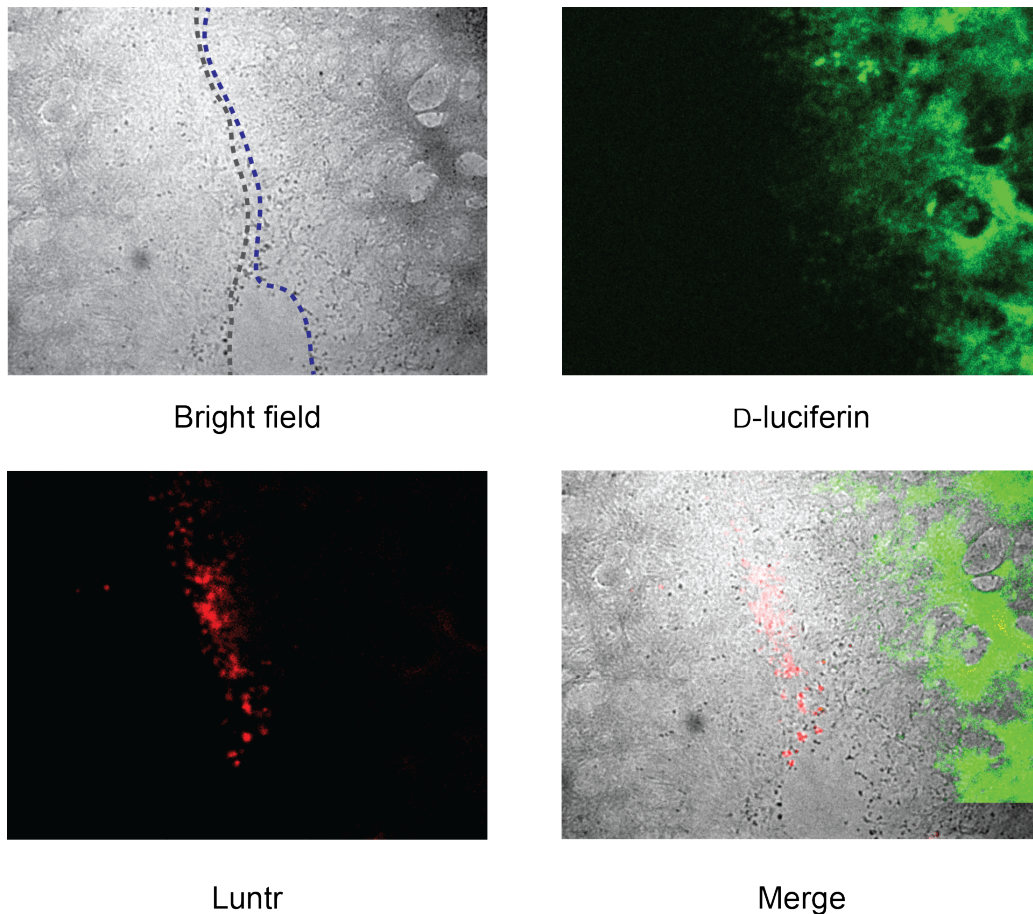


Figure 2-16: Bioluminescence microscopy with Luntr indicates cell-cell interactions. Cells were plated as described in figure 2-10D. The brightfield image shows the rough cell demarcation (DB7 NTR⁺, left, gray; DB7 Fluc⁺, right, blue). D-Luciferin (250 mM) was added and the bioluminescent image was acquired over 10 min (luminescent signal is false-colored green). Luntr (250 mM) was added and bioluminescent images were acquired over 10.5 h (luminescent signal is false colored red). Images were processed using ImageJ software. D-Luciferin and Luntr false color images were overlaid on a brightfield image acquired prior to Luntr imaging. The x, y position of the false color D-luciferin image was manually adjusted in ImageJ to overlay with the corresponding cells.

2.5 Towards creating a permanent record of cell-cell contacts with a caged Cre-recombinase activator

Many cell-cell interactions depend on long-term contact between two populations that can be imaged by the caged luciferin strategy described previously. However, this strategy is restricted to cell-cell contacts occurring during the imaging session.

Interactions that rely on transient cell-cell contacts could be missed. In addition, the sensitivity of the caged luciferin strategy is limited as one uncaged luciferin molecule produces, at most, one photon. A caged probe with increased sensitivity that would induce a permanent turn on of Fluc could capture these missed interactions. Activation of the gene encoding luciferase would boost the sensitivity and also produce a record of any cell-cell contacts.

Activating genes for imaging purposes has been demonstrated with various Cre-recombinase systems in cells and in vivo. Some examples include the stochastic expression of fluorescent proteins [42-43], luciferase expression [44], and even using light to turn on activators of Cre recombinase [21-22,45-46]. Recently Cre-ER was turned into a light activatable gene-editing tool [21-22,46] by using photo-cleavable cages on the phenol of a tamoxifen analogue. Due to this precedent, we pursued a caged tamoxifen analogue that can be activated by nitroreductase (similar to Luntr) [47-49]. This tamoxifen analogue when uncaged binds to a Cre estrogen receptor (Cre-ER) fusion. Binding allows Cre to enter the nucleus and turn a reporter gene on, luciferase in this case. Adapting a caged tamoxifen analogue to report on cell-

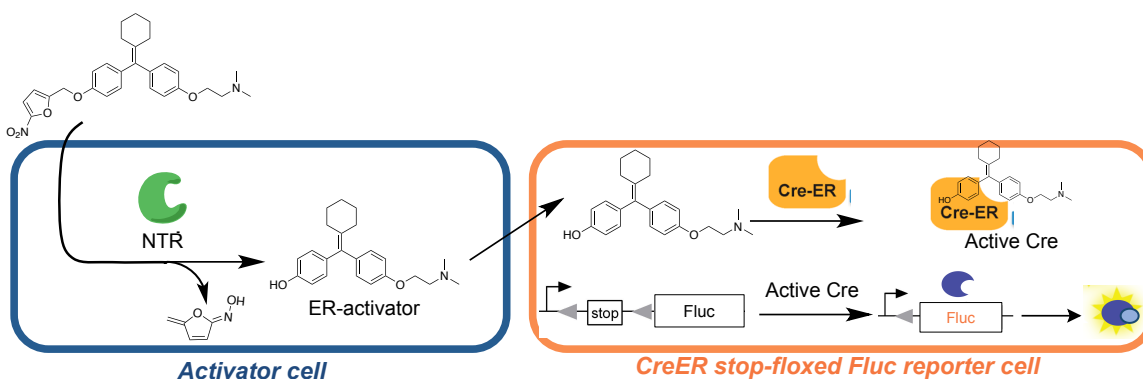
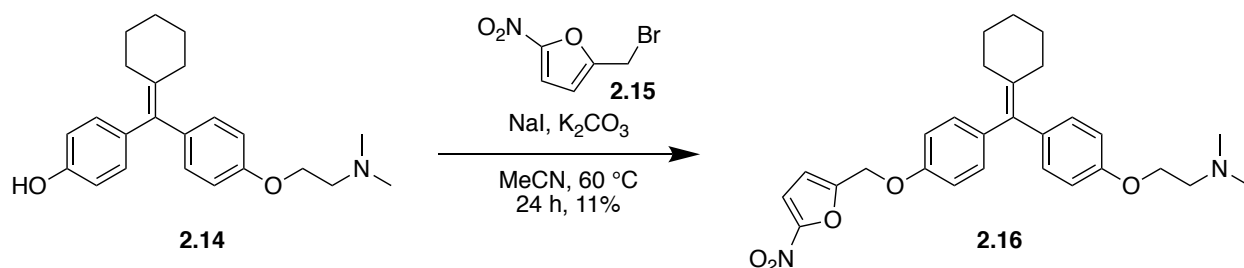


Figure 2-17. Cartoon depiction of cell-cell contact dependent luciferase expression. The caged cyclofen-OH is activated by NTR in the activator cell where it can then diffuse into the CreER stop-flaxed Fluc reporter cell and turn on luciferase expression.

cell contact would rely on a similar strategy as pursued using caged luciferins. The caged Cre activator is enzymatically uncaged (by NTR) and when activating cells are in close proximity to a reporting cell the uncaged molecule will diffuse into reporting cells [5,50]. Once in the reporter cells the molecule can bind to Cre-ER and allow for recombinase to turn the luciferase reporter on (Fig. 2-17).

Scheme 2-3. Synthetic route to cyclo-NTF.



We pursued the tamoxifen analogue cyclofen-OH (**2.14**), due to its ease of synthesis and stability [51]. Caging at the phenol group disrupts binding to the ER domain precluding Cre-ER from entering the nucleus [51-52]. We hypothesized a cleavable cage on the same position would similarly inhibit binding to Cre-ER in the absence of the activating enzyme. 5-Nitrofurans, cleavable by NTR, has been utilized for caging fluorophores and also D-luciferin [38,47-48,53] and is similar in size to cyclofen-OH photocages. With these parameters in mind, we set out to synthesize cyclofen-OH caged by the NTR active 5-nitrofurans-2-ylmethyl, or cyclo-NTF (**2.16**) [48-49].

The synthesis of cyclo-NTF followed literature precedent to access cyclofen-OH [46,51]. We first reduced 5-nitro-2-furaldehyde with aluminum isopropoxide to yield the alcohol at the 2-position. Initial attempts to couple the furan cage and cyclofen-OH under various Mitsunobu conditions were unsuccessful. Eventually a successful route was identified utilizing the known 2-bromomethyl compound (**2.15**). This compound was appended to **2.14** via nucleophilic displacement [47] of the bromide to yield cyclo-NTF (**2.16**). With this compound in hand, its ability to be activated by NTR and subsequently turn on luciferase expression via Cre-ER will be tested.

2.6 Conclusion

In this Chapter, the design and synthesis two novel caged luciferins, Lucy and Luntr, were described. Their stability and uncaging activity was characterized in vitro and in cells. Both Lucy and Luntr were found to be efficiently uncaged by their respective activating enzymes, PLE and NTR. However only Luntr was found to be sufficiently stable for cell studies. NTR was expressed in various cell lines and used to activate Luntr in mixed cell cultures. The limited release and lifetime of uncaged Luntr proved advantageous for visualizing cells in close contact. We anticipate that the technology will be most advantageous for imaging cell-cell interactions where the largest numbers of interacting cells are expressing NTR. The caged luciferin approach is important for capturing real-time cell-cell contacts. However, a different approach is necessary if a permanent record of past cell-cell contacts is desired. Towards this end progress on a Cre recombinase system was also described. This approach will enable more sensitive detection of small populations for cell-cell contacts in addition to a

permanent record of past cell-cell interactions in cell culture and in vivo. The development of imaging tools to visualize cell contacts addresses a void in our ability to “see” microscopic events at the macroscopic level.

2.7 Materials and methods

2.7a General biological methods

Buffer salts and dNTPs were purchased from Fisher Scientific. All restriction enzymes and T4 DNA ligase were purchased from New England Biolabs. DNA polymerase (DreamTaq) was purchased from Thermo Scientific. All PCR reactions were analyzed using agarose gel electrophoresis with ethidium bromide staining. For all experiments, Luntr stock solutions (30 mM) were prepared by dissolving the luciferin in DMSO (100 mM), followed by dilution with PBS to 30 mM (final concentration of 30% DMSO). Luntr solutions were used immediately or stored at -20 °C and sonicated for 30 min prior to use. NADH was purchased as the disodium salt from Research Products International Corp.

2.7b Fluorescent spectra and assays

All fluorescence spectra, compound stability and uncaging assays were performed using clear 96-well microplates (BD-Falcon or Grenier-Bio) and 100 µL sample volumes. The plates were incubated at 37 °C for 0-100 h, and measurements were acquired using a Molecular Devices SpectrMax Gemini XPS microplate reader. All

measurements were performed in triplicate. Fluorescence data were exported and analyzed using Microsoft Excel.

2.7c UV-Vis spectra

Absorption spectra were acquired on a the Thermo-Fisher NanoDrop spectrophotometer (ND-2000c) using 250 μ M luciferin solutions. Absorbance data were exported and analyzed using Microsoft Excel. All absorption spectra are representative of three independent experiments.

2.7d General bioluminescence imaging

All biochemical and cultured cell assays were performed in black 96-well plates (Grenier Bio One). Plates containing luminescent reagents were imaged in a light-proof chamber with an IVIS Lumina (Xenogen) CCD camera chilled to -90 °C. The stage was kept at 37 °C. The camera was controlled using Living Image software. Exposure times were set to 10 s for biochemical assays, with the data binning levels set to medium. The exposure times for cultured cell assays ranged from 10-60 s with data binning levels set to medium or large. Regions of interest were selected for quantification and total flux values were analyzed using Living Image software.

2.7e Cloning and plasmid construction

E. coli nfsB gene isolation and cloning

The NTR gene (*nfsB*) was isolated following the procedure of Friedlos *et al.*¹ In brief, the gene was cloned from the *E. coli* B (Berkeley strain) using the forward primer 5'-

ACTTTCAAAGCTTCCACCATGGATATCATTTCT-3' and the reverse primer 5'-TATGATGAGGATCCAAACAGGGTTATGCAA-3'. The insert contained the restriction sites *Bam*HI and *Hind*III along with a Kozak sequence at the start of the NTR gene. This NTR gene was then cloned into the pcDNA3.1 vector downstream of the CMV promoter.

pET-NTR plasmid

The NTR gene was cloned from the previously made pcDNA3.1-NTR vector using the forward primer 5'-CATGAATTCATGGATATCATTTCTGTGCGCC-3' and the reverse primer 5'-TACCTCGAGGACTTCGGTTAAGGTGATGTTT-3'. The insert contained *Eco*RI and *Xho*I restriction sites and the Kozak sequence was removed. This gene was cloned into the PET vector downstream of the T7 promoter and in-frame with a C-terminal His₆ tag.

Fluc-NTR fusion plasmid

The NTR gene was cloned from the previously described pcDNA3.1-NTR vector using the forward primer 5'-TATAGGATCCATGGATATCATTTCTGTGCGCC-3' and the reverse primer 5'-TATGATGAGGATCCAAACAGGGTTATGCAA-3'. The insert contained *Bam*HI and *Xho*I restriction sites and the Kozak sequence was removed. The insert was then placed into a pcDNA-Fluc plasmid downstream of the Fluc gene and a short ASAAGS linker.

NTR-IRES-eGFP and Fluc-IRES-eGFP (pBMN)

The NTR gene was cloned from the previously described pET-NTR plasmid using the forward primer 5'-TATAGGATCCATGGATATCATTCTGTGCGCC-3' and the reverse primer 5'-TACGCGGCCGCTTAGACTTCGGTTAAGGTGATGTTTTGC-3'. The Fluc gene was amplified from the Fluc (pg14 "luc2")-IRES-eGFP gene (courtesy of the Contag lab, Stanford) with the forward primer 5'-TATGGATCCATGGAAGATGCCAAAAACATTAAGAAGGGCC-3' and the reverse primer 5'-ATAGCGGCCGCTTATTACACGGCGATCTTGCCGCC-3'. The following applies to both inserts (NTR and Fluc): The insert contained *Bam*HI and *Not*I restriction sites. The insert was then placed into a pBMN-IRES-eGFP plasmid upstream of the IRES sequence (courtesy of the Nolan Lab, Stanford (Addgene plasmid #1736)).

2.7f General cell culture and transductions

HEK293 or DB7 cells (American Type Cell Culture) were cultured in complete media: DMEM containing 10% (vol/vol) FBS (Corning), 4.5 g/L glucose, 2 mM glutamine, 100 U/mL penicillin, and 100 µg/mL streptomycin (all from Gibco). Cells were maintained in a humidified incubator at 37 °C. HEK PLE⁺ cells were generated by introducing a PLE-mCherry IRES gene expressed from the pCAG plasmid (courtesy of the Lavis Lab, Janelia Farm Research Campus) via infectious retroviral particles created using Phoenix Ampho cells (a generous gift from the Nolan Lab at Stanford University). Selection was performed using G418 (100 µg/mL) and mCherry expression was confirmed by flow cytometry. Cells (1x10⁶ HEK PLE⁺ or HEK 293) were harvested, centrifuged and suspended in FACS buffer prior to flow cytometry analysis on a BD LSR

II instrument. Flow cytometry data were analyzed using FloJo (TreeStar) software. NTR-expressing (NTR⁺) HEK293 cells were generated by introducing the pcDNA3.1 NTR construct via cationic lipid transfection (Lipofectamine 2000, Invitrogen). Cells stably expressing the transgene were selected over 14 d using G418 (MediaTech Inc., 100 µg/mL). HEK cells expressing the Fluc-NTR fusion were generated by introducing the pcDNA3.1 Fluc-NTR construct via cationic lipid transfection (Lipofectamine 2000, Invitrogen) as above. Cells stably expressing the transgene were selected over 15 d using G418 (MediaTech Inc., 500 µg/mL). NTR⁺ and Fluc⁺ DB7 cells were prepared *via* retroviral transduction (Phoenix-Eco) of pBMN vectors (GFP⁺) containing the relevant genes. The cells were sorted *via* FACS at the CIRM Stem Cell Research Center Flow Cytometry Core (UCI). Fluc⁺ HEK293 cells were courtesy of the Contag laboratory (Stanford University).

2.7g NTR expression and purification

The pET-NTR plasmid was transformed into chemically competent *E. coli* BL21 cells. Overexpression and purification were performed following the procedure of Olekhovich, *et al.*² The purified protein was dialyzed using 50 mM Tris•HCl, 50 mM NaCl (pH 7.5) buffer and lyophilized prior to storage.

2.7h Fluorescence uncaging assays

Luntr uncaging by recombinant NTR

Luntr (100 µM) was incubated with NADH (150 µM) and purified NTR (2 µg/mL). Assays were performed in PBS using clear 96-well microplates. All incubations and readings

were performed at 37 °C. Fluorescence spectra were acquired every 10 min over 170 min. An excitation wavelength of 350 nm was used. Assays were performed in triplicate, and the data shown in Fig. 2-2B and 2-3C are representative of three independent experiments. Data were analyzed using Microsoft Excel.

Luntr uncaging by NTR expressing cells

HEK293, NTR⁺ HEK and Fluc-NTR⁺ HEK cells were plated in a 96-well microplate (100,000 cells/well in PBS containing 5% FBS). Luntr (500 µM) was added to each well and the fluorescence emission at 525 nm (with an excitation wavelength of 350 nm) was recorded every 2.5 min over 180 min. The cells were incubated at 37 °C for the duration of the experiment. Assays were performed in triplicate, and the data shown in Fig. 2-12 are representative of three independent experiments. Data were analyzed using Microsoft Excel.

2.7i Enzymatic bioluminescent assays

Bioluminescent assays with uncaged Luntr and Fluc⁺ HEK293 cells

Luntr (2.75 mM) was incubated in the presence (+) or absence (-) of NTR (20 µg/mL) and NADH (3 mM) for 0-90 min. Aliquots were transferred to a black 96-well plate containing Fluc⁺ HEK cells (50,000 cells/well) and bioluminescent images were acquired over 90 min as described above. Assays were performed in triplicate and the data in Fig. 2-3E are representative of two independent experiments. Data were analyzed using Microsoft Excel.

Bioluminescent assays with synthetic luciferin **2.10** and Fluc

Luciferin **2.10** was purified and immediately used for bioluminescent assays. Stocks of all luciferins (10 mM in PBS) were diluted to the final imaging concentrations of 5, 10, 25, 50 and 100 μ M. Imaging was performed in BLI buffer as described above. Images were acquired in 96-well black plates six minutes after the addition of Fluc (1 μ g/well). Bioluminescent images were otherwise acquired as above. Assays were performed in quadruplicate, and the data shown in Fig. 2-7 are representative of three experiments.

2.7j Cell culture bioluminescent assays

PLE⁺ HEK Lucy uncaging assays

HEK PLE⁺ (activator) or HEK293 (control) cells were plated in 96-well black microplates (500-50,000 cells/well). Lucy was incubated for 0-18 h at 37 °C. Aliquots of media, to result in a final concentration of 250 μ M Lucy, were subsequently transferred to a 96-well plate containing HEK luc2 reporter cells (50,000 per well) and imaged as described previously. Images were acquired 3-5 minutes post substrate administration.

NTR⁺ HEK Luntr uncaging assays (media transfer)

NTR⁺ (activator) or HEK293 (control) cells were plated in black 96-well plates or 24-well clear tissue culture plates (BD Falcon, 500,000 cells/well) and incubated with Luntr (500 μ M) for 0-16 h. Aliquots of the media (50 μ L) were subsequently transferred to a black 96-well plate containing Fluc⁺ HEK reporter cells (50,000 per well) to a final volume of 100 μ L. Images were acquired 5 min post-substrate addition as above. Assays were

performed in triplicate, and the data shown in Fig. 2-9 are representative of at least three independent experiments. Data were analyzed using Microsoft Excel.

NTR⁻ DB7 Luntr stability assay

NTR⁻ cells were plated in black 96-well plates (50,000 cells/well) and incubated with Luntr (500 μ M) for 0-5 h. Aliquots of the media (50 μ L) were subsequently transferred to a black 96-well plate containing Fluc⁺ DB7 reporter cells (50,000 cells/well). Luntr was added to a total volume of 100 μ L (250 μ M Luntr). Images were acquired 5 min post-substrate addition as above. Assays were performed in triplicate, and the data shown in Fig. 2-4B are representative of at least three independent experiments. Data were analyzed using Microsoft Excel.

NTR⁺ HEK lysate Luntr uncaging assays

NTR⁺ HEK (activator) or HEK293 (control) cells were counted, centrifuged and resuspended in PBS (5% FBS) at 0.5×10^6 cells/mL. NTR⁺ or NTR⁻ HEK (0.5×10^6 cells/mL) were lysed by freeze-thaw cycles (3x) and repeated passage through a 30-gauge syringe. The lysate was then plated (100 μ L/well) with or without 1 mM NADH. Luntr (500 μ M) was added and the lysates were incubated for 0-60 min. Media aliquots of lysate (50 μ L) were subsequently transferred to a black 96-well plate containing Fluc⁺ HEK reporter cells (50,000 per well) and imaged as above. The final volume in each well was 100 μ L (250 μ M Luntr). After 60 min, NTR enzyme (10 μ g/mL) was added to the wells and incubated for 15 min prior to image acquisition. All assays were performed

in triplicate, and the data shown in Fig. 2-11 are representative of at least two independent experiments. Data were analyzed using Microsoft Excel.

NTR⁺ HEK Luntr uncaging co-culture assays

Fluc⁺ HEK cells were plated in black 96-well plates (50,000 cells/well) with HEK NTR⁺ HEK or control (HEK293) cells (50,000 cells/well) in a total volume of 100 μ L. Luntr was then added to the relevant wells (final concentration of 250 μ M) to a total volume of 110 μ L. The plate was then incubated for 0-60 min at 37 °C. Bioluminescent images were collected as described above. All assays were performed in triplicate, and the data in Fig. 2-10A are representative of 3 independent experiments.

Fluc-NTR HEK fusion assays

Fluc⁺ HEK cells were plated in black 96-well plates (100,000 cells/well) with NTR⁺ HEK or HEK293 cells (control, 100,000 cells/well) in a total volume of 100 μ L. HEK Fluc-NTR⁺ cells (100,000 cells/well) were also plated. Luntr (250 μ M) was then added to the relevant wells (final concentration of 250 μ M) to a total volume of 110 μ L. Plates were incubated with Luntr at 37 °C over 60 min and bioluminescent images were acquired as above. All assays were performed in triplicate, and the data in Fig. 2-13 are representative of two independent experiments. Data were analyzed using Microsoft Excel.

Distance-dependent uncaging assay with fabricated stencils

Laser-cut, PDMS stencils were kindly provided by the Hui Lab (UCI). The stencils were washed in ethanol and placed in a clear 12-well tissue culture plate (BD Falcon). The stencils were allowed to dry for 24 h in a sterile environment. All cells were plated in the cutouts at 50,000 cells/cutout in a total volume of 25 μ L. The plates were incubated overnight at 37 °C to allow cells to adhere. For imaging, Luntr in DMEM with 10% (vol/vol) FBS was flowed over the stencils to a final volume of 850 μ L and final Luntr concentration of 250 μ M. The plates were then incubated at 37 °C and imaged over 0-90 min as described above. Experiments were performed in duplicate, triplicate or quadruplicate. The data shown in Fig. 2-10B-C are representative of at least five independent experiments. Data were analyzed using Microsoft Excel.

Uncaging assay with fabricated stencils immediately post-plating

PDMS stencils were set up as described above. All cells were plated in the cutouts at 50,000 cells/cutout in a total volume of 30 μ L per cutout. NTR⁺ cells were plated separate or mixed with Fluc⁺ cells. NTR⁻ (control) cells were also plated separate or mixed with Fluc⁺ cells. Immediately post-plating Luntr (10 μ L in PBS) was added to each cutout. The final Luntr concentration in each cutout was 250 μ M. Care was taken to avoid overflow or media exchange between wells. The plates were then incubated at 37 °C and imaged over 0-120 min as described above. Assays were performed in triplicate or quadruplicate. Data shown in Fig. 2-14C are representative of four independent experiments. Data were analyzed using Microsoft Excel.

Uncaging assays with varying cell ratios

DB7 Fluc⁺ cells were plated (50,000 cells/well) in a 96 black-well plate in 50 mL of media. DB7 NTR⁺ or NTR⁻ cells were then added to the wells in varying ratios, 0.5:1 (25,000 DB7 NTR⁺ or NTR⁻ cells), 1:1 (50,000 DB7 NTR⁺ or NTR⁻ cells), or 2:1 (100,000 DB7 NTR⁺ or NTR⁻ cells). Additional media was added to bring the total volume to 100 mL per well. The cells were incubated for 12 h at 37 °C prior to imaging. Luntr (10 µL) was then added to each well (final concentration of 250 µM). Cells were incubated for 0-150 min at 37 °C and images were acquired as described above. All assays were performed in triplicate or quadruplicate, and the data in Fig. 2-14A are representative of 2 independent experiments. Data were analyzed using Microsoft Excel.

Cell-cell contact growth experiment

Fabricated PDMS stencils (with 1 mm separations) were rinsed in ethanol, placed in the wells of a 12-well tissue culture plate and left overnight to dry. The next day, DB7 NTR⁺ or NTR⁻ cells (25,000 cells/cutout) were plated 1 mm apart from DB7 Fluc⁺ cells (25,000 cells/cutout). The total volume in each cutout was 30 mL (0.833×10^6 cells/mL). The cells were left overnight to adhere. The following day, 1 mL of DMEM (10% FBS) was flowed into the wells. The PDMS stencils were then removed, and the cells were allowed to grow into each other over 7-10 d. The cells were imaged each day by adding Luntr (250 mM final concentration) to each well and acquiring photons as above. Regions of interest (ROI's) were drawn in the initial zone of separation (1 mm) and photon counts were analyzed. Data depicted in Figure 2-10E, 2-15 were collected in

triplicate and are representative of three biological replicates. Data were analyzed using Microsoft Excel.

2.7k Bioluminescence microscopy experiments

Fabricated PDMS stencils (with 2 mm separations) were prepared as described above. DB7 Fluc⁺ and DB7 NTR⁺, or NTR⁻ cells were plated 2 mm apart as described above. The following day, the stencils were removed and 1 mL of media was flowed into the wells. The cells were then incubated for 6 d. Prior to imaging the media was replaced with 500 μ L PBS (pH 7.4, 8% FBS). Luntr or D-luciferin (250 μ M) was added to the well and the plate was mounted on the stage of an inverted microscope (Zeiss Axio Observer) set in a custom light-tight black box with the temperature set at 26 °C. Images were collected by a Zeiss 5x FLUAR objective (NA 0.25) and transmitted directly to a CCD camera (XR MEGA-10, Stanford Photonics) with a 700 nm short-pass optical filter mounted on the microscope's bottom port. Bioluminescence images were obtained with 10 minute exposures (Luntr, total acquisition time of 10.5 h) or 2 min exposures (D-luciferin, total acquisition time of 10 min) and 30 fps readouts for recordings at single cell resolution. Brightfield images were obtained with 9 second exposures and 30 fps readouts. Images were obtained and processed using the software Piper (Stanford Photonics) with a cosmic ray filter set to discriminate using the sum of all pixel values above 800 and reject frames greater than 3 standard deviations of the running average (~3-7% of frames). The software ImageJ was used for analysis and stacking (Z project function) of bioluminescence images for each experiment and for bioluminescence-brightfield overlays.

2.7I NMR uncaging and stability studies

Luntr uncaging monitored via ^1H NMR

Luntr (5 mM) and NADH (7.5 mM) were dissolved in deuterated PBS. NTR (2 $\mu\text{g}/\text{mL}$) was added and the mixture was incubated at 37 $^\circ\text{C}$ for 2 h. A synthetic standard of 6-amino luciferin (**S1**) was doped in the NMR sample after 2 h. NMR spectra were obtained prior to NTR addition, 2 h post-NTR addition and directly after addition of 6-amino luciferin. In Figure 2-3D the spectra are referenced to the furthest downfield resonance of Luntr (prior to NTR addition). In Figs. 2-4, 2-5, 2-6 the spectra are referenced to residual solvent.

Doping of synthetic luciferin **2.10** into enzymatic reduction

The enzymatic reduction of Luntr by NTR was performed as described above. After a 3 h incubation at room temperature, a ^1H -NMR spectrum was acquired. A solution of luciferin **2.10** (400 μL of a 10 mM solution in PBS) was then added to 400 μL of the enzymatic reduction mixture. After mixing, a second ^1H -NMR spectrum was acquired.

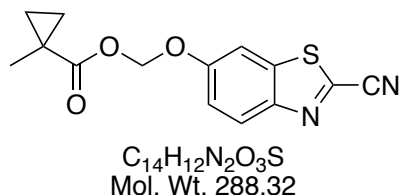
Stability studies with synthetic luciferin **2.10**

Luciferin **2.10** was purified *via* reversed-phase HPLC and dissolved in deuterated PBS (pH 7.4) to a final concentration of 10 mM. DMSO was then added as a reference standard. ^1H -NMR spectra were acquired at 400 MHz over 23 h. The sample was incubated in a 37 $^\circ\text{C}$ water bath for the duration of the study.

2.7m General synthetic procedures

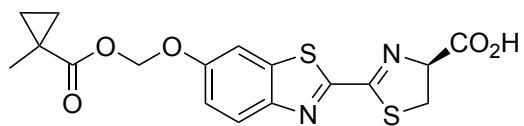
All reagents purchased from commercial suppliers were of analytical grade and used without further purification. Compounds **2.3**, **2.8**, and **2.9** were prepared as previously described [12]. Appel's salt, 4,5-dichloro-1,2,3-dithiazolium chloride (**2.7**), was prepared according to literature precedent [12]. Reaction progress was monitored by thin-layer chromatography (TLC) and preparative TLC performed on EMD 60 F254 plates, visualized with UV light, ninhydrin or KMnO₄ stain. 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 a 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). Anhydrous solvents (acetonitrile (MeCN), dichloromethane (DCM), methanol (MeOH)) were dried by passage over neutral alumina. Dimethylformamide (DMF) was passed over activated molecular sieves and dimethyl sulfoxide (DMSO) was purchased from Acros Organics in AcroSeal bottles and used without further purification. Reaction vessels were either flame- or oven-dried prior to use. NMR spectra were acquired with Bruker Advanced spectrometers. All spectra were acquired at 298 K. ¹H-NMR spectra were acquired at 400 MHz or 500 MHz, and ¹³C-NMR spectra were acquired at 125 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. The abbreviations used can be found in the document JOC Standard Abbreviations and Acronyms, <http://pubs.acs.org/paragonplus>.

2.7n Compound syntheses



((2-Cyanobenzo[d]thiazol-6-yl)oxy)methyl 1-methylcyclopropane-1-carboxylate (2.5). A flame-dried

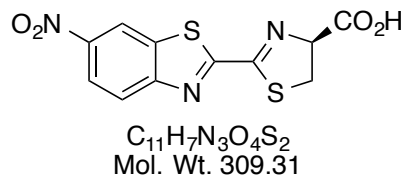
round bottom flask with sodium hydride (0.066 g, 1.7 mmol of a 60% dispersion in mineral oil) was flushed with N_2 . Dry DMF (9 mL) was then added to the flask and the mixture was cooled to 0 °C. 6-Hydroxybenzo[d]thiazole-2-carbonitrile (**3**, 0.226 g 1.28 mmol) was dissolved in 1 mL dry DMF and added dropwise to the mixture. Iodomethyl 1-methylcyclopropane-1-carboxylate (**4**) (0.460 g, 1.92 mmol) was then added dropwise over 5 min. The reaction was allowed to warm to room temperature and stirred for 18 h under N_2 . The reaction was then quenched with H_2O , whereupon a precipitate formed. The solids were filtered, dried under vacuum and reserved. The filtrate was extracted with EtOAc (2 x 30 mL), and the organic fractions were combined and washed with water (3 x 100 mL) and brine (1 x 100 mL), and dried with $MgSO_4$. The organics were then filtered and concentrated *in vacuo* to afford a residue, which was purified by flash column chromatography (eluting with 8:2 to 1:1 hexanes:EtOAc). This solid was combined with the reserved precipitate to afford **5** (0.11 g, 29%) as an off-white solid. 1H NMR (400 MHz, $CDCl_3$) δ 8.13 (d, $J = 9.1$, 1H), 7.56 (d, $J = 2.2$, 1H), 7.34 (dd, $J = 9.1, 2.4$, 1H), 5.84 (s, 2H), 1.31 (s, 3H), 1.27 (app dd, $J = 6.8, 4.1$, 2H), 0.76 (app dd, $J = 6.8, 4.0$, 2H); ^{13}C (125 MHz, $CDCl_3$) δ 174.9, 157.4, 148.1, 137.2, 134.8, 126.3, 119.0, 113.1, 106.5, 85.3, 19.3, 18.8, 17.6; HRMS (GC-Cl) m/z calcd for $C_{14}H_{12}N_2O_3S$ [$M + H$] $^+$ 289.0647, found 289.0638.



C₁₇H₁₆N₂O₅S₂
Mol. Wt. 392.44

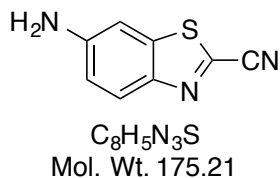
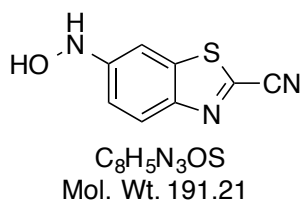
(S)-2-(6-(((1-Methylcyclopropane-1-carbonyl)oxy)methoxy)benzo[d]thiazol-2-yl)-4,5-dihydrothiazole-4-carboxylic acid (2.1, Lucy). D-Cysteine hydrochloride monohydrate

(0.0820 g, 0.521 mmol) and potassium carbonate (0.0959 g, 0.694 mmol) were dissolved in 0.5 mL H₂O and added to a 20 mL vial containing a suspension of **2.5** (0.0850 g, 0.295 mmol) in 1.5 mL MeCN. The reaction mixture was stirred under N₂ and monitored for the disappearance of **2.4** by TLC (3:2 hexanes:acetone). After 1.5 h, the reaction was neutralized to pH 7 with acetic acid, whereupon a light brown precipitate formed. The precipitate was collected *via* filtration and rinsed with chilled MeCN. The filtrate was concentrated *in vacuo* and purified by reversed phase semi-preparative HPLC, eluting with a gradient of 0 to 100% MeCN in water (at a flow rate of 5 mL/min). The precipitate and purified material were combined and concentrated to yield **2.1** (0.050 g, 43%) as a light brown solid. ¹H NMR (400 MHz, DMSO) δ 8.07 (d, *J* = 8.9, 1H), 7.87 (d, *J* = 2.3, 1H), 7.26 (dd, *J* = 9.0, 2.2, 1H), 5.84 (s, 2H), 4.92 (t app, *J* = 8.9, 1H), 3.75 (t app, *J* = 9.3, 1H), 3.50 (t app, *J* = 10, 1H), 1.23 (s, 3H), 1.13 (m, 2H), 0.79 (m, 2H); ¹³C (125 MHz, CDCl₃) δ 174.1, 170.0, 160.7, 159.2, 155.2, 148.4, 136.6, 124.7, 117.1, 108.1, 85.3, 83.9, 36.3, 18.7, 18.3, 16.6; HRMS (ESI-TOF) *m/z* calcd for C₁₇H₁₆N₂O₅S₂ [M – CO₂H][–] 347.0524, found 347.0511.



(S)-2-(6-Nitrobenzo[d]thiazol-2-yl)-4,5-dihydrothiazole-4-carboxylic acid (2.2).

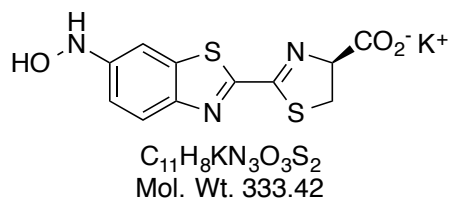
D-Cysteine hydrochloride monohydrate (0.388 g, 2.46 mmol) was dissolved in 5.0 mL H₂O and neutralized to pH 7 with 1.5 M KOH. The solution was added to a round bottom flask containing a suspension of **2.9** (0.500 g, 2.46 mmol) in 10 mL MeCN. The reaction mixture was stirred under N₂ and monitored for the disappearance of **2.9** by TLC (7:3 hexanes:EtOAc). After completion (2.5 h) the reaction mixture was diluted with PBS (20 mL), the aqueous layer was washed with EtOAc (2 x 20 mL) and then acidified to pH 3 with 1 M HCl. The aqueous layer was then extracted with EtOAc (2 x 30 mL). The organic layers were combined, washed with brine (1 x 50 mL), dried over MgSO₄, filtered through Celite and concentrated *in vacuo* to yield **1** as a light brown solid (0.584 g, 77%). Note: Luntr (**2.1**) has limited solubility in most solvents, and in 100% DMSO it changes color rapidly. A mixture of solvents was used for NMR characterization. ¹H NMR (400 MHz, CD₃OD:benzene-*d*₆ 1:1) δ 8.71 (s, 1H), 8.20 (d, *J* = 9.0, 1H), 8.03 (d, *J* = 9.2, 1H), 5.30 (app t, *J* = 9.0, 1H), 3.81 (app t, *J* = 9.8, 1H), 3.68 (app t, *J* = 10.3, 1H); ¹³C (125 MHz, CD₃OD:D₂O 1:1) δ 177.3, 168.6, 165.5, 157.2, 147.1, 137.2, 125.2, 123.2, 120.4, 82.4, 37.7. HRMS (ESI-TOF) *m/z* calcd for C₁₀H₆N₂O₂S₂ [M – CO₂H]⁻ 263.9901, found 263.9879.



6-(Hydroxyamino)benzo[d]thiazole-2-carbonitrile (S2), 6-Aminobenzo[d]thiazole-2-carbonitrile

(2.11). 6-Nitro-1,3-benzothiazole-2-carbonitrile (**2.9**, 0.300 g, 1.46 mmol) and zinc

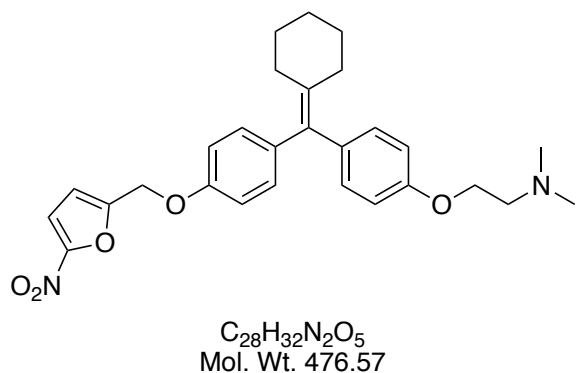
powder (0.201 g, 3.07 mmol) were added to a 250 mL round bottom flask. The flask was flushed with N₂ and 150 mL of EtOH was added. The mixture was stirred until **5** dissolved. Subsequently, 15 mL of H₂O was added to the flask (followed by 1 mL of sat. NH₄Cl solution). The reaction mixture was stirred at room temperature and monitored by TLC (7:3 hexanes to EtOAc) for appearance of a more polar product (*R_f* = 0.2) with a dark purple UV signature. After stirring for 4 h, the reaction was filtered through Celite, diluted with 50 mL brine and extracted with EtOAc (3 x 100 mL). The organic layers were combined and washed with H₂O (2 x 100 mL) and brine (1 x 100 mL) and dried by rotary evaporation to yield a yellow solid. The solid was immediately absorbed to silica gel and isolated *via* flash column chromatography (eluting with 7:3 to 1:1 hexanes:EtOAc) to remove unreacted starting material (**5**). A mixture of **2.11** and **2.12** as a yellow-orange solid was isolated. The products were immediately carried on to the next step without further purification.



(S)-2-(6-(Hydroxyamino)benzo[d]thiazol-2-yl)-4,5-dihydrothiazole-4-carboxylic acid (2.10). A mixture of **2.11** and **2.12** (0.092 g) was dissolved in 20 mL

MeCN and added to a 100 mL round bottom flask. The flask was flushed with N₂. A pH 8 solution (basified with 1.5 M KOH) of D-Cysteine•HCl (0.0928 g, 0.281 mmol) was made in 2.0 mL H₂O and added to the flask containing compounds **2.11** and **2.12**. The reaction mixture was stirred at room temperature and monitored for disappearance of starting material by TLC (7:3 hexanes:EtOAc). After stirring for 2 h the reaction was diluted with 20 mL H₂O, and the aqueous layer was washed with EtOAc (3 x 20 mL).

The aqueous layer was then concentrated *via* rotary evaporation, the resultant bright yellow solid was dissolved in ~ 5 mL H₂O, passed through a 0.2 mm filter and either stored at -80 °C or immediately HPLC purified (semi preparative, reversed phase, eluting with a gradient of 0-90% MeOH in H₂O at a flow rate of 4.2 mL/min). The relevant fractions were combined and concentrated *in vacuo* to produce **2.10** as a yellow solid (0.018 g, 4% over two steps). Purified **2.10** was immediately used for experiments after isolation. ¹H NMR (500 MHz, D₂O) δ 7.78 (d, *J* = 8.8, 1H), 7.57 (d, *J* = 1.7, 1H), 7.12 (dd, *J* = 2.1, *J* = 8.9, 1H), 5.15 (app t, *J* = 9.0, 1H), 3.76 (app t, *J* = 10.6, 1H), 3.58 (dd, *J* = 8.2, 11.2, 1H).



2-(4-(Cyclohexylidene(4-((5-nitrofuran-2-yl)methoxy)phenyl)methyl)phenoxy)-N,N-dimethylethan-1-amine (2.18).

Cyclofen-OH (**2.14**, 0.048 g, 0.14 mmol), potassium carbonate (0.022 g, 0.16 mmol), and sodium iodide (0.024 g, 0.16 mmol),

were added to a flame-dried flask. The flask was flushed with nitrogen and 10 mL of acetonitrile was added via syringe. While stirring at room temperature, 2-(bromomethyl)-5-nitrofuran (**2.15**) was added (0.033 g, 0.16 mmol) to the reaction mixture after being dissolved in acetonitrile. The reaction was then heated to 60 °C and stirred for 48 h. After cooling, the solvent was removed via rotary evaporation. The brown-red solid was then dissolved in EtOAc (20 mL) and washed with water (1 x 20 mL), dried with MgSO₄, and the organic phase was evaporated *in vacuo*. The brown solid was dissolved in DCM:MeOH (9:1) and purified by preparative TLC (8% MeOH in DCM). The separation

was performed in 8% MeOH in DCM and the corresponding band was scraped from the plate and the compound was eluted from the silica with 20% MeOH in DCM. The solvent was evaporated to yield 0.005 g (8%) of cyclo-NTR (**2.16**) as a light brown solid. ^1H NMR (400 MHz, CD_3OD), ~66 % estimated purity δ 7.54 (d, $J = 3.7$, 1H), 7.21 (d, $J = 3.7$, 1H), 7.05 (app d, $J = 8.7$, 2H), 6.96 (app d, $J = 8.6$, 2H), 6.88 (app d, $J = 8.5$, 2H), 6.68 (app d, $J = 8.5$, 2H), 4.94 (s, 2H), 4.57 (m, 2H), 3.92 (m, 2H), 3.32 (s, 6H), 2.23 (m, 4H), 1.58 (m, 7H). HRMS (ESI-TOF) m/z calcd for $\text{C}_{28}\text{H}_{33}\text{N}_2\text{O}_5$ $[\text{M} + \text{H}]^+$ 477.2390, found 477.2401.

References

1. Germain, R. N.; Robey, E. A.; Cahalan, M. D., A Decade of Imaging Cellular Motility and Interaction Dynamics in the Immune System. *Science* **2012**, *336*, 1676.
2. Ntziachristos, V., Fluorescence molecular imaging. *Annu. Rev. Biomed. Eng.* **2006**, *8*, 1.
3. Prescher, J. A.; Contag, C. H., Guided by the light: visualizing biomolecular processes in living animals with bioluminescence. *Curr. Opin. Chem. Biol.* **2010**, *14*, 80.
4. Paley, M. A.; Prescher, J. A., Bioluminescence: a versatile technique for imaging cellular and molecular features. *Medchemcomm* **2014**, *5*, 255.
5. Sellmyer, M. A.; Bronsart, L.; Imoto, H.; Contag, C. H.; Wandless, T. J.; Prescher, J. A., Visualizing cellular interactions with a generalized proximity reporter. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 8567.
6. Cronin, M.; Akin, A. R.; Collins, S. A.; Meganck, J.; Kim, J. B.; Baban, C. K.; Joyce, S. A.; van Dam, G. M.; Zhang, N.; van Sinderen, D.; O'Sullivan, G. C.; Kasahara, N.; Gahan, C. G.; Francis, K. P.; Tangney, M., High resolution in vivo bioluminescent imaging for the study of bacterial tumour targeting. *PLoS One* **2012**, *7*.
7. Badr, C. E.; Tannous, B. A., Bioluminescence imaging: progress and applications. *Trends Biotechnol.* **2011**, *29*, 624.

8. Rabinovich, B. A.; Ye, Y.; Etto, T.; Chen, J. Q.; Levitsky, H. I.; Overwijk, W. W.; Cooper, L. J. N.; Gelovani, J.; Hwu, P., Visualizing fewer than 10 mouse T cells with an enhanced firefly luciferase in immunocompetent mouse models of cancer. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 14342.
9. Sacco, A.; Doyonnas, R.; Kraft, P.; Vitorovic, S.; Blau, H. M., Self-renewal and expansion of single transplanted muscle stem cells. *Nature* **2008**, *456*, 502.
10. 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.
11. 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.
12. McCutcheon, D. C.; Porterfield, W. B.; Prescher, J. A., Rapid and scalable assembly of firefly luciferase substrates. *Org. Biomol. Chem.* **2015**, *13*, 2117.
13. 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.
14. Evans, M. S.; Chaurette, J. P.; Adams, 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. *Nature Methods* **2014**, *11*, 393.
15. Takakura, H.; Kojima, R.; Ozawa, T.; Nagano, T.; Urano, Y., Development of 5'- and 7'-Substituted Luciferin Analogues as Acid-Tolerant Substrates of Firefly Luciferase. *ChemBioChem* **2012**, *13*, 1424.
16. Li, J.; Chen, L. Z.; Du, L. P.; Li, M. Y., Cage the firefly luciferin! - a strategy for developing bioluminescent probes. *Chem. Soc. Rev.* **2013**, *42*, 662.
17. Yao, H.; So, M. K.; Rao, J., A bioluminogenic substrate for in vivo imaging of beta-lactamase activity. *Angew. Chem. Int. Ed.* **2007**, *46*, 7031.
18. Rush, J. S.; Beatty, K. E.; Bertozzi, C. R., Bioluminescent Probes of Sulfatase Activity. *ChemBioChem* **2010**, *11*, 2096.
19. Niwa, K.; Nakajima, Y.; Ohmiya, Y., Applications of luciferin biosynthesis: Bioluminescence assays for L-cysteine and luciferase. *Anal. Biochem.* **2010**, *396*, 316.
20. Van de Bittner, G. C.; Dubikovskaya, E. A.; Bertozzi, C. R.; Chang, C. J., In vivo imaging of hydrogen peroxide production in a murine tumor model with a chemoselective bioluminescent reporter. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 21316.

21. Inlay, M. A.; Choe, V.; Bharathi, S.; Fernhoff, N. B.; Baker, J. R.; Weissman, I. L.; Choi, S. K., Synthesis of a photocaged tamoxifen for light-dependent activation of Cre-ER recombinase-driven gene modification. *Chem. Commun.* **2013**, *49*, 4971.
22. Faal, T.; Wong, P. T.; Tang, S. Z.; Coulter, A.; Chen, Y. M.; Tu, C. H.; Baker, J. R.; Choi, S. K.; Inlay, M. A., 4-Hydroxytamoxifen probes for light-dependent spatiotemporal control of Cre-ER mediated reporter gene expression. *Molecular Biosystems* **2015**, *11*, 783.
23. Lindberg, E.; Mizukami, S.; Iyata, K.; Miyawaki, A.; Kikuchi, K., Development of Luminescent Coelenterazine Derivatives Activatable by beta-Galactosidase for Monitoring Dual Gene Expression. *Chem. Euro. J.* **2013**, *19*, 14970.
24. Tian, L.; Yang, Y. L.; Wysocki, L. M.; Arnold, A. C.; Hu, A.; Ravichandran, B.; Sternson, S. M.; Looger, L. L.; Lavis, L. D., Selective esterase-ester pair for targeting small molecules with cellular specificity. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 4756.
25. Bender, D. M.; Peterson, J. A.; McCarthy, J. R.; Gunaydin, H.; Takano, Y.; Houk, K. N., Cyclopropanecarboxylic acid esters as potential prodrugs with enhanced hydrolytic stability. *Org. Lett.* **2008**, *10*, 509.
26. Demeijere, A., Bonding Properties of Cyclopropane and Their Chemical Consequences. *Angewandte Chemie-International Edition in English* **1979**, *18*, 809.
27. Friedlos, F.; Court, S.; Ford, M.; Denny, W. A.; Springer, C., Gene-directed enzyme prodrug therapy: quantitative bystander cytotoxicity and DNA damage induced by CB1954 in cells expressing bacterial nitroreductase. *Gene Ther.* **1998**, *5*, 105.
28. Knox, R. J.; Friedlos, F.; Boland, M. P., The Bioactivation of Cb-1954 and Its Use as a Prodrug in Antibody-Directed Enzyme Prodrug Therapy (Adept). *Cancer and Metastasis Rev.* **1993**, *12*, 195.
29. Chung-Faye, G.; Palmer, D.; Anderson, D.; Clark, J.; Downes, M.; Baddeley, J.; Hussain, S.; Murray, P. I.; Searle, P.; Seymour, L.; Harris, P. A.; Ferry, D.; Kerr, D. J., Virus-directed, enzyme prodrug therapy with nitroimidazole reductase: A phase I and pharmacokinetic study of its prodrug, CB1954. *Clin. Cancer Res.* **2001**, *7*, 2662.
30. Johansson, E.; Parkinson, G. N.; Denny, W. A.; Neidle, S., Studies on the nitroreductase prodrug-activating system. Crystal structures of complexes with the inhibitor dicoumarol and dinitrobenzamide prodrugs and of the enzyme active form. *J. Med. Chem.* **2003**, *46*, 4009.

31. Helsby, N. A.; Ferry, D. M.; Patterson, A. V.; Pullen, S. M.; Wilson, W. R., 2-Amino metabolites are key mediators of CB 1954 and SN 23862 bystander effects in nitroreductase GDEPT. *Br. J. Cancer* **2004**, *90*, 1084.
32. McCormack, E.; Silden, E.; West, R. M.; Pavlin, T.; Micklem, D. R.; Lorens, J. B.; Haug, B. E.; Cooper, M. E.; Gjertsen, B. T., Nitroreductase, a Near-Infrared Reporter Platform for In Vivo Time-Domain Optical Imaging of Metastatic Cancer. *Cancer Res.* **2013**, *73*, 1276.
33. Bhaumik, S.; Sekar, T. V.; Depuy, J.; Klimash, J.; Paulmurugan, R., Noninvasive optical imaging of nitroreductase gene-directed enzyme prodrug therapy system in living animals. *Gene Ther.* **2012**, *19*, 295.
34. White, E. H.; Worther, H.; Seliger, H. H.; Mcelroy, W. D., Amino Analogs of Firefly Luciferin and Biological Activity Thereof. *J. Am. Chem. Soc.* **1966**, *88*, 2015.
35. 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.
36. Shao, Q.; Jiang, T. T.; Ren, G.; Cheng, Z.; Xing, B. G., Photoactivable bioluminescent probes for imaging luciferase activity. *Chem. Commun.* **2009**, 10.1039/b908346d, 4028.
37. Olekhovich, I. N.; Goodwin, A.; Hoffman, P. S., Characterization of the NAD(P)H oxidase and metronidazole reductase activities of the RdxA nitroreductase of *Helicobacter pylori*. *FEBS J.* **2009**, *276*, 3354.
38. Lee, M. K.; Williams, J.; Twieg, R. J.; Rao, J. H.; Moerner, W. E., Enzymatic activation of nitro-aryl fluorogens in live bacterial cells for enzymatic turnover-activated localization microscopy. *Chem. Sci.* **2013**, *4*, 220.
39. Christofferson, A.; Wilkie, J., Mechanism of CB1954 reduction by *Escherichia coli* nitroreductase. *Biochem. Soc. Trans.* **2009**, *37*, 413.
40. Roldan, M.; Perez-Reinado, E.; Castillo, F.; Moreno-Vivian, C., Reduction of polynitroaromatic compounds: the bacterial nitroreductases. *FEMS Microbiol. Rev.* **2008**, *32*, 474.
41. Brusehaber, E.; Bottcher, D.; Bornscheuer, U. T., Insights into the physiological role of pig liver esterase: Isoenzymes show differences in the demethylation of prenylated proteins. *Biorg. Med. Chem.* **2009**, *17*, 7878.
42. Livet, J.; Weissman, T. A.; Kang, H. N.; Draft, R. W.; Lu, J.; Bennis, R. A.; Sanes, J. R.; Lichtman, J. W., Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **2007**, *450*, 56.

43. Cai, D.; Cohen, K. B.; Luo, T.; Lichtman, J. W.; Sanes, J. R., Improved tools for the Brainbow toolbox. *Nature Methods* **2013**, *10*, 540.
44. Cronican, J. J.; Beier, K. T.; Davis, T. N.; Tseng, J. C.; Li, W. D.; Thompson, D. B.; Shih, A. F.; May, E. M.; Cepko, C. L.; Kung, A. L.; Zhou, Q.; Liu, D. R., A Class of Human Proteins that Deliver Functional Proteins into Mammalian Cells In Vitro and In Vivo. *Chem. Biol.* **2011**, *18*, 833.
45. Kennedy, M. J.; Hughes, R. M.; Peteya, L. A.; Schwartz, J. W.; Ehlers, M. D.; Tucker, C. L., Rapid blue-light-mediated induction of protein interactions in living cells. *Nature Methods* **2010**, *7*, 973.
46. Lu, X.; Agasti, S. S.; Vinegoni, C.; Waterman, P.; DePinho, R. A.; Weissleder, R., Optochemogenetics (OCG) Allows more precise control of genetic engineering in mice with CreER regulators. *Bioconjugate Chem.* **2012**, *23*, 1945.
47. Grimm, J. B.; Gruber, T. D.; Ortiz, G.; Brown, T. A.; Lavis, L. D., Virginia Orange: A versatile, red-shifted fluorescein scaffold for single- and dual-input fluorogenic probes. *Bioconjugate Chem.* **2016**, *27*, 474.
48. Vorobyeva, A. G.; Stanton, M.; Godinat, A.; Lund, K. B.; Karateev, G. G.; Francis, K. P.; Allen, E.; Gelovani, J. G.; McCormack, E.; Tangney, M.; Dubikovskaya, E. A., Development of a Bioluminescent Nitroreductase Probe for Preclinical Imaging. *PLoS One* **2015**, *10*.
49. Berry, J. M.; Watson, C. Y.; Whish, W. J. D.; Threadgill, M. D., 5-nitrofuranyl-methyl group as a potential bioreductively activated pro-drug system. *Journal of the Chemical Society-Perkin Transactions 1* **1997**, DOI 10.1039/a607202j, 1147.
50. Porterfield, W. B.; Jones, K. A.; McCutcheon, D. C.; Prescher, J. A., A "Caged" Luciferin for Imaging Cell-Cell Contacts. *J. Am. Chem. Soc.* **2015**, *137*, 8656.
51. Sinha, D. K.; Neveu, P.; Gagey, N.; Aujard, I.; Benbrahim-Bouzidi, C.; Le Saux, T.; Rampon, C.; Gauron, C.; Goetz, B.; Dubruille, S.; Baaden, M.; Volovitch, M.; Bensimon, D.; Vriza, S.; Jullien, L., Photocontrol of Protein Activity in Cultured Cells and Zebrafish with One- and Two-Photon Illumination. *ChemBioChem* **2010**, *11*, 653.
52. Wong, P. T.; Roberts, E. W.; Tang, S.; Mukherjee, J.; Cannon, J.; Nip, A. J.; Corbin, K.; Krummel, M. F.; Choi, S. K., Control of an Unusual Photo-Claisen Rearrangement in Coumarin Caged Tamoxifen through an Extended Spacer. *ACS Chem. Biol.* **2017**, *12*, 1001.
53. Feng, P.; Zhang, H. T.; Deng, Q. K.; Liu, W.; Yang, L. H.; Li, G. B.; Chen, G.; Du, L. P.; Ke, B. W.; Li, M. Y., Real-Time Bioluminescence Imaging of Nitroreductase in Mouse Model. *Anal. Chem.* **2016**, *88*, 5610.

Chapter 3: Orthogonal luciferase-luciferin pairs for bioluminescence imaging

3.1 Introduction

Bioluminescence has two main limitations as mentioned in chapter 1. The lack of spatial resolution was addressed in chapter 2. The other limitation of bioluminescence is the inability to monitor more than one biological feature at a time. This is due, in part, to a lack of distinguishable luciferase-luciferin pairs for *in vivo* use. The optimal luciferases (from the insect family) use the same substrate, D-luciferin (Fig. 3-1A) [1-2]. Thus, they cannot easily discriminate multiple cell types in a single subject. Additionally, unlike fluorescent protein technologies, a diverse suite of accessible bioluminescent probes does not yet exist. To address this void, D-luciferin analogs have been engineered to emit different colors of light [3-5]. However, these substrates are still utilized by the same luciferases, precluding the distinct genetic tagging of individual cell types. Insect luciferases have also been engineered to emit different colors of light with D-luciferin [6-8]. The observed emission spectra are not sufficiently resolved, though, for routine use in complex tissues or animals. Discriminating among different wavelengths in bioluminescence (and whole body optical imaging, in general) is exceedingly difficult.

Contrasting with these attempts to achieve spectral resolution, we aimed to obtain distinguishable bioluminescent probes via substrate resolution. Substrate-resolved bioluminescence is well preceded in nature, as structurally distinct luciferase-luciferin pairs have been identified across diverse phyla [9-11]. Some of these pairs, including those from the firefly and *Renilla reniformis* have been used extensively [2,12-14]. Firefly (Fluc) and *Renilla* luciferase employ chemically unique substrates (D-

luciferin and coelenterazine, respectively), enabling their tandem application in vivo [15-16]. Coelenterazine is not ideal for use in these environments, though, owing to its suboptimal bioavailability and stability [2,17]. Other, naturally occurring luciferases and

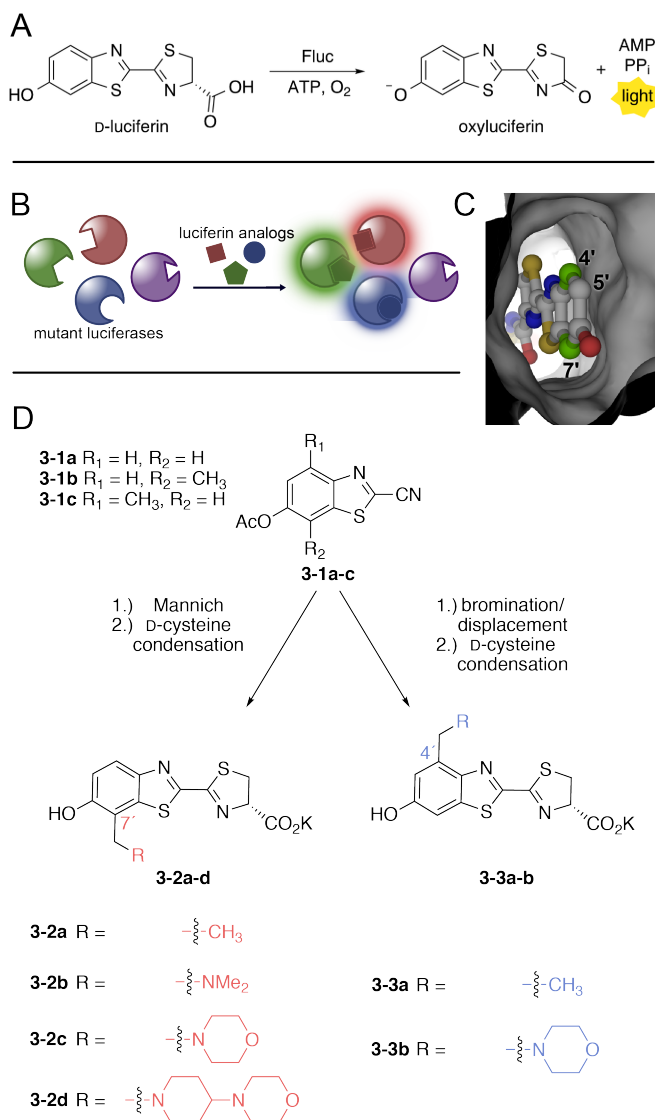


Figure 3-1. Expanding the bioluminescence toolkit with unique enzyme-substrate pairs. (A) Luciferase-mediated light production proceeds via an adenylation–oxidation sequence. (B) Strategy to develop orthogonal luciferase-luciferin pairs via substrate resolution. Genetically engineered luciferases and chemically modified luciferins were screened to identify novel partners. Only complementary enzyme-substrate pairs interact to produce light. (C) Model of D-luciferin bound to firefly luciferase (Fluc). (D) Synthesis of C7' (left) and C4' (right) sterically modified luciferins.

luciferins can be used in combination with Fluc/D-luciferin or other bioluminescent systems [16,18]. However, most of these native pairs remain poorly characterized or ill-suited for routine use.

Artificial (i.e., mutant) luciferases can exhibit altered bioluminescent properties, including tolerance for chemically modified substrates. Fluc itself has been manipulated to process analogs of D-luciferin [19]. In elegant work along these lines, Miller and coworkers prepared a class of non-natural aminoluciferins that were found to be robust light emitters with Fluc, but the products inhibited the enzymatic reaction [20]. Product inhibition was relieved using mutated versions of the enzyme [21]. These same mutations also resulted in sharply reduced emission with D-luciferin, providing key precedent for the development and utilization of orthogonal pairs [22]. The mutant enzymes from these studies, though, were not selective for one analog over another perhaps due to the structural similarities between the luciferin scaffolds. Simultaneous enzyme-substrate manipulation has also been applied to aequorin (a marine photoprotein) and the luciferase from the deep-sea shrimp *Oplophorus gracilirostris* [23-24]. In both cases, altered bioluminescent outputs (e.g., colors and stabilities) were achieved, but orthogonal substrate usage was not realized.

Here we report a general strategy for the *de novo* production of orthogonal luciferase-luciferin pairs. We synthesized a series of sterically modified luciferins that were poor emitters with Fluc, but intrinsically capable of robust light production. We then iteratively screened these analogs with libraries of mutant luciferases and identified substrate-selective enzymes. The “hits” were also biochemically characterized. Importantly, when the mutants and analogs were combined, robust light production was

observed only when complementary enzyme-substrate partners interacted. Sequential administration of substrates enabled unique luciferases to be illuminated (and thus resolved) within cultured cell models. These tools promise to bolster a variety of multi-cellular imaging applications. Importantly, our approach to identifying orthogonal bioluminescence pairs is also general and should enable rapid diversification of the bioluminescence toolkit.

3.2 Designing and constructing sterically modified luciferins

To expediently identify orthogonal bioluminescence tools, we aimed to screen sterically perturbed luciferins against libraries of mutant luciferases (Fig. 3-1B). We used the Fluc/D-luciferin pair as a starting point for several reasons. First, this duo is the most widely used in biomedical imaging applications owing to the non-toxicity of the reagents and bioavailability of the substrate [25-26]. Second, the Fluc/D-luciferin reaction releases the highest percentage of tissue-penetrating light among known bioluminescent families [27]. Thus, new enzymes and substrates based on the firefly pair would be more applicable to *in vivo* studies. Third, a wealth of structural and biochemical information on Fluc could guide our engineering efforts [11,28-31]. Finally, D-luciferin derivatives are arguably the most synthetically tractable luciferin architectures [32-33].

