

# UC Irvine

## UC Irvine Electronic Theses and Dissertations

### Title

Expanding the bioluminescent toolkit for in vivo imaging

### Permalink

<https://escholarship.org/uc/item/61n7p8kg>

### Author

Jones, Krysten

### Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,  
IRVINE

**Expanding the bioluminescent toolkit for *in vivo* imaging**

DISSERTATION

submitted in partial satisfaction of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Krysten Ann Jones

Dissertation Committee:  
Associate Professor Jennifer Prescher, Chair  
Professor Gregory Weiss  
Professor Celia Goulding

2017

Select figures from Chapter 1 © American Physiological Society

Select figures from Chapter 1 © Nature Publishing Group

Select figures from Chapter 1 © American Society for Microbiology

Select figures from Chapter 1 © Company of Biologists LTD

Select figures from Chapter 1 © Sasportas, et al.

Chapter 2 © American Chemical Society

Select figures and material from Chapter 5 © American Chemical Society

All other materials © 2017 Krysten A. Jones

## DEDICATION

To

my parents, Bob and Trisha Jones,  
in recognition of their love and support

“Happiness can be found even in the darkest of times,  
if one only remembers to turn on the light.” – Albus Dumbledore<sup>1</sup>

---

<sup>1</sup>Rowling, J.K. *Harry Potter and the Prisoner of Azkaban*. Scholastic, 2001.

# TABLE OF CONTENTS

	Page
LIST OF FIGURES	vi
LIST OF TABLES	x
ACKNOWLEDGEMENTS	xi
CURRICULUM VITAE	xiii
ABSTRACT OF THE DISSERTATION	xvi
CHAPTER 1: Optical tools for studying biological processes in living systems	
1.1 Introduction	1
1.2 Visualizing cells with combinations of fluorescent	2
1.3 Imaging cells at the macroscale with bioluminescent proteins	8
1.4 Future challenges and applications of bioluminescence imaging	13
1.5 Objectives of this study	15
References	16
CHAPTER 2: Visualizing cell proximity with genetically encoded bioluminescent reporters	
2.1 Introduction	22
2.2 Design of “split” forms of luciferase	24
2.3 “Split” forms of luciferase can reassemble in the extracellular space to enable sensitive imaging of collections of interacting cells	26
2.4 “Split” luciferase fragments enable facile imaging of cell-cell proximity in cultured models	32
2.5 Conclusions and future directions	38
2.6 Materials and methods	38
References	42

CHAPTER 3: Generalizable split reporters for improved visualization of cell-cell interactions	
3.1 Introduction	47
3.2 Fragment assembly is robust but underperforms native Gluc	48
3.3 Orientation and linker length influence Gluc complementation	51
3.4 Alternative dimerization domain motifs and variable repeats	53
3.5 Analyzing split reporters linked to fluorescent proteins or “designer” luciferases	56
3.6 Conclusions and future directions	61
3.7 Materials and methods	62
References	72
CHAPTER 4: Membrane bound split reporters for localizing zones of cell contact	
4.1 Introduction	74
4.2 Evaluating cell proximity reporters <i>in vivo</i>	75
4.3 Transmembrane domains can anchor split luciferase fragments to cell surfaces	77
4.4 Split reporter transmembrane modifications improve Gluc assembly	79
4.5 Ligand-receptor binding can facilitate Gluc assembly for imaging interacting cells	82
4.6 Conclusions and future directions	87
4.7 Materials and methods	88
References	94
CHAPTER 5: Orthogonal luciferase-luciferin pairs for bioluminescence imaging	
5.1 Introduction	96
5.2 Designing and constructing sterically modified luciferins	99
5.3 Analyzing bioluminescent light emission with modified luciferins	101

5.4 Measuring the light-emitting potential of luciferin analogs	103
5.5 Evolving substrate-specific luciferases	104
5.6 Analyzing the origins of orthogonality	110
5.7 Characterization of a lead orthogonal pair	112
5.8 Cellular imaging with orthogonal pairs	121
5.9 <i>In vivo</i> imaging with orthogonal pairs	127
5.10 Conclusions and future directions	130
5.11 Materials and methods	131
References	146
CHAPTER 6: Conclusions and future directions	151
References	156

## LIST OF FIGURES

		<b>Page</b>
Figure 1-1	Fluorescent probes for imaging	3
Figure 1-2	Fluorescent probes can be used to monitor cellular interactions during transendothelial migration	5
Figure 1-3	Confocal microscopy analysis of biofilm architecture	6
Figure 1-4	Motile intratumoral cells can enter tumor-associated vessels	7
Figure 1-5	Mechanism of bioluminescent light production	8
Figure 1-6	Luciferase-reporter influenza viruses can be used to monitor viral infection and evaluate antiviral therapies	12
Figure 1-7	Bioluminescence imaging of multiple aspects of cancer progression	13
Figure 2-1	Imaging cell-cell interactions with bioluminescent probes	23
Figure 2-2	Fos and Jun facilitate the stable assembly of Gluc	25
Figure 2-3	Cys-to-Ser mutations in Jun provide enhanced bioluminescent signal	26
Figure 2-4	Gluc fragments are expressed at comparable levels	27
Figure 2-5	Native Gluc is stable over time	28
Figure 2-6	Gluc complementation provides a rapid and sensitive readout on cell-cell interactions	28
Figure 2-7	Small numbers of cultured cells can be visualized using split Gluc fragments	29
Figure 2-8	Split Gluc technology is generalizable	30
Figure 2-9	Increased bioluminescent signal is observed with smaller volumes	31
Figure 2-10	Larger numbers of Jun-N <sub>Gluc</sub> -secreting cells decrease bioluminescent signal	31
Figure 2-11	Gluc complementation is distance dependent and can report on cell migration	33
Figure 2-12	Distance-dependent light emission is maintained over time	34



Figure 2-13	Gluc complementation is observed under flow	36
Figure 2-14	<i>In vitro</i> cell migration model	37
Figure 3-1	Imaging cell-cell interactions with split luciferase probes	48
Figure 3-2	Gluc complementation is robust and detectable in gel	49
Figure 3-3	Whole gel analysis of Gluc complementation	50
Figure 3-4	Complemented Gluc light output is decreased as compared to holo Gluc	50
Figure 3-5	Gluc fragments fused to the carboxy terminus of Jun or Fos are the optimal orientation for light emission	52
Figure 3-6	Linker length influences Gluc complementation	53
Figure 3-7	Engineered coiled-coils enable Gluc assembly	54
Figure 3-8	Variable repeats result in decreased light emission	56
Figure 3-9	Fluorescent protein fusions facilitate red-shifted light emission upon Gluc complementation	58
Figure 3-10	Venus fusions can be detected with 575-650 nm bandpass filter	59
Figure 3-11	Nluc complementation reports on cell-cell interactions	61
Figure 4-1	Strategy for cell surface localized light emission	74
Figure 4-2	Diffusion of the split reporter fragments is observed <i>in vivo</i>	76
Figure 4-3	Gluc complementation can report on cell-cell interactions <i>in vivo</i>	77
Figure 4-4	CD4 transmembrane domain can anchor luciferase to the cell surface	79
Figure 4-5	Gluc assembly is improved by split reporter transmembrane modifications	81
Figure 4-6	Ligand-receptor binding facilitates the assembly of Gluc	84
Figure 4-7	CD40 split reporters enable sensitive imaging of interacting cells	86
Figure 4-8	Split Gluc-CD40 system localizes signal to cell surface in a distance-dependent manner	87
Figure 5-1	Expanding the bioluminescence toolkit with unique enzyme-substrate pairs	98

Figure 5-2	Docking studies with sterically modified luciferins	100
Figure 5-3	Measuring luciferin light emission	101
Figure 5-4	Fluorescence emission spectra	102
Figure 5-5	Bioluminescence emission spectra	103
Figure 5-6	Generating mutant luciferase libraries and screening for orthogonal pairs	105
Figure 5-7	Representative images of luciferin analogs screened on plate	106
Figure 5-8	Some mutant enzymes provide similar or greater photon outputs than native luciferase	107
Figure 5-9	Sequencing analysis of “hits” from site-directed libraries	108
Figure 5-10	A computer script was used to screen all possible enzyme-substrate pairings <i>in silico</i>	109
Figure 5-11	Substrate-substrate pairing frequency in the top 5,000 <i>in silico</i> screen hits	111
Figure 5-12	Enzyme-enzyme pairing frequency in the top 5,000 <i>in silico</i> screen hits	112
Figure 5-13	Analyzing orthogonal enzyme-substrate pairs	113
Figure 5-14	Orthogonal substrate usage in mutant or Fluc (WT) lysate	114
Figure 5-15	Bioluminescent photon production from luciferin analogs and luciferase enzymes	115
Figure 5-16	Comparative analyses of all analogs with mutants A-C and wild type (WT) luciferase	117
Figure 5-17	Bioluminescent photon production from luciferin analogs with combinatorial enzymes	119
Figure 5-18	Comparative analyses of combinatorial enzymes and Fluc	120
Figure 5-19	Mutant luciferases are expressed at comparable levels in mammalian cells	122
Figure 5-20	Imaging cells with orthogonal luciferase-luciferin pairs	122
Figure 5-21	Luciferase-expressing cells exhibit orthogonality	123

Figure 5-22	Cellular imaging with mutant luciferases and D-luciferin	123
Figure 5-23	Cellular imaging with mutant luciferases and luciferin analogs	124
Figure 5-24	Cellular bioluminescent photon production is sustained	125
Figure 5-25	Cellular imaging with orthogonal pairs	126
Figure 5-26	Cellular imaging with orthogonal pairs patterned with biocompatible stencils	127
Figure 5-27	<i>In vivo</i> imaging with orthogonal pairs	129

## LIST OF TABLES

		<b>Page</b>
Table 1-1	Characteristics of commonly used luciferase-luciferin pairs	10
Table 5-1	Biochemical analyses of orthogonal enzyme-substrate pairs	118

## ACKNOWLEDGEMENTS

Without a doubt, I need to express my sincerest appreciation to my advisor and committee chair, Professor Jennifer Prescher. Her mentorship has been unparalleled. You provided me with endless enthusiasm and support, teaching me what it takes to be a good scientist. It has truly been an honor and a privilege! Thank you!

I would also like to thank my committee members, Professors Gregory Weiss and Celia Goulding. Greg, I will never forget that you went above and beyond to be at my advancement exam. Even having just been in an accident, you still had wonderful insights and advice for my project. Celia, I learned everything I know about protein expression from rotating in your lab and know it is a skill I will continue to use for years to come. Thank you for helpful advice while I was applying for postdoctoral positions.

I also thank the faculty, staff, and students in the Department of Molecular Biology and Biochemistry and the Department of Chemistry. I was at the interface of both departments and received endless support and advice from many of the faculty members. Thank you to Professor Andrej Luptak for guiding me through my first year at UCI. Thank you to Professors Melissa Lodoen and Naomi Morrissette for advice during joint lab meetings and for showing so much interest in my research progress. Thank you to the many students whom have provided experimental advice and helpful discussions, especially to members of the Weiss, Martin, Lodoen, Goulding, Hui, and Fruman labs.

Thank you to my previous mentors who have all helped me reach my goals: Dr. Jane Cavender, Dr. Jodi Yorty, Dr. Dave Bodine, Dr. Fabio Candotti, Dr. David Wyles and Dr. Robert Schooley.

Thank you to my fellow Prescher lab members. The last five years would have been rough without having such wonderful colleagues and friends to work with. Thank you to the first five members of the Prescher group, Dr. Miranda Paley, Dr. Rachel Steinhardt, Dr. Dave McCutcheon, Dr. Lidia Nazarova, and Dr. Dave Patterson. You all welcomed me into the group during my rotation and helped me realize it was the right lab for me. Dave and Lidia, in particular, were amazing to work with on different biorthogonal chemistry projects. Thank you also to Will Porterfield and Colin Rathbun. You both motivated me to work smarter and more critically in almost every step of the luciferase-luciferin project. Thank you to the other members of the luciferase-luciferin team, Brendan Zhang and Zi Yao, especially for remaining enthusiastic about screening all dark plates. Thank you to Dr. Zach Reinert, Dr. Joanna Laird, and Anya Kent for helpful discussions on all things cell-cell contact. Thank you to Dave Row and Sean Nguyen for being excellent labmates and great guys to talk science with. Thank you to Dr. Hui-Wen Shih for providing amazing perspective and advice, along with Cat-Opoly! Thank you to the many undergrads, Bryan Xie, Marian Sagoe, Monique Reyes, and Christine Hua, I had the opportunity to work with.

Thank you to Friend Group™ for helping keep me sane throughout graduate school and sharing so many amazing experiences along the way: Dr. Olivia Cromwell, Dr. Mary Beth Daub, Dr. Ethan Hill, Dr. Aaron Hollas, Dr. Domarin Khago, Dr. Chris Kotyk, Dr. Juliet K. Kotyk, Dr. Greg Lackner, Greg Suryan, and Nick Tallarida. It has definitely been a crazy five years but I'm so glad to have become such great friends with each and every one of you. See you all in September for our next crazy adventure!

Also, thank you to my many friends from home and undergrad: Laura and Bradd Gower, Pam Zaritsky, Andrea Levine, Rachel Share, Emily Ward-Rill, Heather Faulhefer, and Joshua Andrzejewski. I have known all of you for so many years and appreciate all of your encouragement and

Thank you to my parents for being unbelievably supportive every step of the way. You've both pushed me to always do my best and have helped me keep perspective these last five years. I owe every success I've had along the way to you. Thank you to my sister Nicole, I'm glad you could visit California and provide me some much-needed breaks from the lab. Thank you to my Grandma, for always checking in and sending so many great care packages. Thank you to my aunt and uncles, Maureen, Brian, and Kenny, for your encouragement and for always making the time to visit when I'm home.

Finally, thank you, Dave McCutcheon. You have been supportive, loving, and goofy every step of the way. I'm so excited to see what our future holds and to embrace some cold Chicago winters!

Financial support was provided by the Department of Molecular Biology and Biochemistry, University of California, Irvine and an institutional Chemical and Structural Biology Training Grant predoctoral fellowship (T32-GM10856).

## CURRICULUM VITAE

**Krysten A. Jones**

University of California, Irvine  
Irvine, CA 92697

Email: joneska@uci.edu

### ***EDUCATION AND TRAINING***

---

**University of California, Irvine, Irvine, CA** 2011-2017

*Ph.D., Biological Sciences*

**Elizabethtown College, Elizabethtown, PA** 2003-2006

*B.S., Biotechnology*

### ***RESEARCH EXPERIENCE***

---

**University of California, Irvine, Irvine, CA** 2011–present

Graduate research

Advisor: Professor Jennifer A. Prescher

**University of California, San Diego, San Diego CA** 2009–2011

Research associate

Advisor: Dr. Robert T. Schooley

**National Institutes of Health, Bethesda, MD** 2007–2008

*Postbaccalaureate Intramural Research Training Award (IRTA)*

Advisors: Dr. David M. Bodine and Dr. Fabio Candotti

**Elizabethtown College, Elizabethtown, PA** 2005–2006

Undergraduate research

Advisor: Professor Jodi L. Yorty

**Hanson Technologies, Carlisle, PA** June–Aug 2006

Summer intern

Advisor: Dr. Maureen A. Dyer

**Fox Chase Cancer Center, Philadelphia, PA** June–Aug 2005

Student assistant

Advisor: Dr. Eileen K. Jaffe

## TEACHING EXPERIENCE

---

**Research Mentor**, Prescher lab, University of California, Irvine

Involved in training of 4 undergraduate and 5 graduate student researchers, Jan 2013–present

**Graduate Student Instructor**, Department of Chemistry, University of California, Irvine

Chemical Biology (for undergraduate students), Spring 2011

**Graduate Student Instructor**, Department of Molecular Biology and Biochemistry, University of California, Irvine

Microbiology (for undergraduate students), Spring 2014

Molecular Biology (for undergraduate students), Spring 2014

Biochemistry (for undergraduate students), Winter 2012

Microbiology (for undergraduate students), Winter 2012

**Tutor**, Department of Biology, Elizabethtown College

Biology and General Chemistry (for first-year students), 2004-2006

## HONORS AND AWARDS

---

2016	Best Oral Presentation Prize, UC Irvine Immunology Fair
2016	Howard A. Schneiderman Fellowship
2014-present	NIH Chemical and Structural Biology Trainee
2006	Conover Memorial Scholarship (Elizabethtown College)
2005	Inducted into Tri-Beta (National Biology Honor Society)

## PUBLICATIONS

---

8. Jones, K.A.\*; Porterfield, W.B.\*; Rathbun, C.M.\*; McCutcheon, D.C.; Paley, M.A.; Prescher, J.A. Orthogonal luciferase-luciferin pairs for bioluminescence imaging, *J. Am. Chem. Soc.* **2017**, doi: 10.1021/jacs.6b11737. [\* denotes equal contribution]
7. Steinhardt, R.C.; Rathbun, C.M.; Krull, B.T.; Yu, J.; Yang, Y.; Nguyen, B.D.; Kwon, J.; McCutcheon, D.C.; Jones, K.A.; Furche, F.; Prescher, J.A. Brominated luciferins are versatile bioluminescent probes. *Chembiochem.* **2017**, *18*, 96-100.
6. Nazarova, L.A; Ochoa, R.A.; Jones, K.A.; Morrisette, N.M.; Prescher, J.A. Extracellular *Toxoplasma gondii* tachyzoites metabolize and incorporate unnatural sugars into cellular proteins, *Microbes Infect.* **2016**, *18*, 199-210.
5. Porterfield, W.B.; Jones, K.A.; McCutcheon, D.C.; Prescher, J.A. A “caged” luciferin for imaging cell-cell contacts, *J. Am. Chem. Soc.* **2015**, *137*, 8656-8659.
4. Jones, K.A.; Li, D.J.; Hui, E.E.; Sellmyer, M.A.; Prescher, J.A. Visualizing cell proximity with genetically encoded bioluminescent reporters, *ACS Chem. Biol.* **2015**, *10*, 933-938.
3. Patterson, D. M.; Jones, K. A.; Prescher, J. A. Improved glycan imaging with modified cyclopropene chemical reporters, *Mol. BioSys.* **2014**, *10*, 1693-1697.



2. Gutierrez, J.A.; Jones, K.A.; Flores, R.; Singhania, A.; Woelk, C.H.; Schooley, R.T.; Wyles, D.L. Vitamin D metabolites inhibit Hepatitis C Virus and modulate cellular gene expression, *J. Virol. Antivir. Res.* **2014**, *3*, doi:10.4172/2324-8955.1000129
1. Adriani, M.; Jones, K.A.; Uchiyama, T.; Kirby, M.R.; Silvin, C.; Anderson, S.M.; Canotti, F. Defective inhibition of B-cell proliferation by Wiskott-Aldrich syndrome protein-deficient regulatory T cells. *Blood.* **2011**, *117*, 6608-6611.

## PRESENTATIONS

---

15. Graduate Student and Post-Doc Symposium, University of California, Irvine **Jan 2017**.
14. 14<sup>th</sup> Annual UC Irvine Immunology Fair, Irvine, CA **Dec 2016**.
13. Annual Conference of the International Chemical Biology Society, Madison, WI **Oct 2016**.
12. University of California, Irvine, Chemical & Structural Biology Training Program, Irvine, CA **April 2016**.
11. University of California, Irvine, Department of Molecular Biology & Biochemistry, Irvine, CA **Jan 2015**.
10. University of California, Irvine, Cancer Biology & Therapeutics Training Program, Irvine, CA **Nov 2014**.
9. Annual Symposium of The Protein Society, San Diego, CA **July 2014**.
8. Molecular Biology & Biochemistry Department Retreat, Lake Arrowhead, CA **March 2014**.
7. University of California, Irvine, Department of Molecular Biology & Biochemistry, Irvine, CA **March 2014**.
6. Southern California Eukaryotic Pathogen Symposium, Riverside, CA **2013**.
5. Molecular Biology & Biochemistry Department Retreat, Lake Arrowhead, CA **March 2013**.
4. University of California, Irvine, Department of Molecular Biology & Biochemistry, Irvine, CA **April 2013**.
3. Conference on Retroviruses and Opportunistic Infections, Boston, MA **Feb 2011**.
2. NHGRI Annual Scientific Retreat, Gettysburg, PA **June 2008**.
1. ACS (Philadelphia Section) Graduate/Undergraduate Poster Session, Philadelphia, PA **Oct 2006**.

## ABSTRACT OF THE DISSERTATION

Expanding the bioluminescent toolkit for *in vivo* imaging

By

Krysten Ann Jones

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2017

Associate Professor Jennifer A. Prescher, Chair

Imaging tools have revolutionized our understanding of living systems by allowing researchers to visualize biological features in real time. Among the most popular imaging agents of this sort are the bioluminescent proteins (luciferases). These probes catalyze the oxidation of small molecule substrates (luciferins), releasing non-toxic, non-perturbing visible light in the process. Bioluminescence imaging (BLI) with luciferase-luciferin pairs is well suited for use in tissues and small animals. Since no excitation light is required, there is virtually no background signal. Indeed, BLI has been used to track numerous biological phenomena *in vivo*, including immune cell homing, tumor cell proliferation, and pathogen egress. Despite its remarkable versatility, bioluminescence lacks the spatial resolution necessary to monitor cell-cell interactions. Furthermore, BLI has been historically limited to monitoring one cell type or biological feature at a time. To address these issues, this dissertation involves engineering new luciferase tools and comprised two main goals: (1) engineering contact-dependent bioluminescent probes to report on cellular communication and (2) generating novel luciferase-luciferin pairs *via* directed enzyme evolution. Collectively, these tools will enable new studies of immune cell function, host-pathogen dynamics, and numerous other multi-cellular networks *in vivo*.

# CHAPTER 1: Optical tools for studying biological processes in living systems

## 1.1 Introduction

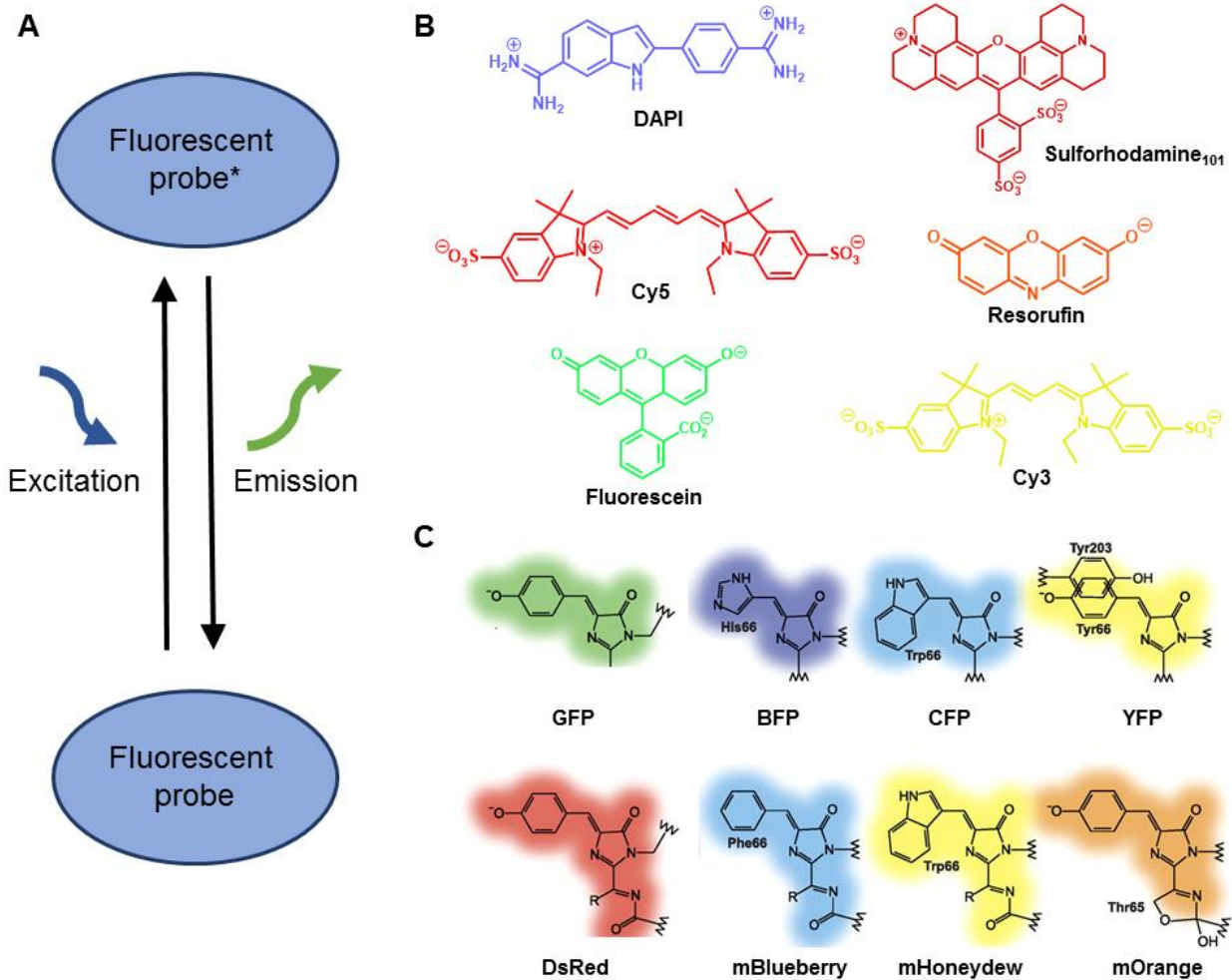
A complete understanding of organismal biology requires methods to monitor dynamic cellular movements and interactions *in vivo*. Immune function, pathogen clearance, and cell migration involve multicellular networks, which are influenced by spatial and environmental cues (1-3). These dynamic conditions are challenging to reproduce in cultured cells or excised tissue models but are necessary to provide a complete picture of biological processes as they occur in living systems. Thus, researchers must often turn to rodents or other living organisms to capture the complexity of the underlying biology. Such models present several challenges for studies, owing to tissue heterogeneity, the depth of biological tissues, and the need for models that closely recapitulate human disease (4, 5).

In recent years, a set of optical imaging technologies has emerged that enables the real-time observation of cellular movements in whole organisms (5-8). These tools produce non-toxic, non-perturbing visible light that can provide readouts on cell movements and other functions (8, 9). Among the most popular imaging tools are the fluorescent and bioluminescent probes (8, 10-12). This chapter introduces these probes and how they have been utilized to understand basic biology and disease mechanisms. I first highlight the utility of fluorescent probes for microscale observations of interacting cells and their limitations for *in vivo* imaging. I then discuss bioluminescence imaging, highlighting commonly used luciferase-luciferin pairs and their utility for visualizing cells at the macroscopic scale. Lastly, I describe existing challenges of

bioluminescence imaging that are not yet addressed with the currently available bioluminescent probes.

## **1.2 Visualizing cells with combinations of fluorescent probes**

One class of imaging tools comprises fluorescent probes. Fluorescence is a process in which an external light source excites a fluorophore to a higher energy state and subsequent emission at a longer wavelength is detected (Figure 1-1A) (5). Fluorescent probes can be divided into two groups: small molecules fluorophores and fluorescent proteins (13-16). Fluorophores commonly contain extended  $\pi$ -systems that dictate their emission wavelength (Figure 1-1B). As such, a wide range of fluorophores has been developed with different colors and properties for cell labeling (17-19). Some fluorophores can selectively label target organelles, such as the mitochondria, Golgi apparatus, or nucleus. DAPI and Hoechst 33258 are two such examples for nuclear staining, both of which fluoresce when bound to DNA (20, 21). Additionally, many fluorophores (DiR, DiI, Atto dyes) can be appended to cell surfaces (22). Cell-targeting probes with antibody-fluorophore conjugates have also been utilized for *in vivo* brain and tumor imaging (23-25). While fluorophores are broadly applicable to labeling and tracking any cell type, they are only useful for short-term imaging (19).

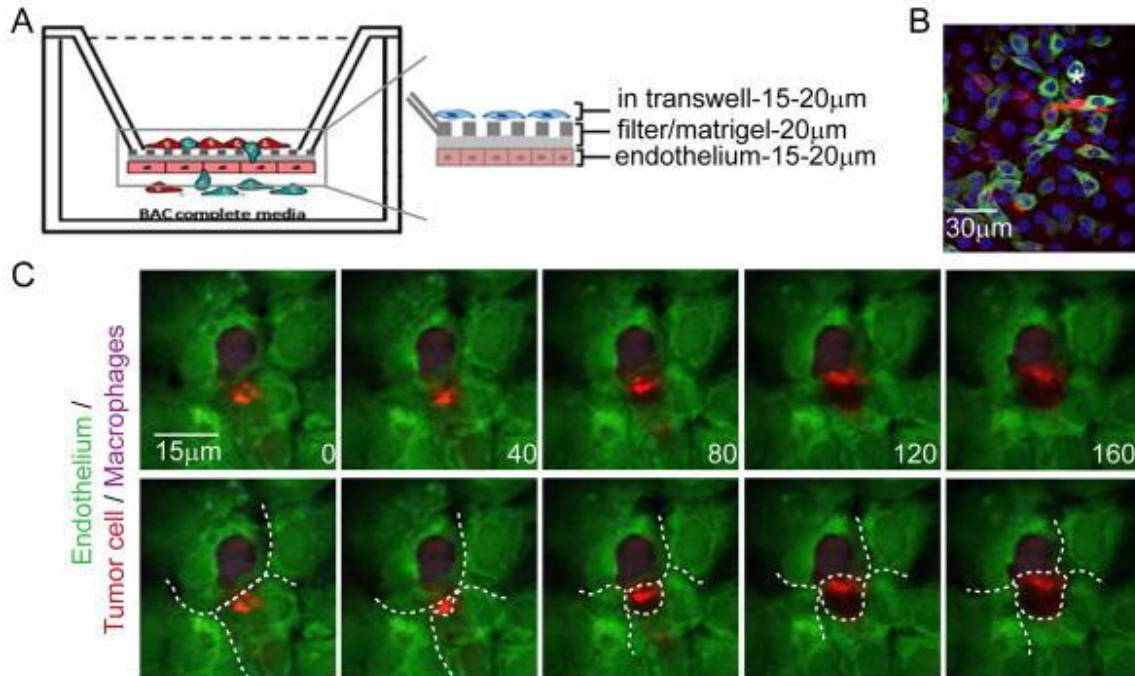


**Figure 1-1.** Fluorescent probes for imaging. (A) Fluorescence requires light to excite electrons in the probe. Relaxation to the ground state results in the emission of a photon of light. (B) Examples of common fluorophores for biomedical imaging. (C) Examples of chromophore mutations to generate GFP-like fluorescent proteins. For (C), image reproduced with permission from (Chudakov et al. *Physiol Rev.* **2010**, *90*, 1103-63).

For long-term, serial tracking of cells, genetically encoded probes, such as fluorescent proteins, are more desirable (26). Green fluorescent protein (GFP) was the first described fully genetically encodable fluorescent label (27). The ability of GFP to label and track cells in living systems spurred engineering efforts to develop GFP mutants with enhanced spectral characteristics and properties. By targeting the three native amino acids that form the internal

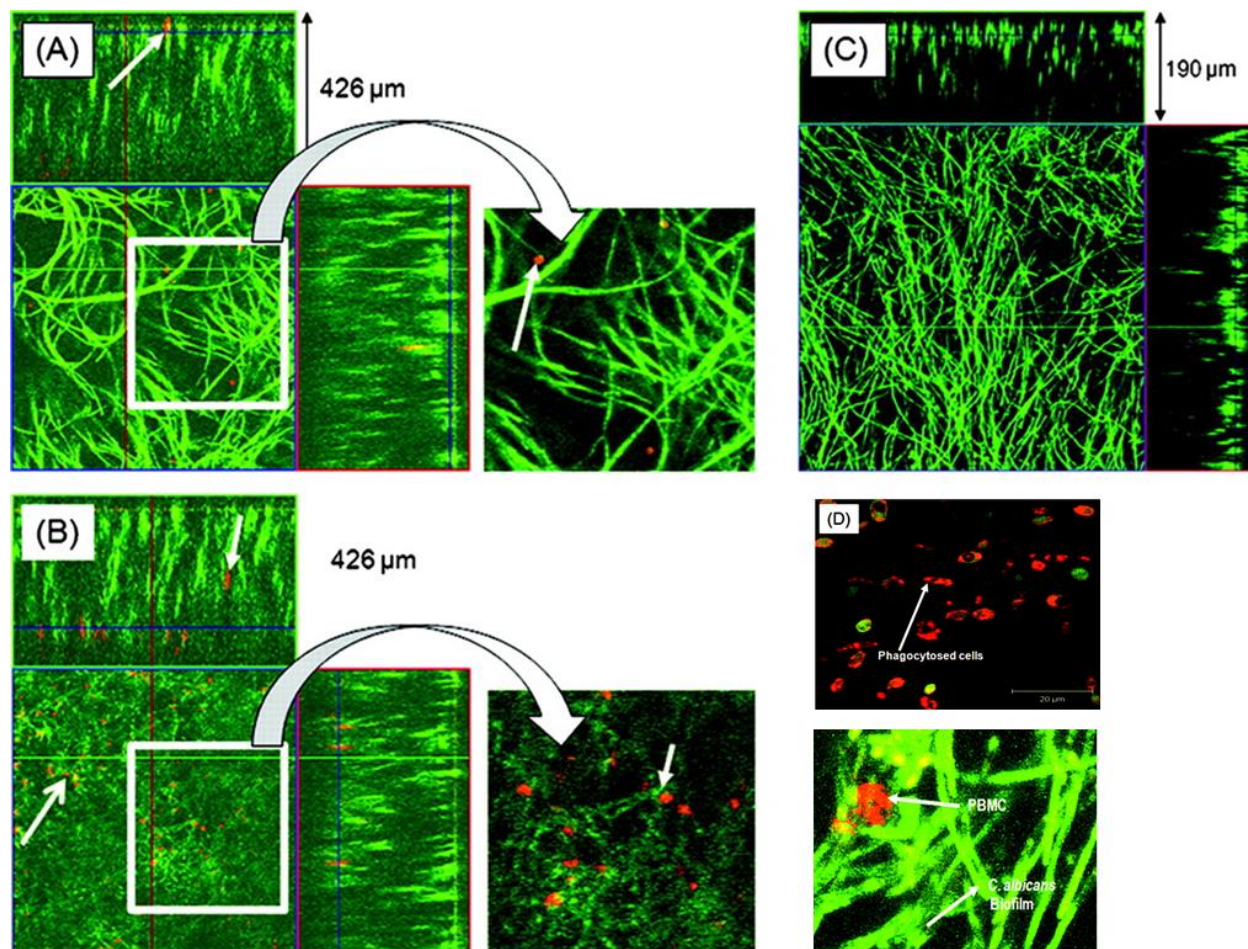
chromophore of GFP during protein folding, a palette of fluorescent proteins was developed (Figure 1-1C) (28). While GFP-like reporters and the related coral fluorescent proteins have been limited to emission maxima <610 nm, new discovery and directed evolution efforts have generated far-red fluorescent proteins useful for *in vivo* imaging (29-31).

Combinations of these fluorescent probes have been widely employed at the microscopic scale to visualize intricate cellular events and interactions. For example, the Condeelis group has examined the interactions between tumor cells, macrophages, and endothelial cells with *in vitro* models of intravasation (32). They first formed endothelial monolayers with primary human microvascular endothelial cells (DAPI-labeled). After confirming intact monolayers, they added the metastatic breast cancer line, MDA-MB-231 (GFP-labeled), and the BAC1.2F5 murine macrophage cell line (CMPTX-labeled) and analyzed cellular interactions (Figure 1-2A). This study found that macrophages and tumor cells physically interact at the endothelium (Figure 1-2B). These interactions influenced migration, as tumor cells underwent preferentially transendothelial migration at sites where macrophages were present (Figure 1-2C).



**Figure 1-2.** Fluorescent probes can be used to monitor cellular interactions during transendothelial migration. (A) Strategy for visualizing tumor cell-macrophage interactions during endothelial cell migration. (B) Physical interactions between tumor cells (MDA-MB-231, green) and macrophages (CMPTX-labeled, red) at the endothelium (blue). (C) Still images from live imaging of tumor cell transmigration. Endothelial cell edges are shown with white dashed lines. Image reproduced with permission from (Roh-Johnson et al. *Oncogene*. **2014**, 33, 4203-12).

Fluorescent probes have also been integral in revealing new mechanistic insights in host-pathogen interactions (33, 34). For example, the Ghannoum group evaluated the role of peripheral blood mononuclear cells (PBMCs) in the ability of *C. albicans* to form biofilms (35). When they cultured GFP-expressing *C. albicans* in the presence of PBMCs (stained with Mitotracker Deep Red 633 dye), the fungal cells formed thicker biofilms (Figure 1-3A-C). The PBMCs were primarily present in the basal and middle layers of the biofilm and phagocytosis of *C. albicans* was not observed. In similar models with planktonic (free-floating) *C. albicans*, the fungal cells were phagocytosed by the PBMCs (Figure 1-3D)

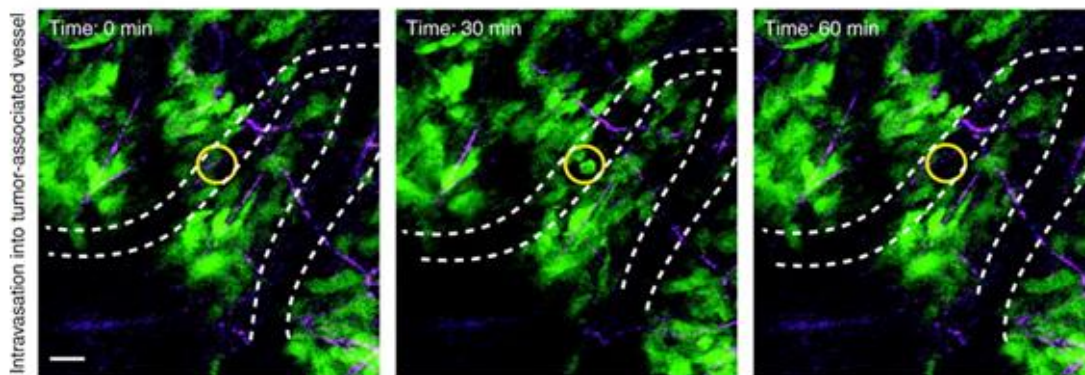


**Figure 1-3.** Confocal microscopy analysis of biofilm architecture. (A) The upper layer and (B) basal/middle layers of GFP-tagged *C. albicans* biofilms and PBMCs (Mitotracker Deep Red 633 dye, red) were imaged. Biofilms formed in the presence of PBMCs are thicker (426 μM) than (C) those formed without PBMCs (190 μM). (D) Planktonic *C. albicans* are phagocytosed by PBMCs (top) but phagocytosis is not observed in biofilms. Image reproduced with permission from (Chandra et al. *Infect. Immun.* **2007**, *75*, 2612-2620).

Fluorescent probes can further be used for *in vivo* imaging applications. *In vivo* imaging presents additional challenges as endogenous chromophores in tissues and blood can absorb and scatter light (5, 36). This results in fewer photons being detected and broadening of the area of detectable signal, respectively. This can be circumvented by installing “windows” into tissue to allow close placement of the microscope objective (37, 38). Thus, intravital microscopy and other techniques have revolutionized our understanding of cellular communication in many cases



(7, 39-41). In cancer biology, Beerling and co-workers used intravital microscopy to visualize motile tumor cells within the primary tumor mass (42). To determine if these cells could enter vasculature and play a role in cancer progression, they monitored a C26 colorectal tumor cell mass over time. A tumor cell (Dendra2-labeled, green) within the tumor mass is shown to enter tumor-associated vasculature and is subsequently transported by the blood (Figure 1-4). The study found that cells within the tumor mass, not just tumor cells at the invasive front of the tumor, could contribute to metastatic disease and poor prognosis.



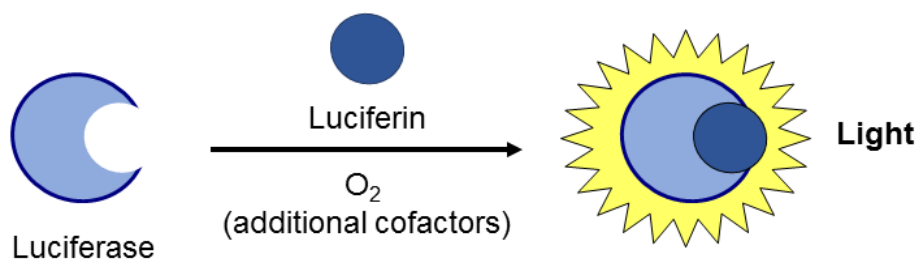
**Figure 1-4.** Motile intratumoral cells can enter tumor-associated vessels. Intravital microscopy imaging with GFP-labeled C26 colorectal tumor cells monitors a tumor cell over time to visualize intravasation into a tumor-associated vessel. At 30 min (middle), the tumor cell is observed in the vessel and is transported in the blood (and out of sight) at 60 min (right). Vasculature is outlined with dashed lines and collagen fibers are shown in purple. Scale bars: 10  $\mu\text{m}$ . Image reproduced with permission from (Beerling et al. *J. Cell Sci.* **2011**, *124*, 299-310).

Despite the remarkable level of spatial detail obtained with fluorescence imaging as showcased in the above examples, these techniques require invasive surgeries and are often not readily accessible to all users. Additionally, the level of resolution often comes at the expense of surveying only small, previously selected areas. It is therefore difficult to utilize conventional fluorescence techniques without knowing “when and where” to look. Furthermore, the excitation light required for fluorescence microscopy typically results in poor signal-to-noise ratios, limiting one’s ability to analyze cell interactions and migration over extended distances and

times (5, 43). While this is beginning to change due to new optical imaging approaches, such as light sheet microscopy and adaptive optics, these techniques are not yet readily accessible to the non-specialist (44-47).

