

# UC Irvine

## UC Irvine Electronic Theses and Dissertations

### Title

Multicomponent bioluminescence imaging via substrate resolution

### Permalink

<https://escholarship.org/uc/item/9fj4780n>

### Author

Rathbun, Colin

### Publication Date

2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,  
IRVINE

Multicomponent bioluminescence imaging via substrate resolution

DISSERTATION

submitted in partial satisfaction of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

in Chemistry

by

Colin Michael Rathbun

Dissertation Committee:  
Professor Jennifer Prescher, Chair  
Professor Scott Rychnovsky  
Professor Chris Vanderwal

2018

Chapter 1 © American Chemical Society  
Chapter 2 © American Chemical Society  
Chapter 3 © American Chemical Society  
All other materials © 2018 Colin M. Rathbun

# DEDICATION

To my parents. For always tolerating my asperiments. I love you.

# TABLE OF CONTENTS

	Page
<b>LIST OF FIGURES</b>	<b>vii</b>
<b>LIST OF TABLES</b>	<b>ix</b>
<b>LIST OF SCHEMES</b>	<b>x</b>
<b>ACKNOWLEDGMENTS</b>	<b>xi</b>
<b>CURRICULUM VITAE</b>	<b>xii</b>
<b>ABSTRACT OF THE DISSERTATION</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Bioluminescence basics . . . . .	2
1.2 Building better bioluminescent tools . . . . .	4
1.2.1 Streamlined and scalable luciferin syntheses . . . . .	4
1.2.2 Brighter and more colorful probes . . . . .	6
1.2.3 In vivo improvements . . . . .	10
1.3 Luciferase-luciferin pairs for multicellular imaging . . . . .	10
1.4 Bioluminescent reporters for cell–cell interactions . . . . .	12
1.5 Bioluminescent reporters for analyte detection in vivo . . . . .	15

1.6	Conclusions and future directions . . . . .	15
1.7	Objectives of this study . . . . .	16
<b>2</b>	<b>Orthogonal luciferase-luciferin pairs for bioluminescence imaging</b>	<b>21</b>
2.1	Introduction . . . . .	21
2.2	Results and discussion . . . . .	24
2.2.1	Designing and constructing sterically modified luciferins . . . . .	24
2.2.2	Analyzing bioluminescent light emission with modified luciferins . . . . .	26
2.2.3	Measuring the light-emitting potential of luciferin analogues . . . . .	28
2.2.4	Evolving substrate-specific luciferases . . . . .	29
2.2.5	Analyzing the origins of orthogonality . . . . .	33
2.2.6	Cellular imaging with orthogonal pairs . . . . .	36
2.3	Conclusions . . . . .	37
2.4	Author contributions . . . . .	38
2.5	Materials and methods . . . . .	38
2.5.1	General bioluminescence imaging . . . . .	38
2.5.2	General chemiluminescence procedure <sup>52</sup> . . . . .	39
2.5.3	Bioluminescence kinetic measurements . . . . .	40
2.5.4	Data mining . . . . .	40
<b>3</b>	<b>Parallel screening for rapid identification of orthogonal bioluminescent tools</b>	<b>47</b>
3.1	Introduction . . . . .	47
3.2	Results and discussion . . . . .	50
3.2.1	Expanding the pool of candidate luciferins and luciferases . . . . .	50
3.2.2	Screening for Orthogonal Luciferase-Luciferin Pairs in Silico . . . . .	53

3.2.3	Imaging with orthogonal pairs in cultured cell and animal models . . .	56
3.2.4	Analyzing trends in orthogonal substrate use . . . . .	58
3.2.5	Added diversity improves orthogonality . . . . .	62
3.3	Conclusions . . . . .	64
3.4	Author contributions . . . . .	65
3.5	Materials and methods . . . . .	65
3.5.1	General bioluminescence imaging . . . . .	65
3.5.2	In silico screens . . . . .	65
3.5.3	Analyzing orthogonality as a function of diversity . . . . .	66
3.6	Note on parallel screening <i>in silico</i> . . . . .	66
<b>4</b>	<b>Brominated luciferins for multicomponent imaging and cross-coupling</b>	<b>75</b>
4.1	Introduction . . . . .	75
4.2	Synthesis of brominated cyclic luciferins . . . . .	77
4.3	Characterization of brominated cyclic luciferins . . . . .	78
4.4	Stille cross-coupling of brominated luciferin precursors . . . . .	79
4.5	Suzuki cross coupling of brominated luciferin precursors . . . . .	80
4.6	Conclusions and future directions . . . . .	82
4.7	Materials and methods . . . . .	82
4.7.1	General synthetic methods . . . . .	82
4.7.2	Compound Syntheses . . . . .	83
<b>5</b>	<b>Multicomponent bioluminescence imaging via substrate unmixing</b>	<b>91</b>
5.1	Introduction . . . . .	91
5.2	Results and discussion . . . . .	94
5.2.1	Linear unmixing algorithms distinguish orthogonal bioluminescent probes	94

5.2.2	Mixed cell populations can be distinguished in vivo. . . . .	99
5.3	Conclusions and future directions . . . . .	101
5.4	Methods and Materials . . . . .	102
5.4.1	General bioluminescence imaging . . . . .	102
5.4.2	Bacterial lysate analysis of luciferase mutants . . . . .	102
5.4.3	CrossCompare algorithm . . . . .	102
5.4.4	Substrate unmixing . . . . .	103
5.4.5	Mammalian plasmid construction . . . . .	103
5.4.6	Mammalian cell culture . . . . .	103
5.4.7	Mammalian cell imaging with luciferase mutants . . . . .	104
5.4.8	<i>In vivo</i> imaging of orthogonal luciferase-luciferin pairs . . . . .	104
<b>Appendices</b>		<b>108</b>
A	NMR spectra . . . . .	108



# LIST OF FIGURES

	Page
1.1 Popular luciferase-luciferin pairs for cellular imaging . . . . .	3
1.2 Recently synthesized luciferin heterocycles . . . . .	5
1.3 Palette of luciferin analogues . . . . .	7
1.4 BRET for deep-tissue imaging . . . . .	9
1.5 Orthogonal luciferin-luciferase pairs for multicomponent imaging . . . . .	12
1.6 Bioluminescent reporters for cellular interactions. . . . .	14
2.1 Expanding the bioluminescence toolkit with unique enzyme-substrate pairs .	22
2.2 Measuring luciferin light emission . . . . .	27
2.3 Generating mutant luciferase libraries and screening for orthogonal pairs . .	30
2.4 Analyzing orthogonal enzyme-substrate pairs . . . . .	32
2.5 Bioluminescent photon production from luciferin analogs with combinatorial enzymes . . . . .	35
2.6 Imaging cells with orthogonal luciferase-luciferin pairs . . . . .	36
3.1 Parallel screening of luciferase mutants and luciferin analogues to identify orthogonal pairs . . . . .	49
3.2 Brighter mutants can be evolved but are not orthogonal. . . . .	51
3.3 Sterically modified luciferins emit varying levels of photons with Fluc . . . .	52
3.4 Uncovering orthogonal pairs in silico . . . . .	54

3.5	Multicellular imaging with orthogonal pairs . . . . .	56
3.6	Noninvasive in vivo imaging with orthogonal pairs . . . . .	57
3.7	Examining the origins of substrate selectivity . . . . .	59
3.8	Selectivities of mutant enzymes for synthetic luciferins . . . . .	60
3.9	Orthogonal pair analysis . . . . .	62
3.10	Improving orthogonality via enzyme-substrate diversity . . . . .	63
4.1	Dose response of brominated cyclic luciferins . . . . .	79
5.1	Bioluminescent substrate unmixing for the study of disease states . . . . .	93
5.2	Initial mutants of interest for multicomponent bioluminescence imaging . . . . .	94
5.3	Design of experiment to test the success of substrate unmixing . . . . .	96
5.4	Orthogonal bioluminescent probes can be distinguished in bacterial lysate . . . . .	97
5.5	Orthogonal bioluminescent probes can be distinguished in mammalian cells . . . . .	98
5.6	Orthogonal mutant screen in mice . . . . .	100
5.7	Orthogonal bioluminescent probes can be distinguished in mice . . . . .	101

## LIST OF TABLES

	Page
2.1 Biochemical analyses of orthogonal enzyme-substrate pairs . . . . .	33
4.1 Evaluation of Stille cross coupling conditions . . . . .	80
4.2 Evaluation of Suzuki cross coupling conditions . . . . .	81
5.1 Target mutants employed for resolution of <b>4'-BrLuc</b> and D-luc . . . . .	95

## LIST OF SCHEMES

	Page
4.1 Cyclic luciferin retrosyntheses. . . . .	76
4.2 Synthesis of <b>4'-BrCycLuc</b> . . . . .	77
4.3 Synthesis of <b>7'-BrCycLuc</b> . . . . .	78
4.4 Preliminary scope of cross coupling with brominated luciferins . . . . .	81

# ACKNOWLEDGMENTS

I would like to thank the National Science Foundation for funding through the Graduate Research Fellowship Program (grant No. DGE-1321846, and Allergan for funding through the Allergan Graduate Fellowship. I would also like to thank the members of the Prescher lab, Bernard Choi, and Bruce Tromberg, for very helpful discussions.

Jenn, thanks for being an excellent mentor and the smartest person I know. I appreciate that you were so welcoming and took me in to the lab during a time when I thought I might give up on my Ph.D. Thanks for always tolerating my shenanigans in and out of the lab.

Crystal, supporting each other through graduate school has made all of this so much easier. Thanks for putting up with my late nights in the lab, or my long weekend days of experiments. I love you, and I am excited about what the next chapter of our lives will be.

To 2208 Meyer Pl, Daniel, and DGRMDPhD: I'll always remember our fun times.

# CURRICULUM VITAE

## CONTACT INFORMATION

**Colin M. Rathbun**

University of California, Irvine  
Chemistry Dept., Natural Sciences II  
Irvine, CA 92617

Voice: (616) 920-2679  
E-mail: crathbun@uci.edu

## EDUCATION

**University of California, Irvine, Irvine, CA**

Ph.D., Chemistry, September 2012 – May 11, 2018

- **Thesis Advisor:** Associate Professor Jennifer Prescher
- 3.98 cumulative GPA

**Hope College, Holland, MI, USA**

B.S., Chemistry (A.C.S. Certified), Minor in Mathematics, May, 2012

- **Thesis Advisor:** Professor Jeffrey B. Johnson
- 3.99 cumulative GPA

## RESEARCH EXPERIENCE

**University of California, Irvine, Irvine, CA**

Under Professor Jennifer A. Prescher

**May 2014 – May 2018**

- *Building better bioluminescent reporters via ab initio calculations* (2014)
- *Orthogonal luciferase–luciferin pairs for bioluminescence imaging* (February 2015 – November 2016)
- *Engineering mutually orthogonal enzyme–substrate pairs for bioluminescence imaging* (November 2016 – November 2017)
- *Multicomponent bioluminescence imaging in vivo* (November 2017 – May 2018)

Under Professor Vy M. Dong

**December 2012 – May 2014**

- *Mechanistic study of a metal-catalyzed, regioselective carbohydrate acylation reaction* (2013)
- *Rhodium-catalyzed retrohydroformylation* (January 2014 – May 2014)

**Hope College, Holland, MI**

Under Professor Jeffrey B. Johnson

**2010 – 2012**

- *Kinetic investigation of C–C bond activation in quinolinyl ketones* (2010)
- *Observing the effects of ligand modification on C–C bond activation* (2011)
- *C–C bond activation promoted by an imine directing group* (Summer 2012)

**University of Buenos Aires, Buenos Aires, Argentina**

Under Professor Fabio Doctorovich

**Summer 2011**

- *Synthesis of an electron-poor, water-soluble porphyrin for the isolation of HNO*

## SELECTED HONORS & AWARDS

**National Science Foundation**

- **Graduate Research Fellowship** (2012)
- **International Research Experience for Undergraduates** (2011)

### Barry M. Goldwater Scholarship Foundation

- Barry M. Goldwater Scholarship (2011)

### University of California, Irvine

- Allergan Graduate Fellowship (2017 – 2018)
- Grad Slam Campus-Wide Finalist (2017)
- UCI NSF GRFP Symposium 3rd-Place Presentation (2017)
- AGS Symposium Judges' Winner (2016)

### Hope College

- Presidential Scholarship (2008)
- Chemistry Department Jaeger Scholarship (2009)
- Chemistry Department John H. Kleinheksel Award (2009)
- Alcor Chapter Mortar Board (2011 – 2012)

### PEER-REVIEWED PUBLICATIONS

9. Rathbun, C. M.; Ionkina, A.; Prescher, J. A. "Multicomponent bioluminescence imaging in vivo." *Manuscript in preparation*.
8. Rathbun, C. M.\*; Porterfield, W. B.\*; Jones, K. A.\*; Sagoe, M. J.; Reyes, M. R.; Hua, C. T.; Prescher, J. A. "Parallel screening for rapid identification of orthogonal bioluminescent tools." *ACS Cent. Sci.*, **2017**, 3, 1254.
7. Rathbun, C. M.; Prescher, J. A. "Bioluminescent Probes for Imaging Biology Beyond the Culture Dish." *Biochemistry*, **2017**, 56, 5178. *Invited review*.
6. Rathbun, C. M.\*; Jones, K. A.\*; Porterfield, W. B.\*; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. "Orthogonal Luciferase-Luciferin Pairs for Bioluminescence Imaging." *J. Am. Chem. Soc.*, **2017**, 139, 2351.
5. Steinhardt, R. C.; Rathbun, C. M.; Krull, B. T.; Yu, J. M.; Yang Y.; Nguyen, B. D.; Kwon, J.; McCutcheon, D. C.; Jones, K. A.; Furche, F.; Prescher, J. A. "Brominated Luciferins are Versatile Bioluminescent Probes." *ChemBioChem*, **2016**, 18, 96.
4. Steinhardt, R. C.; O'Neill, J. M.; Rathbun, C. M.; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. "Design and Synthesis of an Alkynyl Luciferin Analogue for Bioluminescence Imaging." *Chem. Eur. J.*, **2016**, 22, 3671.
3. Chen, I. H.; Kou, K. G. M.; Le, D. N.; Rathbun, C. M.; Dong, V. M. "Recognition and Site-Selective Transformation of Monosaccharides by Using Copper(II) Catalysis." *Chem. Eur. J.*, **2014**, 20, 5013.
2. Lutz, J. P.; Rathbun, C. M.; Stevenson, S. M.; Powell, B. M.; Boman, T. S.; Baxter, C. E.; Zona, J. M.; Johnson, J. B. "Rate-Limiting Step of the Rh-Catalyzed Carboacylation of Alkenes: C-C Bond Activation or Migratory Insertion?" *J. Am. Chem. Soc.*, **2012**, 134, 715.
1. Rathbun, C. M.; Johnson, J. B. "Rhodium-Catalyzed Acylation of Quinolinyl Ketones: Carbon-Carbon Single Bond Activation as the Turnover Limiting Step of Catalysis." *J. Am. Chem. Soc.*, **2011**, 133, 2031.

### PRESENTATIONS

Rathbun, C. M. *Engineered luciferase-luciferin pairs for multicomponent bioluminescence imaging*. Janelia Protein Engineering Workshop, Janelia, Ashburn, VA, March 4–8, 2018. (poster and oral)

Rathbun, C. M. *Engineered luciferase-luciferin pairs for multicomponent bioluminescence imaging*. Gordon Research Conference: Bioorganic Chemistry, Andover, NH, June 11–16, 2017. (poster)

Rathbun, C. M. *Using the firefly to illuminate cancer*. Grad Slam Finals Competition (T.E.D.-style talk), U.C. Irvine, April 11, 2017. (oral)

Rathbun, C. M. *Constructing new bioluminescent tools with minimally perturbed luciferins*. Vertex Day, U.C. Irvine, March 11, 2016. (oral)

Rathbun, C. M. *Constructing new bioluminescent tools with minimally perturbed luciferins*. ACS National Meeting, San Diego, CA, March 14–17, 2016. (oral)

Rathbun, C. M. *Mechanistic study of a metal-catalyzed, regioselective carbohydrate acylation reaction*. ACS National Meeting, Indianapolis, IN, September 9–12, 2013. (poster)

## TEACHING EXPERIENCE

### University of California, Irvine, Irvine, CA

*Mentor to Undergraduate Researchers* **January 2015 – August 2017**

- Oversaw undergraduate Yuhang Yang in synthesis and cross-coupling of brominated luciferins.
- Mentored Yusef Ibrehith in screening a luciferase library for deep mutational scanning analysis.

*Graduate Chemical Biology Teaching Assistant* **January – March 2017**

- Lead weekly workshops and helped draft problem sets and exams for a graduate-level chemical biology course lead by Jennifer Prescher.
- Participated in **GetFIT** program for prospective faculty, which included giving a lecture and soliciting feedback from a faculty mentor.

*Organic Chemistry Teaching Assistant* **September 2012 – May 2013**

- Oversaw undergraduates in general and organic chemistry labs. Lead workshops for organic chemistry lecture classes.

### Hope College, Holland, MI

*Teaching Assistant* **2009 – 2012**

- Oversaw undergraduates in general and organic chemistry labs. Lead workshops for organic and physical chemistry lecture classes.

## PROGRAMMING LANGUAGES

**Python:** Data mining and analysis with Numpy. Algorithms for parallel processing with supercomputing clusters.

**HTML/CSS:** Developed static lab group website from the ground-up. Implemented liquid templating engine for easy future updates.

**L<sup>A</sup>T<sub>E</sub>X:** Used for all major long-form graduate school documents.

**Java:** Working knowledge.

**Basic knowledge:** JavaScript, Unix Shell, Ruby.

## LANGUAGES

**English** (native speaker)

**Spanish:** Conversational. Intermediate classes in college as well as 2011 summer research experience in Buenos Aires, Argentina.

## RECREATIONAL INTERESTS

- Enjoy cycling, rock climbing, and hiking.
- Homebrewing and cooking.
- Keyboardist and vocalist.



# ABSTRACT OF THE DISSERTATION

Multicomponent bioluminescence imaging via substrate resolution

By

Colin Michael Rathbun

Doctor of Philosophy in Chemistry

University of California, Irvine, 2018

Professor Jennifer Prescher, Chair

Genetically-encoded fluorescent probes have revolutionized our understanding of biological systems. However, the transition of fluorescent probes *in vivo* has been hampered by the opacity of tissue and its propensity for autofluorescence. A complementary imaging technology, bioluminescence, does not suffer from these complications because it does not require excitation light. Thus, the technique is exquisitely sensitive-with the ability to see as few as ten cells in a mouse. Bioluminescence relies on luciferase enzymes that catalyze the oxidation of small-molecule substrates (luciferins), releasing photons of light in the process. Unfortunately, the optimal luciferases for *in vivo* use rely on the same luciferins, precluding studies of more than one feature at a time.

To address this issue, I developed and analyzed a number of new luciferin probes, and created a selection platform to find mutually orthogonal luciferases and luciferins for multicomponent imaging. In contrast to the spectroscopic resolution of fluorescent tools, these probes were designed to exhibit substrate resolution. Combining luciferin analogs and mutant enzymes, we tested 20 luciferins with 207 luciferases, generating 4,140 enzyme-substrate combinations, and thus a potential for more than 4 million possible sets. Since it would be impractical to evaluate these manually, I derived a mathematical quantification of orthogonality to score each potential pairing. Next, I wrote a supercomputer algorithm to search this dataset

for the highest-scoring pairs. The software provided a ranked list of mutually orthogonal enzyme-substrate pairs that were biochemically verified. Resolution was maintained when these probes were moved into mouse models, highlighting the speed and accuracy of my approach.

My most recent work focused on increasing the practicality of these tools for preclinical imaging. The major drawbacks of our approach included temporal resolution and background emission. I addressed these issues by utilizing traditional spectral unmixing algorithms to deconvolute substrate signals mathematically. This enabled sequential imaging of substrates, and the ability to resolve smaller numbers of cells. With a top orthogonal pair I “unmixed” gradients of mutant luciferases in bacterial lysate and resolved mixtures of these mutants in mouse tumor models. These tools and algorithm will be useful for visualizing multiple cell types in mouse models.

# Chapter 1

## Introduction

Imaging tools enable researchers to see inside tissues and cells and monitor biological features in real time. While powerful, most of these probes are confined to monitoring cellular behaviors at the micro scale, in culture dishes and on slides. Visualizing cellular behaviors in more authentic environments requires tools that can function across larger spatial and time scales.<sup>1</sup> Few probes fit the bill, considering the demands placed on biocompatibility and sensitivity in whole tissues and animals. Consequently, basic questions regarding multicellular interactions in immune function, cell migration, and other physiological processes remain unanswered.

Bioluminescent probes can address the need for sensitive imaging on the macro scale. These tools derive from naturally glowing organisms (e.g., fireflies). All bioluminescent species produce light via the luciferase-catalyzed oxidation of a small molecule luciferin. Luciferase enzymes and luciferin substrates can be imported into diverse cell types and engineered to report on biological processes.<sup>2</sup> The bioluminescent signal is inherently weak (especially when compared to conventional fluorescence imaging), but there is virtually no background emission. Fluorescence imaging, by contrast, relies on light-based excitation sources that

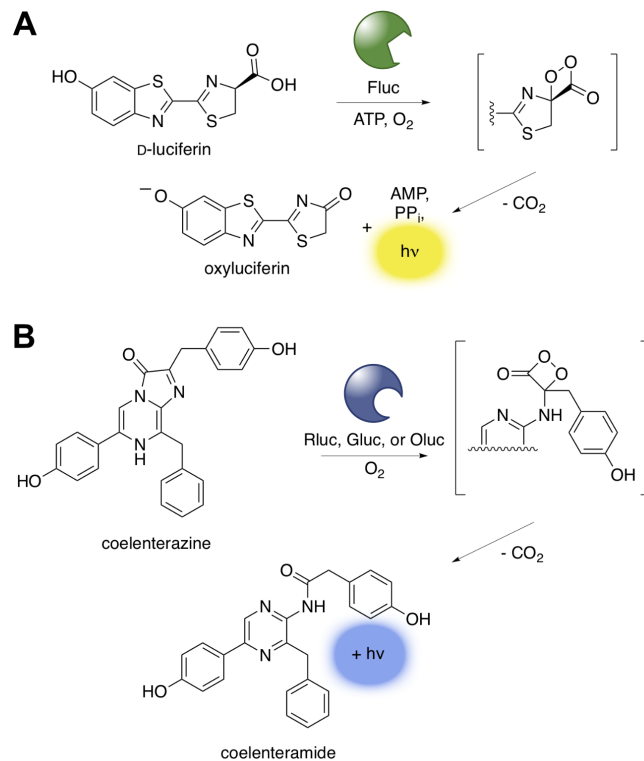
can induce tissue autofluorescence and result in poor signal-to-noise ratios. Because bioluminescence requires no excitation light, it can enable exquisitely sensitive imaging even in heterogeneous tissues. In fact, bioluminescent probes can be preferred to fluorescent tools for long-term cell tracking in rodent and other opaque models. One can serially image luciferase reporters without detriment to organisms and without knowing when and where to apply excitation light.

The versatility of bioluminescence has enabled a broad range of biological studies, although limitations persist.<sup>2</sup> This imaging modality has long been plagued by a lack of bright, easily distinguishable probes and poor spatial resolution. However, advances in luciferin chemistry and luciferase engineering have begun to address these issues. This perspective will highlight recent achievements in developing new and improved imaging tools. Collectively, the bioluminescent probes are addressing long-standing voids in imaging capabilities and are being applied to visualizing biology beyond the culture dish.

## 1.1 Bioluminescence basics

Millions of luciferases exist in the natural world, but phylogenetically related enzymes use the same luciferin.<sup>3</sup> The identities of such luciferins remain exceedingly difficult to characterize, and of those reported to date, only two have found routine application in mammalian cell imaging.<sup>2</sup> The first, D-luciferin (Figure 1.1A), is used by the firefly and a number of other terrestrial organisms to produce light. The second, coelenterazine (Figure 1.1B), is a molecule found in bioluminescent sea creatures, including the sea pansy *Renilla reniformis* and deep-sea copepods *Gaussia princeps* and *Oplophorus gracilirostris*.

Despite their distinct chemical structures, D-luciferin and coelenterazine use a similar mechanism for light production: the cognate luciferases oxidize the small molecules to generate



**Figure 1.1:** Popular luciferase-luciferin pairs for cellular imaging. (A) D-luciferin is oxidized by firefly luciferase (Fluc) to produce oxyluciferin and a photon of light. (B) Coelenterazine is oxidized by a variety of marine luciferases, including Renilla luciferase (Rluc), Gaussia luciferase (Gluc), and Oplophorus luciferase (Oluc). These luciferases, unlike Fluc, require only oxygen as a cofactor in the light-emitting reaction.

excited state oxyluciferins; relaxation of these molecules to the ground state results in photon emission (Figure 1.1).

The color of light released is primarily dictated by the small molecule luciferin. In the case of D-luciferin, the aromatic core is sufficiently extended to provide yellow-green ( $\sim 560$  nm) light. For coelenterazine, the  $\pi$  system of the putative emitter is shorter, resulting in blue wavelengths ( $\sim 475$  nm) of emission. The luciferase environment can further modulate bioluminescent color. Thus, enzymes that use the exact same substrate can emit different wavelengths. In some cases, the emission spectra are sufficiently resolved to enable multicolor imaging.<sup>4</sup> Discriminating among related luciferases *in vivo*, though, remains difficult because of their broad emission spectra and the strong absorption of  $>650$  nm light.<sup>5</sup>

## 1.2 Building better bioluminescent tools

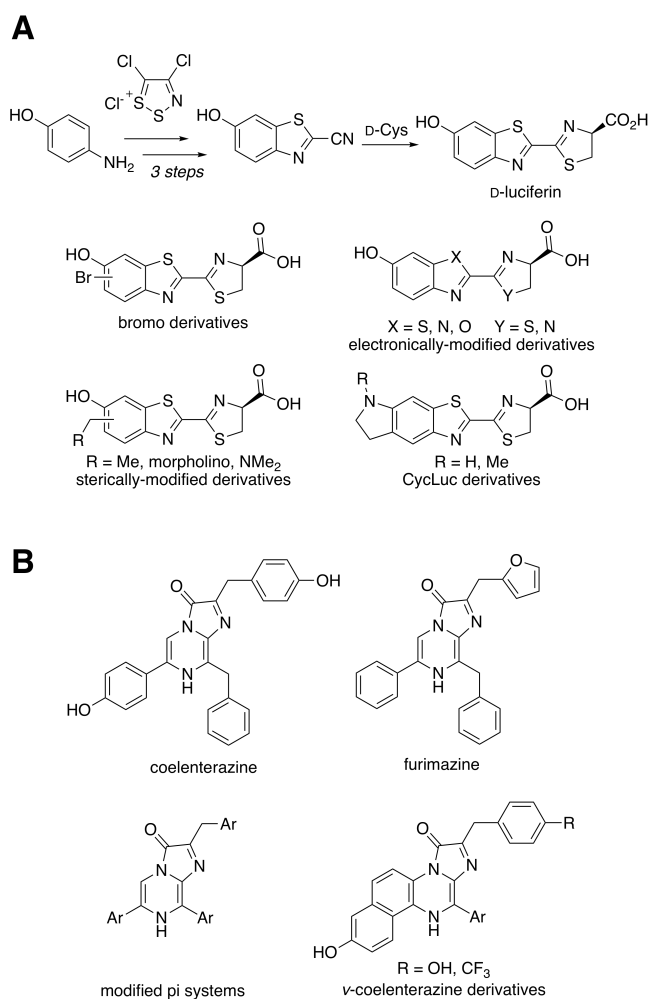
Luciferases and luciferins have been used for decades to image biomolecules, gene expression, and even whole cells, but the most popular probes are not without limitation.<sup>2</sup> The poor tissue penetrance of many luciferins has historically precluded sensitive imaging in hard-to-access tissues (e.g., brain). Additionally, only a limited palette of bioluminescent probes exists, hindering efforts to visualize multicellular or other multicomponent processes. Chemical and biological approaches are beginning to address the need for improved and more diverse collections of bioluminescent probes.

### 1.2.1 Streamlined and scalable luciferin syntheses

Most bioluminescence imaging studies require an exogenously delivered luciferin. For *in vivo* work, the doses can range between 1 and 5 mg per mouse per imaging point. Thus, large—and often prohibitively expensive—quantities of luciferin are required for animal studies.

Accessing luciferins in bulk has been a historic challenge. The chemical complexity of D-luciferin and coelenterazine has also frustrated efforts to rapidly diversify these scaffolds to produce new analogues.

In recent years, improved methods for producing luciferins have been reported. In the case of D-luciferin, we developed a streamlined method for accessing the requisite benzothiazole core using Appel's salt and modern C-H activation (Figure 1.2).<sup>6</sup> This approach improved the



**Figure 1.2:** Recently synthesized luciferin heterocycles. (A) D-luciferin and related analogues can be accessed using Appel's salt and C-H activation chemistry. (B) A diverse array of coelenterazine analogues have been synthesized in recent years.

yield and generalizability of the synthesis, eliminating the need for multiple functional group conversions used in previous routes. The strategy has enabled multigram syntheses of D-

luciferin and has proven to be quite modular. In fact, we and others have used this chemistry to produce more than 30 distinct analogues.<sup>7</sup> Many of these scaffolds also comprise key functional handles for late stage diversification (Figure 1.2). Efforts to expediently synthesize amino luciferins<sup>8</sup> and related analogues<sup>9</sup> are similarly facilitating more widespread use of bioluminescent tools.