Generating an expanded set of bioluminescent tools required access to diverse luciferin scaffolds. A variety of D-luciferin analogs have been synthesized over the past four decades [3,5,34-38], and those capable of robust emission with Fluc harbor common features: an electron-donating group at the 6' position, a carboxylate

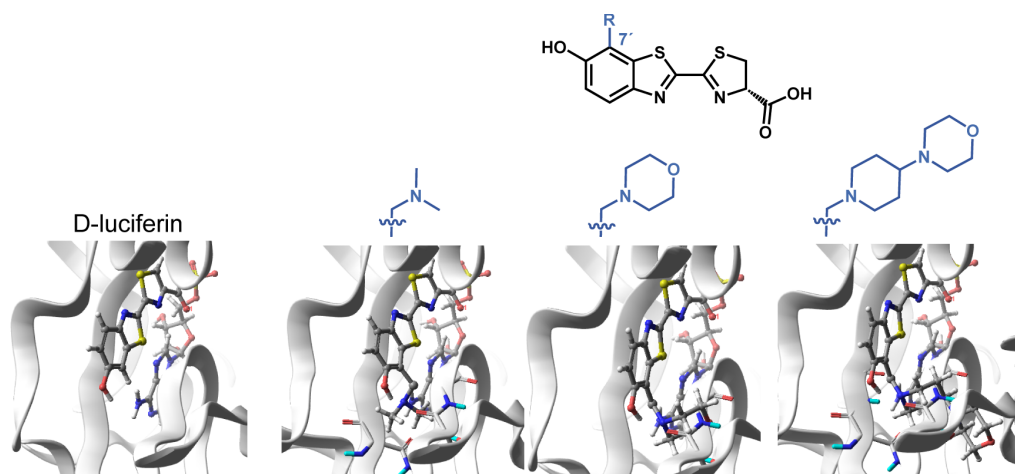


Figure 3-2. Docking studies with sterically modified luciferins. The 7' position of D-luciferin abuts the Fluc active site. Luciferins in Fluc (PDB: 4G36) were modified to contain a dimethyl amino methyl (3-2b), methyl morpholino (3-2c) or methyl morpholino-piperidyl (3-2d) appendage. As the steric bulk increased, more clashes with β -sheets were observed. When analyzed with Maestro software, all three of the analogs shown failed to dock in any conformation under various Glide conditions (SP, XP, or with SMARTS constraints).

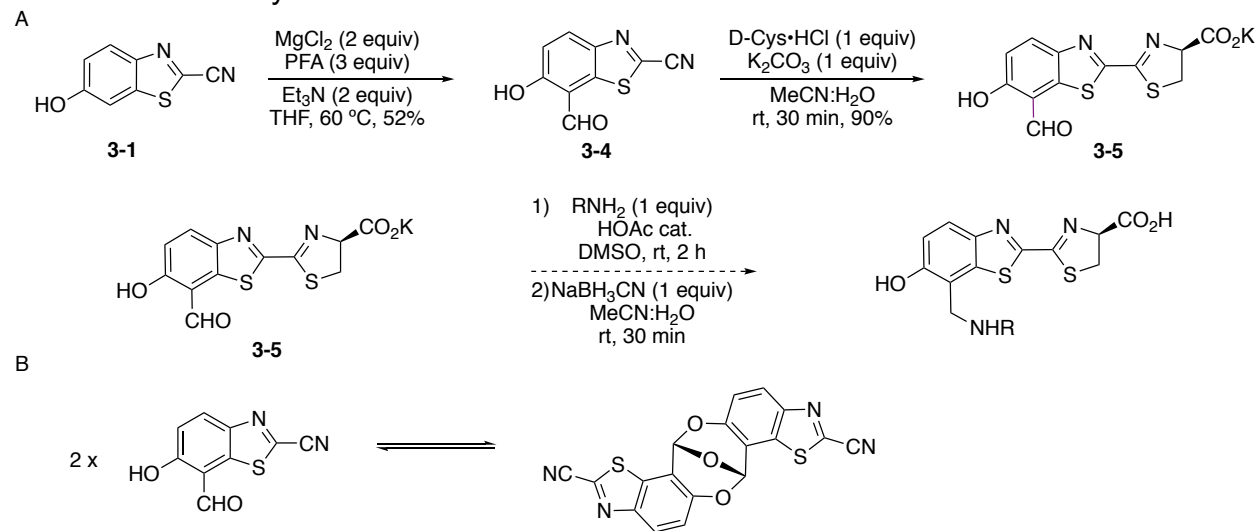
appendage (for adenylation), and an abstractable proton alpha to the carboxylate [39-40]. Beyond these requirements, Fluc can tolerate a surprisingly large variety of modified luciferins [33,35,41], including 6'-amino substituents [19-20,35], alkylated [42-44], and acylated [45] scaffolds, and even luciferins with non-natural chromophores [4,46]. Crystallographic analyses have also corroborated these experimental results, indicating flexibility within the luciferase active site and "space" to accommodate luciferins with appendages at or near the 6'-position [30-31].

Unlike most efforts to produce luciferin analogs reported to date, we were attracted to the 4' and 7' positions of the luciferin core. These positions lie in close proximity to the Fluc backbone (Fig. 3-1C). Substrates with additional steric bulk at these sites would likely be occluded from the Fluc active site and thus good targets for orthogonal probe development: while poor emitters with the native enzyme, the

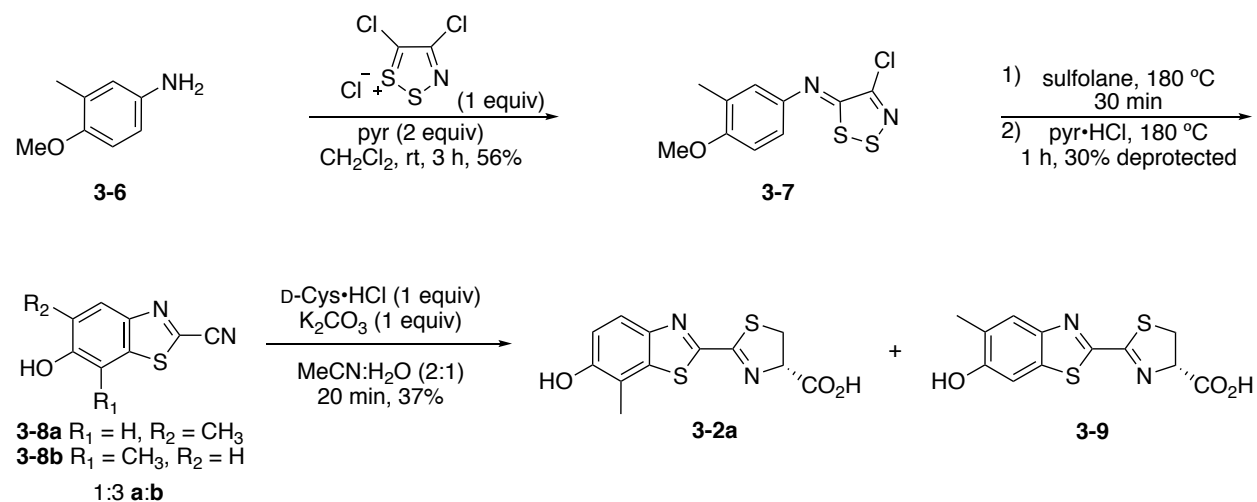
molecules could potentially give off light with designer mutants. Indeed, preliminary docking studies suggested that only analogs with small (e.g., 2-3 atoms) substituents at C4' and C7' could effectively access the active site (Fig. 3-2).

Accessing 4'- and 7'-derivatized luciferins presented an early challenge. These positions have been rarely exploited for analog development, and no prior syntheses were amenable to preparing libraries or large quantities of these probes. Rapid, high-

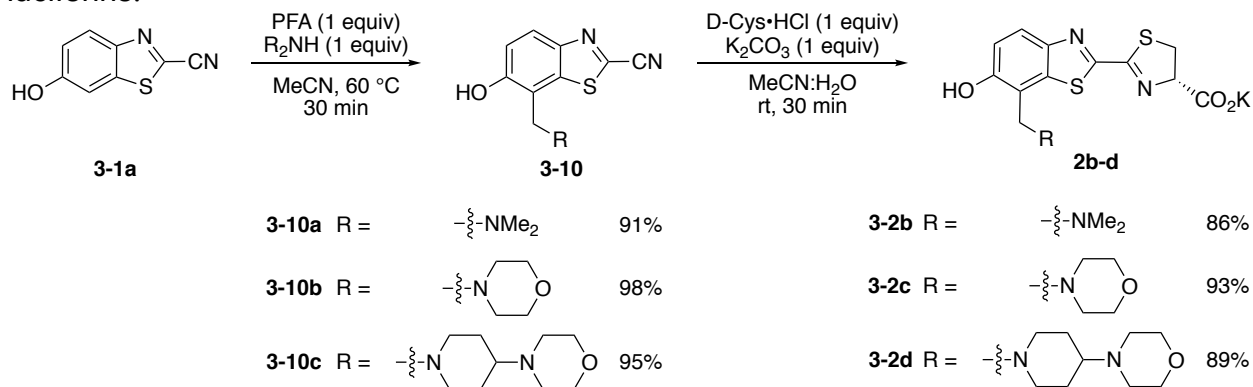
Scheme 3-1. (A) Initial synthetic route to C7'-modified luciferin analogs. (B) Hemiacetal formed from aldehyde intermediate **3-4**.



Scheme 3-2. Synthesis of 7'-methyl luciferin (**3-2a**).



Scheme 3-3: General procedure for synthesis of Betti bases and C7'-modified luciferins.



yielding syntheses were essential, as large quantities of luciferins are required for light emission assays. Fortunately, the core benzothiazole unit (**3-1a-c**) of the desired analogs could be accessed from a common route (Fig. 3-1D) and in multi-gram quantities [32,37]. From this single intermediate, we envisioned installing functional handles at C4' and C7', to rapidly assemble a variety of luciferins. We were initially drawn to an aldehyde group, owing to its ease of diversification under mild conditions (e.g., reductive amination) and broad compatibility. Aldehyde installation on **3-1a** was problematic, though, due to formation of a hydrated hemiacetal (Scheme 3-1) [47]. To circumvent this issue, we turned to more reactive iminium ions. These electrophiles can be readily trapped by electron-rich aromatics in a Mannich-type reaction [48]. Toward this end, benzothiazole **3-1a** was modified with a series of tertiary benzyl amines via *in situ* iminium formation and coupling (Fig. 3-1D and Schemes 3-2 and 3-3). The amino appendages were selected to enhance the water solubility of the luciferin core. Importantly, this synthetic approach was modular and amenable to large scale (1–10 g) syntheses. “Matched” probes with steric modifications at C4' were also prepared (**3-3a-**

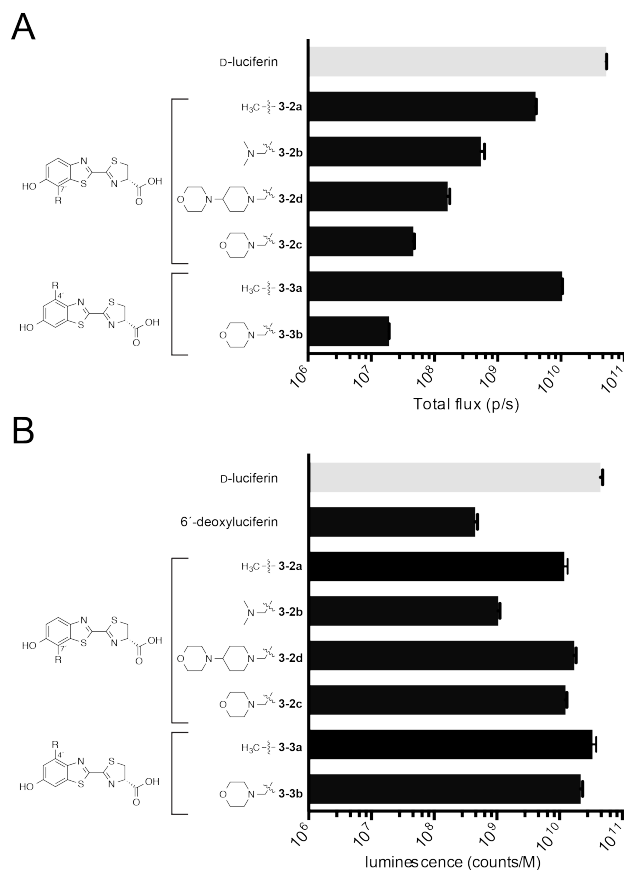


Figure 3-3. Measuring luciferin light emission. (A) Bioluminescence from luciferin analogs (100 μ M) incubated with 1 mg of Fluc. Emission intensities are plotted as total photon flux values on a log scale. Error bars represent the standard deviation of the mean for $n \geq 3$ experiments. (B) Chemiluminescence observed with luciferin analogs. Emission intensities are plotted as counts per molar luciferin on a log scale. Error bars represent the standard error of the mean for $n \geq 3$ experiments.

Consistent with the observed light outputs, the measured kinetic constants for all analogs showed reduced performance relative to D-luciferin (Table 3-1). For example, the measured K_m values were \sim 100-fold larger than the native substrate, with the largest analogs (**3-2c** and **3-2d**) exhibiting the lowest relative binding affinities. Despite their large K_m values, **3-2b-d** exhibited emission spectra similar to D-luciferin (Figs. 3-6 and 3-7). Only the C4'-modified analog **3-3b** emitted noticeably red-shifted bioluminescent light, likely due to poor Fluc binding in the excited state [31] or the luminophore being forced into a more polar environment [49-50].

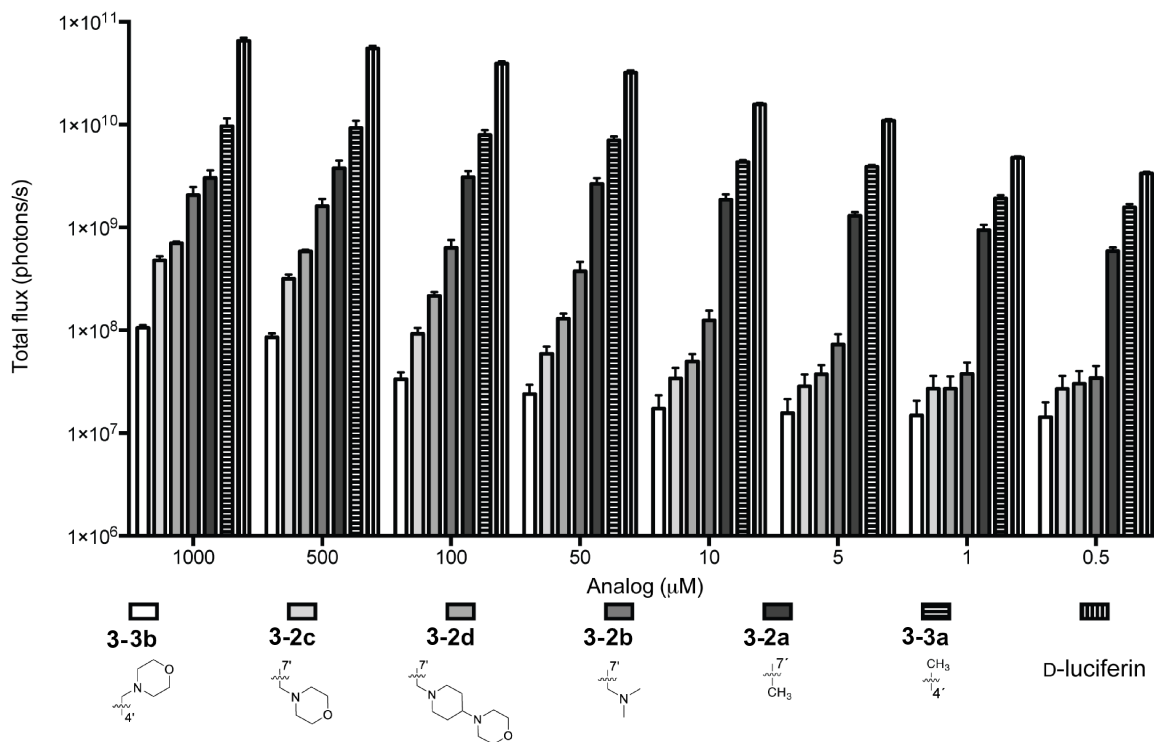


Figure 3-4. Sterically modified luciferin analogs emit light with Fluc. Analogs (0.5-1000 μM) were incubated with Fluc and ATP and light emission was quantified. Error bars represent standard error of the mean for n = 7 experiments.

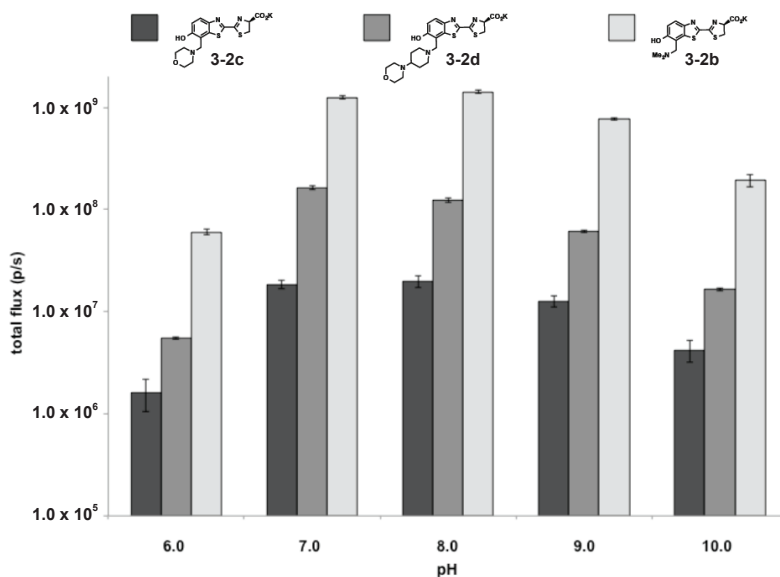


Figure 3-5. Effect of pH on bioluminescent light output with three luciferin analogs. Analogs (100 μM) were incubated with Fluc, 1mM ATP, and 0.5 mM CoA, at various pHs and light emission was quantified. Error bars represent the standard deviation of the mean for n = 3 experiments.

Table 3-1. Enzymatic and optical parameters.

K_m (μM) ^a	λ_{max} (nm)	Compound
6.24 ± 0.31	560	D-luc
329 ± 17	558	3-2b
>750	566	3-2c
>750	556	3-2d
420 ± 112	612	3-3b

^a Kinetic constants are apparent values, determined via measurements of initial rates of light emission over a range of 2 μM to 10 mM. Errors represent the standard error of the mean for $n \geq 3$ measurements.

3.4 Measuring the light-emitting potential of luciferin analogs

We attributed the weak bioluminescence of the analogs to poor utilization by Fluc. It was possible, though, that the luciferins were simply not capable of photon production upon activation and oxidation in the active site. For productive bioluminescence, an analog must be able to reach an electronic excited state (S_1) and relax back to the ground state with concomitant photon release [51-52]. If an analog cannot reach S_1 or emit efficiently from that state, reduced photon outputs would be expected. Such molecules would also be poor candidates for orthogonal probe development. To ensure that our lead analogs were intrinsically capable of light emission, we utilized a previously described chemiluminescence assay [53]. This process mimics the enzymatic reaction itself: formation of an activated ester intermediate, followed by proton abstraction and subsequent reaction with molecular oxygen [40,52,54]. When analogs **3-2a-d** and **3-3a-b** were subjected to the assay, robust light emission was observed (Fig. 3-3B). In fact, photon outputs for some of the weakest bioluminescent emitters (including **3-2c** and **3-3b**) were on par with D-luciferin.

A control compound (6'-deoxyluciferin) lacking an electron-dense residue on the aromatic ring (a key feature of luciferins) exhibited only weak levels of emission. These results provided assurance that while luciferin scaffolds may be poor substrates for Fluc, they are still *capable* of photon production and thus good candidates for orthogonal tool development.

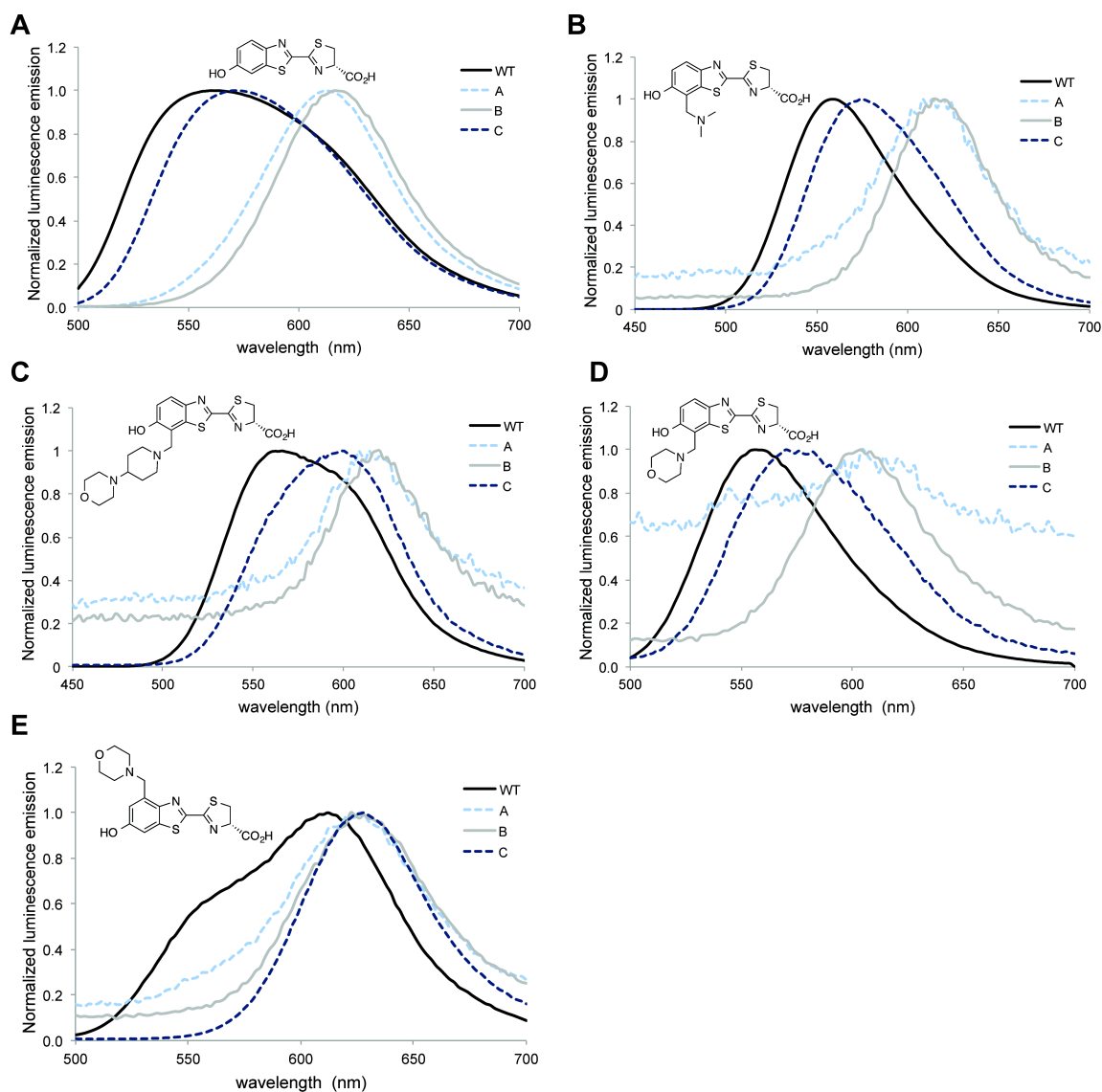


Figure 3-6. Bioluminescence emission spectra. (A)-(E) Luciferin analogs (2-5 mM) were incubated with luciferase enzymes (WT and mutants A-C) in bioluminescence buffer. Spectra were acquired as described in the Materials and Methods section.

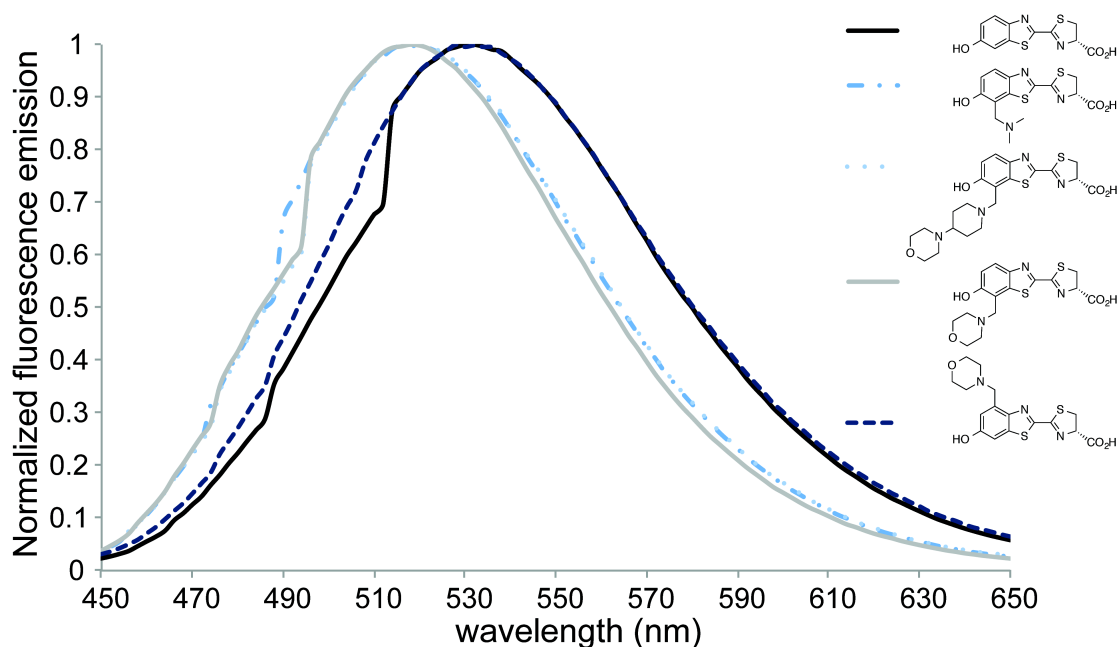


Figure 3-7. Fluorescence emission spectra. Fluorescence emission spectra for luciferin solutions (100 μ M in PBS, pH 7.4). An excitation wavelength of 350 nm was used.

3.5 Evolving substrate-specific luciferases

Having prepared candidate orthogonal luciferins, we set out to identify mutant luciferases that could selectively process the molecules. Predicting enzyme mutations that confer substrate selectivity or otherwise beneficial properties is challenging. Fluc is a highly dynamic enzyme [30,55], complicating the selection of residues from static structural or sequence data. Moreover, amino acids known to play key roles in enzyme function have been identified far from the luciferin binding site [22]; such critical residues are often revealed only by random mutagenesis approaches [56-57]. Screening libraries of completely random mutants was impractical in our case, though, owing to the large library sizes needed to achieve adequate enzyme coverage [58]. Screening in bulk is also difficult as bioluminescent light emission is too weak to detect on conventional cell

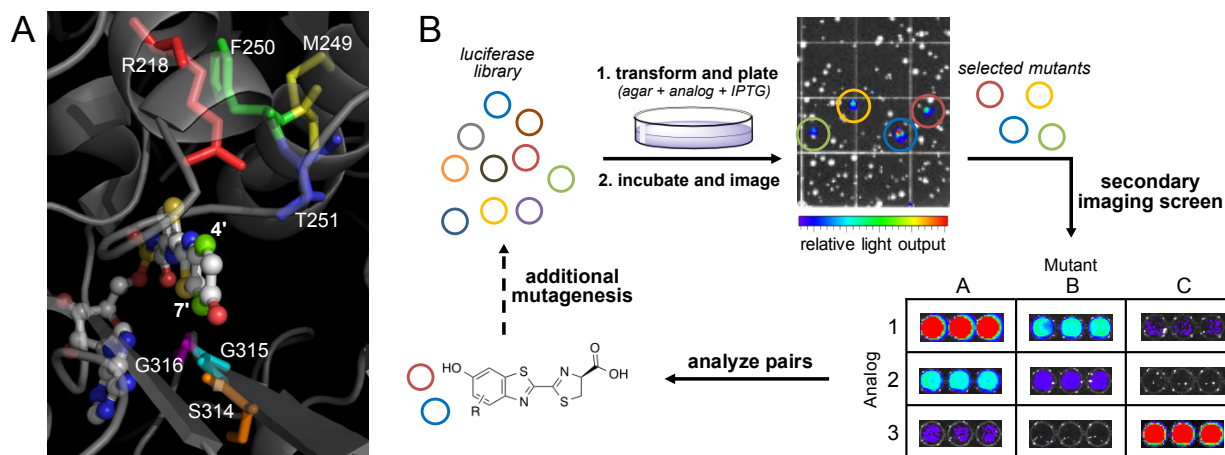


Figure 3-8. Generating mutant luciferase libraries and screening for orthogonal pairs. (A) Amino acids targeted for mutagenesis. These residues were selected based on their proximity to the 4' and 7' positions of luciferin. (B) Library screening strategy. An initial on-plate screen identified functional mutants. These “hits” were subjected to a secondary screen for orthogonality with other mutants and luciferin analogs.

sorters or other high-throughput instruments. Thus, each enzyme-substrate combination must be physically segregated (to a certain extent) and interrogated for light emission with a sensitive camera.

Recognizing that manual screening necessitated the use of smaller libraries, we developed focused, semi-rational libraries where the mutations were confined to regions known to modulate substrate binding [59]. “Hits” from these smaller, individual libraries could then be easily combined and assayed in subsequent library generations for improved function. We initially targeted residues 218, 249-251, and 314-316 for mutagenesis (Fig. 3-8A). These selections were partially based on phylogenetic data gathered from across the insect luciferase family [11,60], along with previous biochemical assays: Arg218 is known to interact with D-luciferin and influence the local structure of the binding pocket [28]; F250 lies in close proximity (~ 3 Å) to the benzothiazole ring of D-luciferin; T251 has been shown to potentiate substrate binding [29]; residues 314-316 line a critical edge near the luciferin phenolate and C7' position.

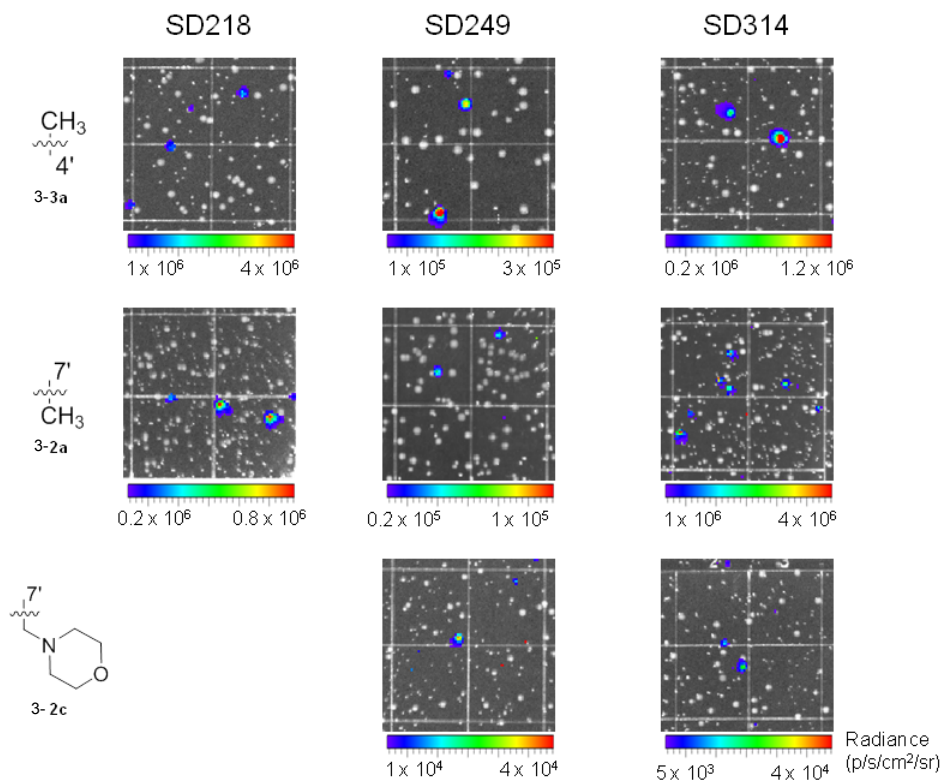


Figure 3-9. Representative images of luciferin analogs screened on plate. Mutant luciferase libraries (SD218, 249, and 314) were introduced into bacteria, and the transformants were spread on agar plates containing the specified luciferins. The plates were imaged and light-emitting colonies were selected for further expansion, screening, and analysis.

Mutations at all of these target sites have been shown to perturb D-luciferin binding (and thus light emission), while preserving the overall structural integrity of the enzyme [7,29,61].

Saturation mutagenesis was used to prepare the desired libraries. The degree of mutation applied at each residue was based on the following considerations: sequence conservation among the insect luciferase family, the identity of the native residue, and the location of the residue. For example, non-conserved residues were mutated to a higher degree compared to conserved residues in the active site. Codon compression methods were further used to eliminate redundancies and reduce the number of library

transformants [62]. The libraries were constructed using synthetic gene assembly [63] in combination with circular polymerase extension cloning (CPEC) [64].

The libraries were screened for orthogonal substrate usage using a two-tiered approach. Library DNA was first introduced into bacteria, and the transformants were

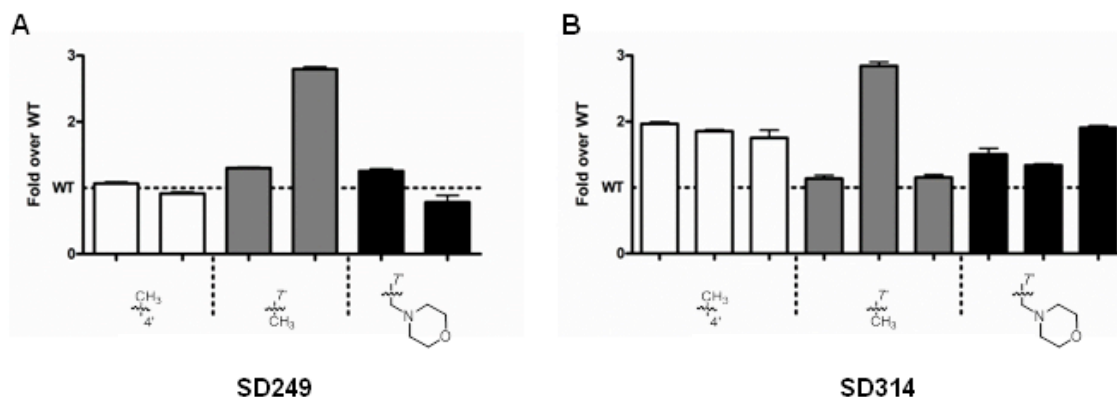


Figure 3-10. Some mutant enzymes provide similar or greater photon outputs than native luciferase. Bacterial colonies expressing mutant enzymes or Fluc (WT) were selected from the primary screen (Fig. 3-9), expanded, and lysed. Lysates were imaged with the analogs shown, and photon outputs were normalized to Fluc. Sample data are provided for transformants originating in the (A) SD249 or (B) SD314 libraries. Error bars represent the standard error of the mean for $n = 3$ experiments.

	SD218	SD249			SD314		
	R218	M249	F250	T251	S314	G315	G316
CH_3 4'	A(3)	V	C H M Y	H I S	L P(3) Q R W	N R S V	D L M V
CH_3 7'	C	–	L	–	M	–	I N Q(2)
Both (4', 7')	K(9, 5)	F(1, 3) L(4, 1)	–	–	A(11, 11) T(12, 11) V(6, 6) C(4, 5)	D	A(8, 2) S(8, 9) T(5, 3)

Figure 3-11. Sequencing analysis of “hits” from site-directed libraries. Unique mutations identified in libraries (SD218, SD249, and SD314) screened with the methyl luciferin analogs. Some of the mutations were identified in both compound screens (both). The frequency of the mutation observed is shown in parentheses. If not specified, the mutation was only observed once.

arrayed across agar plates containing embedded luciferins (Fig. 3-6B). Light-emitting colonies were easily identified (Fig. 3-9) and, in some cases, the light emission values were on par with native Fluc and D-luciferin (Fig. 3-10). A handful of the corresponding mutants were sequenced. Some mutations were observed for multiple analogs, suggesting that they might be selective for bulky luciferins (Fig. 3-11). Other mutations were unique to each compound, which is notable, given the subtle structural differences between some of the analogs.

While initial screens revealed functional mutants (and quickly culled non-functional enzymes), they did not report on selective substrate usage (i.e., orthogonality). The on-plate screens also did not control for overall expression levels and differences in compound transport. To address these parameters, we performed a secondary screen. Colonies emitting detectable levels of light on-plate were selected and expanded overnight. These cultures were then lysed and imaged with analogs. Mutants that provided light emission on par with native Fluc were identified as bona fide “hits” and used to create next-generation sequences. This iterative process was performed to evolve large pools of diverse, but functional enzymes. “Hits” from these subsequent generations were ultimately tested with all luciferin analogs in secondary screens.

To mine the entire collection of imaging data for substrate-selective pairs, we first developed a measure of orthogonality (Fig. 3-12A). The more selective a pair of enzymes for their cognate substrates, the larger the orthogonality rating. The equation returns a high number when mutants react well with one compound and not another, and penalizes pairings that do not have an evenly matched switch in preference. Since

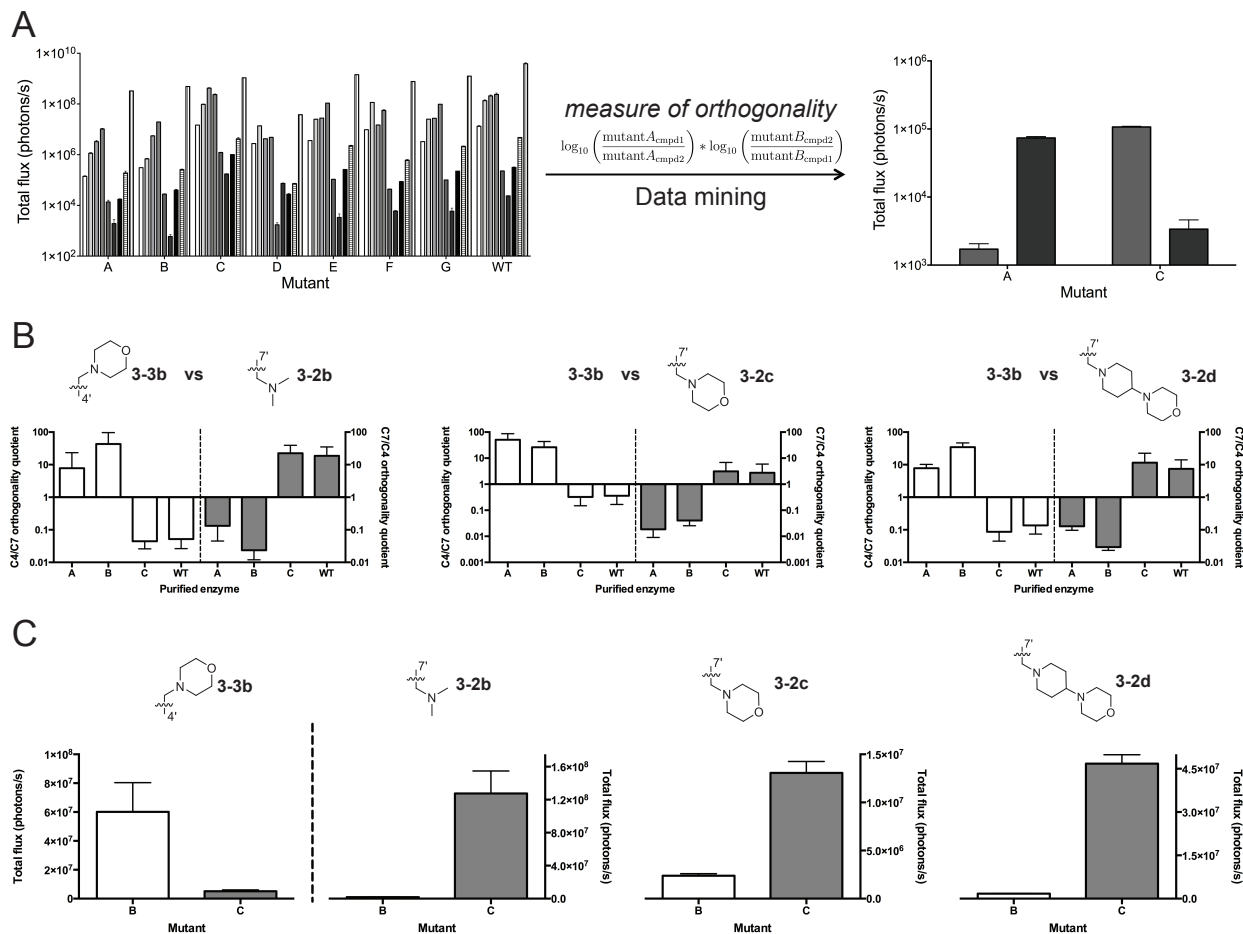


Figure 3-12. Analyzing orthogonal enzyme-substrate pairs. (A) Representative emission of luciferase mutants screened against a panel of luciferin analogs. These data were analyzed with a computer algorithm to determine lead mutants with the strongest orthogonality. (B) Purified mutants exhibit orthogonality. Enzyme (1 μ g) was incubated with 100 μ M of luciferin analogs and emission intensities were used to determine the orthogonality quotient (the ratio of the total flux for the C4/C7 or C7/C4 pairings). The geometric mean is plotted and the error bars represent the 95% confidence intervals for $n > 4$ experiments. (C) Total flux for lead mutants B and C highlights substrate selectivity between C4' and C7' sterically modified luciferins. Error bars represent the standard error of the mean for $n > 4$ experiments.

the number of potential pairings exceeded 3000 in our dataset, we wrote a computer script to rapidly examine all pairs in an unbiased fashion. The script iterated through each possible pairing, calculating the corresponding orthogonality rating for that pairing. Finally, it ranked all pairs, returning a list of hits with the highest rank (and thus utility for multi-component imaging).

The top pairs identified by the script exhibited selectivity for analogs **3-3b** (mutant A and B) and **3-2b-d** (mutant C). The magnitude of each mutant's preference—defined as the orthogonality quotient—was analyzed. As shown in Figures 4B-C, mutants A and B exhibited nearly a 100-fold preference for **3-3b** over other analogs, while mutant C strongly favored C7' modified analogs. Similar trends in orthogonal substrate usage were observed using bacterial lysates (Fig. 3-13) and across a range of luciferin concentrations (Fig. 3-14, 3-15). Biochemical analyses further indicated that the “brightest” mutant enzymes were those capable of most efficient substrate turnover (Table 3-1, 3-2).

3.6 Analyzing the origins of orthogonality

The identities of the mutant “hits” provided some insights into the origins of substrate orthogonality. Mutant A had a single arginine to alanine mutation at amino acid 218. Mutant B comprised the same R218A mutation, but harbored additional mutations at residue 250 (Phe to Met), 314 (Ser to Thr), and 316 (Gly to Thr). These residues are known to play a role in modulating binding and interaction with the luciferin substrate. The R218A mutant is especially interesting, as it is known to greatly reduce light production and red shift emission with D-luciferin [28]. It has been hypothesized that the smaller Ala group allows more water molecules to access the active site, potentially quenching light emission [28]. The bulky morpholino substituent of **3-3b** could fill this active site void to retain photon production. The third mutant (mutant C) was more selective for the C7'-modified luciferins compared to the C4'-modified compound. Mutant C harbored a single mutation, R218K. The R218K mutation may slightly enlarge

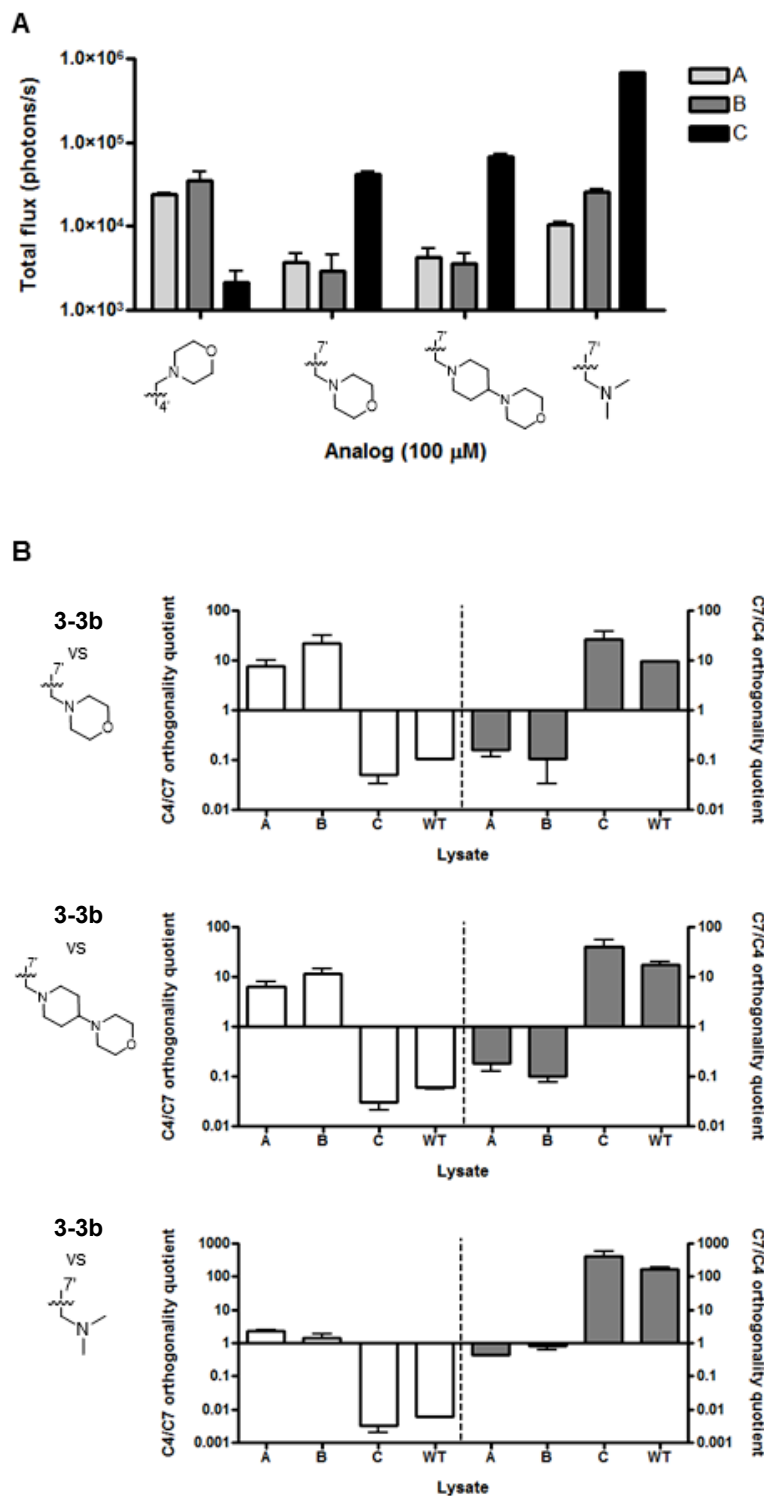


Figure 3-13. Orthogonal substrate usage in mutant or Fluc (WT) lysate. (A) Lysates were imaged with the analogs shown and total bioluminescent light emission is plotted. (B) Orthogonality plots were constructed from data in (A) to compare the indicated analogs with **3-3b**. The geometric mean is plotted and the error bars represent the 95% confidence intervals for triplicate repeats.

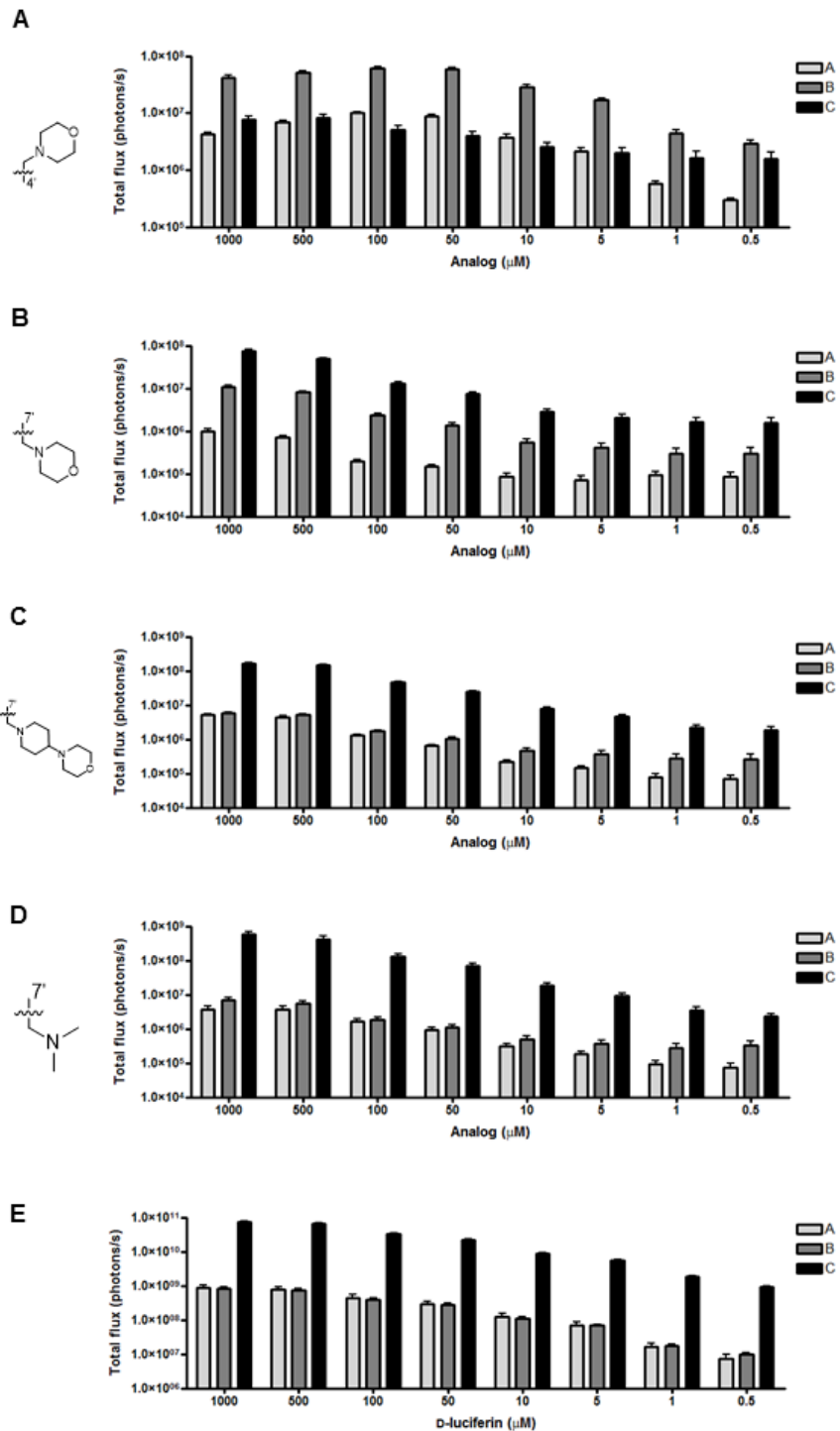


Figure 3-14. Bioluminescent photon production from luciferin analogs and luciferase enzymes. (A)-(D) Luciferin analogs (0.5 – 1000 μM) or (E) D-luciferin were incubated with mutant luciferases in bioluminescence buffer. Images were acquired as described in the Materials and Methods section. Error bars represent the standard error of the mean for $n > 4$ experiments.

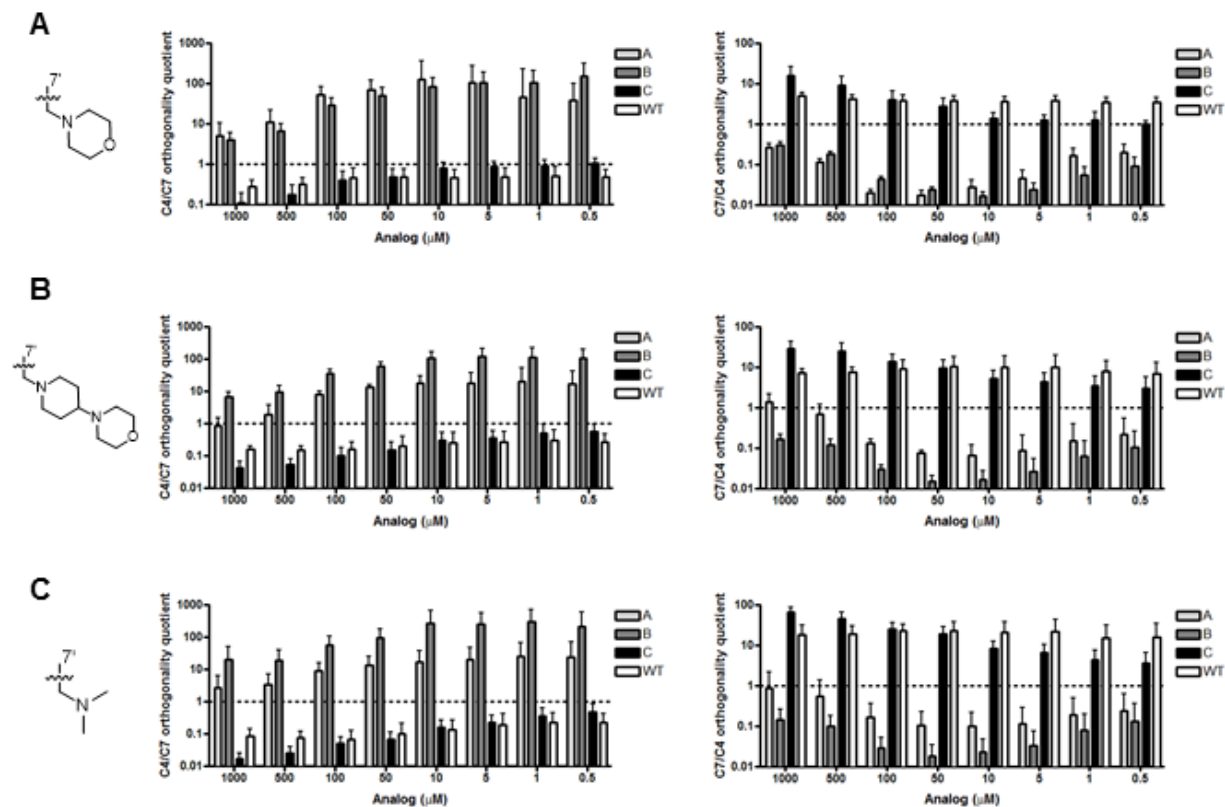
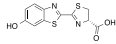
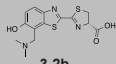
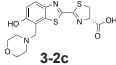
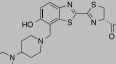
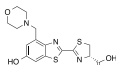


Figure 3-15. Comparative analyses of all analogs with mutants A-C and Fluc (WT). (A)-(C) Analogs (0.5 – 1000 μM) were incubated with mutant luciferase enzymes or Fluc (WT), and images were acquired. Orthogonality plots were constructed to compare **3-3b** with the C7' steric analogs. C4/C7 orthogonality quotient (left) C7/C4 orthogonality quotient (right) indicates preference for C4' or C7' substrates. The geometric mean is plotted and the error bars represent the 95% confidence intervals for $n > 4$ experiments.

the active site of the luciferase. This mutation resulted in improved catalytic activity with some bulkier cyclic amino luciferin analogs [21]. The improved selectivity of mutant C with **2b-d** could be the result of active site positioning. The C7'-substituents could potentially place the luminophore in a more advantageous position in the active site compared to **3-3b**.

To delve into the origins of selectivity, we prepared a small library of additional mutants based on enzyme B (R218A, F250M, S314T, G316T). R218A seemed critical for discriminating the regioisomeric compounds, so this residue was held constant across the series. All possible combinations of the remaining mutations (F250M, S314T,

Table 3-2. Biochemical analyses of orthogonal enzyme-substrate pairs.

Enzyme	% WT light emission ^a	Normalized k_{cat}/K_M ^b	λ_{max} (nm)	Compound
A	1.2 ± 0.35	0.041 ± 0.016	612	 D-luc
B	0.92 ± 0.17	0.013 ± 0.004	616	
C	94 ± 8.4	5.22 ± 0.58	570	
A	0.19 ± 0.02	0.034 ± 0.008	614	 3-2b
B	0.33 ± 0.09	0.050 ± 0.020	614	
C	17 ± 5.2	5.0 ± 1.3	574	
A	0.16 ± 0.02	0.253 ± 0.065	614	 3-2c
B	3.7 ± 0.76	1.09 ± 0.36	618	
C	16 ± 2.3	8.2 ± 2.2	600	
A	0.47 ± 0.01	0.121 ± 0.025	— ^c	 3-2d
B	0.81 ± 0.09	0.155 ± 0.061	604	
C	22 ± 2.3	6.0 ± 1.7	570	
A	38 ± 13	17.1 ± 6.4	622	 3-3b
B	200 ± 41	83 ± 37	628	
C	13 ± 2	13.1 ± 5.7	626	

^a Values normalized to each compound's corresponding emission with WT Fluc. Errors represent standard error of the mean for n = 3 measurements. ^b Kinetic constants are apparent values, determined via measurements of initial rates of light emission over a range of 2 μ M to 10 mM. Errors represent standard error of the mean for n \geq 3 measurements. K_{cat} values are relative to each compound's corresponding value with WT Fluc. Errors represent standard error of the mean for n \geq 3 measurements. ^c λ_{max} value could not be determined due to low level of light emission.

G316T, or native Fluc residues) were then allowed. Imaging analyses of these combinatorial mutants indicated that R218A and F250M were critical for luciferin discrimination (Figs. 3-16, 3-17). Both mutations should result in a larger active site, but why they preferentially accommodate **3-3b** over other analogs remains unknown. It is possible that the mutations disrupt critical binding interactions with the luciferin core, but that steric appendages (e.g., on the C4' side) retain sufficient contacts in the enzyme for subsequent oxidation. Indeed, when **3-3b** was incubated with R218A/F250M, light emission was maintained (as compared to Fluc, Fig. 3-17C). When D-luciferin and the C7' modified analogs were incubated with this same mutant, by contrast, light emission was drastically reduced. Interestingly, the R218A/S314T mutant exhibited an opposite trend in analog selectivity: **3-2b** was preferred to **3-3b** (Fig. 3-17C). Collectively, these results suggest that mutant luciferases can be tuned to respond to unique substrates. It

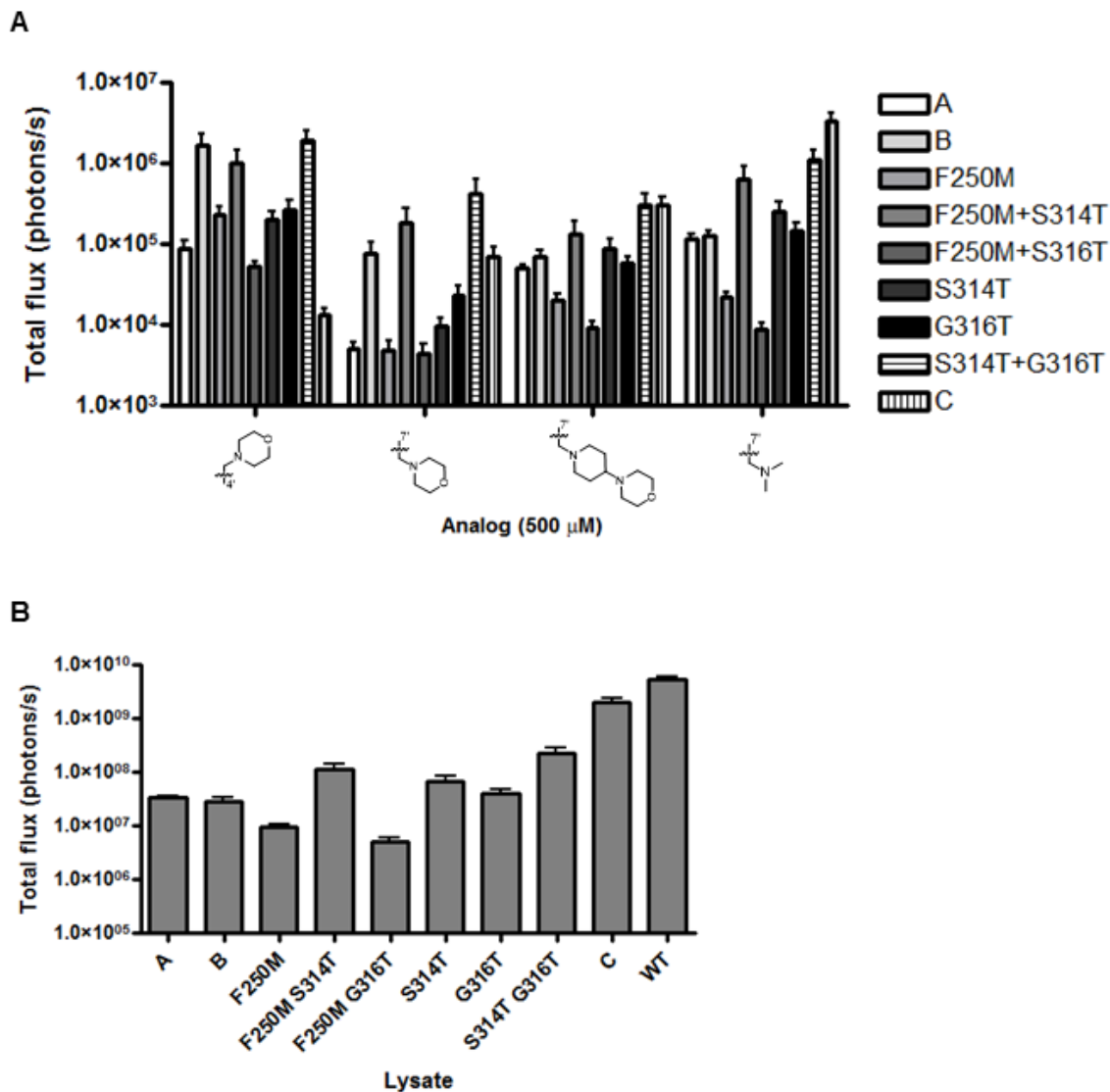


Figure 3-16. Bioluminescent photon production from luciferin analogs with combinatorial enzymes. Combinatorial mutants were prepared, in which R218A was held constant across the library, while the other positions were allowed to code for mutations from mutant B or native Fluc (WT). (A) Luciferin analogs (500 μM) or (B) D-luciferin were incubated with lysates expressing mutant luciferases or Fluc (WT) in bioluminescence buffer. Images were acquired as described in the Materials and Methods section. Error bars represent the standard error of the mean for $n > 4$ experiments.

is also possible that enzyme orthogonality is most readily achieved not by improving the utilization of one substrate, but by diminishing reactivity with all other substrates.

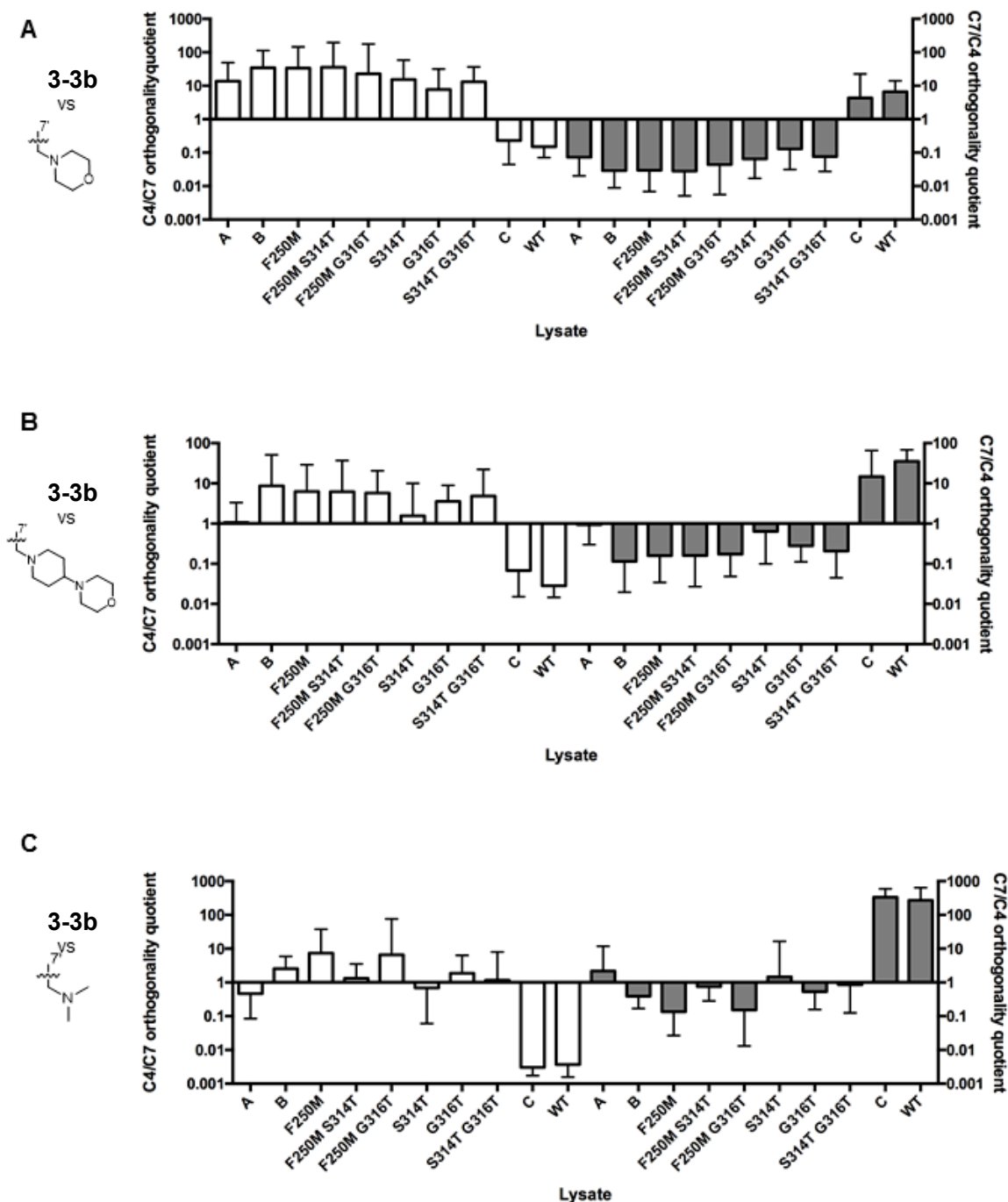


Figure 3-17. Comparative analyses of combinatorial enzymes and Fluc. (A)-(C) Luciferin analogs (500 μ M) were incubated with lysates expressing the indicated mutant enzymes. Images were acquired and orthogonality plots were constructed to compare **3b** with the C7' steric analogs as described in the Materials and Methods section. The geometric mean is plotted and the error bars represent the 95% confidence intervals for $n > 4$ experiments.