### 1.3 Imaging cells at the macroscale with bioluminescent proteins

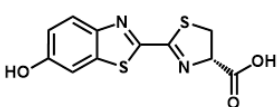
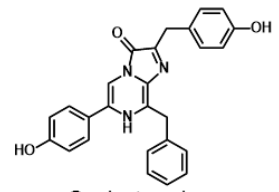
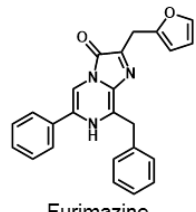
A complementary optical imaging modality, bioluminescence, is well suited for imaging cells noninvasively and on a macroscopic scale. Bioluminescence relies on enzymes (luciferases) that generate light via the chemical oxidation of small molecule substrates (luciferins) (Figure 1-5) (8, 11). When luciferase-expressing cells are exposed to luciferin, the photons produced can penetrate tissues and small organisms and be captured by sensitive cameras (11, 12). No excitation light source is required, and since mammalian tissues do not produce large numbers of photons, there is virtually no background signal. In fact, as few as one to ten luciferase-expressing cells can be observed subcutaneously in live animals (48, 49). Luciferase-expressing cells can also be serially imaged in live animals, enabling biological processes to be monitored noninvasively over both long time and length scales (8, 11).



**Figure 1-5.** Mechanism of bioluminescent light production. All characterized luciferase enzymes utilize a common reaction mechanism to generate light. The substrates and additional co-factors vary based on the luciferase.

Luciferase enzymes have been identified in a wide variety of organisms, including terrestrial and aquatic variants (Table 1-1) (8, 11). Insect luciferases, such as the North American firefly (Fluc) and the click beetle (CBluc) oxidize D-luciferin with molecular oxygen and ATP. The ATP-dependency of Fluc and CBluc has primarily limited their use to intracellular imaging. In contrast, marine luciferases, such as *Gaussia princeps* (Gluc) and NanoLuc (Nluc), catalyze the oxidation of coelenterazine or the designer substrate, furimazine (50-52). Marine luciferases only require molecular oxygen as a co-factor and are therefore routinely utilized in extracellular environments. The different luciferase-luciferin pairs emit light at varying wavelengths and the emission wavelength is influenced by enzyme residues and the luciferin structure (8, 11). All insect luciferases utilize the same substrate and subsequently all emit red light. In contrast, aquatic luciferases emit blue-green light (460-480 nm) (11, 51, 53). For *in vivo* imaging, red light (>600 nm) is most readily transmitted through blood and overlying tissue while blue-green light is readily absorbed by hemoglobin and other chromophores (8, 11, 54). Fluc and CBluc emit the largest percentage of red light and are preferred for *in vivo* applications requiring high sensitivity. While the peak emission of Gluc and Nluc are 460 and 480 nm, their broad emission spectra include wavelengths >600 nm and are therefore still useful for bioluminescence imaging. (8, 11).

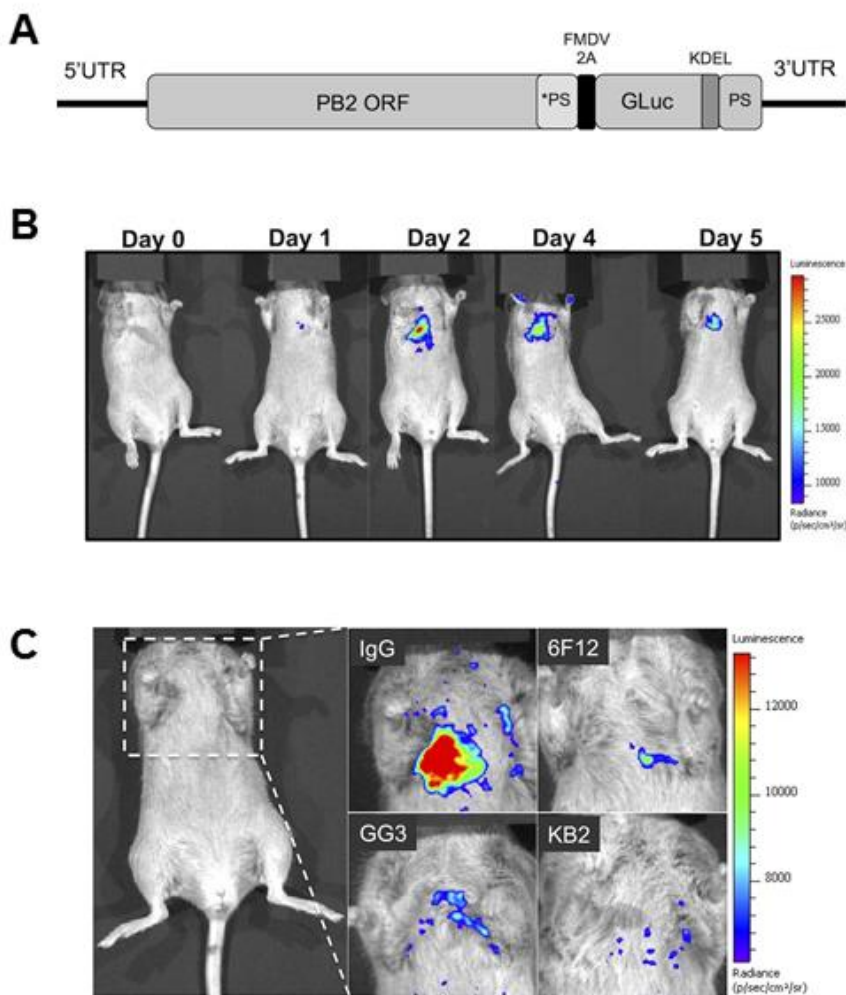
**Table 1-1.** Characteristics of commonly used luciferase-luciferin pairs.

Luciferase	Luciferin	Size (kDa)	Maximum emission $\lambda$ (nm)
Firefly luciferase (Fluc)	 D-Luciferin	61	560
click beetle green (CB green)		61	540
click beetle red (CB red)		61	615
<i>Renilla reniformis</i> (Rluc)	 Coelenterazine	36	480
<i>Gaussia princeps</i> (Gluc)		19.9	480
Nanoluc (Nluc) (Engineered from <i>Oplophorus gracilirostris</i> )	 Furimazine	19	460

Bioavailability and stability of the luciferin substrates are also important factors when selecting a luciferase-luciferin pair for *in vivo* imaging. D-Luciferin is relatively stable and can diffuse through most cell types and tissues (11, 55). These properties afford long-lived luminescent signal in the presence of luciferase-labeled cells. In comparison, coelenterazine is less bioavailable and is more prone to nonspecific oxidation. Coelenterazine is therefore routinely administered via intravenous injection to reach its *in vivo* targets (56, 57). D-Luciferin can be delivered through intraperitoneal injections, which is beneficial for long-term imaging studies and ease of use (58). These collective biochemical properties of bioluminescence have

thus enabled numerous studies of biological phenomena in truly *in vivo* environments, including cell homing, tumor cell proliferation, and pathogen egress.

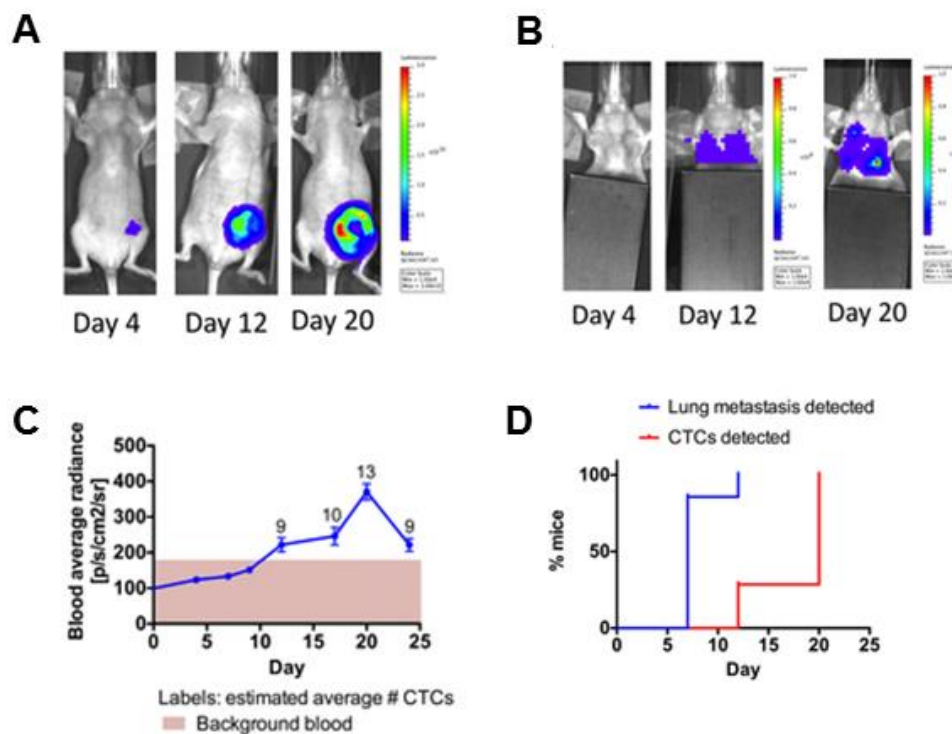
Recently, the Palese lab examined influenza A virus viral replication *in vivo* and evaluated antiviral therapies (59). The authors developed a fully infectious influenza virus that encodes Gluc (PR8-Gluc, Figure 1-6A) and monitored the dynamics of viral infection and spread in real time (Figure 1-6B). Having established a model for visualizing influenza viral infection in living animals, they extended the tool to evaluate novel therapeutic monoclonal antibodies. The antibodies (or an isotype control, IgG) were administered and the mice were subsequently challenged with PR8-Gluc. The study found the monoclonal antibodies can protect from disease and correlated with a significant reduction in viral replication as determined by imaging analysis (Figure 1-6C).



**Figure 1-6.** Luciferase-reporter influenza viruses can be used to monitor viral infection and evaluate antiviral therapies. (A) Gluc was inserted at the C terminus of the viral protein polymerase PB2. (B) Mice were infected with  $10^5$  plaque forming units of PR8-Gluc and imaged over time. (C) Antibody therapies result in reduced PR8-Gluc replication *in vivo*. Reduction of luciferase signal strength and area were observed for antibody treatment groups (6F12, GG3, and KB2) as compared to an isotype control (IgG). Image reproduced with permission from (Heaton et al. *J. Virol.* **2013**, 87, 8272-8281).

The ability to sensitively visualize biological processes over time has similarly proved to be useful to studies of cancer development and progression (60-62). For example, the Gambhir group examined circulating tumor cells (CTCs) in a mouse model of breast cancer metastasis (63). CTCs in the bloodstream are associated with poor patient prognosis but little is known

about CTC numbers during cancer progression. The authors implanted Fluc-expressing 4T1 cells in the murine mammary fat pad and simultaneously monitored for breast tumor growth (Figure 1-7A), lung metastases (Figure 1-7B), and CTCs over 24 days. The sensitivity of this approach allowed for detection of nine CTCs in whole blood, an important criterion since CTCs are rare events. The study found that CTCs increased from day 9 to day 20 after tumor implantation (Figure 1-7C). Furthermore, the timing of CTC detection correlated with the detection of lung metastasis (Figure 1-7D).



**Figure 1-7.** Bioluminescence imaging of multiple aspects of cancer progression. (A) Primary tumor growth and (B) metastases in the upper body (lung) were imaged after implantation of  $2 \times 10^7$  Fluc-labeled 4T1 cells. (C) Blood draws were performed to monitor CTCs dynamics. Blood samples were imaged and quantification is shown. (D) The detection of CTCs correlates with lung metastasis. Image reproduced with permission from (Sasportas et al. *PLoS One* **2014**, *9*, e105079).

#### **1.4 Future challenges and applications of bioluminescence imaging**

Despite its versatility, bioluminescence, as with all macroscopic imaging techniques, suffers from low resolution. While fluorescence imaging can visualize cell movements and interacting cells, the area visualized is confined to a preselected, sub-region of the whole organism. Thus, it remains difficult to monitor spontaneous cell-cell interactions relevant to organismal biology and disease. To address these issues and combine the resolution of fluorescence with the scale of bioluminescence, the Prescher group recently developed a bioluminescent strategy to report on cell-cell proximity in whole animals (64). In this approach, a non-functional luciferin is enzymatically cleaved by “activator” cells resulting in a functional luciferin. Neighboring “reporter” cells can then utilize the luciferin to produce light. In this scenario, the signal thus correlates with the proximity of these two cell populations. This strategy was further employed for visualizing and identifying areas of early metastatic disease. While this approach enabled cell-cell proximity to be readily visualized, it provides only a transient readout on cellular interactions and requires access to a custom synthetic luciferin. Furthermore, while the spatial resolution of this technology (~5mm) is applicable for visualizing global metastatic events, it is insufficient for monitoring direct cell-cell contacts. Thus, there are no general toolsets for the long-term visualization of tumor cell contacts with stromal cells, immune cells, and at metastatic sites.

Additionally, bioluminescence imaging has historically been limited to monitoring one cell type or biological feature at a time. Unlike the large palette of fluorescent proteins, there are only a handful of distinct luciferase-luciferin pairs that have been optimized for bioluminescence imaging. Although these probes emit light with different emission profiles, they are challenging to distinguish in living systems. While spectral resolution can be difficult to achieve, substrate



resolution is well established with native luciferase-luciferin pairs. The different classes of luciferases (terrestrial vs. aquatic) catalyze light emission by oxidizing chemically distinct substrates. These pairs are more readily discerned and have even been used in tandem. For example, T cells expressing both Gluc and CD19-targeted chimeric antigen receptor were shown to traffic to and eradicate Fluc-expressing tumors (65). Sequential administration of coelenterazine and D-luciferin enabled T cell trafficking and tumor burden to be monitored simultaneously. Similar terrestrial-aquatic bioluminescent pairs have been used to monitor metastatic processes, T regulatory and effector T cell functions, TGF $\beta$  signal transduction and drug targeting, and host-pathogen interactions with *Aspergillus* and *Candida* fungal infections (66-72). While informative, these studies are still technically demanding (due to coelenterazine administration) and are limited to monitoring two cell types or biological processes at a time.

### **1.5 Objectives of this study**

Imaging technologies have revolutionized our understanding of living systems by allowing researchers to visualize biological features in real time. Many of the advances in optical imaging have come in the fluorescence realm but the equivalent tools and methods are limited or nonexistent for bioluminescence. To address these issues, my thesis work involved engineering new luciferase tools directed at expanding the bioluminescence toolkit.

I aimed to:

1. Develop genetically encodable bioluminescent probes for visualizing interacting cells.
2. Engineer novel luciferase-luciferin pairs for multi-component imaging via directed enzyme evolution.
3. Apply these tools to *in vivo* imaging models.

## References

1. Batista FD & Dustin ML (2013) Cell:cell interactions in the immune system. *Immunol. Rev.* 251(1):7-12
2. Mayor R & Carmona-Fontaine C (2010) Keeping in touch with contact inhibition of locomotion. *Trends Cell Biol.* 20(6):319-328
3. Ng CT, Snell LM, Brooks DG, & Oldstone MB (2013) Networking at the level of host immunity: immune cell interactions during persistent viral infections. *Cell Host Microbe* 13(6):652-664
4. Perrin S (2014) Preclinical research: Make mouse studies work. *Nature* 507(7493):423-425
5. Troy T, Jekic-McMullen D, Sambucetti L, & Rice B (2004) Quantitative comparison of the sensitivity of detection of fluorescent and bioluminescent reporters in animal models. *Mol. Imaging.* 3(1):9-23
6. Allport JR & Weissleder R (2001) *In vivo* imaging of gene and cell therapies. *Exp. Hematol.* 29(11):1237-1246
7. Germain RN, Robey EA, & Cahalan MD (2012) A decade of imaging cellular motility and interaction dynamics in the immune system. *Science* 336(6089):1676-1681
8. Prescher JA & Contag CH (2010) Guided by the light: visualizing biomolecular processes in living animals with bioluminescence. *Curr. Opin. Chem. Biol.* 14(1):80-89
9. Gheysens O & Gambhir SS (2005) Studying molecular and cellular processes in the intact organism. *Prog. Drug. Res.* 62:117-150
10. Dean KM & Palmer AE (2014) Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nat. Chem. Biol.* 10(7):512-523
11. Paley MA & Prescher JA (2014) Bioluminescence: a versatile technique for imaging cellular and molecular features. *MedChemComm* 5(3):255-267
12. Porterfield WB & Prescher JA (2015) Tools for visualizing cell-cell 'interactomes'. *Curr. Opin. Chem. Biol.* 24:121-130
13. Ai HW (2015) Fluorescent-protein-based probes: general principles and practices. *Anal Bioanal. Chem.* 407(1):9-15
14. Giepmans BN, Adams SR, Ellisman MH, & Tsien RY (2006) The fluorescent toolbox for assessing protein location and function. *Science* 312(5771):217-224

15. Hoffman RM (2013) Fluorescent proteins as visible *in vivo* sensors. *Prog. Mol. Biol. Transl. Sci.* 113:389-402
16. Pak YL, Swamy KM, & Yoon J (2015) Recent progress in fluorescent imaging probes. *Sensors (Basel)* 15(9):24374-24396
17. Lavis LD & Raines RT (2014) Bright building blocks for chemical biology. *ACS Chem. Biol.* 9(4):855-866
18. Resch-Genger U, Grabolle M, Cavaliere-Jaricot S, Nitschke R, & Nann T (2008) Quantum dots versus organic dyes as fluorescent labels. *Nat. Methods* 5(9):763-775
19. Terai T & Nagano T (2013) Small-molecule fluorophores and fluorescent probes for bioimaging. *Pflugers Arch.* 465(3):347-359
20. Kapuscinski J (1995) DAPI: a DNA-specific fluorescent probe. *Biotech. Histochem.* 70(5):220-233
21. Latt SA & Stetten G (1976) Spectral studies on 33258 Hoechst and related bisbenzimidazole dyes useful for fluorescent detection of deoxyribonucleic acid synthesis. *J. Histochem. Cytochem.* 24(1):24-33
22. Hughes LD, Rawle RJ, & Boxer SG (2014) Choose your label wisely: water-soluble fluorophores often interact with lipid bilayers. *PLoS One* 9(2):e87649
23. Ballou B, Ernst LA, & Waggoner AS (2005) Fluorescence imaging of tumors *in vivo*. *Curr. Med. Chem.* 12(7):795-805
24. Chen X, Conti PS, & Moats RA (2004) *In vivo* near-infrared fluorescence imaging of integrin  $\alpha v \beta 3$  in brain tumor xenografts. *Cancer Res.* 64(21):8009-8014
25. Koide Y, *et al.* (2012) Development of NIR fluorescent dyes based on Si-rhodamine for *in vivo* imaging. *J. Am. Chem. Soc.* 134(11):5029-5031
26. Kiessling F (2008) Noninvasive cell tracking. *Handb. Exp. Pharmacol.* (185 Pt 2):305-321
27. Chalfie M, Tu Y, Euskirchen G, Ward WW, & Prasher DC (1994) Green fluorescent protein as a marker for gene expression. *Science* 263(5148):802-805
28. Chudakov DM, Matz MV, Lukyanov S, & Lukyanov KA (2010) Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol. Rev.* 90(3):1103-1163
29. Lin MZ, *et al.* (2009) Autofluorescent proteins with excitation in the optical window for intravital imaging in mammals. *Chem. Biol.* 16(11):1169-1179

30. Rodriguez EA, *et al.* (2016) A far-red fluorescent protein evolved from a cyanobacterial phycobiliprotein. *Nat. Methods* 13(9):763-769
31. Shu X, *et al.* (2009) Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science* 324(5928):804-807
32. Roh-Johnson M, *et al.* (2014) Macrophage contact induces RhoA GTPase signaling to trigger tumor cell intravasation. *Oncogene* 33(33):4203-4212
33. Bain J, Gow NA, & Erwig LP (2015) Novel insights into host-fungal pathogen interactions derived from live-cell imaging. *Semin. Immunopathol.* 37(2):131-139
34. Florentino PT, *et al.* (2014) An historical perspective on how advances in microscopic imaging contributed to understanding the *Leishmania Spp.* and *Trypanosoma cruzi* host-parasite relationship. *Biomed. Res. Int.* 2014:565291
35. Chandra J, McCormick TS, Imamura Y, Mukherjee PK, & Ghannoum MA (2007) Interaction of *Candida albicans* with adherent human peripheral blood mononuclear cells increases *C. albicans* biofilm formation and results in differential expression of pro- and anti-inflammatory cytokines. *Infect. Immun.* 75(5):2612-2620
36. Cheong WF, Prahl SA, & Welch AJ (1990) A review of the optical-properties of biological tissues. *IEEE J. Quantum. Elect.* 26(12):2166-2185
37. Alexander S, Weigelin B, Winkler F, & Friedl P (2013) Preclinical intravital microscopy of the tumour-stroma interface: invasion, metastasis, and therapy response. *Curr. Opin. Cell Biol.* 25(5):659-671
38. Hawkins ED, *et al.* (2016) T-cell acute leukaemia exhibits dynamic interactions with bone marrow microenvironments. *Nature* 538(7626):518-522
39. Murooka TT & Mempel TR (2013) Intravital microscopy in BLT-humanized mice to study cellular dynamics in HIV infection. *J. Infect. Dis.* 208 Suppl 2:S137-144
40. Naumenko V, Jenne C, & Mahoney DJ (2016) Intravital microscopy for imaging the tumor microenvironment in live mice. *Methods Mol. Biol.* 1458:217-230
41. Seynhaeve AL & Ten Hagen TL (2016) High-resolution intravital microscopy of tumor angiogenesis. *Methods Mol. Biol.* 1464:115-127
42. Beerling E, Ritsma L, Vrisekoop N, Derksen PW, & van Rheenen J (2011) Intravital microscopy: new insights into metastasis of tumors. *J. Cell Sci.* 124(Pt 3):299-310
43. Rice BW, Cable MD, & Nelson MB (2001) *In vivo* imaging of light-emitting probes. *J. Biomed. Opt.* 6(4):432-440

44. Booth MJ (2007) Adaptive optics in microscopy. *Philos. Trans. A. Math. Phys. Eng. Sci.* 365(1861):2829-2843
45. Girkin JM, Poland S, & Wright AJ (2009) Adaptive optics for deeper imaging of biological samples. *Curr. Opin. Biotechnol.* 20(1):106-110
46. Santi PA (2011) Light sheet fluorescence microscopy: a review. *J. Histochem. Cytochem.* 59(2):129-138
47. Tomer R, Khairy K, & Keller PJ (2013) Light sheet microscopy in cell biology. *Methods Mol. Biol.* 931:123-137
48. Kim JB, *et al.* (2010) Non-invasive detection of a small number of bioluminescent cancer cells in vivo. *PLoS One* 5(2):e9364
49. Rabinovich BA, *et al.* (2008) Visualizing fewer than 10 mouse T cells with an enhanced firefly luciferase in immunocompetent mouse models of cancer. *Proc. Natl. Acad. Sci. U. S. A.* 105(38):14342-14346
50. Dothager RS, *et al.* (2009) Advances in bioluminescence imaging of live animal models. *Curr. Opin. Biotechnol.* 20(1):45-53
51. Hall MP, *et al.* (2012) Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem. Biol.* 7(11):1848-1857
52. Luker GD & Luker KE (2008) Optical imaging: current applications and future directions. *J. Nucl. Med.* 49(1):1-4
53. Tannous BA, Kim DE, Fernandez JL, Weissleder R, & Breakefield XO (2005) Codon-optimized Gaussia luciferase cDNA for mammalian gene expression in culture and in vivo. *Mol. Ther.* 11(3):435-443
54. Zhao H, *et al.* (2005) Emission spectra of bioluminescent reporters and interaction with mammalian tissue determine the sensitivity of detection in vivo. *J. Biomed. Opt.* 10(4):41210
55. Mofford DM, Adams ST, Jr., Reddy GS, Reddy GR, & Miller SC (2015) Luciferin amides enable *in vivo* bioluminescence detection of endogenous fatty acid amide hydrolase activity. *J. Am. Chem. Soc.* 137(27):8684-8687
56. Gil JS, Machado HB, & Herschman HR (2012) A method to rapidly and accurately compare the relative efficacies of non-invasive imaging reporter genes in a mouse model and its application to luciferase reporters. *Mol. Imaging. Biol.* 14(4):462-471

57. Mezzanotte L, *et al.* (2013) Evaluating reporter genes of different luciferases for optimized *in vivo* bioluminescence imaging of transplanted neural stem cells in the brain. *Contrast Media Mol. Imaging* 8(6):505-513
58. Keyaerts M, Caveliers V, & Lahoutte T (2012) Bioluminescence imaging: looking beyond the light. *Trends Mol. Med.* 18(3):164-172
59. Heaton NS, *et al.* (2013) *In vivo* bioluminescent imaging of influenza A virus infection and characterization of novel cross-protective monoclonal antibodies. *J. Virol.* 87(15):8272-8281
60. Luwor RB, Stylli SS, & Kaye AH (2015) Using bioluminescence imaging in glioma research. *J. Clin. Neurosci.* 22(5):779-784
61. Madero-Visbal RA, *et al.* (2012) Bioluminescence imaging correlates with tumor progression in an orthotopic mouse model of lung cancer. *Surg. Oncol.* 21(1):23-29
62. Ritelli R, *et al.* (2015) Pancreatic cancer growth using magnetic resonance and bioluminescence imaging. *Magn. Reson. Imaging* 33(5):592-599
63. Sasportas LS, Hori SS, Pratz G, & Gambhir SS (2014) Detection and quantitation of circulating tumor cell dynamics by bioluminescence imaging in an orthotopic mammary carcinoma model. *PLoS One* 9(9):e105079
64. Sellmyer MA, *et al.* (2013) Visualizing cellular interactions with a generalized proximity reporter. *Proc. Natl. Acad. Sci. U. S. A.* 110(21):8567-8572
65. Santos EB, *et al.* (2009) Sensitive *in vivo* imaging of T cells using a membrane-bound *Gaussia princeps* luciferase. *Nat. Med.* 15(3):338-344
66. Donat S, *et al.* (2012) Surface display of *Gaussia princeps* luciferase allows sensitive fungal pathogen detection during cutaneous aspergillosis. *Virulence* 3(1):51-61
67. Lewandrowski GK, Magee CN, Mounayar M, Tannous BA, & Azzi J (2014) Simultaneous *in vivo* monitoring of regulatory and effector T lymphocytes using secreted *Gaussia* luciferase, Firefly luciferase, and secreted alkaline phosphatase. *Methods Mol. Biol.* 1098:211-227
68. Maguire CA, *et al.* (2013) Triple bioluminescence imaging for *in vivo* monitoring of cellular processes. *Mol. Ther. Nucleic Acids* 2:e99
69. Serganova I, *et al.* (2009) Multimodality imaging of TGFβ signaling in breast cancer metastases. *FASEB J.* 23(8):2662-2672
70. Stacer AC, *et al.* (2013) NanoLuc reporter for dual luciferase imaging in living animals. *Mol. Imaging* 12(7):1-13

71. Vande Velde G, Kucharikova S, Schrevens S, Himmelreich U, & Van Dijck P (2014) Towards non-invasive monitoring of pathogen-host interactions during *Candida albicans* biofilm formation using in vivo bioluminescence. *Cell Microbiol.* 16(1):115-130
72. Vecchiarelli A & d'Enfert C (2012) Shedding natural light on fungal infections. *Virulence* 3(1):15-17

# **CHAPTER 2: Visualizing cell proximity with genetically encoded bioluminescent reporters**

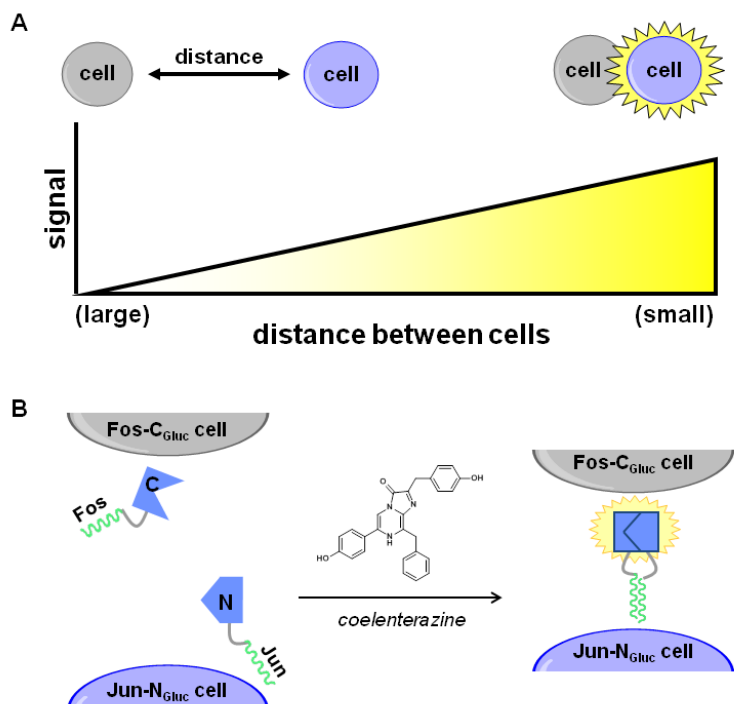
## **2.1 Introduction**

Complex cell–cell interactions regulate diverse facets of human biology, including organ development and immune function (1). Breakdowns in cellular communication also underlie numerous disease states, including cancer metastases (2-4). Given the importance of cell–cell contacts to health and disease, it is surprising that few practical methods exist to globally assay such networks in live organisms. Bioluminescence imaging can be used to track cells across whole tissues and animals, but can only approximate cell locations and interactions owing to its low spatial resolution (5, 6). Cell-to-cell contacts can be directly visualized using histology or intravital microscopy (7-9). However, these methods require invasive procedures and offer only microscopic—and in some cases static—views of biology (10, 11). To monitor cell–cell interactions in real time and across larger length scales, we are developing technologies that blend the exquisite spatial resolution and anatomical detail of microscopic techniques with the broad dynamic range and noninvasive features of bioluminescence. Such methods will be applicable to monitoring cell–cell interactions in tissues and rodent models, including those involved in tumor metastases, cellular immunotherapy, and host–pathogen interactions (12-15).

As noted in chapter 1, bioluminescence is well suited for imaging cells noninvasively and on a macroscopic scale. However, most applications of the technique focus on *single* populations of cells, rather than their interactions (16-18). When luciferase-expressing cells are exposed to luciferin, the photons produced can penetrate tissues and small organisms and be captured by



sensitive cameras. Luciferase-expressing cells can also be serially imaged in live animals, enabling biological processes to be monitored noninvasively over both long time and length scales (5). While ideal for global cell tracking, traditional bioluminescence imaging lacks the spatial resolution to image cell contacts (6). Interacting cells, in theory, could be visualized using unique luciferase–luciferin pairs to selectively illuminate the distinct cell types. In practice, though, these probes cannot be readily distinguished on time scales relevant to most cellular contacts (<days) (5, 6). To retool bioluminescence technology for visualizing cell–cell interactions, we recently reported a method to “turn on” photon production when one cell population interacts with a distinct, second population (Figure 2-1A). This strategy relies on the local release and diffusion of luciferin by “activator” cells and its use by neighboring luciferase-expressing (*i.e.*, “reporter”) cells (19). While suitable for some applications, this approach requires access to custom luciferins and provides only a transient readout on cell proximity.



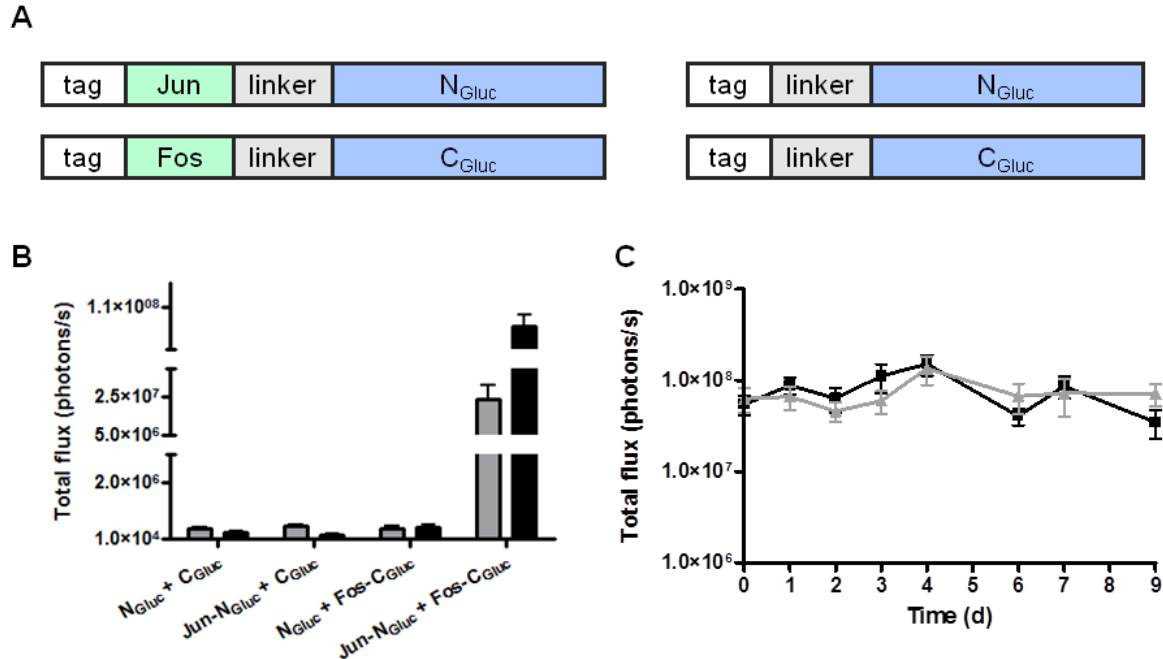
**Figure 2-1.** Imaging cell-cell interactions with bioluminescent probes. (A) Robust signal is observed only when the two distinct cell types are in close proximity. (B) Strategy for visualizing cell-cell interactions with split Gluc. Fos-Jun dimerization drives the assembly of secreted Gluc fragments when cells are in close contact. Complemented Gluc catalyzes light emission with the small molecule coelenterazine.

We aimed to develop a complementary and more accessible method for visualizing cell–cell interactions using bioluminescent protein complementation. Complementation strategies exploit “split” reporters—fragments of proteins or enzymes that assemble and produce signal only when brought into contact (20, 21). Split reporters, including split luciferases, are routinely used to visualize intracellular *protein–protein* interactions, and some have been adapted for use in live animals (22–24). Here, we show that split versions of a secreted luciferase can associate (complement) in the extracellular environment and provide a readout on *cell–cell* interactions. This strategy employs commercially available reagents and is suitable for imaging over both short and long time scales. Thus, the split probes are applicable to examining a broad range of cell–cell communications relevant to basic biology and disease.

## 2.2 Design of “split” forms of luciferase

The reporter that we selected for visualizing cell–cell interactions was *Gaussia* luciferase (Gluc). As noted in chapter 1, Gluc does not require ATP for the light-emitting reaction and has been shown to function in the extracellular space, important characteristics for imaging cell-cell interactions. (Figure 2-1B) (25). Gluc is also among the brightest and most stable luciferases known, making it attractive for imaging small numbers of cells (26, 27). Indeed, Gluc-antibody conjugates and Gluc-CD8 fusions have enabled sensitive imaging of tumor cells and T cells, respectively, in mouse models (28-31). Split versions of Gluc have also been generated and used to visualize protein–protein interactions, including chemokine receptor–ligand binding and amyloid aggregation (32-36).

To adapt Gluc complementation for visualizing cell–cell interactions, we required modular and generalizable protein domains to aid in assembly of the split fragments. Both halves of Gluc fold and remain soluble independent of one another but do not readily associate on their own (24). Thus, tight-binding domains are necessary to bring the fragments into proximity, drive complementation, and ultimately stabilize the assembled product. While several protein pairs can be envisioned for this purpose, we initially focused on the nuclear proteins Fos and Jun. These proteins form a stable, coiled-coil complex ( $K_d \sim 54$  nM) that normally binds DNA. However, Fos–Jun dimerization has been exploited in numerous other contexts, including the assembly of split GFP and other proteins (37-39). Fos-Jun binding affinities can be further tuned via electrostatic modification or other alterations to the coils (40).

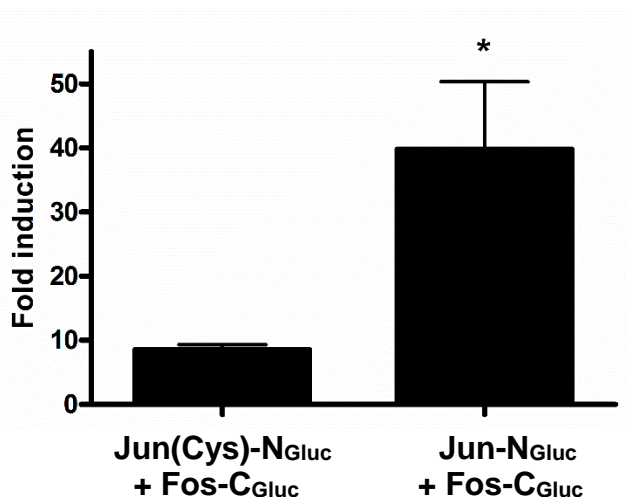


**Figure 2-2.** Fos and Jun facilitate the stable assembly of Gluc. (A) DNA constructs comprising Gluc halves tethered to Fos or Jun (via G4S linkers) and a FLAG affinity tag (Jun-N<sub>Gluc</sub>, Fos-C<sub>Gluc</sub>) were prepared. Control constructs with either no Jun (N<sub>Gluc</sub>) or no Fos (C<sub>Gluc</sub>) were also generated. (B) Fos and Jun are required for split Gluc complementation. HEK293 cells were engineered to express one of the constructs in (A), and the cells were mixed (1:1) as indicated. Supernatants were collected from the cocultures 24 h (gray) or 48 h (black) postplating and imaged with coelenterazine. Photon flux measurements were acquired using a cooled CCD camera. (C) Complemented Gluc is stable over time. HEK293 cells stably secreting Jun-N<sub>Gluc</sub> or Fos-C<sub>Gluc</sub> were mixed together. Media samples from the cocultures were collected and imaged over time (black line). In some cases, the media samples were added to HEK293 cells ( $5 \times 10^3$  cells) and imaged (gray line). For B and C, error bars represent the standard error of the mean for  $n = 3$  (B) or  $n = 4$  (C) experiments.

### 2.3 “Split” forms of luciferase can reassemble in the extracellular space to enable sensitive imaging of collections of interacting cells

To examine whether Fos–Jun dimerization could facilitate split Gluc complementation, we generated HEK293 cells that stably secrete an N-terminal- or C-terminal fragment of Gluc fused to Jun or Fos (Jun-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub>, respectively, Figure 2-2A). Combinations of these cells or control cells were cultured together, and media samples were assayed for Gluc activity using bioluminescence imaging. As shown in Figure 2-2B, robust light emission was observed in

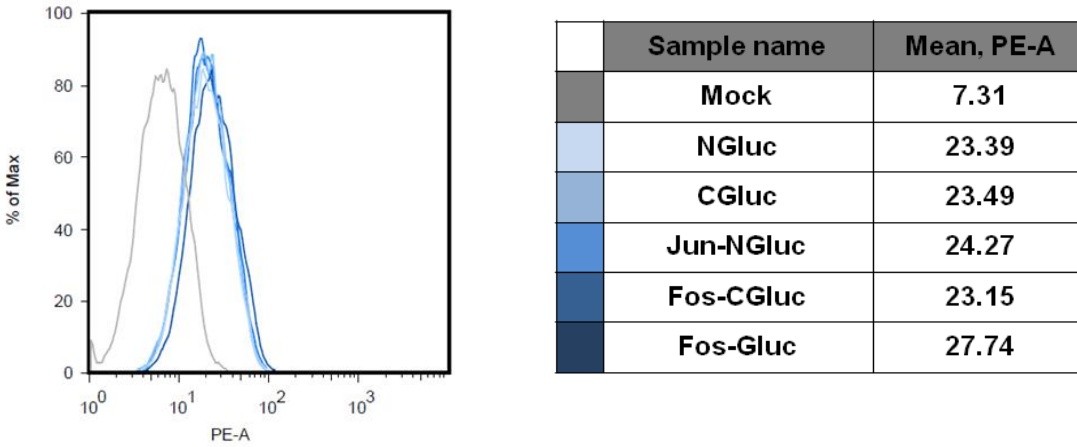
samples containing both Jun-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub>. No signal was observed in the absence of Jun or Fos, confirming that the coils play an essential role in extracellular Gluc assembly.



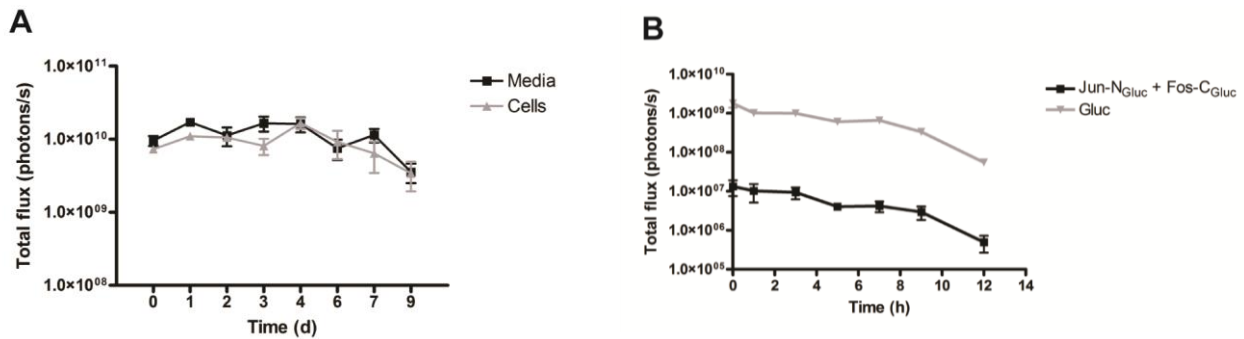
**Figure 2-3.** Cys-to-Ser mutations in Jun provide enhanced bioluminescent signal. HEK293 cells were transfected with plasmids encoding Jun(Cys)-N<sub>Gluc</sub>, Jun-N<sub>Gluc</sub>, or N<sub>Gluc</sub> and mixed with cells secreting Fos-C<sub>Gluc</sub>. Supernatants were collected and imaged 24 h post-plating. The fold inductions in bioluminescent signal versus the control (N<sub>Gluc</sub>) are plotted. Error bars represent the standard error of the mean for n = 3 experiments, \*p < 0.01.