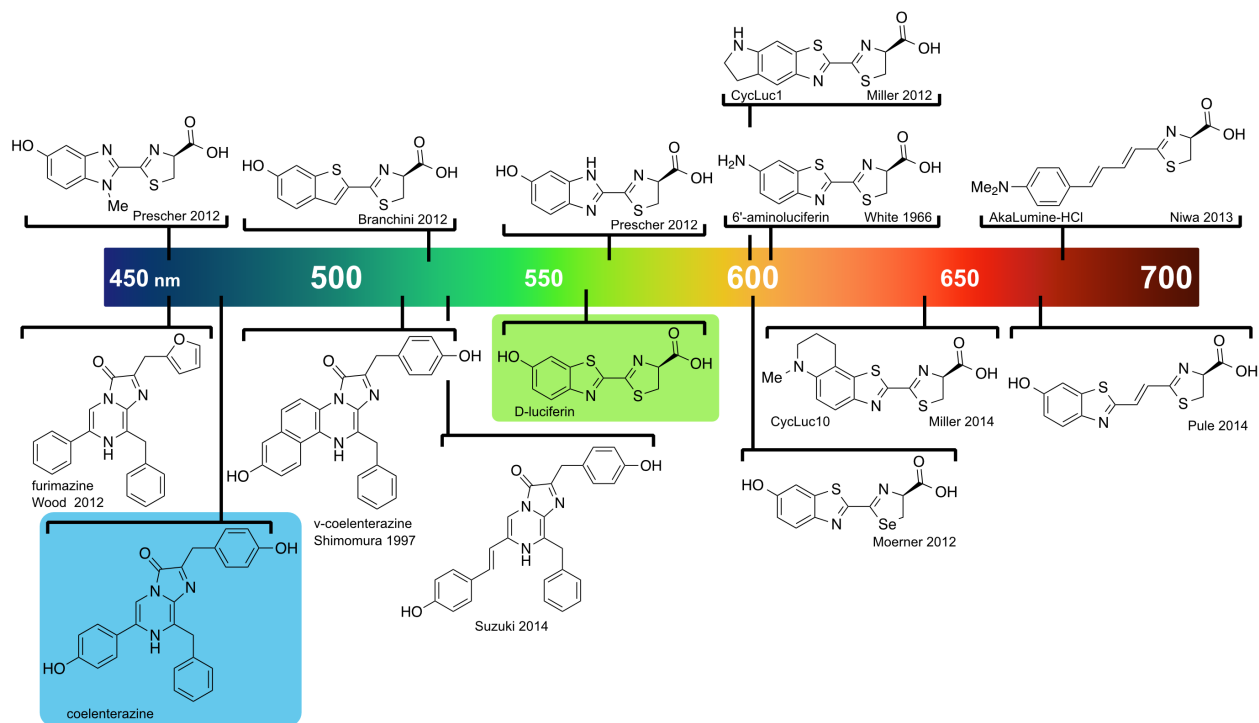
Improved methods for accessing coelenterazine analogues are also advancing imaging studies. Historically, coelenterazine synthesis relied on condensation of glyoxal derivatives with strong acid and elevated temperatures. These conditions are incompatible with a diverse array of desired analogues. Kirkland and colleagues identified milder conditions to access the coelenterazine core via Horner-Wadsworth-Emmons olefination.<sup>10</sup> The method proved to be both scalable and generalizable, providing a range of useful analogues. Recent advances in cross-coupling methodology are similarly enabling access to diverse coelenterazine architectures.<sup>11</sup>

### 1.2.2 Brighter and more colorful probes

Bioluminescence imaging to date has been largely confined to imaging one or two colors at a time. Accessing a full spectrum of emitters has been a long-standing goal, with obvious parallels to fluorescence imaging. Early efforts to modulate color and brightness primarily focused on the luciferase enzyme.<sup>2</sup> Recent years, though, have seen a shift in focus to modulating the small molecule (Figure 1.3).

Because the structure of the luciferin emitter influences bioluminescent output, methods for modulating color and intensity often parallel those in small molecule fluorophore development. For example, rigidifying the aromatic core can provide improved push-pull chromophores and enhanced light emission. Miller and coworkers applied this strategy to D-luciferin, replacing the rotationally flexible 6' electron-donating group with a conforma-

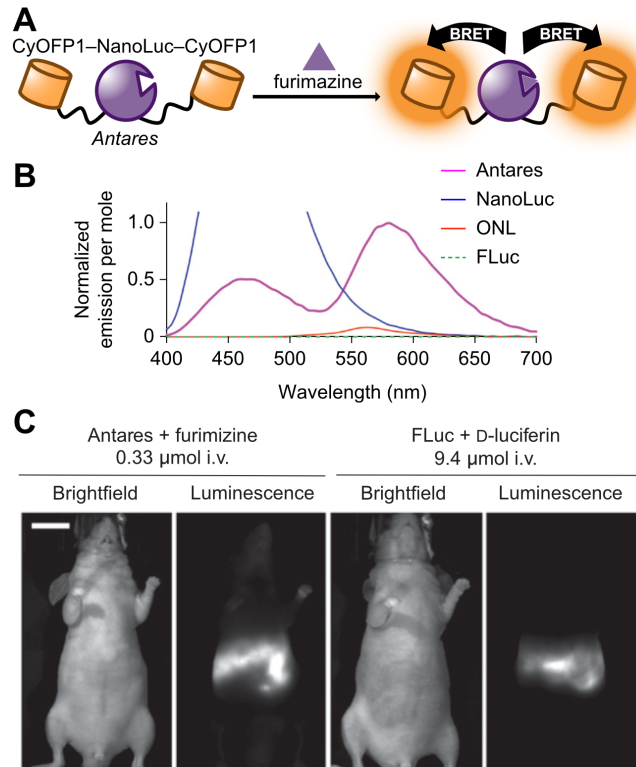




**Figure 1.3:** Palette of luciferin analogues. Native luciferins are highlighted with colored boxes. Brackets denote the wavelength (nanometers) of maximal bioluminescence emission observed upon incubation of the compound with luciferase. While many analogues can provide unique colors of light, most are not efficiently processed by native luciferases.

tionally locked amine.<sup>12</sup> This configuration enabled enhanced donation of electrons to the luciferin core and more robust, red-shifted emission compared to that of related probes.<sup>13</sup> Moreover, many of the cyclic luciferin (CycLuc) scaffolds were viable substrates for Fluc, highlighting the promiscuity of the enzyme. Fluc can also process luciferin analogues with extended aromatic cores and modified heteroatoms, providing additional avenues for wavelength alteration.<sup>6,7,14</sup>  $\pi$ -Extended chromophores typically emit red-shifted light emission.<sup>15</sup> In the case of the vinyl-extended luciferins (Figure 1.3), the shift was >100 nm. It should be noted, though, that these and other extended analogues are often poorly processed by Fluc itself; red-shifted emission comes at the expense of enzyme turnover and thus total photon output. Parallel developments in coelenterazine synthesis have provided scaffolds that emit different colors of light. Red-shifted probes are particularly desirable for imaging with Rluc and Gluc *in vivo*, as the normal blue emission with these enzymes is strongly absorbed by blood.<sup>5</sup> Inouye and colleagues recently developed a streamlined synthesis of a conformationally locked coelenterazine for improved imaging.<sup>11</sup> Others have explored luminophores with extended  $\pi$ -conjugation to achieve more tissue-penetrant wavelengths of light.<sup>16,17</sup> Bioluminescent color and photon output can also be drastically modulated via bioluminescent resonance energy transfer (BRET). A recent fusion of NanoLuc (an optimized mutant luciferase, discussed below) to CyOFP1 [dubbed Antares (Figure 1.4A)] resulted in an emission shift of 114 nm (Figure 1.4B).<sup>13</sup>

Antares is one of the most red-shifted BRET pairs reported to date, enabling enhanced deep-tissue imaging *in vivo* (Figure 1.4C). In related work, Nagai and colleagues linked a mutant Renilla luciferase to Venus, a version of yellow fluorescent protein.<sup>18</sup> Unexpectedly, this construct (dubbed Nano-lantern) produced 6-fold more photons than the luciferase alone and shifted emission into the yellow-green region (530 nm). Additional Nano-lantern colors have been described.<sup>4</sup> Collectively, these and other BRET probes are enabling sensitive imaging of many biological processes, including membrane voltage<sup>19</sup> and gene expression.<sup>20</sup>



**Figure 1.4:** BRET for deep-tissue imaging. (A) Antares comprises a fusion of NanoLuc with two cyan-excitable orange fluorescent proteins (CyOFP1). (B) The BRET fusion exhibits red-shifted emission relative to NanoLuc. ONL denotes Orange Nano-lantern. (C) Antares enabled sensitive imaging with low luciferin levels in mice. Antares and Fluc genes were delivered to the liver via hydrodynamic injection. Luciferins were supplied intravenously. Adapted with permission from ref.<sup>13</sup>

### 1.2.3 In vivo improvements

Luciferins can access most tissues, but their distribution and cell permeability are suboptimal and not uniform.<sup>21</sup> Chemical tinkering and med-chem optimization can improve the bioavailabilities of the substrates. For example, many of the CycLuc scaffolds that comprise more lipophilic amino modifications (vs -OH groups) exhibit improved tissue penetrance in vivo.<sup>22</sup> Far less compound is required for a given imaging session compared to native luciferins. Further amidation of a CycLuc probe enhanced its transport across the blood-brain barrier (Figure 5).<sup>23</sup> Once the probe is in the brain, the amide bond can be cleaved by endogenous fatty acid amide hydrolase (strongly expressed in brain tissue), releasing a functional luciferin.

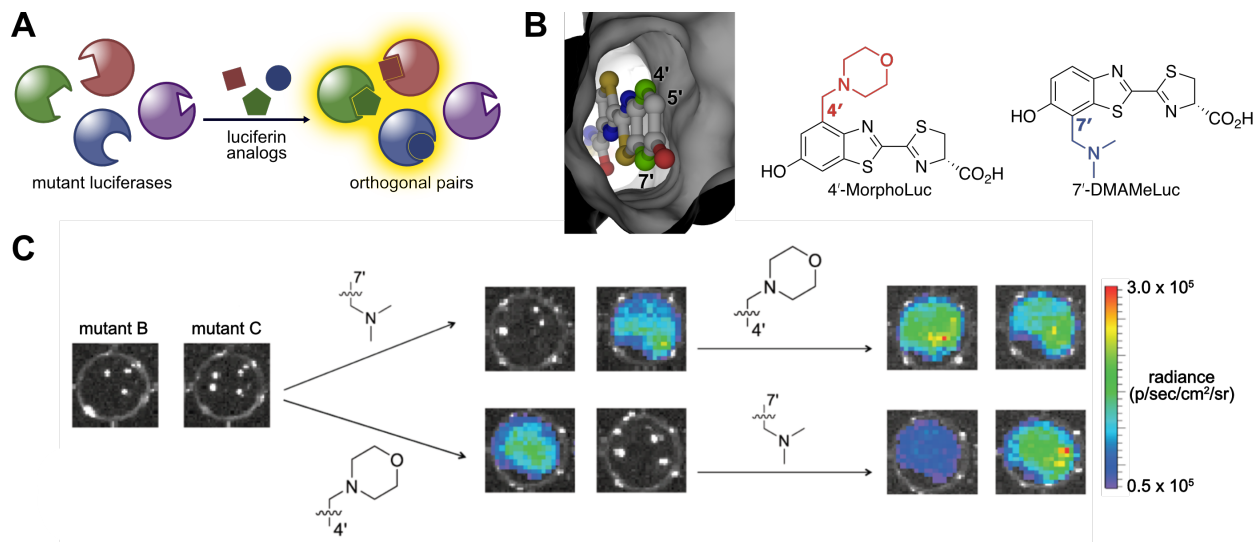
## 1.3 Luciferase-luciferin pairs for multicellular imaging

Many improvements to the luciferin small molecules noted above came at the expense of luciferase turnover (and thus light output). Consequently, recent efforts to build improved bioluminescence tools have focused on identifying new substrates and enzymes in parallel. One source of new imaging tools is nature itself. Bioluminescent organisms (and potentially new luciferase-luciferin pairs) are continually being described,<sup>24</sup> though the need for optimized probes far outpaces their discovery. Others have turned to engineering existing luciferases to better use chemically modified luciferins. In one example, Miller identified mutant luciferases that can more readily process CycLuc derivatives. These analogues were previously demonstrated to be viable substrates for Fluc, although oxyluciferin products inhibited the reaction. Substrate inhibition was relieved and long-lived light emission restored with mutant enzymes.<sup>25</sup> The authors later identified a mutant that preferred a luciferin analogue over D-luciferin, setting the stage for developing substrate-responsive enzymes.<sup>26</sup> Further work by the Miller group revealed latent luciferase activity in a fatty acyl-CoA syn-

thetase from the fruitfly.<sup>27</sup> Interestingly, this enzyme did not emit light with D-luciferin—it was able to use only a CycLuc substrate—opening the door to pairing unnatural luciferin analogues with evolutionary relatives of luciferase. Another engineered bioluminescent enzyme that has seen widespread adoption in recent years is NanoLuc, a derivative of Oplophorus luciferase (Oluc).<sup>28</sup> Early work in this area was motivated by the need for improved coelenterazines (molecules prone to autoxidation and exhibiting poor tissue penetrance) and Oluc subunits (enzyme fragments prone to instability). Seeking a brighter and more stable luciferin, Wood and colleagues replaced the electron-rich phenols of coelenterazine with phenyl and furan groups. The resulting molecule [furimazine (Figure 1.3)] was more stable in media and lysate and less susceptible to nonspecific oxidation. Directed evolution was used to select an enzyme that could readily catalyze light emission from the designer luciferin. The winning mutant (NanoLuc) contained a total of 16 mutations, an impressive number for a 16 kDa protein. NanoLuc exhibits a high turnover rate with furimazine, providing robust signal output. Such high photon flux values are enabling sensitive imaging in complex tissue samples, even point-of-care diagnostics with simple cell phone cameras.<sup>29</sup> We anticipate that the popularity of the NanoLuc-furimazine pair will continue to surge in the near term.

Tandem modification of bioluminescent enzymes and substrates is also enabling multicomponent bioluminescence imaging. In contrast to *in vitro* assays, imaging *in vivo* often precludes spectroscopic resolution of colored probes. Blood and tissue restrict the passage of wavelengths shorter than red,<sup>5</sup> and bioluminescence spectra are broad,<sup>30</sup> necessitating an alternative approach. Substrate resolution is one such strategy (Figure 1.5A).

This approach requires multiple selective, mutually orthogonal luciferin-luciferase pairs. These pairs produce light together but will not react with other mutants or enzymes. Orthogonal pairs already exist in nature (e.g., firefly and marine luciferases, along with their requisite substrates), and many have been adapted for dual-imaging studies. To expand and expedite the search for orthogonal pairs, we turned to producing non-natural analogues and



**Figure 1.5:** Orthogonal luciferin-luciferase pairs for multicomponent imaging. (A) A strategy for multicomponent bioluminescence imaging with luciferin analogues and mutant luciferases. (B) The 4' and 7' sites of D-luciferin were targeted to preclude binding to native Fluc. (C) Substrate resolution was achieved in mouse DB7 cells using 4'- and 7'-modified luciferins and mutant luciferases identified from screening. Adapted with permission from ref.<sup>31</sup>

enzymes.<sup>31</sup> We synthesized a panel of luciferins, with additional steric bulk at the 4' and 7' positions (Figure 1.5B). We screened these analogues against libraries of luciferase mutants to produce more than 3000 potential pairings. To find substrate-resolved hits in this milieu, we mined the data with a computer algorithm. This strategy revealed two enzymes that exhibited substrate resolution with two compounds (one 4' and one 7'). This pair proved to be successful in mammalian cells, as well, demonstrating the robust nature of our screening methodology (Figure 1.5C). In an analogous approach, Kim and colleagues recently reported substrate-resolved bioluminescence imaging with coelenterazine analogues.<sup>32</sup>

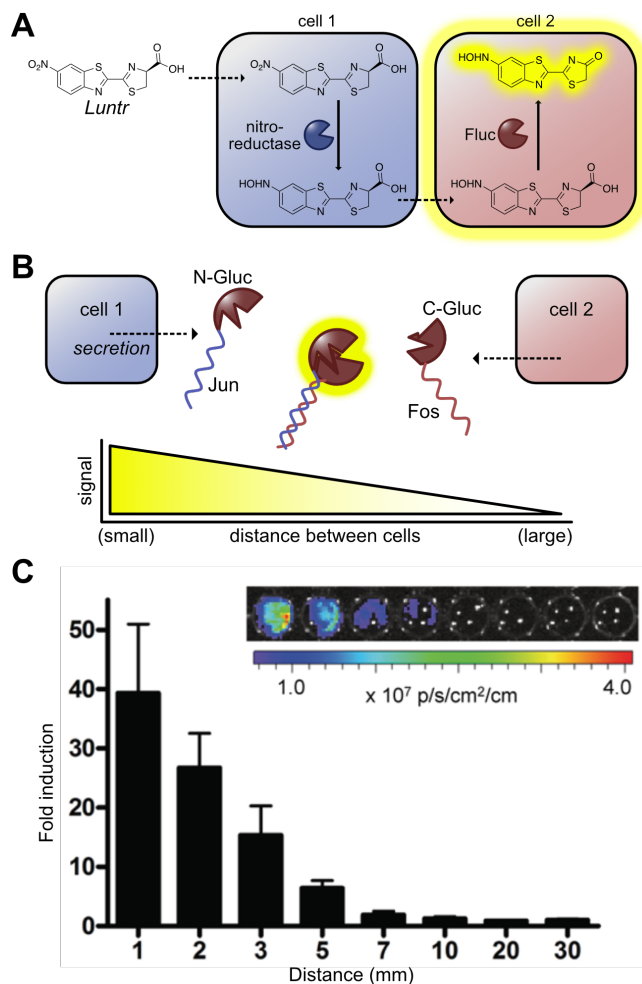
## 1.4 Bioluminescent reporters for cell–cell interactions

A second challenge being addressed by improved bioluminescent tools is monitoring cellular interactions in vivo. Conventional bioluminescence imaging can detect small numbers of

cells, but historically has lacked the spatial resolution to precisely pinpoint their locations or interactions in whole organisms. Target recognition and cell–cell contacts are crucial to numerous physiological processes, including neurotransmission, immune function, and cell migration, so methods to globally assay such interactions are necessary. The development of tools that report on cell–cell contacts has been inspired by classic methods for reporting on biomolecule activities and interactions. For example, caged probes have been used widely to report on small molecule analytes in cells. We and others<sup>33</sup> have shown that such cages can be repurposed to image interactions between cells (Figure 1.6A).

In a recent example, a dark luciferin comprising a 6'- nitro group (Luntr) was used as the cage.<sup>35</sup> This probe could be reduced by one group of cells (expressing nitroreductase) and used by a neighboring group of cells expressing luciferase. Light emission was strongest in areas of cellular contact. Nitroreductase is not expressed endogenously in mammalian cells, making this technology ripe for in vivo applications where high spatial sensitivity is required.

A second class of cell contact sensors uses split reporters, originally developed to image protein-protein interactions (Figure 1.6B). We expanded on this concept to generate split reporters of cell proximity using Gaussia luciferase (Gluc), a secreted protein that functions in the extracellular space. Split fragments of Gluc were fused to leucine zippers Fos and Jun to drive complementation. The N-terminal half was expressed in one cell population and the C-terminal half in another. Light emission in this case tracked with distance between the cell populations (Figure 1.6C).



**Figure 1.6:** Bioluminescent reporters for cellular interactions. (A) Caged luciferins can report on cell proximity between one cell (expressing an uncaging enzyme) and a second cell (expressing luciferase). In one example, Luntr enabled proximity-dependent imaging with nitroreductase- and Fluc-expressing cells. (B) Split luciferase constructs can be used to report on cell-cell interactions. The closer the two cells, the greater the light output. (C) Split Gaussia luciferase can report on cell proximity in vitro. As the distance between the cells increased, light emission decreased. Adapted with permission from ref.<sup>34</sup>



## 1.5 Bioluminescent reporters for analyte detection in vivo

Bioluminescence has long been exploited for detecting enzyme activities and low-abundance metabolites. The majority of these studies, though, have been limited to ex vivo analyses with cultured cells or excised tissues. Continued advances in bioluminescence technology are enabling new probes to be applied as biosensors in vivo.<sup>36,37</sup> In a recent example, Chang and colleagues developed a copper ion sensor for imaging in mice. The probe comprised a caged luciferin, with a bulky chelator group (i.e., the cage) attached to the 6' position of D-luciferin. The sterically encumbered molecule was poorly utilized by luciferase. Upon removal of the cage by copper-(I)-dependent oxidative cleavage, a viable luciferin was liberated and available for light emission. Photon production could thus be correlated to copper ion levels. The caged probe was ultimately used for analyte imaging in mouse models of fatty liver disease.<sup>36</sup>

## 1.6 Conclusions and future directions

Bioluminescence has historically lagged behind fluorescence imaging in terms of the breadth and diversity of available tools. Recent advances in luciferin chemistry and luciferase engineering, though, are beginning to fill this gap. New synthetic methods are providing novel luciferin architectures for improved imaging. Engineered luciferases are enabling the sensitive detection of cells and other analytes in vivo. Combinations of designer substrates and mutant enzymes are furthering the range of potential applications. It is now possible to image multicellular features in live animals, visualize cells in difficult-to-access tissues (e.g., brain tissue), and selectively illuminate cell-cell interactions. Moving forward, we anticipate continued advances in red-shifted probes, tandem luciferase-luciferin engineering, and sensors for cellular metabolites. These tools will influence how researchers conduct experiments

involving multiple cell types and molecular features beyond the culture dish. Additionally, like other useful imaging agents, the tools will likely facilitate discoveries in a diverse range of fields.

## 1.7 Objectives of this study

The power of bioluminescence for sensitive, in vivo cell tracking makes it an excellent tool for preclinical imaging. As discussed above, this has made bioluminescence an area of intense recent interest. Many groups are focused on improvements to (1) the bioavailability<sup>21,23</sup> and emission wavelength<sup>38,39</sup> of the small molecule or (2) the emission wavelength of the luciferase (or constructs thereof).<sup>13,30</sup> At the outset of our work, few were interested in tandem modification of both the luciferin and the luciferase. Further, while most groups focused on the wavelength of emission (to derive both selectivity and brightness), few had studied substrate discrimination. We sought to produce new luciferin analogs and pair them with matched luciferase mutants. This strategy could enable multicomponent imaging in the most complex tissue environments.

I aimed to:

1. Produce libraries of mutant luciferases and identify mutants that utilize luciferin analogs selectively.
2. Develop tools for evaluating orthogonal pairs and identifying new pairs.
3. Synthesize and test new luciferin analogs.
4. Discriminate substrate-resolved bioluminescent probes in vitro and in vivo.

# Bibliography

- [1] Porterfield, W. B.; Prescher, J. A. *Curr. Opin. Chem. Biol.* **2015**, *24*, 121–130.
- [2] Paley, M. A.; Prescher, J. A. *Med. Chem. Comm.* **2014**, *5*, 255–267.
- [3] Martini, S.; Haddock, S. H. D. *Sci. Rep.* **2017**, *7*, 45750.
- [4] Suzuki, K.; Kimura, T.; Shinoda, H.; Bai, G.; Daniels, M. J.; Arai, Y.; Nakano, M.; Nagai, T. *Nat. Commun.* **2016**, *7*, 13718.
- [5] Zhao, H.; Doyle, T. C.; Coquoz, O.; Kalish, F.; Rice, B. W.; Contag, C. H. *J. Biomed. Opt.* **2005**, *10*, 041210.
- [6] McCutcheon, D. C.; Paley, M. A.; Steinhardt, R. C.; Prescher, J. A. *J. Am. Chem. Soc.* **2012**, *134*, 7604–7607.
- [7] Woodroffe, C. C.; Meisenheimer, P. L.; Klaubert, D. H.; Kovic, Y.; Rosenberg, J. C.; Behney, C. E.; Southworth, T. L.; Branchini, B. R. *Biochemistry* **2012**, *51*, 9807–9813.
- [8] Hauser, J. R.; Beard, H. A.; Bayana, M. E.; Jolley, K. E.; Warriner, S. L.; Bon, R. S. *Beilstein J. Org. Chem.* **2016**, *12*, 2019–2025.
- [9] Anderson, J. C.; Grounds, H.; Jathoul, A. P.; Murray, J. A. H.; Pacman, S. J.; Tisi, L. *RSC Adv.* **2017**, *7*, 3975–3982.
- [10] Shakhmin, A.; Hall, M. P.; Walker, J. R.; Machleidt, T.; Binkowski, B. F.; Wood, K. V.; Kirkland, T. A. *Chem.—Eur. J.* **2016**, *22*, 10369–10375.

- [11] Hosoya, T.; Iimori, R.; Yoshida, S.; Sumida, Y.; Sahara-Miura, Y.; Sato, J.-i.; Inouye, S. *Org. Lett.* **2015**, *17*, 3888–3891.
- [12] Reddy, G. R.; Thompson, W. C.; Miller, S. C. *J. Am. Chem. Soc.* **2010**, *132*, 13586–13587.
- [13] Chu, J. et al. *Nat. Biotechnol.* **2016**, *34*, 760–767.
- [14] Conley, N. R.; Dragulescu Andrasi, A.; Rao, J.; Moerner, W. E. *Angew. Chem. Int. Ed.* **2012**, *51*, 3350–3353.
- [15] Kuchimaru, T.; Iwano, S.; Kiyama, M.; Mitsumata, S.; Kadonosono, T.; Niwa, H.; Maki, S.; Kizaka-Kondoh, S. *Nat. Commun.* **2016**, *7*, 11856.
- [16] Nishihara, R.; Suzuki, H.; Hoshino, E.; Suganuma, S.; Sato, M.; Saitoh, T.; Nishiyama, S.; Iwasawa, N.; Citterio, D.; Suzuki, K. *Chem. Commun.* **2014**, *51*, 391–394.
- [17] Grinstead, K. M.; Rowe, L.; Ensor, C. M.; Joel, S.; Daftarian, P.; Dikici, E.; Zingg, J.-M.; Daunert, S. *PLoS ONE* **2016**, *11*, e0158579.
- [18] Saito, K.; Chang, Y.-F.; Horikawa, K.; Hatsugai, N.; Higuchi, Y.; Hashida, M.; Yoshida, Y.; Matsuda, T.; Arai, Y.; Nagai, T. *Nat. Commun.* **2012**, *3*, 1262.
- [19] Inagaki, S.; Tsutsui, H.; Suzuki, K.; Agetsuma, M.; Arai, Y.; Jinno, Y.; Bai, G.; Daniels, M. J.; Okamura, Y.; Matsuda, T.; Nagai, T. *Sci. Rep.* **2017**, *7*, 42398.
- [20] Horikiri, T.; Ohi, H.; Shibata, M.; Ikeya, M.; Ueno, M.; Sotozono, C.; Kinoshita, S.; Sato, T. *PLoS ONE* **2017**, *12*, e0170342.
- [21] Adams, S. T.; Miller, S. C. *Curr. Opin. Chem. Biol.* **2014**, *21*, 112–120.
- [22] Evans, M. S.; Chaurette, J. P.; Adams, S. T.; Reddy, G. R.; Paley, M. A.; Aronin, N.; Prescher, J. A.; Miller, S. C. *Nat. Methods* **2014**, *11*, 393–395.

- [23] Mofford, D. M.; Adams, S. T.; Reddy, G. S. K. K.; Reddy, G. R.; Miller, S. C. *J. Am. Chem. Soc.* **2015**, *137*, 8684–8687.
- [24] Purto, K. V. et al. *Angew. Chem., Int. Ed.* **2015**, *54*, 8124–8128.
- [25] Harwood, K. R.; Mofford, D. M.; Reddy, G. R.; Miller, S. C. *Chemistry & Biology* **2011**, *18*, 1649–1657.
- [26] Mofford, D. M.; Reddy, G. R.; Miller, S. C. *J. Am. Chem. Soc.* **2014**, *136*, 13277–13282.
- [27] Mofford, D. M.; Reddy, G. R.; Miller, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 4443–4448.
- [28] Hall, M. P. et al. *ACS Chem. Biol.* **2012**, *7*, 1848–1857.
- [29] Griss, R.; Schena, A.; Reymond, L.; Patiny, L.; Werner, D.; Tinberg, C. E.; Baker, D.; Johnsson, K. *Nat. Chem. Biol.* **2014**, *10*, 598–603.
- [30] Rumyantsev, K. A.; Turoverov, K. K.; Verkhusha, V. V. *Sci. Rep.* **2016**, *6*, 36588.
- [31] Jones, K. A.; Porterfield, W. B.; Rathbun, C. M.; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. *J. Am. Chem. Soc.* **2017**, *139*, 2351–2358.
- [32] Nishihara, R.; Abe, M.; Nishiyama, S.; Citterio, D.; Suzuki, K.; Kim, S. B. *Sci. Rep.* **2017**, *7*, 908.
- [33] Lindberg, E.; Mizukami, S.; Ibata, K.; Miyawaki, A.; Kikuchi, K. *Chem.—Eur. J.* **2013**, *19*, 14970–14976.
- [34] Jones, K. A.; Li, D. J.; Hui, E.; Sellmyer, M. A.; Prescher, J. A. *ACS Chem. Biol.* **2015**, *10*, 933–938.
- [35] Porterfield, W. B.; Jones, K. A.; McCutcheon, D. C.; Prescher, J. A. *J. Am. Chem. Soc.* **2015**, *137*, 8656–8659.

- [36] Heffern, M. C.; Park, H. M.; Au-Yeung, H. Y.; Van de Bittner, G. C.; Ackerman, C. M.; Stahl, A.; Chang, C. J. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 14219–14224.
- [37] Sekar, T. V.; Foygel, K.; Massoud, T. F.; Gambhir, S. S.; Paulmurugan, R. *Sci. Rep.* **2016**, *6*, 584.
- [38] Jathoul, A. P.; Grounds, H.; Anderson, J. C.; Pule, M. A. *Angew. Chem., Int. Ed.* **2014**, *53*, 13059–13063.
- [39] Iwano, S.; Obata, R.; Miura, C.; Kiyama, M.; Hama, K.; Nakamura, M.; Amano, Y.; Kojima, S.; Hirano, T.; Maki, S.; Niwa, H. *Tetrahedron* **2013**, *69*, 3847–3856.