3.7 Cellular imaging with orthogonal pairs

As a step toward multi-component imaging applications with the orthogonal enzymes, we evaluated the probes in cultured cell models. Mammalian cell lines (HEK293 and DB7) were engineered to express orthogonal mutants A-C. Equivalent expression levels were confirmed using flow cytometry (Fig. 3-18). Cells were then incubated with analogs **3-2b-d** and **3-3b**, and photon outputs were measured. As shown in Figure 3-19, the substrates were able to cross cell membranes and access the relevant luciferases, resulting in sustained emission. Photon production was also confined to cells expressing the complementary luciferase for each orthogonal luciferin: cells with mutant B were only visible upon treatment with analog **3-3b**, while cells with mutant C were only visible upon treatment with analog **2b-d** (Figs. 3-20–3.23). Importantly, the orthogonal pairs could also distinguish unique cell types in a single imaging session. For example, DB7 cells stably expressing mutants B or C could be

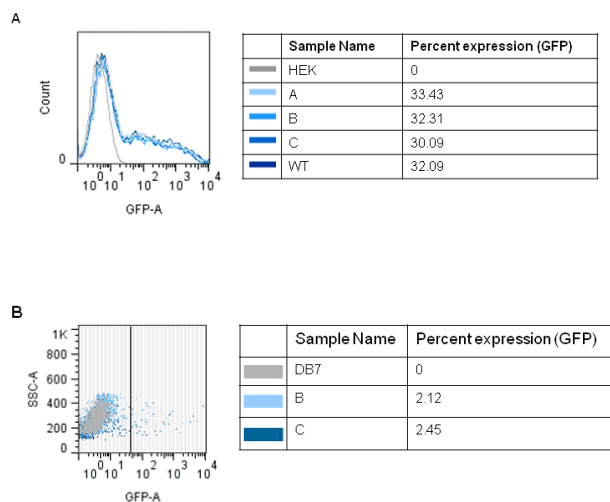


Figure 3-18. Mutant luciferases are expressed at comparable levels in mammalian cells. (A) HEK293 cells or (B) DB7 cells transiently expressing mutant or wild type luciferases were rinsed in FACS buffer and analyzed by flow cytometry. GFP expression was measured. Data shown are representative of (A) $n = 6$; and (B) $n = 2$ replicate experiments.

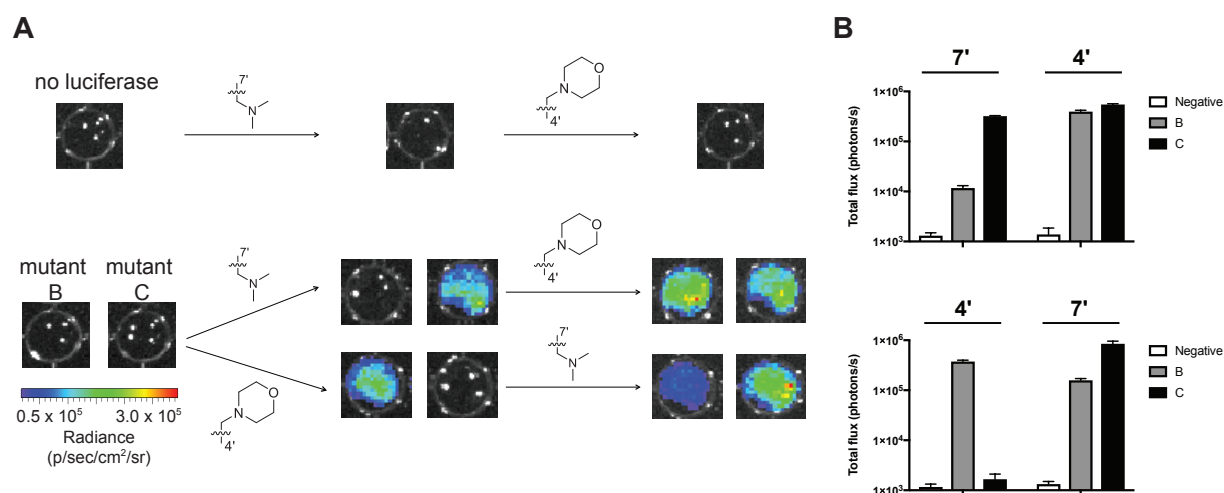


Figure 3-19. Imaging cells with orthogonal luciferase-luciferin pairs. (A) Mutant luciferase-expressing DB7 cells were plated (1.5×10^5 cells/well) in 96-well black plates and sequentially incubated with C4' and C7' sterically modified luciferins (750 μM). Representative bioluminescence images are shown. (B) Quantification of the images from (A). Error bars represent the standard error of the mean for experiments performed in triplicate.

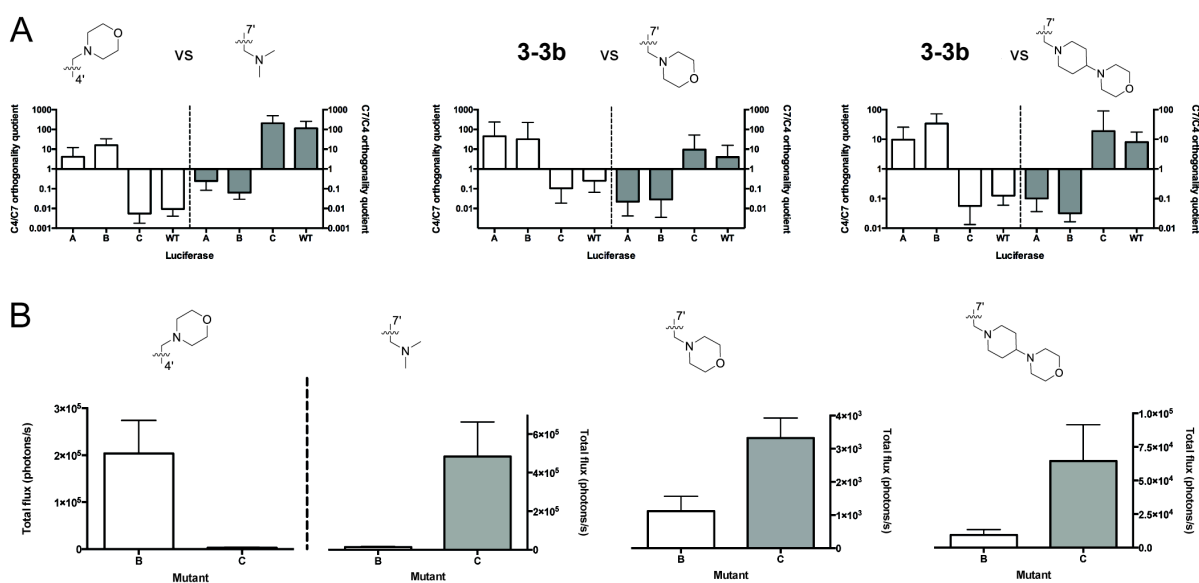


Figure 3-20. Luciferase-expressing cells exhibit orthogonality. (A) HEK293 cells (2×10^5) expressing Fluc or mutant luciferases were incubated with 500 μM of luciferin analogs and the orthogonality quotient was determined. The geometric mean is plotted and the error bars represent the 95% confidence intervals for $n > 4$ experiments. (B) Substrate selectivity between C4' and C7' sterically modified luciferins with lead mutants B and C is maintained in mammalian cells. Error bars represent the standard error of the mean for $n > 4$ experiments.

readily detected via sequential administration of the requisite substrates (Fig. 3-19). Similar trends were observed upon imaging HEK293 cells (Fig. 3-24) and co-cultures (Fig. 2-25). This suggests that cross-reactivity between mutants B and C and their non-orthogonal substrates is minimal. It is also noteworthy that the orthogonal pairs exhibit bioluminescent emission spectra (Fig. 3-26) that can be resolved *in vitro* (Fig. 3-27). Collectively, these data suggest that the enzyme-substrate pairs will be broadly useful for multicomponent imaging.

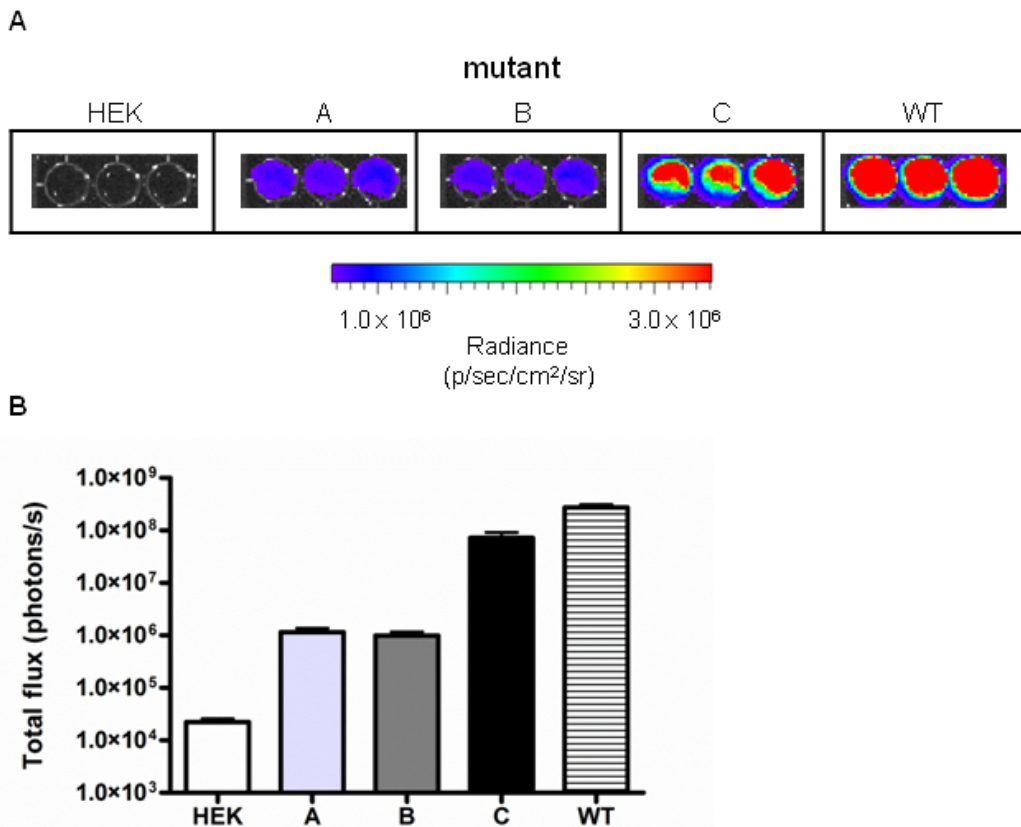


Figure 3-21. Cellular imaging with mutant luciferases and D-luciferin. (A) HEK293 cells expressing Fluc (WT) or mutant luciferases were incubated with D-luciferin (500 μ M). Representative bioluminescence images are shown. (B) Quantification of the images provided in (A). Error bars represent the standard error of the mean for $n = 4$ experiments.

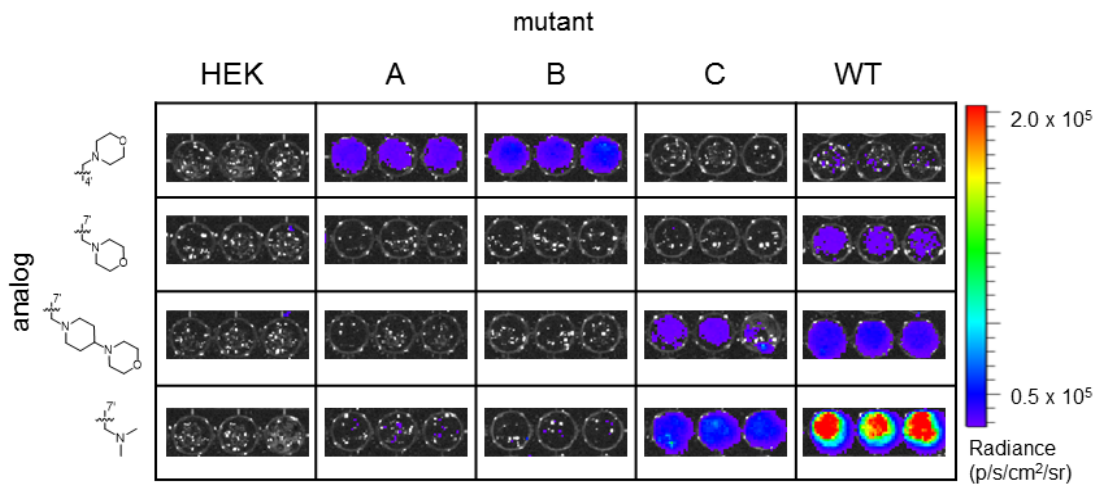


Figure 3-22. Cellular imaging with mutant luciferases and luciferin analogs. HEK293 cells expressing Fluc (WT) or mutant luciferases were incubated with the indicated luciferin analogs (500 μ M). Representative bioluminescence images are shown.

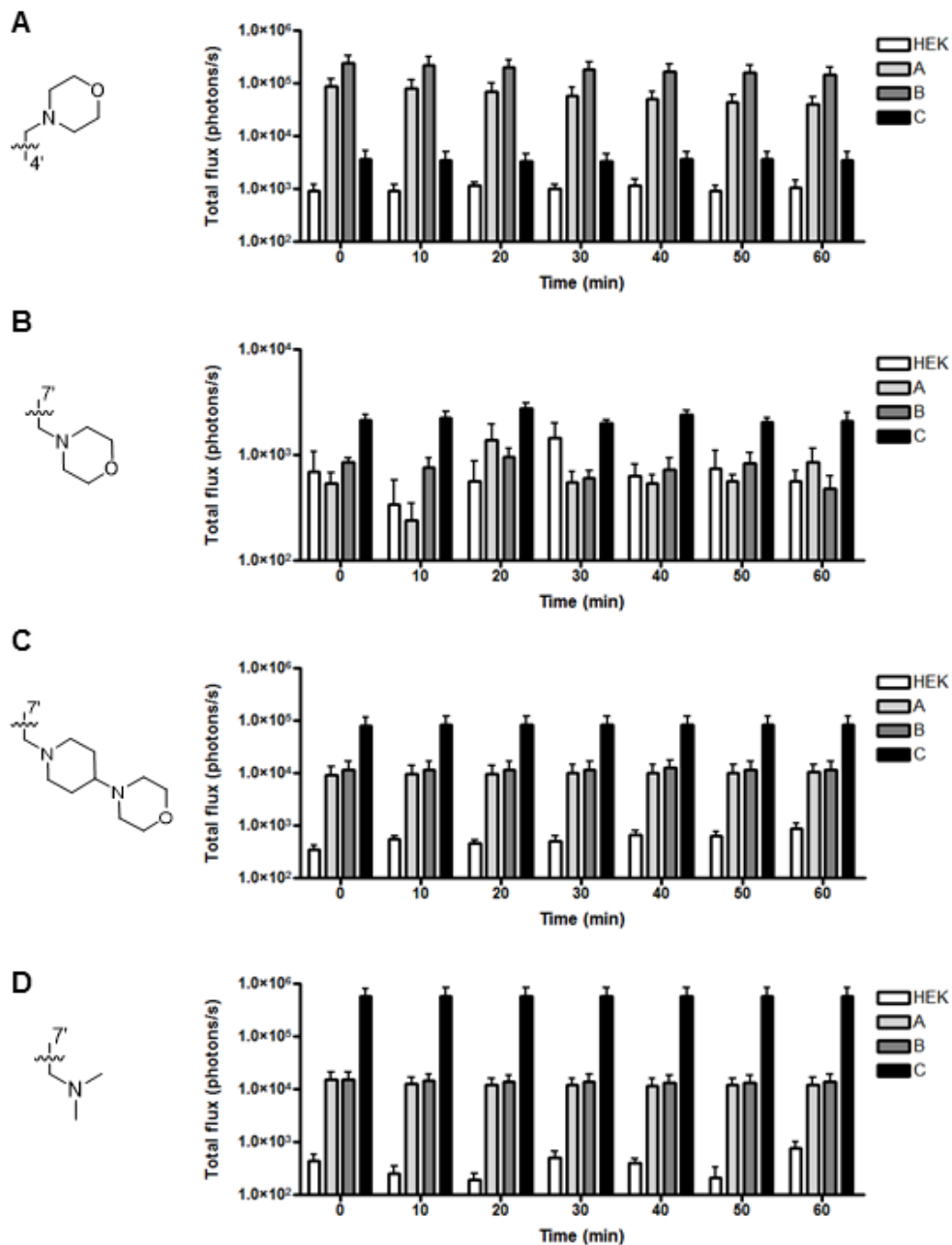


Figure 3-23. Cellular bioluminescent photon production is sustained. (A)-(D) HEK293 cells (200,000 cells/well) expressing mutant luciferases were incubated with 500 μ M of the indicated analog. Images were acquired (0-60 min), and light emission was quantified. Error bars represent the standard error of the mean for $n = 3$ experiments.

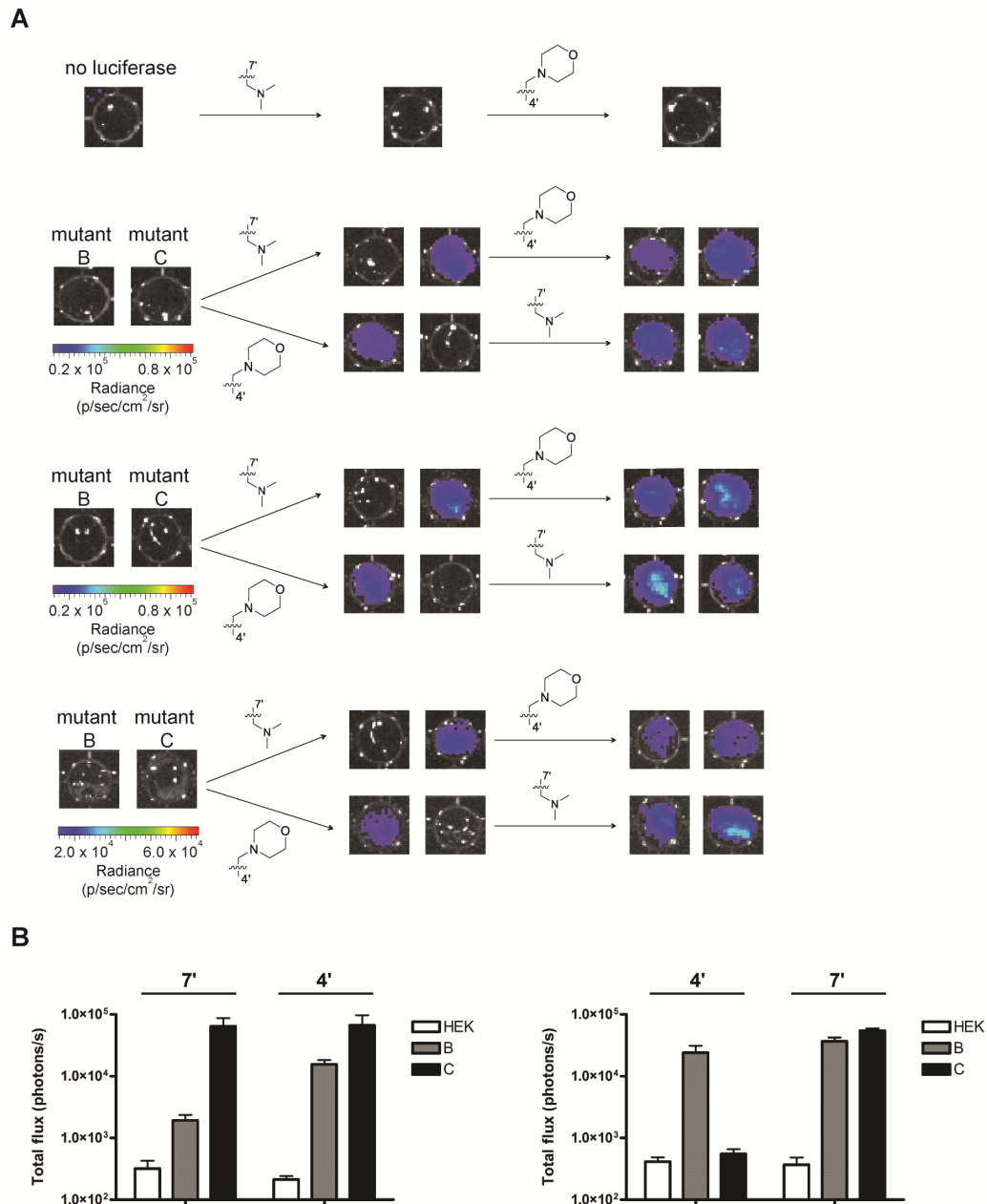


Figure 3-24. Cellular imaging with orthogonal pairs. (A) HEK293 cells expressing mutant B or C (or no luciferase) were imaged with the indicated analog over 10 min. The same cells were then immediately treated with the second orthogonal luciferin and re-imaged. Sample bioluminescent images from 3 independent experiments are shown. (B) Quantification of the images provided from (A). Error bars represent the standard error of the mean for $n = 3$ experiments.

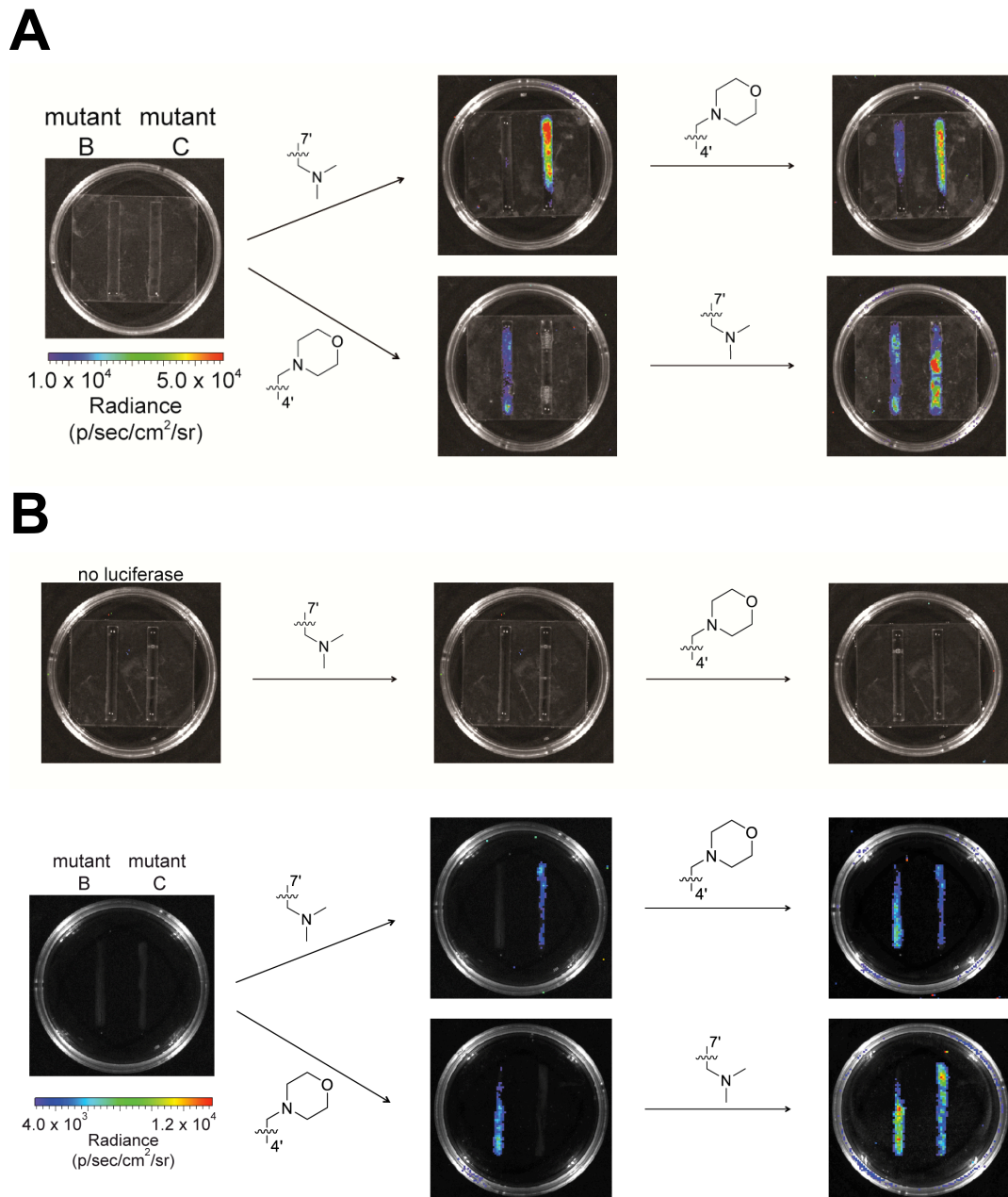


Figure 3-25. Cellular imaging with orthogonal pairs patterned with biocompatible stencils. (A) HEK293 or (B) DB7 cells expressing mutant B or C (or no luciferase) were plated with biocompatible stencils and imaged with the first analog (as shown). The same cells were then immediately treated with the second analog and re-imaged. Sample bioluminescent images from (A) $n = 3$ and (B) $n = 2$ independent experiments are shown.

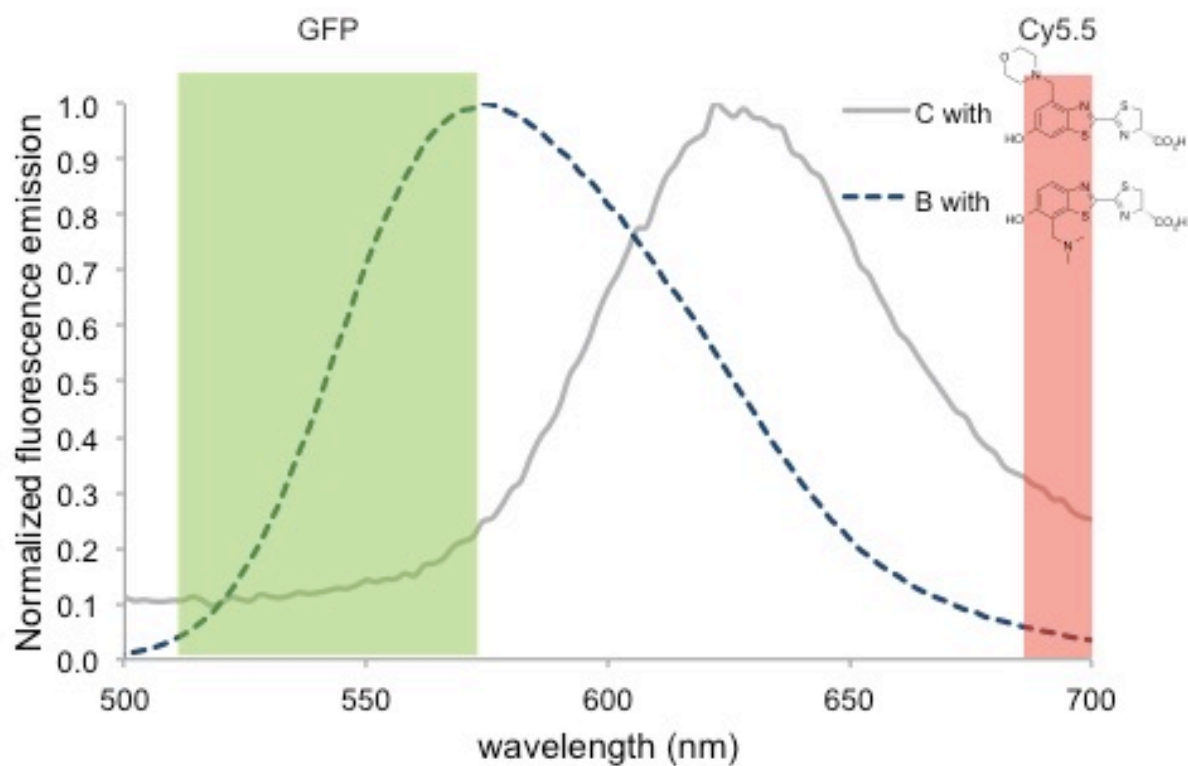


Figure 3-26. Orthogonal bioluminescent pairs can be spectrally resolved. Bioluminescence emission spectra for mutant **C/3-3b** (solid gray) and mutant **B/3-2b** (dotted blue) are shown. The wavelengths collected by GFP and Cy5.5 filters are overlaid.

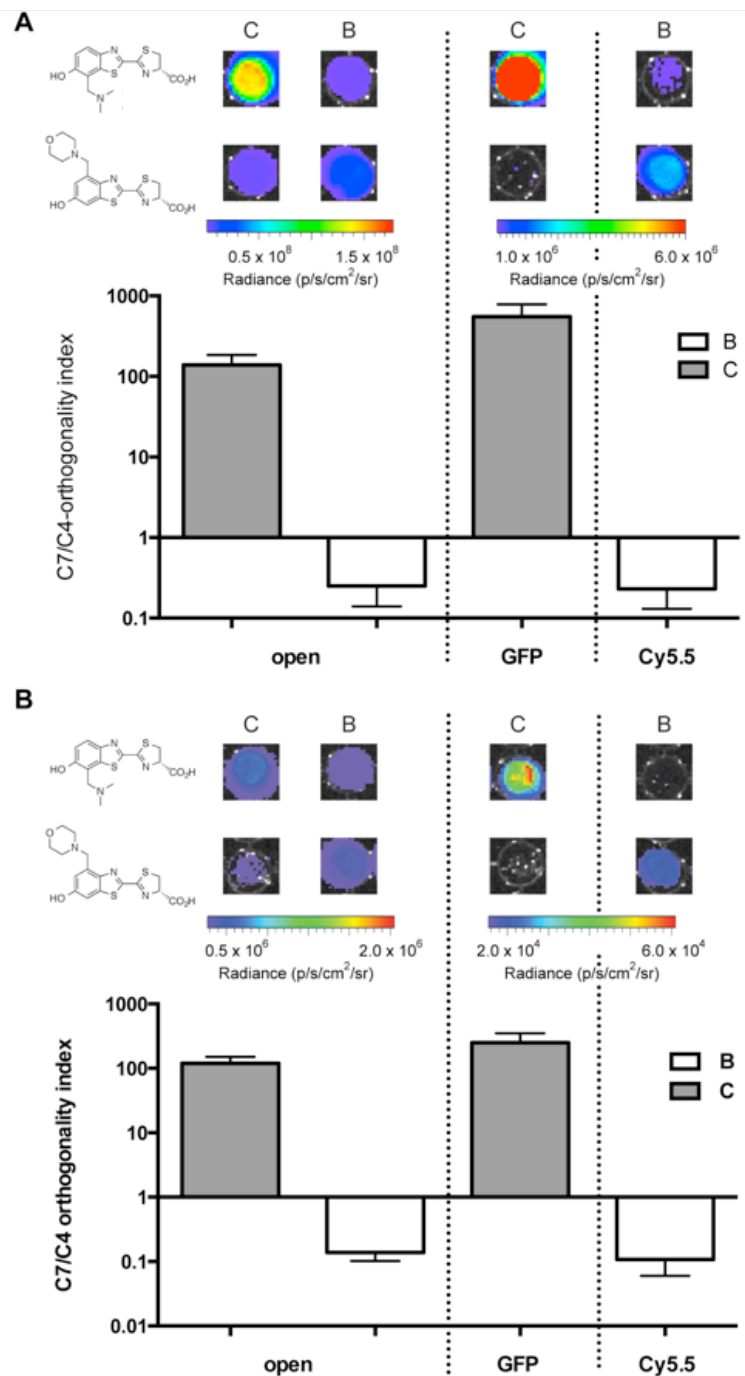


Figure 3-27. Orthogonal luciferase-luciferin pairs can be spectrally resolved *in vitro* and in cells. (A) Bioluminescent orthogonality plots for **3-2b** and **3-3b** with purified mutant enzymes and (B) HEK293 cells expressing mutant B and C. In these experiments, bioluminescent light was collected with either no filter (left), the GFP filter (middle), or the Cy5.5 filter (right). Representative images are also shown. Experiments were performed in triplicate and error bars represent the standard error of the mean for $n = 3$ experiments.

3.8 Conclusions

We developed a general strategy to evolve and identify mutant versions of firefly luciferase that accept distinct, chemically modified luciferins. Bioluminescence has been largely limited to visualizing one biological feature at a time, as the most advantageous luciferases and luciferins for whole animal imaging utilize the same substrate and cannot be distinguished *in vivo*. To address this void, we generated a family of sterically modified luciferins that were poor substrates for firefly luciferase, but inherently capable of producing light. Using an on-plate screen, mutant versions of luciferase were identified that could also catalyze light emission with other analogs. Pools of these functional mutants were then further mined for orthogonal pairs. Some of the mutants could selectively process individual luciferins both *in vitro* and in cells, setting the stage for multi-component *in vivo* imaging.

Future studies will be aimed at generating additional orthogonal pairs and bioluminescent probes with improved brightness. The enzyme-substrate “hits” reported here, while orthogonal and immediately useful, are weaker light emitters than native bioluminescent systems. Improved light outputs can be achieved using additional rounds of mutagenesis and screening. Previous studies have also demonstrated that distant mutations can profoundly influence the architecture of the luciferase active site, and these regions will be incorporated into future libraries. The screening strategy is also broadly applicable to diverse luciferin analogs, and will enable the rapid expansion of the bioluminescence toolkit. Our results suggest that enzymes capable of discriminating even subtle substrate modifications can be readily identified. Such an outcome bodes well for generating additional orthogonal pairs and filling a long-standing

void in imaging capabilities. We anticipate that collections of designer luciferins and luciferases will inspire new discoveries in a variety of disciplines, similar to how fluorescent protein technology enabled seminal advancements in numerous fields.

3.9 Materials and Methods

3.9a General biological methods

Fluorescent spectra and assays

All fluorescence spectra were acquired using clear 96-well microplates (BD-Falcon or Greiner-Bio) and 100 μ L sample volumes. Measurements were recorded on a Molecular Devices SpectraMax Gemini XPS microplate reader. Measurements were acquired in triplicate, and the data were analyzed using Microsoft Excel.

Bioluminescence emission spectra

Emission spectra for all luciferin analogs were recorded on a Horiba Jobin Yvon FluoroMax-4 spectrometer. Each luciferin (2 mM or 5 mM) was incubated in a 10 mm path length cuvette with ATP (2 mM or 5 mM), LiCoA (2 mM) and reaction buffer (20 mM Tris-HCl, 0.5 mg/mL BSA, 0.1 mM EDTA, 1 mM TCEP, 2 mM MgSO₄, pH 7.6) totaling 900 μ L. Purified luciferase enzyme (10-20 μ g) was added immediately prior to data acquisition. The excitation and emission slit widths were adjusted to 0 and 29.4 nm, respectively. Emission data were collected at 2 nm intervals from 400-750 nm at ambient temperature. The acquisition times varied from 0.1-10 sec/wavelength depending on the amount of light produced from each sample. Light emissions were recorded Relative Luminescence Units (RLU), and the intensities were normalized.

UV-Vis spectra

Luciferin absorption spectra were acquired on a Thermo-Fisher NanoDrop spectrophotometer (ND-2000c). Each sample (100 μ M) was prepared by dissolving the luciferin 50 mM PBS pH 7.4. Absorbance data were analyzed using Microsoft Excel.

3.9b General bioluminescence imaging

All analyses were performed in black 96-well plates (Grenier Bio One). Plates containing luminescent reagents were imaged in a light-proof chamber with an IVIS Lumina (Xenogen) CCD camera chilled to -90 °C. The stage was kept at 37 °C during the imaging session, and the camera was controlled using Living Image software. For assays with purified enzymes, exposure times ranged from 0.5-30 s, and data binning levels were set to small or medium. For assays with bacterial cell lysates, the exposure times for ranged from 1 s to 5 min, with data binning levels set to small or medium. Regions of interest were selected for quantification and total flux values were analyzed using Living Image software.

3.9c General chemiluminescence procedure

Carried out as described previously, in Steinhardt *et al.* [53].

3.9d Mutant luciferase library design

Table of site directed libraries: Summary of site-directed library identities, sizes, and relative phylogenetic conservation at each residue. Codon sets defined as in Reetz, *et al.* and Kille, *et al.* [58-59] Conservation designations based on 13 aligned insect luciferase sequences [11].

Library	Residue	Conserved	Codon set	Nucleic Acid Diversity	Amino Acid Diversity
SD249	M249	moderate	Hydrophobic	3388	2800
	F250	low	22c-trick		
	T251	moderate	22c-trick		
SD314	S314	moderate	22c-trick	5808	4800
	G315	high	NDT		
	G316	mod-high	22c-trick		
SD218	R218	high	All besides WT	19	19

Definition of codon sets from.

Name	Codons used	Total codons	Total amino acids	Reference
22c-trick	NDT, VHG, TGG	22	20	Kille <i>et al.</i> [59]
NDT	NDT	12	12	Reetz <i>et al.</i> [58]
Hydrophobic*	GBC, WTK	7	7	Pines <i>et al.</i> [62]
All – Phe*	TGK, VBC, ATG, NAW	20	19	Pines <i>et al.</i> [62]
All – Arg*	KGG, WKT, VHG, NAT	19	19	Pines <i>et al.</i> [62]

*Identified using algorithm provided by Pines *et al.* [62]

3.9e Mutant luciferase library generation

Two sections of the luciferase gene (*pgl4-luc2*), denoted R1 and R2, were targeted for gene assembly. The R1 region comprises amino acids 199-275 and was assembled with primers R1-F0 to R1-F-235 (forward primers table) and R1-R0 to R1-R119 (reverse primers table). The R2 region comprises amino acids 275-346 and was assembled with

primers R1-F0 to F264 (forward) and R1-R0 to R228 (reverse). The gene assembly primers were designed using gene2oligo to identify desirable T_m values (see Tables S3 and S4) [65].

To assemble mutant libraries, primers containing the codon(s) of interest (primer tables S5-7) were used in place of the primer coding for the wild-type sequence. Libraries and primers are named “SD” (site directed), followed by the number of the first residue mutated in the library (e.g., SD218 is the library based on R218). Primers with degenerate codons were mixed in proportion to the number of amino acids encoded by the primer (See Kille, *et al.* for a detailed protocol) [59]. Libraries were assembled according to Bessette, *et al.* [66] All PCR reactions were run using Q5 Hot-start DNA polymerase (New England BioLabs).

3.9f Circular polymerase extension cloning (CPEC)

The libraries generated above were ultimately inserted into a vector lacking the R1 and R2 regions. This deletion vector was generated from pET28a-luc2 using overlapping PCR. Primers Nco-Luc2-For and R1-deletion Rev (PCR1), along with primers R1-deletion-For and Luc2-Not-Rev (PCR2) were used to remove R1. The R2 region was similarly removed using primers Nco-Luc2-For and R2-deletion-Rev (PCR1) and primers R2-deletion-For and Luc2-Not-Rev (PCR2). The PCR1 and PCR2 products were assembled using Nco-Luc2-For and Luc2-Not-Rev (Table S8). The resulting gene products were digested with *NcoI-HF* and *NotI-HF* (New England BioLabs) and inserted into pET28a with T4 ligase (New England BioLabs). The resulting vectors were labeled

pET28-R1del and pET28-R2del, referring to the R1 and R2 deleted regions, respectively.

Library DNA was ultimately inserted into pET28-R1del or pET28-R2del [37]. The pET28-R1del or pET28-R2del were linearized by amplification using primers R1-vector-Fwd and R1-vector-Rev (for pET28-R1del) or R2-vector-Fwd and R2-vector-Rev for pET28-R2del (Table S9). 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* gel electrophoresis to remove any remaining circular template DNA. The resulting linear vector (500 ng) was combined in a 1:1 or 2:1 molar ratio with library insert DNA (targeting region R1 or R2, described above). The insert and vector were combined via CPEC [64] using the following conditions: 1x Q5 Hot start DNA polymerase reaction buffer, dNTPs (0.8 mM), and Q5 Hot start DNA polymerase (1 U) in a total volume of 50 mL. The following thermal cycling conditions were used: initial denaturation at 95 °C for 60 s; 5-10 cycles of denaturation (95 °C, 30 s), slow ramp anneal (72 – 55 °C over 2.5 min), and extension (72 °C, 132 s). The final step involved extension at 72 °C for 5 min. The PCR products were then analyzed using 1% agarose gels and ethidium bromide staining (10 mL of the reaction). A portion of each product (2.5 mL) was utilized directly for transformation into T7 Express *lysY* competent *E. coli* (New England Biolabs) cells. Transformants were plated according to the primary screening protocol described below.

3.9g Primer lists

All primers were purchased from Integrated DNA Technologies, Inc. (San Diego, CA) and are written in the 5′→3′ direction. Upper case letters denote bases coding for the luciferase gene. Lower case letters denote bases added to ensure similar melting temperatures (T_m) for all primers. Bases highlighted in red denote sites targeted for saturation mutagenesis.

Primer tables: Region 1 (R1, wild type *luc2* primers). Lower case letters denote non-luciferase sequences included for maintaining proper T_m 's and do not appear in the final assembled gene.

Forward primers	
F0	GGATCCACCGGATTGCCCAAGGGCGTAGCCCTACC
F35	GCACCGCACCGCTTGTGTCCGATTCAGTCATGCC
F70	GCGACCCCATCTTCGGCAACCAGATCATCCCCGACA
F106	CCGCTATCCTCAGCGTGGTGCCATTCACCACGGC
F141	TTCGGCATGTTACCACGCTGGGCTACTTGATCTGCG
F178	GCTTTCGGGTCGTGCTCATGTACCGCTTCGAGGAG
F213	GAGCTATTCTTGCGAAGCTTcgccctgtacagtcgctg
F235	cagcgactgtacagggc
Reverse primers	
R0	GGCAATCCGGTGGATCC
R18	ACAAGCGGTGCGGTGCGGTAGGGCTACGCCCTT
R51	GCCGAAGATGGGGTCGCGGGCATGACTGAATCGGAC
R87	CCACGCTGAGGATAGCGGTGTGCGGGATGATCTGGTT
R124	CGTGGTGAACATGCCGAAGCCGTGGTGAATGGCA
R159	AGCACGACCCGAAAGCCGAGATCAAGTAGCCAG
R194	cgAAGCTTCGCAAGAATAGCTCCTCCTCGAAGCGGTACATG

Region 2 (R2, wild-type *luc2* primers) Lower case letters denote non-luciferase sequences included for maintaining proper T_m 's and do not appear in the final assembled gene.

Forward primers	
F0	acacatCGAGGAGGAGCTATTCTTGCGAAGCTTGCA
F36	AGACTATAAGATTCAATCTGCCCTGCTGGTGCCACAC

F74	TATTTAGCTTCTTCGCTAAGAGCACTCTCATCGACAAGTACGAC
F117	CTAAGCAACTTGCACGAGATCGCCAGCGGCGGG
F150	GCGCCGCTCAGCAAGGAGGTAGGTGAGGCC
F181	GTGGCCAAACGCTTCCACCTACCAGGCATCCG
F213	CCAGGGCTACGGCCTGACAGAAACAACACTAGTGCCA
F248	TTCTGATCACCCCgctgcctccggaggaga
F264	tctcctccggaggcg
Reverse primers	
R0	TAGCTCCTCCTCGatgtgt
R19	GGGCAGATTGAATCTTATAGTCTTGCAAGCTTCGCAAGAA
R59	GCTCTTAGCGAAGAAGCTAAATAGTGTGGGCACCAGCA
R97	TCGTGCAAGTTGCTTAGGTCGTAAGTGTGCGATGAGAGT
R135	TTGCTGAGCGGCGCCCCGCCGCTGGCGATC
R165	GGAAGCGTTTGGCCACGGCCTCACCTACCTCC
R197	AGGCCGTAGCCCTGGCGGATGCCTGGTAGGT
R228	acGGGGGTGATCAGAATGGCACTAGTTGTTTCTGTC

Primers used to construct site-directed libraries. The bases highlighted in red denote sites targeted for saturation mutagenesis.

SD218 primers	
Forward primers	
SD218-F1	GCACCGCACCGCTTGTGTC KGG TTCAGTCATGCCC
SD218-F2	GCACCGCACCGCTTGTGTC WKT TTCAGTCATGCCC
SD218-F3	GCACCGCACCGCTTGTGTC VHG TTCAGTCATGCCC
SD218-F4	GCACCGCACCGCTTGTGTC NAT TTCAGTCATGCCC
Reverse primers	
SD218-R1	GCCGAAGATGGGGTTCGCGGGCATGACTGAA CM GAC
SD218-R2	GCCGAAGATGGGGTTCGCGGGCATGACTGAA AMW GAC
SD218-R3	GCCGAAGATGGGGTTCGCGGGCATGACTGAA CDB GAC
SD218-R4	GCCGAAGATGGGGTTCGCGGGCATGACTGAA ATN GAC

SD249 primers	
Forward primers	
SD249-F1	CACGGCTTCGGC GBCNDTNDT ACGCTGGGCTACTTGATCTGCGG
SD249-F2	CACGGCTTCGGC GBCNDTVHG ACGCTGGGCTACTTGATCTGCGG
SD249-F3	CACGGCTTCGGC GBCNDTTGG ACGCTGGGCTACTTGATCTGCGG
SD249-F4	CACGGCTTCGGC GBCVHGNDT ACGCTGGGCTACTTGATCTGCGG
SD249-F5	CACGGCTTCGGC GBCVHGVHG ACGCTGGGCTACTTGATCTGCGG
SD249-F6	CACGGCTTCGGC GBCVHGTGG ACGCTGGGCTACTTGATCTGCGG
SD249-F7	CACGGCTTCGGC GBCTGGNDT ACGCTGGGCTACTTGATCTGCGG

SD249-F8	CACGGCTTCGGC GBCTGGVHG ACGCTGGGCTACTTGATCTGCGG
SD249-F9	CACGGCTTCGGC GBCTGGTGG ACGCTGGGCTACTTGATCTGCGG
SD249-F10	CACGGCTTCGGC WTKNDTNDT ACGCTGGGCTACTTGATCTGCGG
SD249-F11	CACGGCTTCGGC WTKNDTVHG ACGCTGGGCTACTTGATCTGCGG
SD249-F12	CACGGCTTCGGC WTKNDTTGG ACGCTGGGCTACTTGATCTGCGG
SD249-F13	CACGGCTTCGGC WTKVHGNDT ACGCTGGGCTACTTGATCTGCGG
SD249-F14	CACGGCTTCGGC WTKVHGVHG ACGCTGGGCTACTTGATCTGCGG
SD249-F15	CACGGCTTCGGC WTKVHGTGG ACGCTGGGCTACTTGATCTGCGG
SD249-F16	CACGGCTTCGGC WTKTGGNDT ACGCTGGGCTACTTGATCTGCGG
SD249-F17	CACGGCTTCGGC WTKTGGVHG ACGCTGGGCTACTTGATCTGCGG
SD249-F18	CACGGCTTCGGC WTKTGGTGG ACGCTGGGCTACTTGATCTGCGG
Reverse primers	
SD249-R1	GVCAHNAHN GCCGAAGCCGTGGTCAAATGG
SD249-R2	GVCAHNCDB GCCGAAGCCGTGGTCAAATGG
SD249-R3	GVCAHNCCA GCCGAAGCCGTGGTCAAATGG
SD249-R4	GVCCDBAHN GCCGAAGCCGTGGTCAAATGG
SD249-R5	GVCCDBCDB GCCGAAGCCGTGGTCAAATGG
SD249-R6	GVCCDBCCA GCCGAAGCCGTGGTCAAATGG
SD249-R7	GVCCCAAHN GCCGAAGCCGTGGTCAAATGG
SD249-R8	GVCCACDB GCCGAAGCCGTGGTCAAATGG
SD249-R9	GVCCACCA GCCGAAGCCGTGGTCAAATGG
SD249-R10	MAWAHNNDT GCCGAAGCCGTGGTCAAATGG
SD249-R11	MAWAHNCDB GCCGAAGCCGTGGTCAAATGG
SD249-R12	MAWAHNCCA GCCGAAGCCGTGGTCAAATGG
SD249-R13	MAWCDBAHN GCCGAAGCCGTGGTCAAATGG
SD249-R14	MAWCDBCDB GCCGAAGCCGTGGTCAAATGG
SD249-R15	MAWCDBCCA GCCGAAGCCGTGGTCAAATGG
SD249-R16	MAWCCAAHN GCCGAAGCCGTGGTCAAATGG
SD249-R17	MAWCCACDB GCCGAAGCCGTGGTCAAATGG
SD249-R18	MAWCCACCA GCCGAAGCCGTGGTCAAATGG

SD314 primers	
Forward primers	
SD314-F1	CTAAGCAACTGCACGAGATCGCC NDTNDTNDT
SD314-F2	CTAAGCAACTGCACGAGATCGCC NDTNDTVHG
SD314-F3	CTAAGCAACTGCACGAGATCGCC NDTNDTTGG
SD314-F4	CTAAGCAACTGCACGAGATCGCC VHGNDTNDT
SD314-F5	CTAAGCAACTGCACGAGATCGCC VHGNDTVHG
SD314-F6	CTAAGCAACTGCACGAGATCGCC VHGNDTTGG
SD314-F7	CTAAGCAACTGCACGAGATCGCC TGGNDTNDT

SD314-F8	CTAAGCAACTTGCACGAGATCGCCTGGNDTVHG
SD314-F9	CTAAGCAACTTGCACGAGATCGCCTGGNDTTGG
Reverse primers	
SD314-R1	TTGCTGAGCGGCGCAHNAHNAHNGGCGATC
SD314-R2	TTGCTGAGCGGCGCCDBAHNAHNGGCGATC
SD314-R3	TTGCTGAGCGGCGCCAAHNAHNGGCGATC
SD314-R4	TTGCTGAGCGGCGCAHNAHNCDBGGCGATC
SD314-R5	TTGCTGAGCGGCGCCDBAHNCDBGGCGATC
SD314-R6	TTGCTGAGCGGCGCCAAHNCDBGGCGATC
SD314-R7	TTGCTGAGCGGCGCAHNAHNCCAGGCGATC
SD314-R8	TTGCTGAGCGGCGCCDBAHNCCAGGCGATC
SD314-R9	TTGCTGAGCGGCGCCAAHNCAGGCGATC

Primers used to create R1 and R2 deletions in pET Luc2

Outside primers	
Nco-Luc2-For	ATACCATGGATGGAAGATGCCAAAAACATTAAGAAGGG
Luc2-Not-Rev	ATAGGCCGCCACGGCGATC
R1 deletion	
R1-deletion-For	GAACAGTAGTGGATCCAAGCTTGCAAGACTATAAGATTCAATCTGCC C
R1-deletion-Rev	TAGTCTTGCAAGCTTGGATCCACTACTGTTCATGATCAGGGC
R2 deletion	
R2-deletion-For	TTCTTGCGAAGCTTACTAGTGCCATTCTGATCACCCCCG
R2-deletion-Rev	CAGAATGGCACTAGTAAGCTTCGCAAGAATAGCTCCTCCTCG

Primers used to amplify inserts for the R1 or R2 regions, along with the pET vector backbone.

R1 insert amplification primers	
R1-insert-Fwd	GACAAAACCATCGCCCTGATCATGAACAGTAGTGGATCCACCGGAT TGCCCAA
R1-insert-Rev	CACCAGCAGGGCAGATTGAATCTTATAGTCTTGCAAGCTTCGCAAGA ATAGCTCCTCCTC
R1- vector amplification primers	
R1-vector-Fwd	ACTACTGTTCATGATCAGGGCGATGGTTTTGTC
R1-vector-Rev	GCAAGACTATAAGATTCAATCTGCCCTGCTGGTG
R2 insert amplification primers	

R2-insert-Fwd	CGAGGAGGAGCTATTCTTGCG
R2-insert-Rev	GGGGGTGATCAGAATGGCACTA
R2- vector amplification primers	
R2-vector-Fwd	TAGTGCCATTCTGATCACCCCC
R2-vector-Rev	CGCAAGAATAGCTCCTCCTCG

3.9h Primary screening protocol

A portion of each library (2.5 – 3.5 mL) was transformed into 50 mL of T7 Express *lysY* competent *E. coli* (New England Biolabs), according to the manufacturer's protocol. During the 60 min recovery step, a solution of 1X kanamycin (Kan, 40 µg/ml, Fisher Scientific), 2X IPTG (1 mM, Gold Biotechnology), and 2X luciferin (0.2 mM to 1 mM) was made and stored in a 42 °C water bath for 10 min. In addition, a pre-made solution of 2X LB broth, 1X Kan and 0.5X agar was liquefied in a microwave (3 min, 30% power). This solution was then cooled in a 42 °C water bath. Once the 60 min bacterial recovery period was complete, the agar and luciferin stocks (1 mL each) were mixed with 20-250 mL of the transformants. This solution was swirled, and immediately spread on square, gridded, LB-Kan plates. The plates were left to sit right side up for 15 min (allowing the agar to solidify), then placed upside down and incubated at 37 °C overnight. The following day, the plates were imaged using the bioluminescence imaging system described previously. Acquisition times ranged from 2-10 min per plate. All colonies emitting detectable levels of light were marked for subsequent analysis in the secondary screen.

3.9i Secondary screening of library members

Light-emitting colonies from the agar plates (primary screen) were picked and added to 5 mL of LB broth (containing 40 µg/mL Kan). Colonies expressing wild-type Fluc (or another luciferase from an earlier round) were also picked for growth and comparative analysis. At least two colonies (per thirty library members) were picked and grown up simultaneously, as controls. All culture tubes were shaken and incubated at 37 °C for 2-5 h until the OD₆₀₀ reached ~0.8. Rapidly growing cultures were stored at 4 °C until all cultures reached the desired OD₆₀₀. A portion of each culture (0.5 mL) was reserved and stored at 4 °C for sequencing analysis. IPTG (0.5 mM) was added to the remainder of the cultures to induce protein expression for 2 h at 30 °C or for 12 h at 25 °C.

After induction, the cultures were centrifuged for 10 min at 3400 x *g* (4 °C). The supernatant was removed, and the pellet was resuspended in 0.4 – 0.6 mL of Fluc lysis buffer (50 mM Tris•HCl, 500 mM NaCl, 0.5% (v/v) Tween, 5 mM MgCl₂, pH = 7.4). Lysed bacterial suspensions (90 µL/well) were added to a 96-well black plate (in triplicate). The lysed cultures were imaged using a stock solution of 10X luciferin and 10X ATP in PBS (1-5 mM luciferin, 10 mM ATP). [59] Library members emitting light equivalent to or greater than wild-type Fluc (first round of screening) or the “hits” from previous rounds (subsequent rounds of screening) were further analyzed. The reserved cultures for these mutants were added to 3.5 mL of LB-Kan and grown in a shaking incubator at 37 °C overnight. DNA from the colonies of interest was harvested and sequenced.

3.9j Identification of luciferase mutant B selectivity

A series of mutants were constructed holding R218A constant and allowing for all possible combinations of native Fluc or mutant sequences at the remaining three residues (F250M, S314T, G316T). The R218A mutation was constructed using overlap PCR with the following primers:

5'-ACCGCTTGTGTC**GCG**TTCAGTCATGCC-3' and

5'-CTTCGGGGGTGATCAGAATGGC-3' for PCR 1

5'-GACAAAACCATCGCCCTGATCATGAAC-3' and

5'-GGCATGACTGAA**GCG**GACACAAGCGGT-3' for PCR 2

Bases highlighted in red denote residue 218 mutated to alanine. The vector backbone was linearized and amplified from pET28a *luc2* encoding wild-type Fluc or a mutant (F250M, S314T, and/or G316T) sequence using primers R1-vector-Fwd and R2-vector-Rev (Primer tables). The insert and vector were combined by CPEC (as above). A portion of the reaction (2.5 mL) was then utilized directly for transformation into XL1 competent *E. coli*. Sequencing analysis confirmed correct mutation and assembly.

3.9k Mammalian plasmid construction

To express the mutant luciferases in mammalian cells, the R1-R2 region of the mutant luciferases was amplified and inserted into pcDNA-Luc2-IRES-GFP[67] or the entire gene was amplified and inserted into pBMNpuro vector (courtesy of Nolan laboratory, Stanford).

The R1 and R2 regions of mutants A-C were amplified with the following primers:

5'-GACAAAACCATCGCCCTGATCATGAAC-3' and

5'-CTTCGGGGGTGATCAGAATGGC-3'

Linearized vector was generated with R1 WT primer R18 (Table S3) and R2 WT primer F181 (Table S4) amplified from pcDNA-Luc2-IRES-GFP. The amplification product was combined with linear vector by CPEC. A portion of the reaction (2.5 mL) was then utilized directly for transformation into XL1 competent *E. coli*. Sequencing analysis confirmed correct mutations and assembly.

The Luc2 gene of mutants A-C were amplified with the following primers:

5'-ATAACGCGTATGGAAGATGCCAAAACAATAAGA-3' and

5'-GAGAGGGATGCATTTATTACACGGCGATCTTGCC-3'

The insert contained *NsiI* and *MluI* restriction sites. The luciferase genes were then cloned into the pBMN vector upstream of the IRES sequence.

3.9I Recombinant protein expression and purification

The pET-luciferase plasmids (WT, mutants A, B, and C) were transformed into chemically competent *E. coli* BL21 cells. Cultures (1 L) were grown in 2 L flasks at 37 °C in LB broth (containing 40 mg/mL Kan) to mid-log phase ($OD_{600} = 0.8 - 1.0$), induced with 0.5 mM IPTG, and incubated at 22 °C for 16 – 18 h. The cells were harvested by centrifugation at 4 °C and resuspended in 20 mL of a solution of 50 mM $NaPO_4$, 300 mM NaCl, 1 mM DTT, and 1 mM PMSF at pH 7.4. Lysozyme (1 mg) was added, and the cells were sonicated and centrifuged at 10000 x *g* for 1 h at 4 °C. WT Fluc and mutant luciferases were purified from clarified supernatants using nickel affinity chromatography (BioLogic Duo Flow Chromatography System, Bio-Rad). Proteins were dialyzed into 25 mM Tris-acetate (pH 7.8) buffer 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 using a standard BCA assay or UV spectroscopy. SDS-PAGE was also performed, and gels were stained with Coomassie R-250.

3.9m Light emission assays with recombinant luciferase

Bioluminescence assays with all luciferin compounds were carried out in triplicate, using solid black, flat-bottom, 96-well plates (BD Bioscience). Assay wells contained purified Fluc (1-2 µg), luciferin substrate (0-1 mM), ATP (Sigma, 0-1 mM), and reaction buffer (20 mM Tris-HCl, 0.5 mg/mL BSA, 0.1 mM EDTA, 1 mM TCEP, 2 mM MgSO₄, pH 7.6), totaling 100 µL. For pH studies, the buffer comprised 20 mM BIS-TRIS propane (with 100 µM MgCl₂, 1 mM ATP, 500 µM Coenzyme A). All non-enzyme assay components were premixed in the wells prior to wild-type or mutant luciferase addition. Images for all assays were acquired as described above and analyzed with Microsoft Excel or GraphPad Prism (version 6.0f for Macintosh, GraphPad Software).

3.9n Spectral imaging with purified enzymes

Bioluminescence imaging was performed essentially as described above. Five minutes post-enzyme addition, sequential images were acquired using an open filter (10 s exposure), a GFP band pass filter (30 s exposure, 515-575 nm), then a Cy5.5 band pass filter (30 s exposure, 695-770 nm). Data were exported to Microsoft Excel and analyzed using GraphPad Prism (version 6.0f for Macintosh, GraphPad Software).

3.9o Bioluminescence kinetic measurements

Measurements were acquired on a Tecan F200 Pro injection port luminometer with a neutral density filter. Reactions were performed in black 96-well flat-bottom plates (Grenier). Solutions of luciferin analog in bioluminescence buffer (20 mM Tris-HCl pH 7.6, 2 mM MgSO₄, 2 mM ATP, 0.1 mM EDTA, 1 mM TCEP, 0.5 mg/mL BSA) were prepared (2 μM - 10 mM analog), and 50 μL were added to each well. The luminescence from each well was measured for 1.5 s prior to the addition of Fluc or mutant in bioluminescence buffer without ATP. For wells containing D-luciferin, a 0.002 mg/mL solution of enzyme was used. For all other compounds, a 0.02 mg/mL solution of enzyme was administered. Following the addition of enzyme, luminescence was 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. The emission maxima were determined by averaging the five maximum photon outputs per run. K_m and relative k_{cat} values were determined using nonlinear regression analyses and robust fit outlier removal in GraphPad Prism (version 6.0f for Macintosh, GraphPad Software).

3.9p Mammalian cell culture

HEK293 (ATCC) and DB7 cells (courtesy of the Contag laboratory, Stanford) were cultured in DMEM (Corning) supplemented with 10% (vol/vol) fetal bovine serum (FBS, Life Technologies), penicillin (100 U/mL), and streptomycin (100 μg/mL). Cells were maintained in a 5% CO₂ water-saturated incubator at 37 °C. Transient transfections of mutant or WT Fluc DNA were performed using cationic lipid formulations (Lipofectamine 2000; Invitrogen). Cells were analyzed for expression or used for imaging analysis 24 –

36 h post transfection. DB7 cells were transduced with ecotropic retrovirus (Phoenix packaging system) followed by selection with puromycin (10 μ M). The cells were sorted via FACS at the Institute for Immunology Flow Cytometry Core (UCI).

3.9q Flow cytometry

Cells transiently expressing luciferase were trypsinized and washed in FACS buffer (PBS with 1% BSA) prior to analysis on a BD Biosciences LSRII. For each sample, 10,000 cells were analyzed and data were analyzed by FloJo software (Tree Star, Inc.).

3.9r Bioluminescence imaging with mammalian cells

HEK293 or DB7 cells transiently expressing Fluc or mutant luciferases were added to black 96-well plates (2×10^5 cells (HEK293) or 3×10^5 (DB7) per well). A stock solution of luciferin (5-10 mM in PBS) was added to each well (500 μ M (HEK) or 1 mM (DB7) final concentration). Sequential imaging in black 96-well plates was performed as above (for HEKs) or with DB7 cells stably expressing mutant luciferases (1.5×10^5 cells per well). A stock solution of luciferin was added to each well (500 μ M (HEK) or 750 μ M (DB7) final concentration) and then a second luciferin solution (500 μ M (HEK) or 750 μ M (DB7) final concentration) was added after the initial imaging session. Bioluminescent-images were then acquired again.

3.9s Sequential bioluminescence imaging with patterned cells

Biocompatible stencils were generated and placed in tissue culture dishes as previously described [68]. HEK293 or DB7 cells expressing mutant B or C were plated (2×10^5 (HEK293) or 3×10^5 (DB7) per stencil chamber). For HEK293 cells, luciferins (500 μ M)

were added sequentially to each stencil chamber immediately after plating. For the DB7 cells, the cells were allowed to adhere for 24 h prior to stencil removal and sequential luciferin (1 mM) addition. Bioluminescent images were acquired as described above except with 10 min acquisitions.