In previous studies, high levels of Jun were found to promote self-dimerization (due to disulfide bond formation) (41, 42). Such homodimerization would inhibit functional Gluc formation and reduce light emission in our case. To mitigate against Jun–Jun binding, and favor Fos–Jun heterodimerization, we engineered a Cys-to-Ser mutation along the Jun coil. When this fusion was evaluated with Fos-C<sub>Gluc</sub> in coculture assays, an ~5-fold enhancement in bioluminescence was observed (Figure 2-3 and 2-4). Complemented Gluc formed via mutant Jun binding to Fos was also long-lived. HEK293 cells secreting mutant Jun-N<sub>Gluc</sub> were cultured with cells secreting Fos-C<sub>Gluc</sub>, and media samples were collected. The aliquots (containing assembled Gluc) were diluted with additional media and assayed for light emission over time. As shown in Figure 2-2C and Figure 2-5, robust bioluminescence was observed for over 1 week, even in dilute solution. Similar light emission trends were observed when complemented Gluc was

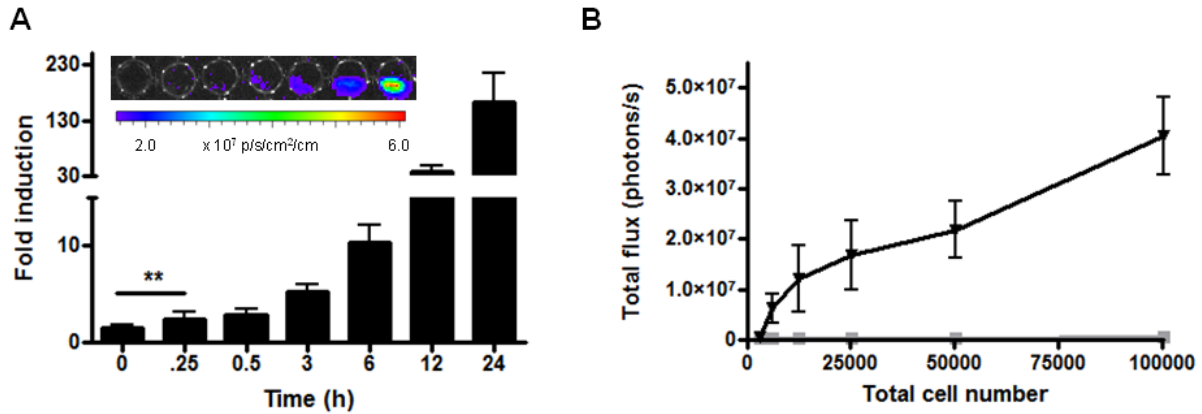
incubated in the presence of live cells. These data underscore the remarkable stability of Gluc, consistent with previous reports (27, 43). The mutant form of Jun was also used in all subsequent studies.



**Figure 2-4.** Gluc fragments are expressed at comparable levels. HEK293 cells stably expressing either the FLAG-tagged Fos/Jun-Gluc fragments or control constructs were fixed, permeabilized, stained with Cy3-anti-FLAG, and analyzed by flow cytometry.

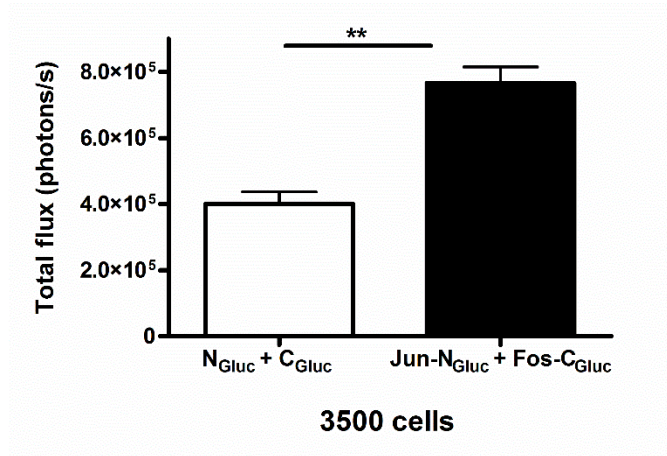


**Figure 2-5.** Native Gluc is stable over time (A) Media samples from Gluc-secreting cells were assayed for bioluminescence activity. The samples were imaged directly (black line), or in some cases, added to HEK293 cells ( $5 \times 10^3$  cells) and imaged (gray line). (B) Gluc degradation is observed in the presence of high concentrations of serum. Jun-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub>-secreting cells were mixed (1:1 ratio). Media samples were collected and incubated with FBS and imaged (black line). Media samples from cells secreting holo Gluc were similarly collected, incubated with FBS, and imaged (gray line). For (A)-(B), error bars represent the standard error of the mean for  $n = 4$  (A) or  $n = 3$  (B) experiments.



**Figure 2-6.** Gluc complementation provides a rapid and sensitive readout on cell-cell interactions. (A) HEK293 cells stably secreting Jun-N<sub>Gluc</sub>, Fos-C<sub>Gluc</sub>, N<sub>Gluc</sub> only, or C<sub>Gluc</sub> only were mixed in 96-well round-bottom plates (1:1, 5 × 10<sup>4</sup> total cells). Supernatants from the cocultures were collected 0-24 h postplating and imaged with coelenterazine. The fold inductions in bioluminescent signal for coplates of Jun-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub>-secreting cells versus N<sub>Gluc</sub>- and C<sub>Gluc</sub>-secreting cells (control cells) are plotted. Samples images for the mixtures of Jun-N<sub>Gluc</sub>- and Fos-C<sub>Gluc</sub>-secreting cells (at each time point) are also shown. (B) HEK293 cells stably secreting Jun-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub> were plated together (1:1, 500-100000 cells total) in round-bottom plates. After 24 h, supernatants from the cocultures were incubated with coelenterazine and imaged (black). Supernatants from mixtures of control cells were similarly collected and imaged (gray). For A and B, error bars represent the standard error of the mean for n = 4 experiments.

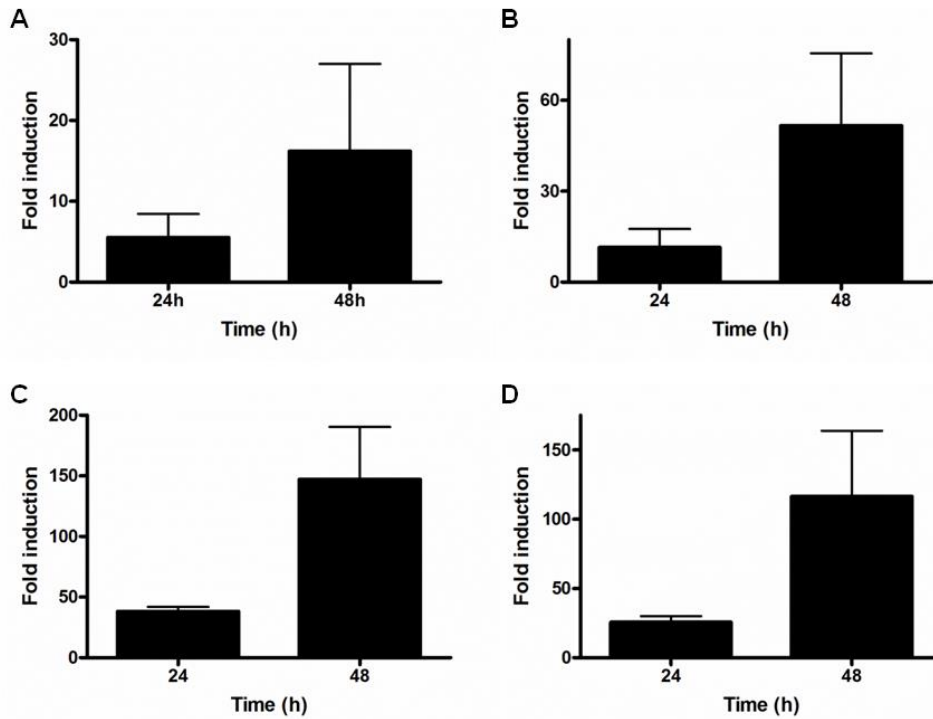
We next examined whether Gluc complementation could provide a fast and sensitive readout on collections of interacting cells. Cells secreting either Jun-N<sub>Gluc</sub> or Fos-C<sub>Gluc</sub> were plated together (50 000 cells total), and media samples were assayed for light emission over time. Photon production was observed 15 min postplating and increased exponentially over a 24-h period as complemented Gluc accumulated in the media (Figure 2-6A). Prolonged coculture also enabled smaller numbers of interacting cells to be detected (Figure 2-6B and 2-7).



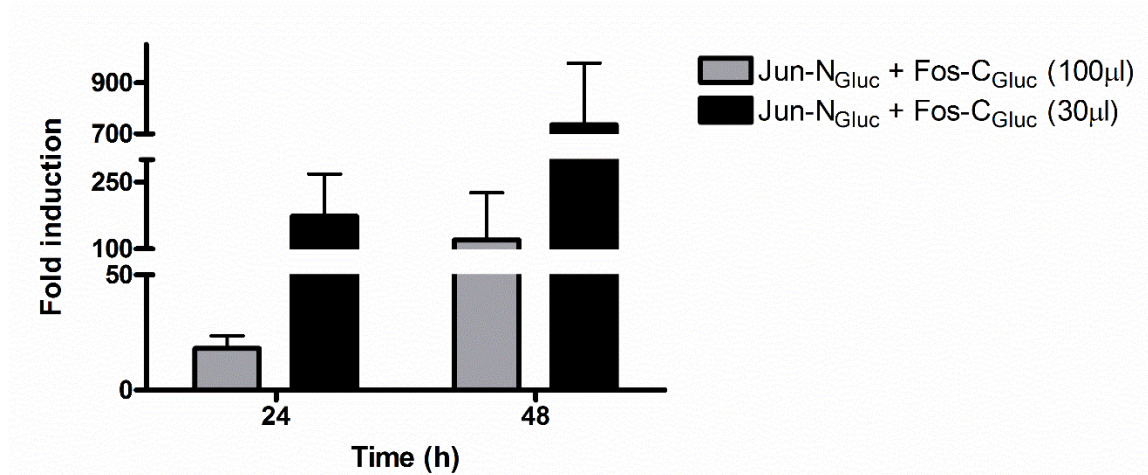
**Figure 2-7.** Small numbers of cultured cells can be visualized using split Gluc fragments. Combinations of HEK293 cells stably secreting Jun- $N_{Gluc}$ , Fos- $C_{Gluc}$ ,  $N_{Gluc}$ , or  $C_{Gluc}$  were plated (1:1) in round-bottom plates. After 24 h, media samples from the co-cultures were incubated with coelenterazine and imaged. As few as 3500 total Jun- $N_{Gluc}$ - and Fos- $C_{Gluc}$ -secreting cells could be detected. Error bars represent the standard error of the mean for  $n = 4$  experiments., \*\*  $p < 0.001$

Similar trends were observed across a variety of cell types (Figure 2-8), indicating that split Gluc assembly is applicable to monitoring a broad range of cell–cell interactions. Stronger bioluminescent outputs were also achieved using smaller volumes of media in the coculture assays (Figure 2-9). Reduced volumes increase the local concentrations of the split fragments and, thus, drive Gluc assembly. It should be noted, though, that high concentrations of Jun- $N_{Gluc}$  diminish bioluminescent output, presumably due to Jun homodimerization (Figure 2-10). Conversely, signal reduction is not observed with increasing concentrations of Fos- $C_{Gluc}$ . While careful pairing of the split fragments and cell types is necessary to ensure maximum light production, our experiments suggest that the split reporters will be suitable for imaging long-lived cell interactions (*e.g.*, T-cell priming by infected cells or cancer cell outgrowth in metastatic sites), along with “permanent” cell–cell interactions (*e.g.*, fusion events) (44-49).

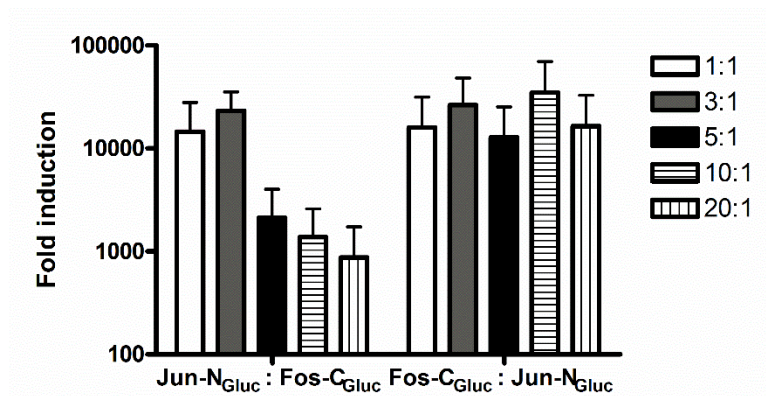




**Figure 2-8.** Split Gluc technology is generalizable (A) 4T1-Luc2 cells and (B) Jurkat cells transiently expressing Fos-C<sub>Gluc</sub>, Jun-N<sub>Gluc</sub>, C<sub>Gluc</sub>, or N<sub>Gluc</sub> were generated. (C) DB7 cells and (D) MDA-MB-231 cells stably expressing Fos-C<sub>Gluc</sub>, Jun-N<sub>Gluc</sub>, C<sub>Gluc</sub>, or N<sub>Gluc</sub> were also generated. For (A)-(D), equivalent expression levels were confirmed by flow cytometry. Combinations of the cells were plated together in round-bottom plates (1:1, 5 x 10<sup>4</sup> cells total). After 24 or 48 h, media samples from the co-cultures were collected and imaged. The fold inductions in bioluminescence signal for Jun-N<sub>Gluc</sub>:Fos-C<sub>Gluc</sub> co-plates versus the control co-plates (N<sub>Gluc</sub>:C<sub>Gluc</sub>) are plotted. For (A)-(D), error bars represent the standard error of the mean for n = 3 experiments.



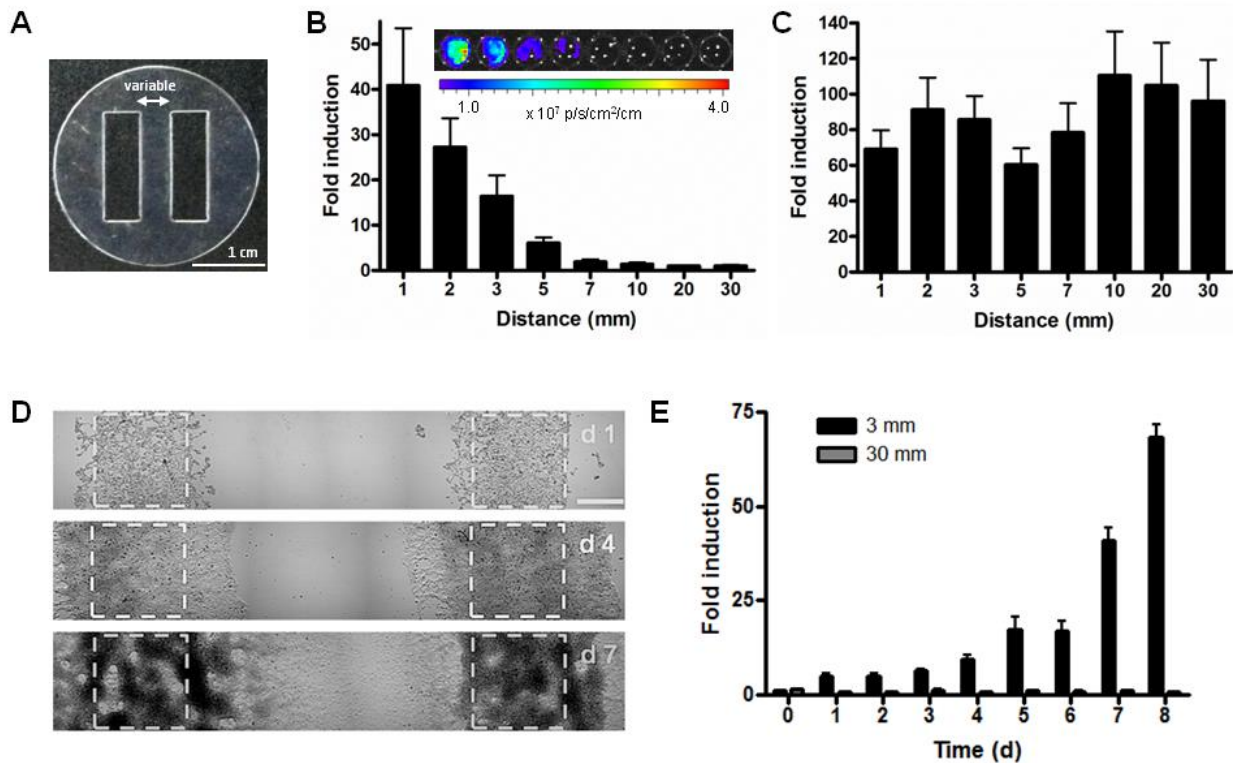
**Figure 2-9.** Increased bioluminescent signal is observed with smaller volumes. Combinations of HEK293 cells stably secreting Fos-C<sub>Gluc</sub>, Jun-N<sub>Gluc</sub>, C<sub>Gluc</sub>, or N<sub>Gluc</sub> were placed in round-bottom plates (1:1, 5 x 10<sup>4</sup> cells total) containing either 30 µL (black) or 100 µL (gray) of media. Equal volumes of media from the co-cultures were then collected and imaged with coelenterazine. The fold inductions in bioluminescent signal for Jun-N<sub>Gluc</sub>:Fos-C<sub>Gluc</sub> co-plates versus the control co-plates (N<sub>Gluc</sub>:C<sub>Gluc</sub>) are plotted. Error bars represent the standard error of the mean for n = 3 experiments.



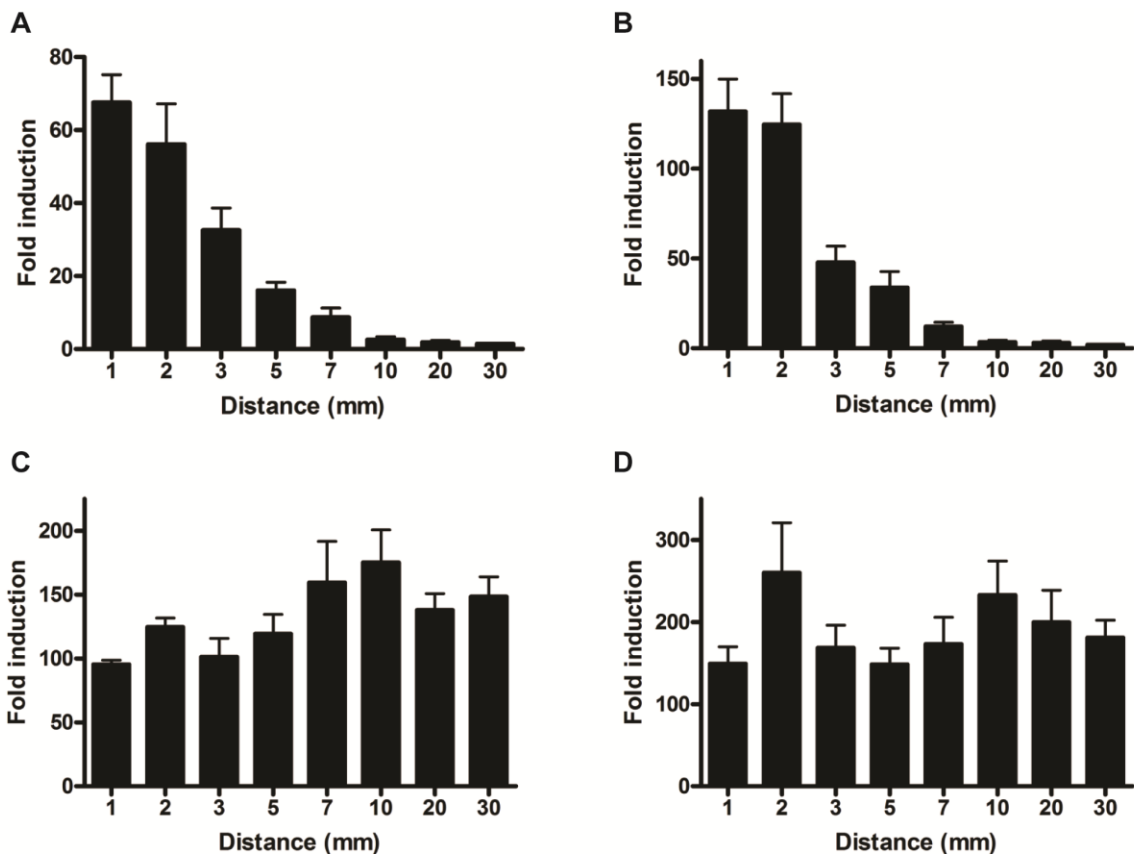
**Figure 2-10.** Larger numbers of Jun-N<sub>Gluc</sub>-secreting cells decrease bioluminescent signal. HEK293 cells stably secreting Jun-N<sub>Gluc</sub> or Fos-C<sub>Gluc</sub> were plated together in varying ratios (1:0–20:1, 2 x 10<sup>5</sup> cells total). After 24 h, media samples from the co-cultures were incubated with coelenterazine and imaged. The fold inductions in bioluminescent signal for Jun-N<sub>Gluc</sub>:Fos-C<sub>Gluc</sub> or Fos-C<sub>Gluc</sub>:Jun-N<sub>Gluc</sub> co-plates versus the control co-plates (N<sub>Gluc</sub>:C<sub>Gluc</sub>) are plotted. Error bars represent the standard error of the mean for n = 3 experiments.

## 2.4 “Split” luciferase fragments enable facile imaging of cell-cell proximity in cultured models

Since Gluc assembly is dependent on the concentrations of the split fragments, we reasoned that luciferase complementation could provide a direct readout on the *distance* between two distinct cell populations. The local concentrations of Jun-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub> are greatest near the cells secreting them; thus, more complemented protein should form as the cell types draw near one another. To test this hypothesis, we patterned Fos-C<sub>Gluc</sub> and Jun-N<sub>Gluc</sub>-secreting cells at defined positions using biocompatible stencils (prepared by David J. Li, Hui lab, UCI) (Figure 2-11A) (50). Cells were seeded in the cutouts and allowed to adhere prior to stencil removal and imaging. As shown in Figure 2-11B, bioluminescent signal increased as the distance between the cells decreased. Millimeter spacings could be readily resolved, and the distance-dependent nature of light emission persisted over time, highlighting the robustness of the strategy (Figure 2-12). Complemented Gluc also formed when media was flowed across the cocultures to mimic an *in vivo* environment (Figure 2-13A). In these cases, the distance-dependent nature of photon production was even more pronounced (Figure 2-13B), suggesting that the probes will provide a faithful readout on direct cell–cell interactions even in more stringent physiological environments. In control experiments, Fos-C<sub>Gluc</sub>-secreting cells and Jun-N<sub>Gluc</sub>-secreting cells were plated together in each stencil cutout. In these cases, equivalent levels of bioluminescent signal were observed regardless of the distances between the cutouts (Figure 2-11C and 2-13B).



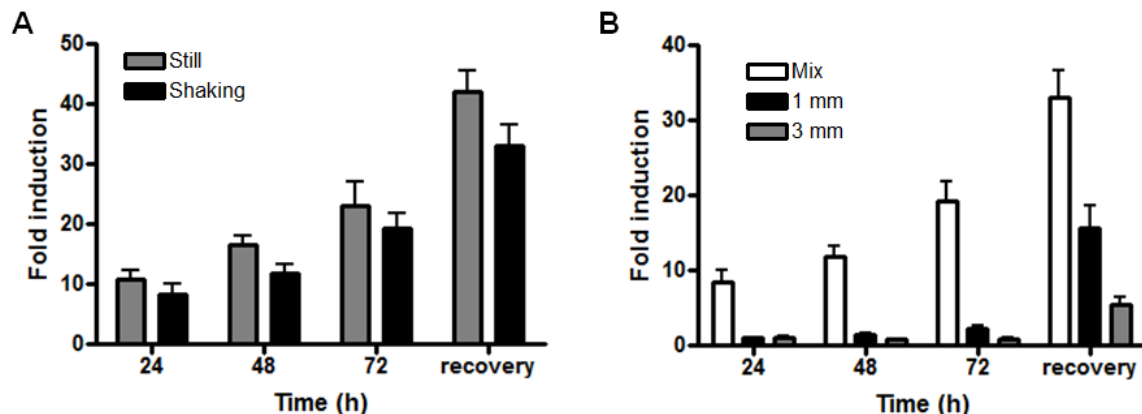
**Figure 2-11.** Gluc complementation is distance dependent and can report on cell migration. (A) Representative PDMS stencil used to plate cells at defined distances. (B) Cells expressing Jun- $N_{Gluc}$  were plated 1-30 mm apart from Fos- $C_{Gluc}$ -expressing cells or control cells (expressing  $C_{Gluc}$  only) using stencils from A. The total volume of media in each cutout was 30  $\mu$ L. After 24 h, fresh media was added to cover the surface of each stencil and enable the supernatants to mix. Media aliquots were then collected and imaged as in Figure 2. The fold inductions in bioluminescent signal for the Jun- $N_{Gluc}$ /Fos- $C_{Gluc}$  samples versus the control samples are plotted. Sample bioluminescence images for the Jun- $N_{Gluc}$ /Fos- $C_{Gluc}$  pairings at each distance are also shown. (C) Cells expressing Jun- $N_{Gluc}$  were plated in a 1:1 ratio with Fos- $C_{Gluc}$ -expressing cells or control cells *in each cutout* (two per stencil). The samples were imaged as in B, and the fold inductions in bioluminescent signal for the Jun- $N_{Gluc}$ /Fos- $C_{Gluc}$  samples versus the control samples are plotted. (D) Cells expressing Jun- $N_{Gluc}$  were plated 3 or 30 mm apart from Fos- $C_{Gluc}$ -expressing cells ( $5.0 \times 10^4$  cells/stencil cutout). After 24 h, the stencils were removed, and fresh media (2 mL) was added to each culture. Cells were allowed to migrate over time and media was replaced every 24 h to prevent the buildup of complemented Gluc. Sample bright-field images of the 3 mm separated cells are shown at different time points. Cells were initially seeded in the dashed boxes. Scale bar: 500  $\mu$ m, 4x magnification. (E) Light emission correlates with the proximity between the two cell populations. The samples in D were imaged, and the fold inductions in bioluminescent signal for Jun- $N_{Gluc}$ -expressing cells plated 3 mm (black) or 30 mm apart (gray) from Fos- $C_{Gluc}$  samples versus control cells are shown. For B and C, error bars represent the standard error of the mean for  $n = 6$  experiments. For E, error bars represent the standard error of the mean for  $n = 3$  replicates. The data are representative of four independent experiments.



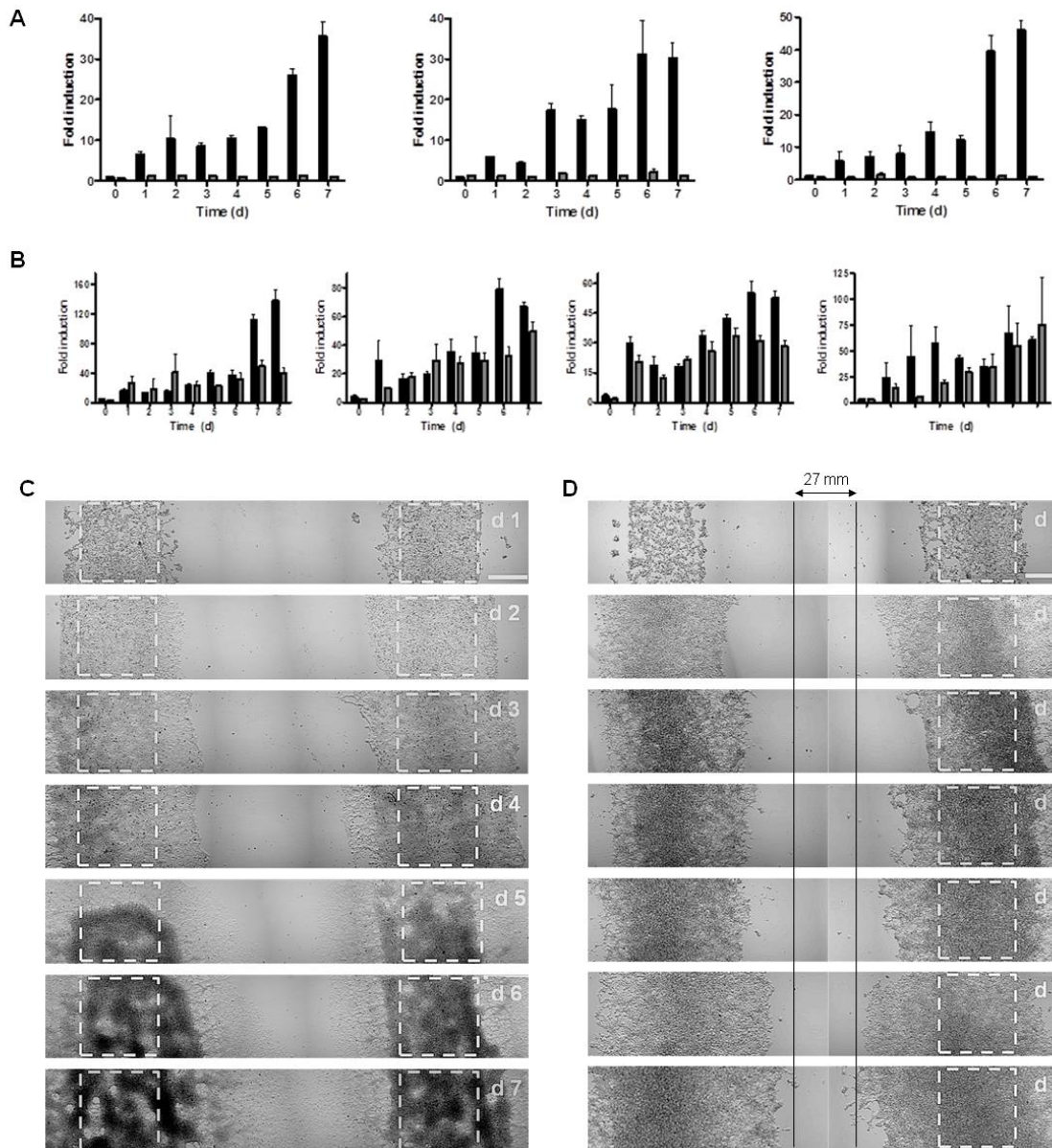
**Figure 2-12.** Distance-dependent light emission is maintained over time. HEK293 cells stably secreting Fos-C<sub>Gluc</sub>, Jun-N<sub>Gluc</sub>, C<sub>Gluc</sub>, or N<sub>Gluc</sub> were plated separately in each stencil cutout. Media was added 24 h after plating to allow the supernatants from each cutout to mix. Media samples from these cultures were then collected 48 h (A) and 72 h (B) after plating and imaged. (C)–(D) Cells secreting Jun-N<sub>Gluc</sub> or Fos-C<sub>Gluc</sub> were plated (1:1 ratio) in the same cutout. Additional media was added 24 h after plating to allow the supernatants from each cutout to mix. Media samples were then collected 48 (C) and 72 h (D) after plating and imaged. For (A)–(D), the fold inductions in bioluminescent signal for Jun-N<sub>Gluc</sub>:Fos-C<sub>Gluc</sub> mixtures versus the control mixtures (N<sub>Gluc</sub>:C<sub>Gluc</sub>) are plotted. Error bars represent the standard error of the mean for n = 6 experiments.

Encouraged by these results, we examined whether the split probes could be used to monitor cell proximity over time. Such assays have immediate relevance to immune cell migration, cancer metastases, and other cell movements. Traditional assays of cell migration employ transwell devices—porous membranes that cells can traverse in response to various stimuli (including other cells). Cells that migrate across the membranes are ultimately detected

using microscopy or other counting methods (51). While well established, such assays rely on static, end-point analyses and cannot capture the dynamic nature of cell migration. Moreover, they are not well suited for monitoring cell migration across larger, macroscopic distances (>2 cm). We envisioned using Gluc complementation to monitor cellular movements in real time. Such assays would enable changes in cell proximity to be detected from media aliquots, without perturbing the cells themselves. Toward this end, we patterned cells secreting either Jun-N<sub>Gluc</sub> or Fos-C<sub>Gluc</sub> at different distances (3 or 30 mm) using the stencils pictured in Figure 2-11A. Upon adherence and stencil removal, cell migration was monitored over time via media sampling and bioluminescence imaging (Figure 2-11D and Figure 2-14). Media was also replaced every 24 h to prevent the buildup of complemented Gluc. Excitingly, as the distance between the cells decreased, a corresponding increase in bioluminescence signal was observed (Figure 2-11E and Figure 2-14). Maximal signal was achieved when the cell populations completely merged. Notably, even though the number of cells increased via cell division throughout these studies, equivalent numbers were present in both the 3- and 30 mm cultures and no cell death was observed. Thus, bioluminescent signal (and split Gluc assembly) could provide a direct readout on the relative distance between two cell populations over time.



**Figure 2-13.** Gluc complementation is observed under flow. (A) DB7 cells stably secreting Jun- $N_{Gluc}$  or Fos- $C_{Gluc}$  were plated in the same cutout (1:1 ratio,  $1 \times 10^5$  cells/stencil cutout). Control cells expressing  $N_{Gluc}$  or  $C_{Gluc}$  were also plated in separate cutouts. The cultures were incubated with no agitation (still, gray) or subjected to shaking (black) and imaged every 24 h. After 72 h, all samples were allowed to stand for 24 h prior to imaging (“recovery”). (B) Cells secreting Jun- $N_{Gluc}$  or Fos- $C_{Gluc}$  were plated together (white), 1 mm apart (black), or 3 mm apart (gray) and incubated with shaking. The cultures were imaged every 24 h as in (A). After 72 h, all samples were allowed to stand for 24 h prior to imaging (“recovery”). For (A)–(B), the fold inductions in bioluminescent signal for Jun- $N_{Gluc}$ :Fos- $C_{Gluc}$  mixtures versus the control ( $N_{Gluc}$ : $C_{Gluc}$ ) mixtures are plotted. Error bars represent the standard error of the mean for samples analyzed in triplicate. Data are representative of three replicate experiments.



**Figure 2-14.** In vitro cell migration model. Cells stably secreting Jun- $N_{Gluc}$ , Fos- $C_{Gluc}$ ,  $N_{Gluc}$ , or  $C_{Gluc}$  were plated ( $5 \times 10^4$  cells/stencil cutout). Stencils were removed 24 h post plating. Additional media was also added to allow secreted fragments from each cell population to mix. Media was replaced every 24 h to avoid the buildup of complemented Gluc. (A) Bioluminescence data from three experiments are shown. Cells were plated 3 mm (black) or 30 mm (gray) apart and allowed to migrate over time. The fold inductions in bioluminescent signal for Jun- $N_{Gluc}$  and Fos- $C_{Gluc}$  mixtures versus control mixtures ( $N_{Gluc}:C_{Gluc}$ ) are plotted. Light emission correlates with the proximity between the two cell populations. (B) Bioluminescence data for three experiments where cells were plated together in each stencil cutout and allowed to migrate over time. The fold inductions in bioluminescent signal for Jun- $N_{Gluc}$  and Fos- $C_{Gluc}$  mixtures versus control mixtures ( $N_{Gluc}:C_{Gluc}$ ) are plotted. (C) Bright-field images of the cells patterned at 3 mm and (D) 30 mm to monitor changes in distance over time (Magnification: 4x). Cells were initially seeded in the dashed boxes (Scale bar: 500  $\mu$ m). For (A)–(B), error bars represent the standard error of the mean for experiments performed in triplicate.



## **2.5 Conclusions and future directions**

In conclusion, we developed a strategy for visualizing cell–cell proximity using engineered bioluminescent probes. These tools comprise split versions of an extracellular luciferase (Gluc) that assemble and form a functional, light-emitting enzyme when cells are in close proximity. Both large numbers of interacting cells and small intercellular distances provide robust bioluminescence in the current approach; these parameters are difficult to distinguish in a single imaging session but can be decoupled in serial imaging studies. Thus, the tools will be immediately applicable to studies of cell–cell interactions that are difficult to observe with conventional toolsets, including prolonged tumor–immune cell contacts and cell fusion events, among others (44, 46, 52, 53). Transient cell contacts cannot be reliably imaged with the current probes, although split reporters with faster folding rates or improved photon outputs could potentially capture such interactions. Intriguingly, since Gluc is long-lived in the extracellular environment, the complemented probe should enable the direct detection of cell–cell interactions in locations that are distinct from, and perhaps far removed from, the actual site of cell–cell interaction *in vivo* (e.g., in the bloodstream). We are currently evaluating this possibility in mouse models. We also envision imaging *direct* cell–cell contacts by anchoring the Gluc fragments to distinct cell membranes or antibody targets. Collectively, such tools will provide a unique vantage point for visualizing cell–cell communication and, as such, improve our mechanistic understanding of biology and disease.

## **2.6 Methods and materials**

### **2.6a Split Gluc constructs**

*Construction of secreted Jun/Fos Gaussia luciferase fusion vectors.* Previously described N-terminal (N<sub>Gluc</sub>) and C-terminal (C<sub>Gluc</sub>) fragments of Gluc (24) were fused to Jun(Cys) or Fos through a (G<sub>4</sub>S)<sub>2</sub> linker and were inserted into the pBMNpuro vector (courtesy of Nolan laboratory, Stanford). Overlap PCR was used to generate Jun-N<sub>Gluc</sub>.

Cysteine to serine mutations were generated by overlap PCR with the following primers:

5'-cctcgatcctccctttatc-3' and 5'-ctgccgccgctatgcatcttg-3' for PCR 1

5'-caagatgcatagcggcggcag-3' and 5'-gcggaatttagctagcgg-3' for PCR2

*Construction of secreted controls.* Secreted, unfused N<sub>Gluc</sub> and C<sub>Gluc</sub> were prepared by overlap PCR to delete either Jun or Fos from the reporter constructs.

Fos and Jun deletions were generated by deletion overlap PCR with the following primers:

5'-cctcgatcctccctttatc-3' and 5'-caccacgcgctatgcatcttg-3' for PCR 1

5'-caagatgcatagcgcgtggtg-3' and 5'-gcggaatttagctagcgg-3' for PCR2

## **2.6b Cell culture**

HEK293 cells (American Type Cell Culture), DB7 cells (courtesy of Contag laboratory, Stanford), and 4T1-luc2 cells (Caliper LifeSciences, now Perkin-Elmer) were cultured in DMEM (Corning) supplemented with 10% (vol/vol) fetal bovine serum (FBS, Life Technologies), penicillin (100 U/mL), and streptomycin (100 µg/mL). Jurkat cells were cultured in RPMI media (Corning) supplemented with 10% (v/v) FBS, penicillin (100 U/mL), and streptomycin (100 µg/mL). Cells were maintained in a 5% CO<sub>2</sub>, water-saturated incubator at 37 °C. Transient

transfections of the reporter (Jun-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub>) and control (N<sub>Gluc</sub> and C<sub>Gluc</sub>) constructs were performed with cationic lipids (Lipofectamine 2000; Invitrogen). DB7 cells stably expressing the fusion protein were selected with puromycin (10 µg/mL; Corning). HEK293 cells were transduced with amphotropic retrovirus (Phoenix packaging system) as previously described,(54) followed by selection with puromycin (10 µg/mL)

### **2.6c Flow cytometry**

Cells transiently or stably expressing the Gluc fragments were trypsinized and washed in FACS buffer (1% BSA, PBS). Cells were fixed with a freshly prepared mixture of methanol:acetone (1:1) treating for 1 min at room temperature for total expression levels. Cells were then washed with FACS buffer (3 x 200 µL) and incubated with monoclonal α-FLAG M2-Cy3 antibody (Sigma Aldrich) at a final concentration of 10 µg/mL on ice for 1 h. Cells were then pelleted and washed (3 x 200 µL) prior to analysis on a BD Biosciences LSRII. For each sample, 10,000 cells were analyzed and cellular fluorescence data was analyzed using FloJo software (Tree Star, Inc.).

### **2.6d Cellular co-culture assays**

*Gluc complementation assay.* Equal numbers of cells expressing Jun(Cys)-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub>, Jun-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub> or N<sub>Gluc</sub> and C<sub>Gluc</sub> were added to round bottom 96-well plates (5 x 10<sup>4</sup> cells per well). Media harvested from each sample at 24 and 48 h post-plating was transferred to black-wall 96-well plates (10 µL/well) for bioluminescence imaging. Native coelenterazine (Nanolight Technology) was reconstituted as a stock solution of 5 mg/mL in ethanol and diluted in DMEM (1000-fold dilution, 3 µL/well) and bioluminescent emission was measured using an

IVIS Lumina II (Xenogen) instrument. Coelenterazine was added to N<sub>Gluc</sub> and C<sub>Gluc</sub> samples and imaged prior to addition to Jun-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub> for imaging.

*Gluc stability assay.* Equal numbers of cells expressing Jun-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub> or N<sub>Gluc</sub> and C<sub>Gluc</sub> were added to round bottom 96-well plates ( $5 \times 10^4$  cells per well). Media was harvested 24 h after plating and transferred to a 1.5 mL centrifuge tube or to a 96-well plate containing HEK293 cells ( $5 \times 10^3$  cells per well). Media was imaged daily as above. In some cases, the media samples were combined with FBS (1:1 v/v) and imaged hourly.