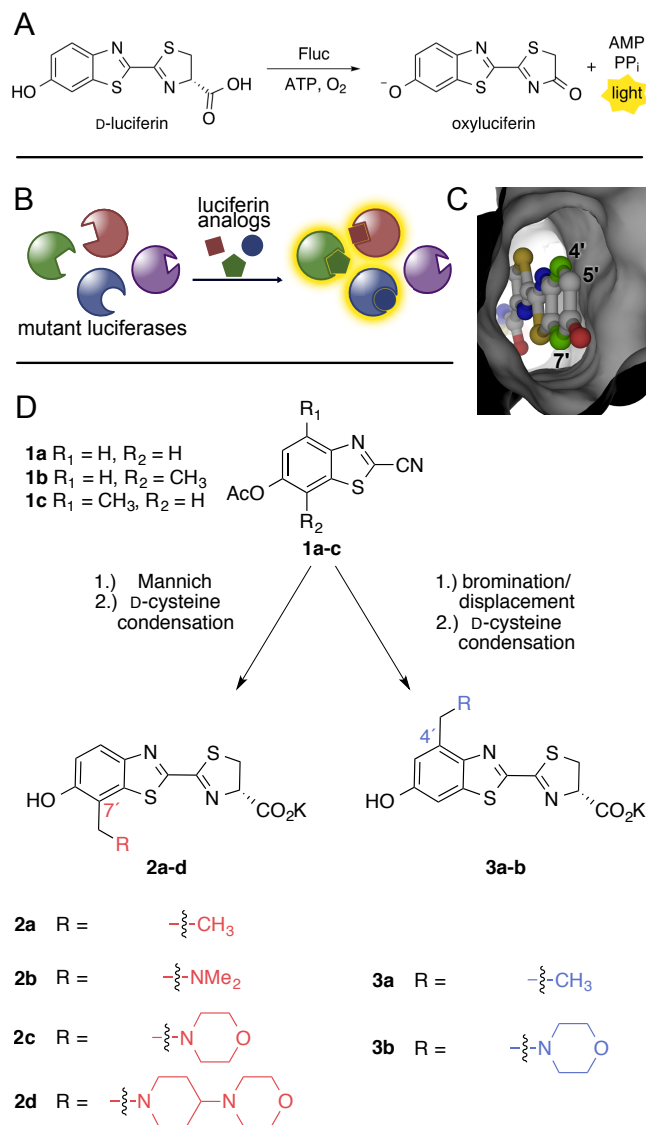
# Chapter 2

## Orthogonal luciferase-luciferin pairs for bioluminescence imaging

### 2.1 Introduction

Bioluminescence imaging is a popular method for visualizing cells and other biological features *in vivo*.<sup>1</sup> This technology relies on enzymes (luciferases) that catalyze the oxidation of small-molecule substrates (luciferins). The oxidation process is accompanied by the release of light (Figure 3.1A). Since mammalian cells and tissues do not emit substantial numbers of photons, bioluminescent light can facilitate sensitive imaging in these environments.<sup>2</sup> Luciferase-labeled cells can also be imaged repeatedly and noninvasively in a variety of pre-clinical models. This broad dynamic range has enabled numerous studies of fundamental biological processes, including cell homing and differentiation, proliferation, and cell-to-cell communication, in physiologically relevant environments.<sup>3</sup>

While versatile, bioluminescence to date has been largely limited to monitoring one cell type or biological feature at a time. This is due, in part, to a lack of distinguishable luciferase-



**Figure 2.1:** 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.



luciferin pairs for in vivo use. The optimal luciferases (from the insect family) use the same substrate, D-luciferin.<sup>1,4</sup> 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 analogues have been engineered to emit different colors of light.<sup>5-7</sup> 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.<sup>8-10</sup> 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.<sup>11-13</sup> Some of these pairs, including those from the firefly and *Renilla reniformis*, have been used extensively.<sup>1,3,14,15</sup> Firefly (Fluc) and *Renilla* luciferase employ chemically unique substrates (D-luciferin and coelenterazine, respectively), enabling their tandem application in vivo.<sup>16,17</sup> Coelenterazine is less ideal for use in these environments, though, owing to its suboptimal bioavailability and stability.<sup>1,18</sup> Other naturally occurring luciferases and luciferins can be used in combination with Fluc/D-luciferin or other bioluminescent systems.<sup>17,19</sup> 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 analogues of D-luciferin.<sup>20</sup> In elegant work along these lines, Miller and co-workers prepared a class of non-natural aminoluciferins that were found to be robust light emitters with Fluc, but the products inhibited the enzymatic reaction.<sup>21</sup> Product inhibition was relieved using

mutated versions of the enzyme.<sup>22</sup> These same mutations also resulted in sharply reduced emission with D-luciferin, providing key precedent for the development and utilization of orthogonal pairs.<sup>23</sup> The mutant enzymes from these studies, though, were less selective for one analogue 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*.<sup>24,25</sup> In both cases, altered bioluminescent outputs (e.g., colors and stabilities) were achieved, but orthogonal substrate usage was not realized.

Here, we report a 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 analogues with libraries of mutant luciferases and identified substrate-selective enzymes. The hits were also biochemically characterized. Importantly, when the mutants and analogues were combined, robust light production was observed 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 enable a variety of multicellular imaging applications. Importantly, our approach to identifying orthogonal bioluminescence pairs is also general and should enable rapid diversification of the bioluminescence toolkit.

## **2.2 Results and discussion**

### **2.2.1 Designing and constructing sterically modified luciferins**

To expediently identify orthogonal bioluminescence tools, we aimed to screen sterically perturbed luciferins against libraries of mutant luciferases (Figure 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 nontoxicity of the reagents and bioavailability of the substrate.<sup>26,27</sup> Second, the Fluc/D-luciferin reaction releases the highest percentage of tissue-penetrating light among known bioluminescent families.<sup>28</sup> 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.<sup>13,29–32</sup> Finally, D-luciferin derivatives are arguably the most synthetically tractable luciferin architectures.<sup>33,34</sup>

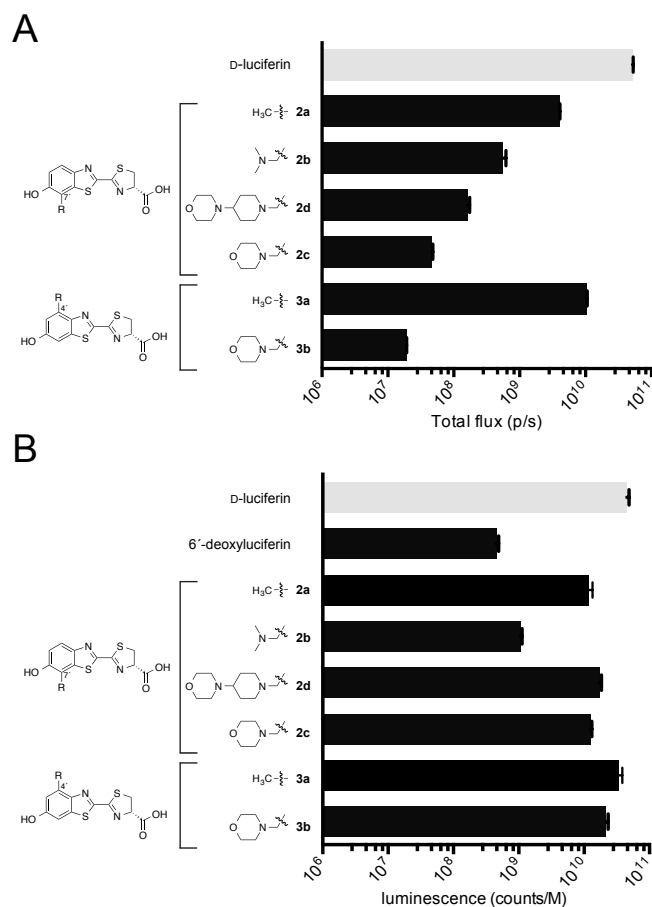
Generating an expanded set of bioluminescent tools required access to diverse luciferin scaffolds. A variety of D-luciferin analogues have been synthesized over the past four decades,<sup>5,7,35–39</sup> and those capable of robust emission with Fluc harbor common features: an electron-donating group at the 6' position, a carboxylate appendage (for adenylation), and an abstractable proton  $\alpha$  to the carboxylate.<sup>36,40</sup> Beyond these requirements, Fluc can tolerate a surprisingly large variety of modified luciferins,<sup>41–43</sup> including 6'-amino substituents,<sup>20,21,42</sup> alkylated<sup>44–46</sup> and acylated<sup>47</sup> scaffolds, and even luciferins with non-natural chromophores.<sup>6,47</sup> 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.<sup>31,32</sup>

Unlike most efforts to produce luciferin analogues 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 (Figure 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 analogues with small (e.g., 2-3 atoms) substituents at C4' and C7' could effectively access the active site.

Generating 4'- and 7'-modified luciferins presented an early challenge. These positions have been rarely exploited for analogue development, and no prior syntheses were amenable to preparing libraries or large quantities of these probes. Rapid, high-yielding syntheses were essential, as large quantities of luciferins are required for light emission assays. Fortunately, the core benzothiazole unit (**1a-c**) of the desired analogues could be accessed from a common route (Figure 3.1D) and in multigram quantities.<sup>34,38</sup> 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 **1a** was problematic, though, due to formation of a hydrated hemiacetal.<sup>48</sup> 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.<sup>49</sup> Toward this end, benzothiazole **1a** was modified with a series of tertiary benzyl amines via in situ iminium formation and coupling (Figure 3.1D). 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 (**3a,b**). A different synthetic approach was necessary, though, as the 4' position cannot be selectively targeted with electrophiles (Figure 3.1D).

### 2.2.2 Analyzing bioluminescent light emission with modified luciferins

With the modified luciferins in hand, we first evaluated their optical properties with Fluc. All analogues were competent light emitters and could be processed by the enzyme (Figure 2.2A). However, the emission intensities were much weaker than those observed with D-luciferin, the native substrate. Interestingly, the largest analogue (**2d**) was not the weakest emitter, suggesting that steric modification alone does not dictate luciferin utilization. Sim-



**Figure 2.2:** Measuring luciferin light emission. (A) Bioluminescence from luciferin analogues (100  $\mu$ M) incubated with 1  $\mu$ g 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 3 or more experiments. (B) Chemiluminescence with luciferin analogues. Emission intensities are plotted as counts per molar luciferin on a log scale. Error bars represent the standard error of the mean for 3 or more experiments.

ilar trends in light emission were observed across a range of physiological pH values (Scheme S3). Consistent with the observed light outputs, the measured kinetic constants for all analogues showed reduced performance relative to D-luciferin. For example, the measured  $K_M$  values were 100-fold larger than the native substrate, with the largest analogues (**2c** and **2d**) exhibiting the lowest relative binding affinities.

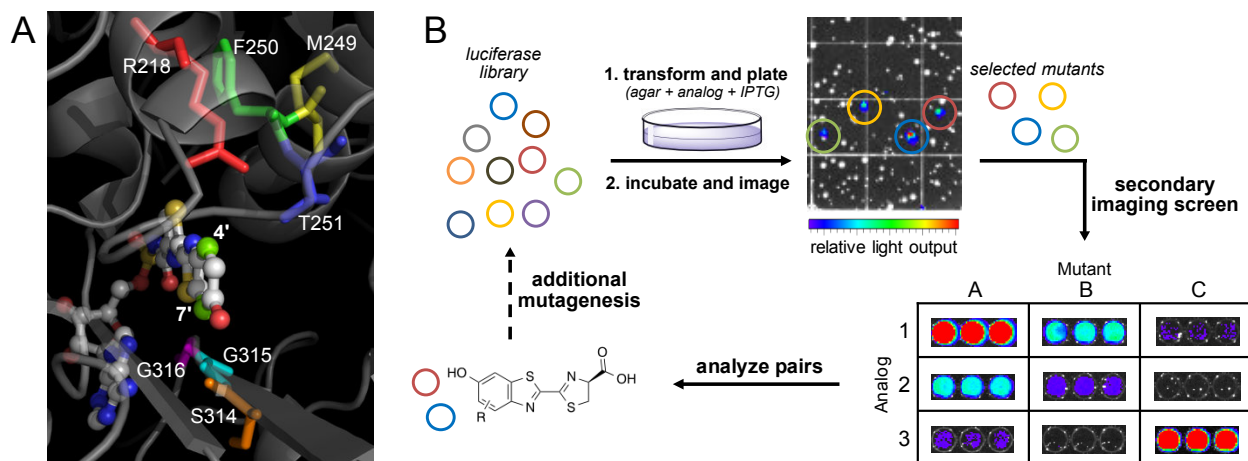
### 2.2.3 Measuring the light-emitting potential of luciferin analogues

We attributed the weak bioluminescence of the analogues 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 analogue must be able to reach an electronic excited state ( $S_1$ ) and relax back to the ground state with concomitant photon release.<sup>50,51</sup> If an analogue 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 analogues were intrinsically capable of light emission, we utilized a previously described chemiluminescence assay.<sup>52</sup> This process mimics the enzymatic reaction itself via formation of an activated ester intermediate, followed by proton abstraction and subsequent reaction with molecular oxygen.<sup>40,53,54</sup> When analogues **2a-d** and **3a,b** were subjected to the assay, robust light emission was observed (Figure 2.2B). In fact, photon outputs for some of the weakest bioluminescent emitters (including **2c** and **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.

## 2.2.4 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,<sup>31,55</sup> 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;<sup>23</sup> such critical residues are often revealed only by random mutagenesis approaches.<sup>56,57</sup> Screening libraries of completely random mutants was impractical in our case, though, owing to the large library sizes needed to achieve adequate enzyme coverage.<sup>57</sup> Screening in bulk is also difficult as bioluminescent light emission is too weak to detect on conventional cell 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, semirational libraries where the mutations were confined to regions known to modulate substrate binding.<sup>58</sup> 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 (Figure 2.3A). These selections were partially based on phylogenetic data gathered from across the insect luciferase family,<sup>13,59</sup> along with previous biochemical assays: Arg218 is known to interact with D-luciferin and influence the local structure of the binding pocket;<sup>29</sup> F250 lies in close proximity to the benzothiazole ring of D-luciferin; T251 has been shown to potentiate substrate binding;<sup>30</sup> residues 314-316 line a critical edge near the luciferin phenolate and C7' position. 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.<sup>9,30,60</sup>



**Figure 2.3:** 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 analogues.

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, nonconserved 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 transformants.<sup>61</sup> The final libraries ranged from 19 to 4800 members in size and were constructed using synthetic gene assembly<sup>62</sup> in combination with circular polymerase extension cloning (CPEC).<sup>63</sup>

The libraries were screened for orthogonal substrate usage using a two-tiered approach. Library DNA was first introduced into bacteria, and the transformants were arrayed across agar plates containing embedded luciferins (Figure 2.3B). Light-emitting colonies were easily identified, and in some cases, the light emission values were on par with native Fluc and D-luciferin. A handful of the corresponding mutants were sequenced. Some mutations were observed for multiple analogues, suggesting that they might be selective for bulky luciferins. Other mutations were unique to each compound, which is notable, given the subtle structural

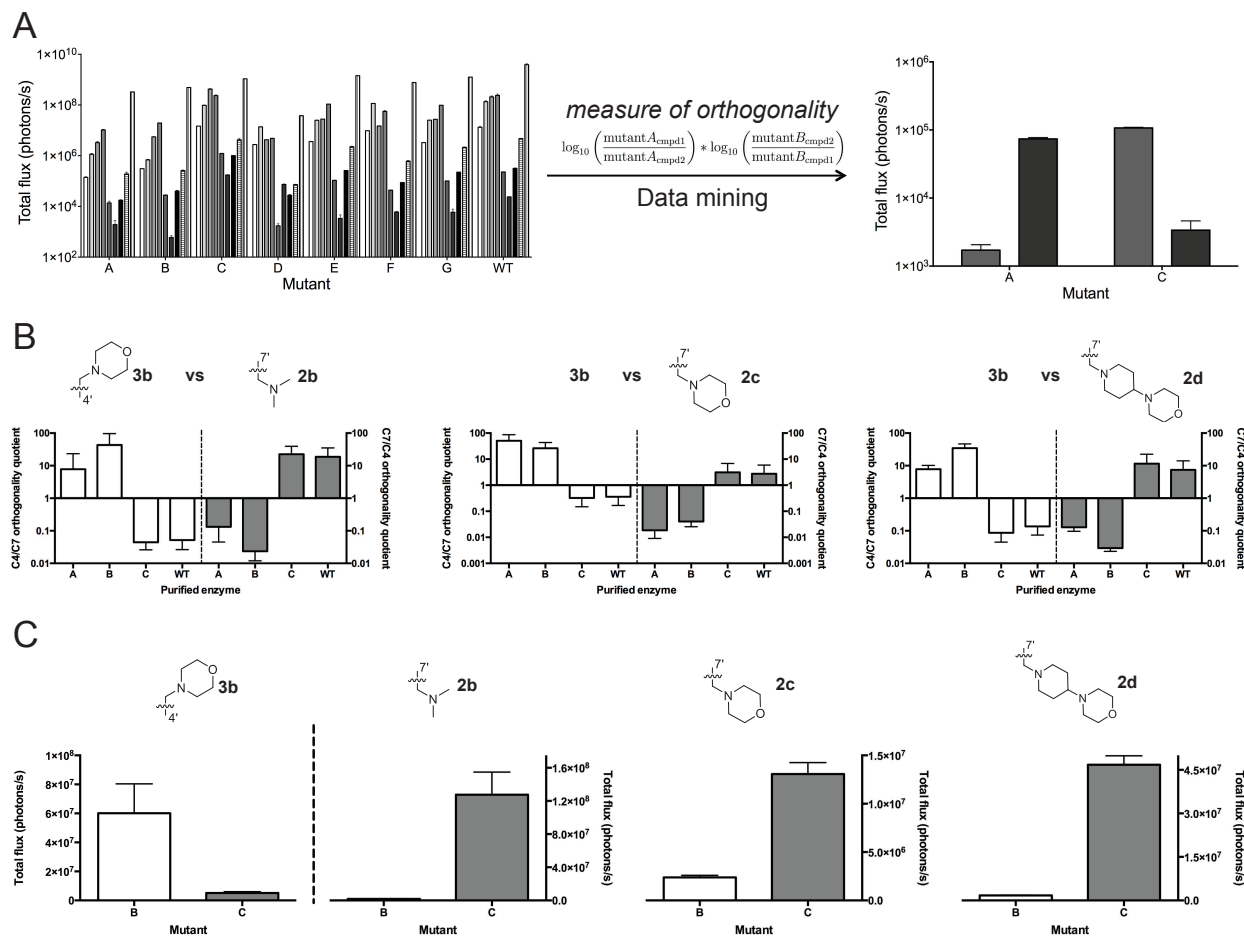


differences between some of the analogues. The number of colonies screened was about 3x the calculated diversity for each library.

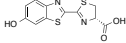
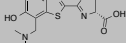
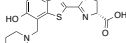
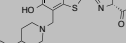
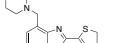
While initial screens revealed functional mutants (and quickly culled nonfunctional enzymes), they did not report on selective substrate usage (i.e., orthogonality). The on-plate screens also did not control overall expression levels and differences in compound transport. To address these parameters, 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 analogues. 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 analogues in secondary screens.

To mine the entire collection of imaging data for substrateselective pairs, we first developed a measure of orthogonality (equation shown in Figure 2.4A). Favorable values are obtained when two mutants (e.g., A and B) react robustly with unique substrates (e.g., compounds **1** and **2**, respectively) in a mutually exclusive manner. Thus, the more selective a pair of enzymes for their cognate substrates, the larger the orthogonality rating. Since the number of potential pairings exceeded 3000 in our data set, we wrote a computer script to rapidly examine all pairs in an unbiased fashion. The program iterated through each possible combination, calculating the corresponding orthogonality rating. The script ultimately returned a list of pairs ranked by their potential for orthogonality (and thus utility for multicomponent imaging).

The top pairs identified by the script exhibited selectivity for analogues **3b** (mutants A and B) and **2b-d** (mutant C). The magnitude of each mutants preference—defined as the orthogonality quotient—was analyzed. As shown in Figure 2.4B and C, mutants A and B exhibited nearly a 100-fold preference for **3b** over other analogues, while mutant C strongly



**Figure 2.4:** Analyzing orthogonal enzyme-substrate pairs. (A) Representative emission of luciferase mutants screened against a panel of luciferin analogues. These data were analyzed with a computer algorithm to determine lead mutants with the strongest orthogonality. (B) Purified mutants exhibit orthogonality. Enzyme (1  $\mu\text{g}$ ) was incubated with 100  $\mu\text{M}$  of luciferin analogues, 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.

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

<sup>a</sup> Analog values normalized to their corresponding emission with Fluc. Errors represent standard error of the mean for  $n = 3$  measurements. <sup>b</sup> Kinetic constants are apparent values, determined via measurements of initial rates of light emission over a range of 2  $\mu$ M to 10 mM. Errors represent standard error of the mean for  $n \geq 3$  measurements.  $k_{cat}$  values are relative to their corresponding value with Fluc. Errors represent standard error of the mean for  $n \geq 3$  measurements. <sup>c</sup>  $\lambda_{max}$  value could not be determined due to low level of light emission.

**Table 2.1:** Biochemical analyses of orthogonal enzyme-substrate pairs

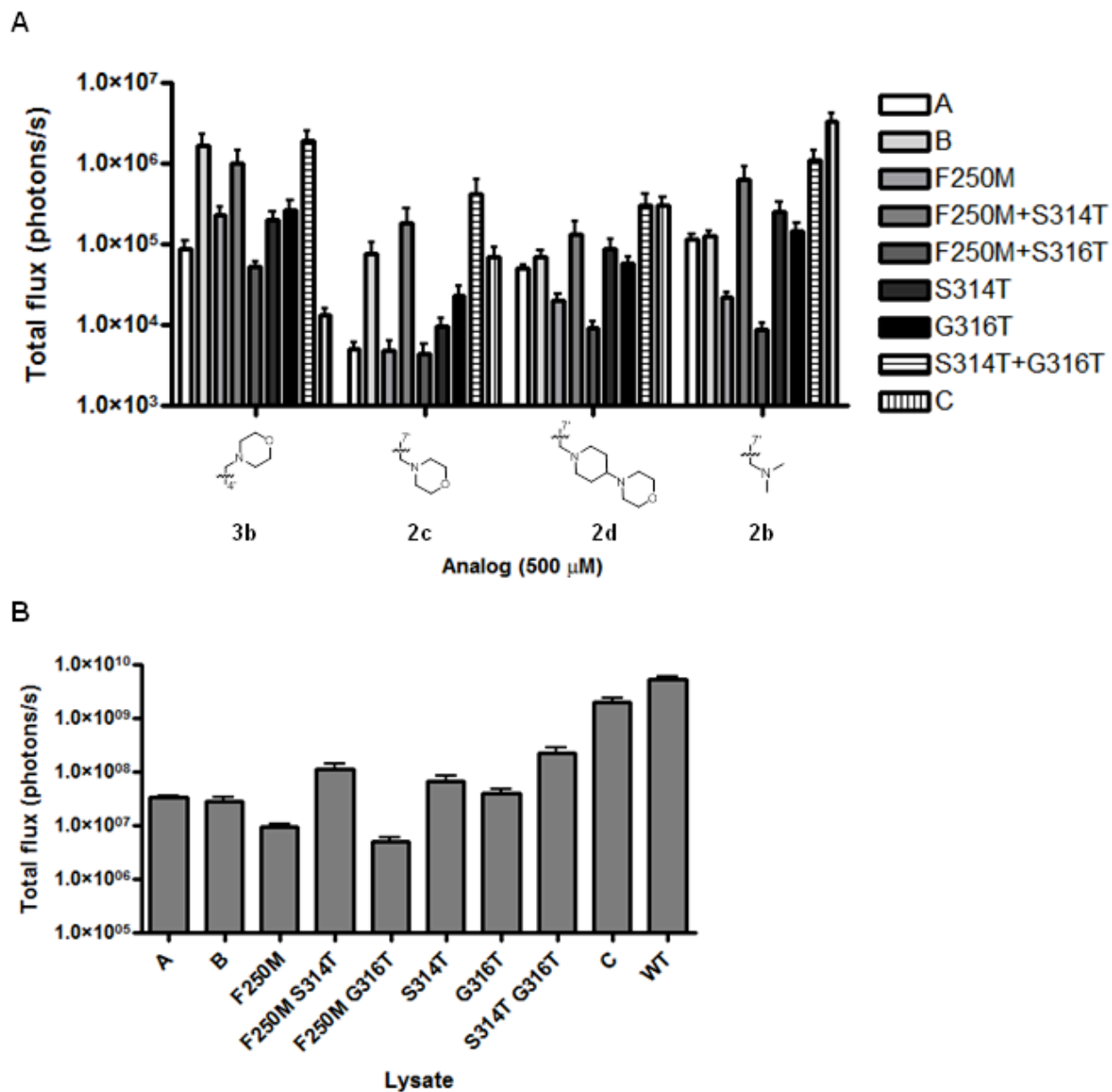
avored C7'-modified analogues. Biochemical analyses further indicated that the brightest mutant enzymes were those capable of most efficient substrate turnover (Table 2.1).

## 2.2.5 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 residues 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.<sup>29</sup> It has been hypothesized that the smaller Ala group allows more water molecules to access the active site, potentially quenching light emission.<sup>29</sup> The bulky morpholino substituent of **3b** could fill this active site void to retain photon production.

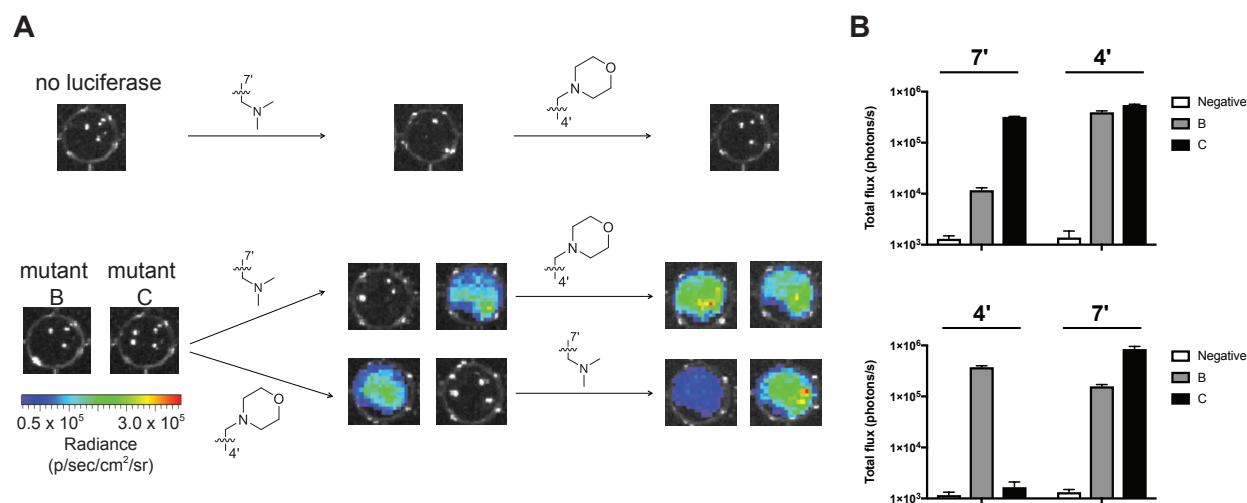
The third mutant (mutant C) was more selective for the C7'-modified luciferins compared to the C4'-modified compound. Mutant C harbored a single mutation, R218K. R218K may slightly enlarge the active site of the luciferase. This mutation has also been shown to boost activity with bulkier cyclic aminoluciferin analogues.<sup>22</sup> The improved selectivity with **2b-d** could be the result of active site positioning. The C7' substituents could potentially place the luminophore in a more advantageous spot for light emission.

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, G316T, or native Fluc residues) were then allowed. Imaging analyses of these combinatorial mutants indicated that R218A and F250M were critical for luciferin discrimination (Figure 2.5). Both mutations should result in a larger active site, but why they preferentially accommodate **3b** over other analogues 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 for subsequent oxidation. Indeed, when **3b** was incubated with R218A/F250M, light emission was maintained (as compared to Fluc, Figure 2.5). When D-luciferin and the C7'-modified analogues were incubated with this same mutant, though, light emission was drastically reduced. Interestingly, the R218A/S314T mutant exhibited an opposite trend in analogue selectivity: **2b** was preferred to **3b** (Figure 2.5). Collectively, these results suggest that mutant luciferases can be tuned to respond to unique substrates. It 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.



## 2.2.6 Cellular imaging with orthogonal pairs

As a step toward multicomponent imaging applications, we evaluated the orthogonal enzymes and probes in cultured cell models. Mammalian cell lines (HEK293 and DB7) were engineered to express orthogonal mutants A-C. Cells were incubated with analogues **2b-d** and **3b**, and photon outputs were measured. As shown in Figure 2.6, 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



**Figure 2.6:** Imaging cells with orthogonal luciferase-luciferin pairs. (A) Mutant luciferase-expressing DB7 cells were plated ( $1.5 \times 10^5$  cells/well) in 96-well black plates and sequentially incubated with C4' and C7' sterically modified luciferins ( $750 \mu\text{M}$ ). Representative bioluminescence images are shown. (B) Quantification of the images from (A) after initial substrate addition. Error bars represent the standard error of the mean for experiments performed in triplicate.

for each orthogonal luciferin: cells with mutant B were only visible upon treatment with analogue **3b**, whereas cells with mutant C were only visible upon treatment with analogues **2b-d**. Importantly, the orthogonal pairs could also distinguish unique cell types in a single imaging session. For example, DB7 cells stably expressing mutant B or C could be readily detected via sequential administration of the requisite substrates (Figure 2.6). These data

suggest that crossreactivity between mutants B and C and their non-orthogonal substrates is minimal.

## 2.3 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 analogues. 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 multicomponent *in vivo* imaging.

Future studies will be aimed at generating additional bioluminescent probes with improved brightness and other optical properties. The enzyme-substrate hits reported here, while 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 luciferins, including analogues with altered chromophores that could provide drastically different colors of light. 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.