3.9t Molecular docking studies

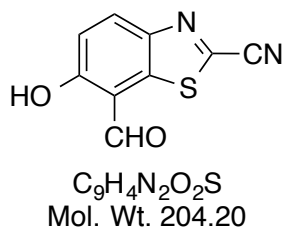
Native Fluc (PDB:4G36) was prepared for docking using the protein prep wizard in Maestro (version 2014-3). The OPLS2005 force field was used for minimization. A Glide-grid was generated using this minimized structure and DSLA was used to provide the coordinates for ligand binding. The luciferin analogs were prepared as the AMP conjugates; their geometries were prepared through the LigPrep module prior to docking. SP docking was used to obtain the input poses for high-level calculations. XP docking was then performed using flexible ligand sampling, sampling of nitrogen inversions and ring conformations. Epik state penalties were used to exclude non-physiologically relevant protonation states and non-planar amide bonds were also penalized. An XP descriptor file was written in order to facilitate post-docking analysis. The docked ligands were evaluated manually via pose-viewer to choose the most relevant poses as well with XP visualizer to quantify the relative contributions of different ligand interactions to the assessed GlideScore.

3.9u General synthetic methods

All reagents purchased from commercial suppliers were of analytical grade and used without further purification. 4,5-Dichloro-1,2,3-dithiazolium chloride, was prepared

according to literature precedent [32]. Reaction progress was monitored by thin-layer chromatography on EMD 60 F254 plates, visualized with UV light, ceric ammonium molybdate (CAM), chloranil, or KMnO_4 stain. Compounds were purified via flash column chromatography using Siliaflash F60 60 Å, 230-400 mesh silica gel (Silicycle), unless otherwise stated. HPLC purifications were performed on a Varian ProStar equipped with a 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), preparative runs were performed using an Agilent Eclipse XD8-C18 PrepHT column (21.2 x 250 mm 7 μm). Anhydrous solvents were dried by passage over neutral alumina with the exception of DMF, which was passed over activated molecular sieves. Reaction vessels were either flame or oven dried prior to use. NMR spectra were acquired with Bruker Advanced spectrometers. All spectra were acquired at 298 K, unless otherwise specified. ^1H -NMR spectra were acquired at either 500 or 400 MHz, and ^{13}C -NMR spectra were acquired at 125 or 100 MHz. Coupling constants (J) are provided in Hz and chemical shifts are reported in ppm relative to either residual non-deuterated NMR solvent, calculated reference, or to a methanol external reference. Low and high-resolution electrospray ionization (ESI) mass spectra were collected at the University of California-Irvine Mass Spectrometry Facility.

3.9v Synthetic Procedures



7-Formyl-6-hydroxy-1,3-benzothiazole-2-carbonitrile (3-4).

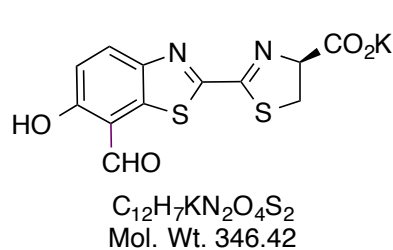
Anhydrous MgCl₂ (1.52 g, 15.9 mmol) and paraformaldehyde (646 mg, 21.3 mmol) were added to a rigorously dried 1 L flask.

The reagents were suspended in anhydrous THF (426 mL), and the vessel was purged with N₂. Triethylamine (1.57 mL, 11.2 mmol) was added, and the reaction stirred at 60 °C until most solids dissolved. Heterocyclic phenol **3-1a** (939 mg, 5.33 mmol) was then added, and the resulting mixture was stirred at 60 °C for an additional 24 h. The reaction was cooled to rt and concentrated *in vacuo*. The residue was dissolved in Et₂O (200 mL), then washed with 1 M Na₂SO₄ (2 x 50 mL), H₂O (2 x 100 mL), and brine (50 mL). The organics were then dried over Na₂SO₄, filtered, and concentrated *in vacuo*. The resulting crude material was purified by flash column chromatography (eluting with 3:1 hexanes:EtOAc) to yield formylated phenol **3-4** as an off-white solid (566 mg, 52%). ¹H NMR (400 MHz, acetone-*d*₆) δ 10.6 (s, 1H), 8.34 (d, *J* = 9.0, 1H), 7.49 (d, *J* = 9.0, 1H); ¹³C NMR (125 MHz, acetone-*d*₆) δ 188.1, 164.0, 148.4, 138.7, 135.2, 133.6, 120.8, 117.7, 114.7; HRMS (ESI⁻) calcd for C₉H₃N₂O₂S [M-H]⁻ = 202.9915, found 202.9921.

General procedure for synthesis of luciferin analogs (3-2a-d, 3-3a-b)

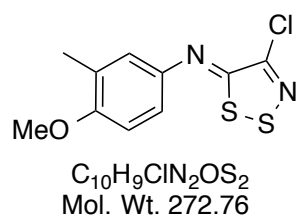
D-Cysteine hydrochloride monohydrate (0.171 mmol) and 2-cyano benzothiazole (0.163 mmol) were suspended in 4:1 MeCN:H₂O (1 mL) in a 20 mL scintillation vial. Potassium carbonate (0.164 mmol) was then added to the mixture, and the resulting bright yellow-

green solutions were stirred under N₂ for 20 min. The potassium salt products were isolated by filtration or purified via reversed-phase HPLC as specified.



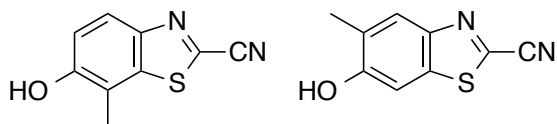
7'-Formyl luciferin (3-5).

Following the general procedure, the potassium salt of **3-5** was obtained as a dark green solid (90%) after HPLC purification (semi-preparative, reversed phase, using the following elution protocol: 100% H₂O for 5 min, followed by a gradient of 0-95% MeCN in H₂O for 15 min. The flow rate was 5 mL/min). ¹H NMR (500 MHz, D₂O) δ 10.0 (s, 1H), 7.63 (d, *J* = 9.2, 1H), 6.71 (d, *J* = 9.2, 1H), 5.17 (app t, *J* = 9.0, 1H), 3.78 (app t, *J* = 10.4, 1H), 3.56 (dd, *J* = 8.5, *J* = 10.7 1H); ¹³C NMR (125 MHz, D₂O/MeOH) δ 189.5, 178.6, 176.1, 166.3, 156.5, 143.3, 135.2, 132.1, 125.4, 117.3, 80.7, 36.8; HRMS (ESI⁻) calcd for C₁₂H₇O₄N₂S₂ [M-K⁺]⁻ = 306.9847, found 306.9843.



(Z)-4-Chloro-N-(4-methoxy-3-methylphenyl)-5H-1,2,3-dithiazol-5-imine (3-7).

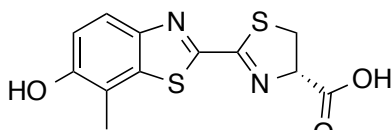
Compound 3-7 was synthesized using a previously published procedure [32], and was isolated as a fluffy, dark orange solid (7.67 g, 57%). ¹H NMR (400 MHz, CDCl₃) δ 7.19-7.16 (m, 2H), 6.90 (d, *J* = 8, 1H), 3.88 (s, 3H), 2.27 (s, 3H); ¹³C NMR (125 MHz, acetone-*d*₆) δ 157.4, 155.8, 149.1, 143.5, 128.4, 124.4, 119.1, 111.4, 56.0, 16.4; HRMS (ESI⁺) calcd for C₁₀H₁₀ClN₂OS₂ [M+H]⁺ = 272.9923, found 272.9913.



$C_9H_6N_2OS$
Mol. Wt. 190.22

6-Hydroxy-7-methylbenzo[d]thiazole-2-carbonitrile (3-8a) and 6-hydroxy-5-methylbenzo[d]thiazole-2-carbonitrile (3-8b).

Following the procedure of McCutcheon *et al.* [32] compounds **3-8a** and **3-8b** were isolated as a mixture of white solids (3:1, **3-8b**:**3-8a**, 56 mg, 30%). Characterized as a mixture of regioisomers, 1H NMR (400 MHz, acetone- d_6) δ 9.39 (br s, 0.6H), 7.96 (s, 0.6H), 7.91 (d, $J = 8.9$, 0.4H), 7.58 (s, 0.6H), 7.32 (d, $J = 8.9$, 0.4H), 2.44 (s, 1.2H), 2.38 (s, 1.8H); ^{13}C NMR (125 MHz, acetone- d_6) δ 158.2, 156.3, 147.4, 146.7, 139.6, 135.8, 133.1, 133.0, 129.2, 126.5, 124.0, 118.8, 116.9, 114.4, 114.2, 106.0, 16.8, 15.3; HRMS (ESI $^-$) calcd for $C_9H_5N_2OS$ [M-H] $^-$ = 189.0123, found 189.0130.

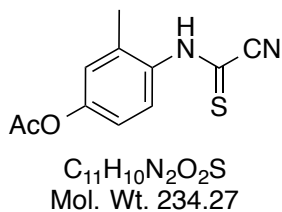


$C_{12}H_{10}N_2O_3S_2$
Mol. Wt. 294.34

7'-Methyl luciferin (3-2a).

Mixture **3-8a** + **3-8b** (0.020 g, 0.11 mmol) was dissolved in 2 mL MeCN and stirred under N_2 in a 20 mL vial. D-Cysteine•HCl•H $_2O$ (0.019 g, 0.11 mmol) and K_2CO_3 (0.014 g, 0.11 mmol) were dissolved in 1.0 mL H $_2O$ and added to the reaction. The reaction was stirred for 30 min, then diluted with H $_2O$ (20 mL), and washed with EtOAc (1 x 20 mL). The aqueous layer was then acidified to pH 3 with 3 M HCl and extracted with EtOAc (2 x 20 mL). The organic layers were combined and washed with H $_2O$ (2 x 20 mL), and brine (1 x 20 mL). The solution was then dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The resulting yellow solid was dissolved in MeOH and purified *via* HPLC (semi-preparative, reversed phase, with the following elution protocol: a gradient of 0-45% MeOH in H $_2O$

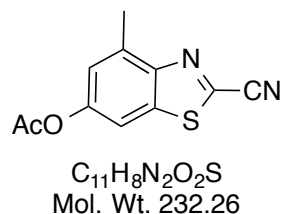
over 5 min, followed by 45% MeOH in H₂O for 10 min, the flow rate was 4 mL/min) to yield **3-2a** (0.003 g, 37% calcd from 3:1 **3-8b**:**3-8a** starting material) as a bright yellow solid. Only compound **3-2a** was isolated for characterization and biochemical ¹H NMR (400 MHz, CD₃OD) δ 7.73 (d, *J* = 8.9, 1H), 7.05 (d, *J* = 8.8, 1H), 5.17 (app t, *J* = 9.4, 1H), 3.73-3.69 (m, 2H), 2.40 (s, 3H); ¹³C NMR (125 MHz, DMSO-*d*₆) δ 171.3, 164.3, 156.3, 154.5, 145.7, 138.1, 122.3, 116.6, 115.7, 78.1, 34.7, 15.1; HRMS (ESI⁻) calcd for C₁₂H₉N₂O₃S₂ [M-H]⁻ = 293.0055, found 293.0059.



***N*-(4-Acetoxy-2-methylphenyl)cyanothioformamide (3-12).**

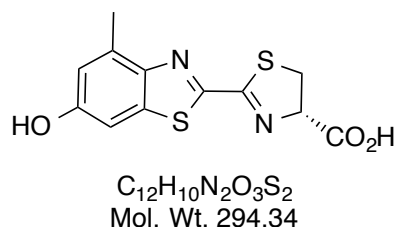
Aniline **3-11** (8.26 g, 50.0 mmol) and Appel's salt (10.8 g, 52.5 mmol) were stirred in 250 mL of anhydrous CH₂Cl₂ under N₂ for ~1 h (until **3-11** was consumed). The reaction volume was then reduced in half under vacuum, and the solids were isolated by vacuum filtration. The isolated solids were resuspended in 2:1 MeCN:THF, and treated with sodium thiosulfate (23.0 g, 145.5 mmol in 80 mL H₂O). The reaction was vigorously stirred for 3 h. The mixture was then filtered to remove elemental sulfur, and the volatiles were removed *in vacuo*. The remaining aqueous mixture was filtered to remove residual solids, and the filtrate was acidified with 1 M NaHSO₄. Cyanothioformamide **3-12** precipitated from solution and was collected by vacuum filtration. The material was washed with additional H₂O and dried to provide **3-12** as a vivid orange solid (10.7 g, 92%). Compound **3-12** was characterized as a mixture of tautomers. ¹H NMR (500 MHz, CDCl₃) δ 7.49-7.37 (m, 1H), 7.06-6.97 (m, 2H), 2.35-2.20 (m, 6H). ¹³C NMR (125 MHz, CDCl₃) δ 170.4, 169.6, 167.5, 164.6, 151.0, 150.2, 135.5, 134.8, 133.5, 132.1, 127.0, 126.9, 124.6, 124.4,

120.7, 120.1, 113.4, 111.9, 21.3, 18.2, 18.0; HRMS (ESI⁻) calcd for C₁₁H₉N₂O₂S [M-H]⁻ = 233.0385, found 233.0388.



6-Acetoxy-4-methyl-1,3-benzothiazole-2-carbonitrile (**3-13**):

Palladium chloride (117 mg, 0.639 mmol), CuI (611 mg, 3.20 mmol), TBAB (4.12 g, 12.8 mmol), and **3-12** (1.50 g, 6.39 mmol) were suspended in anhydrous 1:1 DMF:DMSO (100 mL). The resulting red-brown mixture was placed under N₂ and stirred at 120 °C for 3 h. The reaction was then diluted with ethyl acetate and washed with H₂O (4 x 50 mL). The organics were then dried over Na₂SO₄, filtered, and concentrated *in vacuo*. The crude product was purified by flash column chromatography (eluting with 5:1 hexanes:EtOAc) to provide, (**3-13**) (1.35 g, 91%) as a light beige solid. ¹H NMR (400 MHz, acetone-*d*₆) δ 7.88 (dd, *J* = 2.2, *J* = 0.6, 1H), 7.32 (dd, *J* = 2.2, *J* = 0.9, 1H), 2.75 (s, 3H), 2.32 (s, 3H); ¹³C NMR (125 MHz, acetone-*d*₆) δ 169.6, 152.0, 150.4, 137.1, 137.0, 136.6, 124.1, 113.9, 113.6, 21.0, 18.1. HRMS (ESI⁺) calcd for C₁₁H₈N₂O₂SNa [M+Na]⁺ = 255.0204, found 255.0207.



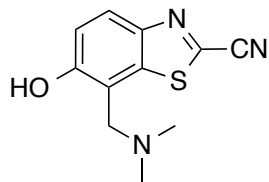
4'-Methyl luciferin (**3-3a**)

Nitrile **3-13** (0.23 g, 1.0 mmol) was dissolved in CH₃CN (4 mL), and a solution of K₂CO₃ (210 mg, 1.5 mmol) in MeOH (81 μL, 2.0 mmol) and H₂O (1 mL) was added. The mixture was stirred at rt until TLC (hexanes:EtOAc, 2:1) showed

consumption of starting material. D-Cysteine•HCl•H₂O was added to the reaction, and the mixture was stirred for an additional 15 min. The reaction was acidified with 1 M NaHSO₄ and the resulting precipitate was collected by vacuum filtration. The precipitate was washed with chilled H₂O and dried under vacuum to give luciferin **3-3a** (270 mg, 92%) as an off-white solid. ¹H NMR (500 MHz, D₂O) δ 6.92 (d, *J* = 2.3, 1H), 6.66 (d, *J* = 1.4, 1H), 5.17 (dd, *J* = 8.4, 9.7 Hz, 1H), 3.80 (dd, *J* = 10, 10.5 Hz, 1H), 3.61 (dd, *J* = 8.5, 11.5 Hz, 1H), 2.40 (s, 3H); ¹³C NMR (125 MHz, DMSO-*d*₆) δ 171.3, 164.5, 157.3, 155.5, 145.9, 137.1, 134.7, 117.3, 104.3, 78.1, 34.7, 17.9; HRMS (ESI⁻) calcd for C₁₂H₉N₂O₃S₂ [M-H]⁻ = 293.0055, found 293.0069.

General procedure for synthesis of C7' modified luciferin intermediates (3-10a-c)

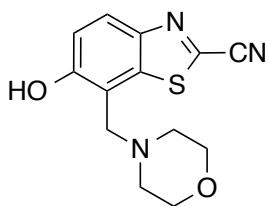
Paraformaldehyde (64 mg, 2.1 mmol) and various amines (2.1 mmol) were suspended in anhydrous MeCN (10 mL) and stirred at 80 °C for 1 h. Heterocyclic phenol **3-1a** (350 mg, 2.0 mmol) was dissolved in anhydrous MeCN (6 mL) and flowed into each reaction. The resulting mixtures were stirred vigorously at 80 °C until TLC showed complete consumption of starting material (~1-2 h). For products that precipitated from solution, purification consisted of filtration followed by recrystallization from warm MeCN. The other reactions were purified by flash column chromatography (eluting with hexanes:EtOAc).



C₁₁H₁₁N₃OS
Mol. Wt. 233.29

7-[(Dimethylamino)methyl]-6-hydroxy-1,3-benzothiazole-2-carbonitrile (3-10a).

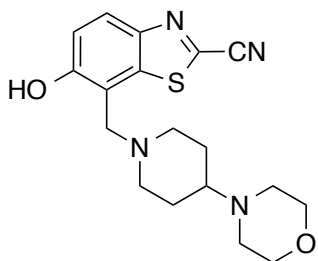
Following the general procedure, benzyl amine **3-10a** was isolated as an off-white solid (365 mg, 91%) upon recrystallization from warm MeCN. ¹H NMR (400 MHz, D₂O) δ 8.02 (d, *J* = 9.1, 1H), 7.32 (d, *J* = 9.1, 1H), 4.53 (s, 2H), 2.94 (s, 6H); ¹³C NMR (125 MHz, D₂O) δ 160.0, 147.9, 142.4, 136.2, 130.0, 120.9, 115.4, 110.9, 59.1, 45.5; HRMS (ESI⁺) calcd for C₁₁H₁₂N₃OS [M+H]⁺ = 234.0701, found 234.0699.



C₁₃H₁₃N₃O₂S
Mol. Wt. 275.33

6-Hydroxy-7-[(morpholin-4-yl)methyl]-1,3-benzothiazole-2-carbonitrile (3-10b).

Following the general procedure, benzyl amine **3-10b** was isolated as a pale tan solid (538 mg, 98%) upon recrystallization from warm MeCN. ¹H NMR (500 MHz, CDCl₃) δ 7.99 (d, *J* = 8.8, 1H), 7.16 (d, *J* = 8.8, 1H), 3.91 (s, 2H), 3.81 (bs, 4H), 2.68 (bs, 4H); ¹³C NMR (125 MHz, CDCl₃) δ 158.7, 146.4, 136.8, 132.1, 125.5, 119.7, 113.5, 111.6, 66.7, 61.2, 53.2; HRMS (ESI⁻) calcd for C₁₃H₁₂N₃O₂S [M-H]⁻ = 274.0650, found 274.0659.

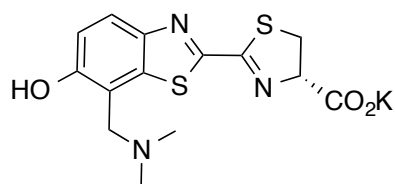


C₁₈H₂₂N₄O₂S
Mol. Wt. 358.46

6-Hydroxy-7-[[4-(morpholin-4-yl)piperidin-1-yl]methyl]-1,3-benzothiazole-2-carbonitrile (3-10c).

Following the general procedure, benzyl amine **3-10c** was isolated as a yellow solid (697 mg, 95%) ¹H NMR (500 MHz, CDCl₃) δ 7.89 (d, *J* = 9.0, 1H), 7.08 (d, *J* = 9.0, 1H), 3.83 (s,

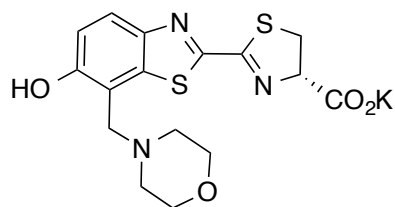
2H), 3.69-3.67 (m, 4H), 3.06 (app d, $J = 11.6$, 2H), 2.51 (bs, 4H), 2.21-2.28 (m, 3H), 1.90 (app d, $J = 12.8$, 2H), 1.61 (q, $J = 11.2$, 2H); ^{13}C NMR (125 MHz, CDCl_3) δ 158.9, 146.0, 136.3, 131.8, 124.9, 119.5, 113.5, 112.3, 67.2, 61.2, 60.5, 52.6, 49.9, 28.1; HRMS (ESI⁺) calcd for $\text{C}_{18}\text{H}_{23}\text{N}_4\text{O}_2\text{S}$ $[\text{M}+\text{H}]^+ = 359.1542$, found 359.1551.



$\text{C}_{14}\text{H}_{14}\text{KN}_3\text{O}_3\text{S}_2$
Mol. Wt. 375.50

7'-[(Dimethylamino)methyl] luciferin (3-2b):

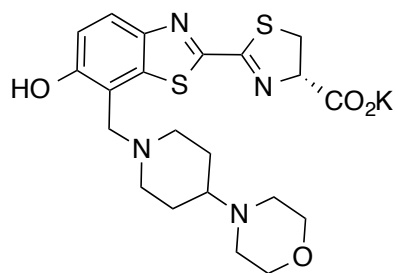
Following the general procedure, the potassium salt of **3-2b** was isolated as a tan solid (289 mg, 86%) following recrystallization from MeCN. ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 7.85 (d, $J = 8.8$, 1H), 7.14 (d, $J = 8.8$, 1H), 5.37 (app t, $J = 9.0$, 1H), 3.75-3.62 (m, 4H), 2.27 (s, 6H); ^{13}C NMR (125 MHz, $\text{DMSO-}d_6$) δ 171.4, 164.4, 158.6, 154.9, 146.6, 135.7, 123.3, 116.8, 116.6, 78.6, 56.1, 44.7, 34.6; HRMS (ESI⁻) calcd for $\text{C}_{13}\text{H}_{14}\text{N}_3\text{OS}_2$ $[\text{M}-\text{CO}_2\text{K}]^- = \text{calcd } 292.0578$, found 292.0588.



$\text{C}_{16}\text{H}_{16}\text{KN}_3\text{O}_4\text{S}_2$
Mol. Wt. 417.54

7'-[(Morpholin-4-yl)methyl] luciferin (3-2c).

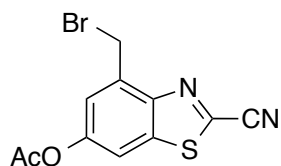
Following the general procedure, the potassium salt of **3-2c** was isolated as a tan solid (325 mg, 93%) following recrystallization from MeCN. ^1H NMR (400 MHz, D_2O) δ 7.54 (d, $J = 8.9$, 1H), 6.94 (d, $J = 8.9$, 1H), 5.17 (app t, $J = 9.2$, 1H), 3.89-3.76 (m, 7H), 3.62 (app t, $J = 9.9$, 1H), 2.81 (s, 4H); ^{13}C NMR (125 MHz, D_2O) δ 180.2, 167.9, 160.1, 158.6, 148.1, 139.7, 126.5, 120.3, 115.1, 83.0, 68.5, 60.1, 55.0, 39.1; HRMS (ESI⁻) calcd for $\text{C}_{15}\text{H}_{16}\text{N}_3\text{O}_2\text{S}_2$ $[\text{M}-\text{CO}_2\text{K}]^- = 334.0684$, found 334.0674.



$C_{21}H_{25}KN_4O_4S_2$
Mol. Wt. 500.67

7'-[[4-(morpholin-4-yl)piperidin-1-yl]methyl] luciferin (3-2d).

Following the general procedure, the potassium salt of **3-2d** was isolated as a tan solid (367 mg, 81%) following HPLC purification (semi-preparative, reversed phase, with the following elution protocol: 100% H₂O for 5 min, followed by 20% MeCN in H₂O for 6 min at 5 mL/min). ¹H NMR (400 MHz, D₂O) δ 7.52 (d, *J* = 8.8, 1H), 6.88 (d, *J* = 8.8, 1H), 5.12 (app t, *J* = 9.3, 1H), 3.94 (dd, *J* = 34.0, *J* = 14.1, 2H), 3.85-3.69 (m, 5H), 3.58 (app t, *J* = 9.9, 1H), 3.30 (app t, *J* = 12.9, 2H), 2.87-2.55 (m, 7H), 2.09 (br s, 2H), 1.64 (app d, *J* = 11.6, 2H); ¹³C NMR (125 MHz, D₂O) δ 180.1, 167.6, 161.9, 159.0, 147.2, 140.6, 127.4, 121.4, 112.9, 83.0, 68.4, 62.6, 59.2, 54.2, 54.1, 51.6, 39.0, 28.3; HRMS (ESI⁻) calcd for C₂₀H₂₅N₄O₂S₂ [M-CO₂K]⁻ = 417.1419, found 417.1412.

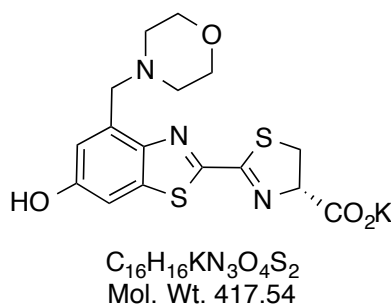


$C_{11}H_7BrN_2O_2S$
Mol. Wt. 311.15

6-Acetoxy-4-[bromomethyl]-1,3-benzothiazole-2-carbonitrile (3-14).

Compound **3-13** (1.1 g, 4.9 mmol) was placed in a rigorously dried vessel and dissolved in freshly degassed dry MeCN. To the stirred solution was added NBS (1.73 g, 9.75 mmol) and benzoyl peroxide (118 mg, 0.487 mmol). The system was purged with N₂ and stirred overnight at 80 °C, in the dark under N₂. The reaction was then diluted with ethyl acetate and washed with H₂O (4 x 50 mL). The organics were then dried over Na₂SO₄, filtered, and concentrated *in vacuo*. The crude product was purified by flash column chromatography (eluting with 10:1 hexanes:EtOAc) to provide **3-14** (1.5 g, 83%) as a light beige solid. ¹H NMR (400 MHz,

acetone- d_6) δ 8.07 (d, $J = 2.2$, 1H), 7.65 (s, $J = 2.2$, 1H), 5.12, (s, 2H), 2.35 (s, 3H); ^{13}C NMR (125 MHz, acetone- d_6) δ 170.0, 152.4, 149.5, 138.6, 138.3, 136.7, 125.2, 117.1, 114.3, 29.1, 21.5. HRMS (ESI $^+$) calcd for $\text{C}_{11}\text{H}_7\text{BrN}_2\text{O}_2\text{SNa}$ $[\text{M}+\text{Na}]^+$ = 332.9309 found 332.9318.



4'-[(Morpholin-4-yl)methyl] luciferin (**3-3b**).

Cyanobenzothiazole **3-14** (564 mg, 1.81 mmol) was dissolved in MeCN (11 mL) and stirred at rt. Morpholine (330 μL , 3.64 mmol) and K_2CO_3 (505 mg, 3.64 mmol) were added to the stirred solution. Upon disappearance of **3-14** (~2 h) a solution of D-cysteine (354 mg, 2.00 mmol) in H_2O (2 mL) was added and the reaction stirred at rt under N_2 for 1 h. The resulting precipitate was isolated by vacuum filtration and washed with chilled MeOH. The residual washes and mother liquor were combined and concentrated *in vacuo*. The isolated residue was then purified by reversed-phase HPLC (preparative, reversed phase, with the following elution protocol: a gradient of 5% MeCN in H_2O over 6 min, followed by 40 % MeCN in H_2O over 6 min, at 20 mL/min). The potassium salt of luciferin **3-3b** (666 mg, 88%) was isolated as a greenish solid and stored at -80 $^\circ\text{C}$. ^1H NMR (500 MHz, D_2O) δ 7.09 (s, 1H), 6.86 (s, 1H), 5.17 (app t, $J = 9.1$, 1H), 3.88 (dd, $J = 37.6$, $J = 13.0$, 2H), 3.78-3.71 (m, 5H), 3.58 (app t, $J = 9.6$, 1H), 2.63 (bs, 4H); ^{13}C NMR (125 MHz, D_2O) δ 180.5, 168.7, 161.0, 158.8, 148.1, 140.5, 133.8, 122.2, 109.5, 82.8, 68.6, 60.3, 55.2, 38.9; HRMS (ESI $^-$) calcd for $\text{C}_{16}\text{H}_{16}\text{N}_3\text{O}_4\text{S}_2$ $[\text{M}-\text{K}]^-$ = 378.0582, found 378.0578.

References

1. Adams, S. T.; Miller, S. C., Beyond D-luciferin: expanding the scope of bioluminescence imaging in vivo. *Curr. Opin. Chem. Biol.* **2014**, *21*, 112.
2. Paley, M. A.; Prescher, J. A., Bioluminescence: a versatile technique for imaging cellular and molecular features. *Medchemcomm* **2014**, *5*, 255.
3. Branchini, B. R.; Hayward, M. M.; Bamford, S.; Brennan, P. M.; Lajiness, E. J., Naphthyl- and Quinolyl-luciferin: Green and Red Light Emitting Firefly Luciferin Analogs. *Photochem. Photobiol.* **1989**, *49*, 689.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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.
9. Haddock, S. H. D.; Moline, M. A.; Case, J. F., Bioluminescence in the Sea. *Ann. Rev. Mar. Sci.* **2010**, *2*, 443.
10. Oba, Y.; Schultz, D. T., Eco-Evo Bioluminescence on land and in the sea. *Adv. Biochem. Eng. Biotechnol.* **2014**, *144*, 3.
11. Viviani, V. R., The origin, diversity, and structure function relationships of insect luciferases. *Cell. Mol. Life Sci.* **2002**, *59*, 1833.
12. Porterfield, W. B.; Prescher, J. A., Tools for visualizing cell-cell 'interactomes'. *Curr. Opin. Chem. Biol.* **2015**, *24*, 121.
13. Badr, C. E.; Tannous, B. A., Bioluminescence imaging: progress and applications. *Trends Biotechnol.* **2011**, *29*, 624.

14. Massoud, T. F.; Paulmurugan, R.; De, A. J.; Ray, P.; Gambhir, S. S., Reporter gene imaging of protein-protein interactions in living subjects. *Curr. Opin. Biotechnol.* **2007**, *18*, 31.
15. 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.
16. 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. Nucleic Acids* **2013**, *2*, e99.
17. Pichler, A.; Prior, J. L.; Piwnica-Worms, D., Imaging reversal of multidrug resistance in living mice with bioluminescence: MDR1 P-glycoprotein transports coelenterazine. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1702.
18. 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.* **2015**, *54*, 7065.
19. Evans, M. S.; Chaurette, J. P.; Adams, 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. *Nature Methods* **2014**, *11*, 393.
20. 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.
21. Harwood, K. R.; Mofford, D. M.; Reddy, G. R.; Miller, S. C., Identification of mutant firefly luciferases that efficiently utilize aminoluciferins. *Chem. Biol.* **2011**, *18*, 1649.
22. Adams, S. T.; Mofford, D. M.; Reddy, G. S. K. K.; Miller, S. C., Firefly Luciferase mutants allow substrate-selective bioluminescence imaging in the mouse brain. *Angew. Chem. Int. Ed.* **2016**, *55*, 4943.
23. Rowe, L.; Dikici, E.; Daunert, S., Engineering bioluminescent proteins: expanding their analytical potential. *Anal. Chem.* **2009**, *81*, 8662.
24. 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. *ACS Chem. Biol.* **2012**, *7*, 1848.
25. Berger, F.; Paulmurugan, R.; Bhaumik, S.; Gambhir, S. S., Uptake kinetics and biodistribution of C-14-D-luciferin-a radiolabeled substrate for the firefly luciferase

- catalyzed bioluminescence reaction: impact on bioluminescence based reporter gene imaging. *Eur. J. Nucl. Med. Mol. Imaging* **2008**, *35*, 2275.
26. 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.
 27. 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*, 041210.
 28. Branchini, B. R.; Magyar, R. A.; Murtiashaw, M. H.; Portier, N. C., The role of active site residue arginine 218 in firefly luciferase bioluminescence. *Biochemistry* **2001**, *40*, 2410.
 29. 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.
 30. 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.
 31. 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.
 32. McCutcheon, D. C.; Porterfield, W. B.; Prescher, J. A., Rapid and scalable assembly of firefly luciferase substrates. *Org. Biomol. Chem.* **2015**, *13*, 2117.
 33. Meroni, G.; Rajabi, M.; Santaniello, E., D-Luciferin, derivatives and analogues: synthesis and in vitro/in vivo luciferase-catalyzed bioluminescent activity. *Arkivoc* **2009**, 265.
 34. White, E. H.; Worther, H.; Field, G. F.; Mcelroy, W. D., Analogs of firefly luciferin. *J. Org. Chem.* **1965**, *30*, 2344.
 35. White, E. H.; Worther, H.; Seliger, H. H.; Mcelroy, W. D., Amino Analogs of Firefly luciferin and biological activity thereof. *J. Am. Chem. Soc.* **1966**, *88*, 2015.
 36. 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.
 37. 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.

38. 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.
39. White, E. H.; Worther, H., Analogs of firefly luciferin. III. *J. Org. Chem.* **1966**, *31*, 1484.
40. 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.
41. Seliger, H. H.; Mcelroy, W. D.; Field, G. F.; White, E. H., Stereospecificity and firefly bioluminescence, a comparison of natural and synthetic luciferins. *Proc. Natl. Acad. Sci. U.S.A.* **1961**, *47*, 1129.
42. Woodroffe, C. C.; Shultz, J. W.; Wood, M. G.; Osterman, J.; Cali, J. J.; Daily, W. J.; Meisenheimer, P. L.; Klaubert, D. H., N-alkylated 6'-aminoluciferins are bioluminescent substrates for Ultra-Glo and QuantiLum luciferase: New potential scaffolds for bioluminescent assays. *Biochemistry* **2008**, *47*, 10383.
43. 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.
44. Takakura, H.; Kojima, R.; Urano, Y.; Terai, T.; Hanaoka, K.; Nagano, T., Aminoluciferins as functional bioluminogenic substrates of firefly luciferase. *Chem. Asian J.* **2011**, *6*, 1800.
45. Shinde, R.; Perkins, J.; Contag, C. H., Luciferin derivatives for enhanced in vitro and in vivo bioluminescence assays. *Biochemistry* **2006**, *45*, 11103.
46. 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.
47. Jones, P. R.; Gelinas, R. M., The first spectral confirmation for the structures of anhydro dimers of *o*-Hydroxybenzaldehydes. *J. Org. Chem.* **1981**, *46*, 194.
48. Phillips, J. P.; Barrall, E. M., Betti reactions of some phenols. *J. Org. Chem.* **1956**, *21*, 692.
49. Hirano, T.; Nagai, H.; Matsushashi, T.; Hasumi, Y.; Iwano, S.; Ito, K.; Maki, S.; Niwa, H.; Viviani, V. R., Spectroscopic studies of the color modulation mechanism of firefly (beetle) bioluminescence with amino-analogs of luciferin and oxyluciferin. *Photochem. Photobio. Sci.* **2012**, *11*, 1281.

50. 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.
51. da Silva, L. P.; da Silva, J. C. G. E., Firefly Chemiluminescence and bioluminescence: efficient generation of excited states. *Chemphyschem* **2012**, *13*, 2257.
52. Hopkins, T. A.; Seliger, H. H.; White, E. H.; Cass, M. W., Chemiluminescence of firefly luciferin. a model for bioluminescent reaction and identification of product excited state. *J. Am. Chem. Soc.* **1967**, *89*, 7148.
53. Steinhardt, R. C.; Rathbun, C. M.; Krull, B. T.; Yu, J.; 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.
54. Kato, D.; Shirakawa, D.; Polz, R.; Maenaka, M.; Takeo, M.; Negoro, S.; Niwa, K., A firefly inspired one-pot chemiluminescence system using n-propylphosphonic anhydride (T3P). *Photochem. Photobio. Sci.* **2014**, *13*, 1640.
55. Mao, Y., Dynamics studies of luciferase using elastic network model: how the sequence distribution of luciferase determines its color. *Protein Eng. Des. Sel.* **2011**, *24*, 341.
56. Chen, M. M. Y.; Snow, C. D.; Vizcarra, C. L.; Mayo, S. L.; Arnold, F. H., Comparison of random mutagenesis and semi-rational designed libraries for improved cytochrome P450 BM3-catalyzed hydroxylation of small alkanes. *Protein Eng. Des. Sel.* **2012**, *25*, 171.
57. Reetz, M. T.; Prasad, S.; Carballeira, J. D.; Gumulya, Y.; Bocola, M., Iterative saturation mutagenesis accelerates laboratory evolution of enzyme stereoselectivity: rigorous comparison with traditional methods. *J. Am. Chem. Soc.* **2010**, *132*, 9144.
58. Reetz, M. T.; Kahakeaw, D.; Lohmer, R., Addressing the numbers problem in directed evolution. *ChemBioChem* **2008**, *9*, 1797.
59. Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z. G.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P., Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. *ACS Synth. Biol.* **2013**, *2*, 83.
60. Amaral, D. T.; Arnoldi, F. G. C.; Viviani, V., Molecular phylogeny of the neotropical bioluminescent beetles. *Luminescence* **2012**, *27*, 96.

61. Viviani, V. R.; Amaral, D. T.; Neves, D. R.; Simões, A.; Arnoldi, F. G. C., The luciferin binding site residues C/T311 (S314) influence the bioluminescence color of beetle luciferases through main-chain interaction with oxyluciferin phenolate. *Biochemistry* **2013**, *52*, 19.
62. Pines, G.; Pines, A.; Garst, A. D.; Zeitoun, R. I.; Lynch, S. A.; Gill, R. T., Codon compression algorithms for saturation mutagenesis. *ACS Synth. Biol.* **2014**, *4*, 604.
63. Ness, J. E.; Kim, S.; Gottman, A.; Pak, R.; Krebber, A.; Borchert, T. V.; Govindarajan, S.; Mundorff, E. C.; Minshull, J., Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently. *Nat. Biotechnol.* **2002**, *20*, 1251.
64. Quan, J.; Tian, J., Circular polymerase extension cloning. *Nat. Protoc.* **2011**, *6*, 242.
65. Rouillard, J. M.; Lee, W.; Truan, G.; Gao, X. L.; Zhou, X. C.; Gulari, E., Gene2Oligo: oligonucleotide design for in vitro gene synthesis. *Nucleic Acids Res.* **2004**, *32*, W176.
66. Bessette, P. H.; Mena, M. A.; Nguyen, A. W.; Daugherty, P. S., Construction of designed protein libraries using gene assembly mutagenesis. In *Directed evolution library creation*, Arnold, F. H.; Georgiou, G., Eds. Humana: 2003; Vol. 231, pp 29.
67. Liu, H. P.; Patel, M. R.; Prescher, J. A.; Patsialou, A.; Qian, D. L.; Lin, J. H.; 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.
68. Jones, K. A.; Li, D. J.; Hui, E.; Sellmyer, M. A.; Prescher, J. A., Visualizing cell proximity with genetically encoded bioluminescent reporters. *ACS Chem. Biol.* **2015**, *10*, 933.

Chapter 4: A general method for identifying selective luciferase-luciferin pairs

4.1 Introduction

Based on the work reported in Chapter 3, we continued our engineering of orthogonal luciferin-luciferase pairs. To realize multi-component imaging in vivo, we needed to access luciferase-luciferin pairs with improved selectivity, and increased total photon output. However, only limited attempts to develop completely orthogonal (substrate-selective) luciferases have been attempted to date [1-2]. Simultaneous engineering of luciferases and luciferins is a straightforward method to achieve selective (i.e., orthogonal) bioluminescence probes, although somewhat more difficult than Fluorescent Protein (FP) engineering based on the requirement for modified proteins and substrates simultaneously. Accessing enzymes with alternative substrate usage is well preceded in directed evolution [3-5]. Traditional applications of this technique, though, have focused on optimizing one enzyme at a time. For example, transcription factors and catalytic domains have been evolved to recognize new substrates [6-7]. Selectivity for one molecule over another is often realized as a result, but is not typically the parameter being optimized. In other efforts, engineering tRNA synthetases for non-canonical amino acid use, negative screens are included to ensure specificity over native amino acid substrates [8]. Orthogonal synthetases have been identified from such approaches, but only when beginning with completely different enzyme starting points [9-10]. Efforts to engineer luciferases for altered substrate usage have resulted in incomplete mutual orthogonality [11], or pairs with low photon output [1]. A more

systematic and broad approach is necessary to realize true bioluminescence based multicomponent imaging.

Here, we present a general and rapid method to engineer orthogonal luciferase-luciferin pairs. This approach relies on developing an initial pool of functional luciferases, and evaluating collections of chemically diverse luciferins across this library (Fig. 4-1A). From these expanded sets of enzymes and small molecules, we developed an algorithm to identify useful orthogonal probes from this large data set. Collectively, we screened and generated a candidate list of greater than 800,000 possible pairings between 157 new mutants and 12 compounds. We also evaluated the selectivity of

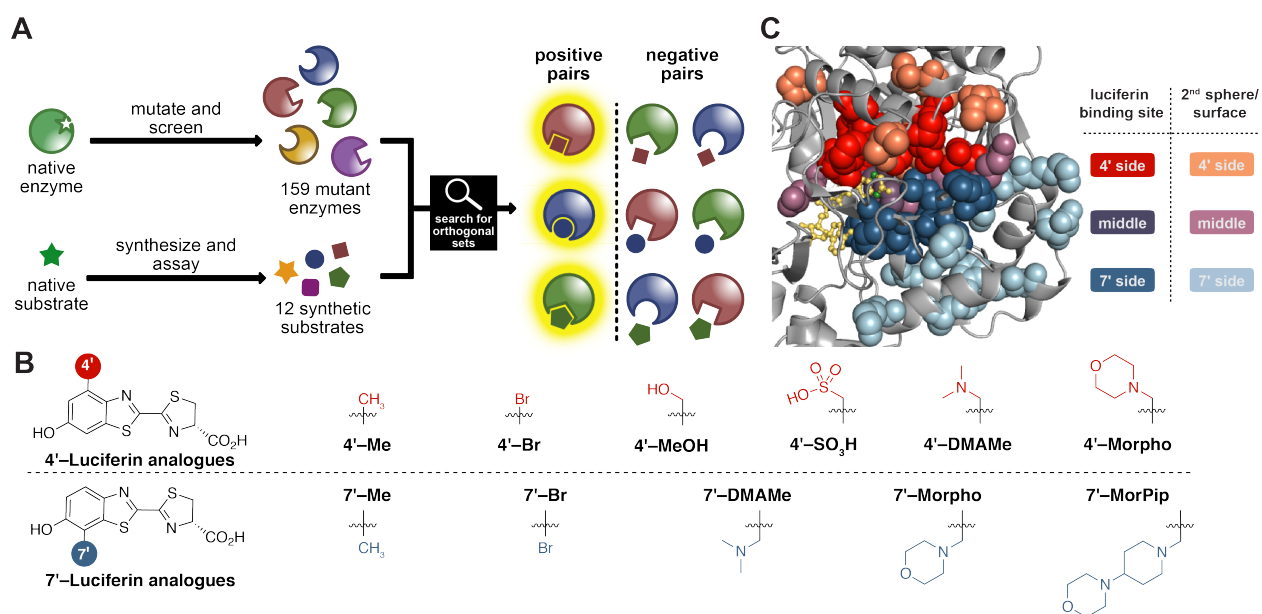


Figure 4-1. General method for accessing novel luciferase-luciferin pairs. **(A)** *Some name for our method* enables rapid search of enzyme and substrate space to find mutually orthogonal enzyme-substrate pairs. *Positive pairs* are defined as enzyme-substrate pairs that effectively turn over to form product. *Negative pairs* are all other possible pairings where product formation is reduced. **(b)** Diverse modifications at the 7' and 4' positions of luciferin can be generated from common cyanobenzothiazole intermediates. **(c)** Residues proximal to the active site were targeted for saturation mutagenesis. Spheres represent targeted residues. Residues of the same color were grouped in single libraries.

many of these pairs in vitro. The selectivities of a subset were further confirmed in cultured cell and animal models, greatly expanding the number of viable imaging probes. We also analyzed the principles governing orthogonal substrate usage and how to identify higher-order sets of selective probes. The methods and algorithms presented here will accelerate the identification of orthogonal probes for bioluminescence, in addition to other areas where selective substrate usage is critical. Importantly, like with other imaging modalities, these new tools will likely spur new discoveries.

4.2 Expanding the pool of candidate luciferins and luciferases

As a starting point for luciferin modification, we focused on derivatives with steric appendages at the 4' and 7' carbons. These positions lie in close proximity to the Fluc backbone [12], and preliminary work suggested that steric modifications do not adversely impact photon emission [1], making them desirable starting points. We also previously identified a pair of luciferases that could discriminate between luciferins with modifications at these positions [1]. A luciferase containing a single mutant used a bulky 7'-modified luciferin selectively, compared to a luciferase with 4 mutations that had preference for a bulky 4'-modified luciferin by screening small library of random mutants. Interestingly, attempts to optimize these probes using traditional directed evolution (one-enzyme and one substrate) did not result in improved selectivity (Fig. 4-2) or provide a starting point for developing new orthogonal probes.

Realizing that screening for *selectivity* could provide a more rapid route to new pairs, we focused on diversifying the number of luciferin and luciferase inputs. Large collections of both new and known [1,13] luciferin analogs were assembled (Fig. 4-1B).

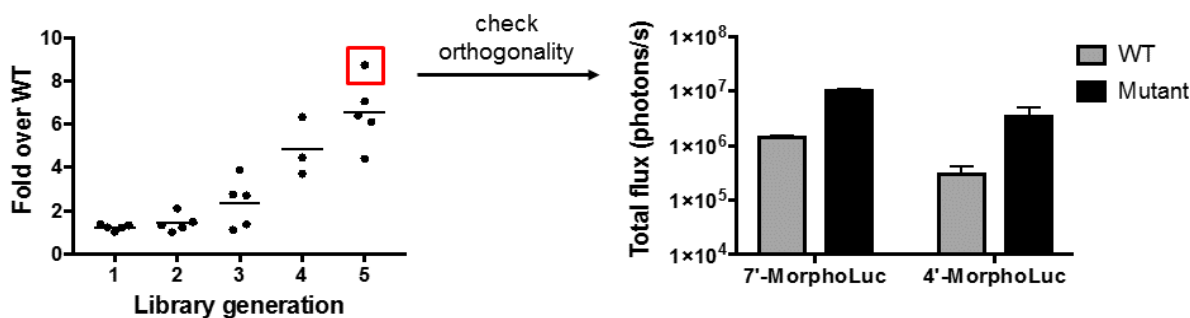


Figure 4-2. “Brighter” mutants can be evolved but may not be the best leads for orthogonality. (A) Fold increase in light emission over five generations of evolving for brightness with 7'-MorphoLuc. (B) Best performing mutant (red box, A) shown in black bars, compared to WT (gray bars) demonstrates an increase in light emission with both compounds, but not in selectivity for 7'-MorphoLuc. Imaging was performed in bacterial lysate (250 μ M luciferin analogue).

All analogs were then benchmarked for light emission with Fluc (Fig. 4-3). Although they varied in terms of their photon outputs, all luciferins were functional light emitters with Fluc. Some level of enzymatic activity is necessary for successful enzyme evolution. In fact, weak enzymatic activities are often good starting points for evolving enzymes [3].

In parallel with luciferin diversification, we targeted a broad set of Fluc sequence space for mutagenesis (Fig. 4-1C). Our goal was to amass a pool of functional luciferases to screen with the library of luciferin analogs in order to efficiently cover a broad swath of chemical and sequence space. Initially, a total of 23 residues in and around the active site were targeted for site-directed mutagenesis in 8 different libraries. The majority of the resulting mutants were likely to be non-functional, and thus not ideal starting points. Luciferase light emission is also weak, and does not interface with rapid identification and segregation, meaning that each mutant-analog pair must be physically interrogated. To quickly traverse the mutants and eliminate non-functional luciferases prior to screening with all analogs, we adapted a high-throughput screening

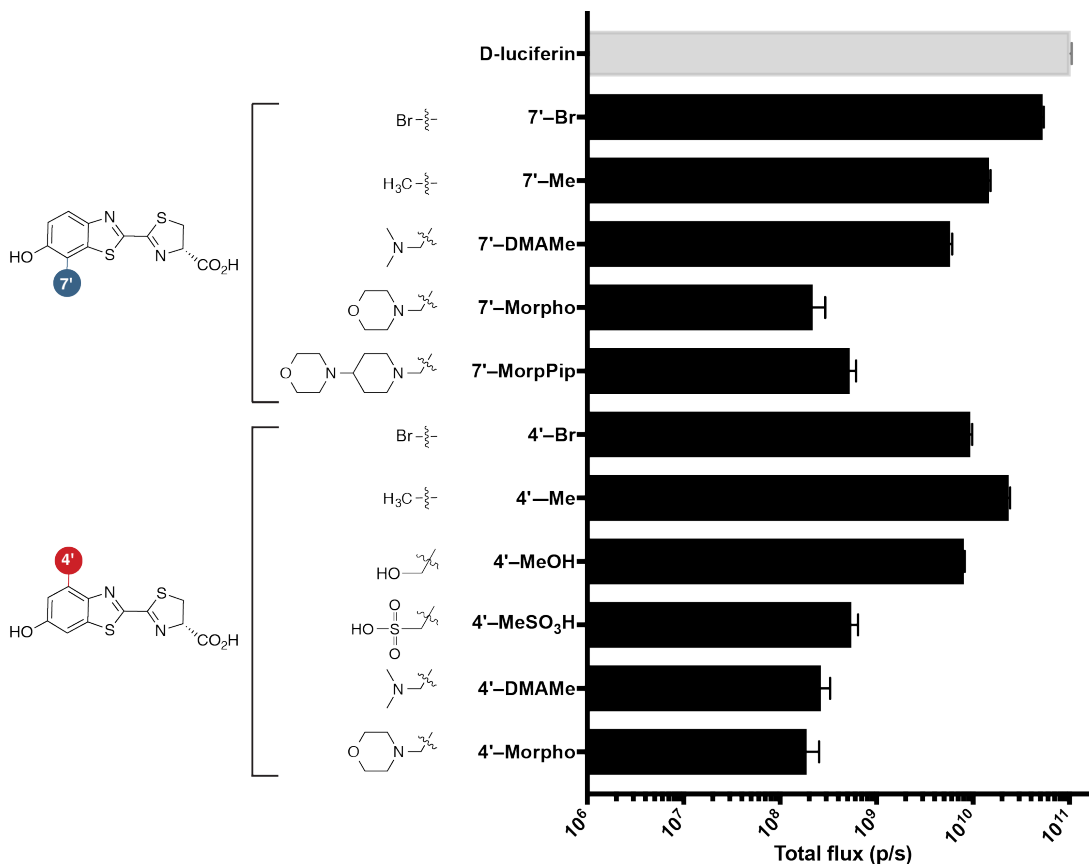


Figure 4-3. All luciferin compounds were imaged at 100 μ M with purified Fluc (1 μ g/well). Error bars represent SEM for $n \geq 3$ independent experiments.

procedure to identify enzymes that still emitted photons [1]. The 4'/7'-BromoLuc, 4'/7'-MeLuc, and 7'-MorphoLuc analogues (Figs. 4-1B, 4-4) were used to eliminate non-functional luciferase mutants. These analogs are the “brightest” and easiest to access in bulk, as large amounts of compound are required for this task. The luciferins were added to agar, and the luciferase libraries were quickly screened on plate to remove the non-functional luciferases (~95-99% of library members). Light emitting colonies were picked and further assessed in lysate assays and by sequencing. A variety of mutants (both common and unique to each structure) were identified (Fig. 4-5). The “hits” were pooled to generate additional libraries (1-3 generations) and further diversify the functional luciferase pool (Fig. 4-4A).

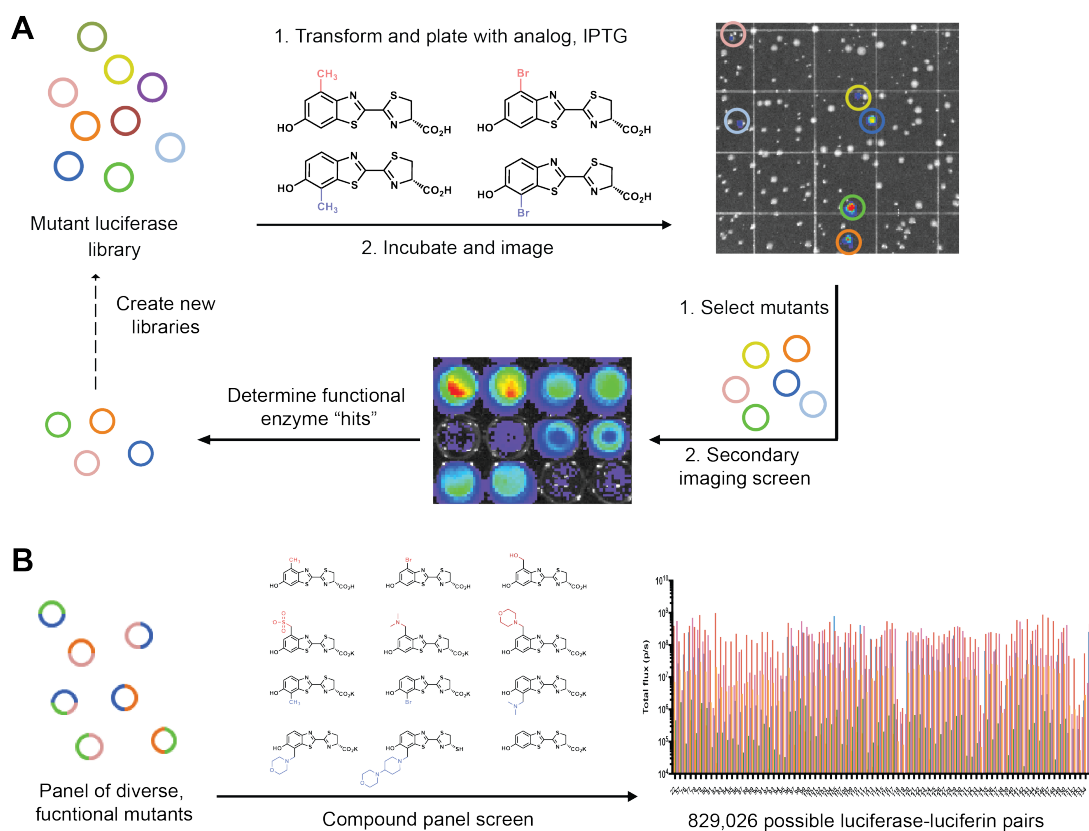


Figure 4-4. Screening for diverse and functional luciferases as starting points for substrate selectivity. **(A)** Mutant luciferase libraries were generated and screened on-plate with minimally perturbed luciferin analogs. Functional mutants were identified and analyzed by a secondary screen to determine “hits” for next-generation libraries. These libraries can undergo further mutagenesis and iterative screening. **(B)** Alternatively, functional enzymes can be imaged with a panel of luciferin analogs to report on substrate selectivity.

4.3 Screening for orthogonal luciferase-luciferin pairs *in silico*

With expanded collections of functional enzymes and substrates in hand, we screened luciferases and analogs for selectivity en route to identifying orthogonal pairs. Testing each combination of two substrates and two mutants would require 829,026 separate experiments—an impractical number (Fig. 4-4B). So, instead we tested each enriched mutant with each luciferin individually and relied on computation to find lead pairs that were mutually selective. With 159 enzymes (See methods section 4.8e for list

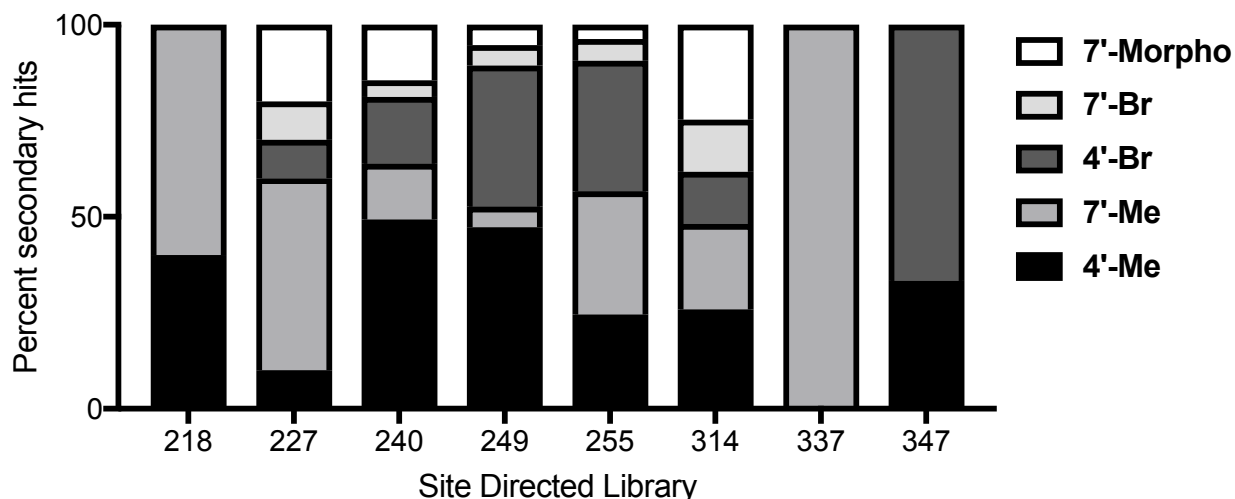


Figure 4-5. Percent unique “hits” per compound and library. For example, 60% of the “hits” from screening the 218 library arose from 7’-MeLuc.

of mutations) and twelve analogs, this resulted in $159 \times 12 = 1908$ individual data points. To mine this data set, we initially turned to an algorithm that quantified orthogonality. This method was not directly applicable, though, to the expanded screening sets. To sift through a greater range of mutant and compound space and provide more streamlined approach to identifying multiple pairs, we modified the algorithm for increased efficiency (Fig. 4-6A). We reasoned that perfect selectivity could be represented by an identity matrix. “Orthogonality” would be maximal if each enzyme was completely selective for its cognate substrate (“1”) and non-functional with other luciferins (“0”). An orthogonality score was then determined by representing each pairing as a square matrix with enzymes in rows and compounds in columns. Each column was normalized and the matrix was compared to the identity matrix via root mean square distance (RMSD). Pairings were sorted by increasing RMSD, with the smallest RMSD value representing the most orthogonal pair. This algorithm had the added benefit of providing faster processing via utilization of multiple computer processors simultaneously. We observed

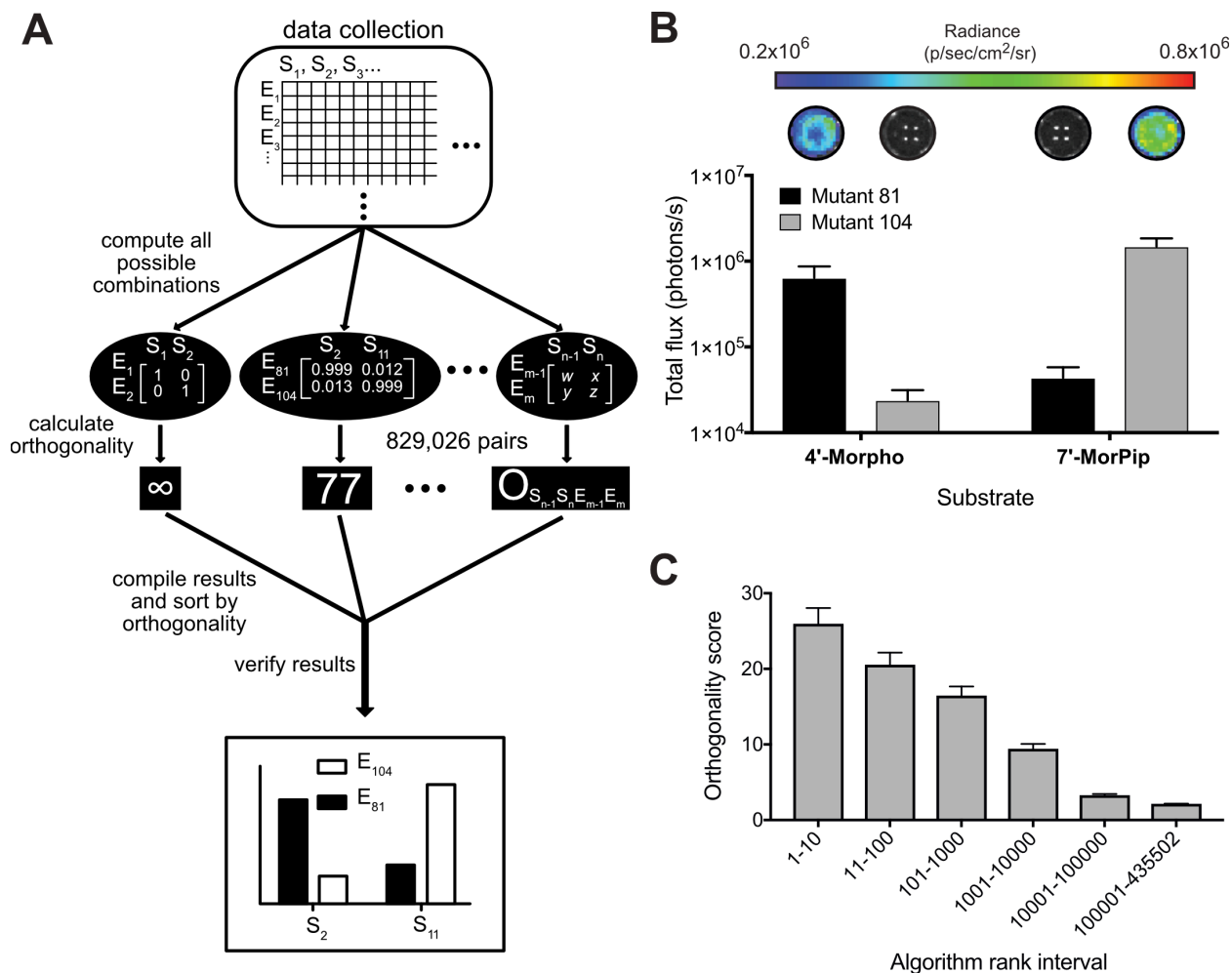


Figure 4-6. Unique luciferase-luciferin pairings maintain substrate selectivity. (A) Example of emission profiles from a top luciferase-luciferin pairing. Representative bioluminescent images are provided. Bacteria expressing mutant enzymes were expanded, lysed, and split for controlled protein levels. Lysate were imaged with substrate resolved analogs. (B) Orthogonality scores correlate with ranking from algorithm analysis. (C) Pairwise enzyme-substrate data was screened *in silico* for adequate orthogonal luciferin-luciferase pairs. For B-C, error bars represent the standard error of the mean for $n \geq 4$ experiments.

a large speed improvements when we utilized up to 64 cores on a supercomputing cluster.

Running the algorithm with the set of mutants and compounds provided a ranked list of 829,026 members. The highest-ranked pairings comprised a diverse set of

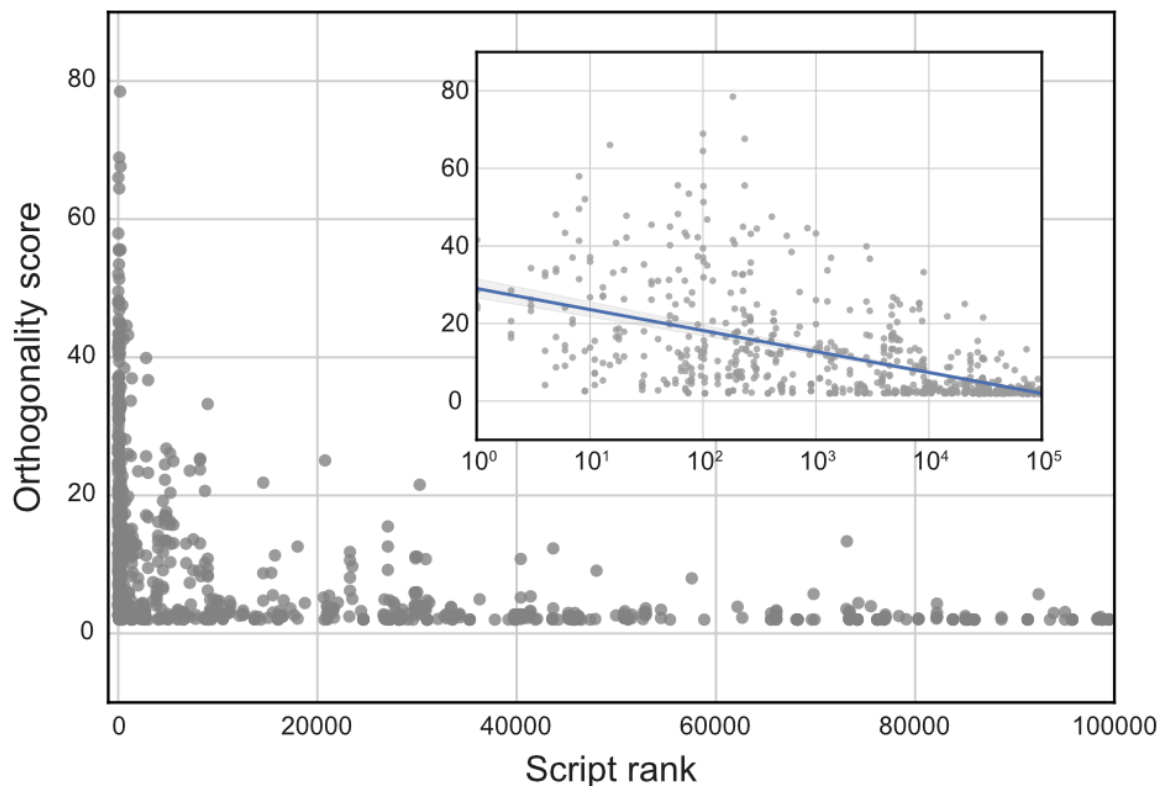


Figure 4-7. A variety of pairings were reproduced across a range of algorithm rankings. Plotted at each point is the calculated orthogonality score for the reproduced pair versus the original rank of the pair as determined by the algorithm. **Inset:** The same data plotted on a logarithmic scale shows a clear positive relationship: generally, as ranking improves, reproduced orthogonality also increases.

enzymes and substrates, greatly expanding the number of known orthogonal luciferase-luciferin pairs. Importantly, the orthogonality of the predicted pairs could be reliably reproduced in cell lysates. The top ranked orthogonal set is plotted in figure 4-6B as an example of the selectivity and photon output. We also analyzed the top ten unique pairings, along with a handful of others in the data set (every tenth rank of the top 100, every 100th rank of the top 1000, and every 1000th rank down to position 5000). Collectively, the reproduced points in the top 1000 hits reliably provided at least 10-fold greater photon outputs with their cognate luciferins in head-to-head comparisons. Diminishments in selectivity were observed farther down the ranked list (past 1000)

(Fig. 4-6C). These results are important, as the data analyzed by the script can be collected at any point. The script assumes that all are comparable/normalized relative to each other. This is a bit of a big assumption, but we have verified that higher rankings in the script correlate with better *reproduced* (on the same day/at the same time) orthogonalities. The rank given by the *in silico* screen is a good predictor of orthogonal substrate usage (Fig. 4-6C, 4-7) and has the ability to cull 99.9% (~828,000 of the total 829,026) of the irrelevant enzyme-substrate sets.