*Time-dependent induction assay.* Reporter and control cells were added to round bottom 96-well plates ( $5 \times 10^4$  cells per well) at a 1:1 ratio for 0-48 h at 37 °C. The media was transferred to a black-wall 96-well plate and imaged as above.

*Cell dose-response assay.* Reporter and control cells were added to a round bottom 96-well plate ( $3.125 \times 10^3$ - $1 \times 10^5$  cells per well) at a 1:1 ratio for 24 h at 37 °C. Media (10 µL/well) was transferred to a black-wall 96-well plate and imaged as above.

*Cell ratio assay.* Cells expressing Jun-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub> were added to 24-well plates at various ratios ( $2 \times 10^5$  total cells per well). Media harvested from each sample at 24 h post-plating was transferred to black-wall 96-well plates for bioluminescence imaging as above.

## **2.6e Distance dependency of Gluc assembly**

Stencils were cut from 250 µm thick Polydimethylsiloxane (PDMS) membranes (Rogers Corp.) using a carbon dioxide laser (VersaLaser VLS-2.30, Universal Laser Systems Inc.) to provide

two chambers separated by distances of 3, 5, 7, 10, 20, and 30 mm. Adhesion of PDMS membranes to the surface of tissue culture plastic is sufficient to block liquid flow and allow pools of media to be seeded in isolation in a single chamber. Stencils were rinsed in 70% ethanol and seeded in tissue culture dishes (12-well plates or 6 cm<sup>2</sup> dishes) for 24 h. Cells were then plated ( $5 \times 10^4$  or  $1.5 \times 10^5$  cells per stencil chamber) and allowed to adhere (24 h) before the media was removed. Fresh DMEM (1 or 3 mL) was added over the entire surface area. Media aliquots (100  $\mu$ L) were then harvested after 24-72 h and analyzed for Gluc activity as above.

### **2.6f *In vitro* migration model**

Stencils cut to 3 and 7 mm were seeded in 6 cm<sup>2</sup> tissue cultures dishes as above. For 30 mm distances, two 7 mm stencils were seeded at the edges of the 6 cm<sup>2</sup> tissue culture dishes. Cells were then plated ( $2.5 \times 10^4$  cells per stencil chamber) and allowed to adhere (24 h) before the media was removed. Fresh DMEM (2 mL) was added over the entire surface area. Every 24 h the cells and media were imaged as above. After imaging, the media was removed and new DMEM (2 mL) was added. Bright-field micrographs of the plates were acquired on a Nikon Eclipse TE200 microscope through a 4x objective. Images were collected using the open source microscopy software Micro-Manager (<http://valelab.ucsf.edu/~MM/MMwiki/>) with a QIClick CCD camera (Qimaging). Images were analyzed using ImageJ software (National Institutes of Health) and Adobe Photoshop CS5.

### **2.6g Statistical analyses**

Data were analyzed using unpaired t-tests (GraphPad Prism 4).

## References

1. Germain RN, Robey EA, & Cahalan MD (2012) A decade of imaging cellular motility and interaction dynamics in the immune system. *Science* 336(6089):1676-1681
2. Roh-Johnson M, *et al.* (2013) Macrophage contact induces RhoA GTPase signaling to trigger tumor cell intravasation. *Oncogene* 33(33):4203-4212
3. Felding-Habermann B, *et al.* (2001) Integrin activation controls metastasis in human breast cancer. *Proc. Natl. Acad. Sci. U. S. A.* 98(4):1853-1858
4. Beerling E, Ritsma L, Vrisekoop N, Derksen PW, & van Rheenen J (2011) Intravital microscopy: new insights into metastasis of tumors. *J. Cell Sci.* 124(Pt 3):299-310
5. Paley MA & Prescher JA (2014) Bioluminescence: a versatile technique for imaging cellular and molecular features. *MedChemComm* 5(3):255-267
6. Prescher JA & Contag CH (2010) Guided by the light: visualizing biomolecular processes in living animals with bioluminescence. *Curr. Opin. Chem. Biol.* 14(1):80-89
7. Weissleder R & Pittet MJ (2008) Imaging in the era of molecular oncology. *Nature* 452(7187):580-589
8. Mishra P, Banerjee D, & Ben-Baruch A (2011) Chemokines at the crossroads of tumor-fibroblast interactions that promote malignancy. *J. Leukoc. Biol.* 89(1):31-39
9. Franco OE, Shaw AK, Strand DW, & Hayward SW (2010) Cancer associated fibroblasts in cancer pathogenesis. *Semin. Cell Dev. Biol.* 21(1):33-39
10. Alexander S, Weigelin B, Winkler F, & Friedl P (2013) Preclinical intravital microscopy of the tumour-stroma interface: invasion, metastasis, and therapy response. *Curr. Opin. Cell Biol.* 25(5):659-671
11. Bonapace L, *et al.* (2012) If you don't look, you won't see: intravital multiphoton imaging of primary and metastatic breast cancer. *J. Mammary Gland Biol. Neoplasia* 17(2):125-129
12. Chaffer CL & Weinberg RA (2011) A perspective on cancer cell metastasis. *Science* 331(6024):1559-1564
13. Conway JR, Carragher NO, & Timpson P (2014) Developments in preclinical cancer imaging: innovating the discovery of therapeutics. *Nat. Rev. Cancer* 14(5):314-328
14. Kalos M, *et al.* (2011) T cells with chimeric antigen receptors have potent antitumor effects and can establish memory in patients with advanced leukemia. *Sci. Transl. Med.* 3(95):95ra73

15. Harrison SC (2008) Viral membrane fusion. *Nat. Struct. Mol. Biol.* 15(7):690-698
16. Gheysens O & Gambhir SS (2005) Studying molecular and cellular processes in the intact organism. *Prog. Drug Res.* 62:117-150
17. Contag CH (2007) *In vivo* pathology: seeing with molecular specificity and cellular resolution in the living body. *Annu. Rev. Pathol.* 2:277-305
18. Dothager RS, *et al.* (2009) Advances in bioluminescence imaging of live animal models. *Curr. Opin. Biotechnol.* 20(1):45-53
19. Sellmyer MA, *et al.* (2013) Visualizing cellular interactions with a generalized proximity reporter. *Proc. Natl. Acad. Sci. U. S. A.* 110(21):8567-8572
20. Michnick SW (2003) Protein fragment complementation strategies for biochemical network mapping. *Curr. Opin. Biotechnol.* 14(6):610-617
21. Villalobos V, Naik S, & Piwnica-Worms D (2007) Current state of imaging protein-protein interactions *in vivo* with genetically encoded reporters. *Annu. Rev. Biomed. Eng.* 9:321-349
22. Luker KE, *et al.* (2012) *In vivo* imaging of ligand receptor binding with Gaussia luciferase complementation. *Nat. Med.* 18(1):172-177
23. Massoud TF, Paulmurugan R, & Gambhir SS (2010) A molecularly engineered split reporter for imaging protein-protein interactions with positron emission tomography. *Nat. Med.* 16(8):921-926
24. Remy I & Michnick SW (2006) A highly sensitive protein-protein interaction assay based on Gaussia luciferase. *Nat. Methods* 3(12):977-979
25. Badr CE & Tannous BA (2011) Bioluminescence imaging: progress and applications. *Trends Biotechnol.* 29(12):624-633
26. Tannous BA (2009) Gaussia luciferase reporter assay for monitoring biological processes in culture and *in vivo*. *Nat. Protoc.* 4(4):582-591
27. Tannous BA, Kim DE, Fernandez JL, Weissleder R, & Breakefield XO (2005) Codon-optimized Gaussia luciferase cDNA for mammalian gene expression in culture and *in vivo*. *Mol. Ther.* 11(3):435-443
28. Venisnik KM, Olafsen T, Gambhir SS, & Wu AM (2007) Fusion of Gaussia luciferase to an engineered anti-carcinoembryonic antigen (CEA) antibody for *in vivo* optical imaging. *Mol. Imaging Biol.* 9(5):267-277

29. Patel KG, *et al.* (2009) Cell-free production of *Gaussia princeps* luciferase--antibody fragment bioconjugates for *ex vivo* detection of tumor cells. *Biochem. Biophys. Res. Commun.* 390(3):971-976
30. van Rijn S, Wurdinger T, & Nilsson J (2014) Multiplex functional bioluminescent reporters using *Gaussia* luciferase fused to epitope tags in an immunobinding assay. *Methods Mol. Biol.* 1098:231-247
31. Santos EB, *et al.* (2009) Sensitive *in vivo* imaging of T cells using a membrane-bound *Gaussia princeps* luciferase. *Nat. Med.* 15(3):338-344
32. Kim SB, Sato M, & Tao H (2009) Split *Gaussia* luciferase-based bioluminescence template for tracing protein dynamics in living cells. *Anal. Chem.* 81(1):67-74
33. Ray P, *et al.* (2012) Secreted CXCL12 (SDF-1) forms dimers under physiological conditions. *Biochem. J.* 442(2):433-442
34. Wille T, Blank K, Schmidt C, Vogt V, & Gerlach RG (2012) *Gaussia princeps* luciferase as a reporter for transcriptional activity, protein secretion, and protein-protein interactions in *Salmonella enterica serovar typhimurium*. *Appl. Environ. Microbiol.* 78(1):250-257
35. Muller M, Cassonnet P, Favre M, Jacob Y, & Demeret C (2013) A comparative approach to characterize the landscape of host-pathogen protein-protein interactions. *J. Vis. Exp.* (77):e50404
36. Hashimoto T, Adams KW, Fan Z, McLean PJ, & Hyman BT (2011) Characterization of oligomer formation of amyloid-beta peptide using a split-luciferase complementation assay. *J. Biol. Chem.* 286(31):27081-27091
37. Hu CD & Kerppola TK (2003) Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat. Biotechnol.* 21(5):539-545
38. Shyu YJ, Suarez CD, & Hu CD (2008) Visualization of AP-1 NF-kappaB ternary complexes in living cells by using a BiFC-based FRET. *Proc. Natl. Acad. Sci. U. S. A.* 105(1):151-156
39. Slavoff SA, Liu DS, Cohen JD, & Ting AY (2011) Imaging protein-protein interactions inside living cells via interaction-dependent fluorophore ligation. *J. Am. Chem. Soc.* 133(49):19769-19776
40. Kaplan JB, Reinke AW, & Keating AE (2014) Increasing the affinity of selective bZIP-binding peptides through surface residue redesign. *Protein Sci.* 23(7):940-953
41. Kohler JJ & Schepartz A (2001) Kinetic studies of Fos.Jun.DNA complex formation: DNA binding prior to dimerization. *Biochemistry* 40(1):130-142



42. Landschulz WH, Johnson PF, & McKnight SL (1988) The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* 240(4860):1759-1764
43. Wurdinger T, *et al.* (2008) A secreted luciferase for *ex vivo* monitoring of *in vivo* processes. *Nat. Methods* 5(2):171-173
44. Friedl P, den Boer AT, & Gunzer M (2005) Tuning immune responses: diversity and adaptation of the immunological synapse. *Nat. Rev. Immunol.* 5(7):532-545
45. Friedl P & Storim J (2004) Diversity in immune-cell interactions: states and functions of the immunological synapse. *Trends Cell Biol.* 14(10):557-567
46. Gupta GP & Massague J (2006) Cancer metastasis: building a framework. *Cell* 127(4):679-695
47. Helming L, *et al.* (2008) Essential role of DAP12 signaling in macrophage programming into a fusion-competent state. *Sci. Signal* 1(43):ra11
48. Shibue T & Weinberg RA (2011) Metastatic colonization: settlement, adaptation and propagation of tumor cells in a foreign tissue environment. *Semin. Cancer Biol.* 21(2):99-106
49. Wetzel SA, McKeithan TW, & Parker DC (2002) Live-cell dynamics and the role of costimulation in immunological synapse formation. *J. Immunol.* 169(11):6092-6101
50. Kane RS, Takayama S, Ostuni E, Ingber DE, & Whitesides GM (1999) Patterning proteins and cells using soft lithography. *Biomaterials* 20(23-24):2363-2376
51. Marshall J (2011) Transwell((R)) invasion assays. *Methods Mol. Biol.* 769:97-110
52. Chen EH, Grote E, Mohler W, & Vignery A (2007) Cell-cell fusion. *FEBS Lett.* 581(11):2181-2193
53. Wickner W & Schekman R (2008) Membrane fusion. *Nat. Struct. Mol. Biol.* 15(7):658-664.
54. Helms MW, Prescher JA, Cao YA, Schaffert S, & Contag CH (2010) IL-12 enhances efficacy and shortens enrichment time in cytokine-induced killer cell immunotherapy. *Cancer Immunol. Immunother.* 59(9):1325-1334

# CHAPTER 3: Generalizable split reporters for improved visualization of cell-cell interactions

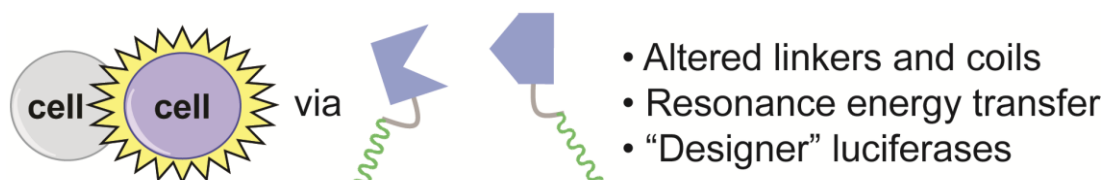
## 3.1 Introduction

As described in chapter 2, we developed an approach to image cell-cell interactions utilizing “split” fragments of a luciferase reporter. When we fused known dimerization domains (i.e., Jun and Fos) to the Gluc fragments robust complementation was observed. Light emission was also detected when cells engineered to secrete either Jun- or Fos-labeled fusions were cultured together.

While our previous studies provided important proof-of-principle for detecting cell-cell interactions, the light emission values remained low. The light emission observed from Fos/Jun-complemented Gluc was ~100-fold below that of the native enzyme. Similar reductions in light emission were observed with Gluc fragments used to monitor intracellular protein-protein interactions (1). While sufficient for imaging cell-cell interactions in excised tissues and other complex environments, such photon outputs would likely require at least  $2.5 \times 10^5$  total interacting cells to produce detectable signal at surface sites *in vivo* (2, 3). Even a 10-fold improvement in total light output could allow for the detection of deep tissue signals, such as tumor-stromal cell interactions in lung or liver models (4, 5).

We thus set out to investigate whether modifications to the split reporter could improve fragment assembly and boost overall photon production. In parallel, we investigated whether interfacing the split reporters to designer luciferases or fluorescent proteins could similarly increase photon counts (Figure 3-1). These latter fusions would also have the potential to red-

shift light emission via resonance energy transfer, making the tools more desirable for applications in whole tissues and animals. The modularity and generalizability of split reporters would also have important ramifications for visualizing a broad range of cellular interactions.

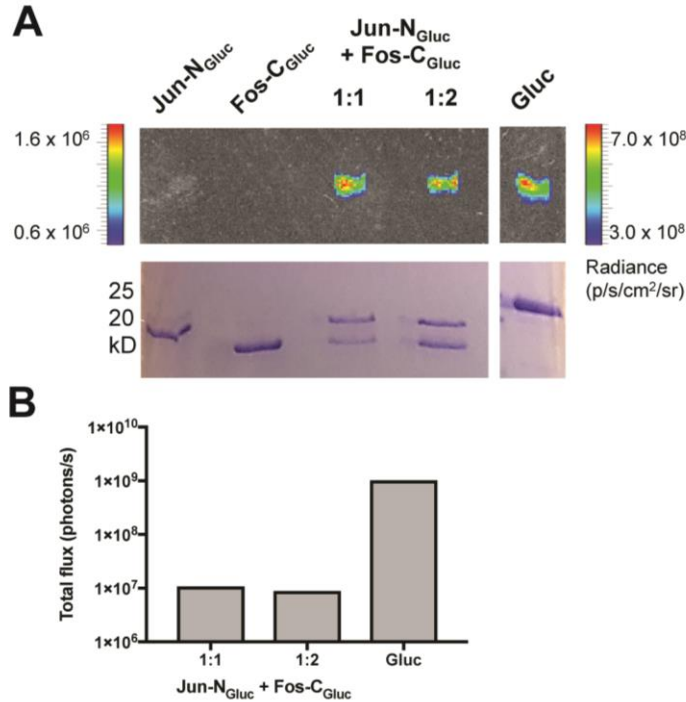


**Figure 3-1.** Imaging cell-cell interactions with split luciferase probes. Strategies for optimizing luciferase complementation were explored, including linker lengths, coil affinity, and local avidity. The split reporters were also interfaced with fluorescent proteins and “designer” luciferases.

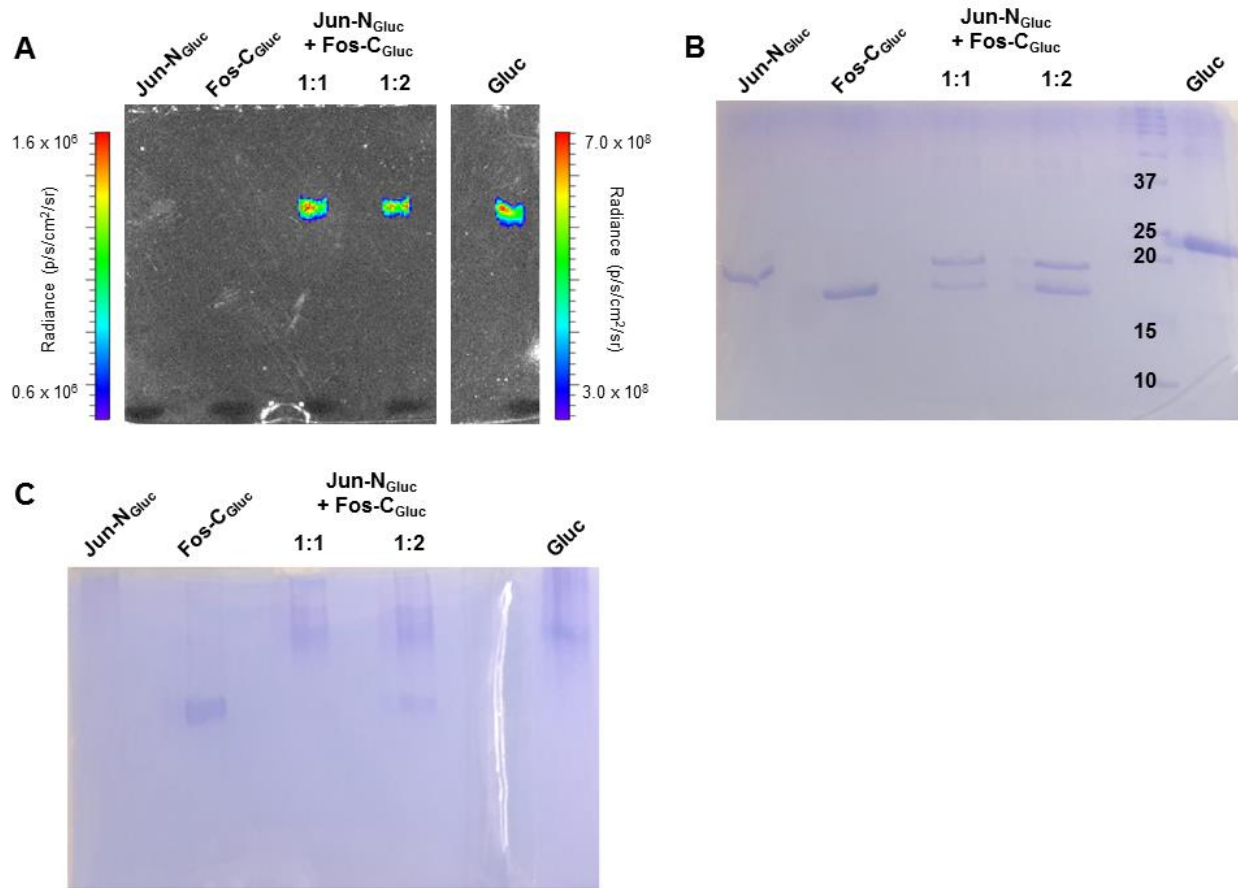
### 3.2 Fragment assembly is robust but underperforms native Gluc

To quantify the efficiency of photon outputs from the Jun/Fos-split Gluc pairs, we first investigated the assembly of purified fragments. The original constructs comprised Gluc fragments (cut between Gly-93 and Gluc-94) from a previously reported optimal cut site (1). Jun and Fos were fused to N- and C-terminal Gluc fragments, respectively, separated by two G<sub>4</sub>S repeats (Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub> and Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub>). We expressed and purified Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub> and Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub> from *E. coli* cells. When mixed, the purified fragments complemented and had strong enough affinities to survive electrophoresis as evidenced by light emission in acrylamide gels (Figure 3-2A, top and 3-3). Increased amounts of Fos did not increase light production, suggesting that full complementation was achieved at equimolar amounts of the split reporter. As expected, though, complementation was disrupted under reducing conditions (Figure 3-2A, bottom). Quantitative analysis of the light output suggested that the complemented enzyme possessed ~1% of the activity of an equimolar amount of holo Gluc (Figure 3-2B and 3-4). These

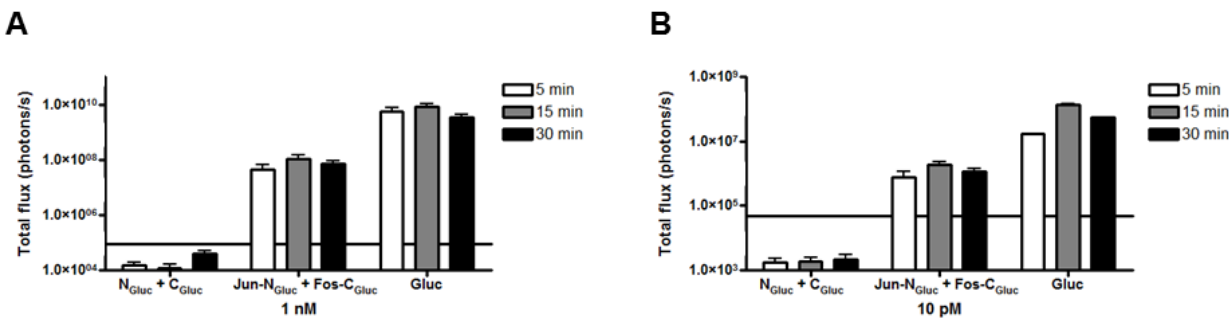
results suggested that fragment assembly remained suboptimal, and perhaps improved light outputs could be obtained with alternate linkers or luciferase domains.



**Figure 3-2.** Gluc complementation is robust and detectable in gel. (A) Native (top) and SDS (bottom) gel of Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>, Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub>, mixed (1:1 and 1:2) split reporters, and Gluc. Native gel was rinsed and soaked in coelenterazine (1 μg/mL) prior to imaging. (B) Quantification of bioluminescent signal from native gel in A. Gels and quantification are representative of n ≥ 4 experiments.



**Figure 3-3.** Whole gel analysis of Gluc complementation. (A) Native and (B) SDS gel of recombinant Jun-(G4S)<sub>2</sub>-N<sub>Gluc</sub>, Fos-(G4S)<sub>2</sub>-C<sub>Gluc</sub>, mixed (1:1 and 1:2) split reporters, and Gluc. Native gel was rinsed and soaked in coelenterazine (1  $\mu$ g/mL) prior to imaging. (C) After imaging, the native gel was rinsed and stained with Coomassie blue.



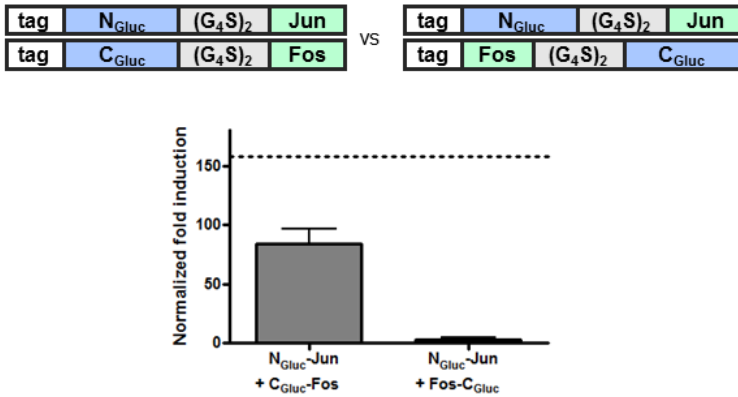
**Figure 3-4.** Complemented Gluc light output is decreased as compared to holo Gluc. (A) 1 nM or (B) 10 pM of mixed N<sub>Gluc</sub> + C<sub>Gluc</sub>, Jun-(G4S)<sub>2</sub>-N<sub>Gluc</sub> + Fos-(G4S)<sub>2</sub>-C<sub>Gluc</sub>, or Gluc were imaged over time (white 5 min, gray 15 min, black 30 min) post mixing. The solid line represents the average background for each split protein sample alone. For A-B, error bars represent the standard error of the mean of n = 3 experiments.

### 3.3 Orientation and linker length influence Gluc complementation

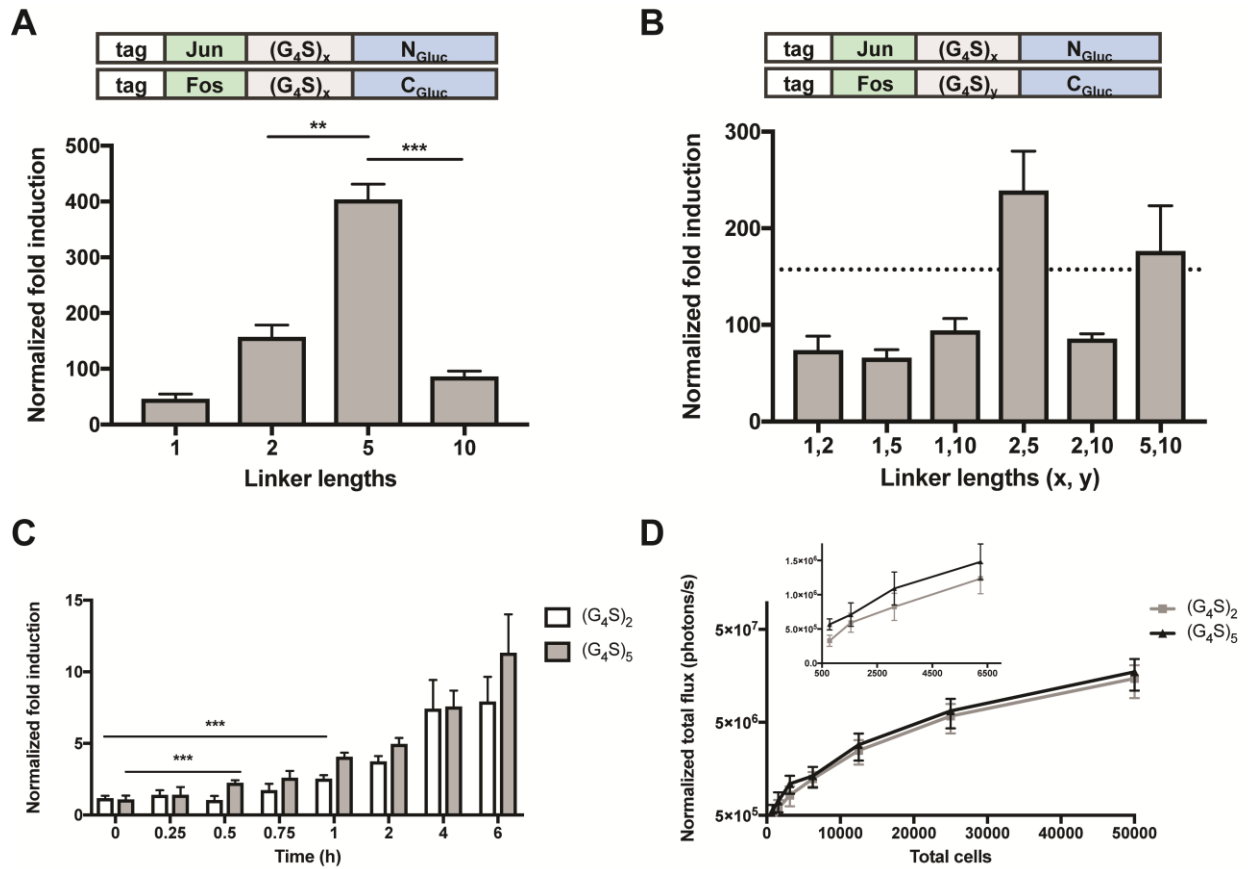
Our first attempts to improve photon output focused on positioning of the Gluc domains. It was possible that the two split fragments were not properly positioned to fully reassemble. To investigate alternative orientations, we fused each Gluc half to the N terminus of Jun and Fos ( $N_{\text{Gluc}}-(G_4S)_2\text{-Jun}$  and  $C_{\text{Gluc}}-(G_4S)_2\text{-Fos}$  and compared their activity to C-terminal fusions ( $\text{Jun}-(G_4S)_2\text{-N}_{\text{Gluc}}$  and  $\text{Fos}-(G_4S)_2\text{-C}_{\text{Gluc}}$ ). The relevant constructs were transiently expressed in HEK293 cells and media was imaged after co-culturing. Fusions to the C terminus of Jun or Fos were found to have the greatest bioluminescence and were selected for further studies (Figure 3-5). We also investigated whether longer linkers could improve Jun-Fos driven Gluc assembly. Toward this end, we generated split reporters with varying numbers of  $G_4S$  repeats (1, 2, 5, or 10). These reporters were transiently expressed in HEK293 cells. Cells expressing fragments of matched or varying linker lengths were cultured together. Media samples were then assayed for complemented Gluc activity via bioluminescence imaging. While light emission was observed for all combinations tested, Gluc fragments comprising  $(G_4S)_5$  repeats provided the greatest levels of light emission (Figure 3-6A and 3-6B). Both shorter and longer linker lengths (1 and 10 repeats, respectively) resulted in decreased levels of light emission. These data suggest that the linker needs to be sufficiently long to enable the dimerization domains to assemble, but not so long as to be entropically disfavored.

The  $(G_4S)_5$  constructs further enabled fast and sensitive readouts on interacting cells. Fragments with  $(G_4S)_2$  or  $(G_4S)_5$  repeats were transiently expressed in HEK293 cells. Cells secreting complementary split probes (with matched linkers) were then cultured together, and media samples were assayed for Gluc activity. As shown in Figure 3-6C, bioluminescent signal was detected 30 min post-plating for cells expressing reporters with the  $(G_4S)_5$  linker. Compared

to the first-generation linker, these probes enabled faster readouts on interacting cells. Additionally, the longer linkers improved overall sensitivity. A smaller number of total cells could also be detected with the (G<sub>4</sub>S)<sub>5</sub> probes compared to the (G<sub>4</sub>S)<sub>2</sub> set (Figure 3-6D). Collectively, these data suggest optimal linker lengths can enable more sensitive detection of cell-cell interactions.



**Figure 3-5.** Gluc fragments fused to the carboxy terminus of Jun or Fos are the optimal orientation for light emission. Constructs were transiently expressed in HEK293 cells and cells were mixed (1:1, 5 x 10<sup>4</sup> total cells) as indicated. Supernatants were collected 24 h post-plating and imaged with coelenterazine. Bioluminescent signal was normalized to expression levels as determined by flow cytometry. The normalized fold inductions in bioluminescent signal for each condition versus N<sub>Gluc</sub>- and C<sub>Gluc</sub>-secretin cells (control cells) are plotted. Dotted line represents average normalized fold induction for mixed Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub> and Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub> secreting cells. Error bars represent the standard error of the mean for n > 3 experiments.



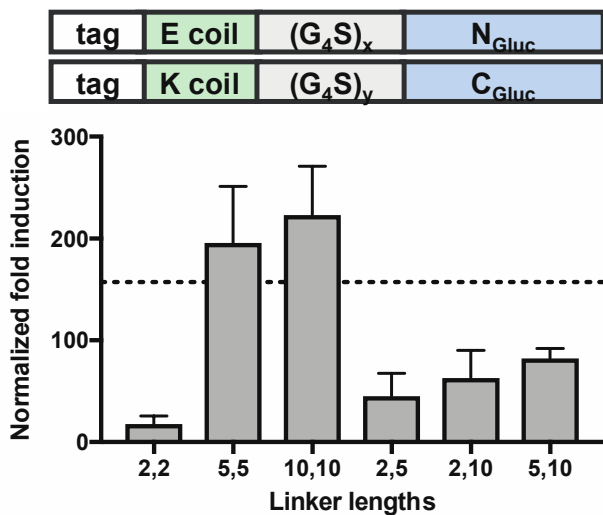
**Figure 3-6.** Linker length influences Gluc complementation. (A) Jun and Fos tethered to Gluc fragments via G<sub>4</sub>S linker repeats (Jun-(G<sub>4</sub>S)<sub>x</sub>-N<sub>Gluc</sub> or Fos-(G<sub>4</sub>S)<sub>x</sub>-C<sub>Gluc</sub>) or (B) varying linker lengths (Jun-(G<sub>4</sub>S)<sub>x</sub>-N<sub>Gluc</sub> or Fos-(G<sub>4</sub>S)<sub>y</sub>-C<sub>Gluc</sub>) were transiently expressed in HEK293 cells. Cells were mixed in 96-well round bottom plates (1:1, 5 × 10<sup>4</sup> total cells). Supernatants were collected from the cocultures 24 h post-plating and imaged with coelenterazine. (C) HEK293 cells transiently expressing Gluc fragments comprising (G<sub>4</sub>S)<sub>2</sub> or (G<sub>4</sub>S)<sub>5</sub> repeats were plated as above. Supernatants were collected 0-6 h postplating and incubated with coelenterazine prior to imaging. (D) HEK293 cells transiently secreting Gluc fragments with (G<sub>4</sub>S)<sub>2</sub> or (G<sub>4</sub>S)<sub>5</sub> were plated together (1:1, 775-50000 cells total) in round bottom plates. After 24 h, supernatants from the cocultures were collected and imaged with coelenterazine. The normalized fold inductions (A-C) or total flux (D) in bioluminescent signal for each condition versus control mixtures (N<sub>Gluc</sub>:C<sub>Gluc</sub>) are plotted. For B, the dotted line represents the average normalized fold induction for mixed Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub> and Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub> secreting cells. For A-D, error bars represent the standard error of the mean for n > 3 experiments, \*\*p < 0.001, \*\*\*p < 0.0001.

### 3.4 Alternative dimerization domain motifs and variable repeats

Beyond the length of the interacting domains, we reasoned that alterations to the coils themselves could be advantageous for examining transient cell-cell interactions. Several engineered domains have been reported to have faster assembly rates, including E and K coils (6-

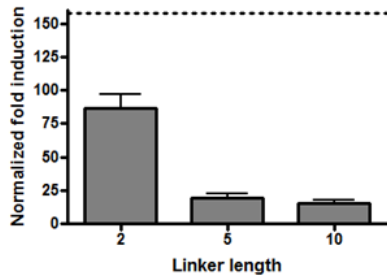
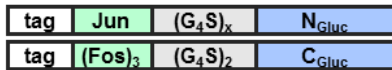
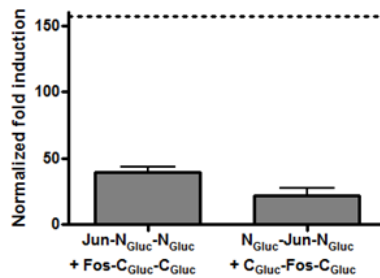
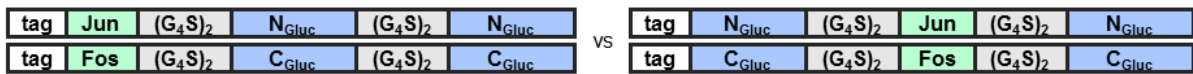


8). The coil motifs comprise a valine and leucine hydrophobic core to prevent homodimerization and promote the formation of heterodimeric coiled-coils. Because of these features, E and K coil form a ~1000-fold more stable complex ( $K_d \sim 60$  pM) than Jun and Fos. To investigate their utility for Gluc complementation, E/K coil coils were fused to  $N_{\text{Gluc}}$  or  $C_{\text{Gluc}}$  (using various linker lengths), and the resulting constructs were expressed in HEK293 cells. After 24 h, media samples were collected and imaged. Surprisingly, photon outputs from E/K-coil split reporters separated by  $(G_4S)_2$  repeats were markedly reduced compared to the analogous Jun and Fos probes (Figure 3-7). This could be due to the slight change in overall length of the leucine zippers. Photon production could be recovered with longer linkers, though, supporting our previous findings that linker lengths are important for fragment assembly. However, the overall photon output was not significantly greater than the original dimerization domains.



**Figure 3-7.** Engineered coiled-coils enable Gluc assembly. E coil and K coil tethered to split Gluc fragments via  $G_4S$  linker repeats (E coil- $(G_4S)_x$ - $N_{\text{Gluc}}$ , K coil- $(G_4S)_y$ - $C_{\text{Gluc}}$ ) were transiently expressed in HEK293 cells. Cells were mixed (1:1,  $5 \times 10^4$  total cells) and supernatants were imaged with coelenterazine 24 h post-plating. The normalized fold inductions in bioluminescent signal for each condition versus control cells ( $N_{\text{Gluc}}:C_{\text{Gluc}}$ ) are plotted. The dotted line represents the average normalized fold induction for mixed Jun- $(G_4S)_2$ - $N_{\text{Gluc}}$  and Fos- $(G_4S)_2$ - $C_{\text{Gluc}}$  secreting cells. Error bars represent the standard error of the mean for  $n \geq 3$  experiments.

We next attempted to improve photon output by targeting repeats of the Fos domain. Complementation strategies routinely employ multiple dimerization domains to increase the avidity of the domains and ultimately drive split protein assembly (9, 10). For instance, multiple copies of FK506-binding protein have previously been utilized to increase the likelihood of dimerization with FKBP12-rapamycin-binding protein (11). Based on this precedent, we reasoned that constructs with multiple repeats of the Fos or Jun domains could increase the likelihood of dimerization. Multiple repeats of the Fos domain were transiently expressed in HEK293 cells and decreased light emission was observed upon media imaging (Figure 3-8A). Reduced photon output was also detected with reporters containing multiple Gluc fragments (Figure 3-8B). These results suggest additional domains could interfere with productive interactions for Gluc reassembly.

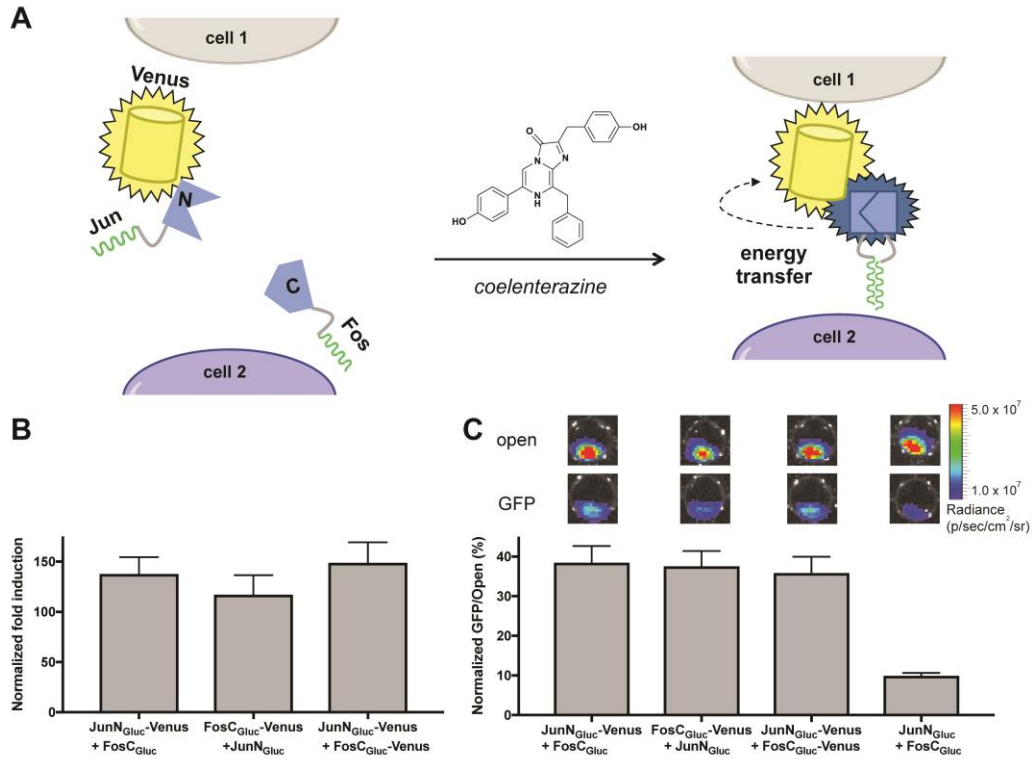
**A****B**

**Figure 3-8.** Variable repeats result in decreased light emission. (A) Multiple Fos domains or (B) Gluc fragment repeats were examined by transient expression in HEK293 cells. Cells were coplated (1:1,  $5 \times 10^4$  total cells) and supernatants were collected 24 h post-plating for imaging. For A-B, the normalized fold inductions in bioluminescent signal for each condition versus control cells (N<sub>Gluc</sub>:C<sub>Gluc</sub>) are plotted. The dotted line represents the average normalized fold induction for mixed Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub> and Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub> secreting cells. Error bars represent the standard error of the mean for  $n > 3$  experiments.