## **2.4 Author contributions**

The work described above was a highly collaborative effort primarily between Will Porterfield, Krysten Jones, and me. My primary contributions in this chapter involve data mining of the enzyme-substrate reactivities to extract orthogonal pairs (Figure 2.4). I also developed the chemiluminescence assay to evaluate the inherent light-emitting ability of our luciferin analogs (Figure 2.2). Finally, I conducted the kinetic analysis of the lead probes identified in this study (Table 2.1). Below are the materials and methods concerning these contributions. For all other data/methods see the full version of the published paper.<sup>64</sup>

## **2.5 Materials and methods**

### **2.5.1 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.



## 2.5.2 General chemiluminescence procedure<sup>52</sup>

Phenyl esters of each luciferin analog were prepared following the basic procedure of Kim et al.<sup>65</sup> In brief, the potassium salt of each luciferin (6.0  $\mu\text{mol}$ ) was added to an oven-dried, two-dram vial containing a small stir bar. DMSO- $d_6$  (0.55 mL) containing a mesitylene internal standard (0.275  $\mu\text{L}$ ) was then added, and the luciferin was dissolved with stirring (5 min). Phenylchloroformate (0.76  $\mu\text{L}$ , 6.0  $\mu\text{mol}$ ) was subsequently added, and a brief color change was observed in most cases. The solutions were stirred for an additional 5 min. A portion of each solution (5  $\mu\text{L}$ ) was reserved, and the remainder was added to an NMR tube for analysis. The NMR sample was kept at ambient temperature until luminometer measurements were acquired (see below). At that point, the NMR sample was frozen ( $-73^\circ\text{C}$ ) to preserve the contents of the tube. At a later time, the tube was thawed and a  $^1\text{H}$  NMR spectrum was immediately acquired (2 scans, 20 s relaxation delay). The concentration of the luciferin phenyl ester was determined via comparison to the internal standard. The reserved portion of each luciferin ester solution was diluted to 0.5 mL with anhydrous DMSO, and 50  $\mu\text{L}$  of this solution was added to six wells of a black 96-well flat-bottom plate (Greiner). Chemiluminescence values were acquired on a Tecan Infinite F200 PRO plate-reading luminometer. Data were acquired for 1.5 s prior to injection of potassium phenoxide solution (50  $\mu\text{L}$  of a 0.1 M solution). The phenoxide solution was prepared via dissolution of potassium *tert*-butoxide (112 mg) and phenol (94 mg) in anhydrous DMSO (10 mL) with stirring (30 min). The total volume in each well was 100  $\mu\text{L}$ . After the addition of base, luminescence data were collected for an additional 50 s (100 ms integration times were used). Relative luminescence yields were determined via trapezoidal integration of the data.

### 2.5.3 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 MgSO<sub>4</sub>, 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).

### 2.5.4 Data mining

See section 3.6 in chapter 3.

# Bibliography

- [1] Paley, M. A.; Prescher, J. A. *Med. Chem. Comm.* **2014**, *5*, 255–267.
- [2] Prescher, J. A.; Contag, C. H. *Curr. Opin. Chem. Biol.* **2010**, *14*, 80–89.
- [3] Badr, C. E.; Tannous, B. A. *Trends Biotechnol.* **2011**, *29*, 624–633.
- [4] Adams, S. T.; Miller, S. C. *Curr. Opin. Chem. Biol.* **2014**,
- [5] Branchini, B. R.; Hayward, M. M.; Bamford, S.; Brennan, P. M.; Lajiness, E. J. *Photochem. Photobiol.* **1989**, *49*, 689–695.
- [6] Jathoul, A. P.; Grounds, H.; Anderson, J. C.; Pule, M. A. *Angew. Chem., Int. Ed.* **2014**, *53*, 13059–13063.
- [7] Mofford, D. M.; Reddy, G. R.; Miller, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 4443–4448.
- [8] Branchini, B. R.; Ablamsky, D. M.; Murtiashaw, M. H.; Uzasci, L.; Fraga, H.; Southworth, T. L. *Anal. Biochem.* **2007**, *361*, 253–262.
- [9] Branchini, B. R.; Ablamsky, D. M.; Rosenman, J. M.; Uzasci, L.; Southworth, T. L.; Zimmer, M. *Biochemistry* **2007**, *46*, 13847–13855.
- [10] Mezzanotte, L.; Fazzina, R.; Michelini, E.; Tonelli, R.; Pession, A.; Branchini, B.; Roda, A. *Mol. Imaging Biol.* **2010**, *12*, 406–414.

- [11] Haddock, S. H. D.; Moline, M. A.; Case, J. F. *Ann Rev Mar Sci* **2010**, *2*, 443–493.
- [12] Oba, Y.; Schultz, D. T. *Adv. Biochem. Eng. Biotechnol.* **2014**, *144*, 3–36.
- [13] Viviani, V. R. *Cellular and Molecular Life Sciences* **2002**, *59*, 1833–1850.
- [14] Massoud, T. F.; Paulmurugan, R.; De, A.; Ray, P.; Gambhir, S. S. *Curr. Opin. Biotechnol.* **2007**, *18*, 31–37.
- [15] Porterfield, W. B.; Prescher, J. A. *Curr. Opin. Chem. Biol.* **2015**, *24*, 121–130.
- [16] Bhaumik, S.; Gambhir, S. S. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 377–382.
- [17] 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. *Mol. Ther.-Nucleic Acids* **2013**, *2*, e99.
- [18] Pichler, A.; Prior, J. L.; Piwnica-Worms, D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1702–1707.
- [19] Petushkov, V. N.; Dubinnyi, M. A.; Tsarkova, A. S.; Rodionova, N. S.; Baranov, M. S.; Kublitski, V. S.; Shimomura, O.; Yampolsky, I. V. *Angew. Chem. Int. Ed.* **2014**, *53*, 5566–5568.
- [20] Evans, M. S.; Chaurette, J. P.; Adams, S. T.; Reddy, G. R.; Paley, M. A.; Aronin, N.; Prescher, J. A.; Miller, S. C. *Nat. Methods* **2014**, *11*, 393–395.
- [21] Reddy, G. R.; Thompson, W. C.; Miller, S. C. *J. Am. Chem. Soc.* **2010**, *132*, 13586–13587.
- [22] Harwood, K. R.; Mofford, D. M.; Reddy, G. R.; Miller, S. C. *Chemistry & Biology* **2011**, *18*, 1649–1657.
- [23] Adams Jr., S. T.; Mofford, D. M.; Reddy, G. S. K. K.; Miller, S. C. *Angew. Chem. Int. Ed.* **2016**, *55*, 4943–4946.

- [24] Rowe, L.; Dikici, E.; Daunert, S. *Anal. Chem.* **2009**, *81*, 8662–8668.
- [25] Hall, M. P. et al. *ACS Chem. Biol.* **2012**, *7*, 1848–1857.
- [26] Berger, F.; Paulmurugan, R.; Bhaumik, S.; Gambhir, S. S. *Eur. J. Nucl. Med. Mol. Imaging* **2008**, *35*, 2275–2285.
- [27] Contag, C. H.; Spilman, S. D.; Contag, P. R.; Oshiro, M.; Eames, B.; Dennery, P.; Stevenson, D. K.; Benaron, D. A. *Photochem. Photobiol.* **1997**, *66*, 523–531.
- [28] Zhao, H.; Doyle, T. C.; Coquoz, O.; Kalish, F.; Rice, B. W.; Contag, C. H. *J. Biomed. Opt.* **2005**, *10*, 041210.
- [29] Branchini, B. R.; Magyar, R. A.; Murtiashaw, M. H.; Portier, N. C. *Biochemistry* **2001**, *40*, 2410–2418.
- [30] Branchini, B. R.; Southworth, T. L.; Murtiashaw, M. H.; Boije, H.; Fleet, S. E. *Biochemistry* **2003**, *42*, 10429–10436.
- [31] Sundlov, J. A.; Fontaine, D. M.; Southworth, T. L.; Branchini, B. R.; Gulick, A. M. *Biochemistry* **2012**, *51*, 6493–6495.
- [32] Nakatsu, T.; Ichiyama, S.; Hiratake, J.; Saldanha, A.; Kobashi, N.; Sakata, K.; Kato, H. *Nature* **2006**, *440*, 372–376.
- [33] Meroni, G.; Rajabi, M.; Santaniello, E. *Arkivoc* **2009**, 265–288.
- [34] McCutcheon, D. C.; Porterfield, W. B.; Prescher, J. A. *Org. Biomol. Chem.* **2015**, *13*, 2117–2121.
- [35] White, E. H.; Wörther, H.; Field, G. F.; McElroy, W. D. *J. Org. Chem.* **1965**, *30*, 2344–2348.
- [36] White, E. H.; Wörther, H. **1966**, *31*, 1484–1488.

- [37] Woodrooffe, C. C.; Meisenheimer, P. L.; Klaubert, D. H.; Kovic, Y.; Rosenberg, J. C.; Behney, C. E.; Southworth, T. L.; Branchini, B. R. *Biochemistry* **2012**, *51*, 9807–9813.
- [38] McCutcheon, D. C.; Paley, M. A.; Steinhardt, R. C.; Prescher, J. A. *J. Am. Chem. Soc.* **2012**, *134*, 7604–7607.
- [39] Conley, N. R.; Dragulescu Andrasi, A.; Rao, J.; Moerner, W. E. *Angew. Chem. Int. Ed.* **2012**, *51*, 3350–3353.
- [40] Branchini, B. R.; Behney, C. E.; Southworth, T. L.; Fontaine, D. M.; Gulick, A. M.; Vinyard, D. J.; Brudvig, G. W. *J. Am. Chem. Soc.* **2015**, *137*, 7592–7595.
- [41] Meroni, G.; Ciana, P.; Maggi, A.; Santaniello, E. *Synlett* **2009**, *2009*, 2682–2684.
- [42] White, E. H.; Worther, H.; Seliger, H. H.; McElroy, W. D. *J. Am. Chem. Soc.* **1966**, *88*, 2015–&.
- [43] Seliger, H. H.; McElroy, W. D.; Field, G. F.; White, E. H. *Proc. Natl. Acad. Sci. U.S.A.* **1961**, *47*, 1129–&.
- [44] Kojima, R.; Takakura, H.; Ozawa, T.; Tada, Y.; Nagano, T.; Urano, Y. *Angew. Chem. Int. Ed.* **2012**, *52*, 1175–1179.
- [45] Wang, J.; Lee, T. S.; Zhang, Z.; Tung, C.-H. *Chem Asian J* **2017**, *12*, 397–400.
- [46] Woodrooffe, C. C.; Shultz, J. W.; Wood, M. G.; Osterman, J.; Cali, J. J.; Daily, W. J.; Meisenheimer, P. L.; Klaubert, D. H. *Biochemistry* **2008**, *47*, 10383–10393.
- [47] Kuchimaru, T.; Iwano, S.; Kiyama, M.; Mitsumata, S.; Kadonosono, T.; Niwa, H.; Maki, S.; Kizaka-Kondoh, S. *Nat. Commun.* **2016**, *7*, 11856.
- [48] Jones, P. R.; Gelinas, R. M. *J. Org. Chem.* **1981**, *46*, 194–196.
- [49] Phillips, J. P.; Barrall, E. M. *J. Org. Chem.* **1956**, *21*, 692–694.

- [50] da Silva, L. P.; Esteves da Silva, J. C. G. *ChemPhysChem* **2012**, *13*, 2257–2262.
- [51] Hopkins, T. A.; Seliger, H. H.; White, E. H.; Cass, M. W. *J. Am. Chem. Soc.* **1967**, *89*, 7148–7150.
- [52] Steinhardt, R. C.; Rathbun, C. M.; Krull, B. T.; Yu, J. M.; Yang, Y.; Nguyen, B. D.; Kwon, J.; McCutcheon, D. C.; Jones, K. A.; Furche, F.; Prescher, J. A. *ChemBioChem* **2016**,
- [53] Hopkins, T. A.; Seliger, H. H.; White, E. H.; Cass, M. W. *J. Am. Chem. Soc.* **1967**, *89*, 7148–7150.
- [54] Kato, D.-i.; Shirakawa, D.; Polz, R.; Maenaka, M.; Takeo, M.; Negoro, S.; Niwa, K. *Photochem. Photobiol. Sci.* **2014**, *13*, 1640–1645.
- [55] Mao, Y. *Protein Eng. Des. Sel.* **2011**, *24*, 341–349.
- [56] Chen, M. M. Y.; Snow, C. D.; Vizcarra, C. L.; Mayo, S. L.; Arnold, F. H. *Protein Eng. Des. Sel.* **2012**, *25*, 171–178.
- [57] Reetz, M. T.; Prasad, S.; Carballeira, J. D.; Gumulya, Y.; Bocola, M. *J. Am. Chem. Soc.* **2010**, *132*, 9144–9152.
- [58] Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z.-G.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P. *ACS Synth. Biol.* **2012**, *2*, 83–92.
- [59] Amaral, D. T.; Arnoldi, F.; Rosa, S. P.; Viviani, V. R. *Luminescence* **2014**, *29*, 412–422.
- [60] Viviani, V. R.; Amaral, D. T.; Neves, D. R.; Simoes, A.; Arnoldi, F. G. C. *Biochemistry* **2013**, *52*, 19–27.
- [61] Pines, G.; Pines, A.; Garst, A. D.; Zeitoun, R. I.; Lynch, S. A.; Gill, R. T. *ACS Synth. Biol.* **2014**, 141030112600008.

- [62] Ness, J. E.; Kim, S.; Gottman, A.; Pak, R.; Krebber, A.; Borchert, T. V.; Govindarajan, S.; Mundorff, E. C.; Minshull, J. *Nat. Biotechnol.* **2002**, *20*, 1251–1255.
- [63] Quan, J.; Tian, J. *Nature Protocols* **2011**, *6*, 242–251.
- [64] Jones, K. A.; Porterfield, W. B.; Rathbun, C. M.; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. *J. Am. Chem. Soc.* **2017**, *139*, 2351–2358.
- [65] Kim, S.; Lee, J. I.; Kim, Y. C. *J. Org. Chem.* **1985**, *50*, 560–565.



# Chapter 3

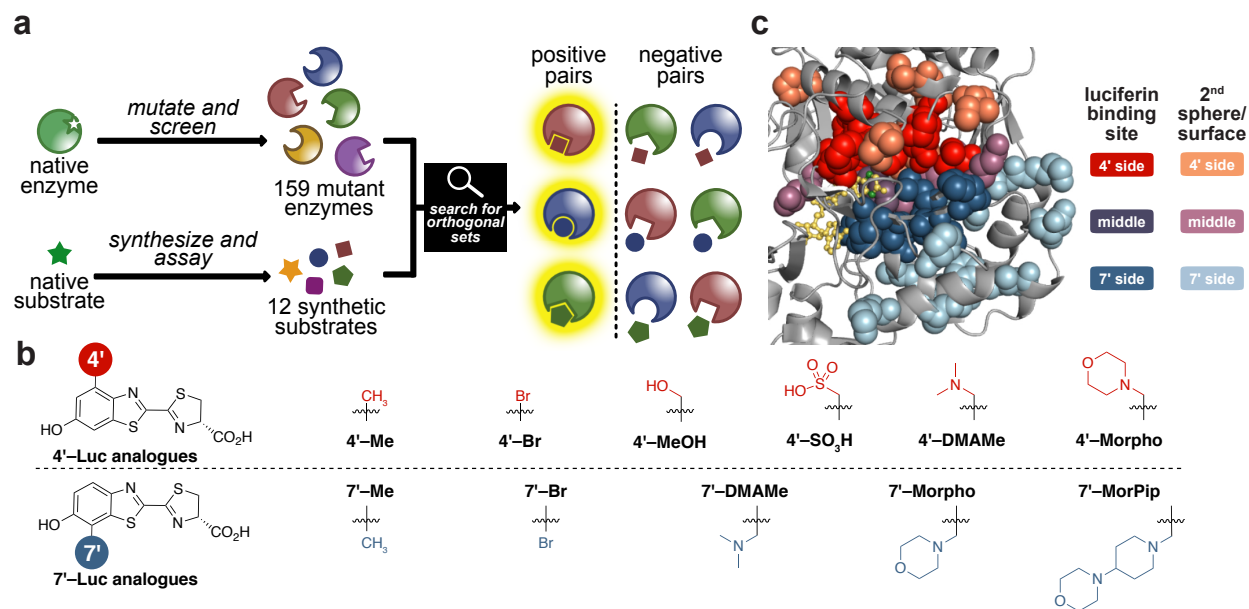
## Parallel screening for rapid identification of orthogonal bioluminescent tools

### 3.1 Introduction

Our understanding of living systems is profoundly shaped by our ability to see biology in action. Central to these efforts are robust and translatable imaging tools.<sup>1,2</sup> Decades of work to engineer and optimize fluorescent proteins have provided a palette of designer probes for cellular studies. Using combinations of these tools, it is now possible to trace the orchestrated behaviors of immune cells,<sup>3</sup> nerve cell connections,<sup>4</sup> and other interactions.<sup>5</sup> Widespread application of fluorescent probes will continue to reveal unanticipated facets of biology. Concurrently, gaps in our knowledge will spur the development of innovative tools. Fluorescent probes endowed with novel functions (altered colors, photoswitches, etc.) are already enabling new pursuits.<sup>6,7</sup>

Biological discoveries will be further bolstered by advances in bioluminescent probe development. Bioluminescence relies on light generation via luciferase enzymes and luciferin small molecules.<sup>8,9</sup> Since no excitation light is required, this modality is attractive for studies involving large length or time scales.<sup>10–12</sup> Indeed, firefly luciferase (Fluc) and its cognate substrate (D-luciferin) are ubiquitously used in rodent models to interrogate molecular and cellular events.<sup>13,14</sup> Fluc and its homologues have been further engineered to provide different colors of light.<sup>15,16</sup> Such tools have been applied for multicomponent imaging *in vitro*. The wavelengths achieved, though, are unsatisfactory for routine use *in vivo*. Multicomponent imaging with even the most spectrally resolved probes remains challenging due to interference from surrounding tissue.<sup>11</sup> As a consequence, bioluminescence has lagged behind fluorescence for multicellular studies in tissues and whole organisms.

To realize multicomponent imaging *in vivo*, we turned to a more tractable parameter: substrate selectivity. Enzymes exhibiting mutually exclusive (i.e., orthogonal) substrate preferences should be readily distinguished in a variety of biological models. Fluc is remarkably tolerant of a variety of luciferin modifications, including both electronic<sup>17–21</sup> and steric<sup>22–26</sup> derivatives. While dozens of luciferin analogues have been crafted, most result in reduced photon outputs relative to D-luciferin, the native substrate, at saturating doses.<sup>27</sup> In some cases, boosts in light emission have been achieved using modified versions of the enzyme.<sup>23,25</sup> These results set the stage for developing designer luciferase-luciferin pairs, but few methods to systematically generate orthogonal sets have been pursued.<sup>25,28</sup> Substrate-selective luciferases are found in nature, and a handful have been coopted for dual imaging (e.g., combinations of D-luciferin- and coelenterazine-utilizing enzymes).<sup>29–32</sup> However, most remain suboptimal for use *in vivo*. The characterization of other naturally occurring luciferases and luciferins has also not kept pace with the demand for new pairs. Consequently, bioluminescence imaging has been limited by a lack of mutually orthogonal enzymes and substrates.



**Figure 3.1:** Parallel screening of luciferase mutants and luciferin analogues to identify orthogonal pairs. (a) General strategy for identifying substrate-selective, mutually orthogonal enzymes. For bioluminescent probes, positive (matched) pairs are enzymesubstrate combinations that provide robust light emission. Negative (mismatched) pairs are combinations that exhibit reduced photon outputs. (b) Collections of 4'- and 7'-modified luciferins used for parallel screening. The molecules were synthesized from a common intermediate. (c) Fluc residues targeted for mutagenesis to accommodate 4'- and 7'-modified luciferins (shaded in red and blue hues, respectively; PDB structure: 4G36). A bound D-luciferin-AMP analogue (yellow) is shown for reference.

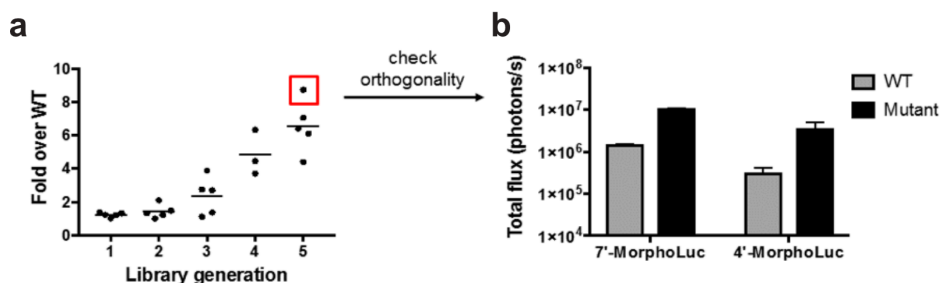
We aimed to expedite the search for luciferases that exhibit unique preferences for distinct luciferin analogues. Accessing enzymes with alternative substrate use is well preceded in directed evolution.<sup>33-35</sup> However, traditional applications of this technique have focused on optimizing one enzyme at a time. Selectivity for one molecule over another is often realized as a consequence, but is not typically the parameter being screened.<sup>36-41</sup> Here, we present a general and rapid approach to achieve substrate selectivity and engineer orthogonal luciferase-luciferin pairs. This strategy relies on parallel screening of functional luciferases with collections of chemically diverse luciferins (Figure 3.1a). The large data sets are then mined for orthogonal combinations using a custom computer script. Enzyme-substrate pairs are deemed orthogonal if robust reactivity is observed when complementary partners interact, but minimal to no reactivity is observed in all other cases (Figure 3.1a). Collectively, we screened 159 mutants and 12 analogues, generating a candidate list of greater than 800,000 possible pairs. We evaluated the orthogonality of 175 pairs in vitro. A subset was successfully applied in cultured cell and animal models, highlighting the feasibility and translatability of the approach. We also analyzed principles governing selective substrate use and identified methods to search for expanded collections of orthogonal imaging agents. Overall, this work greatly expands the number of viable bioluminescence probes for multicomponent imaging and presents a strategy to accelerate the identification of new ones. The parallel screening method is also applicable to other areas where selective substrate use is required.

## **3.2 Results and discussion**

### **3.2.1 Expanding the pool of candidate luciferins and luciferases**

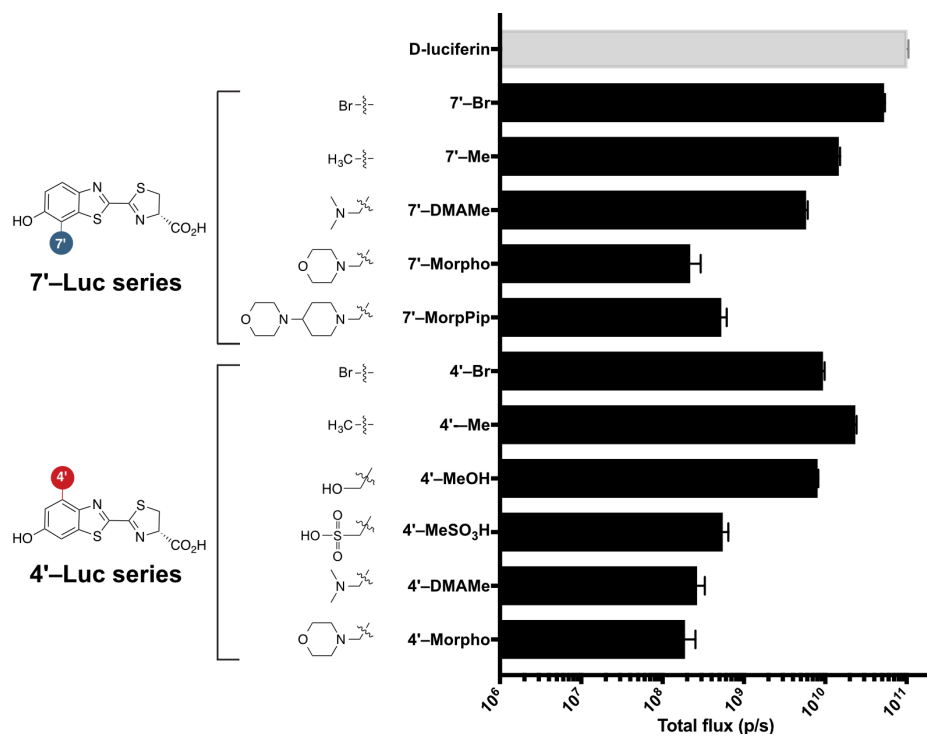
As a starting point for substrate modification, we focused on D-luciferin derivatives with steric appendages at C4' and C7'. These positions lie in close proximity to the Fluc back-

bone,<sup>42</sup> and preliminary work revealed that modifications here do not quench or otherwise impede photon emission.<sup>25</sup> We also previously identified a pair of luciferases that could discriminate between luciferins with modifications at these positions,<sup>25</sup> suggesting that they were good starting points for new probe development. However, attempts to optimize this pair via traditional directed evolution (focusing on one enzyme at a time) did not result in improved substrate selectivity (Figure 3.2).



**Figure 3.2:** Brighter mutants can be evolved but are not orthogonal. (a) Fold increase in light emission tracked over five generations. Mutants were screened for total photon output with 7'MorphoLuc. (b) The brightest mutant (red box in (a)) with 7'MorphoLuc also exhibited increased photon output with 4'MorphoLuc (black bars). Photon flux values for native Fluc (gray bars) are also shown. Imaging was performed in bacterial lysate (250  $\mu$ M luciferin analogue).

We reasoned that screening for selectivity at the outset would provide a more rapid route to new bioluminescent pairs. Engineering luciferases to discriminate among structurally similar compounds can be difficult.<sup>23,43</sup> Thus, we initially focused on diversifying the enzyme and substrate inputs. Collections of both new and known<sup>24,25</sup> luciferins were assembled (Figure 3.1b). These molecules covered a broad range of chemical space and comprised both hydrophilic and hydrophobic functional groups. The luciferins were benchmarked for light emission with Fluc (Figure 3.3). All compounds were functional light emitters, though they varied in terms of photon output. Some level of enzyme activity is necessary for successful evolution, but weak performers can be advantageous starting points for evolving new functions.<sup>33</sup>

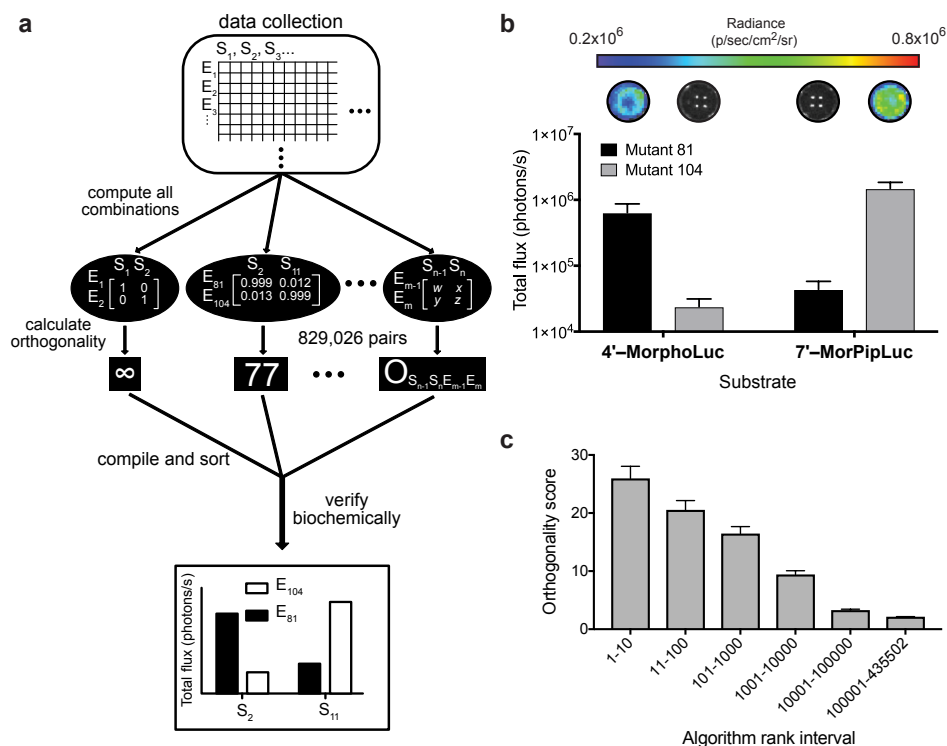


**Figure 3.3:** Sterically modified luciferins emit varying levels of photons with Fluc. Fluc (1  $\mu\text{g}$ ) was incubated with 100  $\mu\text{M}$  luciferin analogue and photon outputs were measured. Error bars represent the standard error of the mean for  $n > 2$  independent experiments.

In parallel with luciferin diversification, we targeted broad sectors of Fluc sequence space for mutagenesis. Twenty-three residues near the active site were selected, and the mutations were covered in 8 libraries (labeled in Figure 3.1c). The majority of the mutants would likely be nonfunctional, and thus not ideal starting points for probe development. We aimed to eliminate these luciferases early on and perform parallel screens with an enriched pool of viable mutants. Such an approach would save time and reagents as luciferases are not amenable to high-volume separations (e.g., FACS) or selections; rather, each mutant must be physically interrogated with a given substrate. We adapted a high-throughput method to traverse the luciferase libraries and cull nonfunctional members.<sup>25</sup> The libraries were transformed into bacteria, and the transformants were grown on agar containing one of four minimally perturbed luciferins: **4'/7'-BrLuc** or **4'/7'-MeLuc** (Figure 3.1b). These analogues were selected for on-plate screens since they are among the brightest emitters and easy to access in bulk. Light-emitting colonies were picked and further assayed in lysate and by sequencing. A variety of mutants were identified, including enzymes that were unique to each luciferin. Some hits were further diversified (1-3 generations) via random mutagenesis to enlarge the pool of luciferase mutants.