4.4 Analyzing trends in orthogonal substrate usage

With several orthogonal luciferases and luciferins identified, we aimed to explore the origins of selectivity. Interestingly, selectivity was achieved not by markedly enhanced turnover of a preferred substrate. Rather, selectivity was achieved from reduced photon production with most other compounds. As shown in Figure 4-8A, matched enzymes and substrates (or “positive” pairs) are on par with native Fluc in terms of photon outputs. Unmatched enzymes and substrates, though, (“negative” pairs), demonstrated activity reductions between 10-1000 fold compared to Fluc. This overall trend demonstrates that, in a given orthogonal pair, selectivity is more readily obtained by reducing light emission with a negatively paired compound versus selectively increasing light emission with the positively paired compound.

Since compound selectivity appears to be achieved by “destroying” enzyme-substrate interactions, structurally related compounds would be expected to exhibit similar trends in orthogonality (Fig. 4-9). For example, enzymes paired with bulky

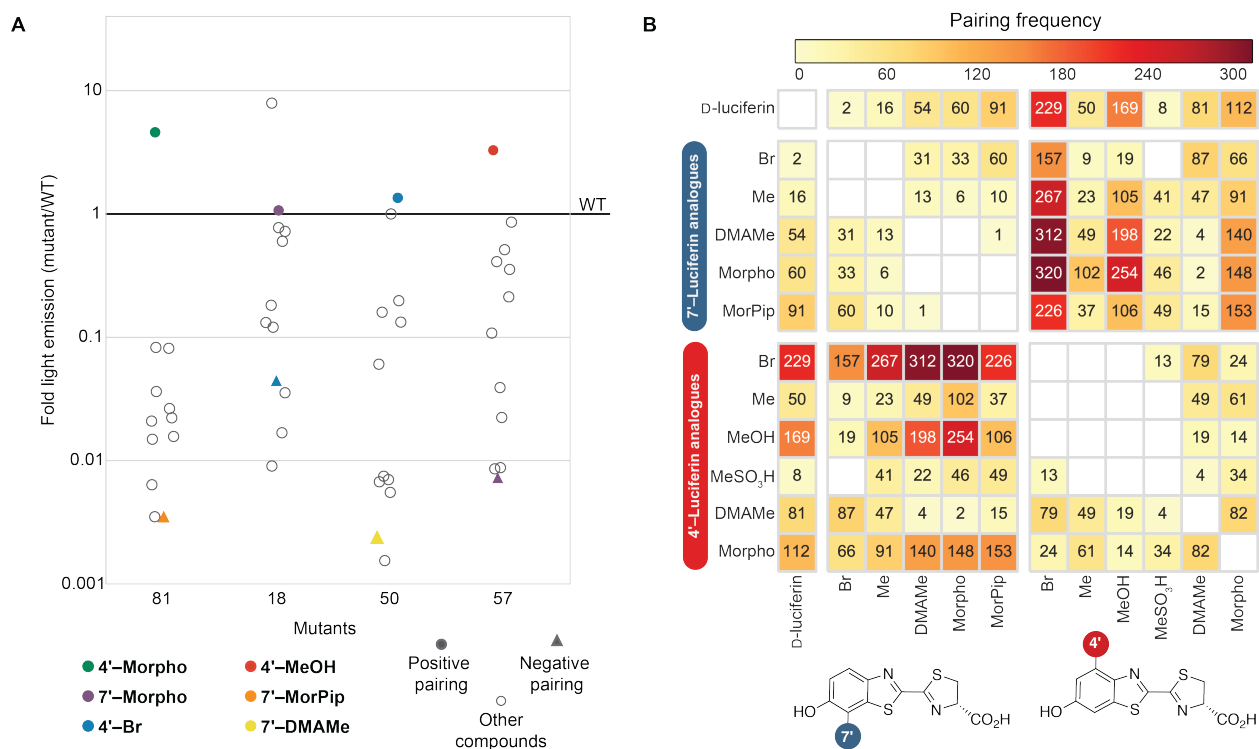


Figure 4-8. (A) Fold light emission over wild-type with each substrate in the four most frequently appearing enzymes in the top 5,000 pairings. Points represent luciferin analogs. **(B)** Compound-compound pairing frequency in the top 5,000 *in silico* hits.

7'-compounds tend to use the other compounds in this class in a similar manner (Fig. 4-9A). Mutant 104 and others in this class likely harbor space in the active site to accommodate steric bulk near the 7'-carbon as all modifications at this position are tolerated well (Fig. 4-9A). More structurally divergent compounds are also more frequently paired with one another (e.g. 4'-BrLuc is paired with 7'-MeLuc 267 times in the top 5,000 pairings, Fig. 4-8B). Conversely, matched analogs (i.e., 4'-4' and 7'-7' pairings) were relatively infrequent. For example, substrates with a modification at the 4' position were rarely orthogonal to other 4' compounds likely due to similar substrate-enzyme interactions in the active site. In contrast, when 4' and 7'-modified substrates are paired it is likely that each substrate interacts differently with the enzyme, making it easier to achieve orthogonality.

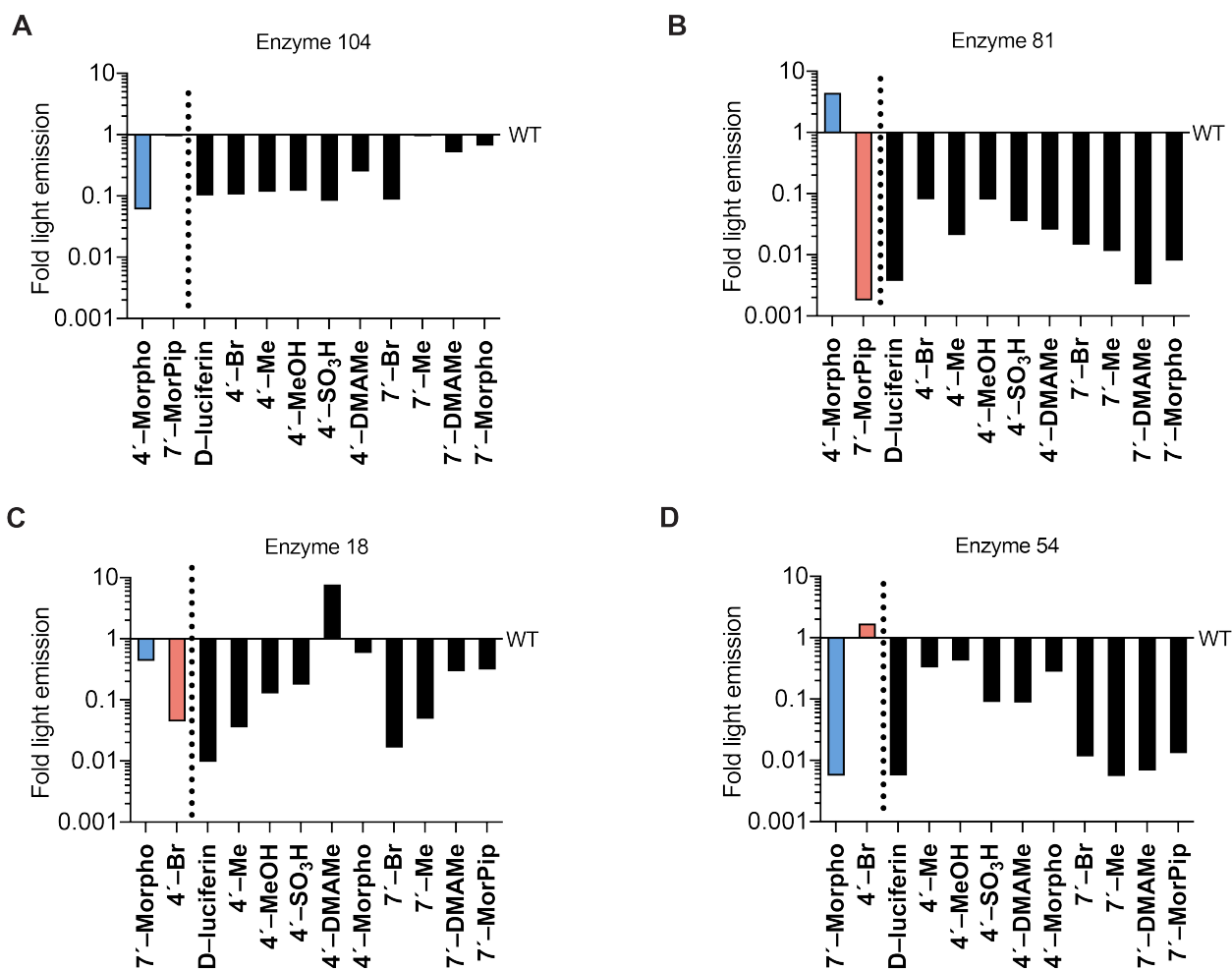


Figure 4-9. Selectivity of mutant enzymes for panel of synthetic luciferins. A-D Light emission of each compound is plotted as the fold of the mutant enzyme light emission over the wild-type luciferase light emission with that compound. The positively/negatively paired compounds are paired in red or blue and are on the left of the x-axis.

Some compounds also appear to be uniquely suited for developing orthogonality. For example, 4'-BrLuc is over-represented in the top 5,000 ranked hits, almost twice as frequently as any other compound (Fig. 4-10). It is unclear why 4'-BrLuc is uniquely selective. The bromine substituent is the same size as the methyl group on 4'-MeLuc negating any steric arguments for 4'-BrLuc uniqueness. Heavy halogen atoms are known to quench the fluorescence of some molecules *via* increased likelihood of

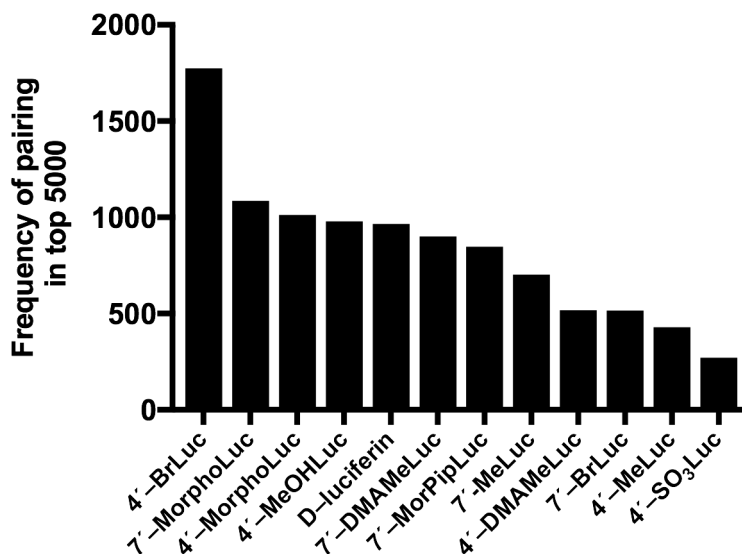


Figure 4-10: Frequency of compounds appearing in the top 0.6% (5,000 out of 829,000) of *in silico* orthogonal sets.

intersystem crossing raising the possibility that 4'-BrLuc is unique due to its electronic properties [14]. That is, perhaps most conformations result in internal quenching, resulting in good negative pairing with most enzymes; ones that can prevent this are uniquely suited for the task of light emission giving a small number of positive pairs. 4'-MeOHLuc also pairs with the same compounds (Fig. 4-8B) that 4'-BrLuc does. This could indicate that a more polarizable substituent at the 4'-position can impart selectivity

To determine if certain libraries correlated with orthogonality, we conducted a similar analysis of enzyme pairing frequency across the top 5,000 hits (Fig. 4-11). Two distinct bands were present in the heat map corresponding to all enzymes from the 240+347 library and three enzymes from the G3-7'-Morpho library. These enzymes showed high pairing frequency to form orthogonal sets, and likely comprise "hotspot" residues. Residues present in the 240 + 347 library are known to modulate substrate binding and light emission spectra with native D-luciferin [15-18]. Other interesting

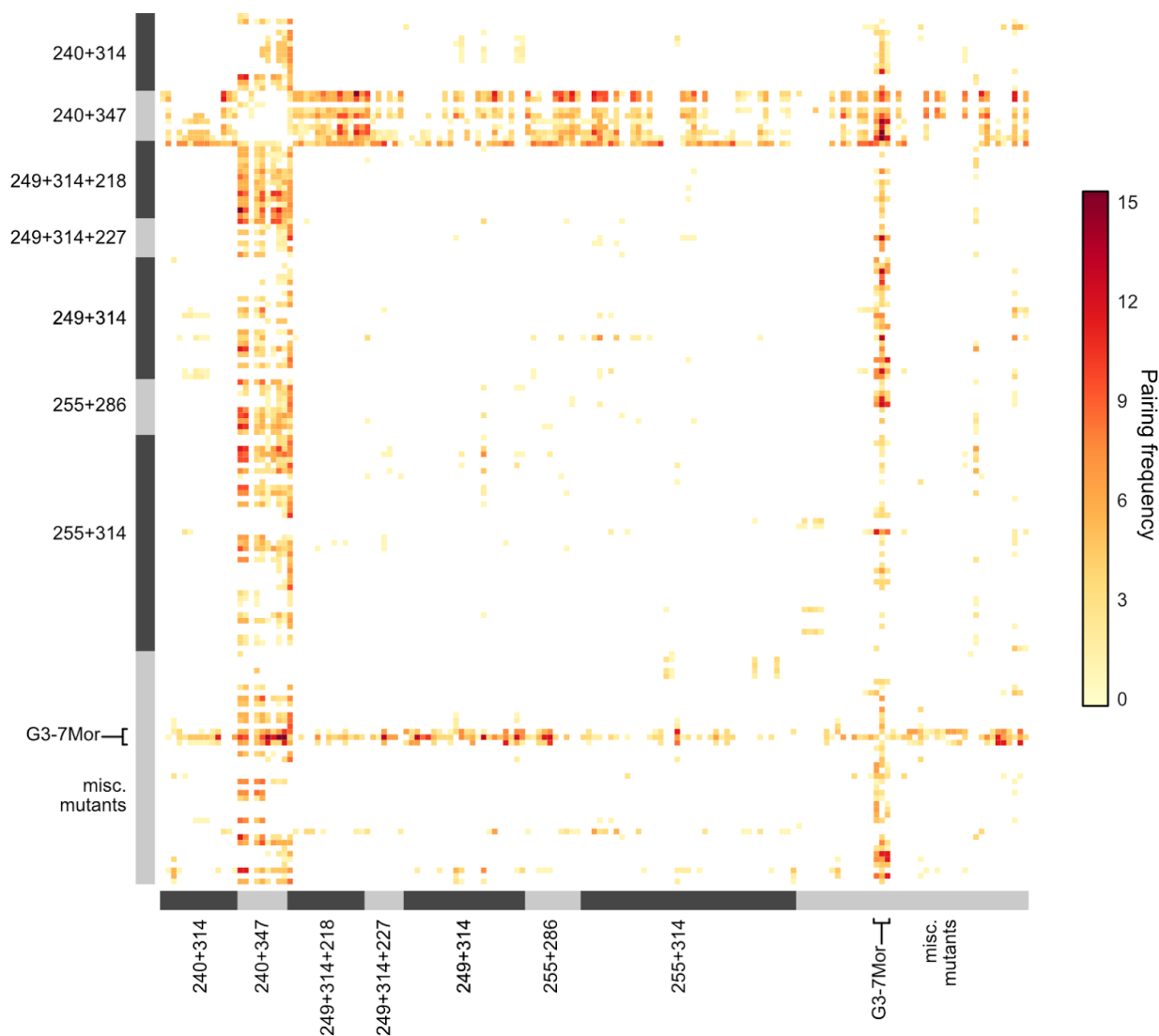


Figure 4-11. Enzyme-enzyme pairing frequency in the top 5,000 *in silico* screen hits. Each square represents a pairing of two enzymes. While certain libraries are very present in the top pairings, most contain one or two isolated hits.

bands are present in the library pairings however, less is known about the roles of the residues in these libraries and their effect on light emission. Further mutagenesis within these libraries (or surrounding residues) could yield additional improvements in orthogonality.

We also analyzed the frequency of positive and negative pairings between luciferins and individual mutated residues. Compounds with 4' modifications often

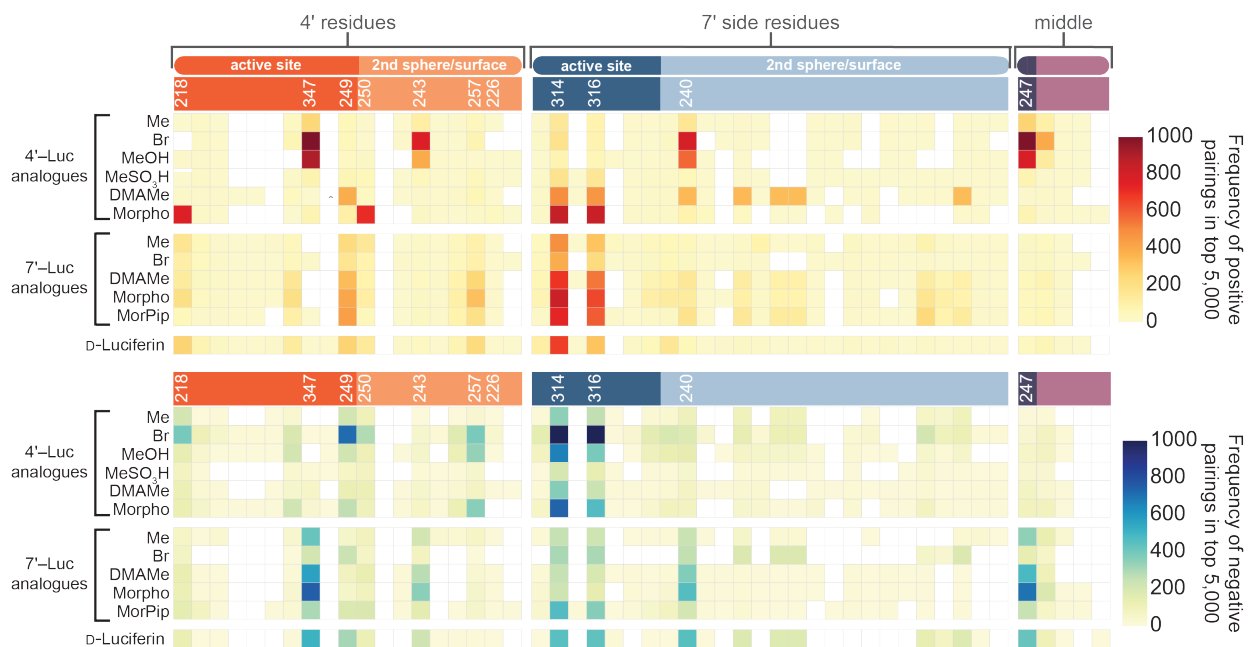


Figure 4-12. Compound-residue pairings frequency in the top 5,000 *in silico* screen hits. The heat maps depict the frequency of pairings between luciferase enzyme residues and the compound with which they have been positively (top) or negatively (bottom) paired with. In both plots the residues are roughly organized by their location in the enzyme, and their relative distance to the 4' or 7' sides of the luciferin substrate. The crystal structure shown in figure 4-1B is colored according to these different regions.

paired positively with mutant residues throughout the libraries examined (Fig. 4-12, top). Residues 240, 247, and 347 seem to prefer 4' modified compounds, in particular 4'-BrLuc and 4'-MeOHLuc. These residues are negatively paired with most of the 7' modified compounds, suggesting that they are good candidates for future orthogonal probe design (Fig 4-12, bottom). Mutations at 218 and 250 are highly selective for 4'-MorphoLuc, as previously observed [1]. It is interesting to note these residues seem to prefer 4'-MorphoLuc only and not other 4'-compounds.

By contrast, the 7' modified compounds were not positively paired with any residues with high frequency, except residues 314 and 316. The backbone of residues 314-316 are in close proximity to the 7'-carbon of the luciferin [12,19]. These residues

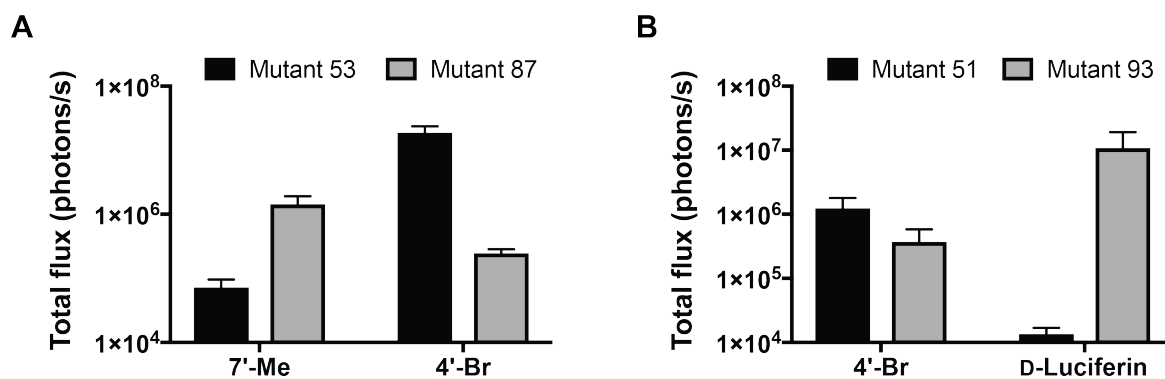


Figure 4-13. Orthogonality of luciferase-luciferin sets is maintained in mammalian cell culture. In both A and B 500 μ M Luciferin analogue was imaged with 100,000 DB7 cells. Error bars represent SEM for experiment performed in triplicate, data are representative of at least two biological replicates.

are positively paired with a variety of sterically large compounds hinting they improve selectivity for compounds with large steric bulk. With the luciferases examined, mutations seem to be biased toward selectivity with 4'-modified luciferins while negatively impacting 7' selectivity. As a consequence, the 7'-modified compounds required luciferases with multiple mutations (and likely dramatically altered active sites). The luciferin-binding region on the 7' face primarily comprises amide bonds of the protein backbone, while the 4' side is comprised of amino acid side chains [12,20]. These differences could account for many of the selectivity differences observed. Identified "hot spot" residues can be targeted in new libraries for screening.

4.5 Imaging with orthogonal pairs in vivo

Three of the top pairs from the script were further verified in cultured cell (Fig. 4-13) and animal models (Fig. 4-14). The analogues and luciferase enzymes chosen were the lead pair from Chapter 3 (4'-MorphoLuc and enzyme 81, and 7'-DMAMMeLuc and enzyme 37, see Chapter 4 for mammalian cell data), and two of the top performing pairs

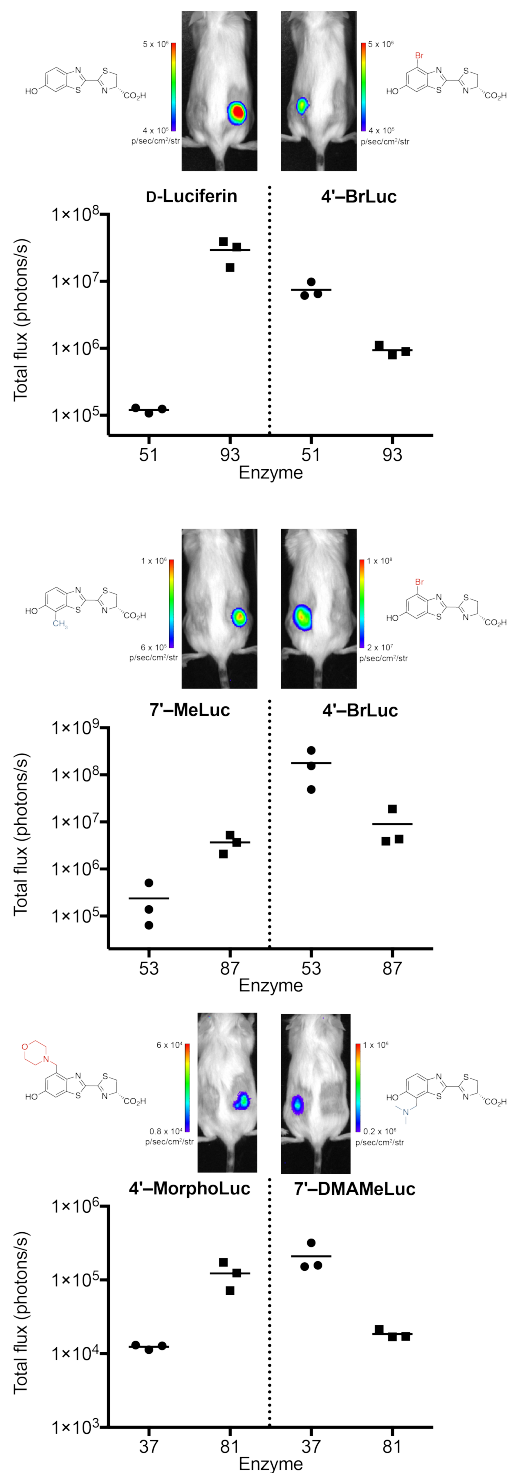


Figure 4-14. Animal model imaging of top orthogonal sets. For each mutant enzyme, 2×10^6 DB7 cells stably expressing the mutant were injected into FVB/NJ mice. Luciferin analogues (67 mM, 100 μ L) were injected i.p. 3 days post tumor injection. Images are representative of three mice for each orthogonal set. Light emission from all mice was quantified and shown with mean emission represented by black bars and individual animals represented by black squares or circles.

identified in this chapter. One set comprised enzyme 87 (R218K, F250Y, S314T, G316T) positively paired with 7'-MeLuc and enzyme 53 (V240I, V241M, F243M, F247Y, S347G) positively paired with 4'-BrLuc (Fig. 4-13A). The second set comprised enzyme 51 (F243M, S347G) positively paired with 4'-BrLuc and enzyme 93 (R218K, M249L, S314T, G316S) positively paired with D-luciferin (Fig. 4-13B). These pairs were selected from the top 20 orthogonal enzyme-substrate pairs due to the ease of accessing the substrates, along with their relative brightness. As demonstrated in Figure 4-13 orthogonality of luciferase-luciferin pairs is maintained in the transition from bacterial lysate to mammalian cell culture. Moving to more complex in vivo models the selectivity of the enzymes is again retained while the brightness of the compounds allows for subcutaneous imaging (Fig. 4-14). We hypothesized some of these sets may not reproduce in more complex environment due to differences in cell permeability or pharmacokinetics, however this was not the case with any of the luciferase-luciferin pairs carried on. Collectively, these data show that script analysis using bacteria cell lysate data is robust enough to transition top orthogonal sets to a variety of biological models.

4.6 Added diversity improves orthogonality

Certain mutants appeared more frequently in the top pairings than others, suggesting that selectivity that we were observing was due to a subset of "privileged" enzymes or substrates within the pool. To investigate this possibility, we selected random subsets of various sizes from the full pool of substrates and enzymes. We reran the algorithm on these random subsets to generate a surface of orthogonality scores (Fig. 4-15A). We observed that regardless of the identity of the enzymes or

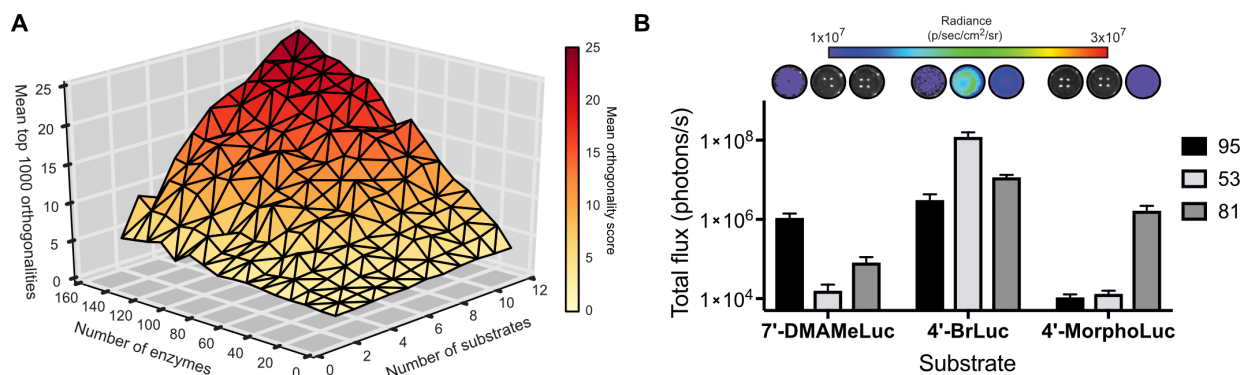


Figure 4-15. (A) Average orthogonality increases with increasing numbers of compounds and mutants. Results were simulated using random subsamples of compounds and mutants from the full dataset. Each point on the surface represents the average of 5 random subsamples. (B) Imaging of orthogonal triple sets. Bacterial cell lysates containing the luciferase mutant (different bars) of interest were imaged with 250 μM of luciferin analogue (x-axis). Images shown above graph are representative of experiments performed in triplicate. Error bars represent SEM for experiments run in triplicate and are representative of three biological replicates.

substrates in the list, orthogonality increased with both greater numbers of enzymes (from 2 to 159) and substrates (from 2 to 12). This result implies that we have not reached a plateau in identifying new luciferase and luciferin pairs. Exploring more sequence space with mutant luciferases and chemical space with modified luciferins will increase the orthogonality of the top pairings.

The modified algorithm could be readily adapted to searching for not just two pairs of orthogonal probes, but also multiple sets. A set of three pairs adds significant complexity, as three positive pairings, but six negative pairings must be identified. From our current data set, this would also require sifting through more than 144 million combinations. We re-evaluated the original dataset in search of three mutually orthogonal enzyme-substrate pairs. Thirty potential orthogonal sets were identified. The top triplet set contained two enzyme-substrate pairs we demonstrated in vivo (enzyme 53/4'-BrLuc and enzyme 81/4'-MorphoLuc). The orthogonality of the top 10 unique

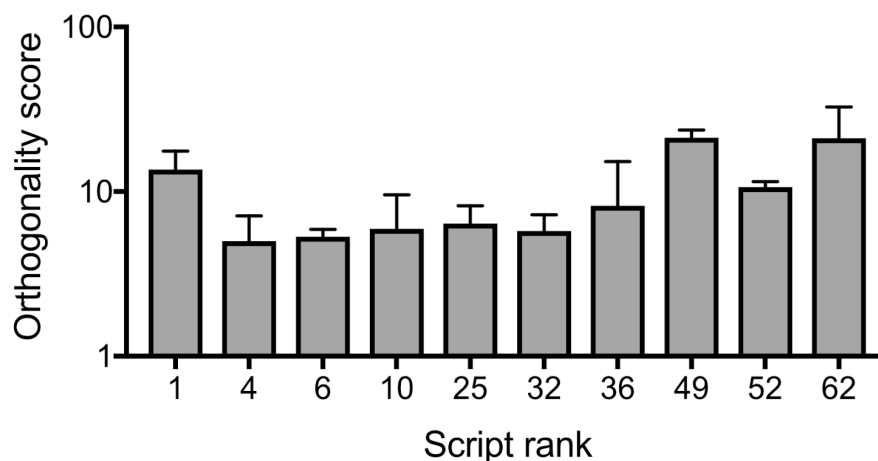


Figure 4-16. Unique sets of three luciferase-luciferin pairs maintain substrate selectivity. The top 10 unique luciferin triple enzyme-substrate pairs, as determined from the algorithm analysis, were verified in secondary screens. Bacteria expressing mutant enzymes were expanded, lysed, and split for controlled protein levels. Lysates were imaged with substrate resolved analogs and analyzed for orthogonality scores.

triplet sets were verified in bacterial lysate and mammalian cells (Figs. 4-15B, 5.16). All reproduced sets maintained orthogonality scores (the geometric mean of fold selectivity for the three luciferase-luciferin pairs) between ~5-20 (Fig. 4-16). The individual light emission for each luciferase and luciferin in the top orthogonal set is shown in Figure 4-15B as an example. These sets of three possessed an overall lower orthogonality score than the top sets of two demonstrated above and would necessitate further evolution to be viable for multicomponent bioluminescence imaging. However, they represent a proof-of-concept, and a starting point for development of sets of three mutually orthogonal luciferin-luciferase pairs.

4.7 Conclusions

We developed a general and rapid method to engineer orthogonal luciferase-luciferin pairs. This approach relies on developing an initial pool of functional

luciferases, and screening this collection with chemically diverse luciferins. We screened and generated a candidate list of greater than 800,000 possible pairings, and developed a computer algorithm to rapidly identify candidate orthogonal luciferase-luciferin pairs. Several of the pairs were translated into cultured cell and animal models, greatly expanding the number of viable imaging probes. We further analyzed the principles governing orthogonal substrate usage and how to identify higher-order sets of selective probes.

It should also be noted that the ranked list of orthogonal probes is a living data set. As new luciferin analogs are synthesized and luciferase mutants are characterized, they can be added. Such additions will increase the sequence and chemical space explored and may result in increased orthogonality. Additionally we show the results of this method can be used to analyze compound pairings and highlight critical residues. This information can subsequently guide future library design and synthetic targets.

The methods and algorithms presented here will accelerate the identification of orthogonal probes for bioluminescence, in addition to other areas where selective substrate usage is critical. The assay was shown to be robust enough to identify enzymes in bacterial lysate that maintained the desired activity *in vivo*. Any quantifiable enzymatic or protein-based output could be used for an amenable data set. This type of analysis could be done readily with troves of freezer stocks in directed evolution labs that work with commonly engineered enzymes and proteins such as P450s, tRNA synthetases, transcription factors, sortases and others.

4.8 Methods

4.8a General bioluminescence imaging

All analyses were performed in black 96-well plates (Grenier Bio One). Plates containing luminescent reagents were imaged in a light-proof chamber with an IVIS Lumina (Xenogen) CCD camera chilled to -90 °C. The stage was kept at 37 °C during the imaging session, and the camera was controlled using Living Image software. For assays with bacterial cell lysates, the exposure times for ranged from 1 s to 5 min, with data binning levels set to small or medium. Regions of interest were selected for quantification and total flux values were analyzed using Living Image software.

4.8b Construction of luciferase libraries

Two sections of the luciferase gene (*pgl4-luc2*), denoted R1 and R2, were targeted for gene assembly as described in Jones *et al.* [1]. To assemble mutant libraries, primers containing the codon(s) of interest (primer tables) were used in place of the primer coding for the wild-type sequence. Other libraries were created using standard quick change pcr techniques. All PCR reactions were run using Q5 Hot-start DNA polymerase (New England BioLabs).

Linearized vector was made from pET28-R1del or pET28-R2del as described in Jones *et al.* Insert containing the library was assembled with the linearized pET vector using Circular Polymerase Extension Cloning (CPEC) as described in Jones *et al.* or by Gibson Assembly. For Gibson Assembly 50 ng of *DpnI* digested, linearized vector and insert at 2:1 to 8:1 (insert:vector) were added to 10 µL of Gibson Assembly master mix

(homemade mix following the Prather recipe on http://www.openwetware.org/wiki/Gibson_Assembly with all enzymes and components purchased from New England Biolabs) and incubated at 50 °C for 20-60 min. 1-3 µL of the reaction mixture was then transformed into chemically competent cells. After transformation into Top10 or DH5α *E. coli*, cells were plated until the number of colonies was equal to or greater than 3 x the library size. Cells were then scraped off of the plates, combined, pelleted and minipreped. The minipreped plasmid DNA was saved for agar-plate screening.

Primer tables: Primers used to construct site-directed libraries. The bases highlighted in red denote sites targeted for saturation mutagenesis.

SD218 primers	
Forward primers	
SD218-F1	GCACCGCACCGCTTGTGTCKGGTTCAGTCATGCCC
SD218-F2	GCACCGCACCGCTTGTGTCKWKTTCAGTCATGCCC
SD218-F3	GCACCGCACCGCTTGTGTCVHGTTCAGTCATGCCC
SD218-F4	GCACCGCACCGCTTGTGTCTNATTCAGTCATGCCC
Reverse primers	
SD218-R1	GCCGAAGATGGGGTCGCGGGCATGACTGAAACCMGAC
SD218-R2	GCCGAAGATGGGGTCGCGGGCATGACTGAAAMWGAC
SD218-R3	GCCGAAGATGGGGTCGCGGGCATGACTGAAACDBGAC
SD218-R4	GCCGAAGATGGGGTCGCGGGCATGACTGAAATNGAC

SD227 primers	
Forward primers	
SD227-F1	GCGACCCCATCTGKGGCAACCAGATCATCCCCGACA
SD227-F2	GCGACCCCATCVBCGGCAACCAGATCATCCCCGACA
SD227-F3	GCGACCCCATCATGGCAACCAGATCATCCCCGACA
SD227-F4	GCGACCCCATCNAWGGCAACCAGATCATCCCCGACA
Reverse primers	
SD227-R1	GCCMCAAGATGGGGTCGCGGGCATGACTGAATCGGAC

SD227-R2	GCC GVB GATGGGGTCGCGGGCATGACTGAATCGGAC
SD227-R3	GCC CAT GATGGGGTCGCGGGCATGACTGAATCGGAC
SD227-R4	GCC WTN GATGGGGTCGCGGGCATGACTGAATCGGAC

SD240 primers	
Forward primers	
SD240-F1	CCGCTATCCTCAGC GBCGBC CCAG BCC ACCACGGC
SD240-F2	CCGCTATCCTCAGC GBCGBC CCAW TK ACCACGGC
SD240-F3	CCGCTATCCTCAGC GBCWTK CCAG BCC ACCACGGC
SD240-F4	CCGCTATCCTCAGC WTKGBC CCAG BCC ACCACGGC
SD240-F5	CCGCTATCCTCAGC WTKWTK CCAG BCC ACCACGGC
SD240-F6	CCGCTATCCTCAGC WTKGBC CCAW TK ACCACGGC
SD240-F7	CCGCTATCCTCAGC GBCWTK CCAW TK ACCACGGC
SD240-F8	CCGCTATCCTCAGC WTKWTK CCAW TK ACCACGGC
SD240-F9	NDT GGCATGTTACCACGCTGGGCTACTTGATCTGCG
SD240-F10	VHG GGCATGTTACCACGCTGGGCTACTTGATCTGCG
SD240-F11	TGG GGCATGTTACCACGCTGGGCTACTTGATCTGCG
Reverse primers	
SD240-R1	GVC GVCCTGAGGATAGCGGTGTCGGGGATGATCTGGTT
SD240-R2	MAW MAWGCTGAGGATAGCGGTGTCGGGGATGATCTGGTT
SD240-R3	MAW GVCCTGAGGATAGCGGTGTCGGGGATGATCTGGTT
SD240-R4	GVC MAWGCTGAGGATAGCGGTGTCGGGGATGATCTGGTT
SD240-R5	CGTGGTGAACATGCC AHN GCCGTGGTG GVC TGG
SD240-R6	CGTGGTGAACATGCC CDB GCCGTGGTG GVC TGG
SD240-R7	CGTGGTGAACATGCC CCA GCCGTGGTG GVC TGG
SD240-R8	CGTGGTGAACATGCC AHN GCCGTGGTG MAW TGG
SD240-R9	CGTGGTGAACATGCC CDB GCCGTGGTG MAW TGG
SD240-R10	CGTGGTGAACATGCC CCA GCCGTGGTG MAW TGG

SD249 primers	
Forward primers	
SD249-F1	CACGGCTTCGGC GBCNDTNDT ACGCTGGGCTACTTGATCTGCGG
SD249-F2	CACGGCTTCGGC GBCNDTVHG ACGCTGGGCTACTTGATCTGCGG
SD249-F3	CACGGCTTCGGC GBCNDTTGG ACGCTGGGCTACTTGATCTGCGG
SD249-F4	CACGGCTTCGGC GBCVHGNDT ACGCTGGGCTACTTGATCTGCGG

SD249-F5	CACGGCTTCGGC GBCVHGVHG ACGCTGGGCTACTTGATCT GCGG
SD249-F6	CACGGCTTCGGC GBCVHGTGG ACGCTGGGCTACTTGATCT GCGG
SD249-F7	CACGGCTTCGGC GBCTGGNDT ACGCTGGGCTACTTGATCT GCGG
SD249-F8	CACGGCTTCGGC GBCTGGVHG ACGCTGGGCTACTTGATCT GCGG
SD249-F9	CACGGCTTCGGC GBCTGGTGG ACGCTGGGCTACTTGATCT GCGG
SD249-F10	CACGGCTTCGGC WTKNDTNDT ACGCTGGGCTACTTGATCT GCGG
SD249-F11	CACGGCTTCGGC WTKNDTVHG ACGCTGGGCTACTTGATCT GCGG
SD249-F12	CACGGCTTCGGC WTKNDTTGG ACGCTGGGCTACTTGATCT GCGG
SD249-F13	CACGGCTTCGGC WTKVHGNDT ACGCTGGGCTACTTGATCT GCGG
SD249-F14	CACGGCTTCGGC WTKVHGVHG ACGCTGGGCTACTTGATCT GCGG
SD249-F15	CACGGCTTCGGC WTKVHGTGG ACGCTGGGCTACTTGATCT GCGG
SD249-F16	CACGGCTTCGGC WTKTGGNDT ACGCTGGGCTACTTGATCT GCGG
SD249-F17	CACGGCTTCGGC WTKTGGVHG ACGCTGGGCTACTTGATCT GCGG
SD249-F18	CACGGCTTCGGC WTKTGGTGG ACGCTGGGCTACTTGATCT GCGG
Reverse primers	
SD249-R1	GVCAHNAHN GCCGAAGCCGTGGTCAAATGG
SD249-R2	GVCAHNCDB GCCGAAGCCGTGGTCAAATGG
SD249-R3	GVCAHNCCA GCCGAAGCCGTGGTCAAATGG
SD249-R4	GVCCDBAHN GCCGAAGCCGTGGTCAAATGG
SD249-R5	GVCCDBCDB GCCGAAGCCGTGGTCAAATGG
SD249-R6	GVCCDBCCA GCCGAAGCCGTGGTCAAATGG
SD249-R7	GVCCCAAHN GCCGAAGCCGTGGTCAAATGG
SD249-R8	GVCCACADB GCCGAAGCCGTGGTCAAATGG
SD249-R9	GVCCACCA GCCGAAGCCGTGGTCAAATGG
SD249-R10	MAWAHNNDT GCCGAAGCCGTGGTCAAATGG
SD249-R11	MAWAHNCDB GCCGAAGCCGTGGTCAAATGG
SD249-R12	MAWAHNCCA GCCGAAGCCGTGGTCAAATGG
SD249-R13	MAWCDBAHN GCCGAAGCCGTGGTCAAATGG
SD249-R14	MAWCDBCDB GCCGAAGCCGTGGTCAAATGG
SD249-R15	MAWCDBCCA GCCGAAGCCGTGGTCAAATGG
SD249-R16	MAWCCAAHN GCCGAAGCCGTGGTCAAATGG
SD249-R17	MAWCCACDB GCCGAAGCCGTGGTCAAATGG

SD249-R18	MAWCCACCAGCCGAAGCCGTGGTCAAATGG
-----------	--------------------------------

SD255 primers	
Forward primers	
SD255-F1	TTCGGCATGTTCAACCACGCTGGGCNDTNDTNDTTGCG
SD255-F2	TTCGGCATGTTCAACCACGCTGGGCNDTNDTVHGTGCG
SD255-F3	TTCGGCATGTTCAACCACGCTGGGCNDTNDTTGGTGCG
SD255-F4	TTCGGCATGTTCAACCACGCTGGGCNDTVHGNDTTGCG
SD255-F5	TTCGGCATGTTCAACCACGCTGGGCNDTVHGVHGTGCG
SD255-F6	TTCGGCATGTTCAACCACGCTGGGCNDTVHGTGGTGCG
SD255-F7	TTCGGCATGTTCAACCACGCTGGGCNDTTGGNDTTGCG
SD255-F8	TTCGGCATGTTCAACCACGCTGGGCNDTTGGVHGTGCG
SD255-F9	TTCGGCATGTTCAACCACGCTGGGCNDTTGGTGGTGCG
Reverse primers	
SD255-R1	AGCACGACCCGAAAGCCGCAAHNAHNAHNGCCCAG
SD255-R2	AGCACGACCCGAAAGCCGCA CDBAHNAHNGCCCAG
SD255-R3	AGCACGACCCGAAAGCCGCA CCAAHNAHNGCCCAG
SD255-R4	AGCACGACCCGAAAGCCGCA AHNCDBAHNGCCCAG
SD255-R5	AGCACGACCCGAAAGCCGCA CDBCDBAHNGCCCAG
SD255-R6	AGCACGACCCGAAAGCCGCA CCACDBAHNGCCCAG
SD255-R7	AGCACGACCCGAAAGCCGCA AHNC CAAHNGCCCAG
SD255-R8	AGCACGACCCGAAAGCCGCA CDBCCAAHNGCCCAG
SD255-R9	AGCACGACCCGAAAGCCGCA CCACCAAHNGCCCAG

SD286 primers	
Forward primers	
SD286-F1	AGACTATAAGATTCAATCTGCCGBCGBCGBCCCCACAC
SD286-F2	AGACTATAAGATTCAATCTGCCGBCWTKWTKCCCACAC
SD286-F3	AGACTATAAGATTCAATCTGCCGBCWTKGBCCCCACAC
SD286-F4	AGACTATAAGATTCAATCTGCCGBCGBCWTKCCCACAC
SD286-F5	AGACTATAAGATTCAATCTGCCWTKWTKWTKCCCACAC
SD286-F6	AGACTATAAGATTCAATCTGCCWTKGBCGBCCCCACAC
SD286-F7	AGACTATAAGATTCAATCTGCCWTKGBCWTKCCCACAC
SD286-F8	AGACTATAAGATTCAATCTGCCWTKWTKGBCCCCACAC
Reverse primers	
SD286-R1	GCTCTTAGCGAAGAAGCTAAATAGTGTGGG GVC GVC GVC
SD286-R2	GCTCTTAGCGAAGAAGCTAAATAGTGTGGG MAW MAW GVC
SD286-R3	GCTCTTAGCGAAGAAGCTAAATAGTGTGGG GVC MAW GVC
SD286-R4	GCTCTTAGCGAAGAAGCTAAATAGTGTGGG MAW GVC GVC

SD286-R5	GCTCTTAGCGAAGAAGCTAAATAGTGTGGGMAWMAWMAW
SD286-R6	GCTCTTAGCGAAGAAGCTAAATAGTGTGGGVVCGVCMW
SD286-R7	GCTCTTAGCGAAGAAGCTAAATAGTGTGGGMAWVCMW
SD286-R8	GCTCTTAGCGAAGAAGCTAAATAGTGTGGGVCMWMAW

SD314 primers	
Forward primers	
SD314-F1	CTAAGCAACTTGCACGAGATCGCCNDTNDTNDT
SD314-F2	CTAAGCAACTTGCACGAGATCGCCNDTNDTVHG
SD314-F3	CTAAGCAACTTGCACGAGATCGCCNDTNDTTGG
SD314-F4	CTAAGCAACTTGCACGAGATCGCCVHGNDTNDT
SD314-F5	CTAAGCAACTTGCACGAGATCGCCVHGNDTVHG
SD314-F6	CTAAGCAACTTGCACGAGATCGCCVHGNDTTGG
SD314-F7	CTAAGCAACTTGCACGAGATCGCCTGGNDTNDT
SD314-F8	CTAAGCAACTTGCACGAGATCGCCTGGNDTVHG
SD314-F9	CTAAGCAACTTGCACGAGATCGCCTGGNDTTGG
Reverse primers	
SD314-R1	TTGCTGAGCGGCGCAHNAHNAHNGGCGATC
SD314-R2	TTGCTGAGCGGCGCCDBAHNAHNGGCGATC
SD314-R3	TTGCTGAGCGGCGCCCAAHNAHNGGCGATC
SD314-R4	TTGCTGAGCGGCGCAHNAHNCDBGGCGATC
SD314-R5	TTGCTGAGCGGCGCCDBAHNCDBGGCGATC
SD314-R6	TTGCTGAGCGGCGCCCAAHNCDBGGCGATC
SD314-R7	TTGCTGAGCGGCGCAHNAHNCAGGCGATC
SD314-R8	TTGCTGAGCGGCGCCDBAHNCAGGCGATC
SD314-R9	TTGCTGAGCGGCGCCCAAHNCAGGCGATC

SD337 primers	
Forward primers	
SD337-F1	NDTNDTGGCNDTGGCCTGACAGAAACAACACTAGTGCCA
SD337-F2	NDTVHGGGCNDTGGCCTGACAGAAACAACACTAGTGCCA
SD337-F3	NDTTGGGGCNDTGGCCTGACAGAAACAACACTAGTGCCA
SD337-F4	VHGNDTGGCNDTGGCCTGACAGAAACAACACTAGTGCCA
SD337-F5	VHGVHGGGCNDTGGCCTGACAGAAACAACACTAGTGCCA
SD337-F6	VHGTGGGGCNDTGGCCTGACAGAAACAACACTAGTGCCA
SD337-F7	TGGNDTGGCNDTGGCCTGACAGAAACAACACTAGTGCCA
SD337-F8	TGGVHGGGCNDTGGCCTGACAGAAACAACACTAGTGCCA
SD337-F9	TGGTGGGGCNDTGGCCTGACAGAAACAACACTAGTGCCA

Reverse primers	
SD337-R1	AGGCCAHNGCCAHNAHNGATGCCTGGTAGGT
SD337-R2	AGGCCAHNGCCAHNCDBGATGCCTGGTAGGT
SD337-R3	AGGCCAHNGCCAHNCCAAGATGCCTGGTAGGT
SD337-R4	AGGCCAHNGCCCDBAHNGATGCCTGGTAGGT
SD337-R5	AGGCCAHNGCCCDBCDBGATGCCTGGTAGGT
SD337-R6	AGGCCAHNGCCCDBCCAAGATGCCTGGTAGGT
SD337-R7	AGGCCAHNGCCCAAHNGATGCCTGGTAGGT
SD337-R8	AGGCCAHNGCCCCACDBGATGCCTGGTAGGT
SD337-R9	AGGCCAHNGCCCCACCAAGATGCCTGGTAGGT

SD347 primers	
Reverse primers	
SD347-R1	CTTCGGGGGTGATCAGAATAHNAHNAAGTTGTTTCTGTCAGG CCG
SD347-R2	CTTCGGGGGTGATCAGAATAHNCDBAGTTGTTTCTGTCAGG CCG
SD347-R3	CTTCGGGGGTGATCAGAATAHNCCAAGTTGTTTCTGTCAGG CCG
SD347-R4	CTTCGGGGGTGATCAGAATCDBAHNAAGTTGTTTCTGTCAGG CCG
SD347-R5	CTTCGGGGGTGATCAGAATCDBCDBGAGTTGTTTCTGTCAGG CCG
SD347-R6	CTTCGGGGGTGATCAGAATCDBCCAAGTTGTTTCTGTCAGG CCG
SD347-R7	CTTCGGGGGTGATCAGAATCAAHNAAGTTGTTTCTGTCAGG CCG
SD347-R8	CTTCGGGGGTGATCAGAATCCACDBGAGTTGTTTCTGTCAGG CCG
SD347-R9	CTTCGGGGGTGATCAGAATCCACCAAGTTGTTTCTGTCAGG CCG

4.8c Agar-plate screening of luciferin analogues

Miniprep library DNA or CPEC/Gibson Assembled DNA was transformed into BL21 or T7 Express *LysY* (New England Biolabs) and plated following the protocol in *Jones et al.* All compounds were added to the top-plating agar to a final concentration of 100 - 200 μ M. Light emitting colonies were picked and grown for further investigation following the protocol described in *Jones et al.* for secondary screening of library members.

4.8d Lysate screening of luciferase mutants and selection of functional mutants

See Chapter 3.8i for bacterial lysate imaging protocol.

4.8e List of mutant luciferase enzymes

Mutant number	Name	Sequence
1	4Me lead	F227C, M249L, S314T, A348G
2	7Br3C lead	R218H, F250T, S314C, Q338C
3	7Br3L lead	R218H, F250T, S314C, Q338L
4	7Br4C lead	R218H, F250T, S314C, G316T, Q338C
5	7Br4V lead	R218H, F250T, S314C, G316T, Q338V
13	G3_7Mor_16	M249F, T252S, F295L, S314T, G316T, A326V, P334S
14	G3_7Mor_18	M249L, S314V, G316S
15	G3_7Mor_21	M249L, S314C, G316A
16	G3_7Mor_24	M249F, Y266H, S293N, S314C, G316A
17	G3_7Mor_3	M249F, F294S, S314T, G316T V240A, M249L, M265T, F295L, I301T, S314T, G316T,
18	G3_7Mor_36	L321R
19	G3_7Mor_4	M249L, F294L, D305E, S314T, G316T
20	G3_7Mor_46	M249L, L267F, S314T, G316T
21	G3_7Mor_5	M249F, S314T, G316T
22	G3_7Mor_57	M249F, Y266H, L295F, I312V
23	G3_7Mor_58	V240A, M249L, L264F, S314T, G316T, K321R
24	G3_7Mor_65	M249L, Q283R, S314T, G316T
25	G3_7Mor_76	I226V, V240A, M249L, I282T, F295L, S314T, G316T
26	G3_7Mor_86	I226V, M249F, D279N, S314T, G316T
27	G3_7Mor_88	I232T, M249L, S314T, G316T
28	G3_7Mor_89	M249F, I282V, H310R, S314T, G316T
29	G5_7Mor_14	F227W, M249L, L286M, L287V, V288M, S314T, G316T
30	G5_7Mor_4	F227W, M249L, L287V, V288M, S314T, G316T
31	G5_7Mor_5	F227W, M249L, L287V, V288L, S314T, G316T
32	G5_7Mor_7	F227W, M249L, V288L, S314T, G316T
33	G5_7Mor_8	F227Y, M249L, L286M, L287V, V288L, S314T, G316T
34	WT	WT
35	R218A	R218A
36	R218H	R218H
37	R218K	R218K
38	rand_4Mor_28	T214A, E269G
39	rand_4Mor_31	K281E, F295L
40	rand_4Mor_38	T290A, F294S
41	rand_4Mor_60	I312V
42	rand_4Mor_71	H310R
43	rand_4Mor_83	E269G

Mutant number	Name	Sequence
44	SD240+314_7Me_11	V240I, V241A, S314T
45	SD240+314_7Me_13	V240A, V241A, F247L, S314T
47	SD240+314_7Me_21	V240A, F243M, S314T
48	SD240+314_7Me_32	F243M, S314T
49	SD240+314_7Me_40	V240A, F247Y, S314T
50	SD240+347_4Br_1	V240I, F247Y, S347G
51	SD240+347_4Br_12	F243M, S347G
52	SD240+347_4Br_17	V240A, V241M, F243M, F247Y
53	SD240+347_4Br_6	V240I, F243M, F247Y, S347G
54	SD240+347_4Me_14	V241A, F247L, S347A
55	SD240+347_4Me_3	V240L, V241L, F247Y, S347A
56	SD240+347_4Me_36	V240M, F243M, F247Y, S347A
57	SD240+347_4Me_39	V240G, F247Y, S347A
58	SD240+347_4Me_9	F243M, F247Y, S347A
59	SD249+314_4Br_14	S314C, G316T
60	SD249+314_4Br_15	S314T
61	SD249+314_4Br_16	M249L, S314C, G316S
62	SD249+314_4Br_18	M249L, F250M, S314C
63	SD249+314_4Br_22	M249L, S314C, G316A
64	SD249+314_4Br_3	S314C, G316S
65	SD249+314_4Br_38	M249L, G316S
66	SD249+314_4Br_43	M249F, S314C, G316S
67	SD249+314_4Me_16	F250Y, S314
68	SD249+314_4Me_2	F250M, S314T, G316T
69	SD249+314_4Me_21	F250Y, S314T, G316T
70	SD249+314_4Me_24	M249L, S314C
71	SD249+314_4Me_47	S314C, G316A
72	SD249+314_4Me_5	S314A, G316S
73	SD249+314_4Me_51	F250T, S314T, G316T
74	SD249+314_7Br_24	F250H, S314C
75	SD249+314_7Br_3	F250T, S314C
76	SD249+314_7Br_4	F250T, S314C, G316T
77	SD249+314_7Me_21	M249L, S314T
78	SD249+314_7Me_25	M249L, S314V
79	SD249+314_7Me_4	M249L, S314T, G316S
80	SD249+314_7Me_5	M249L, S314T, G316T
81	SD249+314+218_4Me_1	R218A, F250M, S314T, G316T
82	SD249+314+218_4Me_13	R218K, F250M, S314T, G316T
83	SD249+314+218_4Me_15	R218K, S314C, G316A
85	SD249+314+218_4Me_25	R218K, F250M, S314T, G316T
86	SD249+314+218_4Me_27	R218K, M249L, S314C
87	SD249+314+218_4Me_34	R218K, F250Y, S314T, G316T
88	SD249+314+218_4Me_35	R218K, F250M, S314C
89	SD249+314+218_4Me_5	R218K, S314T, G316T
90	SD249+314+218_4Me_9	R218K, F250Y, S314C
91	SD249+314+218_7Me_2	R218K, M249L, S314T
92	SD249+314+218_7Me_24	R218K, M249L, S314T, G316S

Mutant number	Name	Sequence
93	SD249+314+218_7Me_27	R218K, M249L, S314V
94	SD249+314+218_7Me_30	R218K, M249L, S314V, G316S
95	SD249+314+227_4Br_1	F227Y, M249L, F250Y, G316S
96	SD249+314+227_4Br_2	F227Y, M249L, S314C
97	SD249+314+227_4Br_3	F227Y, S314C, G316T
98	SD249+314+227_4Br_5	F227Y, S314T
99	SD249+314+227_4Br_6	F227Y, S314C, G316A
100	SD249+314+227_4Br_8	F227Y, M249L, S314C, G316S
101	SD249+314+227_7Br_1	F227Y, F250H, S314C
102	SD255+314_4Br_10	Y255F, L256I, I257V, S314C
103	SD255+314_4Br_11	Y255H, I257V, S314C
104	SD255+314_4Br_2	L256I, I257F, S314T
105	SD255+314_4Br_27	L256I, I257F, S314C, G316S
107	SD255+314_4Br_40	Y255N, I257V, S314C, G316T
108	SD255+314_4Br_42	Y255H, I257V, S314C
109	SD255+314_4Br_49	Y255N, I257V, S314C
110	SD255+314_4Br_5	S347A
111	SD255+314_4Br_50	L256I, I257F, S314C, G316T
112	SD255+314_4Br_59	not sequenced
113	SD255+314_4Br_61	Y255H, I257V, G316S
114	SD255+314_4Br_62	Y255F, L256I, I257V, S314C, G316S
115	SD255+314_4Br_64	L256M, I257C, S314C
116	SD255+314_4Br_9	Y255F, I257L, S314T
117	SD255+314_4Me_1	I257M, S314T
118	SD255+314_4Me_10-1	I257S, 314V, G315H, G316R
119	SD255+314_4Me_10-3	Y255H, S314V
120	SD255+314_4Me_13	I257A, S314C
121	SD255+314_4Me_18	I257C, S314T, G316T
122	SD255+314_4Me_6	L256M, S314V, G316S
123	SD255+314_4Me_7	I257S, S314T, G316T
124	SD255+314_7Br_10	Y255F, I257L, S314C, G316T
125	SD255+314_7Br_15	Y255F, I257L, S314V, G316T
126	SD255+314_7Br_3	Y255F, I257L, S314C, G316A
127	SD255+314_7Br_5	Y255F, I257L, S314C, G316S
128	SD255+314_7Br_6	Y255F, I257L, S314V, G316S
129	SD255+314_7Br_8	Y255F, I257L, S314C
130	SD255+314_7Me_14	I257R, S314T, G316S
131	SD255+314_7Me_21	Y255H, I257R, S314C, G316A
132	SD255+314_7Me_22	Y255H, I257S, S314V
133	SD255+314_7Me_25	Y255F, I257L, G316S
134	SD255+314_7Me_27	not sequenced
135	SD255+314_7Me_29	Y255F, I257M, S314C, G316A
136	SD255+314_7Me_31	Y255H, I257L, S314T, G316T
137	SD255+314_7Me_35	Y255F, I257M, S314V, G316S
138	SD255+314_7Me_36	Y255F, I257L, S314A, G316S
139	SD255+314_7Me_37	Y255F, I257L, G316A
140	SD255+314_7Me_48	Y255H, I257R, S314T, G316S

Mutant number	Name	Sequence
141	SD255+314_7Me_8	L256M, I257F, S314T, G316T
142	SD240+314_4Br_1	V240L, V241F, F247S, S314C, G316A
143	SD240+314_4Br_8	V240A, V241A, S314C, G316S
144	SD240+314_4Me_1	V241A, F247Y, S314T, G316T
145	SD240+314_4Me_6	V241A, F243L, F247Y, S314C
146	SD240+314_4Me_15	V241A, F247L, S314T, G316T
147	SD240+314_7Me_2	V240M, V241M, F247Y, S314V
148	SD240+314_7Me_5	V240A, V241A, F247L, S314C, G316A
149	SD240+314_7Me_13	F243M, G316S
150	SD240+314_7Me_16	F243M
156	SD255+286_4Me113	Y255H, L286V, L287M, V288A
157	SD255+286_4Me143	L256I, I257M, L286V, V288A
158	SD255+286_7Me30	L256M, I257F, L286I, L287V, V288L
159	SD255+286_7Me31	Y255I, I257C, L287V, V288A

4.8e Complete analogue/mutant luciferase screen

Functional and sequentially diverse mutants identified in the secondary screen were retransformed into *E. coli* BL21 cells and aliquots were stored as glycerol stocks at -80 °C. Luciferin compounds were dissolved at 10 mM in 100 mM Phosphate buffer (pH 7.8), aliquoted and stored -80 °C. These accumulated glycerol stocks were then imaged in batches across the entire set of luciferin compounds. Glycerol stocks were grown in 5 mL of LB-Kan media to an OD ~0.8, induced with 500 µM IPTG for 2 h at 30 °C, then pelleted. Cell pellets were lysed in Fluc lysis buffer (50 mM Tris•HCl, 500 mM NaCl, 0.5% (v/v) Tween, 5 mM MgCl₂, pH = 7.4), 600 µL per pelleted culture tube. This cell lysate was then spread across six wells (90 µL/well) on six different 96-well black plates. Wild-type luciferase was always included with each batch as a check on compound integrity. To each well 10 µL of 10 mM luciferin and 1 mM ATP was added and the plate was imaged for 1-60 s. This was repeated until all 12 compounds were imaged with all 159 luciferase mutants for a total of 1908 data points.

4.8f Mammalian plasmid construction

To express the mutant luciferases in mammalian cells, the luciferase gene was amplified and inserted into pBMN-IRES-GFP [21].

The luciferase gene from mutant hits were amplified with the following primers:

5'- ataacgcgatggaagatgccaaaaacattaaga-3' and

5'-gagaggggatgcattattacacggcgatcttgcc-3'

The insert was restriction enzyme digested with *NsiI* and *MluI* (New England Biolabs) and was then cloned into the pBMN-IRES-GFP vector with T4 ligase (New England Biolabs). Sequencing analysis confirmed correct mutations.