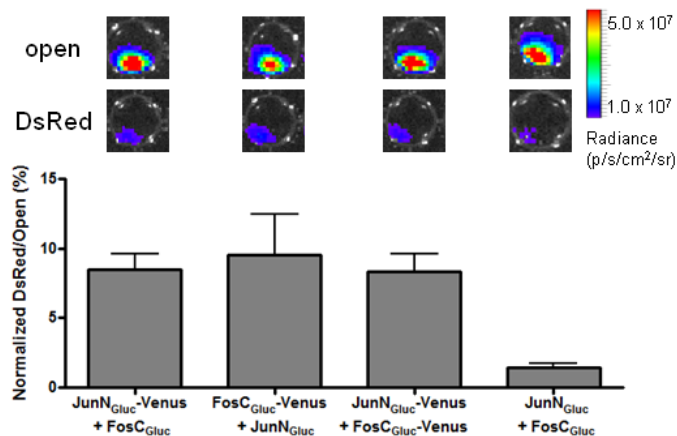
### 3.5 Analyzing split reporters linked to fluorescent proteins or “designer” luciferases

Improvements in split Gluc assembly were observed by systematically modifying our construct design. While fewer numbers of interacting cells could be detected, the total photon output was surprisingly only enhanced 5-fold. This suggested more dramatic design alternations were required to improve total photon output. To determine whether a more significant enhancement in light emission could be achieved, we interfaced the split reporters with designer luciferases or fluorescent proteins. One of the most well-established methods for improving

detection sensitivity in the context of whole cells and tissues is bioluminescence resonance energy transfer (BRET) (12-14). Red-shifted light can more easily penetrate tissues, making these wavelengths more desirable for *in vivo* imaging. By matching our split reporter with Venus fluorescent protein (VFP) as a BRET pair, we could red-shift emission by 48 nm (15-17). We generated the desired split Gluc-VFP fusions (Jun-N<sub>Gluc</sub>-Venus or Fos-C<sub>Gluc</sub>-Venus) and transiently expressed the constructs in HEK293 cells (Figure 3-9A). Cells secreting the fusions were cultured with cells secreting the complementary split reporter (Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub> or Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>), and media samples were collected. The aliquots were imaged and bioluminescent light was collected with and without filters to discern red-shifted emission. Without filters, all mixed cultures resulted in comparable light emission to our original split reporters indicating complementation is not affected by the fluorescent protein fusion (Figure 3-9B). With band-pass filters (515-575 nm or 575-650 nm), the greatest light emission (as compared to no filters) was observed for Venus fusions (Figure 3-9C and Figure 3-10). The fusions resulted in a 30% improvement as compared to the first-generation probes (Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub> and Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub>) when imaged with the 515-575 nm filter. This highlights the utility of linking split reporters with fluorescent proteins for sensitively detecting cell-cell interactions *in vivo*.



**Figure 3-9.** Fluorescent protein fusions facilitate red-shifted light emission upon Gluc complementation. (A) Diagram for BRET system involving split Gluc and Venus fluorescent protein to visualize cell-cell interactions. (B) Split Gluc-Venus fluorescent protein fusions, Jun-(G4S)<sub>2</sub>-N<sub>Gluc</sub>, and Fos-(G4S)<sub>2</sub>-C<sub>Gluc</sub> were transiently expressed in HEK293 cells and were mixed (1:1,  $5 \times 10^4$  total cells) as indicated. After 24 h, supernatants were collected from the cocultures and imaged with coelenterazine. Light emission was detected with no filter or (C) GFP bandpass (515-575 nm) filter. The normalized fold inductions in bioluminescent signal for each condition versus control cells (N<sub>Gluc</sub>:C<sub>Gluc</sub>) are plotted. For B and C, error bars represent the standard error of the mean for  $n = 3$  experiments.

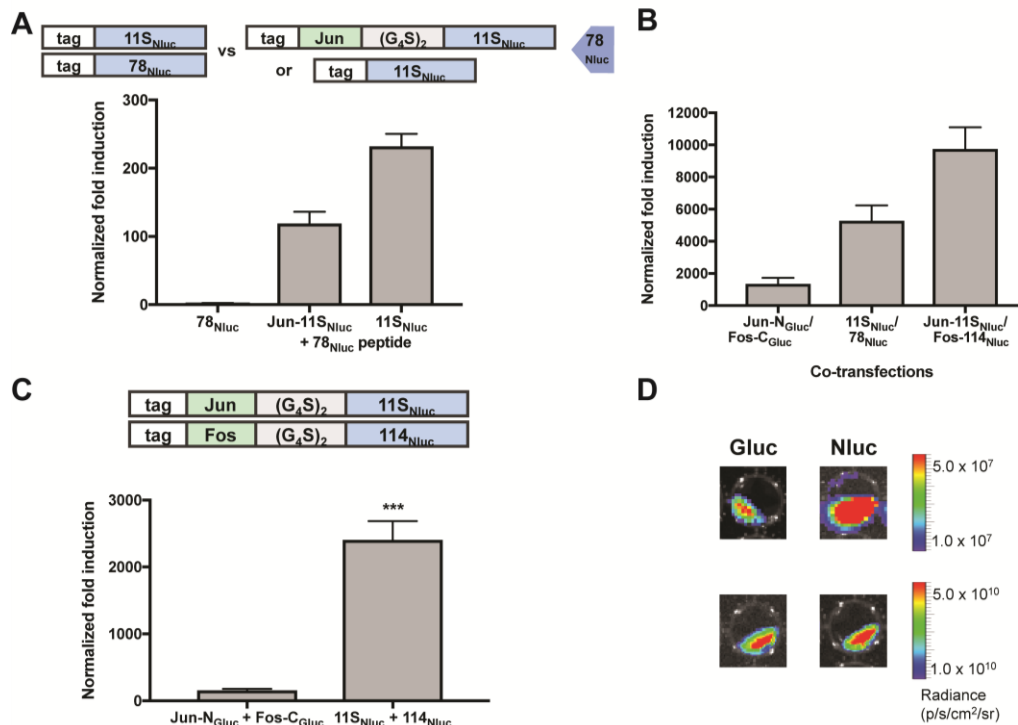


**Figure 3-10.** Venus fusions can be detected with 575-650 nm bandpass filter. Split Gluc-Venus fluorescent protein fusions, Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>, and Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub> were transiently expressed in HEK293 cells and were mixed (1:1, 5 × 10<sup>4</sup> total cells) as indicated. After 24 h, supernatants were collected from the cocultures and imaged with coelenterazine. Light emission was detected with no filter or DsRed bandpass (575-650 nm) filter. The normalized fold inductions in bioluminescent signal for each condition versus control cells (N<sub>Gluc</sub>:C<sub>Gluc</sub>) are plotted. Error bars represent the standard error of the mean for n = 3 experiments.

In addition to BRET pairs, we hypothesized designer luciferases could afford enhanced light emission. One such luciferase is NanoLuc (Nluc) which was engineered for both bright and sustained luminescence (18). Important to our approach, optimized split versions of Nluc (NanoBiT) with varying affinities were recently developed (19). NanoBiT fusions to FRB/FKBP expressed in cells could accurately report on complex association over a range of reporter concentrations. To take advantage of these properties for visualizing cell-cell interactions, we designed two sets of split reporters: one comprising Nluc fragments with low affinity ( $K_D$  190  $\mu$ M) fused to Jun or Fos (Jun-11S<sub>Nluc</sub> and Fos-114<sub>Nluc</sub>), and one comprising high affinity ( $K_D$  3.4 nM) fragments (11S<sub>Nluc</sub> and 78<sub>Nluc</sub>). HEK293 cells transiently expressing the split Nluc probes were cultured together and media samples were assayed for bioluminescence. No signal was observed in media samples containing the high affinity fragments (Figure 3-11A). We reasoned signal was not detectable due to inefficient secretion of 78<sub>Nluc</sub>. Photon production was restored when the synthetic peptide of 78<sub>Nluc</sub> (synthesized by Zach Reinert, Prescher lab) was added to

media from cells secreting Jun-11S<sub>Nluc</sub> or 11S<sub>Nluc</sub>. This was further confirmed by transiently co-expressing 11S<sub>Nluc</sub> and 78<sub>Nluc</sub>. Imaging of supernatant from these cells showed significant signal over background (Figure 3-11B).

Robust light emission was observed in samples containing Jun-11S<sub>Nluc</sub> and Fos-114<sub>Nluc</sub>, with a ~25-fold enhancement in bioluminescence over comparable split Gluc reporters (Figure 3-11C). This increased signal is likely due to more efficient assembly of designer Nluc fragments and not due to increased turnover of Nluc compared to Gluc. Indeed, when Gluc or Nluc expressing cells are cultured under the same conditions and normalized for protein expression levels, comparable total photon production is observed (Figure 3-11D). Regardless, such tools would be broadly useful for imaging cellular interactions relevant to cancer progression, immune function, and numerous other biological processes.



**Figure 3-11.** Nluc complementation reports on cell-cell interactions. (A) High affinity Nluc splits were transiently expressed in HEK293 cells and plated (1:1,  $5 \times 10^4$  total cells). Supernatants were collected from the cultures 24 h post-plating and imaged. 78<sub>Nluc</sub> peptide was also mixed with Jun-(G<sub>4</sub>S)<sub>2</sub>-11S<sub>Nluc</sub> or 11S<sub>Nluc</sub> supernatants and imaged. (B) Split Gluc or Nluc constructs were co-expressed in HEK293 cells and plated ( $5 \times 10^4$  total cells). Supernatants were imaged as above. (C) Jun and Fos tethered to split Nluc via matched (G<sub>4</sub>S)<sub>2</sub> linker repeats were transiently expressed in HEK293 cells. Cells were plated and media was imaged as above. (D) Representative bioluminescent images of split reporters and native enzyme. Gluc and Nluc were expressed in HEK293 cells and plated as above. Supernatants were collected 24 h post-plating and imaged. (For A-C, error bars represent the standard error of the mean for  $n = 3$  experiments, \*\*\* $p < 0.0001$ ).

### 3.6 Conclusions and future directions

In this work, we modified split bioluminescent reporters to sensitively image cell-cell interactions. Multiple design approaches were investigated in the context of our first-generation reporters. While modulating linker length resulted in increased photon outputs, the most significant enhancements were obtained with BRET or Nluc reporters. These improvements are likely to facilitate the imaging of transient cell contacts, such as those involved in the immune response. Additionally, the BRET or split Nluc reporters may aid in the development of *in vivo*



imaging models of interacting cells relevant to early metastatic events. Collectively, these toolsets can inform researchers “when and where” to look for cellular interactions and are likely to be broadly applicable to numerous aspects of biology and disease.

A majority of the design modifications we have described were amenable to monitoring interacting cells. We are currently investigating this generalizability to develop split probes based on Fluc. The red-shifted emission of Fluc is beneficial for sensitive detection of cell-cell interactions *in vivo* but its ATP-dependency presents a challenge for extracellular use. We believe this requirement can be advantageous for visualizing cell proximity in disease models with enhanced extracellular ATP levels, such as cancer and graft-versus-host disease (20-22).

In addition, we envision utilizing recombinant protein sources of holo and split luciferases (as described in section 3.2) to exogenously label cells. These tools, much like fluorescent dyes, could be especially useful for monitoring primary cells and stem cells. General labeling strategies can also be implemented to generate luciferase-antibody conjugates. Current efforts in the lab are aimed at chemically appending a luciferase to affibodies to enable cell specific labeling *in vivo*.

### **3.7 Materials and methods**

#### **3.7a General cloning materials**

Split reporters were amplified from previously reported Gluc/Nluc fusions (19, 23) using DreamTaq DNA polymerase and reaction buffer (ThermoFisher) with 0.2 mM dNTPs (ThermoFisher). PCR reactions were analyzed using agarose gel electrophoresis with ethidium bromide staining. The amplified inserts were isolated and ultimately inserted into pBMN-iPuro

(courtesy of the Nolan laboratory, Stanford) or pcDNA3.1 using T4 DNA ligase (New England Biolabs) or circular polymerase extension cloning with Q5 polymerase (New England Biolabs).

### 3.7b Split reporter constructs

#### *pBMN and pcDNA expression vectors*

DNA sequences encoding flexible (G<sub>4</sub>S)<sub>x</sub> linkers were ligated to sequences encoding FLAG-Jun-N<sub>Gluc</sub> or FLAG-Fos-C<sub>Gluc</sub> in pBMN-iPuro (courtesy of the Nolan laboratory, Stanford). The linkers were constructed using PCR with the following primers:

5'-ggcggaggcgggtctgggggaggctctggtggcgggttc-3' and

5'-actgctctccaccgctccacccccgcgctaccgcccga-3' for PCR 1 to generate (G<sub>4</sub>S)<sub>6</sub>

5'-ataacgctggagggtggaggctcagggggcggaggcagcggcggaggcggg-3' and

5'-atactcgagactcccgccgctccagaacctccccccactgctctccac-3' for PCR 2 to generate (G<sub>4</sub>S)<sub>10</sub>

5'-atcgacgctggagggtggaggctcaggg-3' and 5'-atcgctcgagggaaccaccgcccaccaga-3' to generate (G<sub>4</sub>S)<sub>5</sub>

To generate a G<sub>4</sub>S linker the following primers were mixed 1:1, heated to 85 °C and cooled to room temperature in a heat block before ligating to sequences as above.

5'-ataacgctggtggcgggtggtccctcgagata-3'

5'-atactcgagggaaccaccgcccaccacgctata-3'

DNA sequences encoding E coil and K coil leucine zippers were ligated to sequences encoding FLAG-N<sub>Gluc</sub> or FLAG-C<sub>Gluc</sub> in pBMN-iPuro. The leucine zippers were constructed using sequential overlap PCR with the following primers:

5' gagcgccttgaaaaggaagtgagcgccttgaaaagacgctggtggcg-3' and

5'-gcggaatttagctagcgg-3' for PCR 1

5'-gaagtgagcgccttgaaaaggaagtgagcgccttgaaaag-3' and 5'-gcggaatttacgtagcgg-3' for PCR 2

5'-ggaaaaggaagtgagcgccttgaaaaggaagtgagcgccttg-3' and 5'-gcggaatttacgtagcgg-3' for PCR 3

5'-ataatgcatgaagtgagcgccttgaaaaggaagtgagcgc-3' and 5'-gcggaatttacgtagcgg-3' for PCR 4

Sequenced inserts from PCR 4 were then amplified with:

5'-gaggtgagcgccttcgagaagacgcgtggtggcg-3' and

5'-cttctcagggcgctcacctcctttccaaggcgctc-3' to ultimately generate E coil

5'-cgccttgaaggaaaaggtgagcgccttgaaggaaacgcgtggtggcg-3' and 5'-gcggaatttacgtagcgg-3' for

PCR 1

5'-ggtgagcgccttgaaggaaaaggtgagcgccttgaaggaaaag-3' and 5'-gcggaatttacgtagcgg-3' for PCR 2

5'-ggaaaaggtgagcgccttgaaggaaaaggtgagcgccttgaag-3' and 5'-gcggaatttacgtagcgg-3' for PCR 3

5'-ataatgataaggtgagcgccttgaaggaaaaggtgagcgc-3' and 5'-gcggaatttacgtagcgg-3' for PCR 4

Sequenced inserts from PCR 4 were then amplified with:

5'-aggtctccgctctgaaggaaaaggtgagcgcctc-3' and 5'-ttcctcagagcggagacctttccttcaaggcgctc-3' to

ultimately generate K coil

DNA sequences encoding Fos-(G<sub>4</sub>S)<sub>2</sub> were ligated to sequences encoding FLAG-C<sub>Gluc</sub>. The leucine zipper repeats were constructed using the following primers:

5'-tataatgcatggcggcctgaccg-3' and 5'-ataattaattaagctaccgccgccac-3'

5'-atattaattaatggcggcctgaccg-3' and 5'-atatggccggccgctaccgccgccac-3'

5'-ataggccggccaggcggcctgaccg-3' and 5'-atactcgaggctaccgccgccac-3'

DNA sequences encoding N<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>-Jun or C<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>-Fos were ligated to sequences encoding a FLAG tag in pBMN-iPuro. The altered orientation inserts were constructed using the following primers:

5'-ataatgcataagcccaccgagaacaacg-3' and 5'-atagcgccgcttagtagtcatcaccttctgcttcag-3'

5'-ataatgcatgaggcgatcgtcgagattc-3' and 5'-atagcgccgcttagtagggccaggatgaac-3'

DNA sequences encoding variable repeats of N<sub>Gluc</sub> or C<sub>Gluc</sub> were ligated to sequences encoding FLAG-Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub> or FLAG-Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub> in pBMN-iPuro. The N<sub>Gluc</sub> or C<sub>Gluc</sub> inserts were constructed using the following primers:

5'-atagccggcctaagcccaccgagaacaacg-3' and 5'-atactcgaggctaccgccgccacc-3'

5'-atagccggcctgaggcgatcgtcgacattc-3' and 5'-atactcgaggctaccgccgccacc-3'

5'-ataatgcataagcccaccgagaacaacg-3' and 5'-atatggccggccgctaccgccgccac-3'

5'-ataatgcatgaggcgatcgtcgagattc-3' and 5'-atatggccggccgctaccgccgccac-3'

DNA sequences encoding Venus fused to split Gluc reporters were ligated to sequences encoding GlucA-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>, GlucB-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub>, Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>, or Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub> in pcDNA. The Venus inserts were constructed using the following primers:

5'-atagaattcgtgagcaagggcgaggag-3' and 5'-ataggtacccccggcgggcgtc-3'

5'-ataggtaccgtgagcaagggcgaggag-3' and 5'-atatctagaattattactgtacagctcgtccatgcc-3'

DNA sequences encoding N<sub>Gluc</sub>, C<sub>Gluc</sub>, Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>, Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub>, N<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>-Jun, C<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>-Fos or Gluc were generated using the following primers:

5'-cttaagcttggcaatccggctactgttgtaaagccaccatggccttaccagtaccgcc-3'

5'-gagggggcgatccccgggctctagaattattagcctatccgccctgtg-3'

5'-gagggggcgatccccgggctctagaattattagtcaccaccggcccc-3'

5'-gagggggcgatccccgggctctagaattattagtagtcatcaccttctgcttcagctg-3'

5'-gaggggcgatcccgggctctagaattattagtagggccaggatgaactc-3'

The sequences were combined with linear pcDNA3.1-IRES-GFP via circular polymerase extension cloning (CPEC). The pcDNA3.1-IRES-GFP was linearized with the following primers:

5'-taataattctagagcccgggatccgccctg-3' and 5'-catggtggctttaccaacagtaccggattgccaagcttaag-3'

DNA sequences encoding (G<sub>4</sub>S)<sub>2</sub>-11S<sub>Nluc</sub> and (G<sub>4</sub>S)<sub>2</sub>-114<sub>Nluc</sub> were generated using the following primers:

5'-ggtggcggcggtagcctcgaggtcttcacactcgaagatttcgttggg-3' and

5'-gaggggcgatcccgggctctagaattattagctgttgatggttactcggaacagca-3'

5'-ggtggcggcggtagcctcgaggtgaccggctaccggctgttc-3' and

5'-gaggggcgatcccgggctctagaattattagagaatctcctcgaacagccggtag-3'

The sequences were combined with linear pcDNA3.1-IRES-GFP via CPEC. The pcDNA3.1-IRES-GFP encoding either Jun or Fos was linearized with the following primers:

5'-taataattctagagcccgggatccgccctg-3' and 5'-gctaccgccgccaccggaac-3'

DNA sequences encoding CD8 leader sequence-11S<sub>Nluc</sub> and CD8 leader sequence-78<sub>Nluc</sub> were generated using the following primers:

5'-cgccgccaggccggaattcgtcttcacactcgaagatttcgttggg-3' and

5'-gaggggcgatcccgggctctagaattattagctgttgatggttactcggaacagca-3'

5'-cgccgccaggccggaattcaatgtatcgggctggcgcttg-3' and

5'-gaggggcgatcccgggctctagaattattagttactaattttcttaacaagcgcca-3'

The sequences were combined with linear pcDNA3.1-IRES-GFP via CPEC. The pcDNA3.1-IRES-GFP encoding CD8 leader sequence was linearized with *EcoRI* and *XbaI* (New England Biolabs)

*pCold expression vectors*

DNA sequences encoding N<sub>Gluc</sub>, C<sub>Gluc</sub>, Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>, Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub>, and Gluc were amplified from *E. coli* codon-optimized Gluc (a gift from G.K. Walkup, Astra Zeneca). All inserts were generated by overlapping PCR to encode reporters with a C-terminal His<sub>6</sub> affinity tag using the following primers:

N<sub>Gluc</sub>

5'-ATTAAGAGGTAATACACCATGAATCACAAAGTGcatatgaaaccgaccgaaaacaacg-3'

5'-CCGGTTTCTGGAAGggatccGCCTTGAAAATAAAGATTTTCaccgataccacctgcgca-3'

5'-ctatctagaCTAGTGATGGTGATGGTGATGgaattcACCGCCGGTTTCTGGAAGggatcc-3'

5'-ctgtgctttaagcagagattacctatctagaCTAGTGATGGTGATGGTGATGg-3'

C<sub>Gluc</sub>

5' ATTAAGAGGTAATACACCATGAATCACAAAGTGcatatggaagcgatcgttgacatccg

5' CCGGTTTCTGGAAGggatccGCCTTGAAAATAAAGATTTTCgtcaccaccgcaccttg 3'

5' ctatctagaCTAGTGATGGTGATGGTGATGgaattcACCGCCGGTTTCTGGAAGggatcc 3'

5' ctgtgctttaagcagagattacctatctagaCTAGTGATGGTGATGGTGATGg 3'

Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>

From GlucA

5'-ctgtgcttttaagcagagattacctatctagaCTAGTGATGGTGATGGTGATGg-3'

5'-tgatgaactatggcggcggtggcagcggtagcaaacgaccgaaaacaacg-3'

5'-aacatgctgcgcgaacagggttgcgcagctgaaacagaaagtgatgaactatggcggcggt-3'

5'-aaacctgaaagcgcagaactctgaactggcgagcaccgccaacatgctgcgcgaacagg-3'

5'-ctggcgccgtattgcgcgctggaagaaaaagtgaaaccctgaaagcgcagaactctg-3'

5'-attaagagtaatacaccatgaatcacaaagtgcattgttctggcggccgtattgcg-3'

Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub>

5'-ctgtgcttttaagcagagattacctatctagaCTAGTGATGGTGATGGTGATGg-3'

5'-gccgcgtacggcggcggtggcagcggtagcgaagcgttgacatcccg-3'

5'-attgccaacctgctgaaagaaaaagaaaaactggaattatcctggccgctacggcggc-3'

5'-gctggaagataaaaaagcgcctgcagaccgaaattgccaacctgctgaaagaaaaaga-3'

5'-gcggtctgaccgacacctgcagcggaaaccgatcagctggaagataaaaaagcggc-3'

5'-attaagagtaatacaccatgaatcacaaagtgcattggttctgaccgacacct-3'

Fos-(G<sub>4</sub>S)<sub>2</sub>-LVPRGS-C<sub>Gluc</sub>

5' AGCCTGGTTCCGCGCGGTAGCgaagcgttgacatcccggaaatcccg 3'

5' ttGCTACCGCGCGGAACCAGGCTACCGCCGCCACCGCTGCC 3'

All sequences were combined with linear pCold via CPEC. The pCold vector was linearized with the following primers:

5' CACTTTGTGATTCATGGTGTATTACCTCTTAATAATTAAGTGTGC 3' and

5' TCTAGATAGGTAATCTCTGCTTAAAAGCACAGAATCTAAGATC 3'

### 3.7c Protein expression and purification

The plasmids were transformed into BL21-DE3 cells. A single colony was grown overnight in LB medium containing ampicillin (50  $\mu\text{g/ml}$ ) at 37  $^{\circ}\text{C}$ . This seed culture was transferred into a 1 L culture and incubated until the  $\text{OD}_{600}$  reached 0.6 – 0.7. The culture was chilled in an ice water bath for 1 h before induction with 1 mM IPTG. The induced culture was grown for 24 h at 15  $^{\circ}\text{C}$ . The bacteria were harvested by centrifugation and lysed in 50 mM Tris-HCl pH 8.7, 150 mM NaCl, 10 mM imidazole, 1% NP-40 by sonication. The lysate was purified by centrifugation and filtered before binding to nickel-affinity resin (BioRad). The column was washed with 50 mM Tris-HCl, pH 8.7, 150 mM NaCl, and 20 mM imidazole, and the protein was eluted with 50 mM Tris-HCl, pH 8.7, 150 mM NaCl, and 250 mM imidazole. The imidazole was removed by dialysis against 50 mM Tris-HCl, pH 8.7 and 150 mM NaCl. To obtain  $\text{C}_{\text{Gluc}}$ , FOS-(G<sub>4</sub>S)<sub>2</sub>-LVPRGS- $\text{C}_{\text{Gluc}}$  was expressed and purified as above and cleaved using Thrombin CleanCleave Kit (Sigma Aldrich). After cleavage, the  $\text{C}_{\text{Gluc}}$  fragment was re-purified as above.

### 3.8d Peptide synthesis

A lab member, Zach Reinert, synthesized the  $^{78}\text{Nluc}$  peptide.  $^{78}\text{Nluc}$  (H-NVSGWRLFKKISN-NH<sub>2</sub>) was synthesized on Nova PEG Rink Amide (0.05 mmol) resin via standard Fmoc solid-phase peptide synthesis techniques. Briefly, couplings were performed at room temperature for 30 min with 4 eq of protected Fmoc-amino acid and HCTU and 6 eq of DIPEA. Deprotections were carried out at room temperature with 2 mL of 20 % piperidine in DMF for 15 minutes. Resin was drained and washed three times with DMF after coupling and deprotection reactions. After deprotection of the final residue, the resin was washed three times with DMF, three times with DCM, and three times with MeOH before being dried under vacuum for 20 min. Cleavage of the peptide from resin was accomplished with 5 mL of cleavage cocktail (95% TFA, 2.5%



H<sub>2</sub>O, 2.5% triisopropylsilane) for 3 hours. The solution was filtered and 40 mL cold diethyl ether was added. After centrifugation, the ether was decanted and the peptide was dissolved in 5 mL H<sub>2</sub>O. After sonication for 30 min, the peptide was purified via RP-HPLC using gradients between 0.1% TFA in H<sub>2</sub>O and 0.1% TFA in acetonitrile. Peptide was isolated as a mixture of full-length peptide and a Phe-deletion mutant (~ 30:70). Peptide identities were confirmed by MALDI-TOF mass spectrometry.

### **3.7e Cell culture**

HEK293 (ATCC) cells were cultured in DMEM (Corning) supplemented with 10% (vol/vol) fetal bovine serum (FBS, Life Technologies), penicillin (100 U/mL), and streptomycin (100 µg/mL). Cells were maintained in a 5% CO<sub>2</sub>, water-saturated incubator at 37 °C. Transient transfections were performed using cationic lipid formulations (Lipofectamine 2000; Invitrogen). Cells were plated for imaging analysis or analyzed for expression 24 – 36 h post transfection.

### **3.7f Flow cytometry**

Split reporter expression was verified in transiently transfected cells via flow cytometry. For split reporters expressed from pBMN-iPuro constructs, cells were fixed with a freshly prepared solution of methanol:acetone (1:1, 1 min, on ice). Cells were washed with FACS buffer (PBS with 1% BSA) and incubated with FITC-α-FLAG M2 (Sigma Aldrich, 10 µg/mL) on ice for 1 h. Cells were washed (3 x 200 µL) prior to analysis on a BD Biosciences LSRII. For split reporters expressed from pcDNA3.1-IRES-GFP, cells were trypsinized and washed in FACS buffer prior to analysis. For each sample, 10,000 cells were collected and data were analyzed by FloJo software (Tree Star, Inc.).

### 3.8g Cellular co-culture assays

Combinations of cells expressing complementary split reporters were added to round bottom 96-well plates ( $5 \times 10^4$  cells per well). Media aliquots (10  $\mu$ L/well) were harvested 24 h after plating and transferred to black 96-well plates for imaging. For Gluc reporters, a stock solution of coelenterazine (Nanolight Technology, 5 mg/mL in ethanol) was diluted 1000-fold in PBS and added to each well (3  $\mu$ L/well) prior to imaging. For Nluc reporters, a stock solution of Nano-Glo Luciferase Assay Substrate (Promega) was diluted 50-fold in PBS and added to each well (10  $\mu$ L/well) prior to imaging. Plates were imaged in a light-proof chamber with an IVIS Lumina (Xenogen) CCD camera chilled to -90 °C. The camera was controlled using Living Image software. All data were analyzed using unpaired t-tests (GraphPad Prism 4).

*BRET assay.* Combinations of cells (1:1) expressing split reporters with N-terminal Venus fusions, Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>, or Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub> were added to round bottom 96-well plates ( $5 \times 10^4$  cells per well) in PBS + 10% FBS. Additionally, combinations of cells expressing split reporters with C-terminal Venus fusions, N<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>-Jun, or C<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>-Fos cells were plated as above. Samples were collected 24 h post-plating, transferred to black 96-well plates (10  $\mu$ L/well), and imaged.

*Time-dependent complementation assay.* Combinations of cells (1:1) expressing Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>, Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub>, Jun-(G<sub>4</sub>S)<sub>5</sub>-N<sub>Gluc</sub>, Fos-(G<sub>4</sub>S)<sub>5</sub>-C<sub>Gluc</sub>, N<sub>Gluc</sub>, or C<sub>Gluc</sub> were added to round bottom 96-well plates ( $5 \times 10^4$  cells per well) and incubated at 37 °C. Media samples were collected 0 – 6 h post-plating, transferred to black 96-well plates (10  $\mu$ L/well), and imaged.

*Dose-response assay.* Combinations of cells (1:1) expressing Jun-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>, Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub>, Jun-(G<sub>4</sub>S)<sub>5</sub>-N<sub>Gluc</sub>, Fos-(G<sub>4</sub>S)<sub>5</sub>-C<sub>Gluc</sub>, N<sub>Gluc</sub>, or C<sub>Gluc</sub> were added to round bottom 96-well plates (7.75 x 10<sup>2</sup> – 5 x 10<sup>4</sup> cells per well). Cells were incubated for 24 h at 37 °C. Media samples were transferred to black 96-well plates (10 µL/well) and imaged.

## References

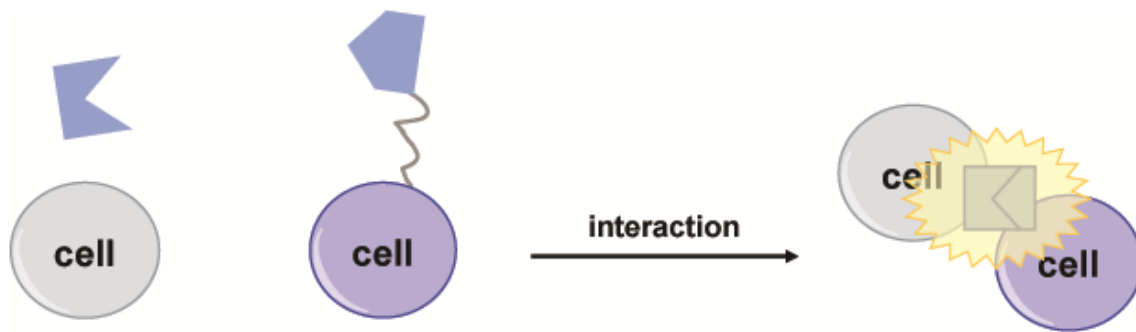
1. Remy I & Michnick SW (2006) A highly sensitive protein-protein interaction assay based on Gaussia luciferase. *Nat. Methods* 3(12):977-979
2. Bovenberg MS, Degeling MH, & Tannous BA (2012) Enhanced Gaussia luciferase blood assay for monitoring of in vivo biological processes. *Anal. Chem.* 84(2):1189-1192
3. Tannous BA & Teng J (2011) Secreted blood reporters: insights and applications. *Biotechnol. Adv.* 29(6):997-1003
4. Rice BW, Cable MD, & Nelson MB (2001) *In vivo* imaging of light-emitting probes. *J. Biomed. Opt.* 6(4):432-440
5. Zhao H, *et al.* (2005) Emission spectra of bioluminescent reporters and interaction with mammalian tissue determine the sensitivity of detection *in vivo*. *J. Biomed. Opt.* 10(4):41210
6. Chao H, Bautista DL, Litowski J, Irvin RT, & Hodges RS (1998) Use of a heterodimeric coiled-coil system for biosensor application and affinity purification. *J. Chromatogr. B: Biomed. Sci. Appl.* 715(1):307-329
7. De Crescenzo G, Litowski JR, Hodges RS, & O'Connor-McCourt MD (2003) Real-time monitoring of the interactions of two-stranded de novo designed coiled-coils: effect of chain length on the kinetic and thermodynamic constants of binding. *Biochemistry* 42(6):1754-1763
8. Tripet B, *et al.* (1996) Engineering a *de novo*-designed coiled-coil heterodimerization domain off the rapid detection, purification and characterization of recombinantly expressed peptides and proteins. *Protein Eng.* 9(11):1029-1042
9. Crabtree GR & Schreiber SL (1996) Three-part inventions: intracellular signaling and induced proximity. *Trends Biochem. Sci.* 21(11):418-422
10. Spencer DM, Wandless TJ, Schreiber SL, & Crabtree GR (1993) Controlling signal transduction with synthetic ligands. *Science* 262(5136):1019-1024

11. Pratt MR, Schwartz EC, & Muir TW (2007) Small-molecule-mediated rescue of protein function by an inducible proteolytic shunt. *Proc. Natl. Acad. Sci. U. S. A.* 104(27):11209-11214
12. De A (2011) The new era of bioluminescence resonance energy transfer technology. *Curr. Pharm. Biotechnol.* 12(4):558-568
13. De A, Jasani A, Arora R, & Gambhir SS (2013) Evolution of BRET Biosensors from Live Cell to Tissue-Scale In vivo Imaging. *Front. Endocrinol. (Lausanne)* 4:131
14. Sun S, Yang X, Wang Y, & Shen X (2016) In Vivo Analysis of Protein-Protein Interactions with Bioluminescence Resonance Energy Transfer (BRET): Progress and Prospects. *Int. J. Mol. Sci.* 17(10)
15. Mo XL & Fu H (2016) BRET: NanoLuc-based bioluminescence resonance energy transfer platform to monitor protein-protein interactions in live cells. *Methods Mol. Biol.* 1439:263-271
16. Nagai T, *et al.* (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* 20(1):87-90
17. Saito K, *et al.* (2012) Luminescent proteins for high-speed single-cell and whole-body imaging. *Nat. Commun.* 3:1262
18. Hall MP, *et al.* (2012) Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem. Biol.* 7(11):1848-1857
19. Dixon AS, *et al.* (2016) NanoLuc complementation reporter optimized for accurate measurement of protein interactions in cells. *ACS Chem. Biol.* 11(2):400-408
20. Wilhem, K, *et al.* (2010) Graft-versus-host disease is enhanced by extracellular ATP activating P2X7R. *Nat. Med.* 16(12):1434-1438
21. Nakamura, M, *et al.* (2006) Cell-surface-localized ATP detection with immobilized firefly luciferase. *Anal. Biochem.* 352(1):61-67
22. Pellegatti, P, *et al.* (2008) Increased level of extracellular ATP at tumor sites: *in vivo* imaging with plasma membrane luciferase. *PLoS One* 3(7):e2599
23. Jones, KA, *et al.* (2015) Visualizing cell proximity with genetically encoded bioluminescent reporters. *ACS Chem. Biol.* 10(4):933-938

# **CHAPTER 4: Membrane bound split reporters for localizing zones of cell contact**

## **4.1 Introduction**

The studies performed with the secreted reporter system (described in Chapter 3) provided proof-of-principle for reporting on cell proximity with split bioluminescent probes. With the successes realized in cultured cell models, I set my sights on translating our system into mouse models. Direct visualization of cell-cell interactions *in vivo* would require the assembled Gluc to remain localized to the sites of production. In the first generation of these tools, both fragments were secreted and culture supernatants were visualized, as compared to the cells themselves, to detect cellular interactions. While certain applications, such as imaging immune cell interactions in the lymph node, could help retain the secreted reporters to a defined location, diffusibility of the split probes could limit the utility of the approach. Thus, we also set out to investigate strategies to tether the Gluc fragments to cell surfaces. By appending one fragment of the luciferase to the cell surface, light emission would be localized to this cell population only when split luciferase-secreting cells are in close proximity (Figure 4-1). In addition to localizing the light emission, combinations of cell surface anchored and secreted reporters could maintain the sensitivity of our approach while also increasing the spatial resolution of the tools.



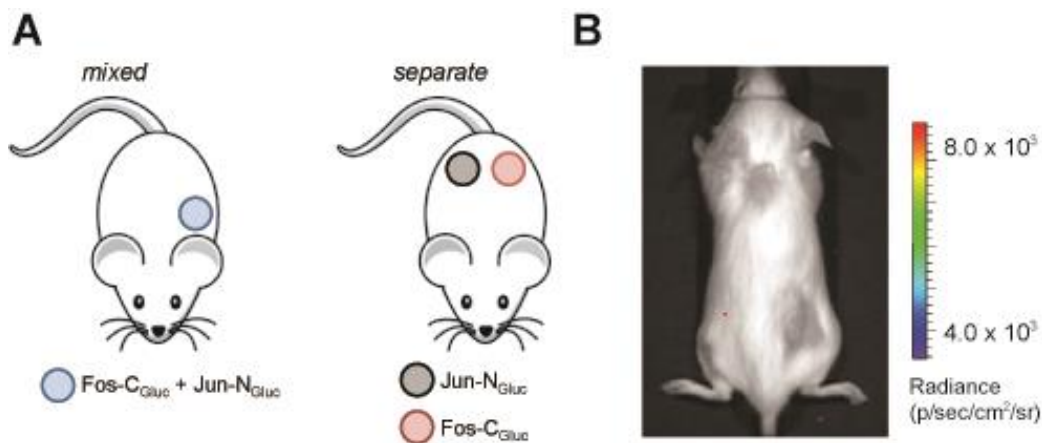
**Figure 4-1.** Strategy for cell surface localized light emission. Bioluminescence is observed when cells secreting one half of the split reporter (gray) interact with cells expressing the complementary split fragment tethered to the cell surface (blue).

#### 4.2 Evaluating cell proximity reporters *in vivo*

In the *in vitro* model described in Chapter 2, the complemented Gluc signal could sensitively detect interacting cells that correlated with the distance between two cell populations. The level of sensitivity achieved suggested that cell-cell interactions could be detected *in vivo*. To evaluate whether Gluc complementation behaved similarly *in vivo*, we aimed to transition the split reporters into mouse models. In live animals, blood flow and interstitial pressures could distribute the split reporter fragments to regions distant from the sites of cellular interactions. This would make it difficult to detect and distinguish the locations of interacting cells.

To determine if two populations of cells expressing each secreted fragment could be detected *in vivo*, I implanted a mixed population of Jun-NGluc and Fos-CGluc expressing DB7 cells in the upper flank of FVB/NJ mice. These coinjected regions contained different ratios (10:1, 1:1, and 0.1:1) of Fos-CGluc and Jun-NGluc cells (totaling  $6 \times 10^7 - 6 \times 10^5 : 6 \times 10^6$  cells per graft). The control group contained mixed NGluc and CGluc expressing DB7 cells at a 1:1 ratio. To determine if Gluc complementation correlates with distance between two cell populations *in vivo*, Jun-NGluc or Fos-CGluc expressing cells alone ( $8 \times 10^6$  cells per graft) were implanted at opposite lower flanks of the mice (Figure 4-2A). Mice were imaged immediately after

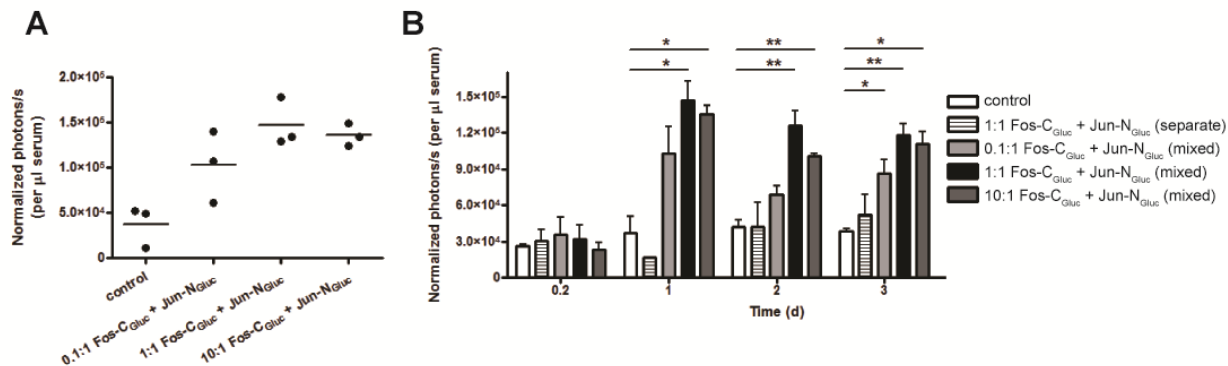
intravenous injection of coelenterazine. For all mice, no bioluminescent signal was detected at the sites of implantation (Figure 4-2B). It is therefore likely that interstitial pressures and blood flow resulted in diffusion of the split reporter fragments and complemented Gluc away from the site of origin.



**Figure 4-2.** Diffusion of the split reporter fragments is observed *in vivo*. (A) Diagram for mice inoculations with mixed (left) or separate (right) subcutaneous injections of DB7 cells expressing Jun-N<sub>Gluc</sub> (gray circles), Fos-C<sub>Gluc</sub> (red circles), or coinjected cells expressing Jun-N<sub>Gluc</sub> and Fos-C<sub>Gluc</sub> (blue circles). The control group contained a 1:1 ratio of DB7 cells expressing N<sub>Gluc</sub> and C<sub>Gluc</sub> or no injected cells. (B) Mice were imaged after intravenous injection of coelenterazine. No signal was detected at sites of implantation for mixed or separate cell injection conditions.