### **3.2.2 Screening for Orthogonal Luciferase-Luciferin Pairs in Silico**

With enriched sets of functional luciferases, we aimed to screen the collection for orthogonal pairs. Testing each combination of two mutants and two substrates would have required 829,026 separate experiments (Figure 3.4a), an impractical number. Instead, we screened each analogue across the same panel of 159 luciferases, generating 1908 (12 substrates x 159 enzymes) individual data points. An ideal orthogonal enzyme would be positively matched with a single substrate and negatively matched with all other luciferins. To identify such enzymes, we established a metric to quantify orthogonality and mine the data. We reasoned that perfect selectivity could be represented by an identity matrix (section 3.6). Orthogo-

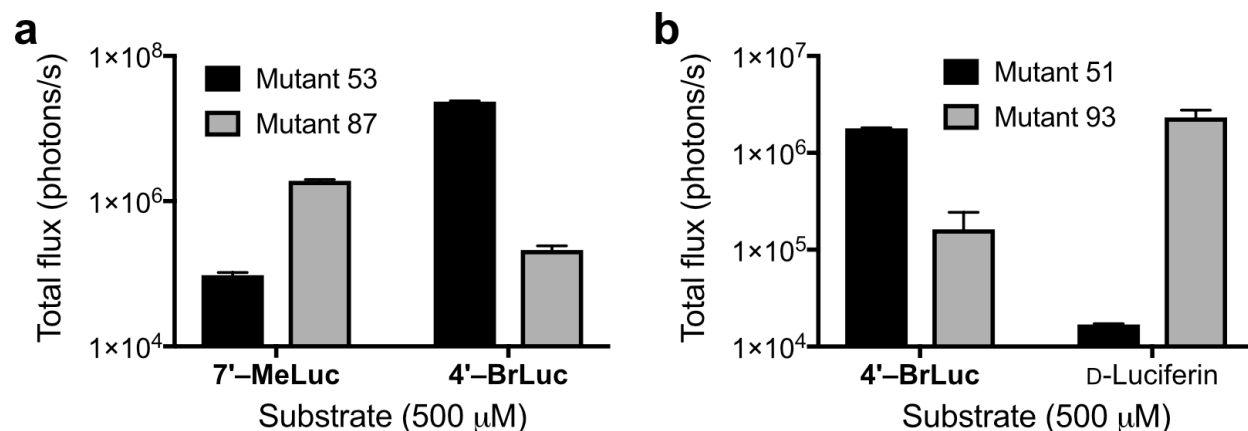


**Figure 3.4:** Uncovering orthogonal pairs in silico. (a) Computational approach to identifying orthogonal sets. Parallel screens of mutant enzymes ( $E_m$ , where  $n$  and  $m$  are integers) and substrate analogues ( $S_n$ , where  $n$  and  $m$  are integers) were performed and light emission values entered into a database. Data were analyzed with a custom computer script to identify orthogonal sets. (b) Sample orthogonal bioluminescent probes. Bacteria expressing mutant enzymes were expanded, lysed, and distributed evenly among replicate wells. Lysates were treated with luciferin analogues and imaged. Representative images are shown, along with quantified photon outputs. (c) Orthogonality scores correlated with computer script rank. Orthogonal sets predicted in silico were verified biochemically as in panel b. Each bar (beyond rank 11) represents >40 unique sets that were evaluated in head-to-head comparisons in vitro. (For interval 1-10, all ten orthogonal sets were examined). For panels b and c, error bars represent the standard error of the mean for 3 or more experiments.



nality would be maximal if each enzyme was completely selective for its cognate substrate (represented by a 1 in the identity matrix) and nonfunctional with other luciferins (0 in the identity matrix). An orthogonality score was determined by representing each set of two luciferases and two luciferins as a square matrix, with enzymes in rows and substrates in columns. These data were compared to the ideal case (identity matrix) via root-mean-square distance (RMSD). The RMSD values were then converted to numeric values (i.e., orthogonality scores) representing the fold resolution between the positive and negative pairings (see Supplementary Note for more details). We wrote a computer script to assemble each possible matrix from the screening data and calculate the orthogonality of each pairing. The pairs were sorted by increasing RMSD, with the smallest value (highest orthogonality score) representing the most orthogonal pair.

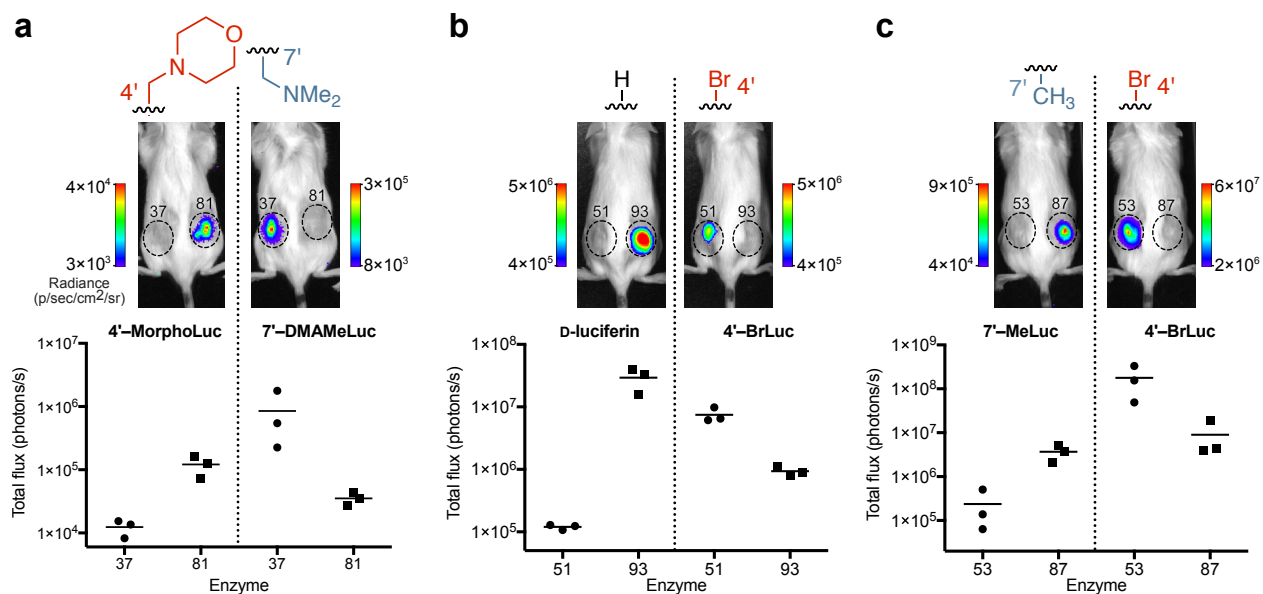
The algorithm provided a ranked list of the 829,026 possible orthogonal sets (Figure 3.4a). The top pair comprised analogues **4'-MorphoLuc** and **7'-MorPipLuc** with mutants 81 and 104 (Figure 3.4a). Selective light emission with these enzymes and substrates was verified in vitro (Figure 3.4b). We further validated the top ten unique pairings on the ranked list, along with a handful of others in the data set (every tenth rank among the top 100, every 100th rank among the top 1000, and every 1000th rank down to position 5000). In all cases, orthogonality scores were measured in bacterial lysate (Figure 3.4c). Among the top 1000 pairs, >10-fold photon outputs were observed with the positively paired luciferase-luciferin set compared to the negatively paired set (Figure 3.4c). Diminishments in selectivity were observed farther down the list. These results suggest that the in silico rank order is a good predictor of orthogonal substrate use. The method also culled 99.9% (more than 828,000 of the total 829,026) of irrelevant enzyme-substrate pairings, enabling fast convergence on important hits. As more luciferases and luciferins are screened, the data set can be expanded and continually mined for new orthogonal pairs.



**Figure 3.5:** Multicellular imaging with orthogonal pairs. Lead pairings of (a) mutants 53 and 87 and (b) 51 and 93 maintain orthogonality in mammalian cell culture. In each case, 100,000 DB7 cells were treated with 500  $\mu$ M luciferin analogue and imaged. Error bars represent SEM for experiments performed in triplicate.

### 3.2.3 Imaging with orthogonal pairs in cultured cell and animal models

We aimed to transition lead pairs from the screening analyses to mammalian cell imaging. In these more complex environments, issues of enzyme stability, substrate biocompatibility, and compound transport are of paramount concern. Fortunately, our approach to enriching functional luciferases preselects for luciferases and luciferins that are well behaved. Three of the top pairs from the script were analyzed in cultured cell (Figure 3.5) and animal models (Figure 3.6): (1) 4'-MorphoLuc/enzyme 81 (R218A, F250M, S314T, G316T) with 7'-DMAMeLuc/enzyme 37 (R218K), (2) 7'-MeLuc/enzyme 87 (R218K, F250Y, S314T, G316T) with 4'-BrLuc/enzyme 53 (V240I, V241M, F243M, F247Y, S347G), and (3) 4'-BrLuc/enzyme 51 (F243M, S347G) with D-luciferin/enzyme 93 (R218K, M249L, S314T, G316S). These pairs were selected due to the ease of accessing the substrates, along with their relative brightness. The mutants were stably expressed in DB7 mouse mammary carcinoma cells. The cells were treated with relevant luciferins and imaged (Figure 3.5). Substrate specificity was maintained in all cases, highlighting the success of the parallel screening method.

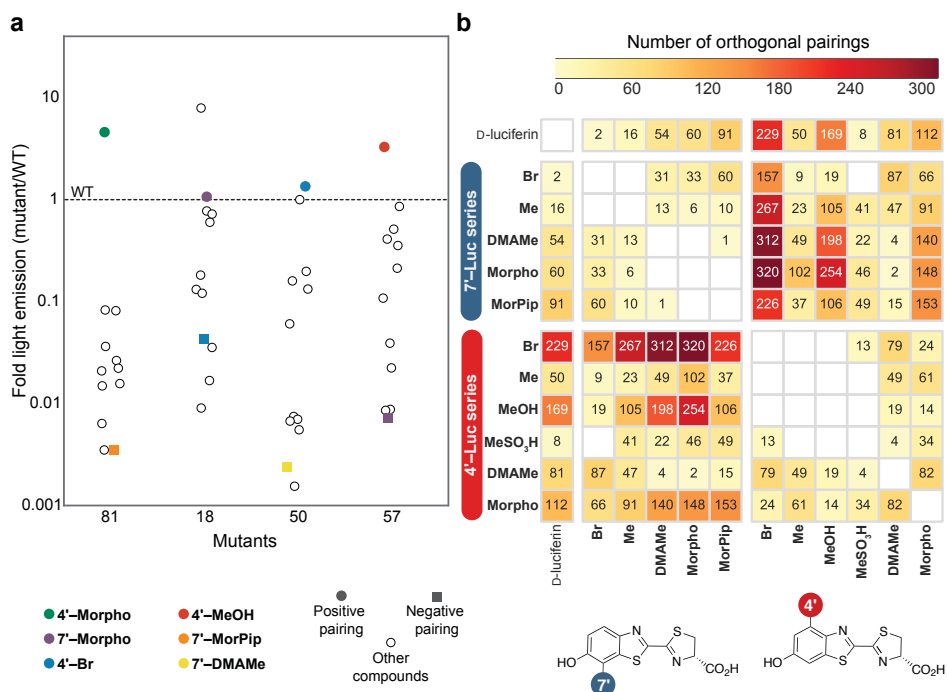


**Figure 3.6:** Noninvasive in vivo imaging with orthogonal pairs. DB7 cells expressing mutants 37 and 81 (a), 51 and 93 (b), or 53 and 87 (c) were inoculated in opposing flanks of FVB/NJ mice. The sites of implantation are indicated with dashed circles. Luciferin analogues were administered ip, and light emission was recorded. Representative bioluminescence images are shown for each set. For panel a, images were acquired 5 days post cell implantation. For panels b and c, images were acquired 3 days post cell implantation. Photon outputs were quantified and plotted. Black lines represent mean photon intensities for  $n = 3$  mice in each set.

Selectivity was also maintained in vivo. DB7 cells expressing the relevant mutants were implanted in FVB/NJ mice. Subsequent administration of the complementary luciferin analogues resulted in light emission for positively paired compounds with minimal cross reactivity (Figure 3.6). These images mark an initial demonstration of dual imaging with purely engineered luciferase-luciferin pairs. It is also important to note that perfect resolution is not required for multicomponent imaging applications. Rather, patterns of substrate use can serve as diagnostic fingerprints.<sup>44</sup> Photon outputs from the orthogonal pairs are in a useful range for monitoring bulk cell populations. The dimmest set (enzyme 37/**7'-DMAMeLuc** and enzyme 81/**4'-MorphoLuc**) emits enough photons to visualize  $6 \times 10^6$  cells in subcutaneous models. The other orthogonal sets are substantially brighter and can enable more sensitive imaging. Collectively, these data show that parallel screens and in silico analyses can be used to identify and transition orthogonal sets to a variety of biological models.

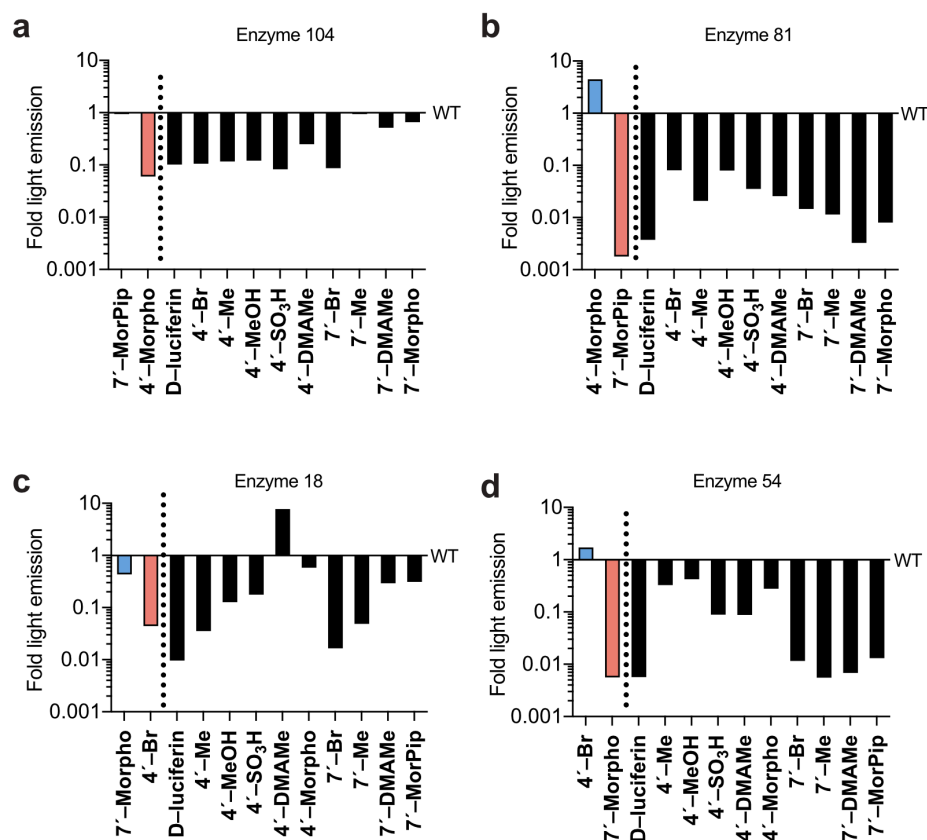
### 3.2.4 Analyzing trends in orthogonal substrate use

To gain insight into principles governing orthogonality, we undertook a detailed analysis of the screening results. The highest-ranked pairs comprised a variety of enzymes and substrates. Seven unique luciferins (from both the 4' and 7' series) were found among the top 10 pairs, along with luciferases comprising mutations at 18 unique sites. The diversity in hits implies that there are a variety of paths to achieve substrate resolution. Among the pairs, orthogonality was primarily realized not by markedly enhanced turnover of a preferred substrate. Rather, selectivity arose from reduced photon production with other compounds. As shown in Figure 3.7a, matched enzymes and substrates (positive pairs) were on par with native Fluc in terms of photon output. The unmatched enzymes and substrates (negative pairs), by contrast, demonstrated reduced activities (10-1000-fold lower). Thus, in a given orthogonal pair, selectivity is mostly achieved by reducing light emission with the negatively



**Figure 3.7:** Examining the origins of substrate selectivity. (a) In a given orthogonal pair, each luciferase retains activity with the matched luciferin (positive pairing, colored circle), while losing activity with the mismatched analogue (negative pairing, colored square). Mutant luciferases were also poorly reactive with all other mismatched analogues examined (open circles). These data suggest that orthogonality arises from selective retention of activity with a single compound. Light emission values are plotted relative to native Fluc and the indicated luciferin. (b) Frequency of luciferin analogue pairings predicted to be orthogonal. The majority of orthogonal sets from the top 5000 pairs (0.6%) comprise structurally divergent compounds (i.e., 4'-modified luciferins paired with 7'-modified luciferins).

paired compound versus selectively increasing light emission with the positively paired compound. For example, mutant 81 provides 4-fold enhanced light output with **4'-MorphoLuc** compared to Fluc. With every other luciferin screened, including the negatively paired compound **7'-MorPipLuc**, mutant 81 emits greater than 10-fold fewer photons than the native enzyme. So while light output with **4'-MorphoLuc** is slightly improved with mutant 81, the decrease in light emission observed with **7'-MorPipLuc** (greater than 100-fold) contributes more to orthogonality.



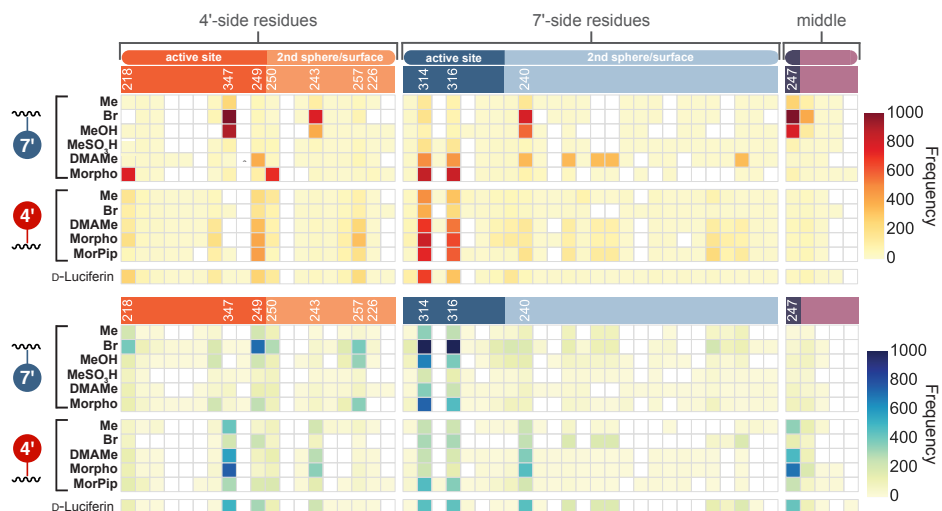
**Figure 3.8:** Selectivities of mutant enzymes for synthetic luciferins. Lead orthogonal enzymes (a) 104, (b) 81, (c) 18, and (d) 54 were treated with a panel of luciferin analogues and imaged. Light emission values are plotted as fold changes from Fluc with the respective compound (for normalization). The positive (blue) and negative (red) pairs are shown on the left side in each case.

Since compound selectivity appears to be achieved by destroying enzyme-substrate interactions, structurally related compounds would be expected to exhibit similar trends in orthogonality. Indeed, bulky 7'-modified compounds tend to positively pair with the same types of

enzymes (Figure 3.8). Many of these mutants (e.g., mutant 104) likely harbor space in the active site to accommodate 7' substituents (Figure 3.8a). Conversely, 4'-modified luciferins tend to produce less light with these same mutants and are thus negatively paired. Structurally divergent compounds were also more likely to comprise an orthogonal pair (Figure 3.7b). For example, substrates with a modification at the 4' position were rarely orthogonal to other 4' compounds. It is probably difficult to destroy activity with one 4'-modified compound without impacting others in the same series. When 4'- and 7'-modified substrates are paired, though, each substrate likely interacts differently with the enzyme, making it easier to achieve orthogonality. These results suggest a strategy for continued orthogonal bioluminescent probe development: incorporate more diverse analogues in parallel screens.

Some compounds appear uniquely suited for orthogonal probe development. For example, **4'-BrLuc** shows up nearly twice as often in the top 1,000 hits compared to other compounds. The mechanistic basis for this preference is unclear. The bromine substituent is roughly the same size as the methyl group in **4'-MeLuc**, negating a pure steric argument. **4'-MeOHLuc** is predicted to form orthogonal pairs (albeit less frequently) with similar compounds as **4'-BrLuc** (Figure 3.7), suggesting that polarizable substituents might be preferred. Heavy halogen atoms (e.g., Br) are also known to quench the fluorescence of some molecules via intersystem crossing.<sup>45</sup> Thus, certain **4'-BrLuc** conformations could result in internal quenching (and poor light emission) and thus pair negatively with several mutants. Additional compound screens and analyses will be necessary to discriminate among these possibilities and gain more insight.

We further analyzed the frequency of positive and negative pairings between luciferins and individual residues (Figure 3.9). Luciferases with mutations at residues 240, 247, or 347 seemed to prefer 4'-modified compounds. These residues are known to modulate the binding and light emission of the native substrate, D-luciferin.<sup>43,46-48</sup> Docking studies corroborated these findings, suggesting that the mutations (e.g., S347G in mutants 51 and 53) likely create



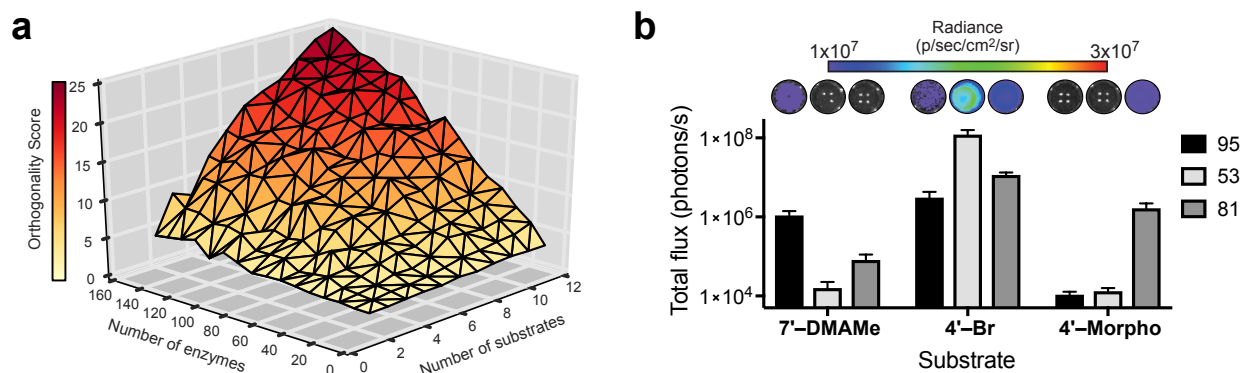
**Figure 3.9:** Orthogonal pair analysis. Heat map of mutation frequency (at a given residue) for enzymes positively paired (top) or negatively paired (bottom) with luciferin analogues. For each plot, the top 5000 pairs from in silico analyses were examined. Enzyme residues are organized by their relative proximity to C4' or C7' of luciferin (as in Figure 3.1b).

space for bulky substituents. These residues are also negatively paired with most of the 7'-modified compounds, suggesting that they are good candidates for future orthogonal probe design. Surprisingly few hot spot residues correlated with selective 7' analogue use (Figure 3.9). Fluc residues near C7' primarily comprise backbone amides.<sup>42,49</sup> Thus, it is unclear how specificity for these analogues might arise.

### 3.2.5 Added diversity improves orthogonality

Multicomponent imaging requires not just pairs of orthogonal enzymes and substrates, but also triplets, quadruplets, and higher order sets. Identifying such expanded collections requires structurally diverse enzyme and substrate architectures. If only a few privileged luciferases or luciferins from our data set could provide the desired selectivities, it would be difficult to achieve larger collections of orthogonal probes. To assess whether we were approaching an upper limit on orthogonality, we performed simulations within the existing data set. Random subsets of various sizes were selected from the full pool of substrates and





**Figure 3.10:** Improving orthogonality via enzyme-substrate diversity. (a) Orthogonality scores increase as more enzymes and substrates are considered. Computational analyses were performed on random subsets of luciferin analogues and mutant enzymes (from the entire data set). Orthogonality scores for all inputs were calculated as before, and the top orthogonal hits were averaged. (b) Validating an orthogonal triplet set. Bacteria expressing mutant luciferases 95, 53, or 81 were lysed and incubated with their corresponding luciferin (250  $\mu$ M). Sample images are shown. Photon outputs were quantified and error bars represent the standard error of the mean for  $n = 3$  experiments.

enzymes. The sets were analyzed using the algorithm from above, and orthogonality scores were generated (Figure 3.10a). Regardless of the identities of enzymes or substrates used, scores increased with greater numbers of both enzymes (from 2 to 159) and substrates (from 2 to 12). This result implies that we have not reached a plateau in identifying orthogonal pairs. Exploring more sequence space with mutant luciferases and chemical space with modified luciferins should also improve the orthogonality of the top pairings.

As a next step, we modified the algorithm to search for not just two pairs of orthogonal probes but also triplets and multiple sets in general. A set of three adds significant complexity, as not only three positive pairings, but also six negative pairings, must be identified. From our current data set, this required sifting through >144 million combinations. We combed the original data set in search of three mutually orthogonal enzyme-substrate pairs. A total of 6171 potential sets were identified. The top triplet set is shown in Figure 3.10b and comprises two enzyme-substrate pairs previously validated *in vivo*. The overall orthogonality score for this set was lower than that of the individual pairs from above. However, this result represents key proof-of-concept and a starting point for the development of larger collections

of mutually orthogonal luciferase-luciferin sets. Perfect selectivity is also not required for using the probes in biological environments. Rather, patterns in substrate use are most important and can be discerned using standard imaging equipment and rates of change in photon output.

### 3.3 Conclusions

We developed a general and rapid strategy to engineer orthogonal luciferase-luciferin pairs. The method relies on developing an initial pool of functional enzymes and screening the collection with chemically diverse luciferins. Using this approach, we generated >800,000 possible pairings and mined the data for orthogonal pairs with a custom computer algorithm. Dozens of candidates were identified and validated *in vitro*. A handful of hits were further translated into cultured cell and animal models, greatly expanding the number of bioluminescent probes for multicomponent imaging.

We further analyzed the principles governing orthogonal substrate use. Chemical and sequence diversity was key to eliciting high levels of selectivity. Thus, the addition of more luciferins and libraries to our living data set should improve orthogonality and lower the barrier to identifying higher-order sets. A fleet of sensitive, selective pairs will bolster imaging capabilities and push the boundaries of what we can see and learn about biological systems. The methods reported here are also applicable beyond the field of bioluminescence. Parallel screens and *in silico* analyses can expedite the search for other orthogonal enzyme-substrate or protein-ligand pairs relevant to optogenetics, cell signaling, and other disciplines.

## 3.4 Author contributions

The work described above was another collaborative effort between Will Porterfield, Krysten Jones, and me. My contributions in this chapter include developing the computer code for the *in silico* screening (Figure 3.4). I also analyzed the resulting orthogonal rankings (Figure 3.7, Figure 3.9, Figure 3.10). Below are the materials and methods concerning these contributions. For all other data/methods see the full version of the published paper.<sup>50</sup>

## 3.5 Materials and methods

### 3.5.1 General bioluminescence imaging

All analyses were performed in black 96-well plates (Greiner 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. Exposure times 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.

### 3.5.2 In silico screens

Scripts to search for orthogonal sets were written in Python3 and executed on either a MacBook Pro (for testing), or the High Performance Computing Cluster at UC Irvine (<https://hpc.oit.uci.edu/>). Searches for orthogonal pairs were run on 16 cores. Searches for triplet sets were run on 32 or 64 cores. The Python code and further documentation is

freely available online: <https://www.chem.uci.edu/jpresche/resources.html>. See section 3.6 for additional information regarding the mathematical evaluation of orthogonality.

### 3.5.3 Analyzing orthogonality as a function of diversity

Data were generated using Python3 in the Jupyter Notebook environment (code available at <https://www.chem.uci.edu/jpresche/resources.html>). For each point on the surface, the algorithm chose 5 random subsets of luciferins and luciferases from the pool of enzyme-substrate data. These subsets were then subjected to the *in silico* screen (see above) and the averages of the top 1,000 orthogonality scores were recorded. All scores were averaged if the total number was less than 1,000. The five subset averages were then further averaged together to give the value of the point.

## 3.6 Note on parallel screening *in silico*

Previously,<sup>25</sup> we evaluated the orthogonality of a set of two enzyme-substrate pairs using the following equation:

$$O = \log_{10} \left( \frac{E_A S_1}{E_A S_2} \right) * \log_{10} \left( \frac{E_B S_2}{E_B S_1} \right) \quad (3.1)$$

Where  $E_A S_1$  is the signal produced when enzyme A receives substrate 1,  $E_A S_2$  is the signal produced with substrate 2 and enzyme A, *etc.* Though functional for pairs of enzymes and substrates, this equation does not allow searches for sets of greater than two. Thus, we sought a new way to quantify orthogonality that could utilize any number of enzyme-substrate pairs.