4.8g Mammalian cell culture

DB7 cells (courtesy of the Contag laboratory, Stanford) were cultured in DMEM (Corning) supplemented with 10% (vol/vol) fetal bovine serum (FBS, Life Technologies), penicillin (100 U/mL), and streptomycin (100 µg/mL). Cells were maintained in a 5% CO₂ water-saturated incubator at 37 °C. To create stable lines expressing mutant luciferases DB7 cells were transduced with ecotropic retrovirus (Phoenix packaging system) as previously described,(Helms Prescher Contag 2010) followed by selection with puromycin (10 µg/mL). The cells were sorted via FACS at the Institute for Immunology Flow Cytometry Core (UCI).

4.8h Mammalian cell imaging with luciferase mutants

DB7 cells stably expressing Fluc or mutant luciferases were added to black 96-well plates (1×10^5 cells per well). A stock solution of luciferin (5 mM in PBS) was added to each well (500 μ M final concentration). Sequential imaging in black 96-well plates was performed as above.

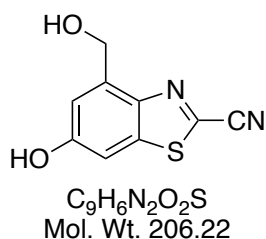
4.8i In vivo imaging of orthogonal luciferase-luciferin pairs

FVB/NJ mice (The Jackson Laboratory) received subcutaneous dorsal injections of 2×10^6 or 6.5×10^6 DB7 mutant luciferase expressing cells. After 24 h, animals received an i.p. injection of luciferin (67 mM or 100 mM, 100 μ L per mouse). Mice were anesthetized (2% isoflurane) and placed on the warmed (37 °C) IVIS stage for imaging. Bioluminescent photons were quantified for the designated regions of interest. Before collecting data for the second administered luciferin, mice were imaged for residual signal from the previous imaging session. Mice were imaged every other day over 5 days.

4.8j General synthetic methods

All reagents purchased from commercial suppliers were of analytical grade and used without further purification. 4,5-Dichloro-1,2,3-dithiazolium chloride, was prepared according as previously reported [22]. Reaction progress was monitored by thin-layer chromatography on EMD 60 F254 plates, visualized with UV light, ceric ammonium molybdate (CAM), chloranil, or KMnO_4 stain. Compounds were purified via flash column

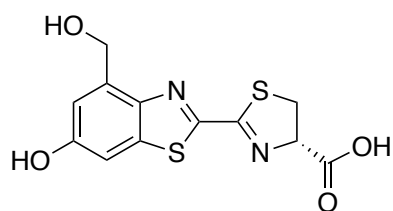
chromatography using Siliaflash F60 60 Å, 230-400 mesh silica gel (Silicycle), unless otherwise stated. HPLC purifications were performed on a Varian ProStar equipped with a 325 Dual Wavelength UV-Vis detector. Semi-preparative runs were performed using an Agilent Prep-C18 Scalar column (9.4 x 150 mm, 5 µm), preparative runs were performed using an Agilent Eclipse XD8-C18 PrepHT column (21.2 x 250 mm 7 µm). Anhydrous solvents were dried by passage over neutral alumina. Reaction vessels were either flame or oven dried prior to use. NMR spectra were acquired with Bruker Advanced spectrometers. All spectra were acquired at 298 K. ¹H-NMR spectra were acquired at either 500 or 400 MHz, and ¹³C-NMR spectra were acquired at 125 MHz. Coupling constants (*J*) are provided in Hz and chemical shifts are reported in ppm relative to either residual non-deuterated NMR solvent, calculated reference, or to a methanol external reference. Low and high-resolution electrospray ionization (ESI) mass spectra were collected at the University of California- Irvine Mass Spectrometry Facility.



6-hydroxy-4-(hydroxymethyl)benzo[d]thiazole-2-carbonitrile.

Following the procedure of Kulangiappar *et al.* [23] 6-acetoxy-4-[bromomethyl]-1,3-benzothiazole-2-carbonitrile (**3-14**) (0.150 g, 0.482 mmol), was dissolved in MeCN (10 mL) and stirred at rt in a round bottom flask. Sodium nitrate was dissolved in H₂O (10 mL) and added to the reaction mixture. The flask was flushed with N₂ and then stirred at 80 °C for 22 h. The volatiles were evaporated *in vacuo*, and the resulting aqueous solution was diluted with 1 M NaHSO₄ (50 mL) and extracted with EtOAc (3 x 50 mL). The organic layers were

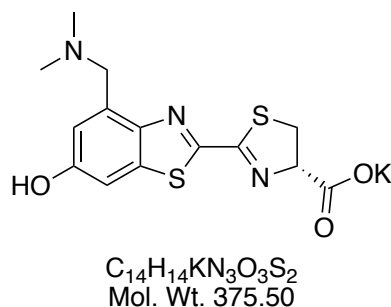
combined, washed with brine (1 x 50 mL), dried over MgSO₄, and concentrated *in vacuo*. The resulting white solid was carried on without further purification or characterization.



C₁₂H₁₀N₂O₄S₂
Mol. Wt. 310.34

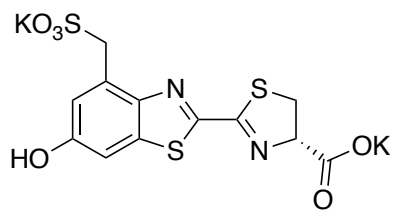
(S)-2-(6-hydroxy-4-(hydroxymethyl)benzo[d]thiazol-2-yl)-4,5-dihydrothiazole-4-carboxylic acid (4'-MeOHLuc).

The crude product (6-hydroxy-4-(hydroxymethyl)benzo[d]thiazole-2-carbonitrile), was dissolved in MeCN (5 mL) and added to a scintillation vial and stirred under N₂. K₂CO₃ (0.067 g, 0.49 mmol) and D-Cysteine•HCL•H₂O (0.086 g, 0.49 mmol) were dissolved in H₂O (1 ml) then added to the reaction mixture. Upon disappearance of starting material the volatiles were evaporated *in vacuo*. The resulting aqueous solution was diluted with H₂O (25 mL), acidified to pH 2, and extracted with EtOAc (5 x 25 mL). The organic layers were combined, dried over MgSO₄, and concentrated *in vacuo* to provide **4'-MeOHLuc**, as a dark orange solid (0.087 g, 60%). ¹H NMR (400 MHz, D₂O) δ 7.14 (d, *J* = 2.2, 1H), 6.99 (d, *J* = 2.2, 1H), 5.20 (dd, *J* = 8.3, *J* = 9.7, 1H), 4.93 (app dd, *J* = 14.0, *J* = 20.8, 2H), 3.81 (app t, *J* = 10.6, 1H), 3.61 (dd, *J* = 8.4, *J* = 11.2, 2H); ¹³C NMR (125 MHz, D₂O) δ 177.8, 165.9, 157.6, 155.8, 144.0, 137.4, 136.3, 115.0, 105.8, 80.1, 60.4, 36.4. HRMS (ESI⁻) calcd for C₁₂H₉O₄N₂S₂ [M-H]⁻ = 309.0004, found 308.9994.



(S)-2-(4-((Dimethylamino)methyl)-6-hydroxybenzo[d]thiazol-2-yl)-4,5-dihydrothiazole-4-carboxylate, potassium salt (4'-DMAMeLuc).

To a solution of 6-acetoxy-4-[bromomethyl]-1,3-benzothiazole-2-carbonitrile (**3-14**) (0.196 g, 0.630 mmol) in MeCN (2 mL) stirring in a scintillation vial, K_2CO_3 (0.261 g, 1.89 mmol) in H_2O (1 mL), and dimethylamine (40 wt. % in H_2O , 0.24 mL, 1.89 mmol) were added. The reaction was stirred at rt until disappearance of starting material (30 min), followed by addition of D-cysteine•HCl• H_2O (0.111 g, 0.630 mmol) to the reaction mixture. The reaction was further stirred at rt until disappearance of the intermediate dimethylamine cyanobenzothiazole. Upon completion the reaction mixture was concentrated *in vacuo* to yield the potassium salt as a yellow solid (0.13 g, 54%) following HPLC purification (preparative, reversed phase, with the following elution protocol: 100% H_2O for 5 min, followed by a gradient of 0-90 % MeOH in H_2O for 15 min. The flow rate was 20 mL/min). 1H NMR (500 MHz, CD_3OD) δ 7.24 (s, 1H), 7.04 (s, 1H), 5.16 (t, $J = 9.3$, 1H), 4.04 (app dd, $J = 13.2$, $J = 20.8$, 2H), 3.68 (app pentet, $J = 9.1$, $J = 11$, 2H) 2.34 (s, 6H); ^{13}C NMR (125 MHz, CD_3OD) δ 177.6, 165.8, 159.2, 158.3, 147.7, 139.3, 134.6, 119.1, 106.9, 83.2, 59.6, 45.3, 37.1. HRMS (ESI $^-$) calcd for $C_{13}H_{14}N_3OS_2$ $[M-CO_2K]^-$ = 292.0578, found 292.0573.



$C_{12}H_8K_2N_2O_6S_3$
Mol. Wt. 450.58

(S)-2-(6-Hydroxy-4-(sulfonatomethyl)benzo[d]thiazol-2-yl)-4,5-dihydrothiazole-4-carboxylate dipotassium salt (4'-SO₃HLuc).

6-Acetoxy-4-[bromomethyl]-1,3-benzothiazole-2-

carbonitrile (**3-14**) (0.200 g, 0.643 mmol) was dissolved in acetone (5 mL) and stirred in a round bottom flask. A solution of sodium sulfite (0.243 g, 1.929 mmol) in H₂O (5 mL) was then added to the flask. The reaction mixture was then stirred at RT for 24 h. The acetone was removed *via* rotary evaporation. K₂CO₃ (0.090, 0.649 mmol) and D-cysteine•HCL•H₂O (0.114 g, 0.649 mmol) were dissolved in H₂O (1 ml) then added to the reaction mixture. The reaction was stirred at RT for 18 h. Upon completion the reaction mixture was concentrated *in vacuo* to yield the potassium salt as a yellow solid (0.13 g, 45%) following HPLC purification (preparative, reversed phase, with the following elution protocol: 100% H₂O for 5 min, followed by a gradient of 0-90 % MeCN in H₂O for 15 min). The flow rate was 20 mL/min. ¹H NMR (500 MHz, D₂O) δ 7.29 (d, *J* = 2.4, 1H), 7.11 (d, *J* = 2.4, 1H), 5.22 (dd, *J* = 8.2, *J* = 9.8, 2H), 4.63 (s, 2H), 3.83 (dd, *J* = 9.9, *J* = 11.1, 1H), 3.61 (dd, *J* = 8.2, *J* = 11.1, 1H). ¹³C NMR (125 MHz, solvent CD₃OD) δ 177.6, 165.7, 158.5, 158.2, 148.1, 139.0, 131.3, 119.2, 106.7, 83.1, 53.5, 37.3. HRMS (ESI⁻) calcd C₁₂H₉N₂O₆S₃ [M-H]⁻ = 372.9623, found 372.9619.

References

1. 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.
2. 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*.
3. Renata, H.; Wang, Z. J.; Arnold, F. H., Expanding the enzyme universe: Accessing non-natural reactions by mechanism-guided directed evolution. *Angew. Chem. Int. Ed.* **2015**, *54*, 3351.
4. Goldsmith, M.; Tawfik, D. S., Directed enzyme evolution: Beyond the low-hanging fruit. *Curr. Opin. Struct. Biol.* **2012**, *22*, 406.
5. Packer, M. S.; Liu, D. R., Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **2015**, *16*, 379.
6. Taylor, N. D.; Garruss, A. S.; Moretti, R.; Chan, S.; Arbing, M. A.; Cascio, D.; Rogers, J. K.; Isaacs, F. J.; Kosuri, S.; Baker, D.; Fields, S.; Church, G. M.; Raman, S., Engineering an allosteric transcription factor to respond to new ligands. *Nature Methods* **2016**, *13*, 177.
7. Feng, J.; Jester, B. W.; Tinberg, C. E.; Mandell, D. J.; Antunes, M. S.; Chari, R.; Morey, K. J.; Rios, X.; Medford, J. I.; Church, G. M.; Fields, S.; Baker, D., A general strategy to construct small molecule biosensors in eukaryotes. *eLife* **2015**, *4*.
8. Lang, K.; Chin, J. W., Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem. Rev.* **2014**, *114*, 4764.
9. Neumann, H.; Wang, K. H.; Davis, L.; Garcia-Alai, M.; Chin, J. W., Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **2010**, *464*, 441.
10. Wan, W.; Huang, Y.; Wang, Z. Y.; Russell, W. K.; Pai, P. J.; Russell, D. H.; Liu, W. R., A facile system for genetic incorporation of two different noncanonical amino acids into one protein in escherichia coli. *Angew. Chem. Int. Ed.* **2010**, *49*, 3211.
11. 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.

12. 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.
13. Steinhardt, R. C.; Rathbun, C. M.; Krull, B. T.; Yu, J.; 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.
14. Solovyov, K. N.; Borisevich, E. A., Intramolecular heavy-atom effect in the photophysics of organic molecules. *Physics-Uspekhi* **2005**, *48*, 231.
15. 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.
16. 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.
17. Viviani, V. R.; Prado, R. A.; Neves, D. R.; Kato, D.; Barbosa, J. A., A route from darkness to light: Emergence and evolution of luciferase activity in amp-coa-ligases inferred from a mealworm luciferase-like enzyme. *Biochemistry* **2013**, *52*, 3963.
18. Harwood, K. R.; Mofford, D. M.; Reddy, G. R.; Miller, S. C., Identification of mutant firefly luciferases that efficiently utilize aminoluciferins. *Chem. Biol.* **2011**, *18*, 1649.
19. Viviani, V. R.; Amaral, D. T.; Neves, D. R.; Simões, A.; Arnoldi, F. G. C., The luciferin binding site residues c/t311 (s314) influence the bioluminescence color of beetle luciferases through main-chain interaction with oxyluciferin phenolate. *Biochemistry* **2013**, *52*, 19.
20. 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.
21. Liu, H. P.; Patel, M. R.; Prescher, J. A.; Patsialou, A.; Qian, D. L.; Lin, J. H.; 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.
22. McCutcheon, D. C.; Porterfield, W. B.; Prescher, J. A., Rapid and scalable assembly of firefly luciferase substrates. *Org. Biomol. Chem.* **2015**, *13*, 2117.

23. Kulangiappar, K.; Kulandainathan, M. A.; Raju, T., Conversion of benzylic bromides to benzaldehydes using sodium nitrate as an oxidant. *Ind. Eng. Chem. Res.* **2010**, *49*, 6670.

Chapter 5: Pursuing more orthogonal luciferase-luciferin pairs with improved selectivity

5.1 Introduction

The last chapter highlighted the success of using diverse luciferin structures and large pools of luciferase mutants to identify orthogonal sets. This approach identified at least three pairs exhibiting 10-fold substrate resolution in animal models. While exciting, it would be ideal to have pairs with enhanced selectivity, in order to sufficiently distinguish cell-types in vivo. Improving the resolution would enable important biological studies, such as tracking tumor-immune cell interactions.

This chapter describes the development and screening of additional luciferins that can be used in conjunction with earlier analogues for improved orthogonal imaging. A total of 8 additional luciferins were synthesized and characterized (Fig. 5-1) and 207 sequentially diverse luciferases were assessed. Collectively, these tools provide a total of 3.6 million potential orthogonal sets of luciferase-luciferins. All pairs were analyzed using the cross-compare algorithm described in Chapter 4 to identify interesting orthogonal sets. From this large comparison, one set caught our eye, as it demonstrated an approximately 1000-fold difference in substrate usage compared to native Fluc change from wild-type Fluc preference. This provided an excellent starting point for evolving a set of luciferase-luciferin pairs with increased orthogonality.

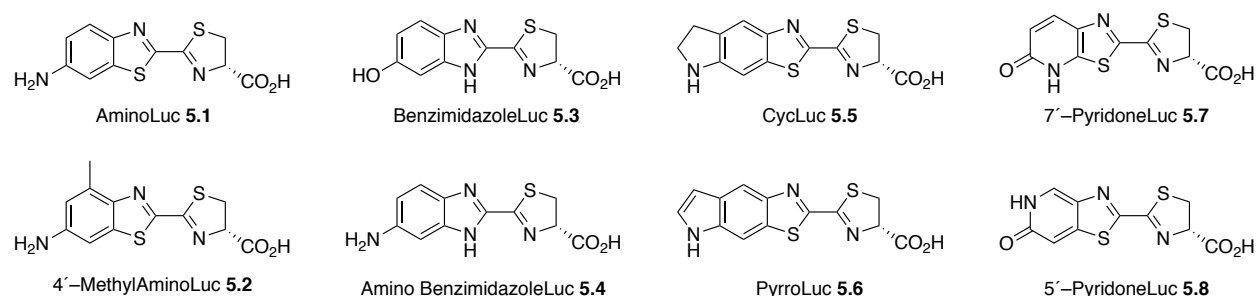


Figure 5-1. Structures of electronically altered luciferin analogues

5.2 Expanding the compound base

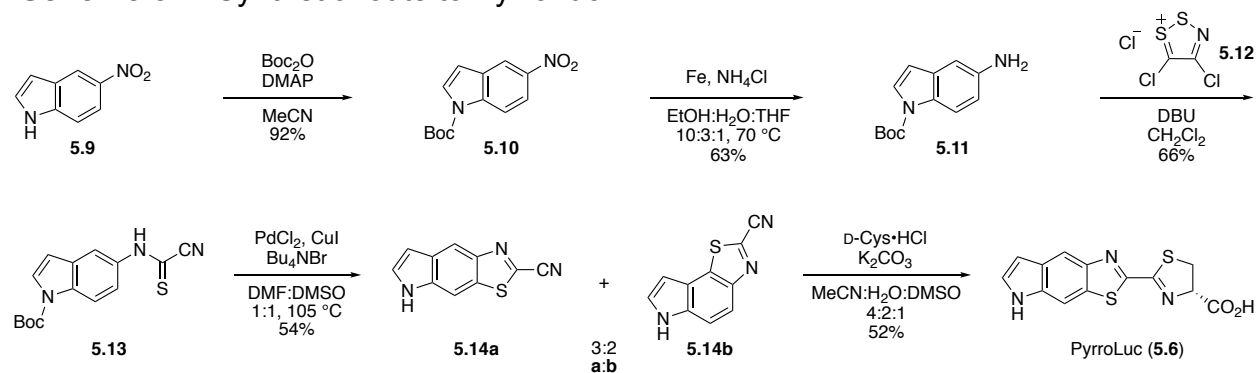
The luciferin analogues described in the previous chapters were mainly limited to steric modifications. We reasoned that these compounds would maintain light-emitting capabilities, unlike electronically perturbed compounds [1-2]. Scaffolds with altered heteroatoms or pi systems electronic modifications could be more likely to shift the wavelength of emission. However, a variety of electronically modified luciferins have been synthesized by us [3-5] and others [5-12]. Since we had already established a robust screening platform, we turned our attention to these more challenging targets. Many of these molecules (Fig. 5-1) are competent light emitters and present different binding interactions that can be modified in order to achieve selective binding. For example, Miller and colleagues demonstrated the cyclic amino luciferin (CycLuc) scaffold could achieve selective binding with a small number of mutations [12-13]. Combining electronically modified luciferins with sterically modified luciferin analogues has the potential to increase the selectivity achieved with mutant luciferase enzymes.

We initially prepared luciferins with altered heteroatoms within the benzothiazole ring, in addition to an analogue based off of the CycLuc structure. Others in the Prescher lab and I have synthesized 6'-amino luciferins (**5.1**, **5.2**), benzimidazole

scaffolds (**5.4**, **5.4**), CycLuc (**5.5**) and pyrrolo (**5.6**) scaffolds, and pyridone luciferins (**5.7**, **5.8**). 6'-Amino luciferin is a well-characterized luciferin [5,14] and a small steric group, 4'-methyl, was added to increase diversity. The benzimidazole luciferin comprises a nitrogen atom (in place of a sulfur) in the core ring and has been reported to be a competent light emitter [3,6]. Other modified aromatic architectures, such as the pyridones, were also explored. Pyridones are known to have interesting tautomeric and electronic properties [15-16] and both the 5' and 7'-PyridoneLuc compounds were accessed from a similar synthetic route. The pyrrolo luciferin (PyrroLuc, **5.6**) was also prepared as we hypothesized the altered aromatic core could also generate unique colors of bioluminescent light emission [7,9,17]. This compound likely has improved cell permeability with unsaturated CycLuc core. Also, There is a wealth of chemistries available to derivatize the indole like motif in this structure [18-19].

To synthesize the target scaffold **5.6**, I used the Appel's salt condensation method described earlier (Scheme 5.1). The starting indole **5.9** in this case required a protecting a group on nitrogen to avoid undesired cross-reactivity. Initially, a tosyl group was explored, but due to difficulties in removing the motif at later stages, this route was abandoned. I next examined a Boc group due to its relative ease of removal and

Scheme 5-1. Synthetic route to PyrroLuc.



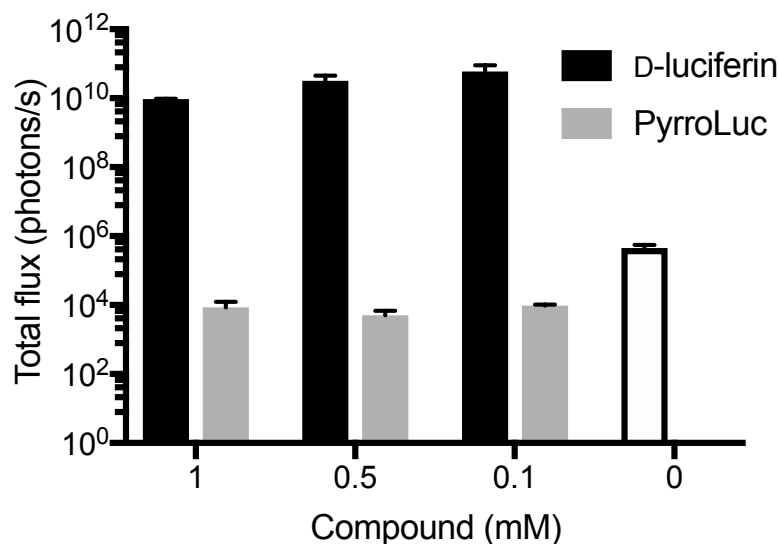


Figure 5-2. Light emission of PyrroLuc with Fluc (1 $\mu\text{g}/\text{well}$) at varying concentrations (gray bars) as compared to D-Luciferin (black bars) and no substrate (white bar). Error bars represent standard error of the mean for data run in triplicate, data are representative of three independent experiments.

compatibility with nitro group reduction. Addition of the Boc group proceeded smoothly. The resulting protected nitroindole (**5.10**) was then reduced with iron and ammonium chloride to yield the amino indole (**5.11**) in good yield [20]. With the protected compound in hand, I condensed the material with **5.12** and fragmented the product to afford **5.13**. Subsequent cyclization and deprotection of this molecule provided a mixture of regioisomeric cyanobenzothiazoles (**5.14a/5.14b**). The regioisomers were inseparable at this stage, but could be isolated via HPLC following cysteine condensation. The pyrrolo[3,2-*f*] luciferin regioisomer (**5.6**) was characterized and evaluated for light emission with recombinant Fluc. PyrroLuc was found to be a poor light emitter (Fig. 5-2). We did not use PyrroLuc for screening, but moved forward with the other electronically modified luciferins. PyrroLuc could potentially serve as a specific luciferase inhibitor in future studies, or derivatization of PyrroLuc could yield an active luciferin with a unique structure.

5.3 Identification and characterization of a selective luciferase-luciferin pair

The construction of eight electronically modified analogues increased the total number of luciferins in our analysis to 20. These analogues were screened against the same panel of 207 mutant luciferases, resulting in a possible 3.6 million orthogonal sets. Out of these, the top 1,000 pairs had an average orthogonality score of 50, which represents an approximate two-fold increase from the data set described in Chapter 4. One of the most interesting luciferase-luciferin sets comprised 7'-PyridoneLuc (5.7) and 4'-MeOHLuc (5.15) (Fig. 5-3A), positively paired with enzymes 146 and 50, respectively. Enzyme 50 is a luciferase mutant that demonstrated an impressive 1,000 fold preference for 4'-MeOHLuc over 7'-PyridoneLuc. The 7'-Pyridone analogue displayed a drastic reduction in light emission with enzyme 50 while 4'-MeOHLuc essentially maintained wild-type levels (Fig. 5-3B). In the other direction, enzyme 146 demonstrated a slight switched preference for 7'-PyridoneLuc over 4'-MeOHLuc. Since the desired outcome was achieved halfway, we reasoned that evolving mutant 146 for more selective processing of 7'-PyridoneLuc was a good strategy.

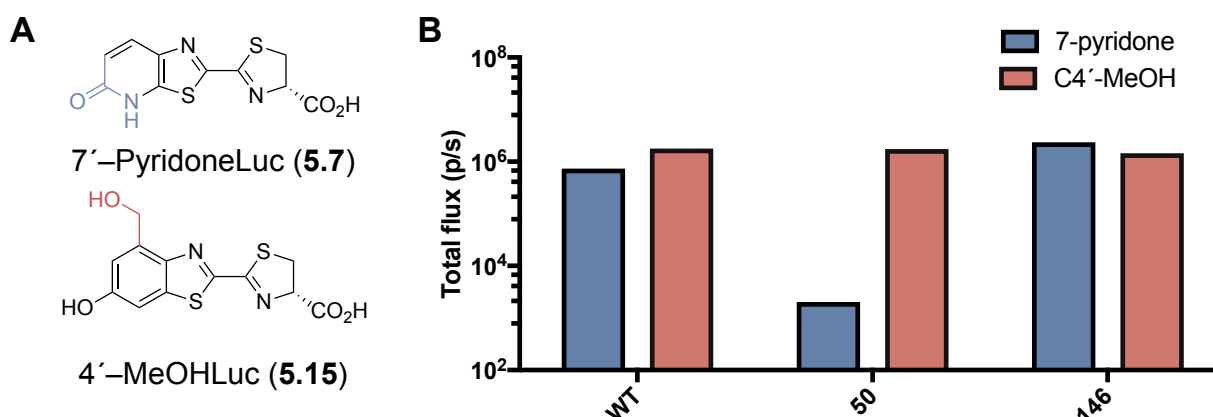


Figure 5-3. (A) Structures of luciferin analogues discussed in this chapter. (B) Mutant luciferases work selectively with luciferin analogues. Total photon output shown as compared to native luciferase (WT) in bacterial cell lysate. All compounds were imaged at 100 μ M concentrations.

Enzyme 50 is an interesting mutant as it contained a single mutation of serine to glycine at residue 347. This residue is known to be important for substrate binding and forms a hydrogen bond with the native substrate (D-luciferin) [11,21]. To further confirm its remarkable substrate selectivity, the enzyme was purified and light emission was assessed using various concentrations of 7'-PyridoneLuc, 4'-MeOHLuc (Fig. 5-4A-B), and D-luciferin. As shown in Figure 5.4, the selectivity of the S347G mutant for 4'-MeOHLuc was maintained in these assays. The substrate preference ranged between ~200-1,000 fold, depending on the concentration of luciferin (Fig. 5-4B). D-Luciferin also demonstrated a significant reduction in light emission with the S347G mutant, but not as drastic as 7'-PyridoneLuc (Fig. 5-4C). The hydrogen bond between 7'-PyridoneLuc and Ser347 is likely important for light emission and/or substrate binding particularly to the pyridone luciferin.

Since the S347G mutant with 4'-MeOHLuc maintains light emission similar to Fluc, we hypothesize that the hydroxymethyl group compensates for the hydrogen bond function of Ser347. Based on the crystal structure of Fluc, the hydroxyl group of 4'-MeOHLuc is in an ideal position to fill the void left behind by the S347G mutation (Fig. 5-5B) [22]. This appendage could increase binding affinity to the luciferase active site, or could replace a hydrogen bond contact to the benzothiazole nitrogen necessary for robust light emission [13,21,23]. Efforts to further elucidate the role of this mutation are underway. Others in the lab are attempting to crystallize this mutant with 7'-PyridoneLuc or 4'-MeOHLuc bound to WT Fluc and the mutated enzyme. Interestingly, a S347A mutation has been shown to be important for selective light emission with cyclic amino luciferins over the native D-luciferin [11,13].

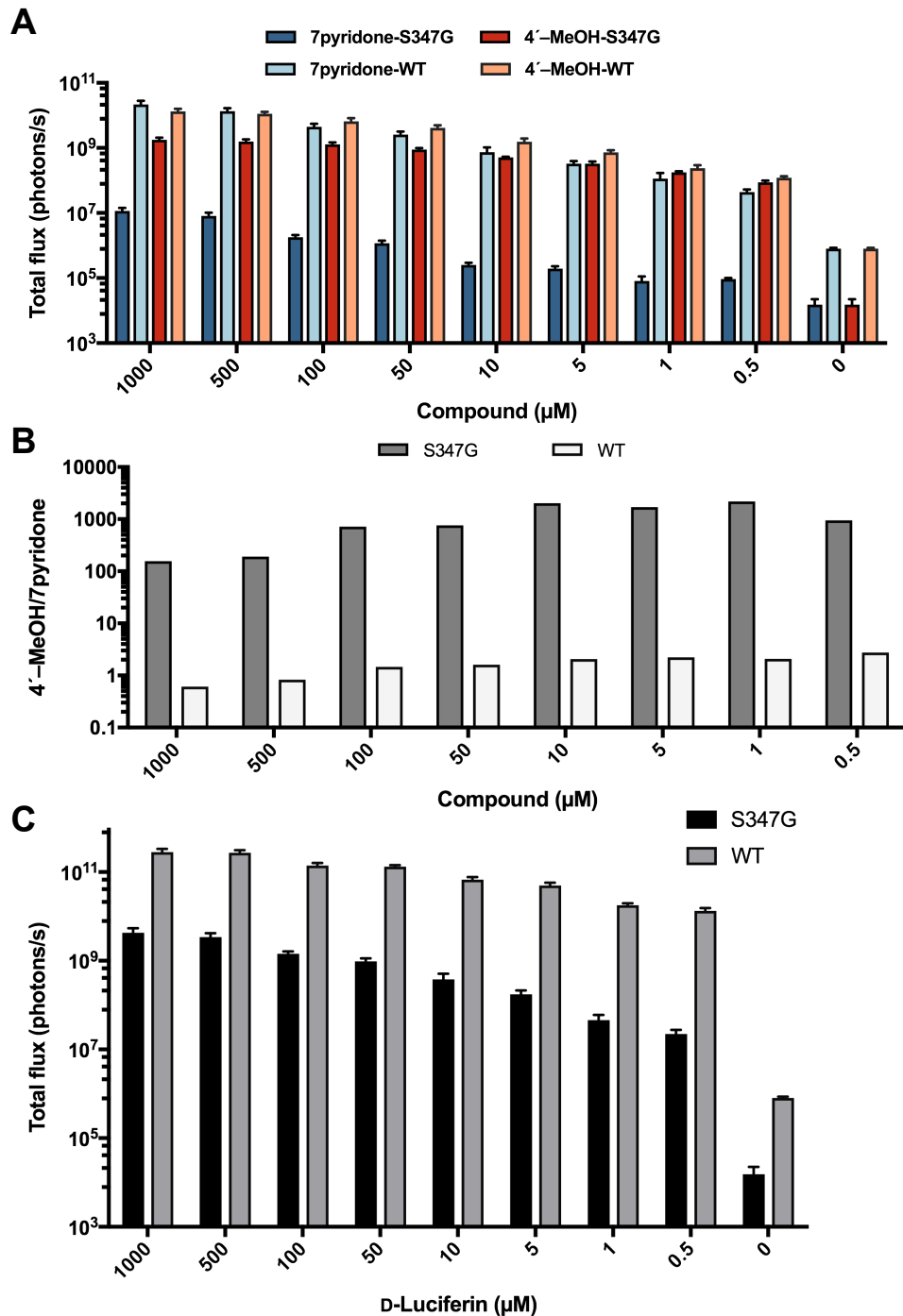


Figure 5-4. Purified enzyme light emission with luciferin analogues. (A) Total flux of 7'-PyridoneLuc with WT Fluc (light blue), the S347G mutant (dark blue) and 4'-MeOHLuc with WT Fluc (light red), and S347G (dark red). (B) Fold light emission of 4'-MeOHLuc/7'-PyridoneLuc with the S347G (gray) and WT Fluc (white). (C) Light emission of WT and S347G luciferase enzymes with D-luciferin. Error bars represent standard error of the mean for three individual experiments.

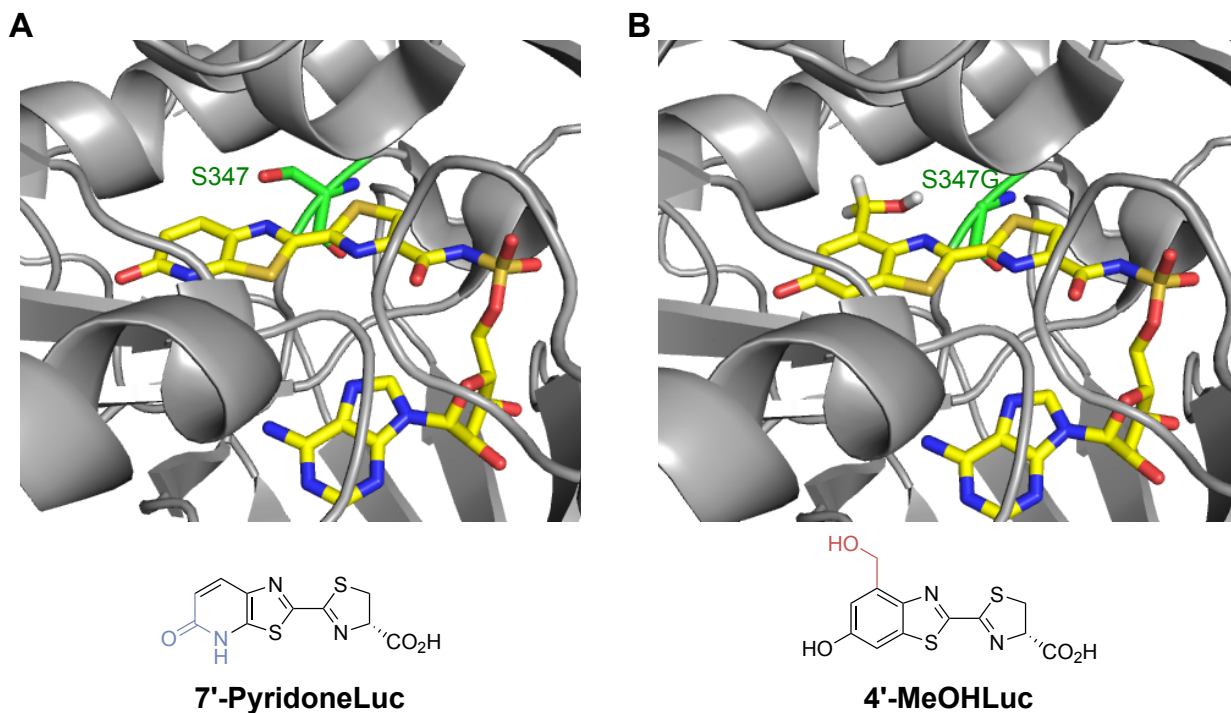


Figure 5-5. (A) 7'-PyridoneLuc depicted in the active site of Fluc. S347 shown in green is in close proximity to the benzothiazole nitrogen of the luciferin analogue. (B) 4'-MeOHLuc is depicted in the active site of luciferase mutant S347G. The hydroxyl group is shown potentially replacing the void left by the mutation to glycine. Images generated by Pymol's mutagenesis function and adapted from crystal structure published by Sundlov, J.A. *et al.* [22].

5.4 Directed evolution of a 7'-PyridoneLuc selective luciferase

The S347G mutation provided remarkable selectivity for 4'-MeOHLuc over 7'-PyridoneLuc. The mutant luciferase identified with the opposite substrate preference (Fig. 5-3B), enzyme 146, provided a starting point for evolving an enzyme that would have similar selectivity. Enzyme 146 comprised 4 mutations (V241A, F247V, S314T, G316T), with no obvious steric or electronic interactions with the analogue. These were located mainly outside of the binding pocket, suggesting that they most likely restructure the active site to favor 7'-PyridoneLuc [24-25]. New site directed

mutations were added to this quadruple mutant and the resulting library was screened for increased selectivity with 7'-PyridoneLuc.

Toward this end residues 284, 311, and 337 were targeted for the library using 22c-trick [26] or NDT codons. Screening identified an additional mutation that reproducibly improved selectivity for 7'-PyridoneLuc (Fig 5-6). The mutation comprised a glutamate to cysteine at residue 311 creating mutant G2-75. Residue 311 is known to be important for a hydrogen bond network in the luciferase active site [27]. However, it is difficult to speculate on specific interactions impacted by the mutation. While the E311C mutant exhibited increased preference for 7'-PyridoneLuc, it also provided decreased total light emission as compared to enzyme 146 (Fig. 5-6). To look for further improvements, the five mutations in mutant G2-75 were shuffled to eliminate any unnecessary mutations. These same five mutations were also shuffled with another 5 mutations (M249F, I257F, F195L, A326V, P334S) found in an enzyme exhibiting

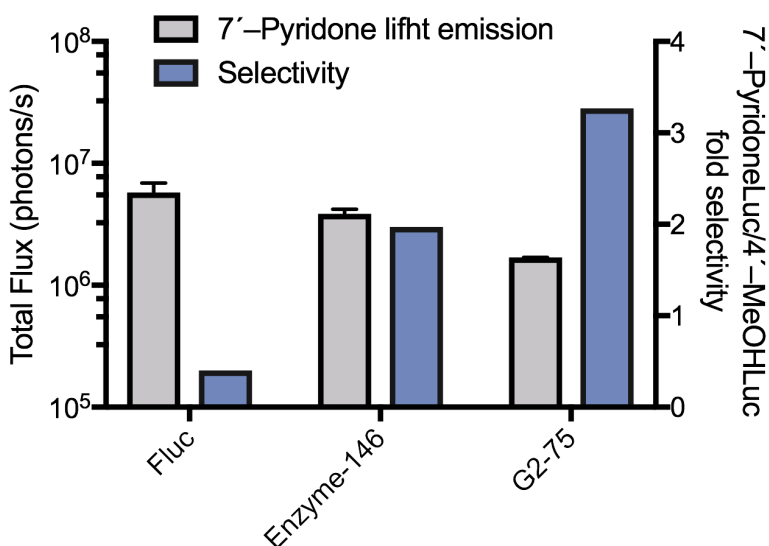


Figure 5-6. Mutant G2-75 increases selectivity for 7'-PyridoneLuc (right y-axis, blue bars), but decreases light emission (left y-axis, gray bars). Error bars represent SEM and data are representative of $n \geq 3$ experiments.

increased light emission with 7'-PyridoneLuc (identified previously in our lab). Unfortunately, no combination of mutations with improved light emission and retention of increased selectivity were found (Fig. 5-7). Additional libraries targeting residues 255-257, 337, and 347 were then screened. Unfortunately, no significant improvements in selectivity were identified (Fig. 5-7).

We next aimed to use the analyses of positive and negative pairings shown in Chapter 4 (Fig. 4-12). Narrowing the results to residues that positively pair with 7'-PyridoneLuc and negatively pair with 4'-MeOHLuc provided the results shown in Figure 5-8. The residues with the most interesting pairing pattern are 286-288 (black arrows). These positions are known to modulate binding and emission wavelength due to

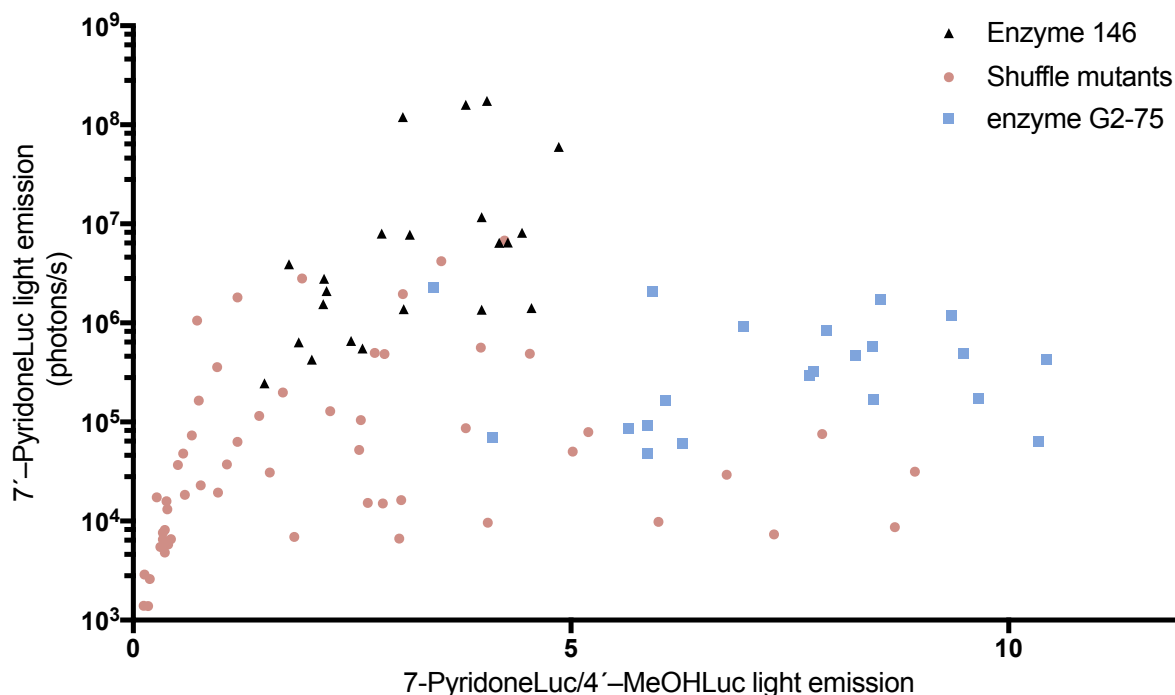


Figure 5-7. Distribution of library hits for the shuffle library. Black triangles represent replicates of enzyme 146. Blue squares are replicates of G2-75 mutant and red circles are individual shuffle mutants screened in lysate. All mutants emit less light (y-axis) than the original G2-75 or are less selective (x-axis).

increased or decreased packing of the luciferin in the active site [23]. In addition, residue 311 has been shown to modulate the positioning of residues 286-288 making it plausible that mutations at these sites could be synergistic with the E311C mutation found in generation two.

Following this analysis a site directed library targeting these residues was assembled with 22c-trick codons at site 286, along with NDT codons at sites 287 and 288. These sites were selected based on homology analysis [28]. Screening of the assembled library resulted in two interesting mutants (G3-48 and G3-87) that increase light emission while retaining or improving the selectivity shown by mutant G2-75. Interestingly, G3-48 lost the E311C mutation found in G2-75 but included a single mutation from leucine to valine at residue 286 (Fig. 5-9A). Mutant G3-87 contained the E311C mutation and two additional mutations at 287 (Leu to Cys) and 288 (Val to His) (Fig 5-9B). No steric or electronic interactions are apparent in either mutant, but the non-conservative mutations L287C and V288H in mutant G3-87 are notable. Light

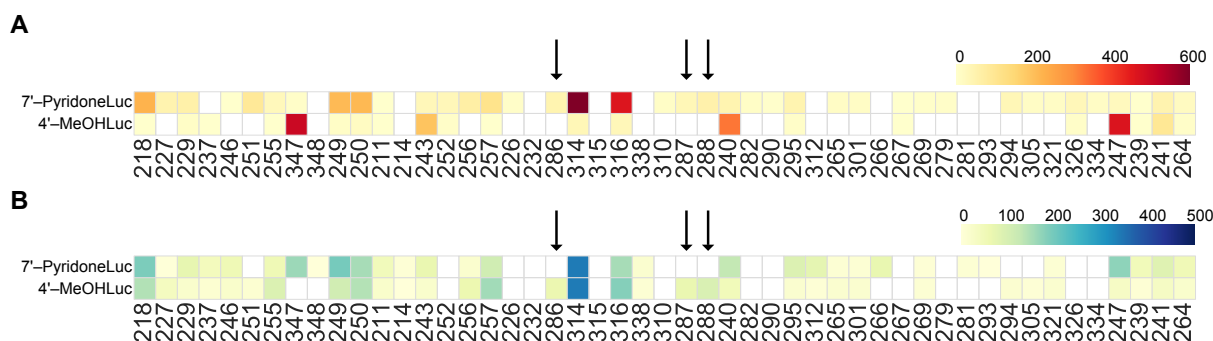


Figure 5-8. The top 5,000 luciferase-luciferin sets out of 3.6 million total were analyzed for their residue-compound interactions. The heat map is corresponds to the frequency of mutations at each residue that positively (A) or negatively (B) pair with 7'-PyridoneLuc or 4'-MeOHLuc. Arrows point out residues 286-288 that show more positive pairings with 7'-PyridoneLuc and negative pairings with 4'-MeOHLuc.

emission trends of 7'-PyridoneLuc and 4'-MeOHLuc were reproduced with these mutants in bacterial lysate. Collectively an 8 fold improvement in light emission (G3-48) was achieved along with a 14 fold improvement in selectivity over 3 rounds of mutagenesis (Fig. 5-10). Additional residues in and around the active site are currently being targeted along with random mutagenesis in hopes of further improving selectivity.

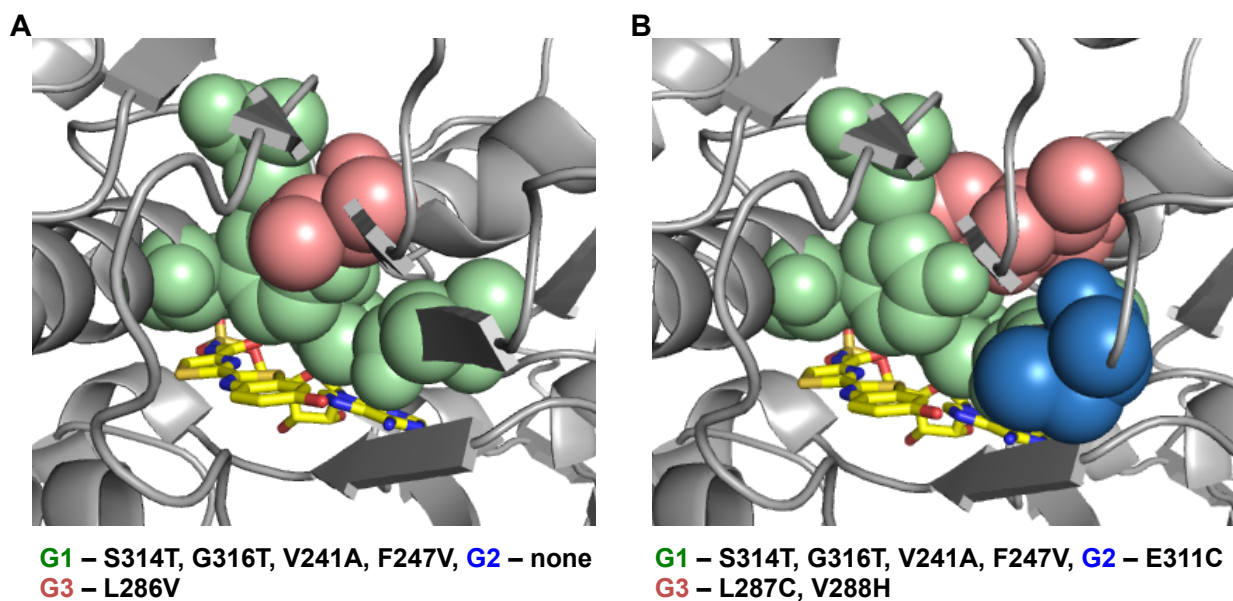


Figure 5-9. Mutations found in the generation 3 enzymes, G3-48 (A) and G3-87 (B), with improved selectivity and brightness. Spheres represent residues that are mutated from WT Fluc. Green spheres are mutations found in the lead enzyme from the cross compare algorithm. Blue spheres represent mutations found in the second generation. Mutations found in the third generation are shown in red. Pymol images adapted from crystal structure in Sundlov *et. al* [22].

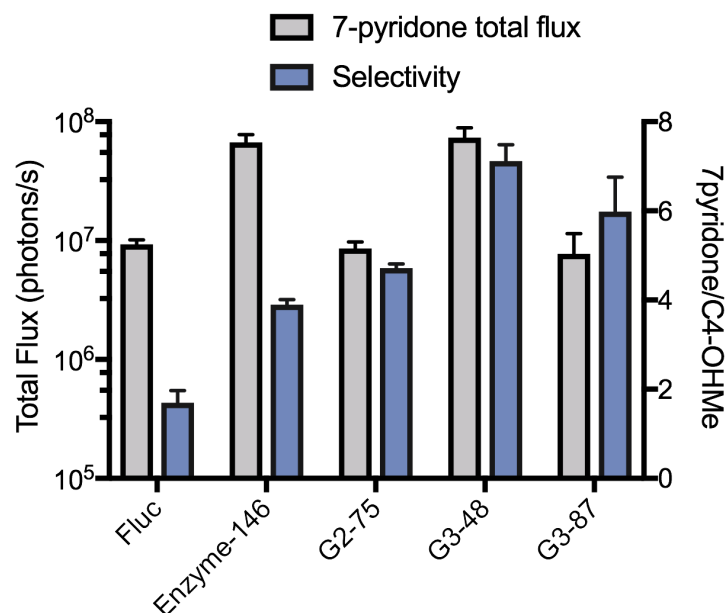


Figure 5-10. Mutants G3-48 and G3-87 maintain or increase selectivity for 7'-PyridoneLuc (right y-axis, blue bars), while maintaining or increasing light emission (left y-axis, gray bars). Error bars represent SEM and data are representative of $n \geq 3$ experiments.

5.5 Conclusions

As demonstrated in Chapters 3 and 4, bright and selective luciferase-luciferin pairs are necessary to enable truly multicomponent bioluminescent imaging. Building libraries of luciferin analogues and diverse luciferase enzymes is an excellent way to identify orthogonal sets of luciferin-luciferase pairs. However, to be more selective and brighter, further evolution of the luciferase enzymes is necessary. In this chapter I have shown that iterative mutagenesis and screening can lead to improved selectivity and light emission with the desired compounds. While improvement was made, a mutant with more selectivity for 7'-PyridoneLuc is desired. In the future, increasing the number of libraries screened, using computationally guided libraries or increasing the luciferase diversity through phylogenetic shuffling are all promising strategies to improve the desired enzymatic activity.

5.6 Materials and methods

5.6a General bioluminescence imaging

All analyses were performed in black 96-well plates (Grenier Bio One). Plates containing luminescent reagents were imaged in a light-proof chamber with an IVIS Lumina (Xenogen) CCD camera chilled to -90 °C. The stage was kept at 37 °C during the imaging session, and the camera was controlled using Living Image software. For assays with bacterial cell lysates, the exposure times for ranged from 1 s to 5 min, with data binning levels set to small or medium. Regions of interest were selected for quantification and total flux values were analyzed using Living Image software.

5.6b Bacterial lysate imaging

See Jones *et al.* [2] for procedure and section 3.8i in Chapter 3.

5.6c Recombinant protein expression and purification

The pET-luciferase plasmids (WT, 347G mutant) containing an N-terminal His₆ tag were transformed into chemically competent *E. coli* BL21 cells. Starter cultures (50 mL) in LB broth (containing 40 µg/mL Kan) were inoculated and grown for 12 h at 37 °C. Overnight cultures (15-20 mL) were transferred to 1 L of LB broth (containing 40 µg/L Kan) and grown in 2 L flasks at 37 °C to mid-log phase (OD₆₀₀ = 0.8 - 1.0), induced with 0.5 mM IPTG, and incubated at 22 °C for 16 – 18 h. The cells were harvested by centrifugation at 4 °C and resuspended in 20 mL of a solution of 50 mM NaPO₄, 300 mM NaCl, 1 mM DTT, and 1 mM PMSF at pH 7.4. Lysozyme (1 mg) was added, and the cells were

sonicated and centrifuged at 10000 x *g* for 1 h at 4 °C. WT Fluc and mutant S347G were purified from clarified supernatants using nickel affinity chromatography (BioLogic Duo Flow Chromatography System, Bio-Rad). Proteins were dialyzed into 25 mM Tris-acetate (pH 7.8) buffer 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 using UV spectroscopy. SDS-PAGE was also performed, and gels were stained with Coomassie R-250. Yield of the S347G mutant was 15 mg/L.

5-6d Light emission assays with recombinant luciferase

Bioluminescence assays with all luciferin compounds were carried out in triplicate, using solid black, flat-bottom, 96-well plates (BD Bioscience). Assay wells contained purified Fluc (1-2 µg), luciferin substrate (0-1 mM), ATP (Sigma, 0-1 mM), and reaction buffer (20 mM Tris-HCl, 0.5 mg/mL BSA, 0.1 mM EDTA, 1 mM TCEP, 2 mM MgSO₄, pH 7.6), totaling 100 µL. For pH studies, the buffer comprised 20 mM BIS-TRIS propane (with 100 µM MgCl₂, 1 mM ATP, 500 µM coenzyme A). All non-enzyme assay components were premixed in the wells prior to wild-type or mutant luciferase addition. Images for all assays were acquired as described above and analyzed with Microsoft Excel or GraphPad Prism (version 6.0f for Macintosh, GraphPad Software).

5.6e Construction of luciferase libraries

Two sections of the luciferase gene (*pgl4-luc2*), denoted R1 and R2, were targeted for gene assembly as described in Jones *et al.* [2]. Other libraries were created using

standard quick-change pcr techniques. All PCR reactions were run using Q5 Hot-start DNA polymerase (New England BioLabs).

Linearized vector was made from pET28-R1del or pET28-R2del as described in Jones *et al.* [2]. Insert containing the library was assembled with the linearized pET vector using circular polymerase extension cloning (CPEC) as described in Jones *et al.* [2] or by Gibson assembly. For Gibson assembly 50 ng of *DpnI* digested, linearized vector and insert at 2:1 to 8:1 (insert:vector) were added to 10 μ L of Gibson assembly master mix (homemade mix following the Prather recipe on http://www.openwetware.org/wiki/Gibson_Assembly with all enzymes and components purchased from New England BioLabs) and incubated at 50 °C for 20-60 min. The reaction mixture (1-3 μ L) was then transformed into chemically competent Top10 or DH5 α *E. coli*, cells were plated until the number of colonies was equal to or greater than 3 x the library size. Cells were then scraped off of the plates, combined, pelleted and minipreped. The minipreped plasmid DNA was saved for agar-plate screening (below).

5.6f Primer lists

All primers were purchased from Integrated DNA Technologies, Inc. (San Diego, CA) and are written in the 5'-3' direction. Upper case letters denote bases coding for the luciferase gene. Lower case letters denote bases added to ensure similar melting temperatures (T_m) subscript is off for all primers. Bases highlighted in red denote sites targeted for saturation mutagenesis.

Residue 284, 311, 337 library		
Forward primers	Residue targeted	
G2-146-R2-F36	284	AGACTATAAGATTCA NDT GCCCTGCTGGTGCCACAC
G2-146-R2-F117-1	311	CTAAGCAACTTGCAC NDT ATCGCCACCGGCACC
G2-146-R2-F117-2	311	CTAAGCAACTTGCAC VHG ATCGCCACCGGCACC
G2-146-R2-F117-3	311	CTAAGCAACTTGCAC TGG ATCGCCACCGGCACC
G2-146-R2-F337	337	GCATC NDT CAGGGCTACG
Reverse primers		
G2-146-R2-R135-1	311	TTGCTGAGCGGCGCGGTGCCGGTGGCGAT AHN GTG
G2-146-R2-R135-2	311	TTGCTGAGCGGCGCGGTGCCGGTGGCGAT CDB GTG
G2-146-R2-R135-3	311	TTGCTGAGCGGCGCGGTGCCGGTGGCGAT CCA GTG
G2-146-R337	337	CGTAGCCCTG AHN GATGC
G2-146-R2-R19	284	GGGC AHN TTGAATCTTATAGTCTTGCAAGCTtCGCAAGAA

G3-shuffle primers		
Forward primers	Residue targeted	
G3-146-sh-R1-F106	241	CCGCTATCCTCAGCGTG GYC CCATTTACCACGGC
G3-146-sh-R1-F141-1	247-249-257	GARGGCATG TTACCACGCTGGGCTACTTG WTT TGC G
G3-146-sh-R1-F141-2	247-249-257	GARGGCTTC TTACCACGCTGGGCTACTTG WTT TGC G
G3-146-sh-R2-F74	295	TATTTAGCTTC YTC GCTAAGAGCACTCTCATCGACAA GTACGAC
G3-146-sh-R2-F117-1	311-314-316	CTAAGCAACTTGCAC GAA ATCGCC ACCGGCACC
G3-146-sh-R2-F117-2	311-314-316	CTAAGCAACTTGCAC TGC ATCGCC ACCGGCACC
G3-146-sh-R2-F150	326	GCGCCGCTCAGCAAGGAGGTAGGTGAG GYC
G3-146-sh-R2-F181	334	GTGGCCAAACGCTTCCACCTA YCT GGCATCCG
Reverse primers		
G3-146-sh-R1-R87	241	CCACGCTGAGGATAGCGGTGTCGGGGATGATCTGGT T
G3-146-sh-R1-R124-1	241-247-249	CGTGGTGAAC CAT GCCG GAR CCGTGGTGAATGG GR
G3-146-sh-R1-R124-2	241-247-249	CGTGGTGAAG GAA GCCG GAR CCGTGGTGAATGG GR
G3-146-sh-R1-R159	257	AGCACGACCCGAAAGCCGCA AAW CAAGTAGCCAG GCTCTTAGC GAR GAAAGCTAAATAGTGTGGGCACCAG CA
G3-146-sh-R2-R59	295	TTC GTGCAAGTTGCTTAGGTCGTA CTT GT CGAT GAGA GT
G3-146-sh-R2-R97-1	311	GCA GTGCAAGTTGCTTAGGTCGTA CTT GT CGAT GAGA GT
G3-146-sh-R2-R135	314-316	TTGCTGAGCGGCGC GGT GCC GGT GGCGAT
G3-146-sh-R2-R165	326	GGAAGCGTTTGGCCAC GR CTCACCTACCTCC
G3-146-sh-R2-R197	334	AGGCCGTAGCCCTGGCGGATGCC AGR TAGGT

G3-286-288 library		
Forward primers	Residue targeted	
G3-146-286-F-1	286-287-288	TGCAAGACTATAAGATTCAATCTGCC NDTNDT NDTCC CACACTATTTAGCTTCTTCG
G3-146-286-F-2	286-287-288	TGCAAGACTATAAGATTCAATCTGCC VHGNDTNDT CC CACACTATTTAGCTTCTTCG
G3-146-286-F-3	286-287-288	TGCAAGACTATAAGATTCAATCTGCC TGGNDTNDT CC CACACTATTTAGCTTCTTCG
Reverse primers		
G3-146-286-R-1	286-287-288	CGAAGAAGCTAAATAGTGTGGG AHNAHNAH GGCAG ATTGAATCTTATAGTCTTGCA
G3-146-286-R-2	286-287-288	CGAAGAAGCTAAATAGTGTGGG AHNAHNCDB GGCAG ATTGAATCTTATAGTCTTGCA
G3-146-286-R-3	286-287-288	CGAAGAAGCTAAATAGTGTGGG AHNAHNCCA GGCAG ATTGAATCTTATAGTCTTGCA

5.6g Screening protocols

Agar plate screening protocol

Plates were screened as described in Jones *et al.* [2]. See chapter 3.8h for protocol. 7'–
PyridoneLuc was screened on agar plate at a final concentration of 200-400 μ M.

Lysate screening of library members

Following the protocol of Studier [29] autoinduction broth (Na_2HPO_4 (6 g/L), KH_2PO_4 (3 g/L), tryptone (20 g/L), yeast extract (5 g/L), NaCl (5 g/L)) was made before hand and stored at rt. Sugar stocks (60% w/v glycerol, 10% w/v glucose, and 8% w/v lactose) were made deionized water, filtered through 2 μ m filters and stored at 4 °C. Prior to use 50 mL aliquots of autoinduction media were prepared by adding premade sugar solutions (500 μ L glycerol, 250 μ L glucose, 1.25 mL lactose), and Kanamycin (to 40 μ g/mL) to 48 mL of premade autoinduction broth.

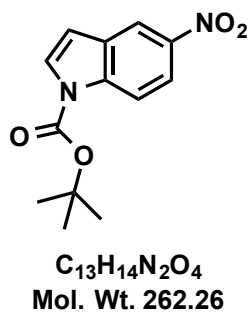
Light emitting colonies identified from the agar plate screen were picked immediately or stored at 4 °C until imaging was performed. These colonies or Fluc/lead mutant colonies were inoculated into 96-well deep-well plates containing 400 µL/well autoinduction media. Deep-well plates were covered, shaken at 225 rpm, 30 °C, for 24 h. After induction the plates were centrifuged (3400 rpm, 5 min) and supernatant was removed. Pellets were either stored at – 20 °C or imaged immediately. To each well 200 µL of lysis buffer (50 mM Tris•HCl, 500 mM NaCl, 0.5% (v/v) Tween, 5 mM MgCl₂, pH = 7.4) was added and the pellets were resuspended using a multichannel pipette. Next 90 µL of lysate from each well was transferred to a 96-well black plate. This step was repeated in order to have duplicate wells of the same lysate on the imaging plate for simultaneous imaging of two luciferin compounds. The appropriate luciferin stock solution (10 µL) was added to each well, 12 wells at a time (final [luciferin] = 100 µM and [ATP] = 1 mM). Images were acquired over 10 min as described above. Bioluminescence readings acquired 5 min post-luciferin addition were used for analyses. Data were analyzed using Microsoft Excel and/or GraphPad Prism.

5.6h General synthetic methods

All reagents purchased from commercial suppliers were of analytical grade and used without further purification. Appel's salt, 4,5-dichloro-1,2,3-dithiazolium chloride (**5.12**), was prepared according to literature precedent. (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

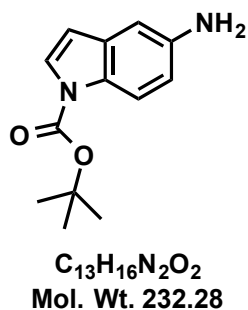
UV-Vis Detector. Semi-preparative runs were performed using an Agilent Prep-C18 Scalar column (9.4 x 150 mm, 5 μ m). Anhydrous solvents (acetonitrile (MeCN), dichloromethane (DCM), methanol (MeOH)) were dried by passage over neutral alumina, dimethyl formamide (DMF) was passed over activated molecular sieves and dimethyl sulfoxide (DMSO), which was purchased from Acros Organics in AcroSeal bottles. Reaction vessels were either flame or oven dried prior to use. NMR spectra were acquired with Bruker Advanced spectrometers. All spectra were acquired at 298 K. ^1H -NMR spectra were acquired at 400 MHz, and ^{13}C -NMR spectra were acquired at 125 MHz. For ^{13}C -NMR data obtained via distortionless enhancement by polarization transfer (DEPTQ). 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. The abbreviations used can be found in the document JOC Standard Abbreviations and Acronyms, <http://pubs.acs.org/paragonplus>.

5.6i Compound syntheses



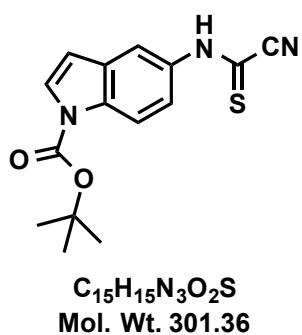
tert-butyl 5-Nitro-1H-indole-1-carboxylate (5.10) Following the procedure of Furuya *et al.* [30], 5-nitroindole **5.9** (1.00 g, 6.17 mmol) and DMAP (0.0754 g, 0.617 mmol) were added to a flame-dried round bottom flask and dissolved in 12 mL of anhydrous MeCN. Di-*tert*-butyl dicarbonate (1.7 mL, 7.4 mmol) was then added and the

mixture was stirred at room temperature for 30 min. The precipitate was collected by filtration and washed with chilled ethanol to yield **5.10** as a light brown solid (1.5 g, 95%). The NMR spectra were consistent with previous reports [31].



tert-butyl 5-Nitro-1H-indole-1-carboxylate (5.11). Following the procedure of Wissner *et al.* [20]; iron powder (0.145 g, 2.60 mmol), NH₄Cl (0.179 g, 3.34 mmol), and 1-Boc-5-nitroindole (**5.10**, 0.0973 g, 0.373 mmol) were added to a round bottom flask and suspended in 8:2:1 EtOH:H₂O:THF (5 mL). The suspension was then stirred

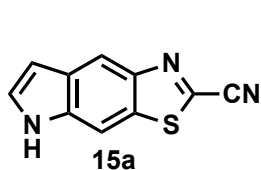
under N₂ at 70 °C for 40 min. The mixture was cooled and filtered through Celite. The filtrate was then extracted with EtOAc (2 x 20 mL). The combined extracts were washed with water (2 x 40 mL) and brine (2 x 40 mL), dried over MgSO₄, filtered and concentrated *in vacuo* to yield **5.11** as a dark brown oil (0.053 g, 63%). The NMR spectra were consistent with previous reports [31].