While the fully secreted split reporter system precluded our ability to image cells in defined locations, we reasoned the stability of complemented Gluc should allow for detection of cell-cell interaction in the blood. In fact, secreted forms of Gluc have previously been monitored in blood samples for tumor growth and response to cancer therapeutics (1, 2). When we sampled blood one day post-engraftment, bioluminescence signal was indeed detected in mice bearing 1:1 and 10:1 ratios of Fos-C<sub>Gluc</sub> to Jun-N<sub>Gluc</sub> expressing cells (Figure 4-3A). Photon production was maintained over the three day imaging period. Excitingly, when split reporter expressing cells were engrafted separately, no significant signal was observed (Figure 4-3B). This suggests *in vivo* Gluc complementation is occurring preferentially at sites of direct cell-cell interactions and

not at more distant sites (e.g., in the blood stream). These tools could be useful for determining relative cell-cell interactions from a simple blood draw.



**Figure 4-3.** Gluc complementation can report on cell-cell interactions *in vivo*. Serum samples from blood draws were adjusted to 10  $\mu\text{l}$  final volume with phosphate buffered saline and imaged with coelenterazine. Quantification of bioluminescence from serum samples on (B) day 1 and (C) for 3 d post-cell implantation. All error bars represent the standard error of the mean for  $n = 3$  mice, \* $p < 0.01$ , \*\* $p < 0.001$ .

### 4.3 Transmembrane domains can anchor split luciferase fragments to cell surfaces

To develop a tool in which the signal does remain localized, I aimed to tether the split Gluc fragments to cell surfaces using transmembrane domains. The initial design was based on work from Santos and colleagues (3). In their work, the CD8 transmembrane domain was used to append full-length Gluc to the surface of mouse and human T cells. The small size of a single transmembrane domain (such as CD8) was beneficial for expression levels but was also unable to completely retain the enzyme at the cell surface. In hopes of improving cell retention while still maintaining high expression levels, I was drawn to a truncated version of the T cell protein CD4 (residues 25-242) (4). This fragment of CD4 comprised the transmembrane domain, seven amino acids on the cytosolic tail, and two (of the four) immunoglobulin domains on the extracellular face. Important to split Gluc assembly, the CD4 transmembrane domain has previously been used to tether split fragments of GFP to the surface of distinct neuronal cell

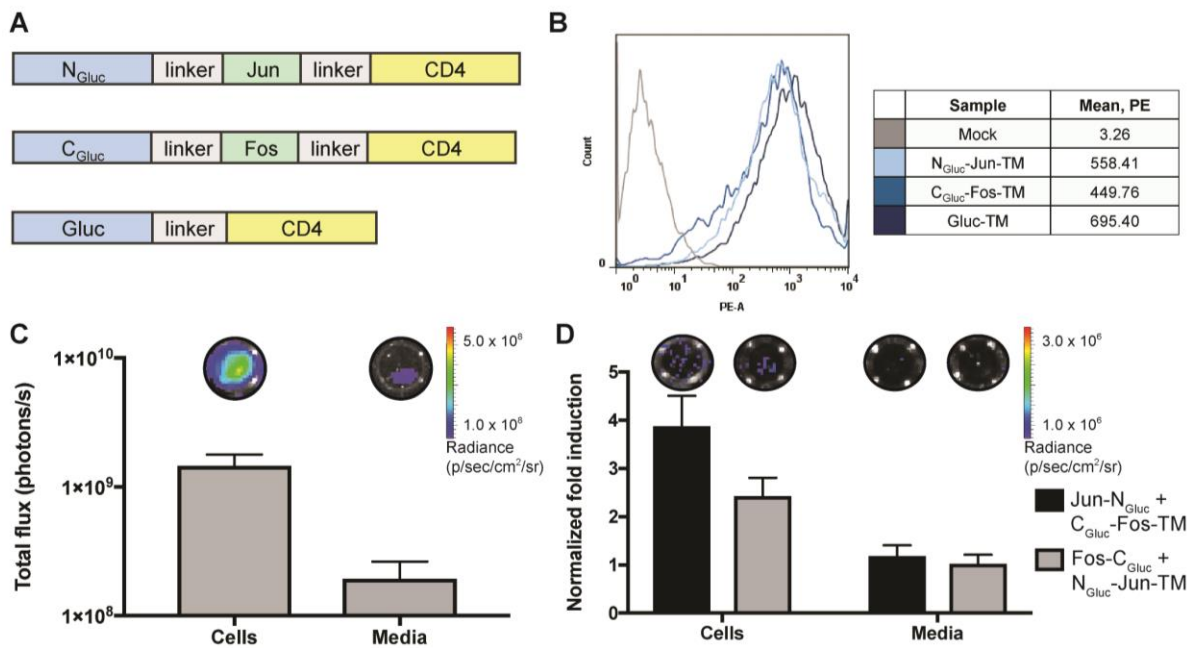


populations. This strategy enabled synaptic connections to be visualized at distances less than 100 nm (5, 6).

To capitalize on these features, I engineered the addition of the truncated CD4 to the C-termini of C<sub>Gluc</sub>-Fos, N<sub>Gluc</sub>-Jun, and Gluc (Figure 4-4A). As with the previously developed secreted reporters, the Gluc fragments were fused to Jun or Fos through a (G<sub>4</sub>S)<sub>2</sub> linker. Standard molecular biology procedures were used to assemble the fragments. An additional linker N-terminal to the CD4 domain was also installed to enable more facile Jun/Fos dimerization (4).

To verify the Gluc-CD4 fusions could anchor the luciferase probes, I expressed the constructs in HEK293 cells and analyzed CD4 surface expression via flow cytometry (Figure 4-4B). Robust expression was observed, indicating CD4 was present at the cell surface. I next examined whether the CD4 transmembrane domain was sufficient to retain the reporters at the cell surface. I cultured cells expressing Gluc-TM and assayed the media and cells for bioluminescence activity. Approximately 10-fold greater signal was detected at the cell surface than in the media (Figure 4-4C). As compared to the Gluc-CD8 fusion, the CD4 transmembrane domain resulted in ~15% better retention at the cell surface, likely due to the larger size of the CD4 domain. Encouraged by these results, I co-plated cells expressing N<sub>Gluc</sub>-Jun-TM or C<sub>Gluc</sub>-Fos-TM with cells secreting the complementary split reporter and assayed as above (Figure 4-4D). No signal was observed in samples collected from the media, indicating the transmembrane domain sufficiently retained the split reporters to prevent complementation in the cellular supernatant. Instead, signal was exclusively observed at the cell surface. It was also noted that cells expressing N<sub>Gluc</sub>-Jun-TM resulted in decreased overall photon output, likely due to Jun self-dimerization at the cell surface. We therefore selected C<sub>Gluc</sub>-Fos-TM for subsequent studies.

While Fos and Jun were still able to drive Gluc complementation, the bioluminescent signal significantly decreased as compared to the secreted reporters. This suggests with secreted and anchored reporters, Fos and Jun are not able to dimerize as efficiently as in the secreted only case. Based on the signal we obtained with our system, we would likely require approximately  $2.5 \times 10^6$  cells implanted subcutaneously to detect signal *in vivo* (3, 7). Therefore, I aimed to improve the sensitivity of this approach.

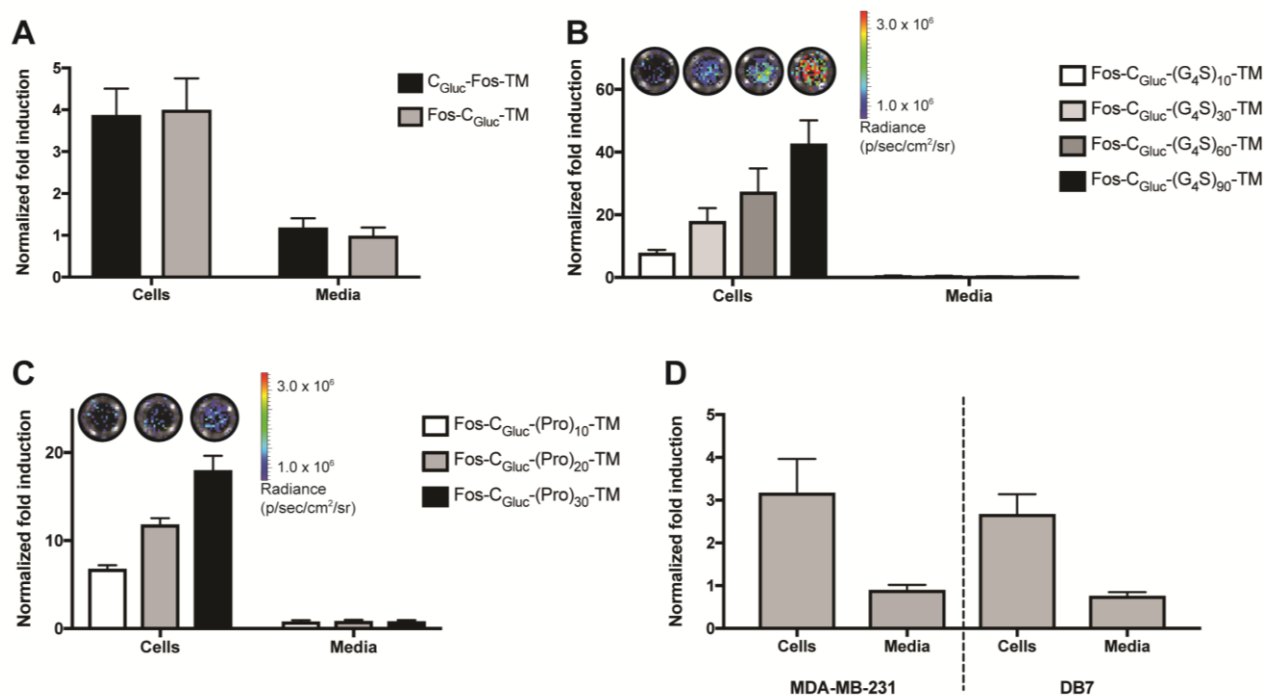


**Figure 4-4.** CD4 transmembrane domain can anchor luciferase to the cell surface. (A) DNA constructs for localizing reporters to the cell membrane consist of  $N_{Gluc}$ -Jun,  $C_{Gluc}$ -Fos, and Gluc attached to a truncated CD4 protein. (B) Gluc fragments are localized to the cell surface. HEK293 cells stably expressing constructs in (A) were stained with PE-anti-CD4 and analyzed by flow cytometry. (C) HEK293 cells expressing Gluc-TM were plated in 96-well round bottom plates ( $5 \times 10^4$  cells). Media from the cultures were collected 24 h after plating and imaged with coelenterazine. Cells were similarly imaged. Sample images from the cell and media imaging are shown. (D) HEK293 cells expressing  $C_{Gluc}$ -FOS-TM or  $N_{Gluc}$ -Jun-TM were mixed with cells secreting the complementary split reporter. Media and cells were imaged as in (B). Fold inductions were normalized to expression levels based on flow cytometry analysis and the bioluminescent signal versus control cells secreting  $N_{Gluc}$  or  $C_{Gluc}$  are plotted. For C and D, error bars represent the standard error of the mean for  $n = 3$  experiments.

#### **4.4 Split reporter transmembrane modifications improve Gluc assembly**

Modifications to the anchored constructs could facilitate improved Gluc assembly to provide a more sensitive readout on interacting cells. If the transmembrane split reporters are too close to the cell surface, other receptors, glycan structures and proteins could be “masking” the reporter fragments and reducing the probability of interactions. To improve the accessibility of the split reporter at the cell surface, I investigated the orientation of the transmembrane reporter and the linker region adjacent to the transmembrane domain.

Our first attempts to improve photon output focused on the orientation of the split reporters at the cell surface. By altering the orientations to encode Fos at the N-terminus (Fos-C<sub>Gluc</sub>-TM), the dimerization domains could be more accessible to drive Gluc complementation. The constructs were transiently expressed in HEK293 cells and co-cultured with cells secreting Jun-N<sub>Gluc</sub>. After 24 h, the media and cells were imaged. The overall photon output was not significantly greater than the original orientation (Figure 4-5A). We therefore selected to continue optimization with the Fos-C<sub>Gluc</sub>-TM construct as this orientation was optimal for the secreted reporters.



**Figure 4-5.** Gluc assembly is improved by split reporter transmembrane modifications. (A) Orientation of the transmembrane domain constructs does not affect complementation. HEK293 cells expressing  $C_{Gluc}$ -Fos TM (black) or Fos- $C_{Gluc}$ -TM (gray) were coplated with cells secreting Jun- $N_{Gluc}$ . After 24 h, media aliquots and cells were imaged with coelenterazine. (B-C) HEK293 cells expressing Fos- $C_{Gluc}$  tethered to the CD4 domain via varying (B) G<sub>4</sub>S linker repeats (10 (white), 30 (light gray), 60 (dark gray), 90 (black)) or (C) Pro linker repeats (10 (white), 20 (gray), 30 (black)) were mixed with cells secreting Jun- $N_{Gluc}$ . Media and cells were imaged 24 h after coculture. Sample images from cell imaging are shown. (D) CD4 transmembrane approach is not generalizable. MDA-MB-231 cells (left) or DB7 cells (right) expressing Fos- $C_{Gluc}$ -(G<sub>4</sub>S)<sub>60</sub>-TM were cultured with Jun- $N_{Gluc}$  secreting cells. After 24 h, cells and media were imaged as above. For A-D, fold inductions were normalized to expression levels based on flow cytometry analysis and the bioluminescent signal versus control cells secreting  $N_{Gluc}$  or  $C_{Gluc}$  are plotted. Error bars represent the standard error of the mean for  $n \geq 3$  experiments.

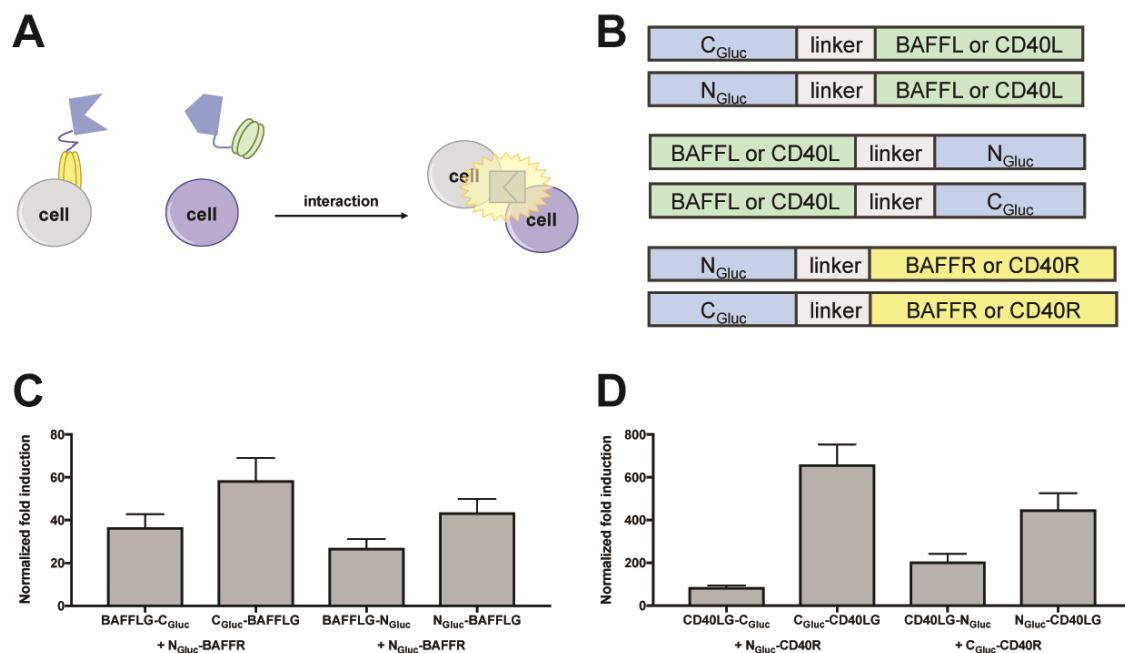
I also investigated whether longer linkers adjacent to the CD4 domain could enhance Gluc assembly. Increasing the linker length by 150-500 amino acids should recapitulate the CD4 structure and allow cell interactions beyond the cell surface to be detected. The original design encoded for Fos- $C_{Gluc}$ -(G<sub>4</sub>S)<sub>2</sub>-TM and I generated new reporters with varying numbers of G<sub>4</sub>S repeats (comprising 10, 30, 60, and 90 repeats) adjacent to the CD4 domain (Fos- $C_{Gluc}$ -(G<sub>4</sub>S)<sub>x</sub>-TM). These reporters were transiently expressed in HEK293 cells and cultured with cells

secreting Jun-N<sub>Gluc</sub>. Media and cell samples were then assayed for Gluc assembly via bioluminescence imaging. With longer linkers, the light emission at the cell surface increased (Figure 4-5B). The transmembrane reporters were still sufficiently held at the cell surface as no significant signal was detected in the media. Unfortunately, the (G<sub>4</sub>S)<sub>90</sub> linker was observed to affect cell growth and morphology. Therefore, longer G<sub>4</sub>S linkers were not investigated further. However, we rationalized that shorter, rigid proline linkers could provide the same, if not better, improvements to photon outputs. To investigate their utility for Gluc complementation, a panel of proline linkers (comprising 10, 20, and 30 repeats) was generated (Fos-C<sub>Gluc</sub>-Pro<sub>x</sub>-TM). The resulting constructs were expressed in HEK293 cells and the cells were mixed with Jun-N<sub>Gluc</sub> secreting cells. After 24 h, cells and media samples were imaged. Transmembrane reporters comprising Pro<sub>10</sub> repeats provided a 2-fold enhancement in signal as compared to reporters with (G<sub>4</sub>S)<sub>2</sub> repeats. As with G<sub>4</sub>S linkers, longer proline linkers also resulted in increased levels of light emission (Figure 4-5C). The (G<sub>4</sub>S)<sub>60</sub> linker was subsequently examined in a variety of cell types (DB7 and MDA-MB-231) but similar trends in light emission were not observed (Figure 4-5D). These cells are more mucinous than HEK293 cells and the thicker glycocalyx is likely to preclude Gluc fragment assembly. While the tools did not translate across a variety of cell lines, these studies did provide important proof of concept for localizing split Gluc bioluminescent signal to the cell surface.

#### **4.5 Ligand-receptor binding can facilitate Gluc assembly for imaging interacting cells**

We aimed to improve Gluc complementation at the cell surface for higher overall signal. Our collective studies with CD4 anchored split reporters suggested a large impediment to Gluc complementation was limited accessibility of the split reporter at the cell surface. We therefore

envisioned utilizing naturally occurring ligand-receptor interactions to drive Gluc assembly and monitor interacting cells (Figure 4-6A). Indeed, Luker and coworkers have previously monitored ligand-receptor binding in a mouse model of breast cancer with split Gluc appended to CXCL12 and CXCR4 (8). These studies utilized  $1 \times 10^6$  total cells in mixed populations for visualization. Improved total photon outputs, and thus more sensitive visualization of interacting cells, is likely to be achieved with a higher affinity ligand-receptor pair. To test this hypothesis, I selected two ligand-receptor pairs from the tumor necrosis factor superfamily: B-cell activating factor (BAFF) ligand and receptor and CD40 ligand and receptor. The BAFF complex ( $K_D$  0.14 nM) is approximately an order of magnitude higher affinity than CXCL12-CXCR4, while the CD40 system ( $K_D$  5.2 nM) is on par with the CXC complex (9, 10). Important to our approach, the Wajant group determined these affinities via full length-Gluc ligand fusions and cell surface expressed receptors. This suggested split Gluc-ligand and Gluc-receptor fusions should not significantly interfere with binding and can be readily expressed in mammalian cells.



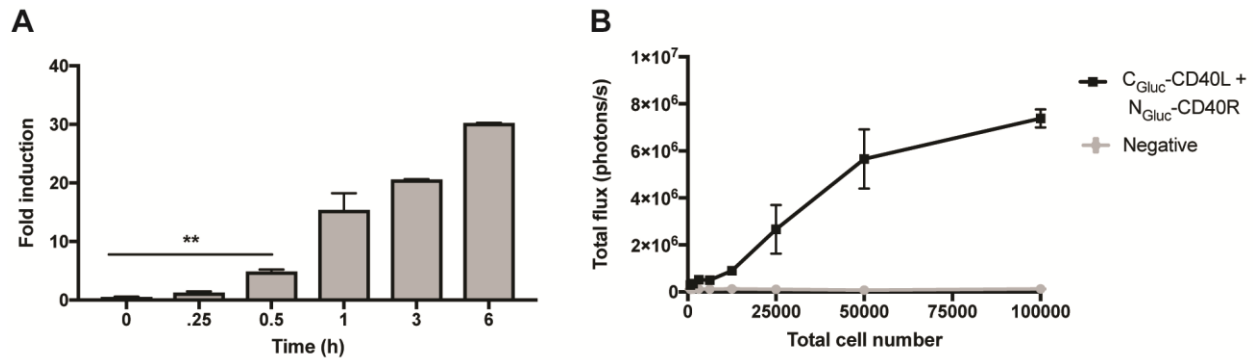
**Figure 4-6.** Ligand-receptor binding facilitates the assembly of Gluc. (A) Strategy for cell surface localized light emission via ligand-receptor binding. (B) DNA constructs comprising Gluc halves tethered to BAFFL or CD40L (via  $G_4S$ ) linkers in various orientations were prepared.  $N_{Gluc}$  and  $C_{Gluc}$  attached to BAFFR or CD40R were also generated. (C-D) N-terminal Gluc-ligand fusions are optimal for Gluc complementation. HEK293 cells transiently expressing various orientations of (C) split Gluc-BAFFL and -BAFFR reporters or (D) split Gluc-CD40L and -CD40R reporters were cocultured in 96-well round bottom plates as indicated. After 24 h, media aliquots from the co-cultures and cells were imaged with coelenterazine. For C and D, expression levels were normalized based on flow cytometry analysis and the fold inductions in bioluminescent signal versus control cells secreting  $N_{Gluc}$  or  $C_{Gluc}$  are plotted. Error bars represent the standard error of the mean for  $n = 3$  experiments.

To identify the optimal orientation of the fusion proteins, we fused  $N_{Gluc}$  or  $C_{Gluc}$  via a  $(G_4S)_2$  linker to the C and N terminus of either BAFF ligand ( $N_{Gluc}$ -BAFFL,  $C_{Gluc}$ -BAFFL, BAFFL- $N_{Gluc}$ , and BAFFL- $C_{Gluc}$ ) or CD40 ligand ( $N_{Gluc}$ -CD40L,  $C_{Gluc}$ -CD40L, CD40L- $N_{Gluc}$ , and CD40L- $C_{Gluc}$ ). Split Gluc fragments were only fused to the N terminus of the receptors ( $N_{Gluc}$ -BAFFR,  $C_{Gluc}$ -BAFFR,  $N_{Gluc}$ -CD40R, and  $C_{Gluc}$ -CD40R) to position Gluc in the extracellular space (Figure 4-6B). We transiently expressed each split reporter in HEK293 cells and co-cultured complementary pairs of cells. Cells and media were assayed for bioluminescence after 24 h. For both systems, no significant signal was detected in the media samples suggesting

the receptors were sufficiently retaining the split Gluc at the cell surface. Robust signal was observed when the cells were imaged. N-terminal Gluc ligand fusions provided the greatest levels of light emission for both BAFF and CD40 systems (Figure 4-6C and 4-6D). Surprisingly, the optimal CD40 ligand-receptor pairing ( $C_{\text{Gluc-CD40L}}$  and  $N_{\text{Gluc-CD40L}}$ ) generated the largest total photon production, with a ~10-fold enhancement in bioluminescence over the optimal BAFF ligand-receptor pair ( $N_{\text{Gluc-BAFFL}}$  and  $C_{\text{Gluc-BAFFR}}$ ). This suggests that binding affinity alone does not predict the total light emission. It is likely that intracellular ligand-receptor association, rates of membrane turnover, or the overall receptor size are likely to influence photon outputs.

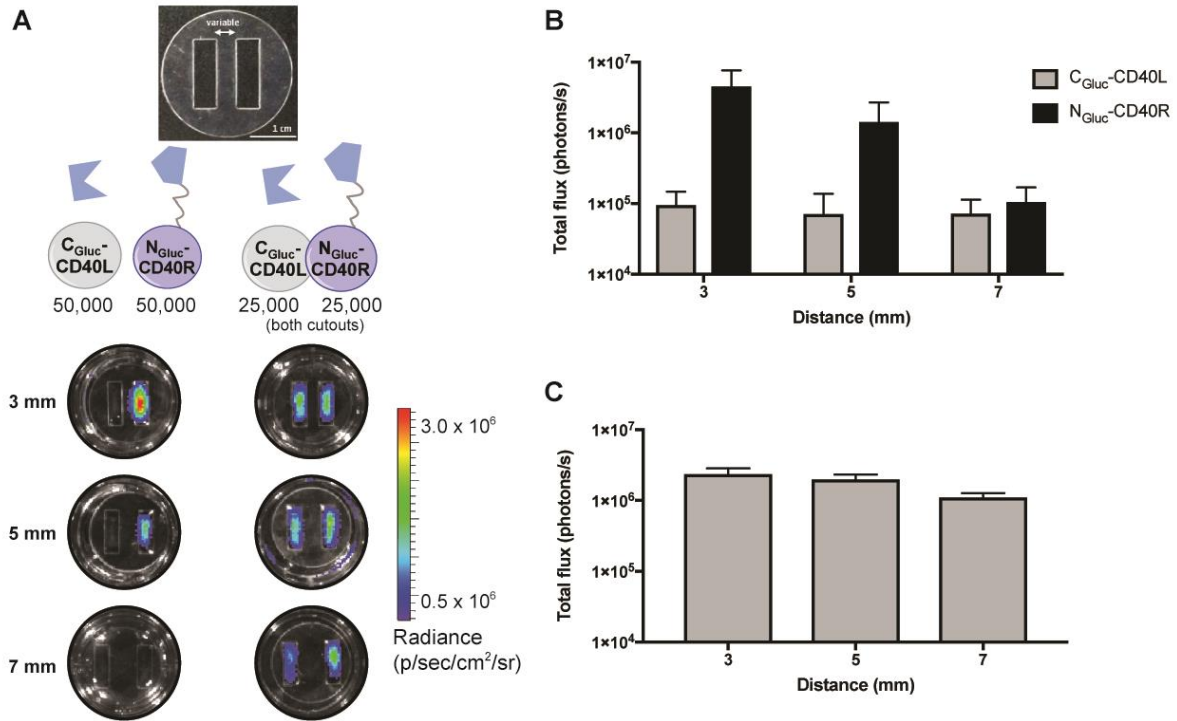
To determine if the split reporters could be utilized for reporting on small numbers of interacting cells, I analyzed the time- and dose-dependency of Gluc complementation with  $C_{\text{Gluc-CD40L}}$  secreting cells and  $N_{\text{Gluc-CD40R}}$  expressing cells. Combinations of the cells were cultured together and media samples from the cultures were assayed at various time points post plating. Signal was detected within 30 min of co-culturing, suggesting cellular interactions could be imaged rapidly (Figure 4-7A). Co-culture studies with different cell densities revealed that as few as 6250 total interacting cells could be detected (Figure 4-7B). These data demonstrate the utility of the reporters for visualizing sensitive cell-cell interactions.





**Figure 4-7.** CD40 split reporters enable sensitive imaging of interacting cells. (A) HEK293 cells stably expressing  $C_{\text{Gluc-CD40L}}$ ,  $N_{\text{Gluc-CD40R}}$ , or  $C_{\text{Gluc}}$  were mixed in 96-well round bottom plates (1:1  $5 \times 10^4$  total cells). Supernatants from the cocultures were removed and cells were imaged with coelenterazine 0-6 h post-plating. The fold inductions in bioluminescent signal for  $C_{\text{Gluc-CD40L}}$  and  $N_{\text{Gluc-CD40R}}$  versus  $C_{\text{Gluc}}$  and  $N_{\text{Gluc-CD40R}}$  (control cells) are plotted. (B) HEK293 cells stably expressing  $C_{\text{Gluc-CD40L}}$  and  $N_{\text{Gluc-CD40R}}$  were plated together (1:1 750-100000 cells total) in 96-well round bottom plates. After 24 h, supernatant from the cocultures was removed and cells were imaged (black). Mixtures of control cells (gray) were similarly imaged. For A and B, error bars represent the standard error of the mean for  $n = 2$  experiments.

The CD40 ligand-receptor system is advantageous for imaging interacting cells, as it should localize signal to one distinct cell population.  $N_{\text{Gluc-CD40R}}$  expressing cells located further away from  $C_{\text{Gluc-CD40L}}$  secreting cells should remain dark due to insufficient quantities of the split Gluc-ligand from reaching distant areas.  $N_{\text{Gluc-CD40R}}$ - and  $C_{\text{Gluc-CD40L}}$ -expressing cells were plated at increasing distances and signal was only detected where receptor-expressing cells were patterned (Figure 4-8). Importantly, light emission also correlated with the distance between the two cell populations. This spatial resolution is likely to be applicable to visualizing events relevant to host-pathogen interactions and metastatic disease.



**Figure 4-8.** Split Gluc-CD40 system localizes signal to cell surface in a distance-dependent manner. (A) Fabricated stencils (top) were used to separate (left) or mix (right)  $C_{Gluc}$ -CD40L (gray) and  $N_{Gluc}$ -CD40R (blue) expressing cells. Media was added 24 h after plating to allow supernatants from each cell population to mix. After 24 h, wells were imaged with coelenterazine. Representative images from three replicate experiments are shown. (B) Quantification of  $C_{Gluc}$ -CD40L (gray) and  $N_{Gluc}$ -CD40R (black) expressing cells separated by 3, 5, and 7 mm. (C) Quantification of mixed  $C_{Gluc}$ -CD40L and  $N_{Gluc}$ -CD40R expressing cells. For B and C, regions of interest were defined by the stencil chambers. Error bars represent the standard error of the mean for  $n = 3$  experiments.

#### 4.6 Conclusions and future directions

In our initial attempts to detect interacting cells with split Gluc reporters *in vivo*, we were not able to visualize cells in a defined location. However, when we assayed for complemented Gluc *via* blood draws, signal was detected only when cells were mixed prior to engraftment. This suggests sites of direct cell-cell interactions can be preferentially monitored with this approach. To address the need for localized signal, we utilized transmembrane domains to tether split-Gluc to the cell surface. Our initial anchoring strategy with CD4 provided readouts on interacting cells and retained the signal at the surface. Subsequent optimization studies with longer linkers

enhanced light emission and suggested pushing the luciferase further from the membrane facilitates fragment assembly. Indeed, split Gluc fusions to the CD40 ligand and receptor provide robust signal while maintaining distant-dependent light emission.

The sensitivity of the ligand-receptor approach suggests these tools are capable of detecting interacting cells *in vivo*. Initial studies are aimed at detecting and spatially resolving two distinct cell populations. Based on these findings, these imaging tools could be used to generate metastasis-specific animal models. These models could enable experiments aimed at identifying how metastatic cells escape primary tumors and infiltrate secondary tissues. Additionally, the CD40 system suggests native interacting transmembrane proteins are good candidates for driving Gluc assembly to visualize direct cell-cell interactions.

## **4.7 Materials and methods**

### **4.7a General cloning materials**

Split reporters were amplified from previously generated Gluc fusions (Chapter 2 and 3) using DreamTaq DNA polymerase and reaction buffer (ThermoFisher) or Q5 hot start high fidelity DNA polymerase and reaction buffer (New England Biolabs). dNTPs were purchased from ThermoFisher. All primers were generated by Integrated DNA Technologies. Amplified PCR products were restriction enzyme digested (New England Biolabs) and inserted into the pBMNpuro vector (courtesy of Nolan laboratory, Stanford) or pcDNA3.1-IRES-GFP vector using T4 DNA ligase (New England Biolabs). Plasmid construction was confirmed by DNA sequencing.

#### 4.7b Split Gluc CD4 transmembrane domain constructs

DNA sequences encoding CD4 were amplified from pDEST-HemmarG (a gift from Yuh-Nung Jan, Addgene plasmid #31221) using the following primers:

5'-ctcgagtccagaaggcctccag-3' and 5'-atagcggccgcttagcgccttcggtgcc-3'

DNA sequences encoding N<sub>Gluc</sub>-(G<sub>4S</sub>)<sub>2</sub>-Jun-(G<sub>4S</sub>)<sub>2</sub>, C<sub>Gluc</sub>-(G<sub>4S</sub>)<sub>2</sub>-Fos-(G<sub>4S</sub>)<sub>2</sub>, or Gluc-(G<sub>4S</sub>)<sub>2</sub> were generated using overlap PCR the following primers:

5'-atagaattcaagcccaccgagaac-3' and 5'-gccgccaccggaaccaccgccaccattaattaagcctatgccgcc-3'

5'-ggtggttccggtggcggcggttagcatgcatagcggc-3' and 5'-ggccttctggaactcgaggctaccgccgcc-3'

5'-atagaattcgaggcgatcgtcgac and 5'-gccgccaccggaaccaccgccaccattaattaagtcaccaccggc-3'

5'-ggtggttccggtggcggcggttagcatgcatggcggc-3' and 5'-ggccttctggaactcgaggctaccgccgcc-3'

5'-atagaattcaagcccaccgagaac-3' and 5'-gccgccaccggaaccaccgccaccattaattaagtcaccaccggc-3'

The Gluc reporters and CD4 products were fused using overlap PCR and were ligated to sequences containing CD8 leader sequence-FLAG in pBMN iPuro.

DNA sequences encoding Fos-(G<sub>4S</sub>)<sub>2</sub> and C<sub>Gluc</sub>-(G<sub>4S</sub>)<sub>2</sub> were simultaneously ligated to sequences encoding the CD4 domain in pBMN-iPuro. The reporter Fos and C<sub>Gluc</sub> inserts were constructed using the following primers:

5'-atagaattcggcggcctgaccgac-3' and 5'-ataatgcatgctaccgccgccacc-3'

5'-ataatgcatgaggcgatcgtcgagattc-3' and 5'-atactcgaggctaccgccgccacc-3'

DNA sequences encoding (G<sub>4S</sub>)<sub>10</sub> were ligated to sequences encoding Fos-(G<sub>4S</sub>)<sub>2</sub>-C<sub>Gluc</sub>-CD4 or C<sub>Gluc</sub>-(G<sub>4S</sub>)<sub>2</sub>-Fos-CD4 in pBMN-iPuro. The linkers were generated by overlap PCR using (G<sub>4S</sub>)<sub>2</sub> as a template using the following primers:

5'-ggcggaggcgggtctggggcgaggctctggtggcggtgttc-3' and 5'-

actgectctccaccgcttccacccccgcgctaccgccgcca-3'

5'-atattaattaatggaggtggaggctcagggggcggaggcagcggcggaggcggg-3' and 5'-  
atactcgagactcccgccctccagaacctccgccccactgcctctccac-3'

5'-ataacgcgtggaggtggaggctcagggggcggaggcagcggcggaggcggg-3' and 5'-  
atactcgagactcccgccctccagaacctccgccccactgcctctccac-3'

To generate (G<sub>4</sub>S)<sub>20</sub> and (G<sub>4</sub>S)<sub>30</sub> linkers, DNA sequences encoding (G<sub>4</sub>S)<sub>10</sub> were simultaneously ligated to sequences encoding Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub>-CD4 or C<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>-Fos-CD4 in pBMN-iPuro. The linkers were generated by overlap PCR using (G<sub>4</sub>S)<sub>10</sub> as a template using the following primers:

5'-atagccggccaggaggtggaggc-3' and

5'-atactcgagactcccgccctccagaacctccgccccactgcctctccac-3'

5'-atattaattaatggaggtggaggctcagggggcggaggcagcggcggaggcggg and

5'-atatggccggcactcccgccctc-3'

5'-ataagcgtggaggtggaggc and 5'-atactcgagactcccgccctccagaacctccgccccactgcctctccac-3'

5'-atagccggccaggaggtggaggc and 5'-ataagcgtactcccgccctc-3'

To generate (G<sub>4</sub>S)<sub>60</sub> and (G<sub>4</sub>S)<sub>90</sub> linkers, DNA sequences encoding (G<sub>4</sub>S)<sub>30</sub> were simultaneously ligated to sequences encoding Fos-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub>-CD4 or C<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>-Fos-CD4 in pBMN-iPuro. The linkers were generated by overlap PCR using (G<sub>4</sub>S)<sub>30</sub> as a template using the following primers:

5'-ataatgcatgaggcgcagtcgagattc-3' and 5'-atatggcgcgccaactcccgccctc-3'

5'-ataatgcatgaggcgcagtcgagattc-3' and 5'-atatgtttaaacactcccgccctc-3'

5'-atagcgcgccaggaggtggaggctcagg-3' and 5'-atagcggccgcttagcgccttcggtgcc-3'

5'-atagtttaaacaggaggtggaggctcagg-3' and 5'-atagcggccgcttagcgccttcggtgcc-3'

DNA sequences encoding Fos-(G<sub>4</sub>S)<sub>2</sub>-(Pro)<sub>10</sub> and C<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>-(Pro)<sub>10</sub> were ligated to sequences encoding C<sub>Gluc</sub>-CD4 or Fos-CD4 in pBMN iPuro. The inserts were generated with the following primers:

5'-tataatgcatggcggcctgaccg-3' and

5'-atactcgagtggaggggggtggcggaggtgggggagggggagcgcctgctaccgccccacc-3'

5'-ataatgcatgaggcgatcgtcgagattc-3' and

5'-atactcgagtggaggggggtggcggaggtgggggagggggagcgcctgctaccgccccacc-3'

DNA sequences encoding (Pro)<sub>20</sub> and (Pro)<sub>30</sub> linkers were ligated to sequences encoding C<sub>Gluc</sub>-CD4 or Fos-CD4 in pBMN iPuro. The inserts were generated with the following primers:

5'-ataagcgtccaccgccaccacctccccctccccacc-3'

5'-atactcgagtggcgggggagggcggtggaggggggtggcg-3'

5'-ataagcgtccgcctccacccccgccaccgccaccacctccccctccccacc-3'

5'-atactcgagcgggtggaggaggcggtggcgggggagggcggtggaggggggtggcg-3'

5'-ataagcgtccccccccctccccgcctccacccccgccacc-3'

5'-atactcgagtggcggaggtgggggaggtggaggaggcggtggcgg-3'

#### **4.7c Split Gluc tumor necrosis factor superfamily constructs**

DNA sequences encoding BAFFR and BAFFL (Rual Genome Res 2004) and CD40R and CD40L were obtained from DNASU (Rual Genome Res 2004). DNA sequences encoding CD40R and CD40L (Ramachandran Science 2004) were obtained from DNASU Plasmid Repository.

DNA sequences encoding chicken tenascin C (TNC) fused to BAFFL and CD40L were ligated to sequences encoding CD8LS-N<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>, CD8LS-C<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>, CD8LS-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>, and CD8LS-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub> in pcDNA3.1-IRES-GFP. The inserts were generated with the following primers:

5'-ggctctggtgtccagcctgagagagcagGGCGCGCCAagtgatcagaatcctcaattgcg-3'

5'-cgacatcaaggatctgctgagcagactggaggagctggagggtctggtgtccagcctgag-3'

5'-ataGGCCGGCCAgcctgtggctgcgccgccccgacatcaaggatctgctgagc-3'

5'-ataATTAATTAAGAGTTTGAGTAAGCCAAAGGACGTG-3'

5'-ataTCTAGAattattaGAGTTTGAGTAAGCCAAAGGACGTG-3'

5'-ggctctggtgtccagcctgagagagcagGGCGCGCCAagtcagaagaacagtcactcaa-3'

5'-ataATTAATTAACAGCAGTTTCAATGCACCAAAAA-3'

5'-ataTCTAGAattattaCAGCAGTTTCAATGCACCAAAAA-3'

DNA sequences encoding BAFFR and CD40R were ligated to sequences encoding N<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub> and C<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub> in pcDNA3.1-IRES-GFP. The inserts were generated with the following primers:

5'-ataggccggccagaaccaccactgcatgagc-3' and 5'-ataTCTAGAattattactgtctctctgactgagatgc-3'

5'-ataggccggccaaggcgaggccccggagcctg-3' and 5'-atatctagaattattattgttctcaggccggccgtctt-3'

DNA sequences encoding CD8LS-N<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>-TNC-CD40L, CD8LS-C<sub>Gluc</sub>-(G<sub>4</sub>S)<sub>2</sub>-TNC-CD40L, CD8LS-TNC-CD40L-(G<sub>4</sub>S)<sub>2</sub>-N<sub>Gluc</sub>, and CD8LS-TNC-CD40L-(G<sub>4</sub>S)<sub>2</sub>-C<sub>Gluc</sub> were ligated in pBMN-IRES-GFP iPuro. The inserts were generated with the following primers:

5'-gactgccggatccaccatggccttaccagtgaccg-3'

5'-ATGCATtattagcctatgccgcctgtgc-3'

5'-ATGCATtattagtcaccaccggcccc-3'

5'-ATGCATttattaGAGTTTgAGTAAGCCAAAGGACGTGA-3'

DNA sequences encoding (G<sub>4</sub>S)<sub>2</sub>-CD40R-IRES-GFP were ligated to sequences encoding CD8LS-N<sub>Gluc</sub> or CD8LS-C<sub>Gluc</sub> in pBMN-IRES-GFP iPuro. The insert was generated with the following primers:

5'-atattaattaatggtggcgggtggttccgg-3' and 5'-tacgcggccgcttactgtacagctcgtccatgc-3'

#### **4.7d Cell culture**

HEK293 cells (American Type Cell Culture) and MDA-MB-231 cells were cultured in DMEM (Corning) supplemented with 10% (vol/vol) fetal bovine serum (FBS, Life Technologies), penicillin (100 U/mL), and streptomycin (100 µg/mL). Cells were maintained in a 5% CO<sub>2</sub>, water-saturated incubator at 37 °C. Transient transfections of the reporter constructs were performed with cationic lipids (Lipofectamine 2000; Invitrogen). HEK293 cells were transduced with amphotropic retrovirus followed by selection with puromycin (10 µg/mL).