For rapid iteration and data sorting, it is important to represent each set of enzyme-substrate pairs as a scalar value. Such an operation can be completed using basic linear algebra. For

the sake of this example, a set of three enzymes and three substrates will be used:

$$E_1, E_2, E_3 \text{ and } S_1, S_2, S_3$$

Consider that the reactivity of each substrate in each enzyme,  $r$ , can be represented in matrix form

$$M = \begin{matrix} & \begin{matrix} S_1 & S_2 & S_3 \end{matrix} \\ \begin{matrix} E_1 \\ E_2 \\ E_3 \end{matrix} & \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \end{matrix} \quad (3.2)$$

with enzymes arrayed into rows and substrates arrayed into columns. In this form, the ideal orthogonal case could be represented by the identity matrix.  $M_{ideal}$  shows three substrate-enzyme pairs where each substrate gives signal in only one enzyme.

$$M_{ideal} = \begin{matrix} & \begin{matrix} S_1 & S_2 & S_3 \end{matrix} \\ \begin{matrix} E_1 \\ E_2 \\ E_3 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{matrix} \quad (3.3)$$

Comparison of this ideal case to a matrix of experimental data via evaluation of the RMSD (root-mean-square displacement) will yield a scalar value that can be used to rank enzyme-substrate pairings: the lower the RMSD, the better the pairing.

$$\sqrt{\langle (M_{exp} - M_{ideal})^2 \rangle} = O_{\text{RMSD}} \quad (3.4)$$

In order to compare the ideal case to an experimental case, we must normalize and symmetrize the experimental matrix. Column-wise normalization is carried out to ensure that

each column vector is the same length. This is done by dividing each vector component by the total vector length (determined via the Pythagorean Theorem), thus producing a set of unit vectors:

$$\vec{S}_1 = \begin{matrix} & S_1 \\ E_1 & \left[ \begin{array}{c} r_{11} \\ r_{21} \\ r_{31} \end{array} \right] \\ E_2 & \\ E_3 & \end{matrix} \qquad \hat{S}_1 = \frac{\vec{S}_1}{|S_1|}$$

Where  $|S_1|$  is the length of  $\vec{S}_1$ , and  $\hat{S}_1$  is the unit vector of  $\vec{S}_1$ . When this process is repeated for  $\vec{S}_2$  and  $\vec{S}_3$ , the matrix can then be symmetrized via the equation:

$$M \cdot M^T = M_{sym} \tag{3.5}$$

Where ‘ $\cdot$ ’ denotes the matrix dot product,  $\hat{M}$  is the fully normalized matrix, and  $\hat{M}^T$  is the matrix transform.

Finally, as indicated in equation 3.4, the orthogonality of the set of vectors can be evaluated via comparison to the ideal case via RMSD.

$$\sqrt{\langle (M_{sym} - M_{ideal})^2 \rangle} = O_{RMSD}$$

For simpler readability, RMSD can be converted to a positive orthogonality value representing the geometric mean of the resolution of each substrate by each compound. This is calculated from comparison of the experimental RMSD,  $O_{RMSD}$  and the worst possible

RMSD (where  $M_{sym}$  is filled with ones),  $O_{worst}$  via a simple quotient.

$$M_{worst} = \begin{matrix} & S_1 & S_2 & S_3 \\ E_1 & \left[ \begin{array}{ccc} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{array} \right] \\ E_2 & \\ E_3 & \end{matrix} \quad O = 2 * \frac{O_{worst}}{O_{RMSD}} \quad (3.6)$$

# Bibliography

- [1] Specht, E. A.; Braselmann, E.; Palmer, A. E. *Annu. Rev. Physiol.* **2017**, *79*, 93–117.
- [2] Lavis, L. D.; Raines, R. T. *ACS Chem. Biol.* **2014**, *9*, 855–866.
- [3] Germain, R. N.; Robey, E. A.; Cahalan, M. D. *Science* **2012**, *336*, 1676–1681.
- [4] Cai, D.; Cohen, K. B.; Luo, T.; Lichtman, J. W.; Sanes, J. R. *Nat. Methods* **2013**, *10*, 540.
- [5] Porterfield, W. B.; Prescher, J. A. *Curr. Opin. Chem. Biol.* **2015**, *24*, 121–130.
- [6] Rodriguez, E. A.; Campbell, R. E.; Lin, J. Y.; Lin, M. Z.; Miyawaki, A.; Palmer, A. E.; Shu, X.; Zhang, J.; Tsien, R. Y. *Trends Biochem. Sci.* **2017**, *42*, 111–129.
- [7] Betzig, E.; Patterson, G. H.; Sougrat, R.; Lindwasser, O. W.; Olenych, S.; Bonifacino, J. S.; Davidson, M. W.; Lippincott-Schwartz, J.; Hess, H. F. *Science* **2006**, *313*, 1642–1645.
- [8] Prescher, J. A.; Contag, C. H. *Curr. Opin. Chem. Biol.* **2010**, *14*, 80–89.
- [9] Branchini, B. R.; Behney, C. E.; Southworth, T. L.; Fontaine, D. M.; Gulick, A. M.; Vinyard, D. J.; Brudvig, G. W. *J. Am. Chem. Soc.* **2015**, *137*, 7592–7595.
- [10] Troy, T.; Jekic-McMullen, D.; Sambucetti, L.; Rice, B. *Mol. Imaging* **2004**, *3*, 9–23.



- [11] Zhao, H.; Doyle, T. C.; Coquoz, O.; Kalish, F.; Rice, B. W.; Contag, C. H. *J. Biomed. Opt.* **2005**, *10*, 041210.
- [12] Rumyantsev, K. A.; Turoverov, K. K.; Verkhusha, V. V. *Sci. Rep.* **2016**, *6*, 36588.
- [13] Adams, S. T. J.; Miller, S. C. *Curr. Opin. Chem. Biol.* **2014**, *21*, 112–120.
- [14] Xu, T.; Close, D.; Handagama, W.; Marr, E.; Sayler, G.; Ripp, S. *Front. Oncol.* **2016**, *6*, 41210.
- [15] Branchini, B. R.; Ablamsky, D. M.; Murtiashaw, M. H.; Uzasci, L.; Fraga, H.; Southworth, T. L. *Anal. Biochem.* **2007**, *361*, 253–262.
- [16] Nakatsu, T.; Ichiyama, S.; Hiratake, J.; Saldanha, A.; Kobashi, N.; Sakata, K.; Kato, H. *Nature* **2006**, *440*, 372–376.
- [17] Conley, N. R.; Dragulescu Andrasi, A.; Rao, J.; Moerner, W. E. *Angew. Chem. Int. Ed.* **2012**, *51*, 3350–3353.
- [18] Branchini, B. R.; Hayward, M. M.; Bamford, S.; Brennan, P.; Lajiness, E. J. *Photochem. Photobiol.* **1989**, *49*, 689–695.
- [19] Woodroffe, C. C.; Meisenheimer, P. L.; Klaubert, D. H.; Kovic, Y.; Rosenberg, J. C.; Behney, C. E.; Southworth, T. L.; Branchini, B. R. *Biochemistry* **2012**, *51*, 9807–9813.
- [20] McCutcheon, D. C.; Paley, M. A.; Steinhardt, R. C.; Prescher, J. A. *J. Am. Chem. Soc.* **2012**, *134*, 7604–7607.
- [21] Kuchimaru, T.; Iwano, S.; Kiyama, M.; Mitsumata, S.; Kadonosono, T.; Niwa, H.; Maki, S.; Kizaka-Kondoh, S. *Nat. Commun.* **2016**, *7*, 11856.
- [22] Woodroffe, C. C.; Shultz, J. W.; Wood, M. G.; Osterman, J.; Cali, J. J.; Daily, W. J.; Meisenheimer, P. L.; Klaubert, D. H. *Biochemistry* **2008**, *47*, 10383–10393.
- [23] Mofford, D. M.; Reddy, G. R.; Miller, S. C. *J. Am. Chem. Soc.* **2014**, *136*, 13277–13282.

- [24] Steinhardt, R. C.; Rathbun, C. M.; Krull, B. T.; Yu, J. M.; Yang, Y.; Nguyen, B. D.; Kwon, J.; McCutcheon, D. C.; Jones, K. A.; Furche, F.; Prescher, J. A. *ChemBioChem* **2016**,
- [25] Jones, K. A.; Porterfield, W. B.; Rathbun, C. M.; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. *J. Am. Chem. Soc.* **2017**, *139*, 2351–2358.
- [26] Jathoul, A. P.; Grounds, H.; Anderson, J. C.; Pule, M. A. *Angew. Chem., Int. Ed.* **2014**, *53*, 13059–13063.
- [27] Rathbun, C. M.; Prescher, J. A. *Biochemistry* **2017**, *56*, 5178–5184.
- [28] Nishihara, R.; Abe, M.; Nishiyama, S.; Citterio, D.; Suzuki, K.; Kim, S. B. *Sci. Rep.* **2017**, *7*, 908.
- [29] Kaskova, Z. M. et al. *Sci. Adv.* **2017**, *3*, e1602847.
- [30] Paley, M. A.; Prescher, J. A. *Med. Chem. Comm.* **2014**, *5*, 255–267.
- [31] Stacer, A. C.; Nyati, S.; Moudgil, P.; Iyengar, R.; Luker, K. E.; Rehemtulla, A.; Luker, G. D. *Mol. Imaging* **2013**, *12*, –13.
- [32] Petushkov, V. N.; Dubinnyi, M. A.; Tsarkova, A. S.; Rodionova, N. S.; Baranov, M. S.; Kublitski, V. S.; Shimomura, O.; Yampolsky, I. V. *Angew. Chem. Int. Ed.* **2014**, *53*, 5566–5568.
- [33] Renata, H.; Wang, Z. J.; Arnold, F. H. *Angew. Chem. Int. Ed.* **2015**, *54*, 3351–3367.
- [34] Packer, M. S.; Liu, D. R. *Nat. Rev. Genet.* **2015**, *16*, 379–394.
- [35] Goldsmith, M.; Tawfik, D. S. *Curr. Opin. Struct. Biol.* **2012**, *22*, 406–412.
- [36] 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. *Nat. Methods* **2015**, *13*, 177–183.

- [37] 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.; Kelly, J. W. *eLife* **2015**, *4*, e10606.
- [38] Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W. *Nature* **2010**, *464*, 441–444.
- [39] Lang, K.; Chin, J. W. *Chem. Rev.* **2014**, *114*, 4764–4806.
- [40] Owens, A. E.; Grasso, K. T.; Ziegler, C. A.; Fasan, R. *ChemBioChem* **2017**, *18*, 1109–1116.
- [41] Wan, W.; Huang, Y.; Wang, Z.; Russell, W. K.; Pai, P.-J.; Russell, D. H.; Liu, W. R. *Angew. Chem. Int. Ed.* **2010**, *49*, 3211–3214.
- [42] Sundlov, J. A.; Fontaine, D. M.; Southworth, T. L.; Branchini, B. R.; Gulick, A. M. *Biochemistry* **2012**, *51*, 6493–6495.
- [43] Harwood, K. R.; Mofford, D. M.; Reddy, G. R.; Miller, S. C. *Chemistry & Biology* **2011**, *18*, 1649–1657.
- [44] Peacor, B. C.; Ramsay, C. M.; Waters, M. L. *Chem. Sci.* **2017**, *8*, 1422–1428.
- [45] Solovyov, K. N.; Borisevich, E. A. *Physics-Uspekhi* **2005**, *48*, 231–253.
- [46] Viviani, V. R.; Prado, R. A.; Neves, D. R.; Kato, D.; Barbosa, J. A. *Biochemistry* **2013**, *52*, 3963–3973.
- [47] Branchini, B. R.; Southworth, T. L.; Murtiashaw, M. H.; Boije, H.; Fleet, S. E. *Biochemistry* **2003**, *42*, 10429–10436.
- [48] Branchini, B. R.; Ablamsky, D. M.; Rosenman, J. M.; Uzasci, L.; Southworth, T. L.; Zimmer, M. *Biochemistry* **2007**, *46*, 13847–13855.

- [49] Viviani, V. R.; Amaral, D. T.; Neves, D. R.; Simoes, A.; Arnoldi, F. G. C. *Biochemistry* **2013**, *52*, 19–27.
- [50] Rathbun, C. M.; Porterfield, W. B.; Jones, K. A.; Sagoe, M. J.; Reyes, M. R.; Hua, C. T.; Prescher, J. A. *ACS Cent. Sci.* **2017**, *3*, 1254–1261.

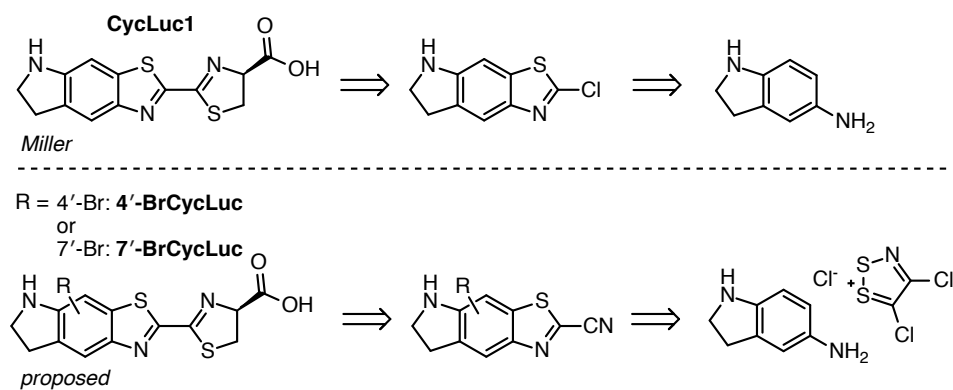
# Chapter 4

## Brominated luciferins for multicomponent imaging and cross-coupling

### 4.1 Introduction

Luciferin analog synthesis and characterization is central to orthogonal probe development. This is due to the inherently molecular basis of bioluminescence; changes to the luciferin core can drastically effect catalysis and light emission.<sup>1,2</sup> With a method for selection of orthogonal luciferin-luciferase pairs in hand, we turned to optimization of these pairings for in vivo use, and expanding the number of luciferin architectures for orthogonal probes.

To design probes that would be more amenable to in vivo imaging, we set out to improve the enzymatic parameters of our lead compounds. Kinetic studies of initial pairings indicated that our sterically modified luciferins had drastically increased  $K_M$  values (Table 2.1). Not only were these values inflated for the negative pairings, they were also increased in the



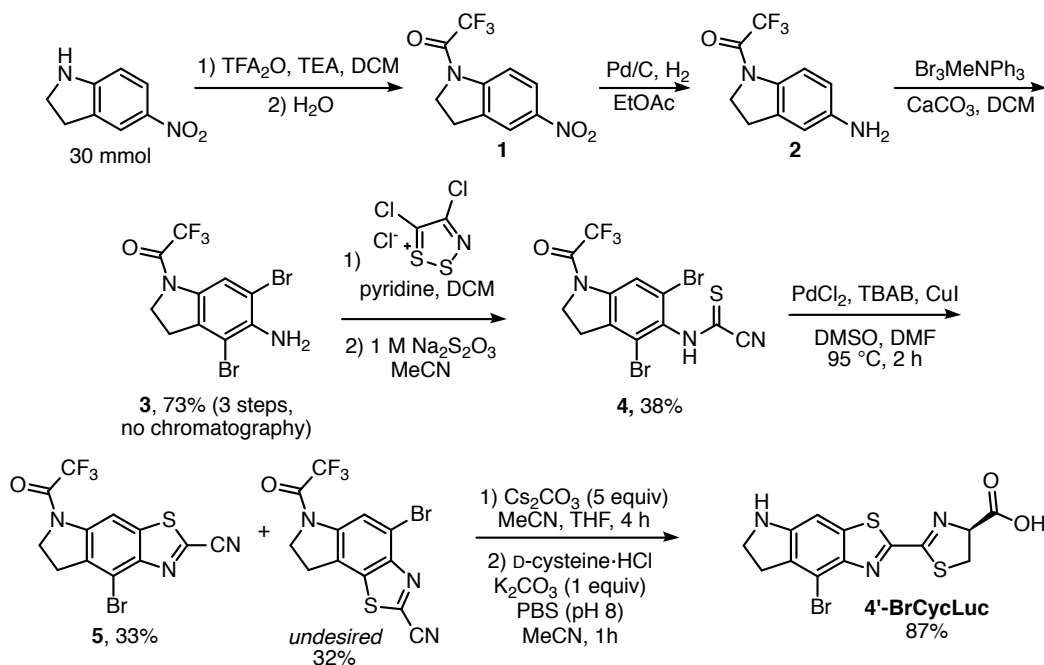
Scheme 4.1: Cyclic luciferin retrosyntheses. *Top*: Miller's retrosynthesis of CycLuc1. *Bottom*: Proposed retrosynthesis of modified CycLuc scaffolds using our Appel's salt chemistry.

positively-paired combination. Reduction in luciferin  $K_M$  values has proven beneficial for in vivo imaging.<sup>3</sup> One example from the Miller group and our lab involved **CycLuc1**, a constrained aminoluciferin with improved in vivo properties. Conveniently, **CycLuc1** contained modifications to only the 5' and 6' positions of the luciferin scaffold, leaving the 4' and 7' positions open.<sup>4,5</sup> We hypothesized that addition of our lead modifications to these sites on CycLuc1 may decrease  $K_M$ , yet retain selectivity for their respective mutants. As demonstrated in chapter 2 and chapter 3, the most promising compounds in terms of orthogonality and brightness were the methyl and bromo series (in addition to D-luc). Methyl- and bromo-modified CycLuc scaffolds could likely be accessed via our previously reported route to other functionalized luciferins (Scheme ??).<sup>6</sup> In this strategy, Appel's salt would be used to install the fused thiazole ring to the indole, while simultaneously installing the cyano group for final condensation with cysteine.

Prior to cysteine condensation, these molecules could also be good substrates for cross-coupling. Metal-catalyzed cross coupling is an excellent technique for divergent synthesis. With a single halogenated precursor, numerous commercially available coupling partners can be appended in a reliable fashion.<sup>7,8</sup> Such chemistries would enable rapid diversification of the luciferin arsenal. As shown in the preceding chapters, greater numbers of compounds and enzymes bolster the pursuit of orthogonality.

## 4.2 Synthesis of brominated cyclic luciferins

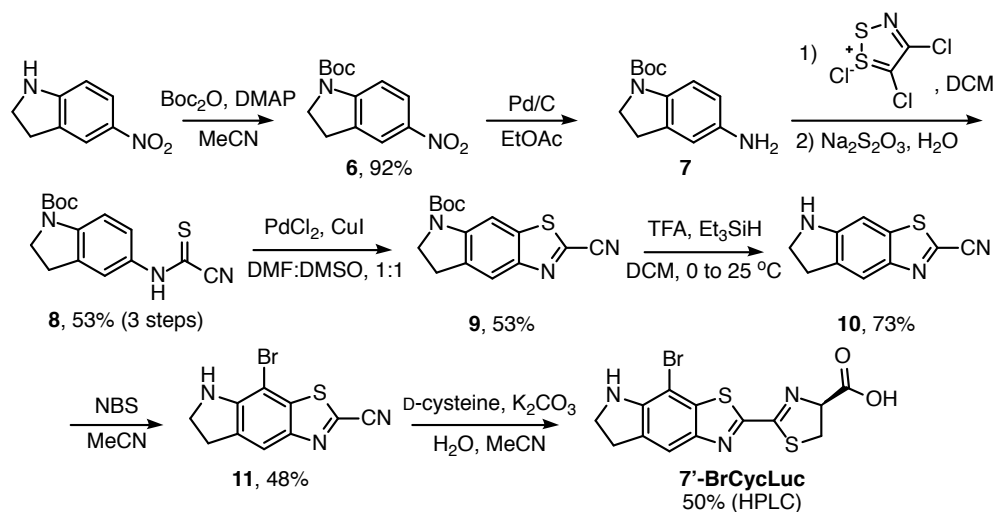
I initially targeted the 4' and 7' bromoluciferins, as these were well vetted compounds from our previous work, and were amenable to further diversification via cross coupling. Indoline bearing a nitro group at the necessary position was commercially available, and was easily TFA protected to give **1** (Scheme 4.2).<sup>4,5</sup> Reduction with palladium on carbon proceeded smoothly to provide **2** without need for purification. Indoline **2** was then dibrominated (**3**), and condensed with Appel's salt (**4**), via our previously described sequence.<sup>6</sup> Palladium-mediated cyclization of thioformamide **4** gave a mixture of two regioisomers, which were separable by column chromatography. Taking the desired isomer (**5**) forward, a subsequent one-pot deprotection and condensation with D-cysteine gave **4'-BrCycLuc**.



Scheme 4.2: Synthesis of **4'-BrCycLuc**. Our route, including Appel's salt condensation, proved general for the CycLuc scaffolds.

Synthesis of **7'-BrCycLuc** proceeded similarly. Compounds **6-10** were isolated following procedures already reported in the lab. Unexplored transformations began with bromination at the 7' position later in the sequence following deprotection of benzothiazole **10**. Bromi-

nation with NBS proceeded as expected via conditions previously described in our group.<sup>6</sup> Condensation of **11** with D-cysteine provided **7'-BrCycLuc** following HPLC purification.

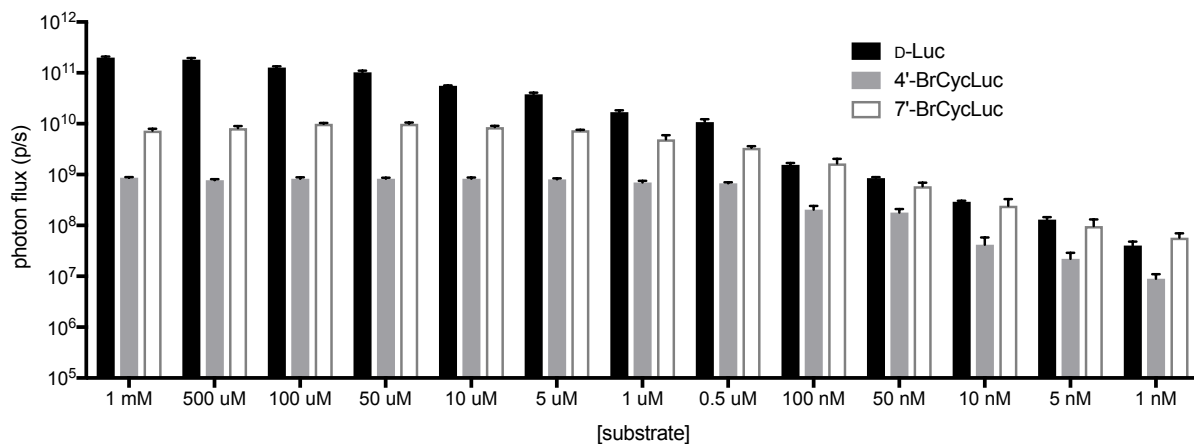


Scheme 4.3: Synthesis of **7'-BrCycLuc**. Our route, including Appel's salt condensation, proved general for the CycLuc scaffolds.

### 4.3 Characterization of brominated cyclic luciferins

Once synthesized, the cyclized compounds were evaluated with purified Fluc in a dose-response assay (Figure 4.1). Due to the CycLuc core, we expected a decrease in apparent  $K_M$  relative to **4'-** and **7'-BrLuc**. Gratifyingly, we observed this to be the case. While photon output with D-luciferin began to decrease after 500  $\mu\text{M}$ , **4'-** and **7'-BrCycLuc** did not begin to lose light emission until 1  $\mu\text{M}$ . This difference in binding caused **7'-BrCycLuc** light output to match that of D-luciferin at 100 nM. In comparison to our original brominated luciferins, this is a vast improvement. Previous work determined **4'-** and **7'-BrLuc**  $K_M$  values to be 2 and 17  $\mu\text{M}$  respectively. Our initial measurements in Figure 4.1 indicate that the brominated CycLucs could have  $K_M$ 's up to two orders of magnitude lower. With decreased  $K_M$  values, these luciferins could be more amenable to in vivo imaging, as substrate concentrations can tend to be low (1-10  $\mu\text{M}$ ) at the site of luciferase expressing cells.<sup>9</sup>



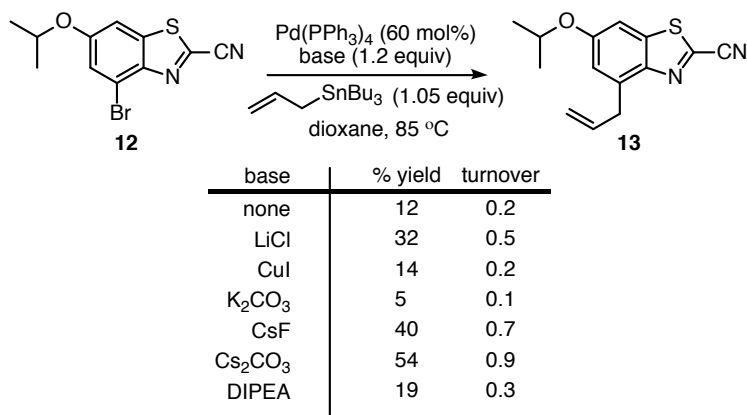


**Figure 4.1:** Dose response of brominated cyclic luciferins. The cyclic scaffolds exhibited lower  $K_M$  values. Each luciferin was imaged in triplicate with 1  $\mu\text{g}$  of Fluc enzyme and 100  $\mu\text{M}$  ATP.

## 4.4 Stille cross-coupling of brominated luciferin precursors

It had not escaped our attention that brominated probes could also be excellent cross coupling substrates. Initial work in the lab by Rachel Steinhardt showed Stille couplings were possible at the 4' position of brominated luciferin **12** (Table 4.1). Rachel used Stille conditions to install an allyl group to give **13** in low yield. Though these reactions were not catalytic in palladium, they demonstrated that cross coupling could be used to diversify brominated luciferins.

Inspired by Rachel's initial work, I screened a panel of reaction additives, hypothesizing that counterion was important for product formation (Table 4.1). I found that optimal reactivity was observed with  $\text{Cs}_2\text{CO}_3$  as a base. However, I did not observe catalyst turnover (with a yield less than my catalyst loading), implying that the palladium tetrakis was unable to complete the catalytic cycle. Stille reactions also require the use of toxic organotin reagents, limiting user friendliness and commercially available coupling partners. Thus, I tabled these

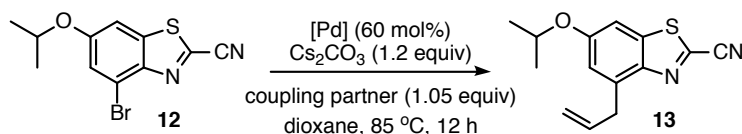


**Table 4.1:** Evaluation of Stille cross coupling conditions. Percent yield was determined by <sup>1</sup>H NMR. Turnover is defined as percent yield divided by catalyst loading.

conditions to work on a reaction that is compatible with boronic acid and ester coupling partners.

## 4.5 Suzuki cross coupling of brominated luciferin precursors

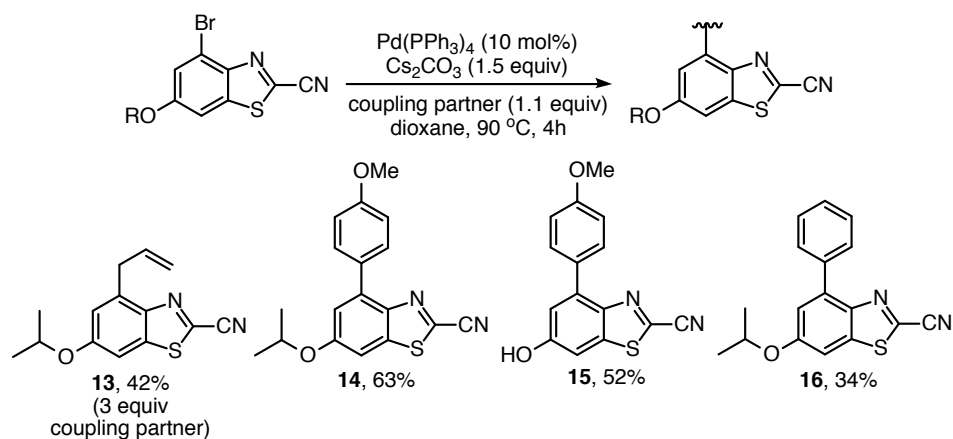
With the initial success of palladium tetrakis in the Stille methodology, I was struck by the similarity of the reaction conditions to Suzuki cross coupling. Thus I attempted a screen of similar reaction conditions using allyl boron coupling partners. Though an initial attempt with conditions optimized for allyl coupling partners did not yield product (Table 4.2 entry 1), a switch to palladium tetrakis (used in the previous conditions with tin) yielded product as evidenced by <sup>1</sup>H NMR spectroscopy (entry 2). It is possible that more coordinately saturated palladium species perform better in these reactions due to product chelation of the metal. Switching to a boronic ester further improved yield, and increasing the temperature of the reaction provided enough energy for catalyst turnover (entry 4). Additionally, increasing the equivalents of coupling partner further improved yield and turnover (entry 5).



entry	catalyst	coupling partner	conditions	% yield	turnover
1	Pd <sub>2</sub> DBa <sub>3</sub>	allyl-BF <sub>4</sub>	K <sub>3</sub> PO <sub>4</sub> , PCy <sub>3</sub> , 85 °C	NR	NR
2	Pd(PPh <sub>3</sub> ) <sub>4</sub>	allyl-BF <sub>4</sub>	85 °C	17	0.3
3	"	allyl-Bpin	85 °C, 10% v H <sub>2</sub> O	30	0.5
4	"	"	11 mol%, 120 °C, 10% v H <sub>2</sub> O, 12 h	31	2.8
5	"	allyl-Bpin (3 equiv)	10 mol%, 1.5 equiv base, 125 °C, 4 h	51	5.1

**Table 4.2:** Evaluation of Suzuki cross coupling conditions. Percent yield was determined by <sup>1</sup>H NMR. Turnover is defined as percent yield divided by catalyst loading.

With these conditions in hand, we were eager to test the generality of the cross coupling with other luciferin precursors and coupling partners. Conditions proved to be general with aryl coupling partners, even resulting in a small improvement in yield (compounds **14** and **16**). These cross coupling conditions also were tolerant of a deprotected luciferin substrate (**15**).



**Scheme 4.4:** Preliminary scope of cross coupling with brominated luciferins. We observed promising yields with both protected and unprotected luciferin precursors. Isolated yields are listed.

## 4.6 Conclusions and future directions

In an effort to improve brightness of our lead compounds, we hypothesized that rigidified amines would offer a reduction in  $K_M$  values. Brominated CycLuc scaffolds were thus synthesized and assayed with Fluc. We found that these molecules exhibited improved  $K_M$  values characteristic of the CycLuc family of compounds. Future work will evaluate these compounds with a range of mutant luciferases, as well as in mammalian cells and in vivo.