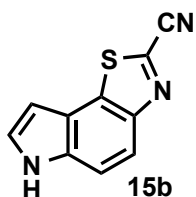
tert-butyl 5-((Cyanocarbonothioyl)amino)-1H-indole-1-carboxylate (5.13). 4,5-Dichloro-1,2,3-dithiazole (**5.12**, 0.614 g, 2.95 mmol) was added to a flame-dried flask and flushed with N₂. DCM (10 mL) was then added to the flask. In a separate flask, Boc-protected 5-aminoindole (**5.11**, 0.684 g, 2.95 mmol)

was dissolved in 11 mL and purged with N₂. This solution was then transferred to the solution of **5.12**, and the resulting red-brown mixture was stirred under N₂ for 30 min. The reaction mixture was cooled to 0 °C and DBU (1.3 mL, 8.8 mmol) was added

dropwise. After 1 h at 0 °C, two additional equivalents of DBU (0.88 mL, 5.9 mmol) were added and the reaction was warmed to room temperature and stirred for 1 h. The crude reaction mixture was concentrated *in vacuo* to remove most of the volatiles. The viscous dark brown residue was then added directly to a column of silica gel and eluted with DCM, followed by 4:1 DCM:MeOH. The relevant fractions were combined and concentrated *in vacuo* to afford **5.13** (0.59 g, 66%) as a yellow–brown solid. ¹H NMR (400 MHz, acetone-*d*₆, mixture of rotamers) δ 8.40 (d, *J* = 1.8, 0.75H), 8.26 (d, *J* = 8.5, 0.25H), 8.22 (d, *J* = 9, 0.75H), 7.83 (d, *J* = 2.2, 0.25H), 7.79 (d, *J* = 3.7, 0.25H), 7.75–7.72 (m, 1.5H), 7.48 (dd, *J* = 8.7, 2.2, 0.25H), 6.77 (d, *J* = 3.3, 0.25H), 6.74 (d, *J* = 3.7, 0.75H), 1.70 (app s, 9H); ¹³C (125 MHz, acetone-*d*₆) (mixture of rotamers, not all carbons distinctly observed) δ 167.2, 163.0, 150.0, 135.4, 134.7, 134.2, 133.7, 131.5, 128.9, 128.49, 120.9, 120.30, 117.3, 116.7, 116.3, 116.2, 114.7, 113.6, 108.2, 108.0, 85.2, 85.0, 28.16. HRMS (ESI–TOF) *m/z* calcd for C₁₅H₁₄N₃O₂S [M – H][–] 300.0807, found 300.0801.



C₁₀H₅N₃S
Mol. Wt. 199.23



6H-Thiazolo[5,4-*e*]indole-2-carbonitrile (5.14a)

and **5H-thiazolo[4,5-*f*]indole-2-carbonitrile**

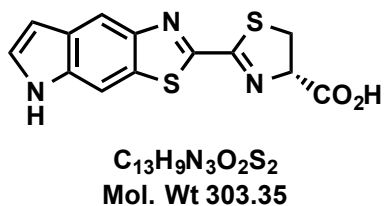
(5.14b). Palladium chloride (0.0110 g, 0.0624

mmol), CuI (0.0448 g, 0.310 mmol), TBAB (0.400

g, 1.24 mmol) and **5.13** were placed in a flame-

dried flask, flushed with N₂ and then suspended in 9.5 mL of anhydrous DMF:DMSO (1:1). The resulting mixture was stirred at 110 °C for 2.5 h under N₂. The reaction was then diluted with EtOAc and washed with H₂O (4 x 100 mL) and brine (2 x 100 mL). The

organic layer was then dried over MgSO₄, filtered and concentrated *in vacuo*. The crude material was purified by flash column chromatography, eluting with 8:2 to 1:1 hexanes:EtOAc to provide a 3:2 mixture of regioisomers (**5.14a**:**5.14b**) as a bright yellow solid (0.12 g, 54%). ¹H NMR (400 MHz, acetone-*d*₆) (mixture of isomers) δ 11.05 (br s, 0.7H), 10.70 (br s, 1H), 8.45 (s, 1H), 8.21 (s, 1.7H), 7.90 (d, *J* = 8.9, 0.7H), 7.62 (d, *J* = 8.9, 0.7H), 7.63-7.60 (m, 1.7H), 6.86 (s app, 0.7H), 6.76 (s app, 1H); ¹³C (125 MHz, DMSO-*d*₆) δ 146.8, 146.1, 137.5, 134.4, 132.5, 130.4, 129.9, 129.7, 129.5, 129.2, 128.1, 127.2, 119.4, 117.0, 115.2, 114.1, 114.0, 109.5, 103.6, 101.7; HRMS (ESI-TOF) *m/z* calcd for C₁₀H₄N₃S [M – H][–] 198.0126, found 198.0125.



(S)-2-(5H-Thiazolo[4,5-f]indol-2-yl)-4,5-

dihydrothiazole-4-carboxylic acid (5.6, PyrroLuc). A

mixture of **5.14a** and **5.14b** was suspended in 1.3 mL MeOH:acetonitrile (3:1) in a 20 mL vial. D-Cysteine hydrochloride monohydrate (0.0192 g, 0.122 mmol) and potassium carbonate (0.0167 g, 0.122 mmol) were dissolved in 0.20 mL water and added to the suspension. The reaction was stirred for 30 min under N₂. Upon consumption of **15** as observed by TLC, the solvent was removed *in vacuo* to yield a yellow solid. The solid was dissolved in PBS (1 mL) and purified by reversed phase semi-preparative HPLC, eluting with 18% MeOH in H₂O (at a flow rate of 3 mL/min). The desired fractions were pooled and concentrated to afford **5.16** as a bright yellow solid (0.018 g, 52%). ¹H NMR (400 MHz, CD₃OD) δ 8.26 (s, 1H), 7.94 (s, 1H), 7.41 (d, *J* = 3.2, 1H), 6.64 (d, *J* = 3.2, 1H), 5.20 (app t, *J* = 9.4, 1H), 3.77-3.69 (m, 2H); ¹³C (125 MHz, CD₃OD) δ 177.5, 175.6, 165.6,

159.6, 148.7, 138.5, 131.6, 130.8, 128.7, 115.5, 103.6, 102.8, 83.3, 37.3; HRMS (ESI-TOF) m/z calcd for $C_{13}H_8N_3O_2S_2$ $[M - H]^-$ 302.0058, found 302.0061.

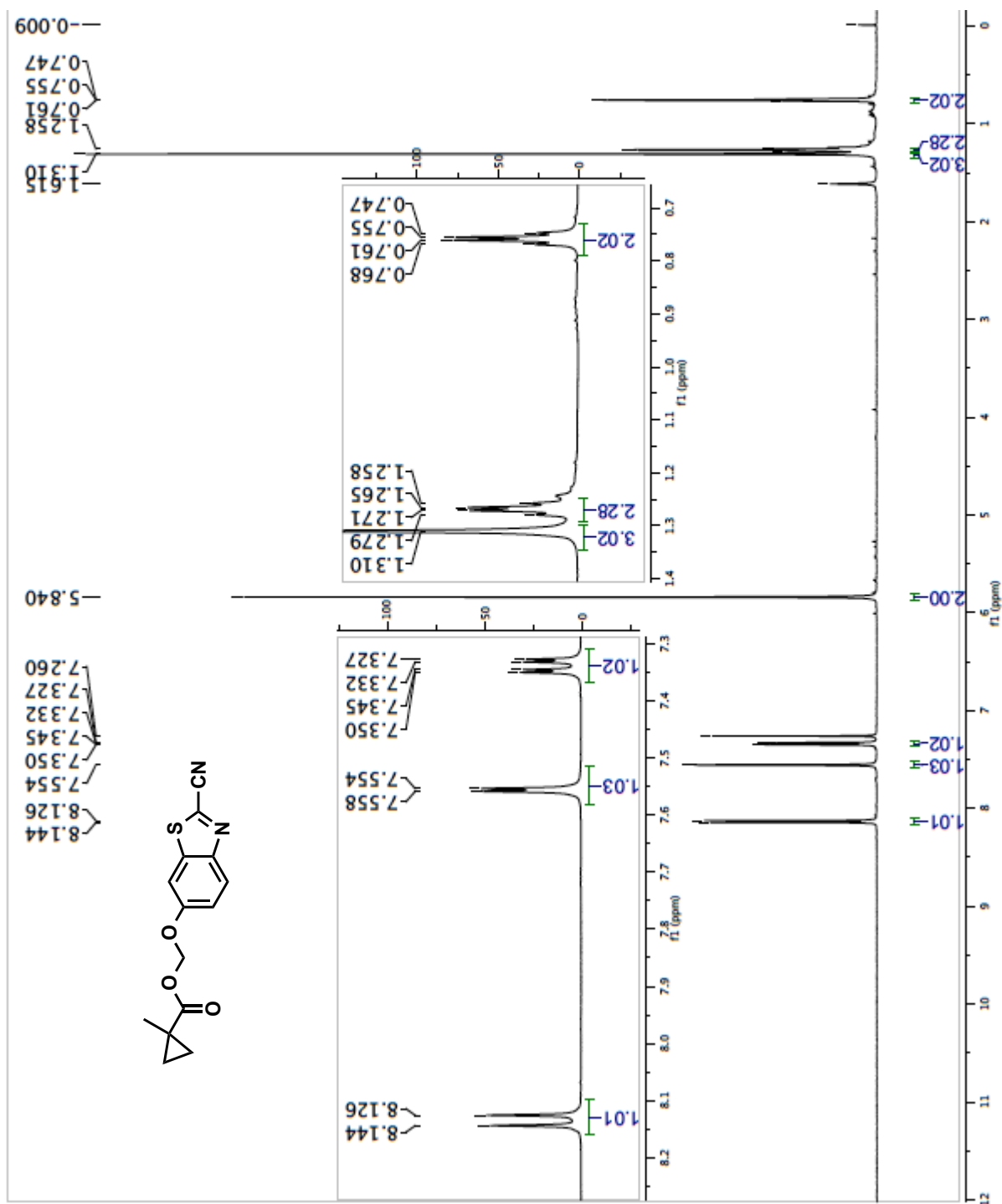
References

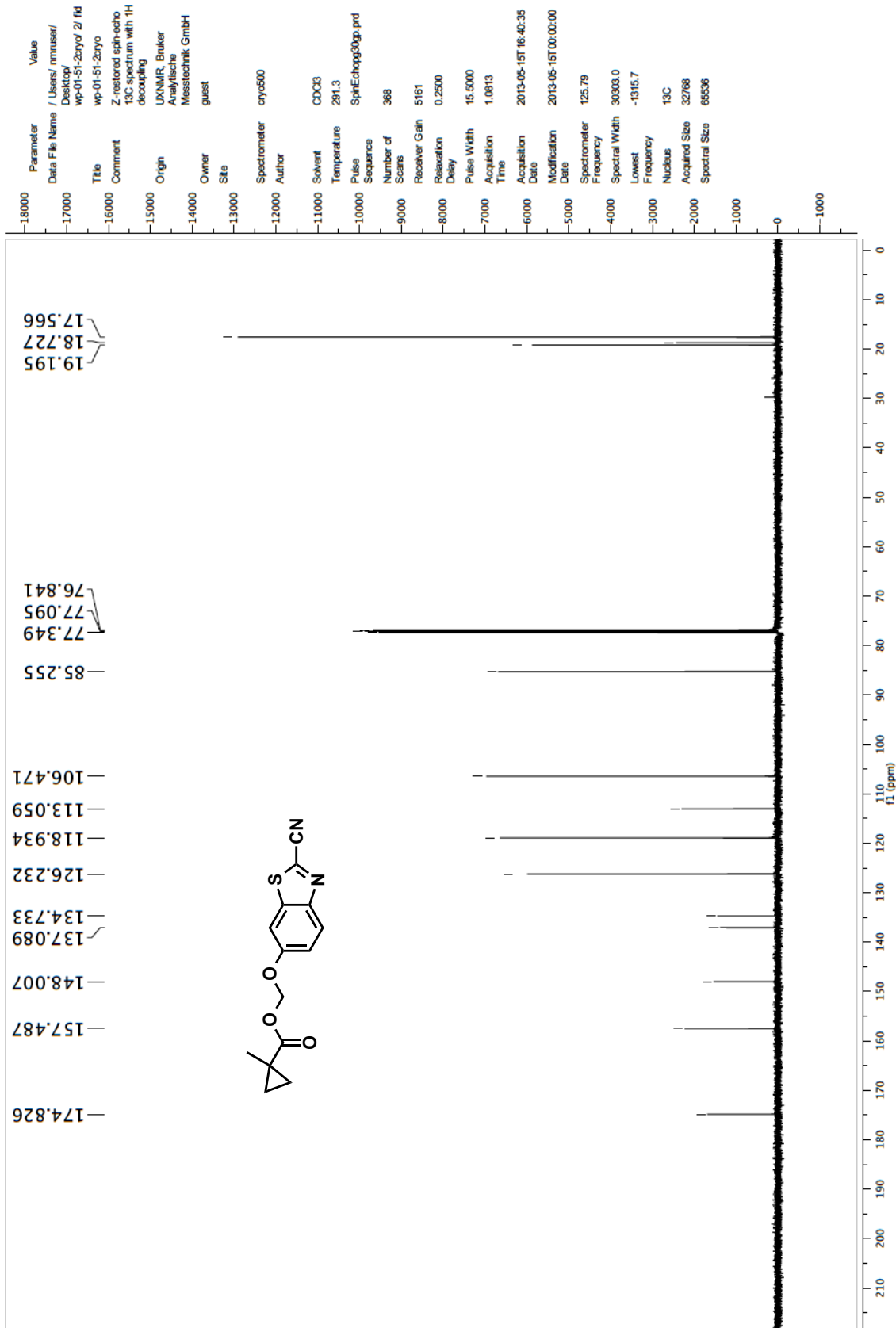
1. Steinhardt, R. C.; Rathbun, C. M.; Krull, B. T.; Yu, J.; Yang, Y.; Nguyen, B. D.; Kwon, J.; McCutcheon, D. C.; Jones, K. A.; Furche, F.; Prescher, J. A., Brominated luciferins are versatile bioluminescent probes. *ChemBioChem* **2017**, *18*, 96.
2. Jones, K. A.; Porterfield, W. B.; Rathbun, C. M.; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A., Orthogonal luciferase-luciferin pairs for bioluminescence imaging. *J. Am. Chem. Soc.* **2017**, *139*, 2351.
3. 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.
4. McCutcheon, D. C.; Porterfield, W. B.; Prescher, J. A., Rapid and scalable assembly of firefly luciferase substrates. *Org. Biomol. Chem.* **2015**, *13*, 2117.
5. White, E. H.; Worther, H.; Seliger, H. H.; Mcelroy, W. D., Amino analogs of firefly luciferin and biological activity thereof. *J. Am. Chem. Soc.* **1966**, *88*, 2015.
6. 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.
7. Branchini, B. R.; Hayward, M. M.; Bamford, S.; Brennan, P. M.; Lajiness, E. J., Naphthyl- and quinolyluciferin: green and red light emitting firefly luciferin analogs. *Photochem. Photobiol.* **1989**, *49*, 689.
8. 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.
9. 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.
10. 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.

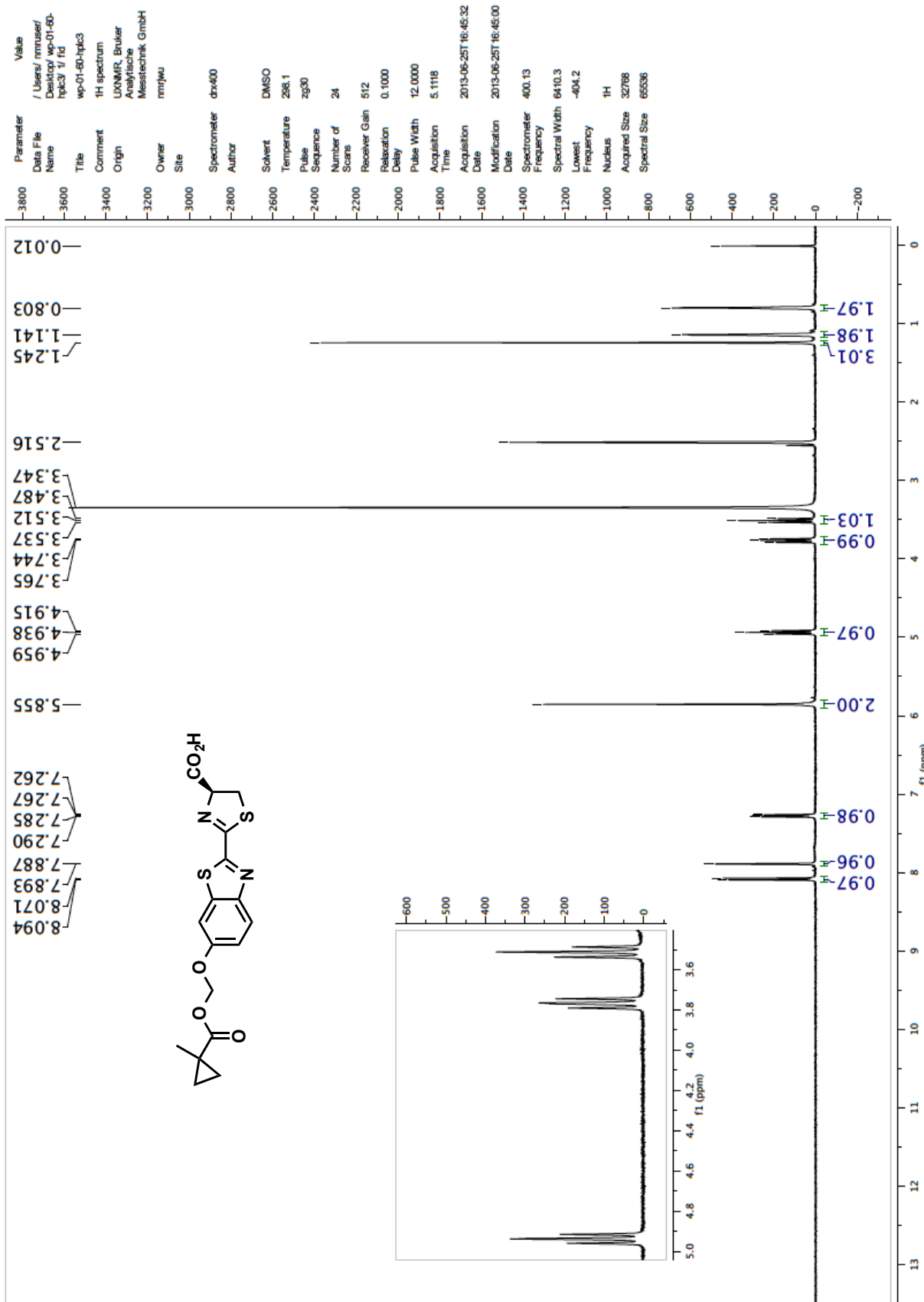
11. Mofford, D. M.; Reddy, G. R.; Miller, S. C., Aminoluciferins extend firefly luciferase bioluminescence into the near-infrared and can be preferred substrates over D-Luciferin. *J. Am. Chem. Soc.* **2014**, *136*, 13277.
12. Reddy, G. R.; Thompson, W. C.; Miller, S. C., Robust light emission from cyclic alkylaminoluciferin substrates for firefly luciferase. *J. Am. Chem. Soc.* **2010**, *132*, 13586.
13. Harwood, K. R.; Mofford, D. M.; Reddy, G. R.; Miller, S. C., Identification of mutant firefly luciferases that efficiently utilize aminoluciferins. *Chem. Biol.* **2011**, *18*, 1649.
14. 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.
15. Wu, D. L.; Liu, L.; Liu, G. F.; Jia, D. Z., Ab initio/DFT and AIM studies on dual hydrogen-bonded complexes of 2-hydroxypyridine/2-pyridone tautomerism. *J. Phys. Chem. A* **2007**, *111*, 5244.
16. Wong, M. W.; Wiberg, K. B.; Frisch, M. J., Solvent Effects 3. Tautomeric equilibria of formamide and 2-pyridone in the gas-phase and solution - an ab initio SCRF study. *J. Am. Chem. Soc.* **1992**, *114*, 1645.
17. 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.
18. Taber, D. F.; Tirunahari, P. K., Indole synthesis: a review and proposed classification. *Tetrahedron* **2011**, *67*, 7195.
19. Bandini, M.; Eichholzer, A., Catalytic Functionalization of indoles in a new dimension. *Angew. Chem. Int. Ed.* **2009**, *48*, 9608.
20. Wissner, A.; Overbeek, E.; Reich, M. F.; Floyd, M. B.; Johnson, B. D.; Mamuya, N.; Rosfjord, E. C.; Discafani, C.; Davis, R.; Shi, X. Q.; Rabindran, S. K.; Gruber, B. C.; Ye, F.; Hallett, W. A.; Nilakantan, R.; Shen, R.; Wang, Y. F.; Greenberger, L. M.; Tsou, H. R., Synthesis and structure-activity relationships of 6,7-disubstituted 4-anilinoquinoline-3-carbonitriles. The design of an orally active, irreversible inhibitor of the tyrosine kinase activity of the epidermal growth factor receptor (EGFR) and the human epidermal growth factor receptor-2 (HER-2). *J. Med. Chem.* **2003**, *46*, 49.
21. Viviani, V. R.; Amaral, D.; Prado, R.; Arnoldi, F. G. C., A new blue-shifted luciferase from the Brazilian *Amydetes fanestratus* (Coleoptera: Lampyridae)

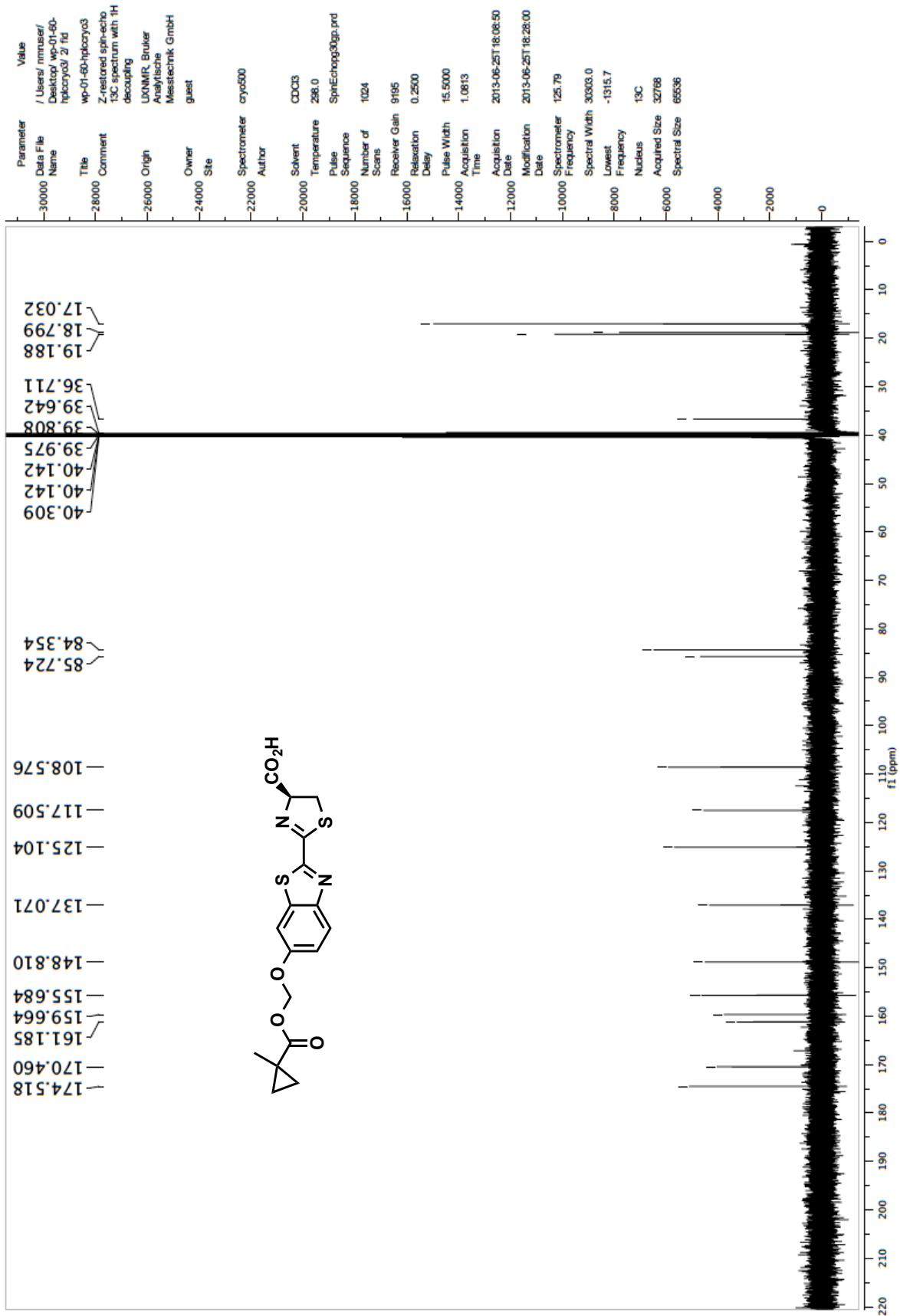
- firefly: molecular evolution and structural/functional properties. *Photochem. Photobio. Sci.* **2011**, *10*, 1879.
22. Sundlov, J. A.; Fontaine, D. M.; Southworth, T. L.; Branchini, B. R.; Gulick, A. M., Crystal structure of firefly luciferase in a second catalytic conformation supports a domain alternation mechanism. *Biochemistry* **2012**, *51*, 6493.
 23. 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.
 24. Tokuriki, N.; Jackson, C. J.; Afriat-Jurnou, L.; Wyganowski, K. T.; Tang, R. M.; Tawfik, D. S., Diminishing returns and tradeoffs constrain the laboratory optimization of an enzyme. *Nat. Commun.* **2012**, *3*, 1257.
 25. Campbell, E.; Kaltenbach, M.; Correy, G. J.; Carr, P. D.; Porebski, B. T.; Livingstone, E. K.; Afriat-Jurnou, L.; Buckle, A. M.; Weik, M.; Hollfelder, F.; Tokuriki, N.; Jackson, C. J., The role of protein dynamics in the evolution of new enzyme function. *Nat. Chem. Biol.* **2016**, *12*, 944.
 26. Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z. G.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P., Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. *ACS Synth. Biol.* **2013**, *2*, 83.
 27. Viviani, V. R.; Amaral, D. T.; Neves, D. R.; Simões, A.; Arnoldi, F. G. C., The luciferin binding site residues C/T311 (S314) influence the bioluminescence color of beetle luciferases through main-chain interaction with oxyluciferin phenolate. *Biochemistry* **2013**, *52*, 19.
 28. Reetz, M. T.; Wu, S., Greatly reduced amino acid alphabets in directed evolution: making the right choice for saturation mutagenesis at homologous enzyme positions. *Chem. Commun.* **2008**, 10.1039/b813388c, 5499.
 29. Studier, F. W., Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **2005**, *41*, 207.
 30. Furuya, T.; Strom, A. E.; Ritter, T., Silver-mediated fluorination of functionalized aryl stannanes. *J. Am. Chem. Soc.* **2009**, *131*, 1662.
 31. Chakrabarty, M.; Kundu, T.; Harigaya, Y., Mild deprotection of tert-butyl carbamates of NH-heteroarenes under basic conditions. *Synth. Commun.* **2006**, *36*, 2069.

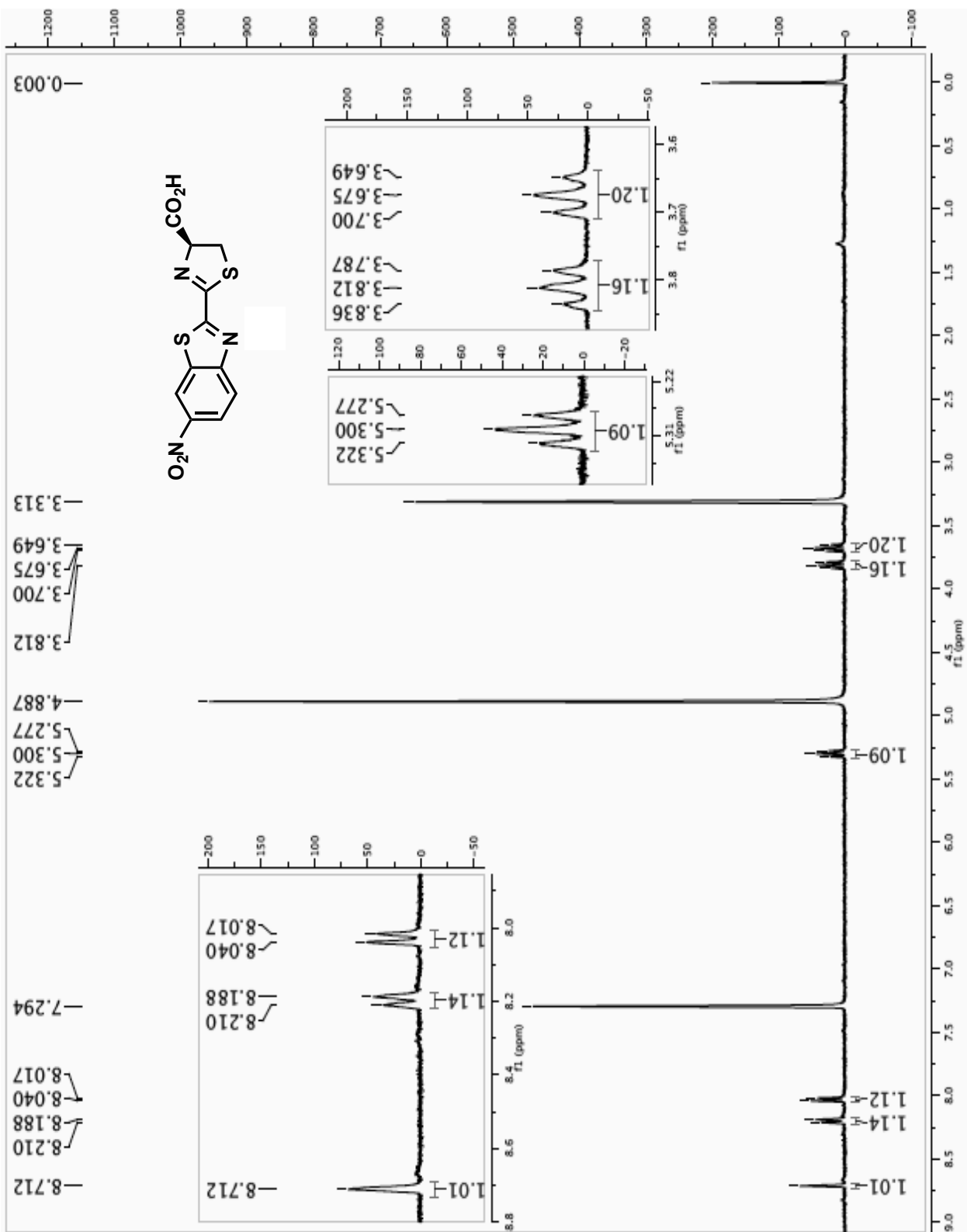
APPENDIX A: NMR spectra

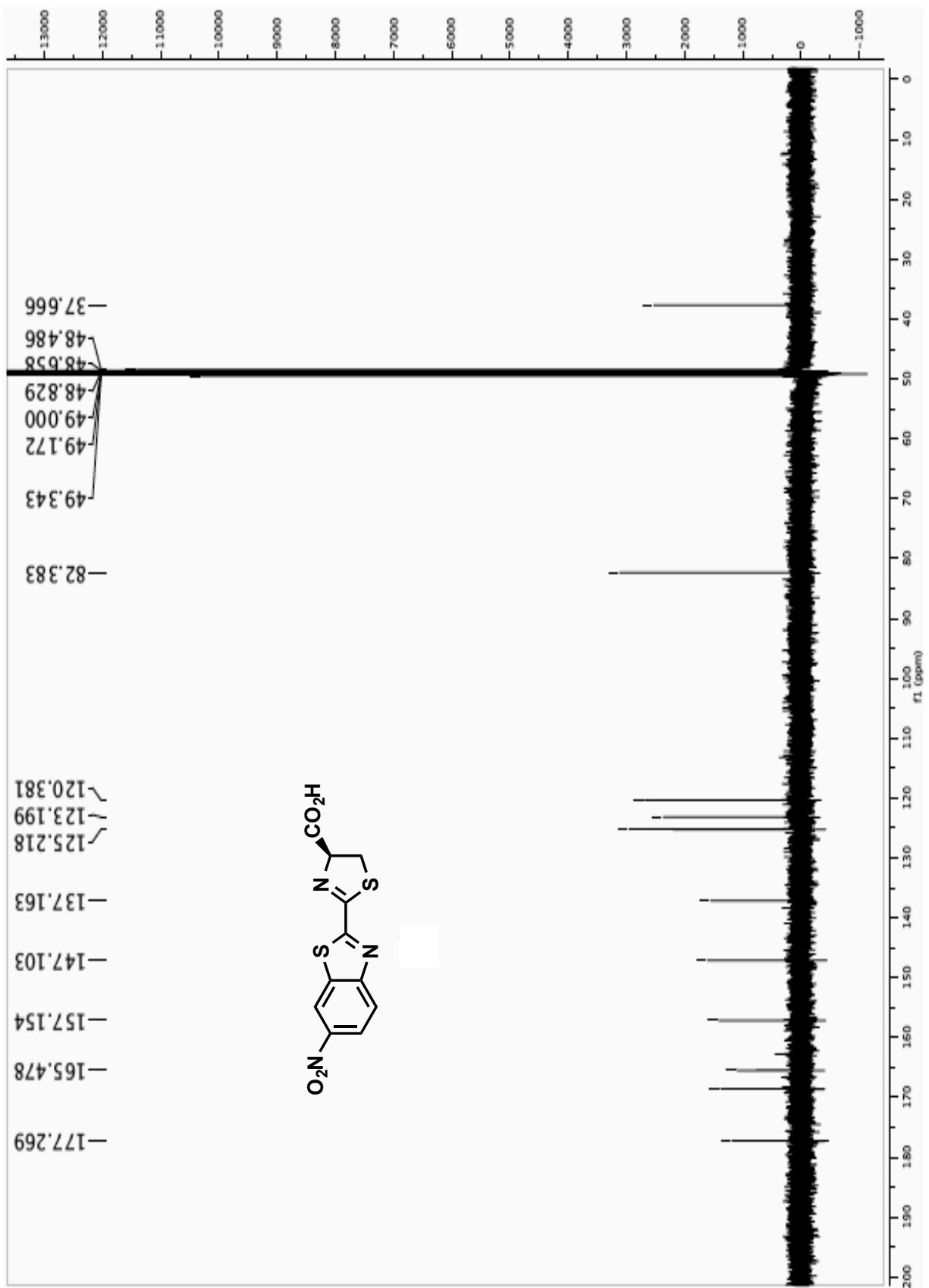


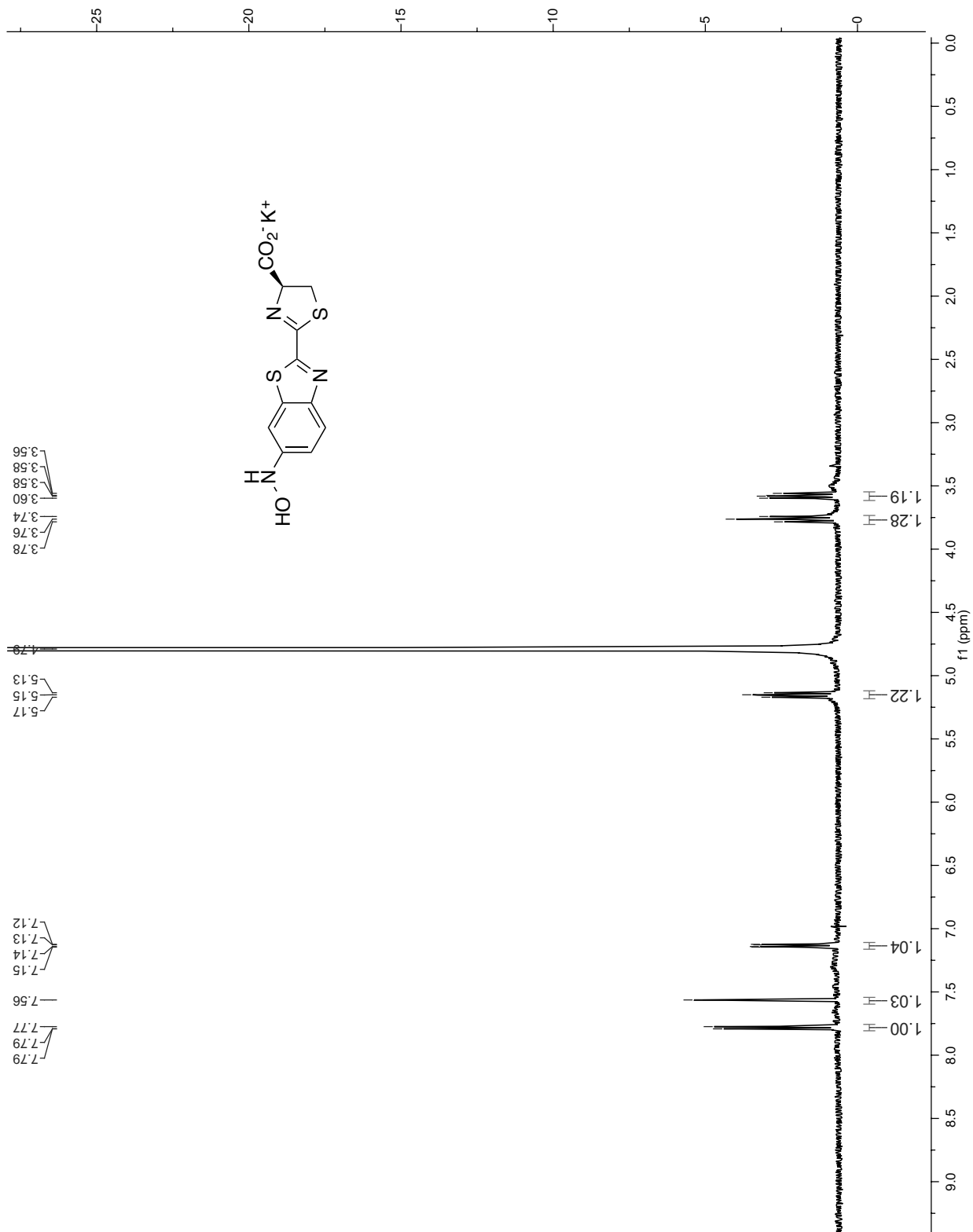


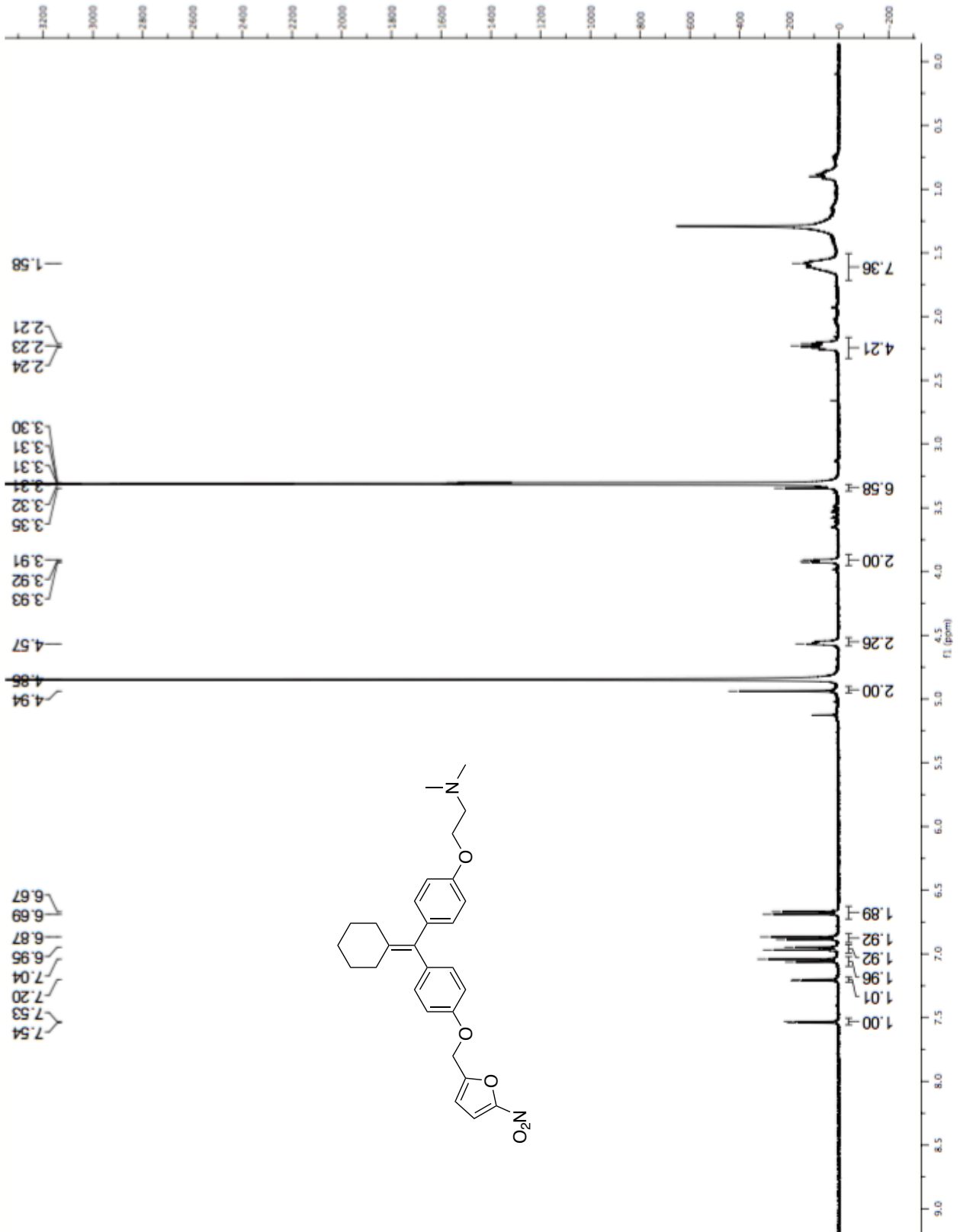












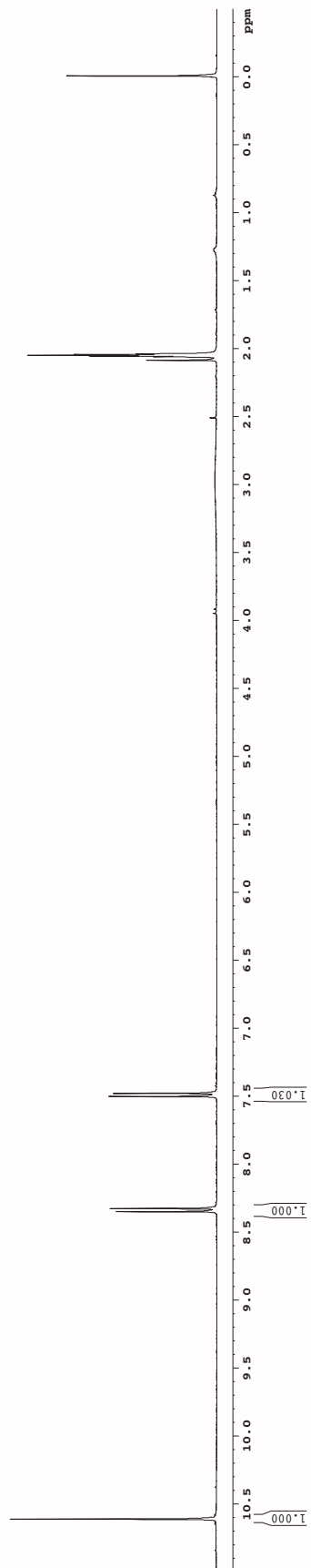
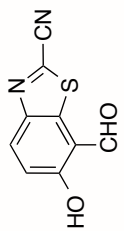
900.006

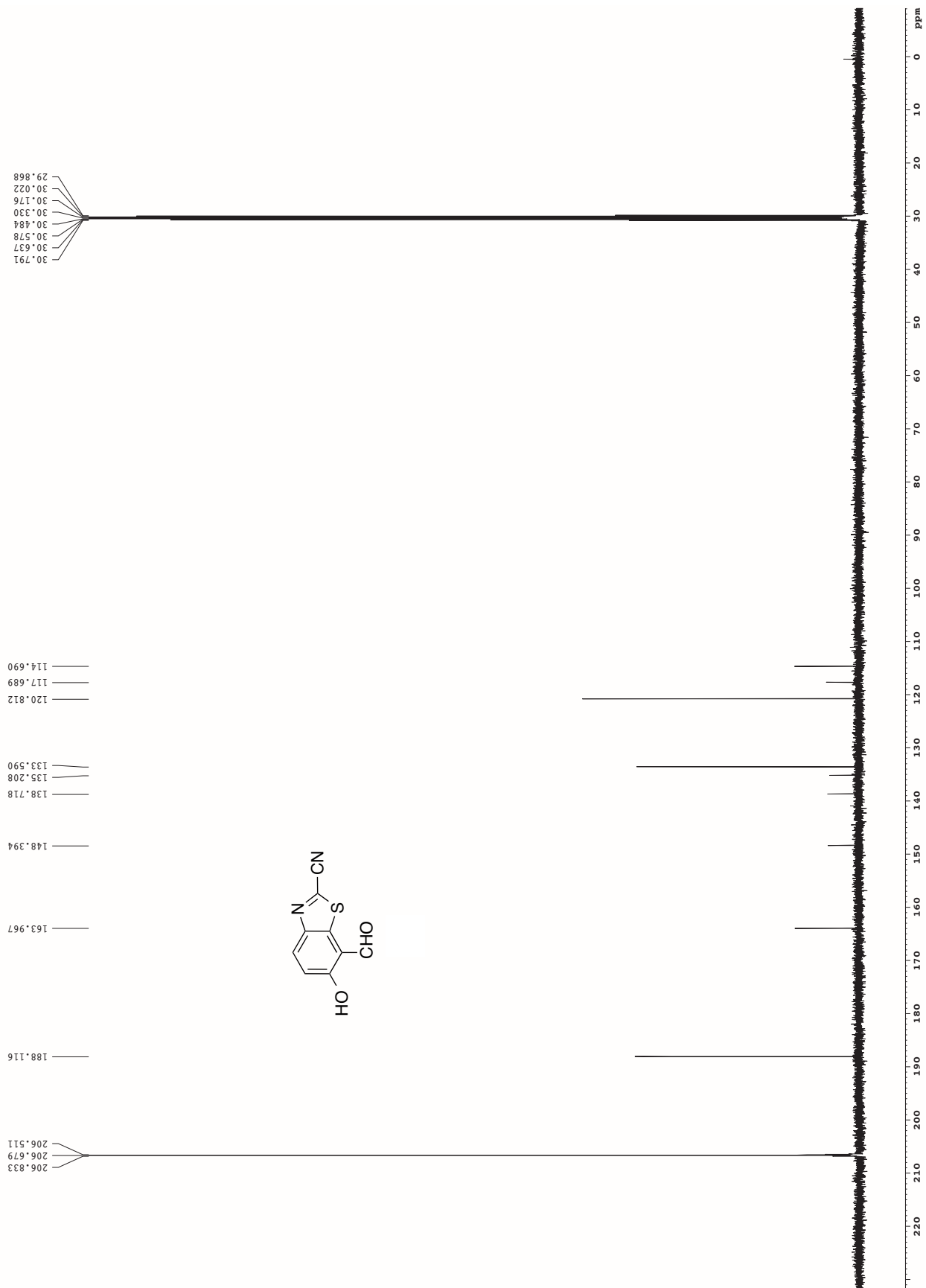
2.086
2.061
2.055
2.050
2.045
2.039

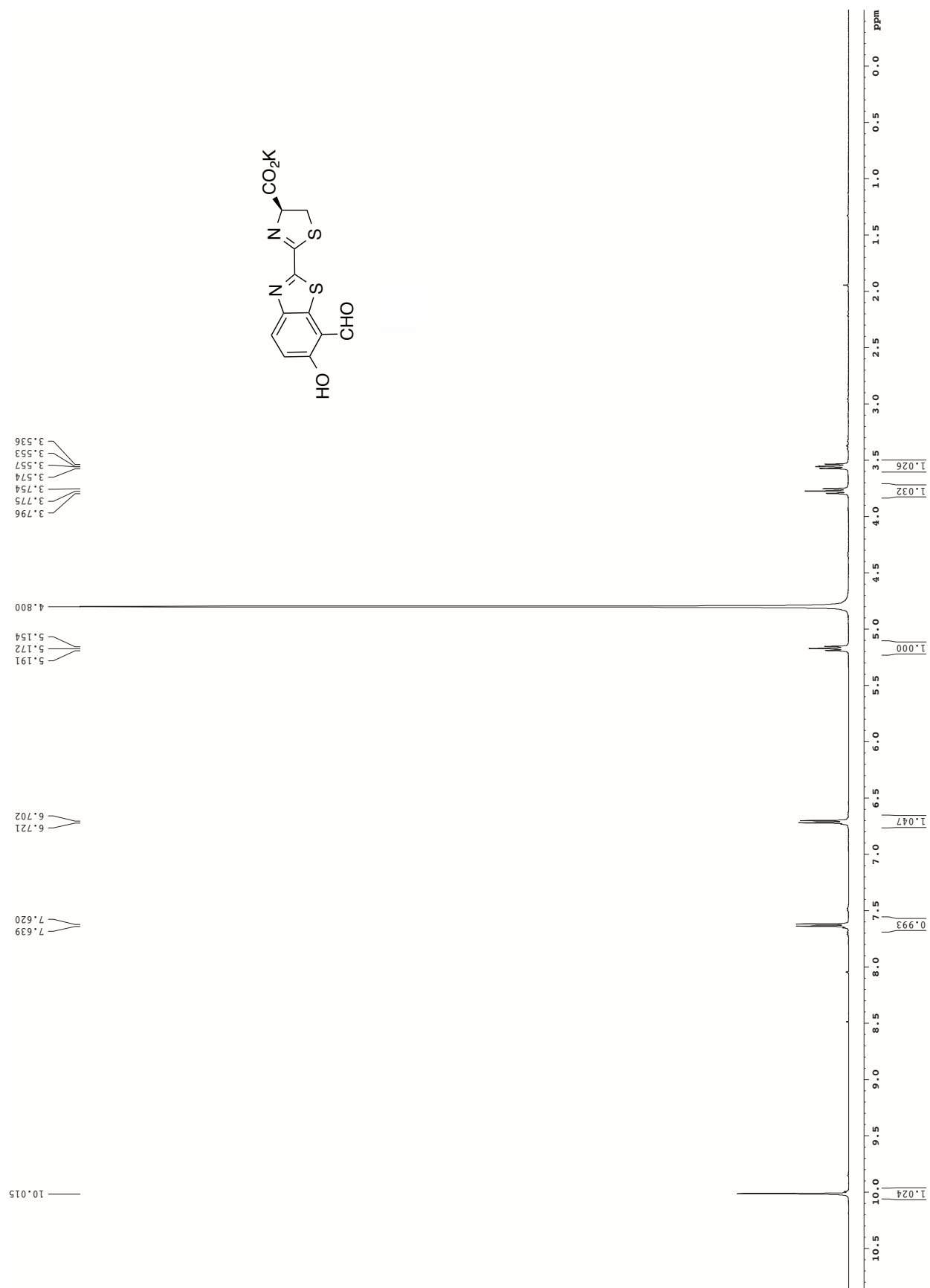
7.504
7.481

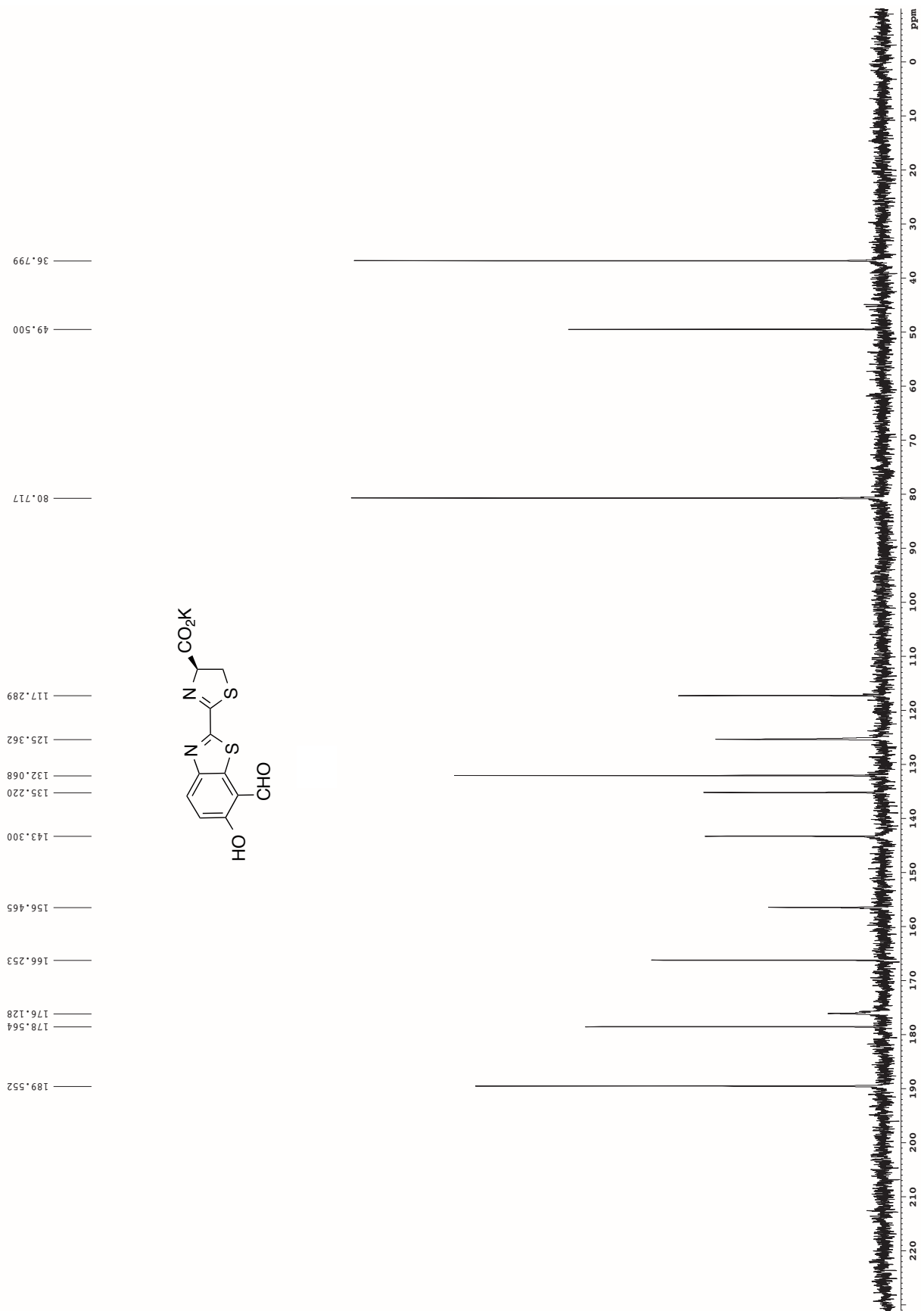
8.329
8.351

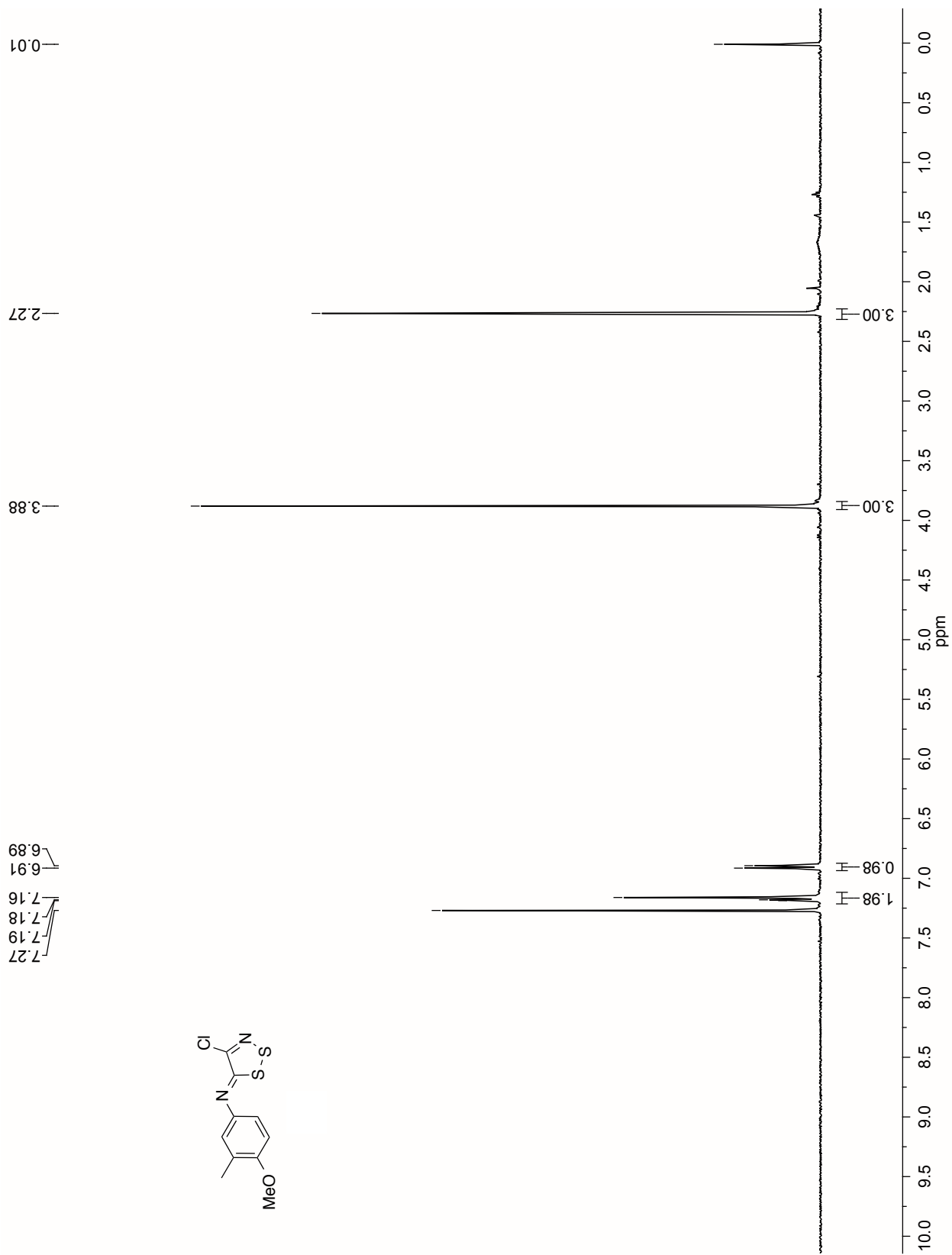
10.610

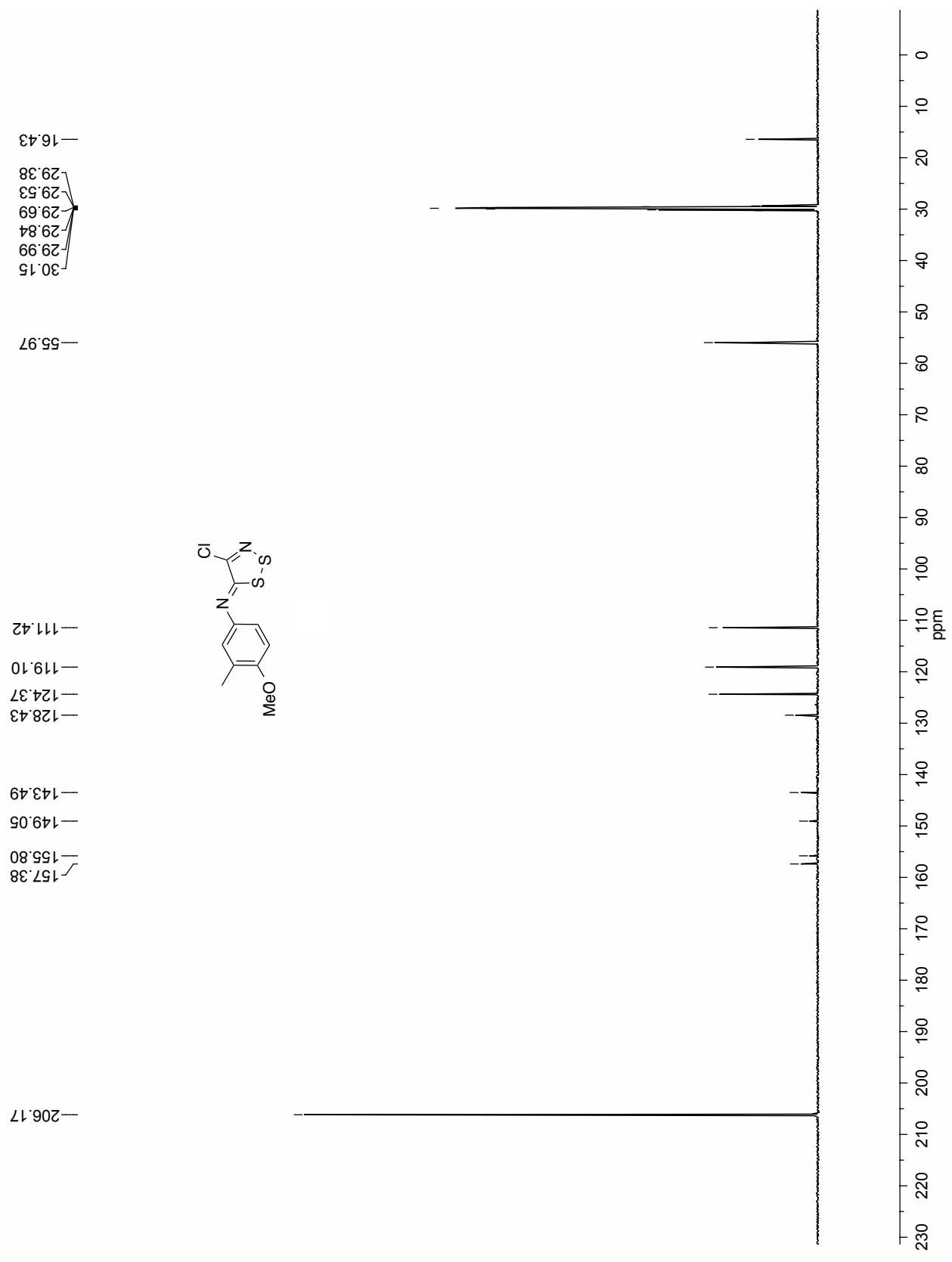


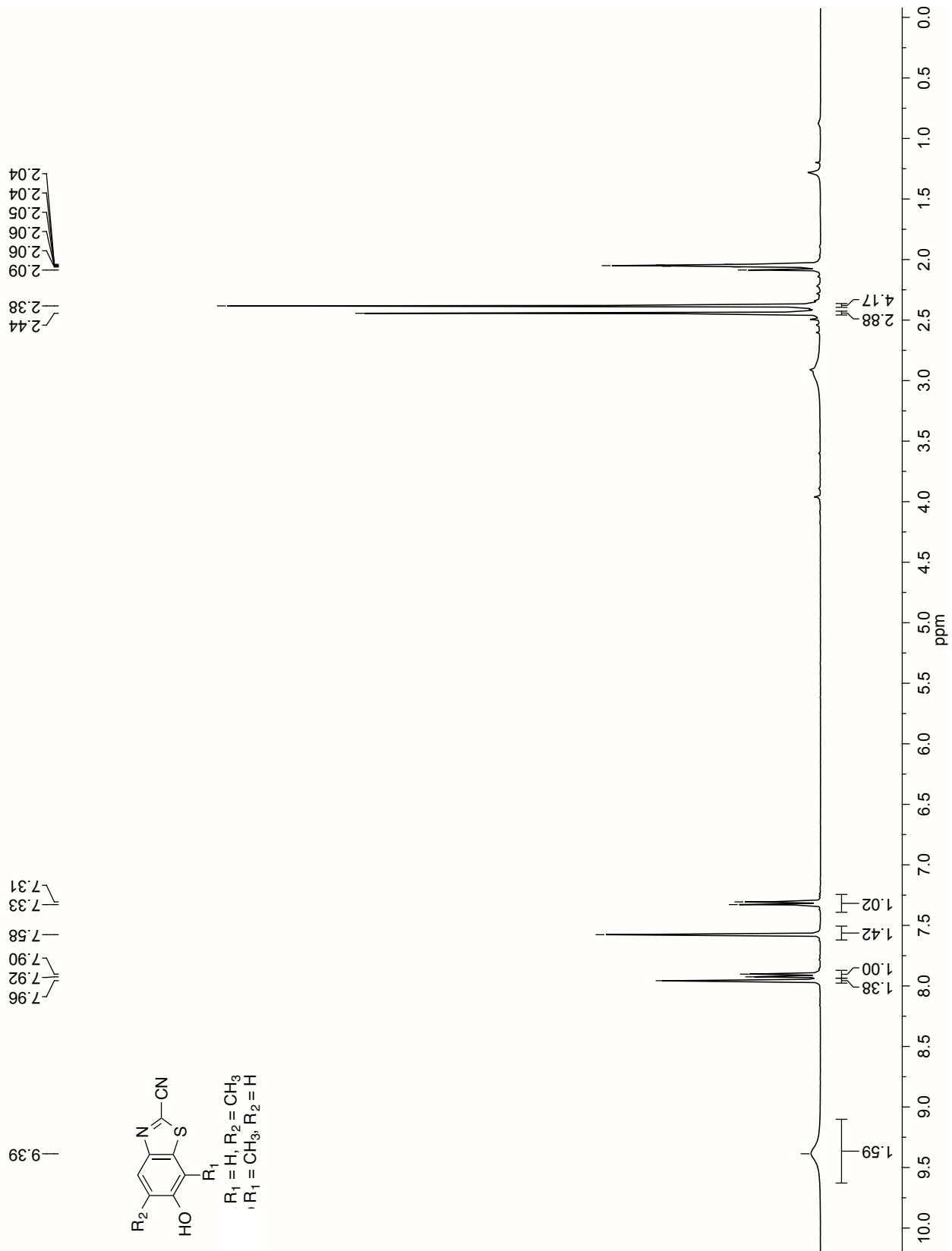








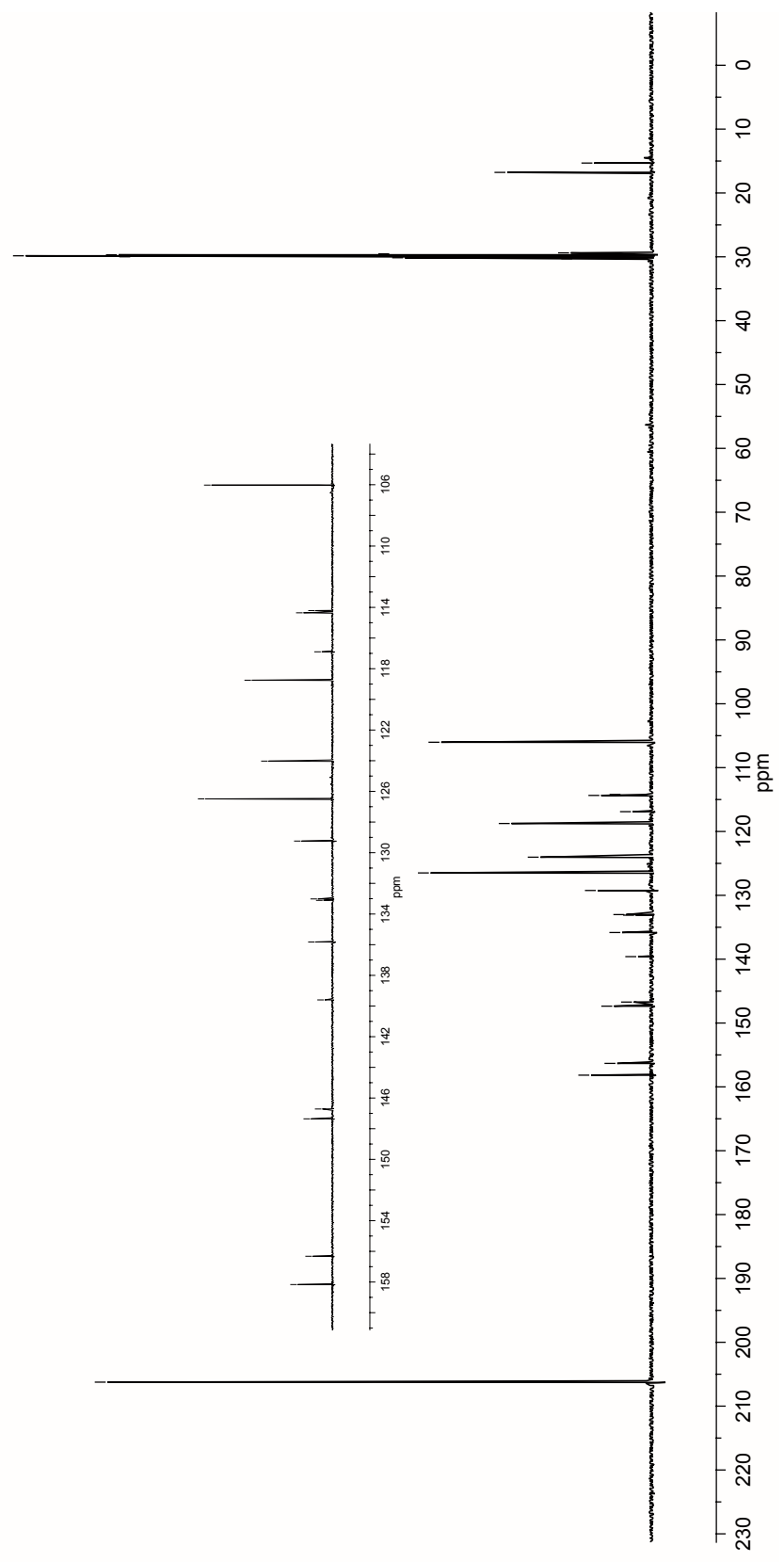
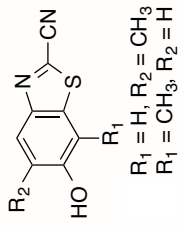