#### **4.7e Flow cytometry**

Split reporter expression was verified in transiently transfected cells via flow cytometry. Cells expressing split-CD4 reporters were incubated with PE-α-CD4 (Biolegend, 5 µL/1 x 10<sup>6</sup> cells) on ice for 1 h. Cells were washed (3 x 200 µL) prior to analysis on a BD Biosciences LSRII. For split reporters expressed from pcDNA3.1-IRES-GFP or pBMN-IRES-GFP, cells were trypsinized and washed in FACS buffer prior to analysis. For each sample, 10,000 cells were collected and data were analyzed by FloJo software (Tree Star, Inc.).

#### **4.7f Cellular co-culture assays**



Combinations of cells expressing complementary secreted and transmembrane split reporters were added to round bottom 96-well plates ( $5 \times 10^4$  cells per well). Media aliquots (10  $\mu$ L/well) were harvested 24 h after plating and transferred to black 96-well plates for imaging. A stock solution of coelenterazine (Nanolight Technology, 5 mg/mL in ethanol) was diluted 1000-fold in PBS and added to each well (3  $\mu$ L/well) prior to imaging. Plates were imaged in a light-proof chamber with an IVIS Lumina (Xenogen) CCD camera chilled to -90 °C. The camera was controlled using Living Image software. All data were analyzed using unpaired t-tests (GraphPad Prism 4).

*Time-dependent complementation assay.* Combinations of cells (1:1) expressing C<sub>Gluc</sub>-CD40L, N<sub>Gluc</sub>-CD40R, or C<sub>Gluc</sub> were added to round bottom 96-well plates ( $5 \times 10^4$  cells per well) and incubated at 37 °C. Media samples were collected 0 – 24 h post-plating, transferred to black 96-well plates (10  $\mu$ L/well), and imaged.

*Cell dose-response assay.* Combinations of cells (1:1) expressing C<sub>Gluc</sub>-CD40L, N<sub>Gluc</sub>-CD40R, or C<sub>Gluc</sub> were added to round bottom 96-well plates ( $7.5 \times 10^2$  –  $1 \times 10^5$  cells per well). Cells were incubated for 24 h at 37 °C. Media samples were transferred to black 96-well plates (10  $\mu$ L/well) and imaged.

#### **4.7g Distance dependency of Gluc assembly**

Biocompatible stencils were generated as previously described (11). HEK293 cells expressing C<sub>Gluc</sub>-CD40L, N<sub>Gluc</sub>-CD40R, or C<sub>Gluc</sub> were plated ( $5 \times 10^4$  cells per stencil chamber) and allowed

to adhere (24 h) before the media was removed. Fresh DMEM (750  $\mu$ L) was added and cells were analyzed 24 h later for Gluc activity.

## References

1. Tannous BA & Teng J (2011) Secreted blood reporters: insights and applications. *Biotechnol. Adv.* 29(6):997-1003
2. Bovenberg MS, Degeling MH, & Tannous BA (2012) Enhanced Gaussia luciferase blood assay for monitoring of *in vivo* biological processes. *Anal. Chem.* 84(2):1189-1192
3. Santos EB, *et al.* (2009) Sensitive *in vivo* imaging of T cells using a membrane-bound Gaussia princeps luciferase. *Nat. Med.* 15(3):338-344
4. Feinberg EH, *et al.* (2008) GFP Reconstitution Across Synaptic Partners (GRASP) defines cell contacts and synapses in living nervous systems. *Neuron* 57(3):353-363
5. Gordon MD & Scott K (2009) Motor control in a *Drosophila* taste circuit. *Neuron* 61(3):373-384
6. Yamagata M & Sanes JR (2012) Transgenic strategy for identifying synaptic connections in mice by fluorescence complementation (GRASP). *Front. Mol. Neurosci.* 5:18
7. Wurdinger T, *et al.* (2008) A secreted luciferase for *ex vivo* monitoring of *in vivo* processes. *Nat. Methods* 5(2):171-173
8. Luker KE, *et al.* (2012) *In vivo* imaging of ligand receptor binding with Gaussia luciferase complementation. *Nat. Med.* 18(1):172-177
9. Balabanian K, *et al.* (2005) The chemokine SDF-1/CXCL12 binds to and signals through the orphan receptor RDC1 in T lymphocytes. *J. Biol. Chem.* 280(42):35760-35766
10. Lang I, *et al.* (2016) Binding Studies of TNF Receptor Superfamily (TNFRSF) Receptors on Intact Cells. *J. Biol. Chem.* 291(10):5022-5037
11. Jones, KA, *et al.* (2015) Visualizing cell proximity with genetically encoded bioluminescent reporters. *ACS Chem. Biol.* 10(4):933-938

# CHAPTER 5: Orthogonal luciferase-luciferin pairs for bioluminescence imaging<sup>1</sup>

## 5.1 Introduction

As mentioned in Chapter 1, bioluminescence 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-luciferin pairs for *in vivo* use. The optimal luciferases (from the insect family) use the same substrate, D-luciferin (Figure 5-1A) (1, 2). Thus, they cannot easily discriminate multiple cell types in a single subject. Additionally, unlike fluorescent protein technologies, a diverse suite of accessible bioluminescent probes does not yet exist. To address this void, D-luciferin analogs have been engineered to emit different colors of light (3-5). However, these substrates are still utilized by the same luciferases, precluding the distinct genetic tagging of individual cell types. Insect luciferases have also been engineered to emit different colors of light with D-luciferin (6-8). The observed emission spectra are not sufficiently resolved, though, for routine use in complex tissues or animals. Discriminating among different wavelengths in bioluminescence (and whole body optical imaging, in general) is exceedingly difficult.

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

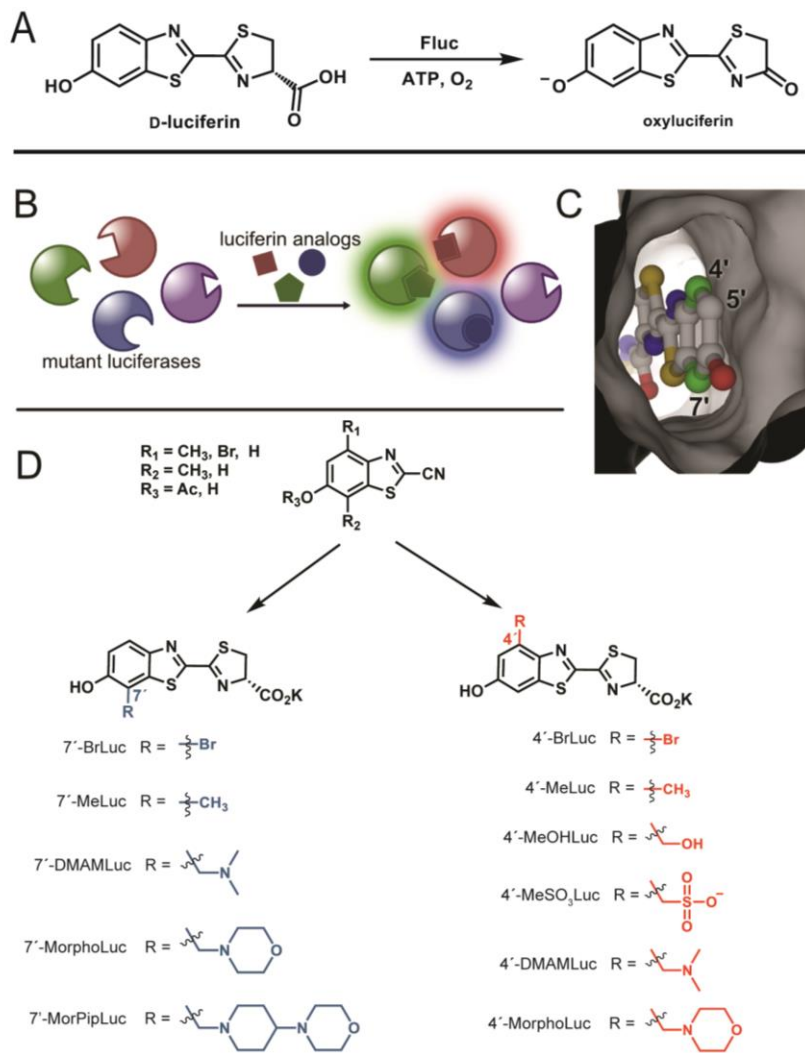
---

<sup>1</sup> William B. Porterfield, Colin M. Rathbun, David C. McCutcheon, and Miranda A. Paley contributed to the work presented in this chapter.

enabling their tandem application *in vivo* (15, 16). Coelenterazine is less ideal for use in these environments, though, owing to its suboptimal bioavailability and stability (2, 17). Other naturally occurring luciferases and luciferins can be used in combination with Fluc/D-luciferin or other bioluminescent systems (16, 18). However, most of these native pairs remain poorly characterized or ill-suited for routine use. Artificial (i.e., mutant) luciferases can exhibit altered bioluminescent properties, including tolerance for chemically modified substrates. Fluc itself has been manipulated to process analogs of D-luciferin (19). In elegant work along these lines, Miller and coworkers prepared a class of non-natural aminoluciferins that were found to be robust light emitters with Fluc, but the products inhibited the enzymatic reaction (20). Product inhibition was relieved using mutated versions of the enzyme (21). These same mutations also resulted in sharply reduced emission with D-luciferin, providing key precedent for the development and utilization of orthogonal pairs (22). The mutant enzymes from these studies, though, were less selective for one analog over another perhaps due to the structural similarities between the luciferin scaffolds. Simultaneous enzyme-substrate manipulation has also been applied to aequorin (a marine photoprotein) and the luciferase from the deep-sea shrimp *Oplophorus gracilirostris* (23, 24). In both cases, altered bioluminescent outputs (e.g., colors and stabilities) were achieved, but orthogonal substrate usage was not realized.

Here we report a strategy for the *de novo* production of orthogonal luciferase-luciferin pairs. We synthesized a series of sterically modified luciferins that were poor emitters with Fluc, but intrinsically capable of robust light production. We then iteratively screened these analogs with libraries of mutant luciferases and identified substrate-selective enzymes. A lead set of “hits” was also biochemically characterized. Importantly, when the lead mutants and analogs 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 multi-cellular imaging applications. Importantly, our approach to identifying orthogonal bioluminescence pairs is also general and should enable rapid diversification of the bioluminescence toolkit.

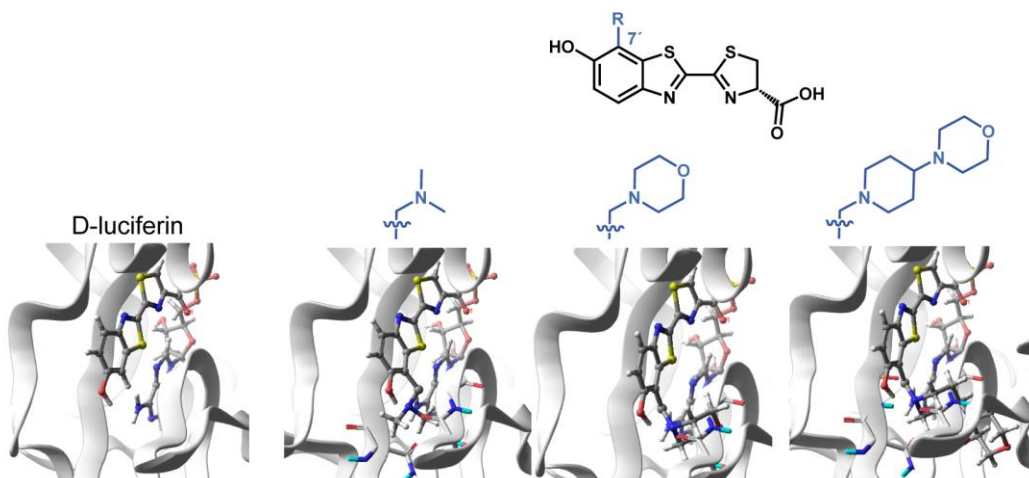


**Figure 5-1.** Expanding the bioluminescence toolkit with unique enzyme-substrate pairs. (A) Luciferase-mediated light production proceeds via an adenylation-oxidation sequence. (B) Strategy to develop orthogonal luciferase-luciferin pairs via substrate resolution. Genetically engineered luciferases 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.

## 5.2 Designing and constructing sterically modified luciferins

To expediently identify orthogonal bioluminescence tools, we aimed to screen sterically perturbed luciferins against libraries of mutant luciferases (Figure 5-1B). We used the Fluc/D-luciferin pair as a starting point for several reasons. First, this duo is the most widely used in biomedical imaging applications owing to the non-toxicity of the reagents and bioavailability of the substrate (25, 26). Second, the Fluc/D-luciferin reaction releases the highest percentage of tissue penetrating light among known bioluminescent families (27). Thus, new enzymes and substrates based on the firefly pair would be more applicable to *in vivo* studies. Third, a wealth of structural and biochemical information on Fluc could guide our engineering efforts (11, 28-31). Finally, D-luciferin derivatives are arguably the most synthetically tractable luciferin architectures (32, 33).

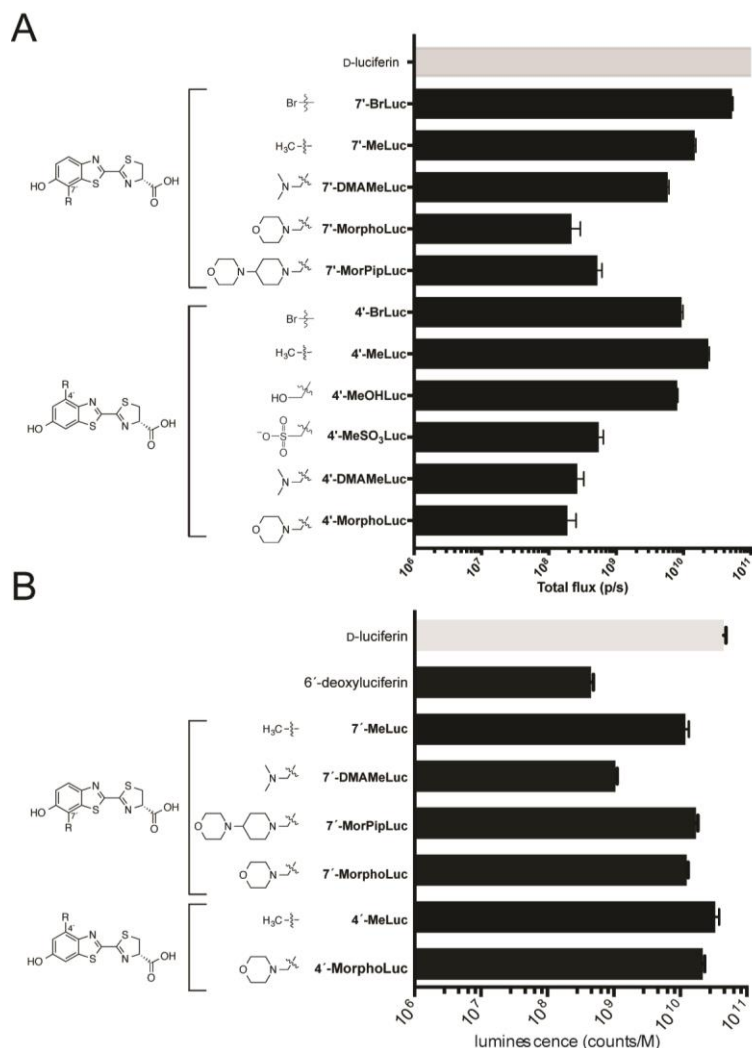
Generating an expanded set of bioluminescent tools required access to diverse luciferin scaffolds. A variety of D-luciferin analogs have been synthesized over the past four decades, and those capable of robust emission with Fluc harbor common features: an electron-donating group at the 6' position, a carboxylate appendage (for adenylation), and an abstractable proton alpha to the carboxylate (3, 5, 34-40). Beyond these requirements, Fluc can tolerate a surprisingly large variety of modified luciferins, including 6'-amino substituents, alkylated and acylated scaffolds, and even luciferins with non-natural chromophores (4, 19, 21, 33, 38, 39, 41-46). Crystallographic analyses have also corroborated these experimental results, indicating flexibility within the luciferase active site and “space” to accommodate luciferins with appendages at or near the 6' - position (30, 31).



**Figure 5-2.** Docking studies with sterically modified luciferins. The 7' position of D-luciferin abuts the Fluc active site. Luciferins in Fluc (PDB: 4G36) were modified to contain a dimethyl amino methyl, methyl morpholino, or methyl morpholino-piperidyl appendage. As the steric bulk increased, more clashes with  $\beta$ -sheets were observed. When analyzed with Maestro software, all three of the analogs shown failed to dock in any conformation under various Glide conditions (SP, XP, or with SMARTS constraints).

Unlike most efforts to produce luciferin analogs reported to date, we were attracted to the 4' and 7' positions of the luciferin core. These positions lie in close proximity to the Fluc backbone (Figure 5-1C). Substrates with additional steric bulk at these sites would likely be occluded from the Fluc active site and thus good targets for orthogonal probe development: while poor emitters with the native enzyme, the molecules could potentially give off light with designer mutants. Indeed, preliminary docking studies suggested that only analogs with small (e.g., 2-3 atoms) substituents at C4' and C7' could effectively access the active site (Figure 5-2). Generating 4'- and 7'-modified luciferins presented an early challenge. These positions have been rarely exploited for analog development, and no prior syntheses were amenable to preparing libraries or large quantities of these probes. Rapid, high-yielding syntheses were essential, as large quantities of luciferins are required for light emission assays. Fortunately, the core benzothiazole unit of the desired analogs could be accessed from a common route (Figure 5-1D) and in multi-gram quantities (32, 36). From a small number of intermediates, we installed

functional handles at C4' and C7', to ultimately assemble a panel of luciferins on large scale (1–10 g).



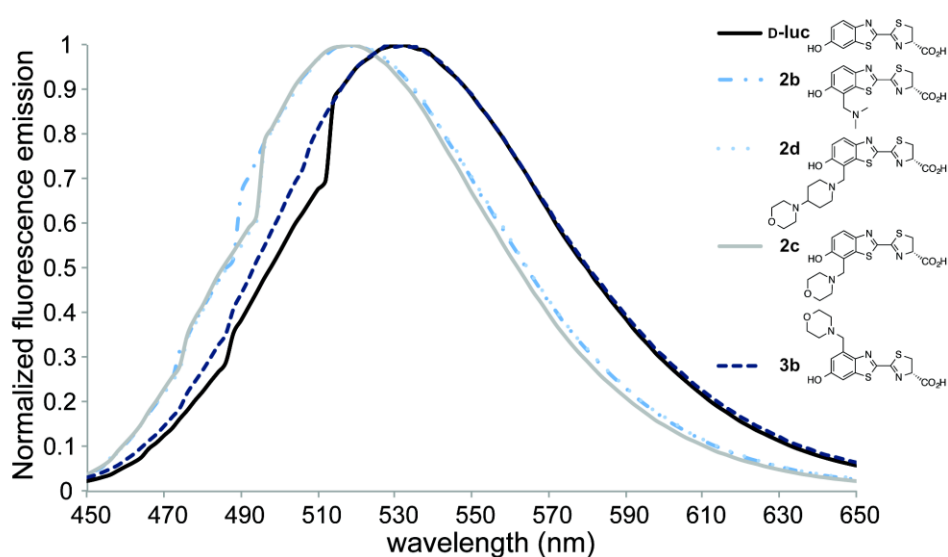
**Figure 5-3.** Measuring luciferin light emission. (A) Bioluminescence from luciferin analogs (100  $\mu$ M) incubated with 1  $\mu$ g of Fluc. Emission intensities are plotted as total photon flux values on a log scale. Error bars represent the standard error of the mean for  $n = 3$  experiments. (B) Chemiluminescence with luciferin analogs. Emission intensities are plotted as counts per molar luciferin on a log scale. Error bars represent the standard error of the mean for  $n \geq 3$  experiments.

### 5.3 Analyzing bioluminescent light emission with modified luciferins

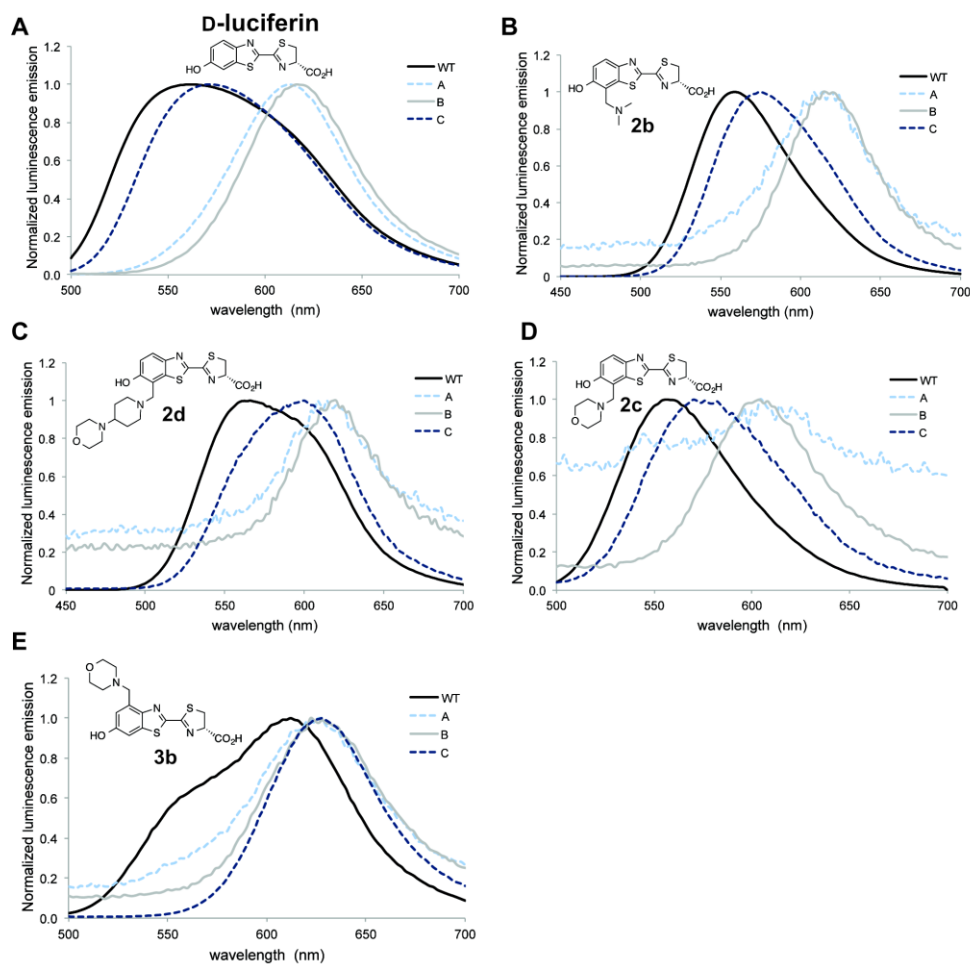
With the modified luciferins in hand, we first evaluated their optical properties with Fluc. All analogs were competent light emitters and could be processed by the enzyme (Figure 5-3).



However, the emission intensities were much weaker than those observed with D-luciferin, the native substrate. Interestingly, the largest analog (7'-MorPipLuc) was not the weakest emitter, suggesting that steric modification alone does not dictate luciferin utilization. Consistent with the observed light outputs, the measured kinetic constants for a subset of tested analogs showed reduced performance relative to D-luciferin (Table 5-1). For example, the measured  $K_m$  values were ~100-fold larger than the native substrate, with the largest analogs (7'-MorphoLuc and 7'-MorPipLuc) exhibiting the lowest relative binding affinities. Despite their large  $K_m$  values, 7'-DMAMeLuc, 7'-MorphoLuc and 7'-MorPipLuc exhibited emission spectra similar to D-luciferin (Figures 5-4 and 5-5). Only the C4'-modified analog 4'-MorphoLuc emitted noticeably redshifted bioluminescent light, likely due to poor Fluc binding in the excited state or the luminophore being forced into a more polar environment (30, 47).



**Figure 5-4.** Fluorescence emission spectra. Fluorescence emission spectra for luciferin solutions (100  $\mu$ M in PBS, pH 7.4). An excitation wavelength of 350 nm was used.



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

#### 5.4 Measuring the light-emitting potential of luciferin analogs

We attributed the weak bioluminescence of the analogs to poor utilization by Fluc. It was possible, though, that the luciferins were simply not capable of photon production upon activation and oxidation in the active site. For productive bioluminescence, an analog must be able to reach an electronic excited state (S1) and relax back to the ground state with concomitant photon release (48, 49). If an analog cannot reach S1 or emit efficiently from that state, reduced photon outputs would be expected. Such molecules would also be poor candidates for orthogonal probe development. To ensure that our lead analogs were intrinsically capable of light emission,

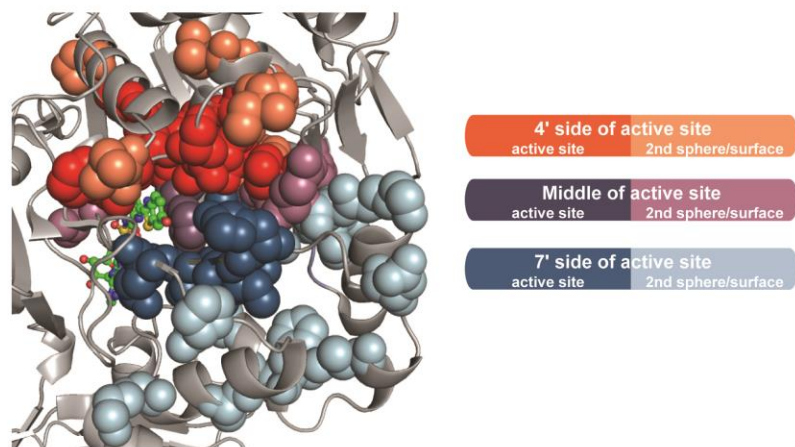
we utilized a previously described chemiluminescence assay (50). This process mimics the enzymatic reaction itself via formation of an activated ester intermediate, followed by proton abstraction and subsequent reaction with molecular oxygen (34, 49, 51). When a subset of the analogs was subjected to the assay, robust light emission was observed (Figure 5-3B). In fact, photon outputs for some of the weakest bioluminescent emitters (including 7'-MorphoLuc and 4'-MorphoLuc) 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.

## **5.5 Evolving substrate-specific luciferases**

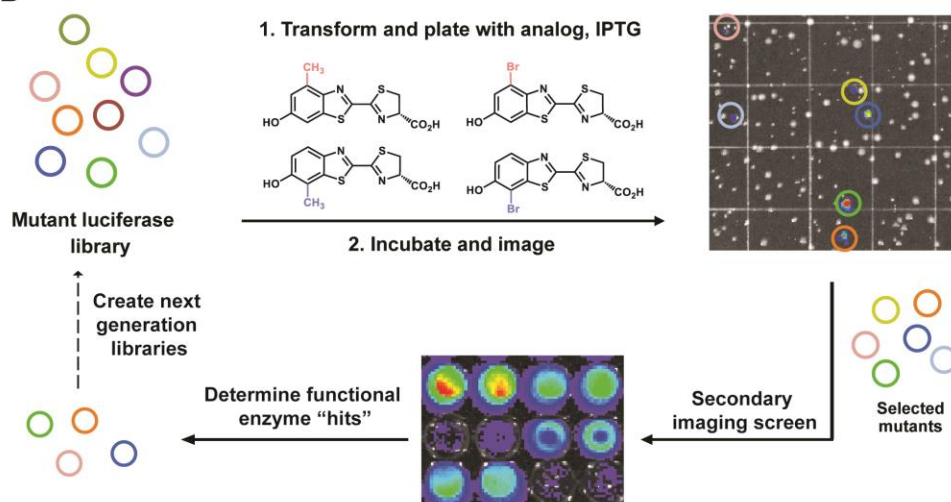
Having prepared candidate orthogonal luciferins, we set out to identify mutant luciferases that could selectively process the molecules. Predicting enzyme mutations that confer substrate selectivity or otherwise beneficial properties is challenging. Fluc is a highly dynamic enzyme, complicating the selection of residues from static structural or sequence data (31, 52). Moreover, amino acids known to play key roles in enzyme function have been identified far from the luciferin binding site; such critical residues are often revealed only by random mutagenesis approaches (22, 31, 52-54). Screening libraries of completely random mutants was impractical in our case, though, owing to the large library sizes needed to achieve adequate enzyme coverage (55). 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.

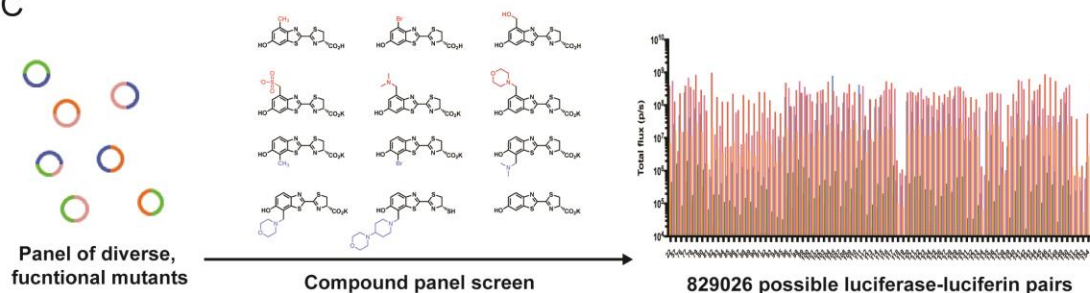
A



B

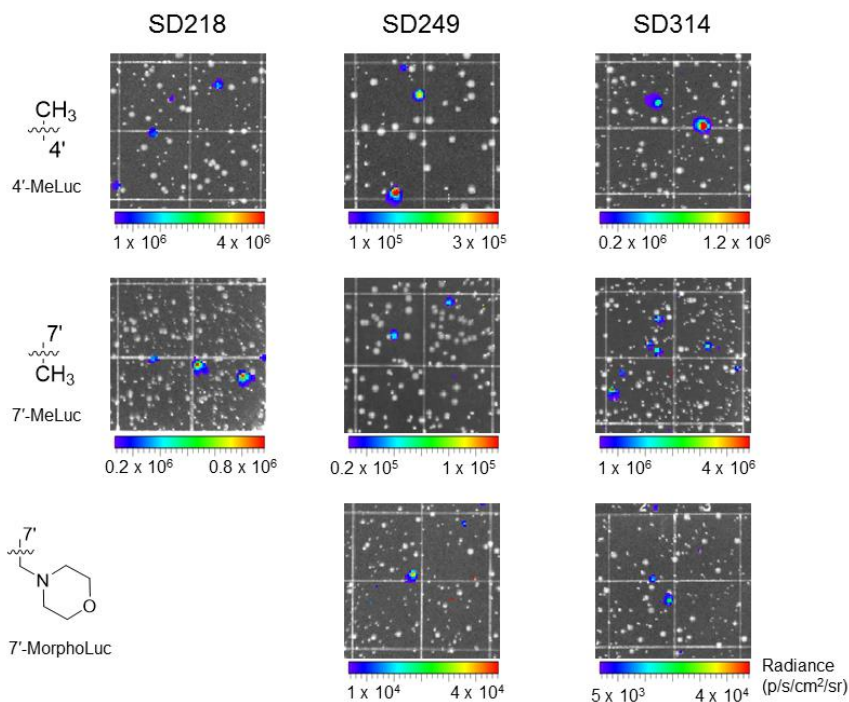


C



**Figure 5-6.** 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 to develop next-generation libraries. (C) "Hits" from next-generation libraries were subjected to a secondary screen for orthogonality with panel of luciferin analogs.

Recognizing that manual screening necessitated the use of smaller libraries, we developed focused, semi-rational libraries where the mutations were confined to regions known to modulate substrate binding (56). “Hits” from these smaller, individual libraries could then be easily combined and assayed in subsequent library generations for improved function. We initially targeted residues proximal to the active site for mutagenesis (Figure 5-6A). These selections were partially based on phylogenetic data gathered from across the insect luciferase family, along with previous biochemical assays (11, 57). Mutations at many of these target sites have been shown to perturb D-luciferin binding (and thus light emission), while preserving the overall structural integrity of the enzyme (7, 29, 58).

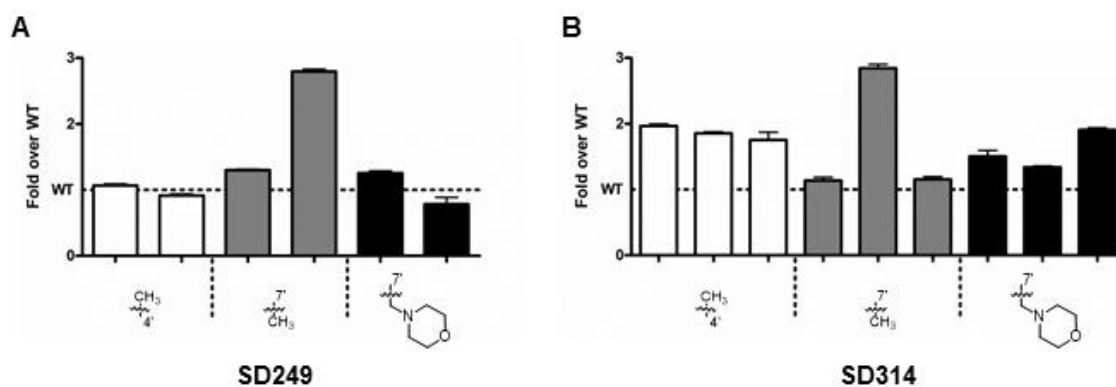


**Figure 5-7.** Representative images of luciferin analogs screened on plate. Mutant luciferase libraries (SD218, 249, and 314) were introduced into bacteria, and the transformants were spread on agar plates containing the specified luciferins. The plates were imaged and light-emitting colonies were selected for further expansion, screening, and analysis.

Saturation mutagenesis was used to prepare the desired libraries. The degree of mutation applied at each residue was based on the following considerations: sequence conservation among

the insect luciferase family, the identity of the native residue, and the location of the residue. For example, non-conserved residues were mutated to a higher degree compared to conserved residues in the active site. Codon compression methods were further used to eliminate redundancies and reduce the number of transformants (59). The final libraries ranged from 19-19600 members in size, and were constructed using synthetic gene assembly in combination with circular polymerase extension cloning (CPEC) (60, 61).

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 5-6B). Light-emitting colonies were easily identified (Figure 5-7) and, in some cases, the light emission values were on par with native Fluc and D-luciferin (Figure 5-8). A handful of the corresponding mutants were sequenced. Some mutations were observed for multiple analogs, suggesting that they might be selective for bulky luciferins (Figure 5-9). Other mutations were unique to each compound, which is notable, given the subtle structural differences between some of the analogs. The number of colonies screened was ~3X the calculated diversity for each library.

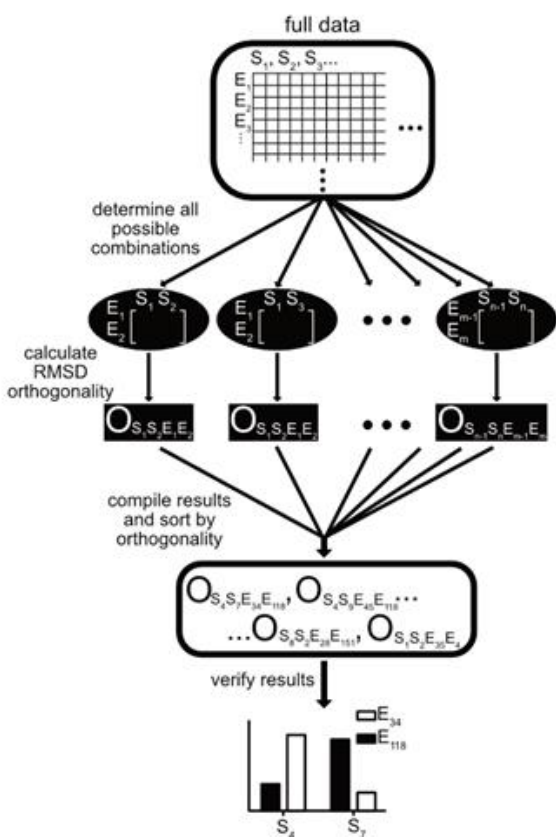


**Figure 5-8.** Some mutant enzymes provide similar or greater photon outputs than native luciferase. Bacterial colonies expressing mutant enzymes or Fluc (WT) were selected from the primary screen (above), expanded, and lysed. Lysates were imaged with the analogs shown, and photon outputs were normalized to Fluc. Sample data are provided for transformants originating in the (A) SD249 or (B) SD314 libraries. Error bars represent the standard error of the mean for  $n = 3$  independent experiments.

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

**Figure 5-9.** Sequencing analysis of “hits” from site-directed libraries. Unique mutations identified in libraries (SD218, SD249, and SD314) screened with the methyl luciferin analogs. Some of the mutations were identified in both compound screens (both). The frequency of the mutation observed is shown in parentheses. If not specified, mutation was only observed once.

The on-plate screen quickly culled non-functional enzymes but did not control for overall expression levels and differences in compound transport. To address these parameters, we performed a secondary screen. Colonies emitting detectable levels of light on-plate were selected and expanded overnight. These cultures were then lysed and imaged with analogs. Mutants that provided light emission on par with native Fluc were identified as bona fide “hits” and used to create next-generation sequences. This iterative process was performed to evolve large pools of diverse, but functional enzymes (Figure 5-6C). “Hits” from these subsequent generations were ultimately tested with all luciferin analogs in secondary screens to report on selective substrate usage (i.e., orthogonality).



**Figure 5-10.** A computer script was used to screen all possible enzyme-substrate pairings *in silico*. An algorithm analyzed the imaging dataset for substrate resolution of orthogonal luciferase-luciferin pairs.

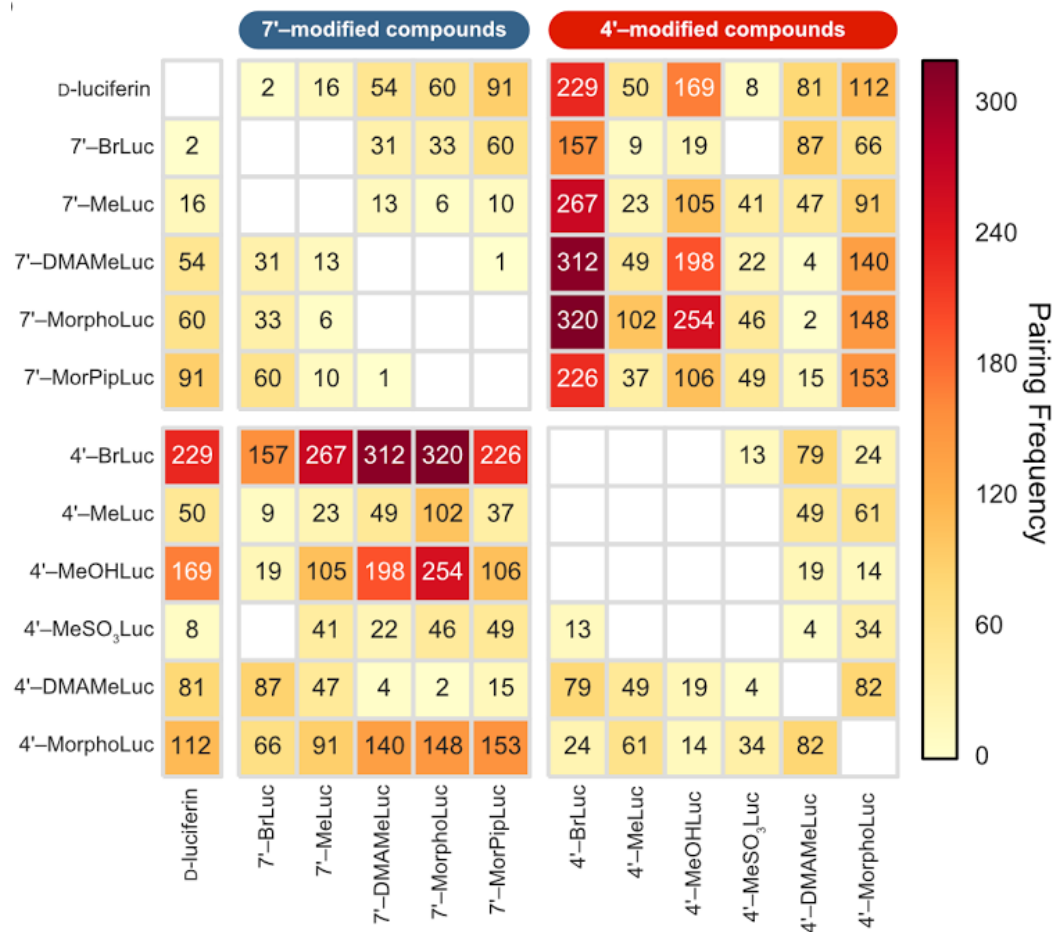
Since the entire collection of imaging data for substrate-selective pairs exceeded 800,000 combinations, we wrote a computer script to rapidly examine all pairs in an unbiased fashion. The algorithm iterated through each possibility and ranked the pairings by orthogonality (Figure 5-10). Orthogonality for each possible pairing was determined by representing each pairing as a set of normalized vectors. Each set was then symmetrized and compared to the identity matrix (the ideal case for orthogonality) via root mean square distance (RMSD). Favorable RMSD scores are obtained when two mutants (e.g., A and B) react with unique substrates (e.g., substrate 1 and 2, respectively), in a mutually exclusive manner. By utilizing this approach, we aimed to



not only determine the top orthogonal pairings but also identify trends in enzyme-substrate selectivity.