In tandem with rational probe design, the ability to cover a large swath of luciferin structure-space would accelerate our efforts in identification of orthogonal bioluminescent probes. Metal cross coupling of brominated luciferin precursors shows promise in this regard. We have demonstrated that cross coupling is possible under standard Stille and Suzuki conditions. Modest yields and catalyst turnovers were observed with the cross coupling of allyl groups. Further, a variety of luciferin backbones and substitutions gave coupling product. Future work will identify conditions that provide increases in reaction efficiency so that probes can be synthesized in large quantities.

## 4.7 Materials and methods

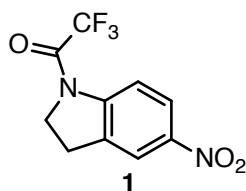
### 4.7.1 General synthetic methods

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 as previously reported.<sup>10,11</sup> Reaction progress was monitored by thin-layer chromatography on EMD 60 F254 plates, visualized with UV light, ceric ammonium molybdate (CAM), chloranil, or  $\text{KMnO}_4$  stain. Compounds were purified via flash column chromatography using SiliaFlash F60, 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  $\mu$ m), preparative runs were performed using an Agilent Eclipse XD8-C18 PrepHT column (21.2 x 250 mm 7  $\mu$ m). 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.  $^1\text{H}$ -NMR spectra were acquired at either 500 or 400 MHz, and  $^{13}\text{C}$ -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.

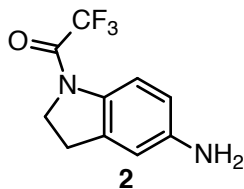
#### 4.7.2 Compound Syntheses

##### 2,2,2-trifluoro-1-(5-nitroindolin-1-yl)ethan-1-one



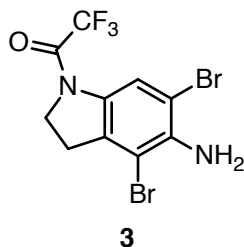
Procedure was used directly from Miller *et al.*<sup>4</sup> Spectral data matched those previously published.

### 1-(5-aminoindolin-1-yl)-2,2,2-trifluoroethan-1-one



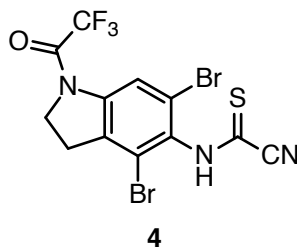
To a flask containing **1** (6.1 mmol, 1.6 g) was added 60 mL ethyl acetate and stirred. The flask was sealed and purged with a steady flow of nitrogen. Palladium on carbon (160 mg) was carefully added to the purged flask. Hydrogen gas was bubbled through the solution via a needle-connected balloon. Once the reaction was determined complete by TLC, the solution was slowly filtered through celite and reduced. The crude material was taken on directly.

### 1-(5-amino-4,6-dibromoindolin-1-yl)-2,2,2-trifluoroethan-1-one



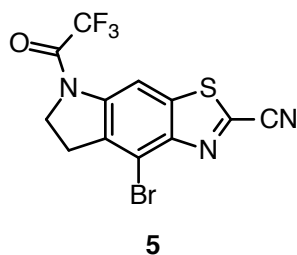
Procedure was adapted from Steinhard *et al.*<sup>6</sup> To a stirring solution of **2** (1.3 mmol, 300 mg) and CaCO<sub>3</sub> (2.1 mmol, 270 mg) in DCM was slowly added PhNMe<sub>3</sub>Br<sub>3</sub>. A precipitate was observed. Water was added (100 mL) and the solution turned violet. After stirring for an additional hour, the organic layer was washed with Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> (2x100 mL), and brine (2x100 mL). The solution was dried over MgSO<sub>4</sub> and concentrated by rotary evaporation to yield a light orange powder (420 mg, 82%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.33 (s, 1H), 4.28 (dd, *J* = 12.1, 4.4 Hz, 2H), 3.18 (t, *J* = 8.2 Hz, 2H). <sup>19</sup>F NMR (301 MHz, CDCl<sub>3</sub>) δ -73.33.

(4,6-dibromo-1-(2,2,2-trifluoroacetyl)indolin-5-yl)carbamothioyl cyanide



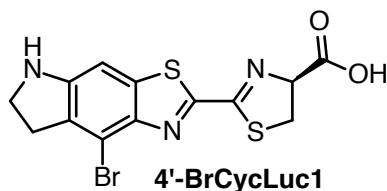
A dry round bottom flask was charged with **3** (0.52 mmol, 200 mg) and dry DCM (50 mL). Appel's salt (1.56 mmol, 325 mg) was added, and the solution was stirred under nitrogen for 30 min. Pyridine (3.12 mmol, 252  $\mu$ L) was then added dropwise. After full consumption of starting material was observed by TLC (about 1 hour), a 1 M solution of  $\text{Na}_2\text{S}_2\text{O}_3$  (10 mL) was added. Following full consumption of intermediate (observed by TLC, about 3 hours), the solution was washed with additional  $\text{Na}_2\text{S}_2\text{O}_3$  (2x50 mL) and brine (3x50 mL), dried over  $\text{MgSO}_4$  and concentrated by rotary evaporation. Flash column chromatography (10% ethyl acetate in hexanes) yielded a dark brown-red oil (107 mg, 45%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.86 (s, 1H), 8.64–8.50 (m, 1H), 4.41 (dd,  $J = 18.4, 8.9$  Hz, 2H), 3.39–3.21 (m, 2H).  $^{19}\text{F}$  NMR (301 MHz,  $\text{CDCl}_3$ )  $\delta$  -73.09.

8-bromo-5-(2,2,2-trifluoroacetyl)-6,7-dihydro-5H-thiazolo[4,5-f]indole-2-carbonitrile



A dry round bottom flask was charged with **4** (4.2 mmol, 1.9 g), PdCl<sub>2</sub> (0.84 mmol, 150 mg), CuI (2.1 mmol, 400 mg), and tetrabutylammonium bromide (8.4 mmol, 2.7 g). The flask was sealed with a septum and placed under nitrogen. Dry DMSO (40 mL) and DMF (40 mL) were then added, and the mixture was stirred at 95 °C for 3 h. Reaction progress was monitored by <sup>1</sup>H NMR. Upon consumption of starting material, ethyl acetate was added (100 mL) and the solution was washed with water (5x100 mL) and brine (2x100 mL), and dried over MgSO<sub>4</sub> and concentrated by rotary evaporation. Flash column chromatography (5 to 30% ethyl acetate in hexanes) gave the final product as a white powder (520 mg, 33%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.78 (s, 1H), 4.49 (t, *J* = 8.3 Hz, 2H), 3.46 (t, *J* = 8.3 Hz, 2H). <sup>19</sup>F NMR (301 MHz, CDCl<sub>3</sub>) δ -73.11.

**(*S*)-2-(8-bromo-6,7-dihydro-5*H*-thiazolo[4,5-*f*]indol-2-yl)-4,5-dihydrothiazole-4-carboxylic acid**

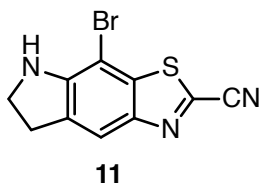


To a scintillation vial containing a stir bar and **5** (10 mg, 0.027 mmol) was added MeCN (0.5 mL) and THF (0.5 mL). The solution was stirred. Once dissolved, Cs<sub>2</sub>CO<sub>3</sub> (45 mg, 0.13 mmol) was added and the solution was stirred for 4 h or until the full consumption of starting material was observed by TLC. In a separate vial, D-cysteine·HCl (4.9 mg, 0.028 mmol), K<sub>2</sub>CO<sub>3</sub> (3.9 mg, 0.028 mmol) and phosphate buffered saline (pH 8, 0.5 mL) were combined. The solution of D-cysteine was then added dropwise to the stirring scintillation vial. An immediate color change to a bright yellow-red solution was observed. The solution was stirred for an additional 1 h, and transferred to a separatory funnel. The basic solution was gently washed with ethyl acetate (2x5 mL) and then acidified to pH 2. The product



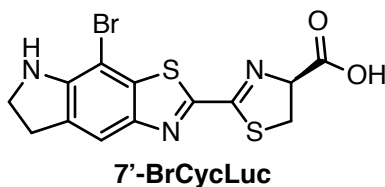
was extracted into an organic layer of ethyl acetate (3x5 mL), and washed with water (2x, 5 mL). The solution was concentrated via rotary evaporation to yield a red film that required no further purification (9 mg, 87%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 6.88 (d, *J* = 10.8 Hz, 1H), 5.34 (dd, *J* = 9.4, 8.7 Hz, 1H), 3.76–3.65 (m, 4H), 3.14 (t, *J* = 8.3 Hz, 2H).

#### 4-bromo-6,7-dihydro-5*H*-thiazolo[4,5-*f*]indole-2-carbonitrile



Procedure was adapted from Steinhard *et al.*<sup>6</sup> A dry round bottom flask was charged with **10** (43 mg, 0.21 mmol) followed by freshly recrystallized NBS (42 mg, 0.24 mmol). The mixture was taken up in MeCN (5 mL) and stirred at room temperature for an hour. The crude reaction mixture was loaded directly onto silica and purified via flash column chromatography (5 to 20% ethyl acetate in hexanes) to yield a white solid (29 mg, 48%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.74 (s, 1H), 3.87–3.78 (m, 2H), 3.39–3.29 (m, 2H).

#### (*S*)-2-(4-bromo-6,7-dihydro-5*H*-thiazolo[4,5-*f*]indol-2-yl)-4,5-dihydrothiazole-4-carboxylic acid



Condensation was conducted as described in the synthesis of **4'-BrCycLuc**. Following isolation, preparatory HPLC purification was carried out with the following elution protocol:

a gradient of 50% MeOH in 25 mM  $\text{PO}_4\text{K}_3\text{H}$  (pH 7.8) over 5 min, increasing to 90% MeCN over 20 min, at 4 mL/min. An orange-red film was isolated (20 mg, 50%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.07 (s, 1H), 5.30–5.22 (m, 1H), 3.75–3.60 (m, 5H), 3.07 (t,  $J = 8.7$  Hz, 2H).

### **General procedure for Suzuki cross coupling of brominated luciferin precursors**

To a flame-dried Schlenk tube was added luciferin precursor (0.034 mmol), palladium tetrakis (4 mg, 0.003 mmol), and  $\text{Cs}_2\text{CO}_3$  (17 mg, 0.051 mmol). The tube was evacuated and backfilled with nitrogen three times. Dry dioxane (0.5 mL) was added via syringe. The solution was stirred under nitrogen for 10 min. Finally, boronic acid or ester coupling partner was added (0.1 mmol), the tube was sealed, and the reaction was heated to 120 °C for 2 h. Reaction mixtures were run through a silica plug prior to analysis of the crude material by  $^1\text{H}$  NMR.

# Bibliography

- [1] Rathbun, C. M.; Prescher, J. A. *Biochemistry* **2017**, *56*, 5178–5184.
- [2] Adams, S. T.; Miller, S. C. *Curr. Opin. Chem. Biol.* **2014**, *21*, 112–120.
- [3] Evans, M. S.; Chaurette, J. P.; Adams, S. T.; Reddy, G. R.; Paley, M. A.; Aronin, N.; Prescher, J. A.; Miller, S. C. *Nat. Methods* **2014**, *11*, 393–395.
- [4] Reddy, G. R.; Thompson, W. C.; Miller, S. C. *J. Am. Chem. Soc.* **2010**, *132*, 13586–13587.
- [5] Mofford, D. M.; Reddy, G. R.; Miller, S. C. *J. Am. Chem. Soc.* **2014**, *136*, 13277–13282.
- [6] Steinhardt, R. C.; Rathbun, C. M.; Krull, B. T.; Yu, J. M.; Yang, Y.; Nguyen, B. D.; Kwon, J.; McCutcheon, D. C.; Jones, K. A.; Furche, F.; Prescher, J. A. *ChemBioChem* **2016**, *18*, 96–100.
- [7] Miyaura, N.; Suzuki, A. *Chem. Rev.* **1995**, *95*, 2457–2483.
- [8] Nicolaou, K. C.; Bulger, P. G.; Sarlah, D. *Angew. Chem. Int. Ed.* **2005**, *44*, 4442–4489.
- [9] Berger, F.; Paulmurugan, R.; Bhaumik, S.; Gambhir, S. S. *Eur. J. Nucl. Med. Mol. Imaging* **2008**, *35*, 2275–2285.
- [10] McCutcheon, D. C.; Porterfield, W. B.; Prescher, J. A. *Org. Biomol. Chem.* **2015**, *13*, 2117–2121.

- [11] McCutcheon, D. C.; Paley, M. A.; Steinhardt, R. C.; Prescher, J. A. *J. Am. Chem. Soc.* **2012**, *134*, 7604–7607.

# Chapter 5

## Multicomponent bioluminescence imaging via substrate unmixing

### 5.1 Introduction

A cornerstone of fluorescence imaging is the ability to monitor multiple subjects in tandem.<sup>1</sup> Indeed, spatiotemporal information about a species of interest is of limited use without the ability to also visualize its interacting partners. Multicolored fluorescent probes have brought a multitude of cellular interactions to light, from the interactions between immune and cancer cells,<sup>2</sup> to the contacts formed between neurons.<sup>3</sup> However, these studies have been historically limited to short length scales. Fluorescence imaging in thick tissues, or in entire opaque organisms remains a challenge. Excitation light causes autofluorescence of endogenous fluorophores (such as flavin and heme), and all but red emission light is absorbed and scattered by tissue.<sup>4</sup>

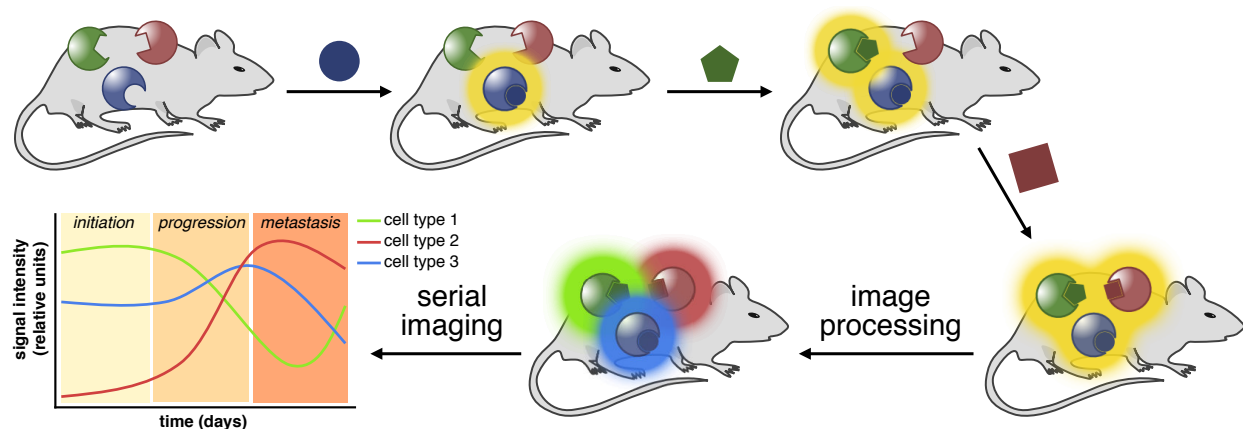
As highlighted in previous chapters, bioluminescence imaging (BLI) is well suited for use in complex tissues. BLI utilizes genetically encoded luciferase enzymes and exogenously

administered luciferin substrates to emit photons of light. Because tissue does not normally glow, this technique is exquisitely sensitive, with the ability to detect 1-10 cells in a mouse. This sensitivity has made BLI the go-to technique for monitoring cell proliferation and migration in mouse models.<sup>5</sup> Despite the utility of the tool, multicomponent imaging has yet to be developed for BLI. With such a capability, longstanding questions regarding cellular interactions in an organismal context could be answered. Different classes of immune cells could be tracked as they respond to disease states, stem cells labeled with different probes could be visualized as they form organoids, or the cellular makeup early metastases could be monitored.

Early efforts in multicomponent BLI capitalized on the natural substrate resolution present in nature. Initial reports paired firefly luciferase with Gaussia or Renilla luciferases.<sup>6-8</sup> These studies used the resolved probes in the same cell to simultaneously monitor cell localization and proliferation as well as expression of a gene of interest. Three component imaging was recently realized with firefly luciferase, Gaussia luciferase, and Vargula luciferase.<sup>9</sup> The authors used the three probes to study the effects of gene therapy on glioma proliferation. Glioma cells were monitored with firefly luciferase, gene delivery was quantified via Vargula luciferase, and Gaussia luciferase monitored NF- $\kappa$ B expression (a cell death marker). However, substrates were administered on subsequent days, limiting time resolution to 72 hours, and Vargula luciferin is not widely available to the community.

Color resolution has also been pursued as a means of multicomponent BLI. Click beetle luciferases developed by Promega have been popular for such applications.<sup>10-12</sup> Applications of color-resolved BLI have been shown *in vitro*,<sup>11</sup> and *in vivo*.<sup>10,12</sup> However, *in vivo* studies were conducted at shallow tissue depth, or with large amounts of cells to mitigate red-shifting of signal due to tissue absorption. Discriminating colors in thick tissue and live organisms is challenging, even with the best spectrally resolved probes.

We set out to develop robust, multicomponent bioluminescence imaging using orthogonal enzyme-substrate pairs. First, we developed a variety new synthetic techniques to enable diversification of firefly luciferin.<sup>13–17</sup> Next, we mutated firefly luciferase to selectively process these luciferin analogs.<sup>14,18</sup> These orthogonal luciferin-luciferase pairs retained selectivity in mouse models.<sup>18</sup> In this work, we demonstrate the ability of these probes to distinguish mixed populations, and apply this methodology to a cancer metastasis model.



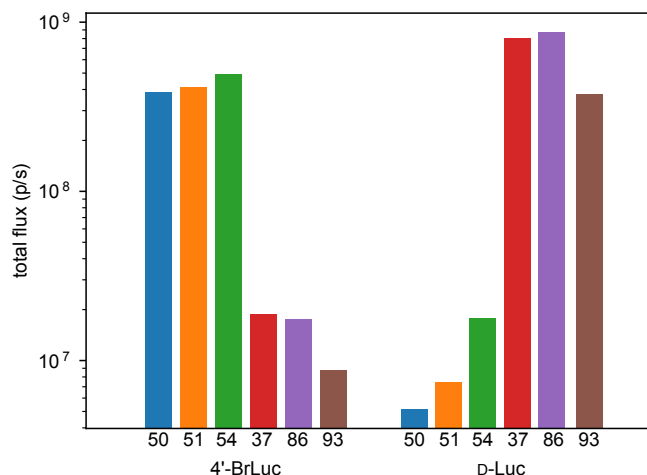
**Figure 5.1:** Bioluminescent substrate unmixing for the study of disease states. Sequential substrate addition enables unmixing of bioluminescent signal. With the ability to track multiple cell types simultaneously, the dynamics of disease states can be tracked over time.

Because we are using substrate resolution (not color resolution) for multicomponent imaging, we could not rely on optical filters or spectral unmixing algorithms to distinguish signals. Thus, we developed a sequential imaging approach coupled with a *substrate* unmixing algorithm to distinguish orthogonal pairs (Figure 5.1). We envisioned sequential substrate injection, with images acquired after each was administered. Substrate light output would “build” as images were acquired, and a processing step would unmix the signals to give the final relative abundance of each probe. Imaging in this way over time would enable study of disease states throughout the body of a mouse.

## 5.2 Results and discussion

### 5.2.1 Linear unmixing algorithms distinguish orthogonal bioluminescent probes

To realize our goal, we needed to select an ideal pair of luciferin analogs and luciferase mutants that would provide reliable substrate resolution. Bioluminescent probes for *in vivo* multicomponent imaging must be bright, bioavailable molecules. They must also be synthesized in large quantities. With this in mind, we examined top pairings from our parallel screens<sup>18</sup> and chose D-luciferin (D-luc) and **4'-BrLuc** as our lead pair. D-Luciferin is a molecule that is familiar to the community, commercially available, and has proved robust in a vast range of imaging studies.<sup>5</sup> **4'-BrLuc** is also a well-vetted molecule that has been the subject of several studies in our lab. Its synthesis is scalable and straightforward,<sup>17</sup> and initial *in vivo* studies have shown that it retains selectivity and brightness *in vivo*.<sup>18</sup>



**Figure 5.2:** Initial mutants of interest for multicomponent bioluminescence imaging. Target compound-mutant pairs were identified via parallel screening.<sup>18</sup> Each mutant selected demonstrated greater than 10-fold resolution in bacterial lysate. Compounds were administered at 100  $\mu$ M.

Next, we set out to identify mutant luciferases that would be amenable to multicomponent imaging *in vivo*. We aimed for two mutants that demonstrated greater than ten-fold se-

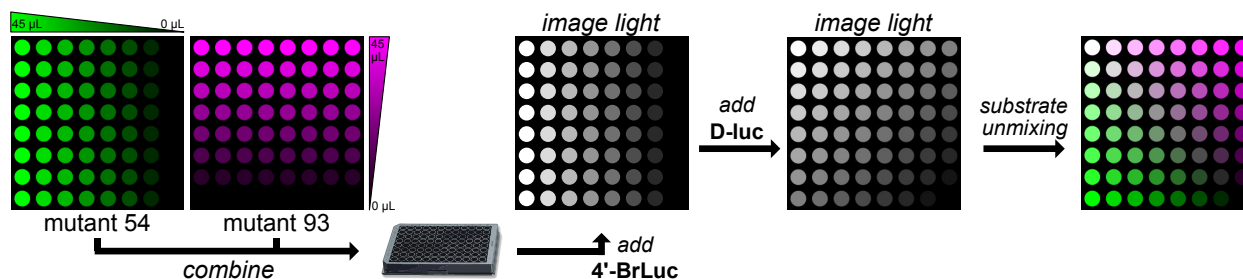


Enzyme	218	240	241	243	247	249	250	314	316	347
WT	<i>R</i>	<i>V</i>	<i>V</i>	<i>F</i>	<i>F</i>	<i>M</i>	<i>F</i>	<i>S</i>	<i>G</i>	<i>S</i>
50	R	<b>I</b>	V	F	<b>Y</b>	M	F	S	G	<b>G</b>
51	R	V	V	<b>M</b>	F	M	F	S	G	<b>G</b>
53	R	<b>I</b>	V	<b>M</b>	<b>Y</b>	M	F	S	G	<b>G</b>
54	R	V	<b>A</b>	F	<b>L</b>	M	F	S	G	<b>A</b>
37	<b>K</b>	V	V	F	F	M	F	S	G	S
85	<b>K</b>	V	V	F	F	M	<b>M</b>	<b>T</b>	<b>T</b>	S
86	<b>K</b>	V	V	F	F	<b>L</b>	F	<b>C</b>	G	S
87	<b>K</b>	V	V	F	F	M	<b>Y</b>	<b>T</b>	<b>T</b>	S
93	<b>K</b>	V	V	F	F	<b>L</b>	F	<b>V</b>	G	S

**Table 5.1:** Target mutants employed for resolution of **4'-BrLuc** and D-luc. Mutants 50, 51, 53, and 54 preferred **4'-BrLuc** while 37, 85, 86, 87, and 93 preferred D-luc.

lectivity with their matched luciferin, and retained robust light output. From our parallel screening data, we identified six mutants; three paired with **4'-BrLuc** (50, 51, 53, and 54) and three with D-luc (37, 85, 86, 87, and 93) (Figure 5.2 and Table 5.1). Mutants 50, 51, 53, and 54 harbored similar mutations, with each changing Ser 347 to Gly or Ala. It is likely that the removal of the serine hydroxyl group opens space in the active site for the bromine. D-Luc was positively paired with mutants 37, 85, 86, 87, and 93; each with a mutation at Arg 218 to Lys. This mutation alone (in the case of mutant 37) is enough to drastically reduce light emission with **4'-BrLuc** and has been characterized in several other studies.<sup>19-21</sup>

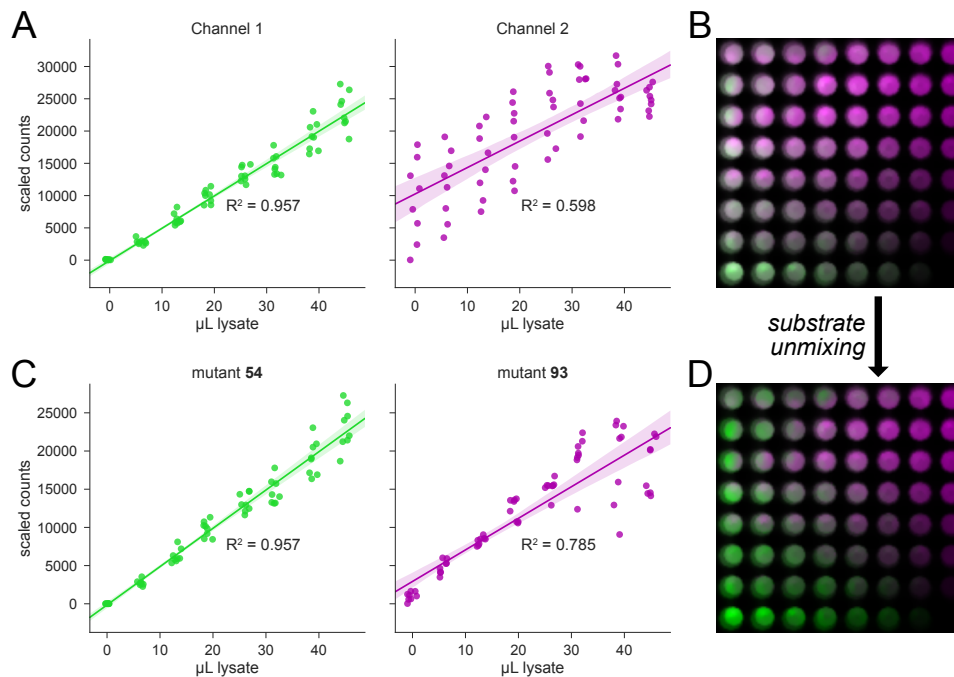
For a proof-of-concept of our technique, we designed an experiment to simulate multicomponent imaging in vivo at a variety of probe ratios (Figure 5.3). *E. coli* cells expressing the mutants of interest were lysed, and lysates were mixed in a gradient of 64 concentrations. **4'-BrLuc** was administered to each well, and an image of the plate was acquired. D-Luciferin was next added to the same 64 wells, and a second image was acquired. These images were then overlaid and linear unmixing algorithms were applied to determine the relative quantities of each mutant.



**Figure 5.3:** Design of experiment to test the success of substrate unmixing. Gradients of mutant luciferases in bacterial lysate were plated in an 8x8 square. **4'-BrLuc** was first added to each of the 64 wells, and light emission was measured. Second, D-luc was added to each of the same 64 wells, and light emission was again measured. The images were overlaid, and substrate unmixing was applied to the raw images. Shown in these images is a simulation of the expected results.

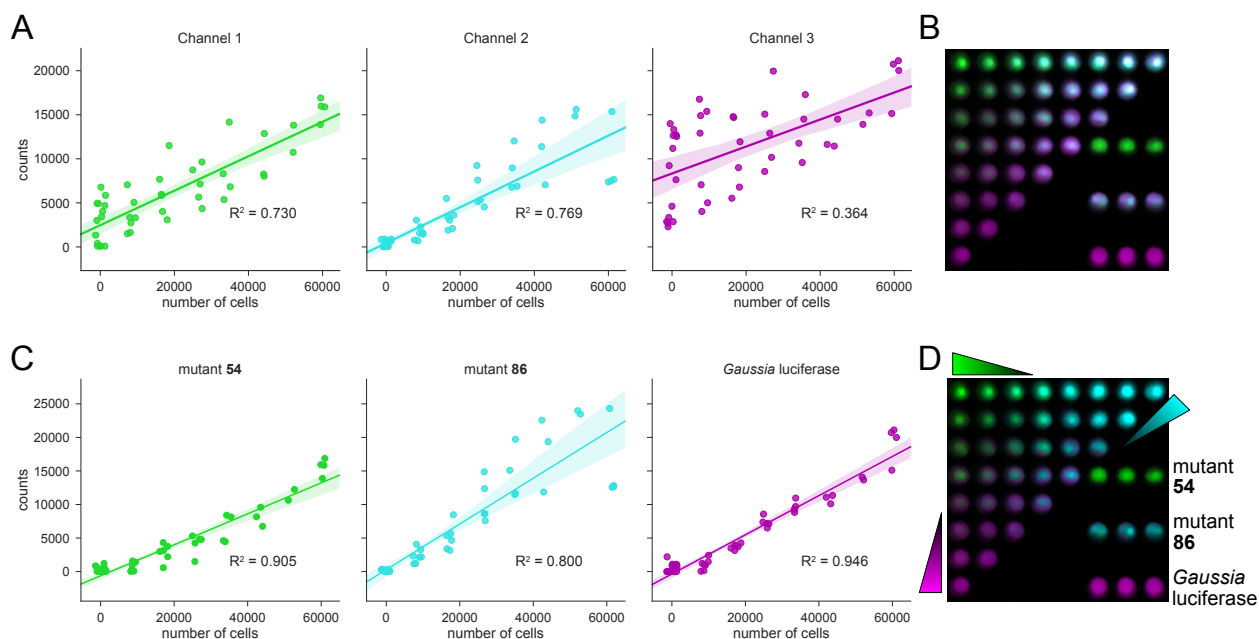
When these images were normalized and overlaid (Figure 5.4C), it was clear that the corresponding signal was not reflexive of the lysate ratio (Figure 5.4B). While channel 1 (the first image that was acquired following addition of **4'-BrLuc**), showed a linear relationship across the entire range of mutant concentrations and ratios, the second channel showed a high error ( $R^2 = 0.598$ ). This result makes sense because residual light from the addition of **4'-BrLuc** was present in the sample, so the second image reflected a combination of both signals.