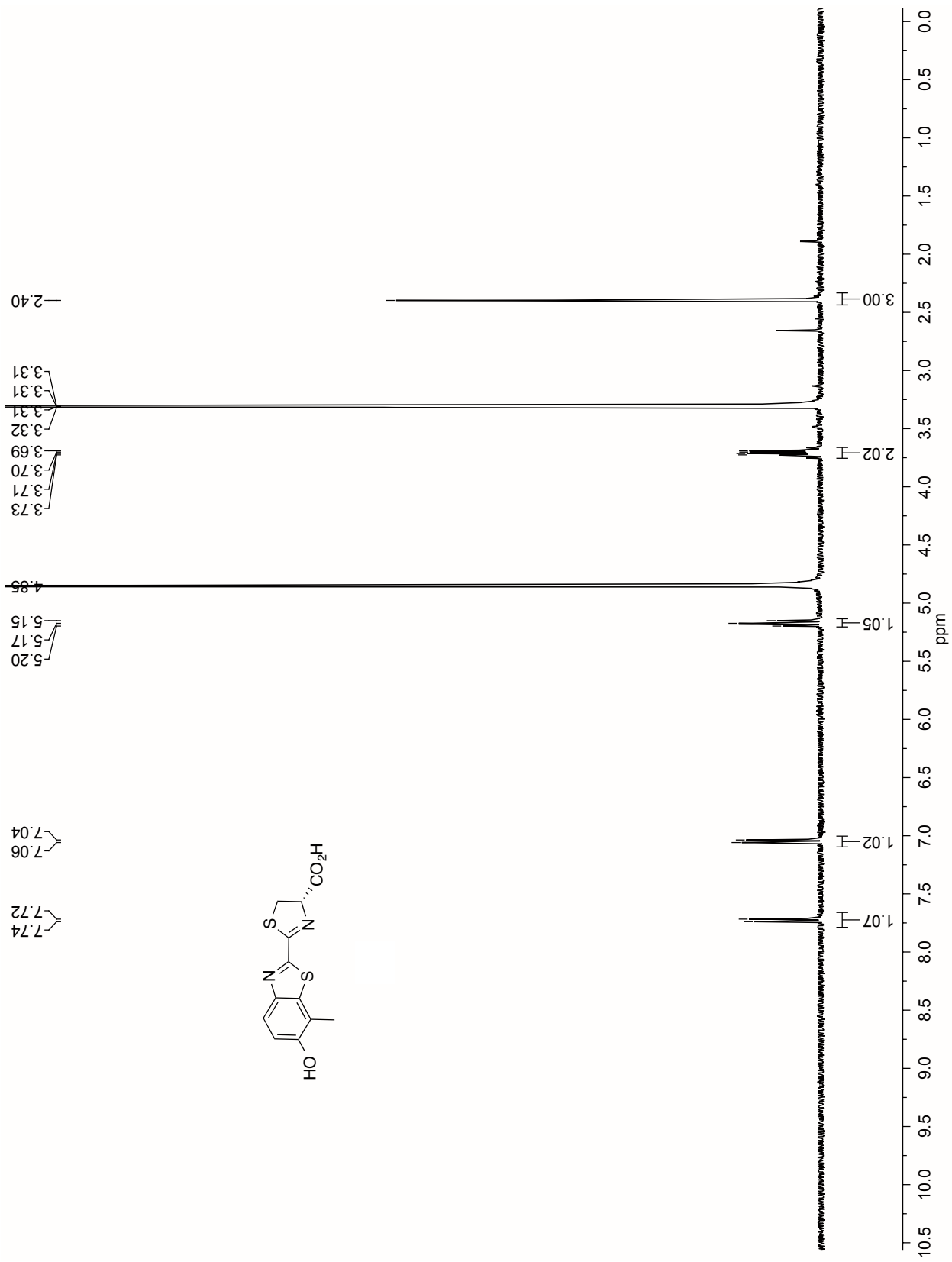


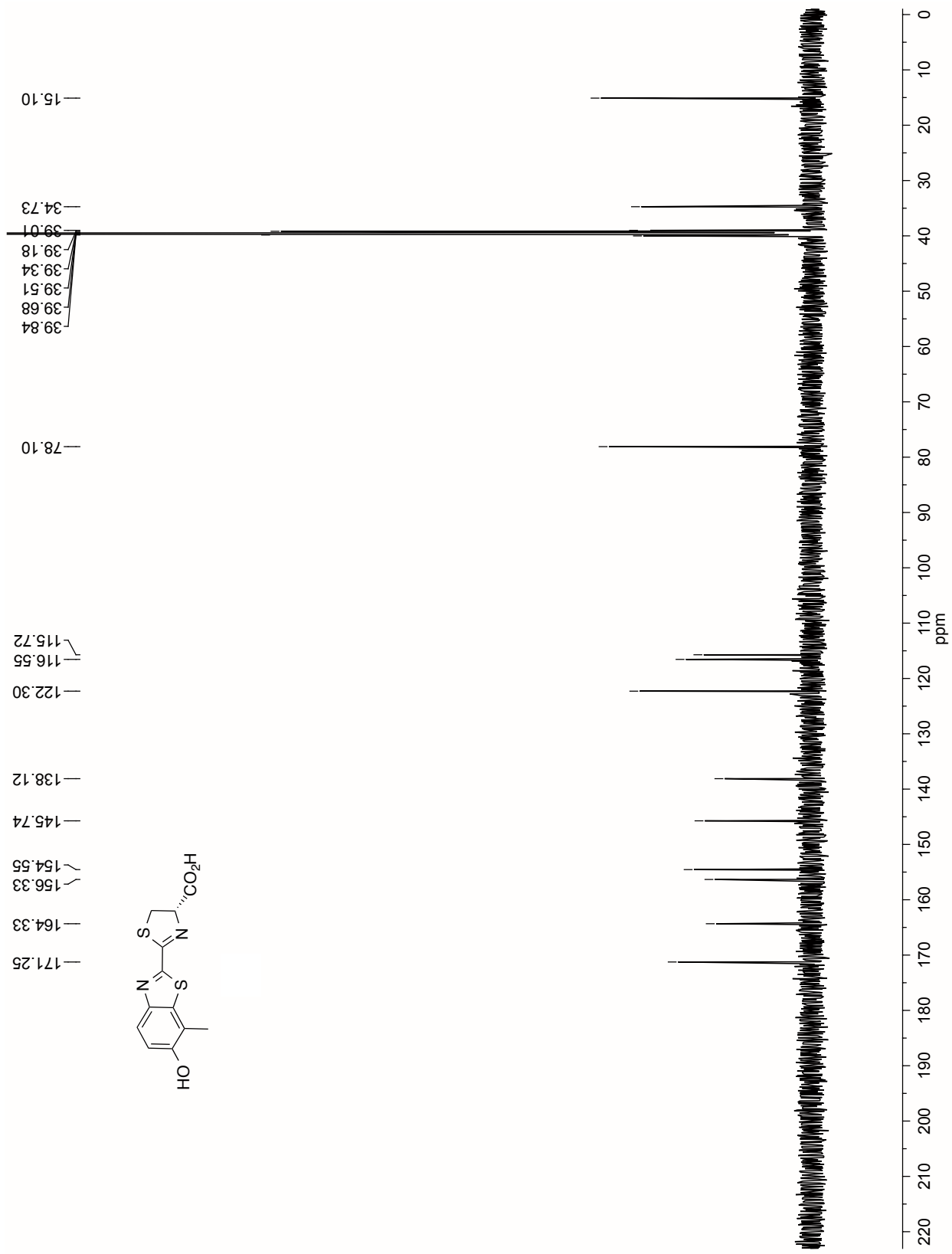
15.31
 16.78
 29.38
 29.53
 29.69
 29.84
 29.99
 30.15
 30.30

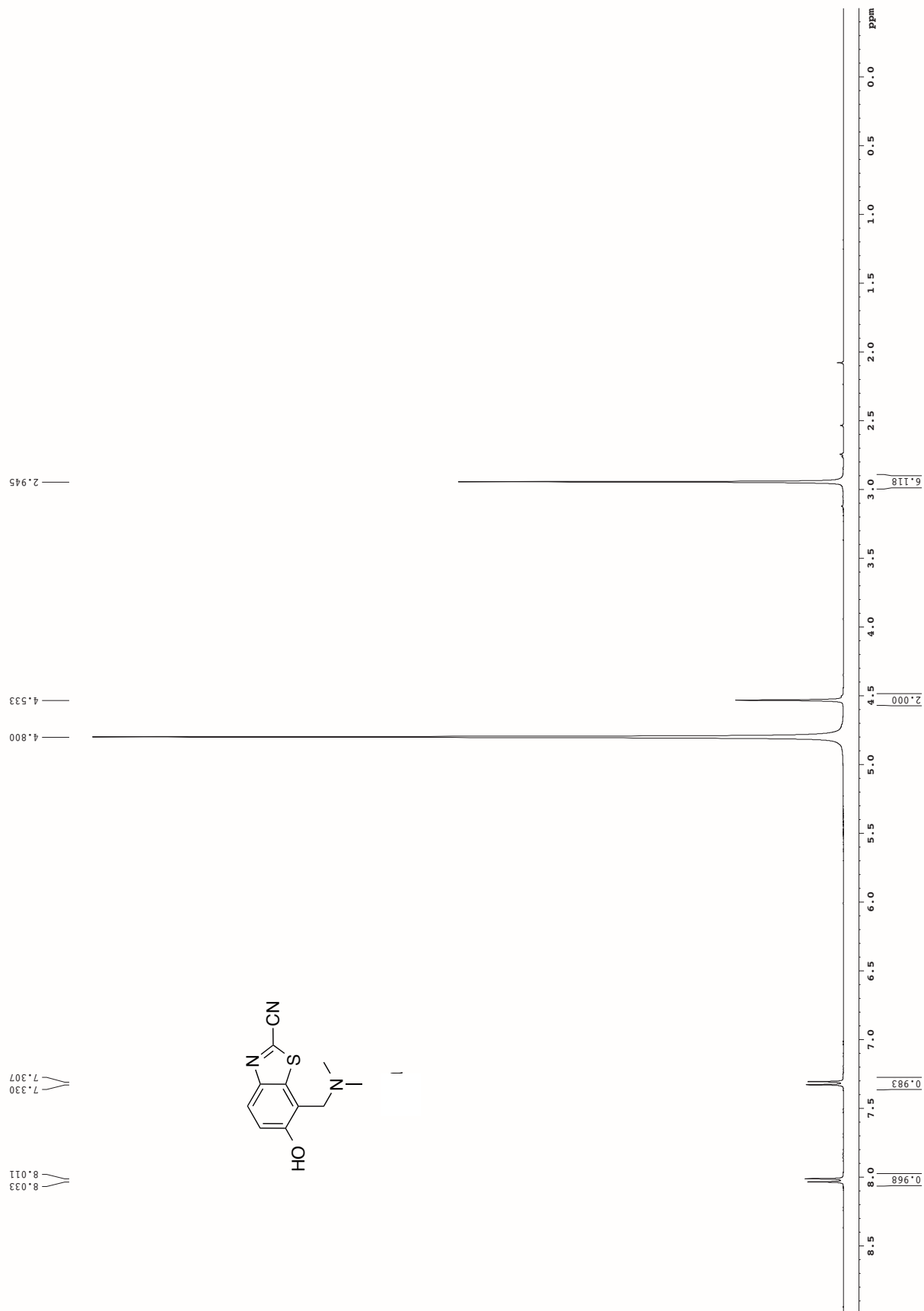
106.02
 114.21
 114.35
 116.90
 118.75
 124.03
 126.49
 128.49
 129.24
 124.03
 118.75
 124.03
 126.49
 133.02
 133.11
 133.02
 135.82
 139.60
 146.72
 147.36
 146.72
 147.36
 148.72
 156.32
 158.17

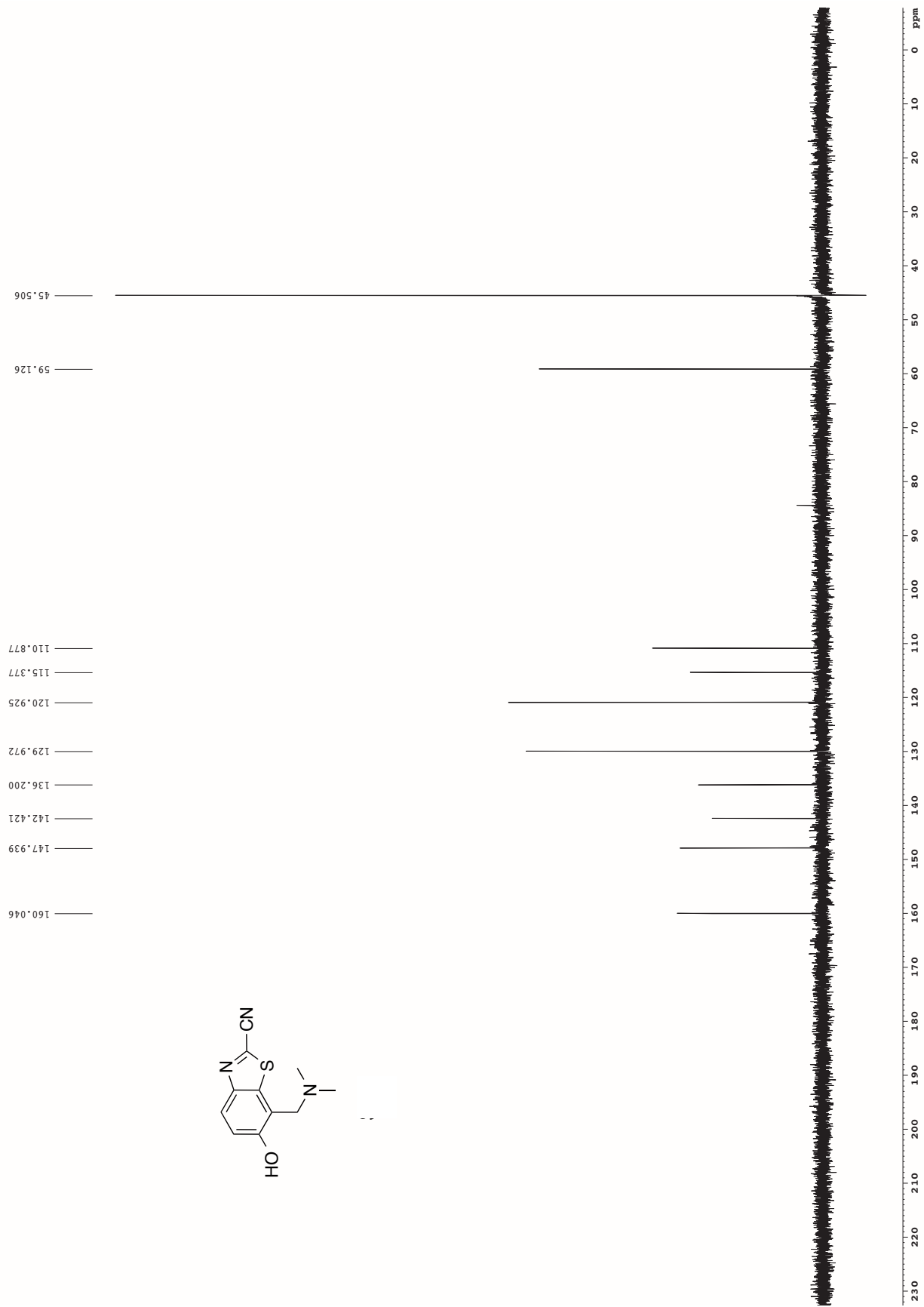
206.24

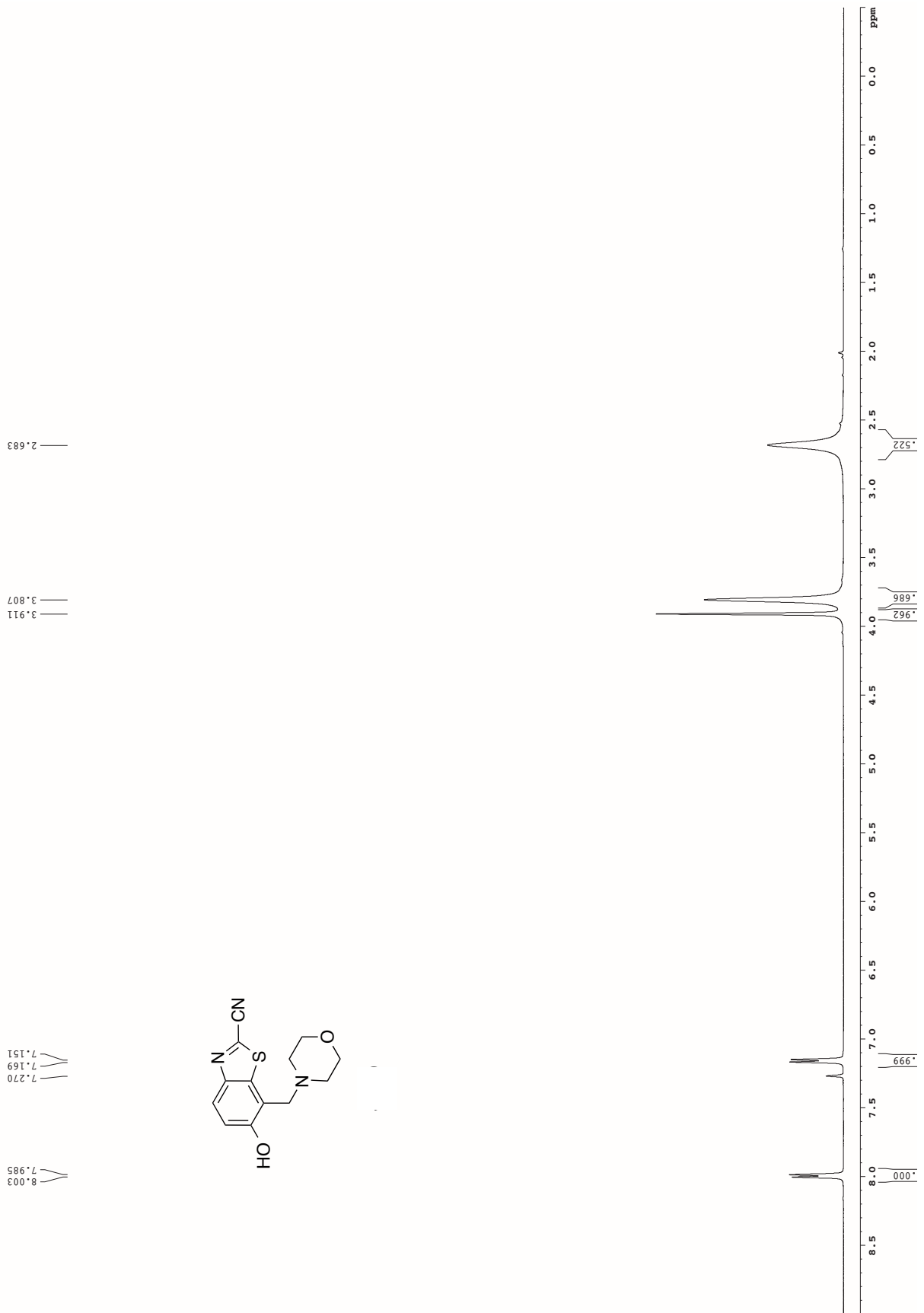


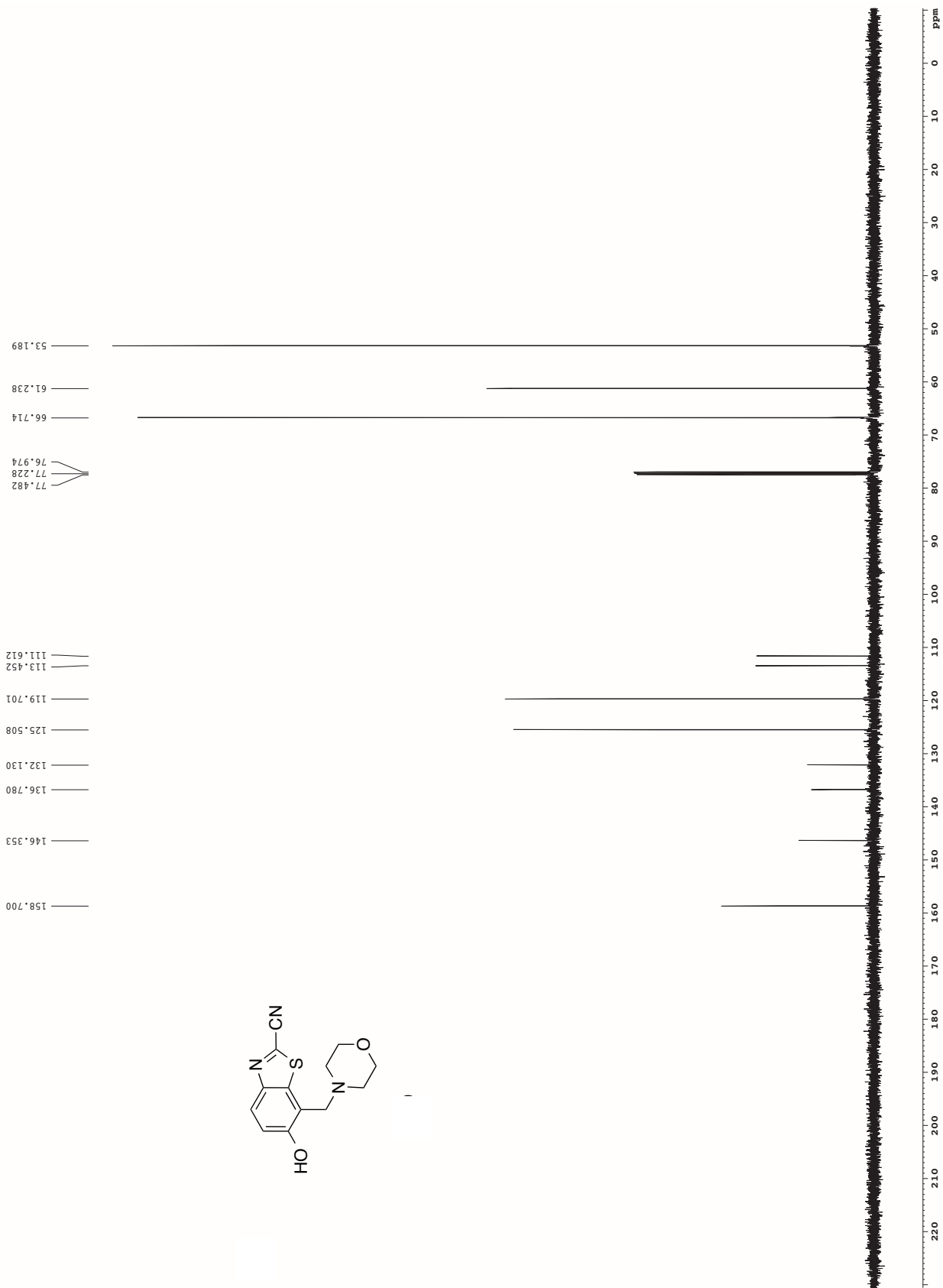
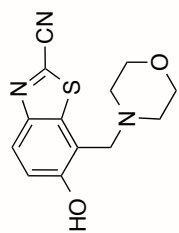












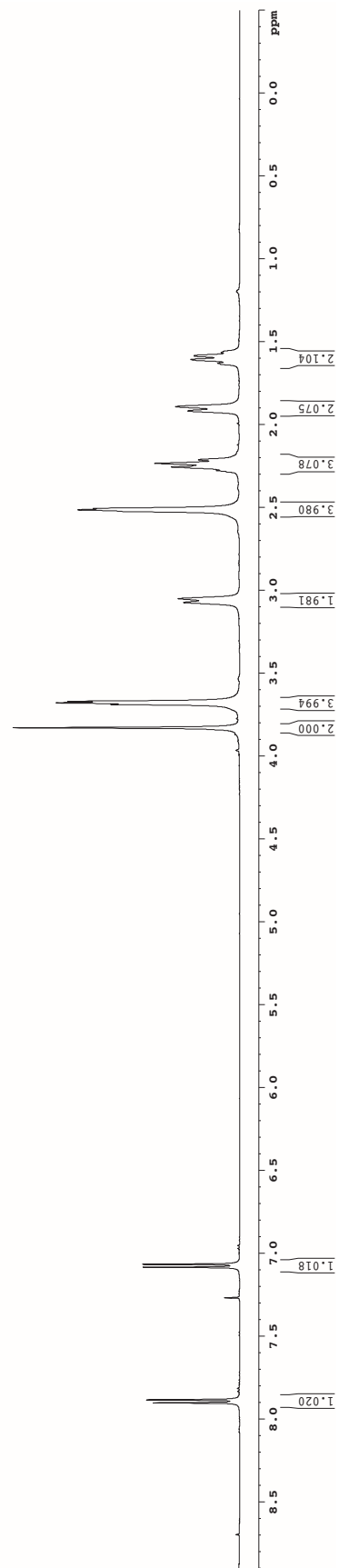
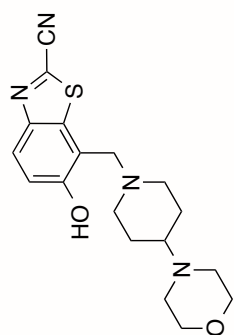
1.586
1.631
1.608
1.893
1.918
2.213
2.235
2.257
2.279
2.507
2.515

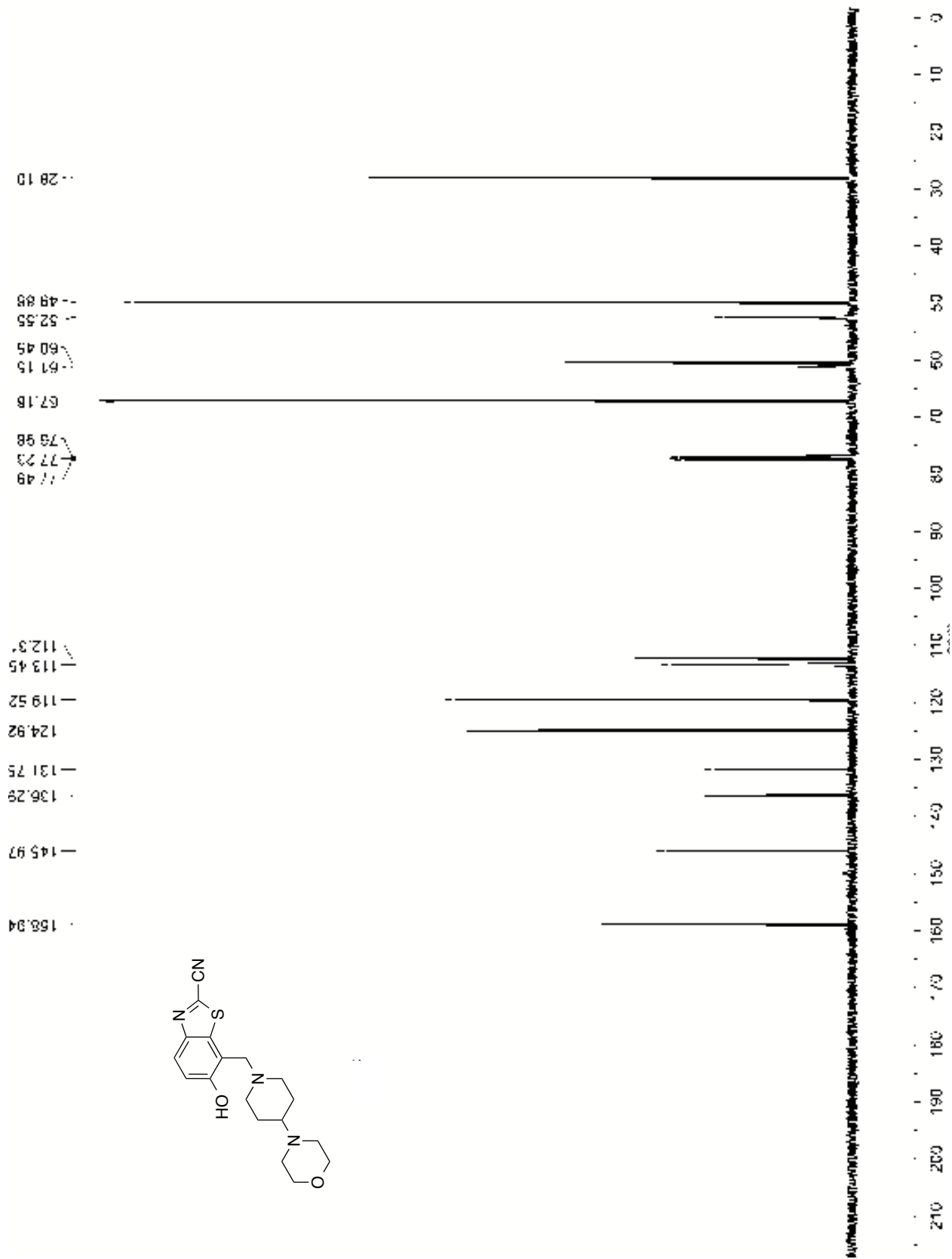
3.051
3.075

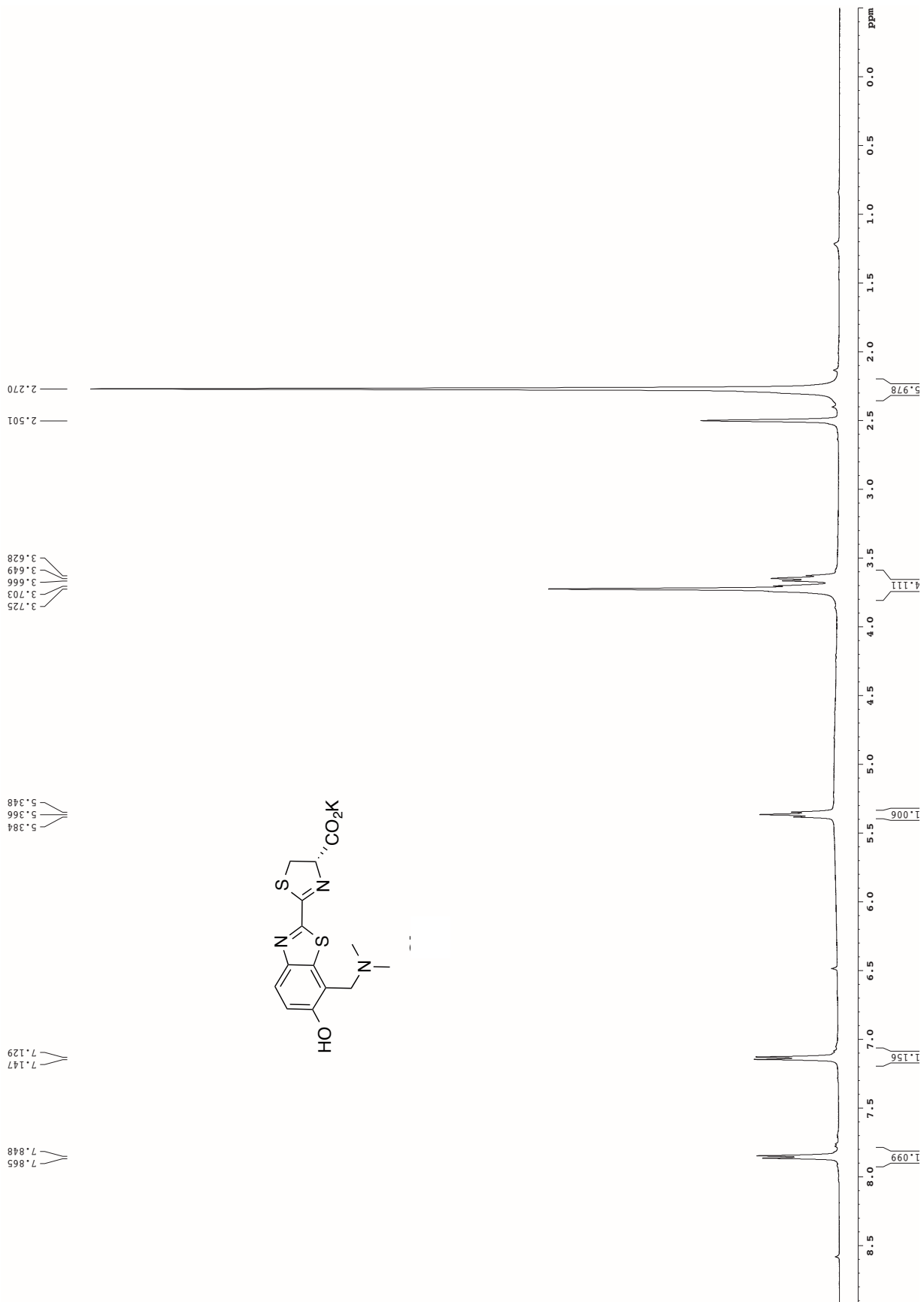
3.671
3.680
3.689
3.830

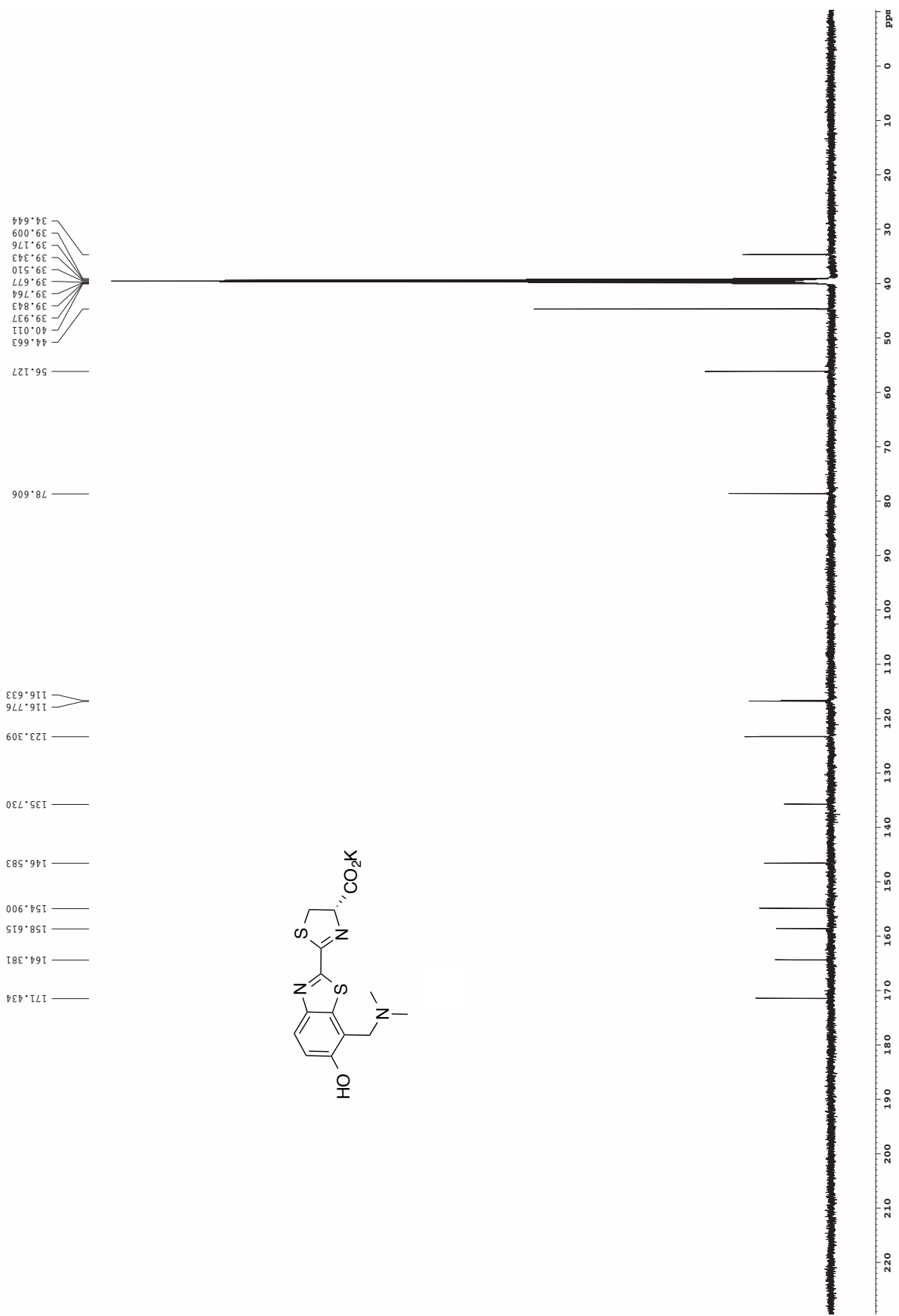
7.068
7.086
7.270

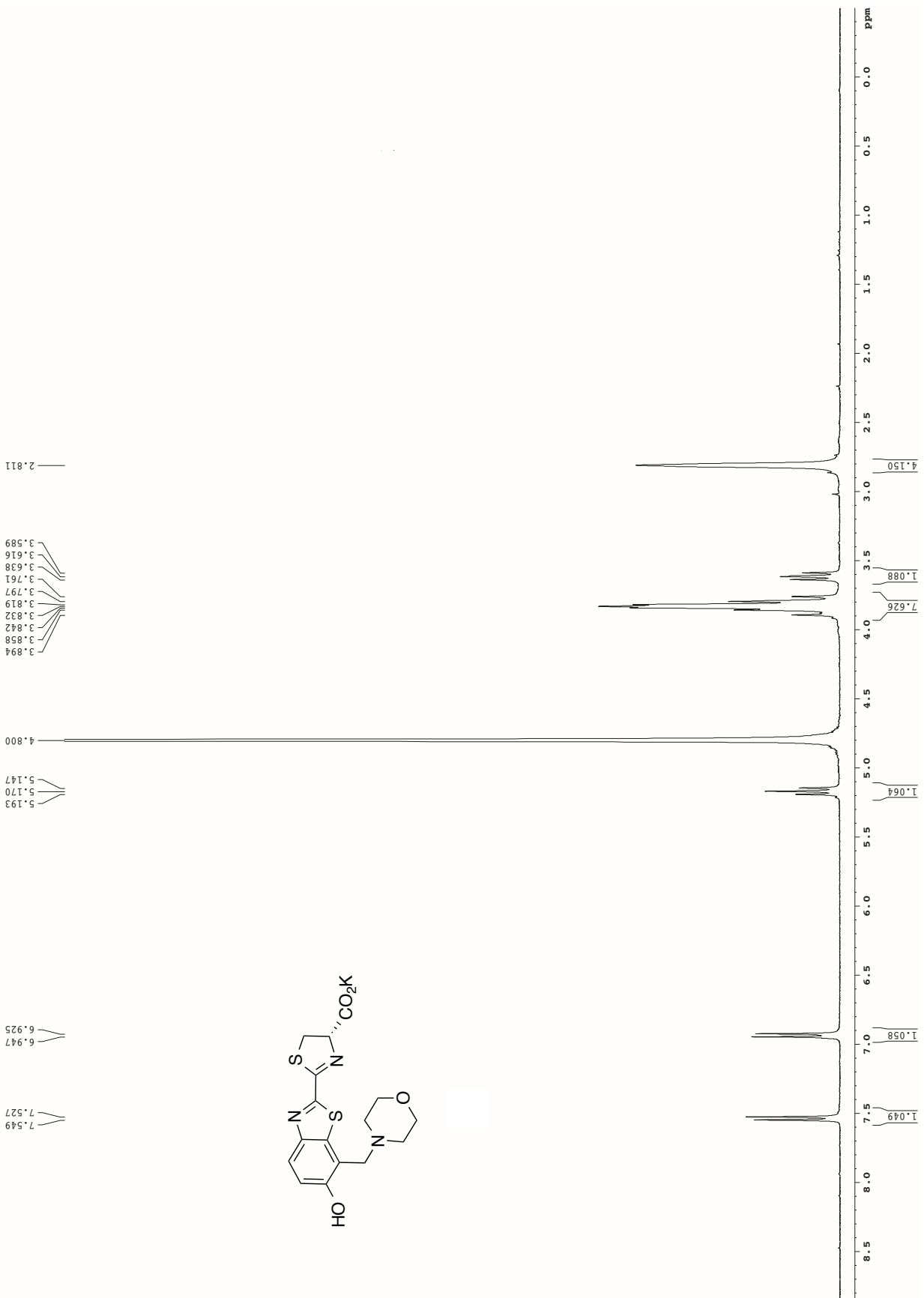
7.886
7.904

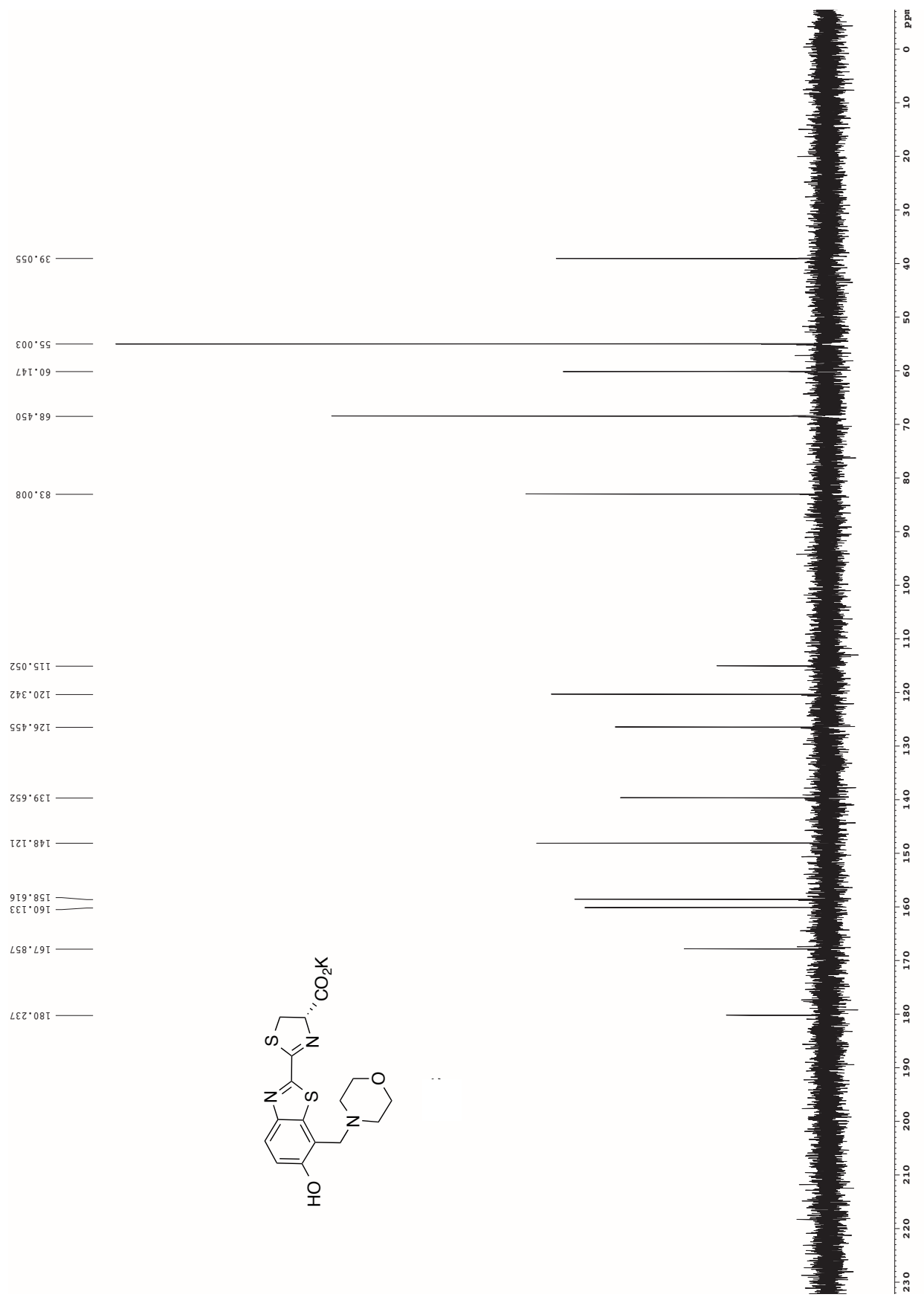


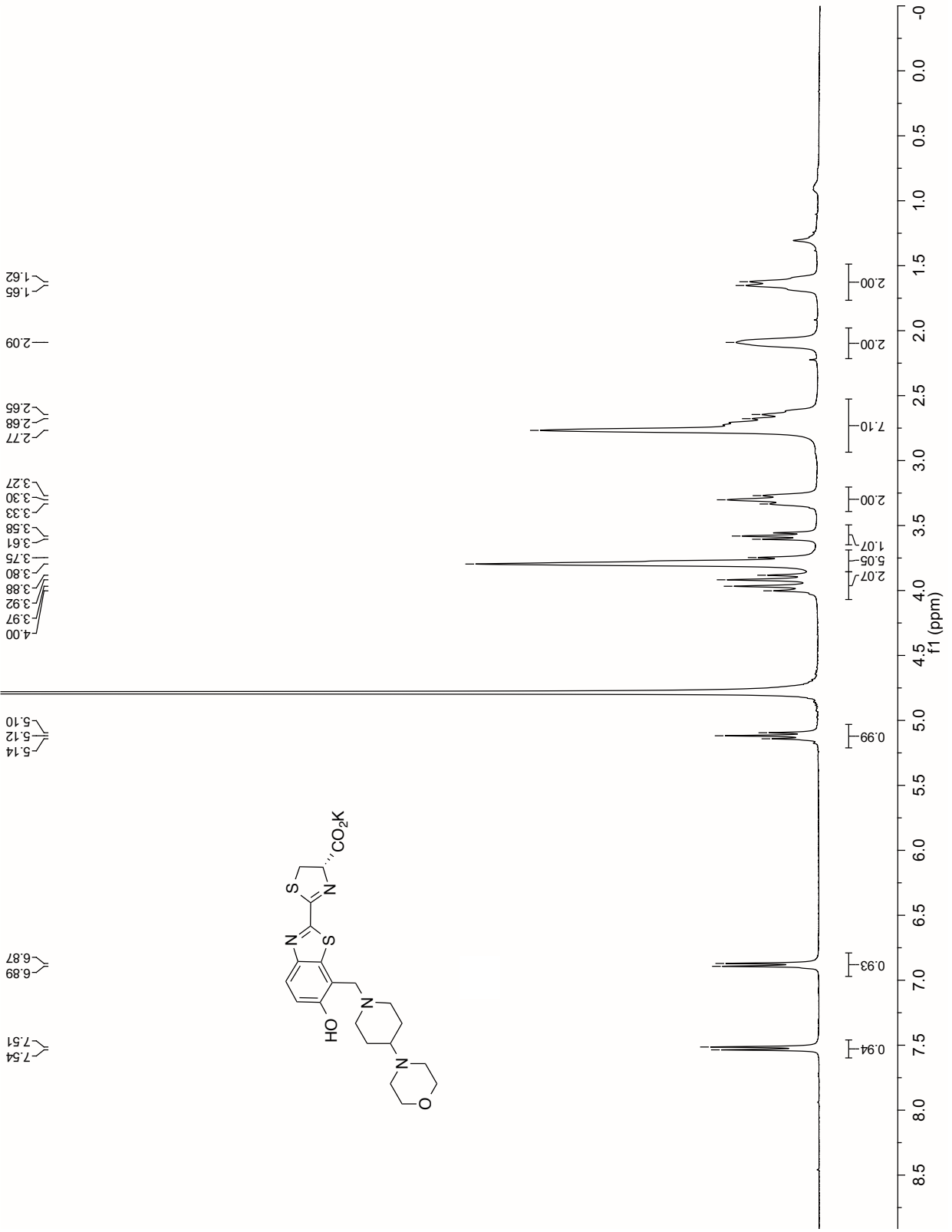


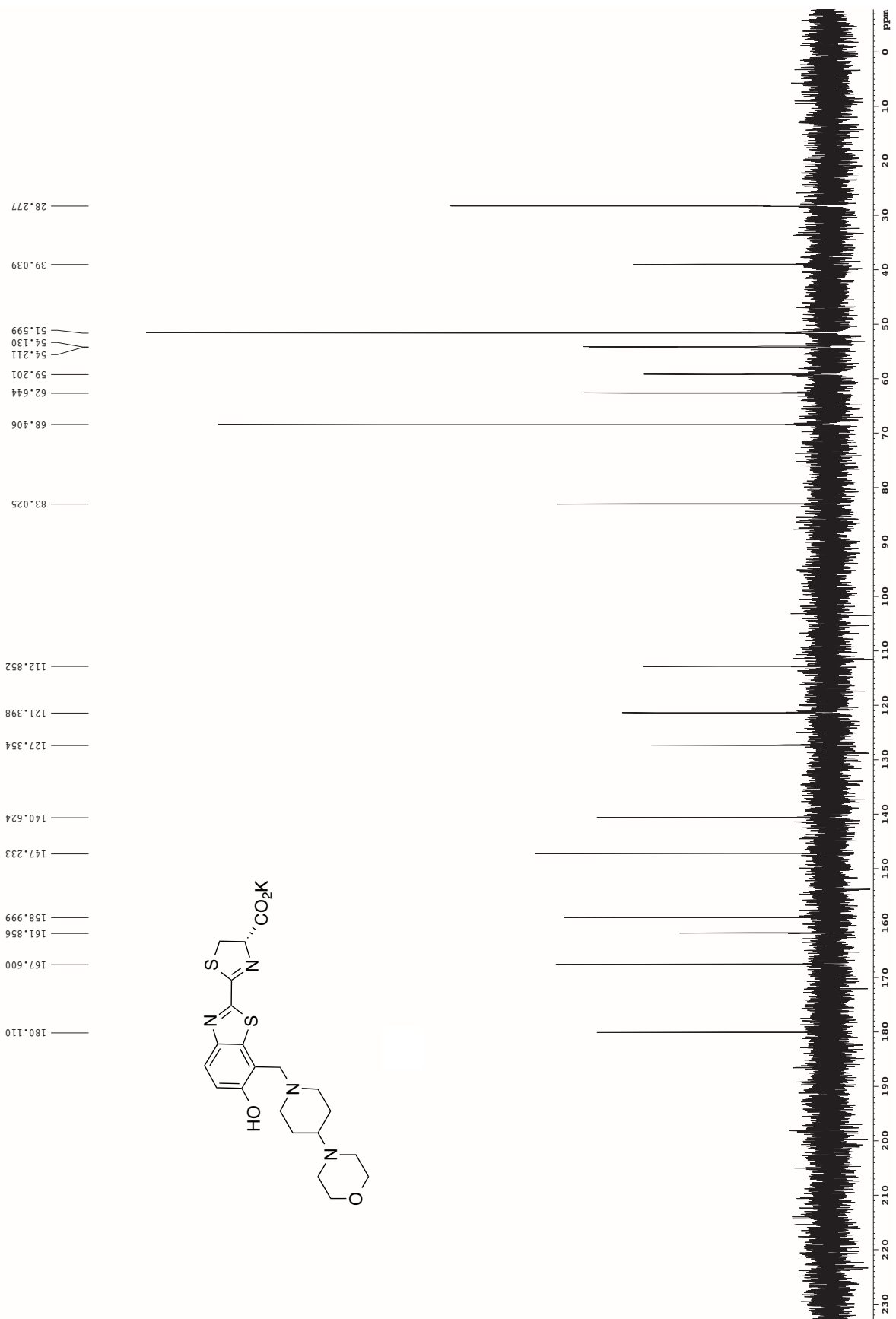


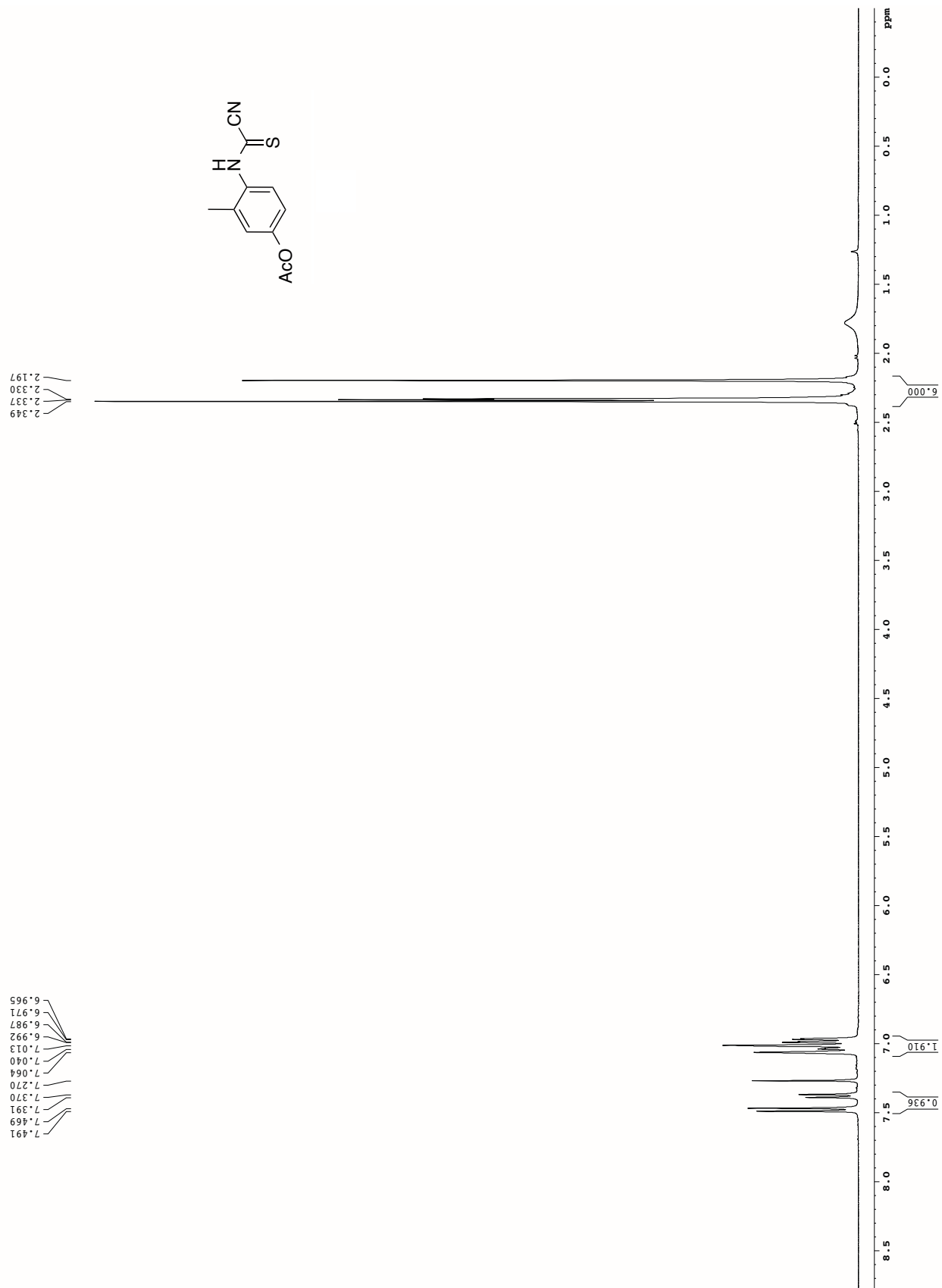




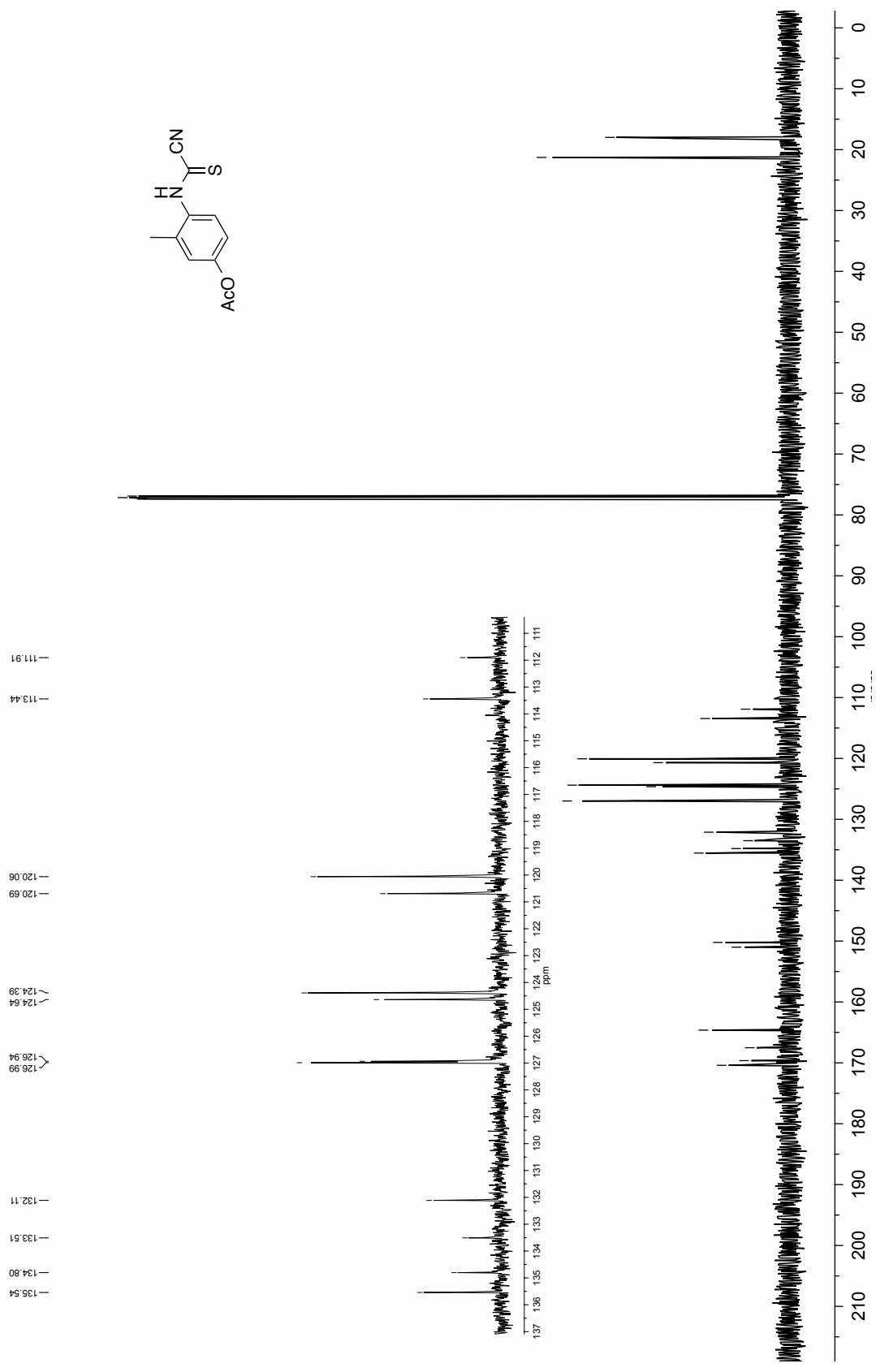
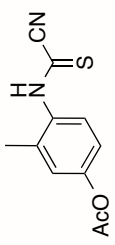


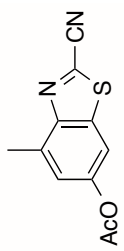






170.41
 169.64
 167.52
 164.62
 150.99
 150.22
 135.54
 134.80
 133.51
 132.11
 124.39
 120.06
 120.06
 113.44
 111.91
 77.41
 77.16
 76.91
 21.28
 18.17
 17.99





2.045
2.050
2.324
2.745
2.747
2.748

7.315
7.318
7.321
7.323
7.877
7.878
7.882
7.884