To produce an output that was more reflective of the abundance of each mutant in the mixed samples, we turned to linear unmixing, a technique typically used in the field of fluorescence imaging.<sup>22</sup> Linear unmixing algorithms are commonly used in fluorescence microscopy when visualizing two fluorophores that have partially overlapping spectra. Overlap in emission wavelengths can result in convolution of signal that makes quantification of relative abundance difficult. We faced a similar situation with our substrate resolved probes. Though we were not concerned with the color of the bioluminescent emission, we thought that these same algorithms could be used to unmix *substrate* overlap. Indeed, when unmixing is applied to the gradient of mixed lysate, channel 2 is distinguished much more accurately (Figure 5.4D and E). This is apparent visually in Figure 5.4E, as well as with the  $R^2$  value of channel 2



**Figure 5.4:** Orthogonal bioluminescent probes can be distinguished in bacterial lysate. Based solely on the intensity of the light output, linear unmixing algorithms can distinguish orthogonal probes. **(A)** Quantification of signal before substrate unmixing fit via linear regression. The shaded area represents the 95% confidence interval of the fit. **(B)** Overlaid signal from raw images taken after each substrate addition. **(C)** Quantification of signal after substrate unmixing, fit via linear regression. The shaded area represents the 95% confidence interval of the fit. **(D)** Overlaid signal after substrate unmixing.

in Figure 5.4D. With this technique we were able to detect the presence of 6  $\mu\text{L}$  of mutant **54** lysate in 45  $\mu\text{L}$  of mutant **93** lysate (and the reverse).



**Figure 5.5:** Three orthogonal probes can be distinguished in mammalian cells. Gradients of cells expressing firefly luciferase mutants **54** and **86**, and *Gaussia* luciferase were plated in a triangle, with 60,000 cells per well. **4'-BrLuc** (500  $\mu\text{M}$ ), D-luciferin (500  $\mu\text{M}$ ), and coelenterazine (40  $\mu\text{M}$ ) were added in sequence. **(A)** Quantification of each channel from **(B)** fit via linear regression. The shaded area represents the 95% confidence interval of the fit. **(B)** Overlay of raw signal from mixed images. **(C)** Quantification of each channel from the unmixed image in **(D)** fit via linear regression. The shaded area represents the 95% confidence interval of the fit. **(D)** Overlay of the unmixed channels.

Next, we sought to test this method in mammalian cells, and determine whether we could unmix three luciferin signals (Figure 5.5). For these experiments we added *Gaussia* luciferase and coelenterazine (CTZ) as our third luciferase-luciferin pair (channel 3), with mouse DB7 cells stably expressing mutants **54**, **86**, and **Gluc**. Mixtures of cells totaling 60,000 were plated in a triangle of gradients and imaged between substrate additions (**4'-BrLuc**, D-Luc, and CTZ respectively). Overlay of the channels showed significant error in channels 1 and 3 ( $R^2$  values of 0.730 and 0.364 respectively, Figure 5.5A), with the effect of suppression of channel 2 (visually apparent through the lack of cyan in Figure 5.5B). Upon substrate unmixing, correlations improved considerably, with the proper gradient readily visualized in

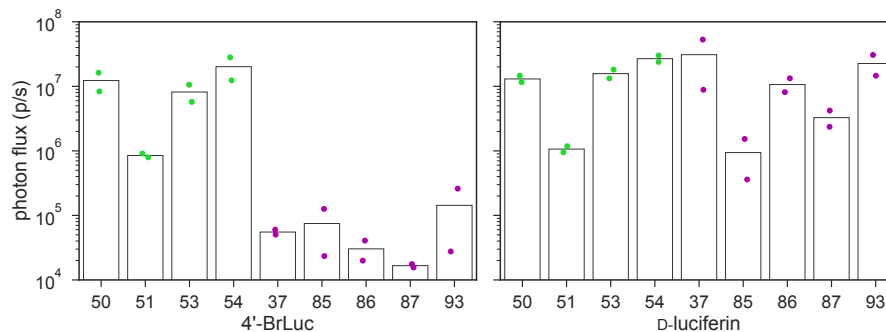
the overlaid image (Figure 5.5C and D). Note that the cyan hue assigned to mutant 86 is now visible.

This experiment highlights the utility of substrate unmixing, as even perfectly substrate resolved pairs (in the case of firefly and *Gaussia* luciferases) are not perfectly resolvable without applying unmixing algorithms.

### 5.2.2 Mixed cell populations can be distinguished in vivo.

Encouraged by our results in vitro, we sought to test our probes in vivo. To identify the best set of probes for our mouse studies, we first screened a panel of luciferase mutants in vivo. Optimal mutants would not only behave reproducibly, but also provide a building of bioluminescent signal, with signal response after addition of the first substrate remaining low enough to enable good overlay of the signal from the second substrate. With the help of Anastasia Ionkina, DB7 cells stably expressing a range of 9 different mutants (Table 5.1) were implanted into the backs of mice. The next day, substrates were administered sequentially, and images were acquired between injections. Signals at each tumor site were quantified, averaged, and passes through the CrossCompare script (chapter 3). The algorithm identified the most orthogonal pair to employ mutants 51 and 37. Inspired by the naming of fluorescent proteins after fruits, we coined mutant 37 as Cashew and mutant 51 as Pecan. We chose the theme of seeds/nuts because all tend to be the same color, with shape being the defining feature.

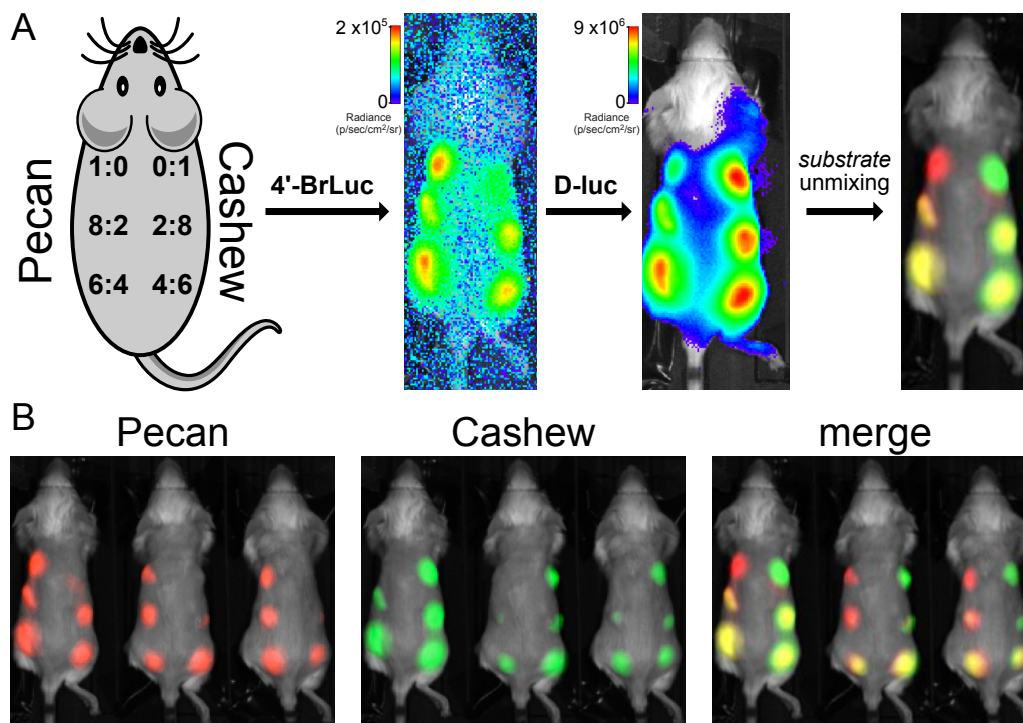
With optimum pairs identified, Anastasia and I designed an experiment to probe our ability to resolve mixed cell populations in vivo (Figure 5.7). In a similar fashion to the experiments conducted in vitro, we examined six different cell mixtures with varying relative concentrations of cells expressing either Pecan or Cashew. A total of one million cells were injected at each site, and images were acquired the next day. Upon injection of **4'-BrLuc** (first image



**Figure 5.6:** A panel of orthogonal mutants were screened in mice to choose an optimal pair for further experimentation. Our CrossCompare script identified mutants 37 and 51 as the best orthogonal pair. Points represent separate subcutaneous tumors implanted in different mice. Bars lie at the mean value of the replicates.

in Figure 5.7), the expected signal response was observed. Tumors on the left of the mouse (containing majority Pecan cells) gave robust signal, while less signal was observed in the right side of the mouse (tumors containing mostly Cashew). After 35 min of image acquisition, we administered D-luc and observed signal shift to the right side of the mouse (second image in Figure 5.7). Similarly, the tumors containing majority Cashew gave the largest photon counts. Finally, we subjected these images to the substrate unmixing algorithm and assigned pseudocolors to each luciferase (Pecan: red, Cashew: green in Figure 5.7). As expected, the signals from the top tumors reflect populations of 100% Pecan, or 100% Cashew. Additionally, the cell mixtures show resolution of the majority mutant; all signal on the left of the mouse is either yellow or red, and all signal on the right is either yellow or green. Though it does not appear that mixtures of 80/20 or 60/40 are resolvable (the middle and bottom tumors look similar), there is an obvious difference between left and right, even in the case of the bottom tumors. This indicates that we are able to resolve the makeup of mixed tumors versus those that contain a single cell type. Additionally, we can determine the cell type that is in the majority in the mixed cases.

Future experiments will incorporate fluorescent proteins of different colors into these constructs. These secondary markers will enable us to remove the tumors and analyze them via flow cytometry to determine relative populations *ex vivo*.



**Figure 5.7:** Orthogonal bioluminescent probes can be distinguished in mice. Cells were implanted at the ratios indicated. Compounds were administered sequentially via i.p. injection at 65 mM. Images were acquired 35 min post-injection.

### 5.3 Conclusions and future directions

We introduced a new technique for multicomponent bioluminescence imaging in vivo called substrate unmixing. This method takes advantage of both *substrate* and *intensity* resolution to resolve mixtures of luciferase expressing cells. We validated substrate unmixing in bacterial lysate, mammalian cell culture, and in mouse models. Future work will apply this technique to a cancer metastasis model. We plan to image the makeup of metastatic tumors that result from primary tumors of mixed cell populations. This technique will also have a wide range of applications in the field of in vivo imaging.

## 5.4 Methods and Materials

### 5.4.1 General bioluminescence imaging

All analyses were performed in black 96-well plates (Greiner 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. Exposure times 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 or ImageJ (NIH).

### 5.4.2 Bacterial lysate analysis of luciferase mutants

Bacterial cell stocks stored in glycerol containing the mutations of interest (in pET) were streaked on an agar plate containing kanamycin. After overnight growth, a colony was picked and expanded overnight. A portion of this culture (100  $\mu$ L) was added to 5 mL of LB (kan) and luciferase expression was induced as described previously in chapter 2.

### 5.4.3 CrossCompare algorithm

Scripts to search for orthogonal sets were written in Python3 and executed on either a MacBook Pro (for testing), or the High Performance Computing Cluster at UC Irvine (<https://hpc.oit.uci.edu/>). Searches for orthogonal pairs were run on 16 cores. Searches for triplet sets were run on 32 or 64 cores. The Python code and further documentation is freely available online: <https://www.chem.uci.edu/~jpresche/resources.html>. See chapter 2 for additional information.



#### 5.4.4 Substrate unmixing

Substrate unmixing was conducted using ImageJ (installed under the FIJI package). Luminescence images containing the raw CCD counts were loaded into FIJI and subjected to a 2 pixel median filter to remove any cosmic noise. Next, the signal at each pixel was min-max scaled to lie between 0 and 65535 (the maximum value that can be stored in a 16-bit image). In this way, the brightest pixel in each image should have a value of 65535, and the dimmest should have a value of 0. Next, images were stacked, and an additional image containing the maximum value each of the stacked images was computed (as a Z projection). This new image was added to the stack, and signals were unmixed using the ImageJ plugin developed by Gammon *et al.*<sup>11</sup> Pseudocolors were assigned in FIJI through the “Merge Channels” tool.

#### 5.4.5 Mammalian plasmid construction

To express the mutant luciferases in mammalian cells, the luciferase gene was amplified and inserted into pBMN-IRES-GFP. The following primers were used in the amplification: 5- ataacgcgtatggaagatgccaaaacattaaga-3 and 5-gagagggatgcatttattacacggcgatcttgcc-3 The PCR product was digested with *Nsi*I and *Mlu*I (New England BioLabs) and inserted into the pBMN-IRES-GFP vector with T4 ligase (New England BioLabs). Sequencing analysis was used to confirm the plasmid construct.

#### 5.4.6 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 g/mL). Cells were maintained in a 5% CO<sub>2</sub> 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 in chapter 2. Transduced cells were then treated with puromycin (10  $\mu\text{g}/\text{mL}$ ) and ultimately sorted via FACS at the Institute for Immunology Flow Cytometry Core (UCI).

#### **5.4.7 Mammalian cell imaging with luciferase mutants**

DB7 cells stably expressing Fluc or mutant luciferases were added to black 96-well plates (1 x 10<sup>5</sup> cells per well). A stock solution of luciferin (5 mM in PBS) was added to each well (500  $\mu\text{M}$  final concentration). Sequential imaging was performed as described in the General bioluminescence imaging section.

#### **5.4.8 *In vivo* imaging of orthogonal luciferase-luciferin pairs**

FVB/NJ mice (The Jackson Laboratory) received subcutaneous dorsal injections of 1x10<sup>6</sup> DB7 mutant luciferase expressing cells. After 24 h, animals were anesthetized (2% isoflurane) and placed on the warmed (37 °C) IVIS stage for imaging. To eliminate movement between injections, three legs were fixed to the imaging platform with tape. Each mouse received an i.p. injection of luciferin (67 mM, 100  $\mu\text{L}$  per mouse) under the untaped leg. Images were acquired with 5 min exposure times for 35 min. For sequential imaging, mice were immediately injected with the second substrate and imaged for an additional 35 min. Bioluminescent photons were quantified for the designated regions of interest.

# Bibliography

- [1] Rodriguez, E. A.; Campbell, R. E.; Lin, J. Y.; Lin, M. Z.; Miyawaki, A.; Palmer, A. E.; Shu, X.; Zhang, J.; Tsien, R. Y. *Trends Biochem. Sci.* **2017**, *42*, 111–129.
- [2] Germain, R. N.; Robey, E. A.; Cahalan, M. D. *Science* **2012**, *336*, 1676–1681.
- [3] Choi, J.-H.; Sim, S.-E.; Kim, J.-i.; Choi, D. I.; Oh, J.; Ye, S.; Lee, J.; Kim, T.; Ko, H.-G.; Lim, C.-S.; Kaang, B.-K. *Science* **2018**, *360*, 430–435.
- [4] Zhao, H.; Doyle, T. C.; Coquoz, O.; Kalish, F.; Rice, B. W.; Contag, C. H. *J. Biomed. Opt.* **2005**, *10*, 041210.
- [5] Paley, M. A.; Prescher, J. A. *Med. Chem. Comm.* **2014**, *5*, 255–267.
- [6] Vilalta, M.; Jorgensen, C.; Dégano, I. R.; Chernajovsky, Y.; Gould, D.; Noël, D.; Andrades, J. A.; Becerra, J.; Rubio, N.; Blanco, J. *Biomaterials* **2009**, *30*, 4986–4995.
- [7] Wendt, M. K.; Molter, J.; Flask, C. A.; Schiemann, W. P. *JoVE (Journal of Visualized Experiments)* **2011**, e3245–e3245.
- [8] Bhaumik, S.; Lewis, X. Z.; Gambhir, S. *JBO, JBOPFO* **2004**, *9*, 578–587.
- [9] 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. *Mol. Ther.-Nucleic Acids* **2013**, *2*, e99.

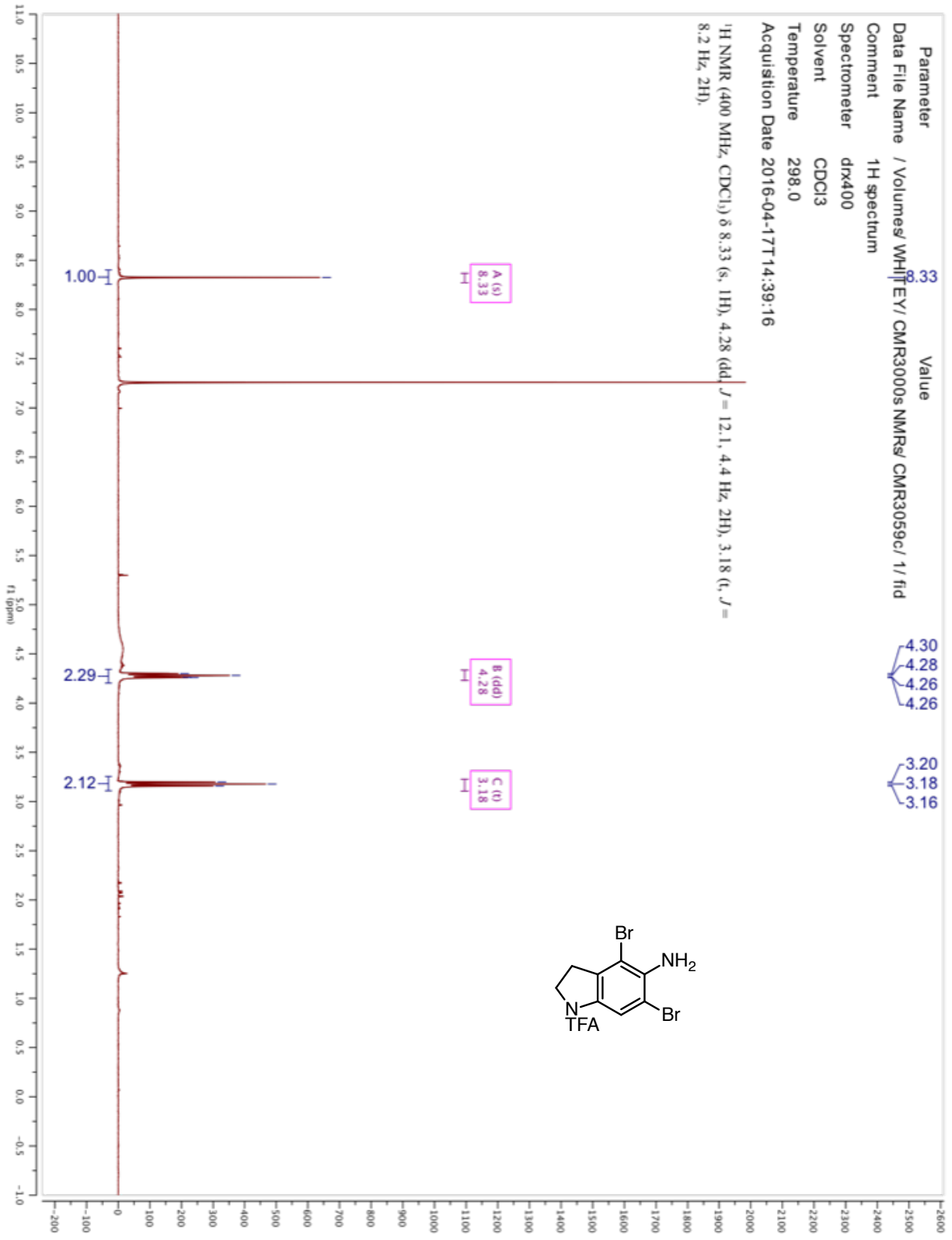
- [10] Mezzanotte, L.; Que, I.; Kaijzel, E.; Branchini, B.; Roda, A.; Lowik, C. **2011**, *6*, e19277.
- [11] Gammon, S. T.; Leevy, W. M.; Gross, S.; Gokel, G. W.; Piwnica-Worms, D. *Anal. Chem.* **2006**, *78*, 1520–1527.
- [12] Daniel, C.; Poiret, S.; Dennin, V.; Boutillier, D.; Lacorre, D. A.; Foligné, B.; Pot, B. *Appl. Environ. Microbiol.* **2015**, *81*, 5344–5349.
- [13] Steinhardt, R. C.; O’Neill, J. M.; Rathbun, C. M.; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. *Chem.—Eur. J.* **2016**, *22*, 3671–3675.
- [14] Jones, K. A.; Porterfield, W. B.; Rathbun, C. M.; McCutcheon, D. C.; Paley, M. A.; Prescher, J. A. *J. Am. Chem. Soc.* **2017**, *139*, 2351–2358.
- [15] McCutcheon, D. C.; Porterfield, W. B.; Prescher, J. A. *Org. Biomol. Chem.* **2015**, *13*, 2117–2121.
- [16] McCutcheon, D. C.; Paley, M. A.; Steinhardt, R. C.; Prescher, J. A. *J. Am. Chem. Soc.* **2012**, *134*, 7604–7607.
- [17] Steinhardt, R. C.; Rathbun, C. M.; Krull, B. T.; Yu, J. M.; Yang, Y.; Nguyen, B. D.; Kwon, J.; McCutcheon, D. C.; Jones, K. A.; Furche, F.; Prescher, J. A. *ChemBioChem* **2016**, *18*, 96–100.
- [18] Rathbun, C. M.; Porterfield, W. B.; Jones, K. A.; Sagoe, M. J.; Reyes, M. R.; Hua, C. T.; Prescher, J. A. *ACS Cent. Sci.* **2017**, *3*, 1254–1261.
- [19] Branchini, B. R.; Magyar, R. A.; Murtiashaw, M. H.; Portier, N. C. *Biochemistry* **2001**, *40*, 2410–2418.
- [20] Branchini, B. R.; Southworth, T. L.; Murtiashaw, M. H.; Boije, H.; Fleet, S. E. *Biochemistry* **2003**, *42*, 10429–10436.

- [21] Harwood, K. R.; Mofford, D. M.; Reddy, G. R.; Miller, S. C. *Chem. Biol.* **2011**, *18*, 1649–1657.
- [22] Dickinson, M. E.; Bearman, G.; Tille, S.; Lansford, R.; Fraser, S. E. *BioTechniques* **2001**, *31*, 1272, 1274–1276, 1278.

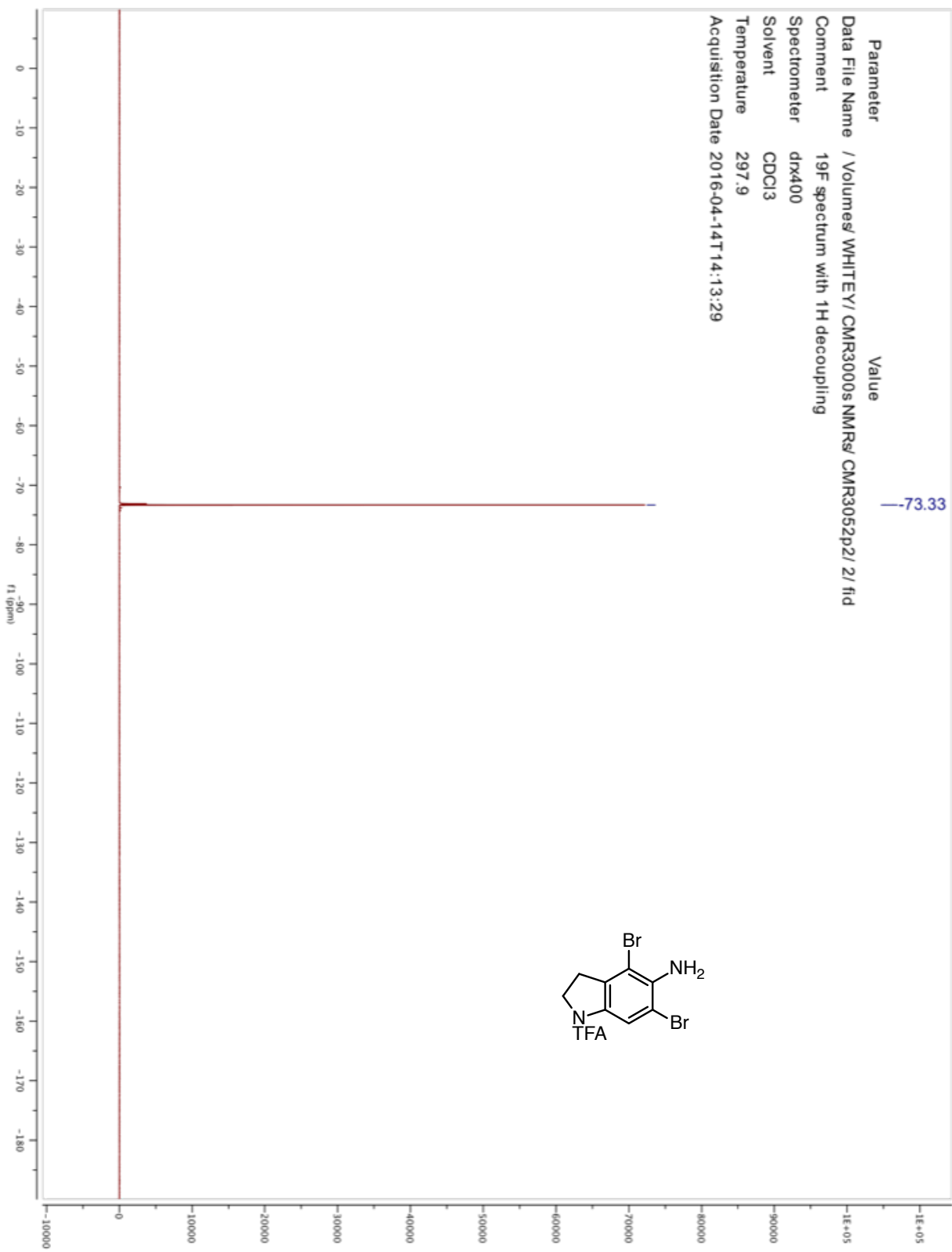
# Appendices

## A NMR spectra

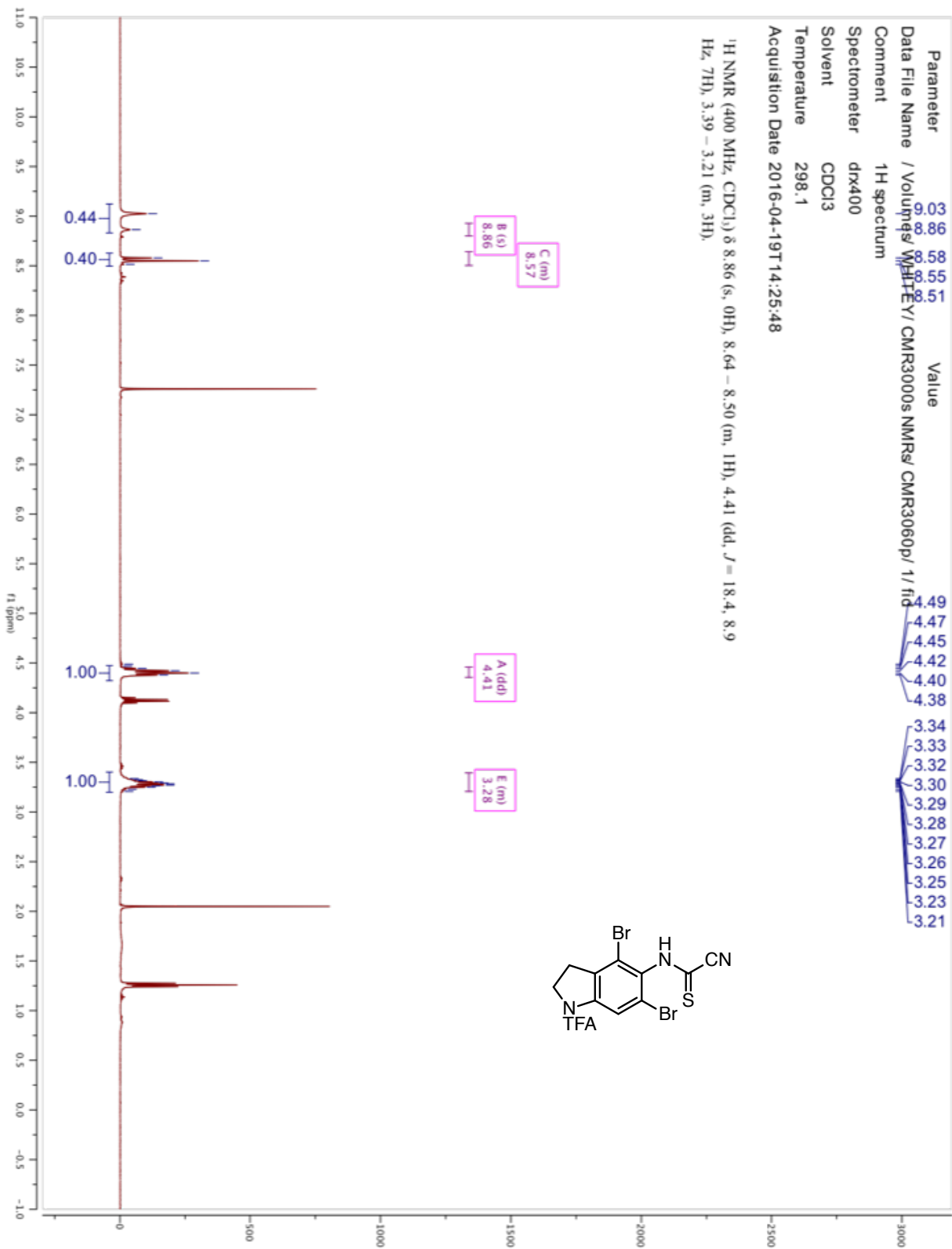
NMR spectra were acquired with Bruker Advanced spectrometers. All spectra were acquired at 298 K.  $^1\text{H}$ -NMR spectra were acquired at either 500 or 400 MHz, and  $^{13}\text{C}$ -NMR spectra were acquired at 125 MHz.



Parameter Value  
Data File Name /Volumes/WHITEY/CMR3000s NMRs/CMR3052p2/2/ftd  
Comment 19F spectrum with 1H decoupling  
Spectrometer drx400  
Solvent CDCl3  
Temperature 297.9  
Acquisition Date 2016-04-14T14:13:29







Parameter Value  
Data File Name / Volumes/ WHITEY/ CMR3000s NMR5/ CMR3049p/ 2/ fid  
Comment 19F spectrum with 1H decoupling  
Spectrometer dx400  
Solvent CDCl3  
Temperature 298.0  
Acquisition Date 2016-04-06T 16:57:48

—73.09