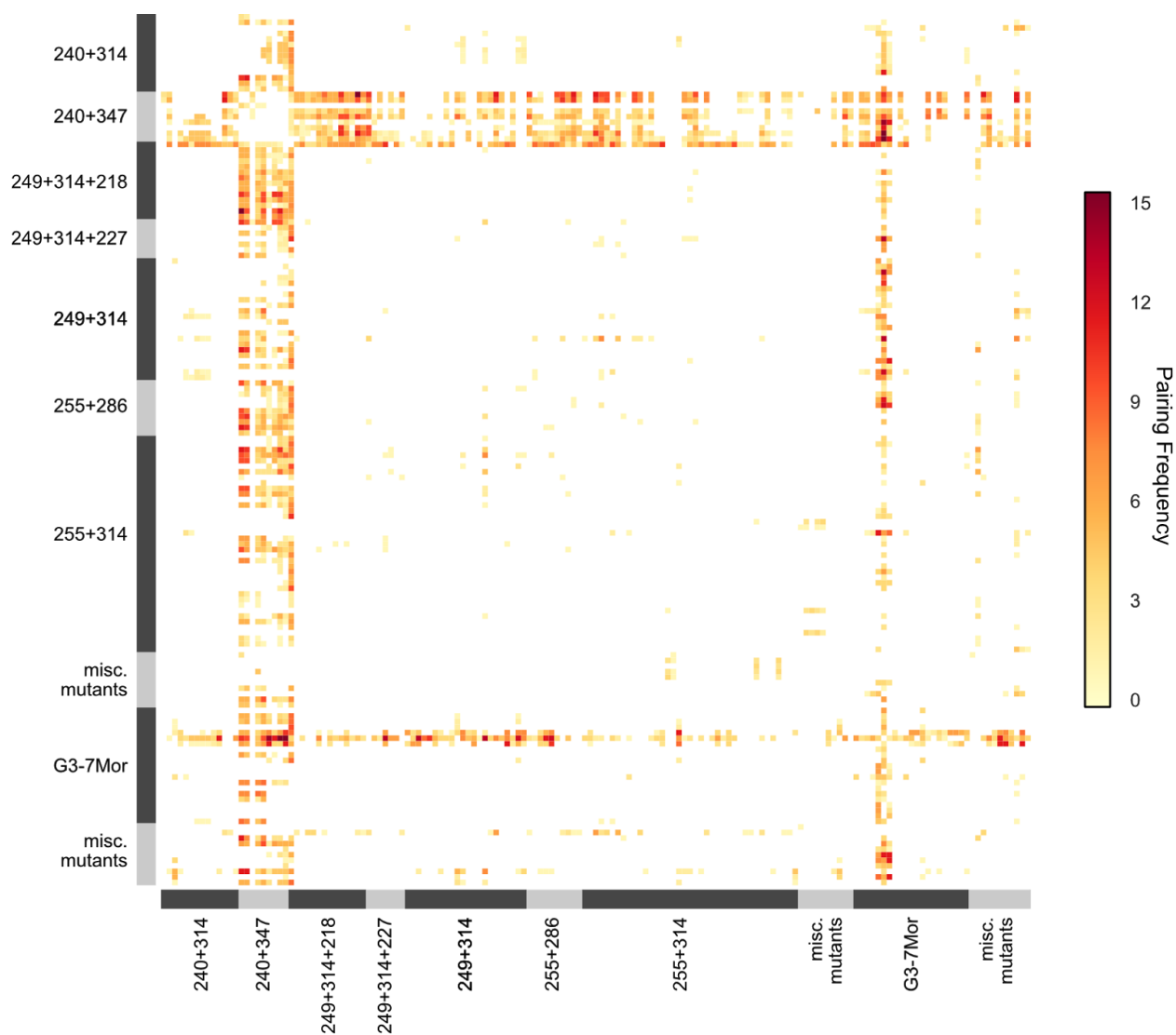
### **5.6 Analyzing trends in top orthogonal pairs identified by computer script**

To analyze the data for orthogonality trends, we restricted our analysis to the top 5,000 algorithm hits. These hits produced at least 10-fold substrate selectivity with their respective enzyme. First, we examined the frequency of substrate-substrate pairings among these hits. Substrates with modifications at the 4' and 7' positions were most commonly paired (Figure 5-11, upper right and lower left quadrants). Matched analogs (i.e., 4'-4' and 7'-7') were less frequently paired suggesting the screening approach selected for “bump-hole” substrate-enzyme interactions (62-64). Substrates with a “bump” (i.e, steric bulk) at the 4' position were rarely orthogonal to other 4' analogs likely due similar substrate-enzyme interactions in the “hole” generated on the protein. In contrast, when 4' and 7'-bumped substrates are paired it is possible for each substrate to interact differently with the enzyme resulting in orthogonality. These results are informative for determining substrates for on-plate screens and for future luciferin analog design.



**Figure 5-11.** Substrate-substrate pairing frequency in the top 5,000 *in silico* screen hits. Compounds of differing substitution (4' and 7') are paired more frequently than those with substitution at the same site.

Based on these results, we conducted a similar analysis with frequency of enzyme pairings to determine if certain residues correlate with orthogonality (Figure 5-12). Two distinct bands were present in the heatmap corresponding to all enzymes from the 240+347 library and three enzymes from the G3-7Mor (249+314+random mutagenesis) library. These enzymes showed high pairing frequency to form orthogonal sets with a majority of the enzymes analyzed. While there are other isolated library-library pairings, the predominance of these two libraries suggests the enzymes contain important mutations for analog selectivity. Further library design or mutagenesis at these positions could afford improvements in substrate resolution.

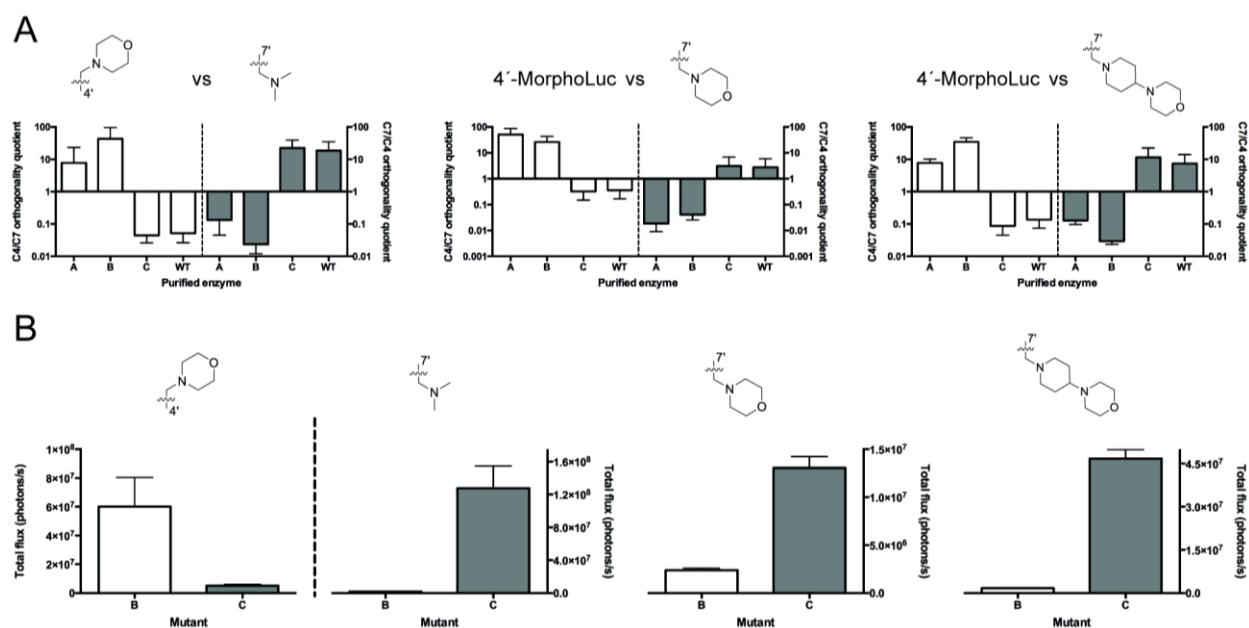


**Figure 5-12.** Enzyme-enzyme pairing frequency in the top 5,000 *in silico* screen hits. Each square represents a pairing of two enzymes. Certain libraries (240+347 and G3-7Mor) pair frequently with all other mutants.

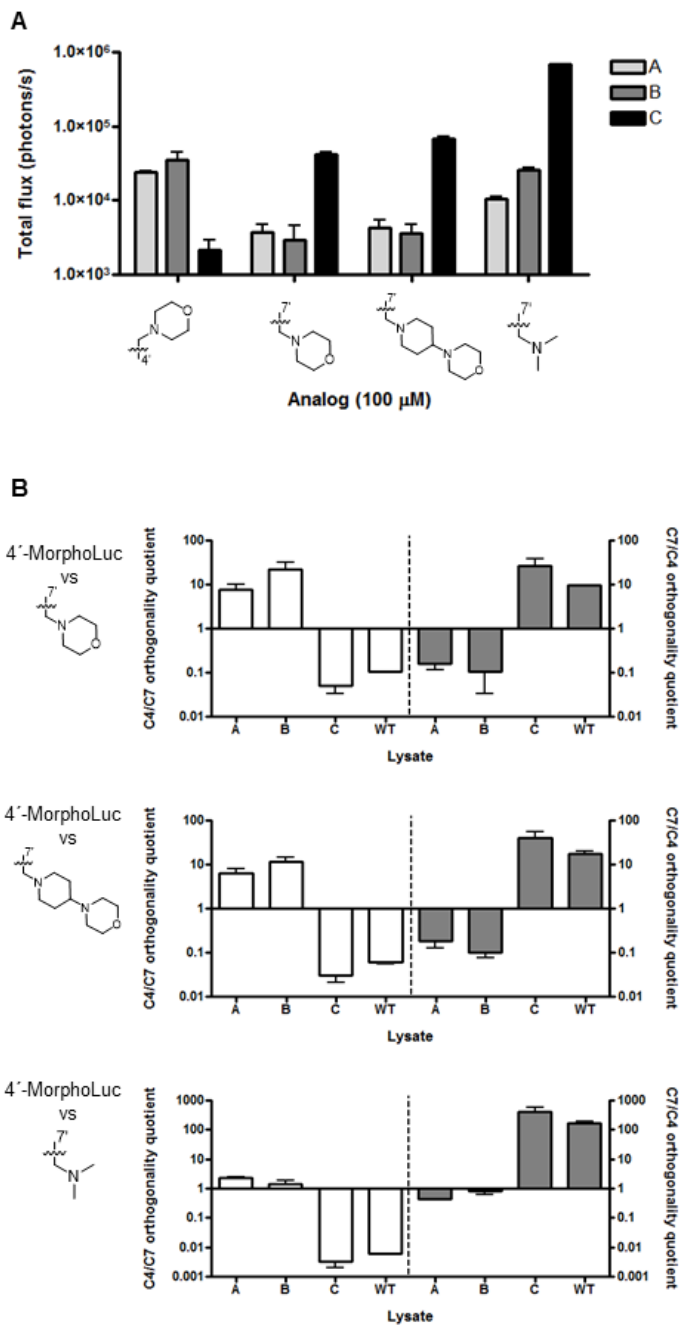
### 5.7 Characterization of a lead orthogonal pair

One top pair identified by the script exhibited selectivity for analogs 4'-MorphoLuc (mutant A and B) and 7'-DMAMeLuc, 7'-MorphoLuc, or 7'-MorPipLuc (mutant C). The magnitude of each mutant's preference—defined as the orthogonality quotient—was analyzed.

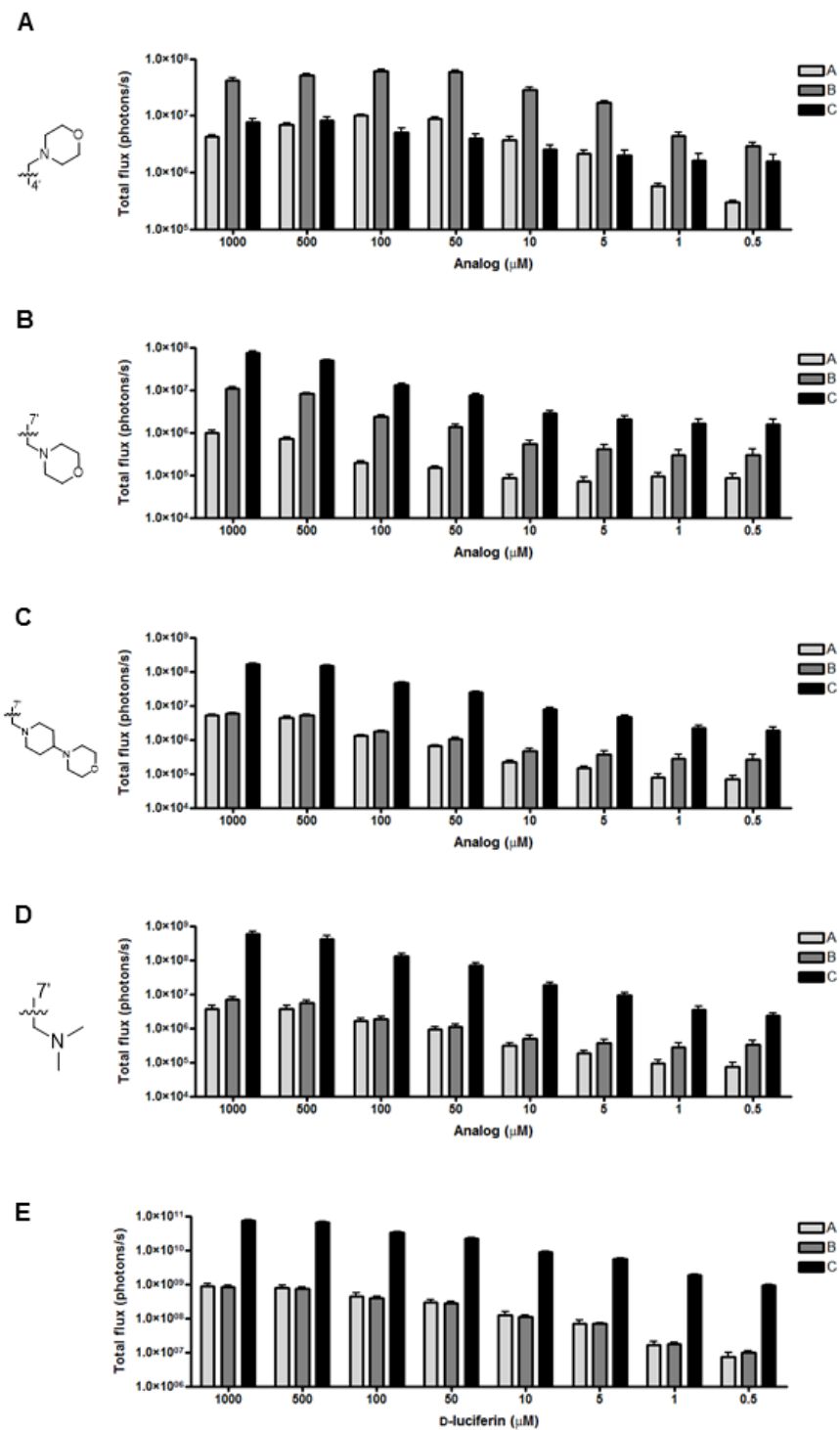
As shown in Figures 5-13A and 5-13B, mutants A and B exhibited nearly a 100-fold preference for 4'-MorphoLuc over other analogs, while mutant C strongly favored C7' modified analogs. Similar trends in orthogonal substrate usage were observed using bacterial lysates (Figures 5-14) and across a range of luciferin concentrations (Figure 5-15 and 5-16). Biochemical analyses further indicated that the “brightest” mutant enzymes with their matched substrate showed the most efficient substrate turnover (Table 5-1).



**Figure 5-13.** Analyzing orthogonal enzyme-substrate pairs. Purified mutants exhibit orthogonality. Enzyme (1  $\mu\text{g}$ ) was incubated with 100  $\mu\text{M}$  of luciferin analogs and emission intensities were used to determine the orthogonality quotient (the ratio of the total flux for the C4/C7 or C7/C4 pairings). The geometric mean is plotted and the error bars represent the 95% confidence intervals for  $n > 4$  experiments. (C) Total flux for lead mutants B and C highlights substrate selectivity between C4' and C7' sterically modified luciferins. Error bars represent the standard error of the mean for  $n > 4$  experiments

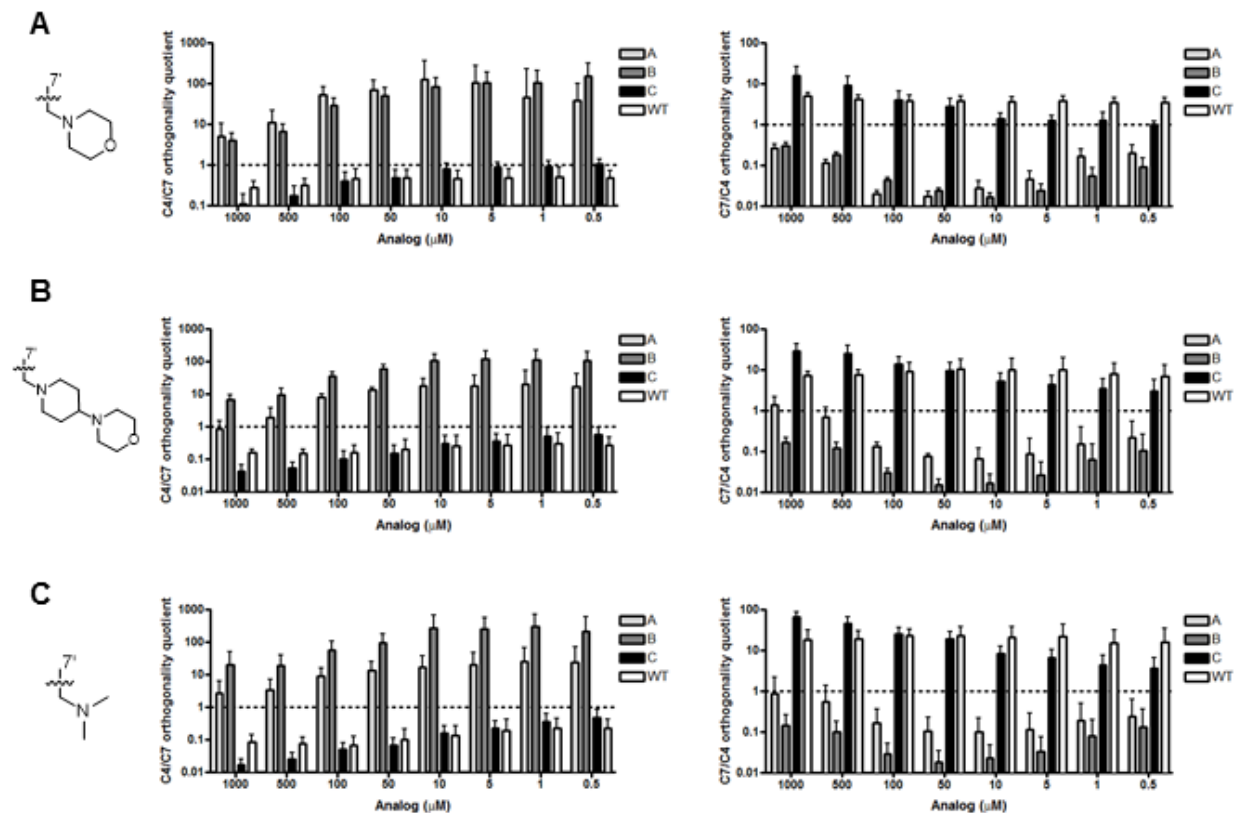


**Figure 5-14.** Orthogonal substrate usage in mutant or Fluc (WT) lysate. (A) Lysates were imaged with the analogs shown and total bioluminescent light emission is shown. (B) Orthogonality plots were constructed from data in (A) to compare the indicated analogs with C4'-MorphoLuc. The geometric mean is plotted and the error bars represent the 95% confidence interval for  $n = 3$  experiments.



**Figure 5-15.** Bioluminescent photon production from luciferin analogs and luciferase enzymes. (A)-(D) Luciferin analogs (0.5 – 1000  $\mu\text{M}$ ) or (E) D-luciferin were incubated with mutant luciferases (A-C) in bioluminescence buffer. Images were acquired as described in the Materials and Methods section. Error bars represent the standard error of the mean for  $n > 4$  experiments.

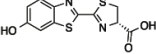
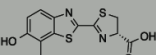
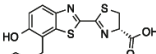
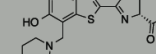
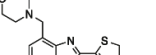
The identities of the mutant “hits” provided some insights into the origins of substrate orthogonality. Mutant A had a single arginine to alanine mutation at amino acid 218. Mutant B comprised the same R218A mutation, but harbored additional mutations at residue 250 (Phe to Met), 314 (Ser to Thr), and 316 (Gly to Thr). These residues are known to play a role in modulating binding and interaction with the luciferin substrate. The R218A mutant is especially interesting, as it is known to greatly reduce light production and red shift emission with D-luciferin (28). It has been hypothesized that the smaller Ala group allows more water molecules to access the active site, potentially quenching light emission (28). The bulky morpholino substituent of 4'-MorphoLuc 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 analogs (21). The improved selectivity with 7'-DMAMeLuc, 7'-MorphoLuc, and 7'-MorPipLuc could be the result of active site positioning. The C7'-substituents could potentially place the luminophore in a more advantageous spot for light emission.



**Figure 5-16.** Comparative analyses of all analogs with mutants A-C and wild type (WT) luciferase. Analogs (0.5 – 1000  $\mu\text{M}$ ) were incubated with mutant luciferase enzymes (A)-(C) or Fluc (WT), and images were acquired. Orthogonality plots were constructed to compare C4'-MorphoLuc with the C7' steric analogs. C4/C7 orthogonality quotient (left) C7/C4 orthogonality quotient (right) indicates preference for C4' or C7' substrates. The geometric mean is plotted and the error bars represent the 95% confidence interval for  $n > 4$  experiments.



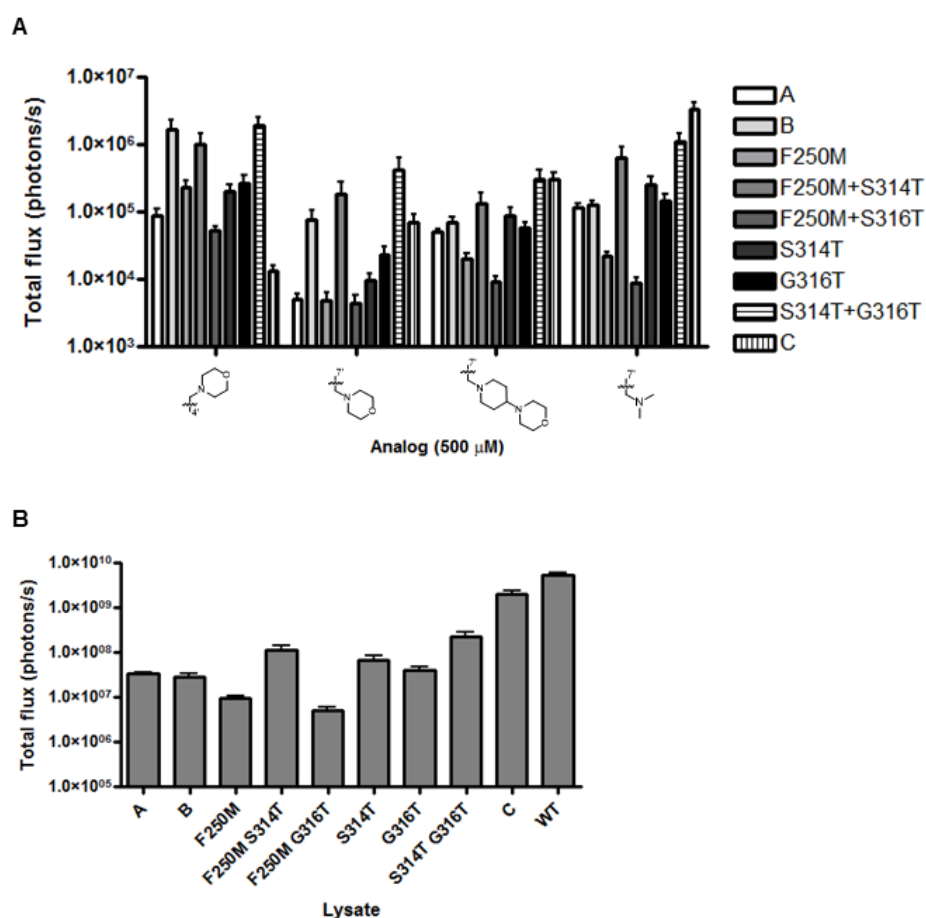
**Table 5-1.** Biochemical analyses of orthogonal enzyme-substrate pairs.

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	 7'-DMAMeLuc
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	 7'-MorphoLuc
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>	 7'-MorpPipLuc
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	 4'-MorphoLuc
B	<b>200 ± 41</b>	<b>83 ± 37</b>	<b>628</b>	
C	13 ± 2	13.1 ± 5.7	626	

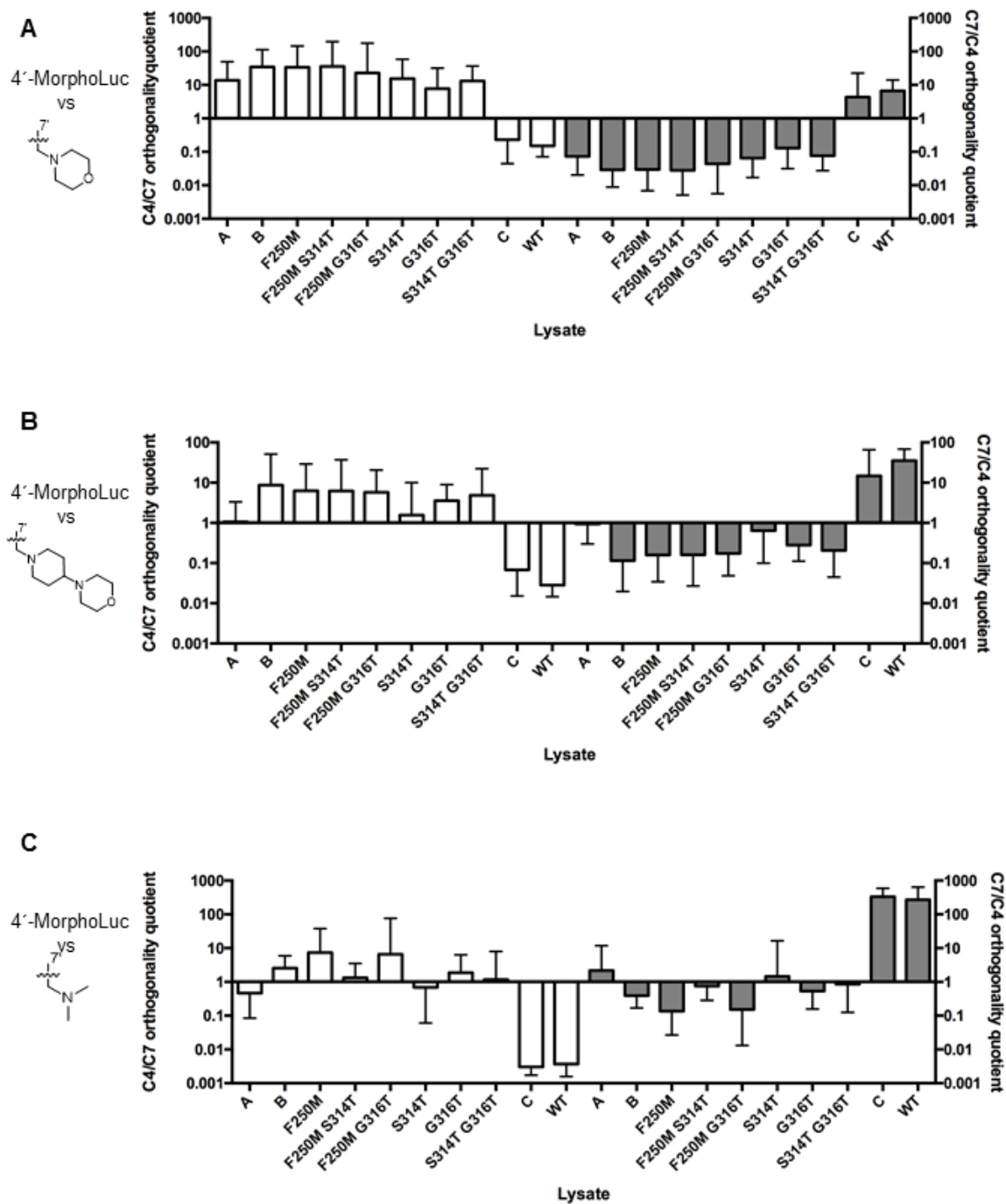
<sup>a</sup> Values normalized to each compound's corresponding emission with WT Fluc. Errors represent standard error of the mean for  $n = 3$  measurements. <sup>b</sup> Kinetic constants are apparent values, determined via measurements of initial rates of light emission over a range of 2  $\mu$ M to 10 mM. Errors represent standard error of the mean for  $n \geq 3$  measurements.  $K_{cat}$  values are relative to each compound's corresponding value with WT 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.

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 (Figures 5-17 and 5-18). Both mutations should result in a larger active site, but why they preferentially accommodate 4'-MorphoLuc over other analogs remains unknown. It is possible that the mutations disrupt critical binding interactions with the luciferin core, but that steric appendages (e.g., on the C4' side) retain sufficient contacts for subsequent oxidation. Indeed, when 4'-MorphoLuc was incubated with R218A/F250M, light

emission was maintained (as compared to Fluc, Figure 5-18C). When D-luciferin and the C7'-modified analogs were incubated with this same mutant, though, light emission was drastically reduced. Interestingly, the R218A/S314T mutant exhibited an opposite trend in analog selectivity: 7'-DMAMeLuc was preferred to 4'-MorphoLuc (Figure 5-18C). 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.



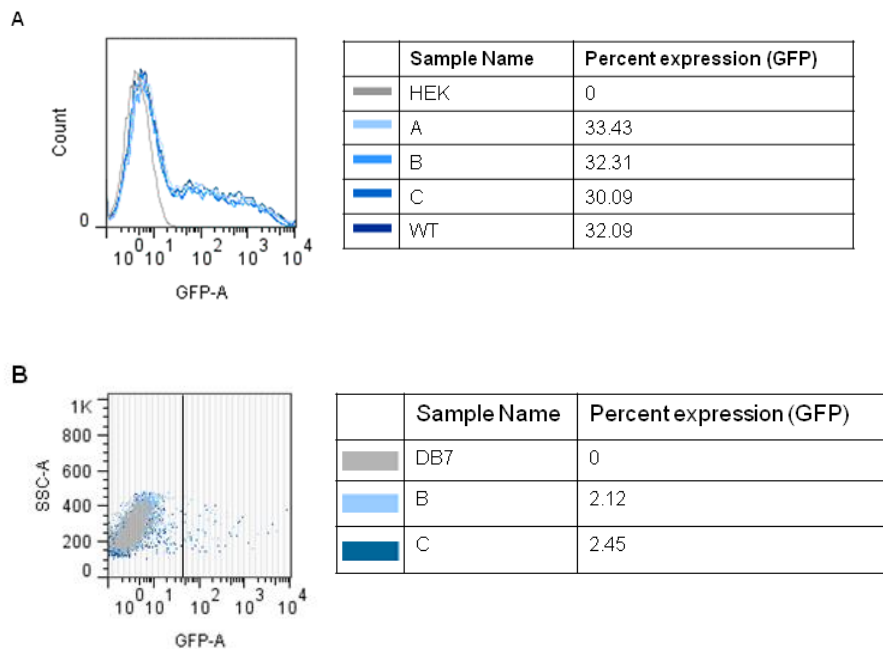
**Figure 5-17.** Bioluminescent photon production from luciferin analogs with combinatorial enzymes. Combinatorial mutants were prepared, in which R218A was held constant across the library, while the other positions were allowed to code for mutations from mutant B or native Fluc (WT). (A) Luciferin analogs (500  $\mu$ M) or (B) D-luciferin were incubated with lysates expressing mutant luciferases or Fluc (WT) in bioluminescence buffer. Images were acquired as described in the Materials and Methods section. Error bars represent the standard error of the mean for  $n > 4$  experiments.



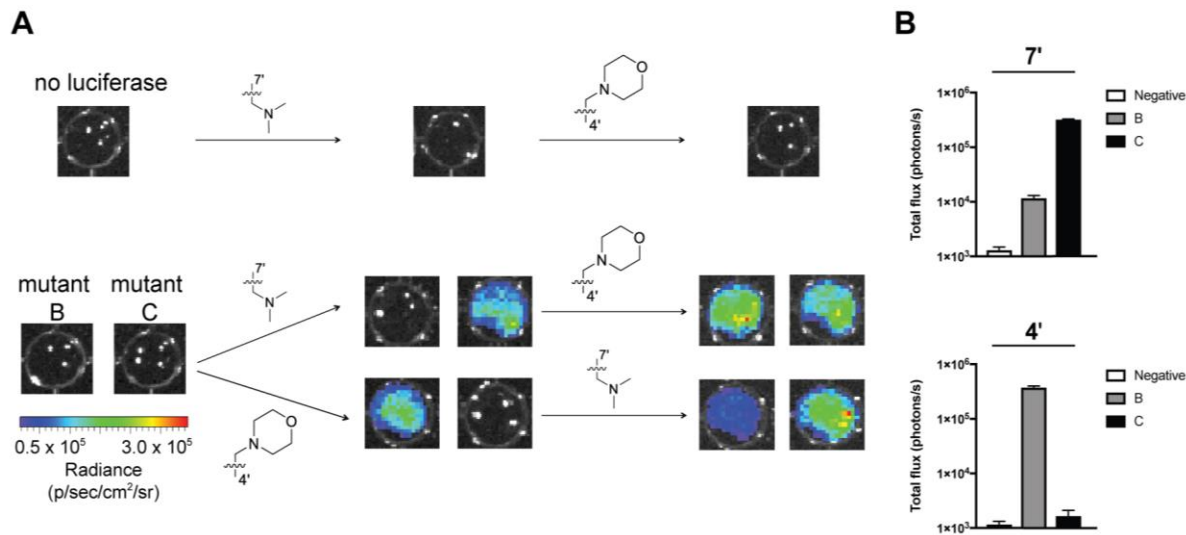
**Figure 5-18.** Comparative analyses of combinatorial enzymes and Fluc. (A)-(C) Luciferin analogs (500  $\mu$ M) were incubated with lysates expressing the indicated mutant enzymes. Images were acquired and orthogonality plots were constructed to compare C4'-MorphoLuc with the C7' steric analogs as described in the Materials and Methods section. The geometric mean is plotted and the error bars represent the 95% confidence interval for  $n > 4$  experiments.

## 5.8 Cellular imaging with orthogonal pairs

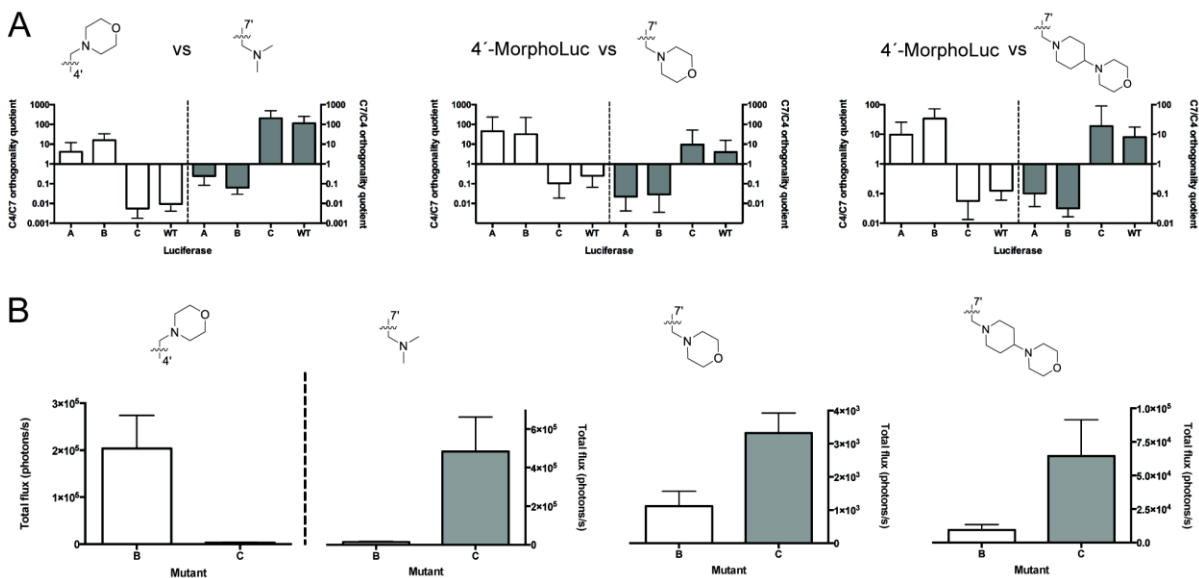
As a step toward multi-component imaging applications, we evaluated the orthogonal enzymes and probes in cultured cell models. Mammalian cell lines (HEK293 and DB7) were engineered to express orthogonal mutants A-C. Equivalent expression levels were confirmed using flow cytometry (Figure 5-19). Cells were then incubated with analogs 7'-DMAMeLuc, 7'-MorphoLuc, 7'-MorPipLuc, and 4'-MorphoLuc, and photon outputs were measured. As shown in Figure 5-20, the substrates were able to cross cell membranes and access the relevant luciferases, resulting in sustained emission. Photon production was also confined to cells expressing the complementary luciferase for each orthogonal luciferin: cells with mutant B were only visible upon treatment with analog 4'-MorphoLuc, while cells with mutant C were only visible upon treatment with analog 7'-DMAMeLuc, 7'-MorphoLuc, and 7'-MorPipLuc (Figure 5-21 to 5-24). Importantly, the orthogonal pairs could also distinguish unique cell types in a single imaging session. For example, DB7 cells stably expressing mutants B or C could be readily detected via sequential administration of the requisite substrates (Figure 5-20). Similar trends were observed upon imaging HEK293 cells (Figure 5-25) and co-cultures (Figure 5-26). These data suggest that cross-reactivity between mutants B and C and their non-orthogonal substrates is minimal.



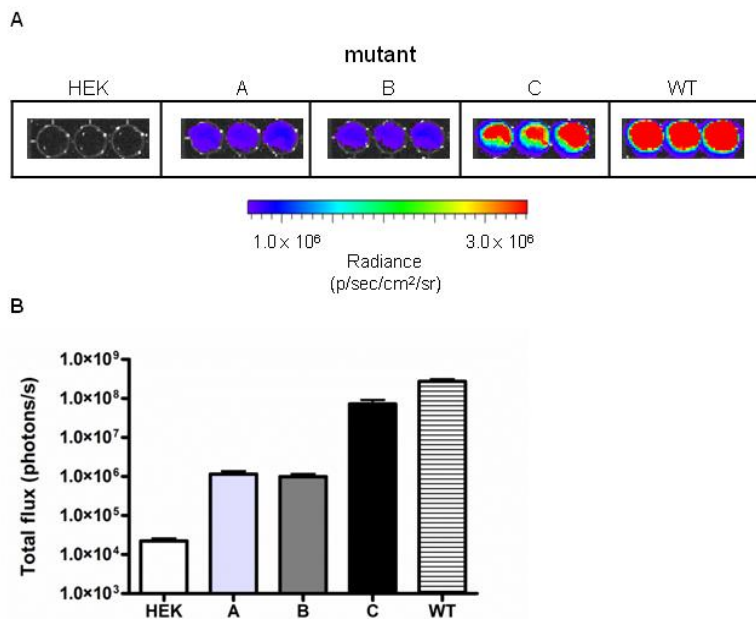
**Figure 5-19.** Mutant luciferases are expressed at comparable levels in mammalian cells. (A) HEK293 cells or (B) DB7 cells transiently expressing mutant or wild type luciferases were rinsed in FACS buffer and analyzed by flow cytometry. GFP expression was measured. Data shown are representative of (A)  $n = 6$ ; and (B)  $n = 2$  replicate experiments.



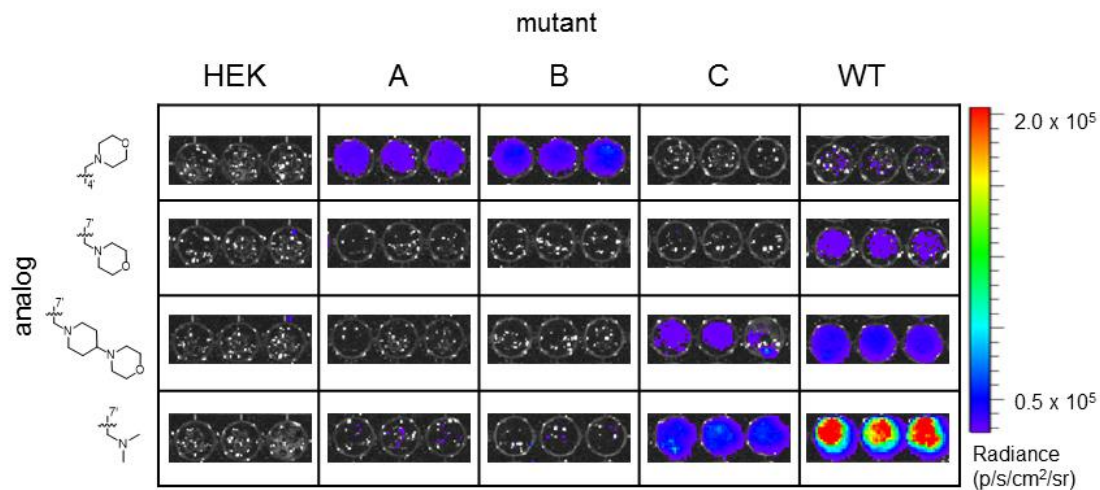
**Figure 5-20.** Imaging cells with orthogonal luciferase-luciferin pairs. (A) Mutant luciferase expressing DB7 cells were plated ( $1.5 \times 10^5$  cells/well) in 96-well black plates and sequentially incubated with C4' and C7' sterically modified luciferins ( $750 \mu\text{M}$ ). Representative bioluminescence 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.



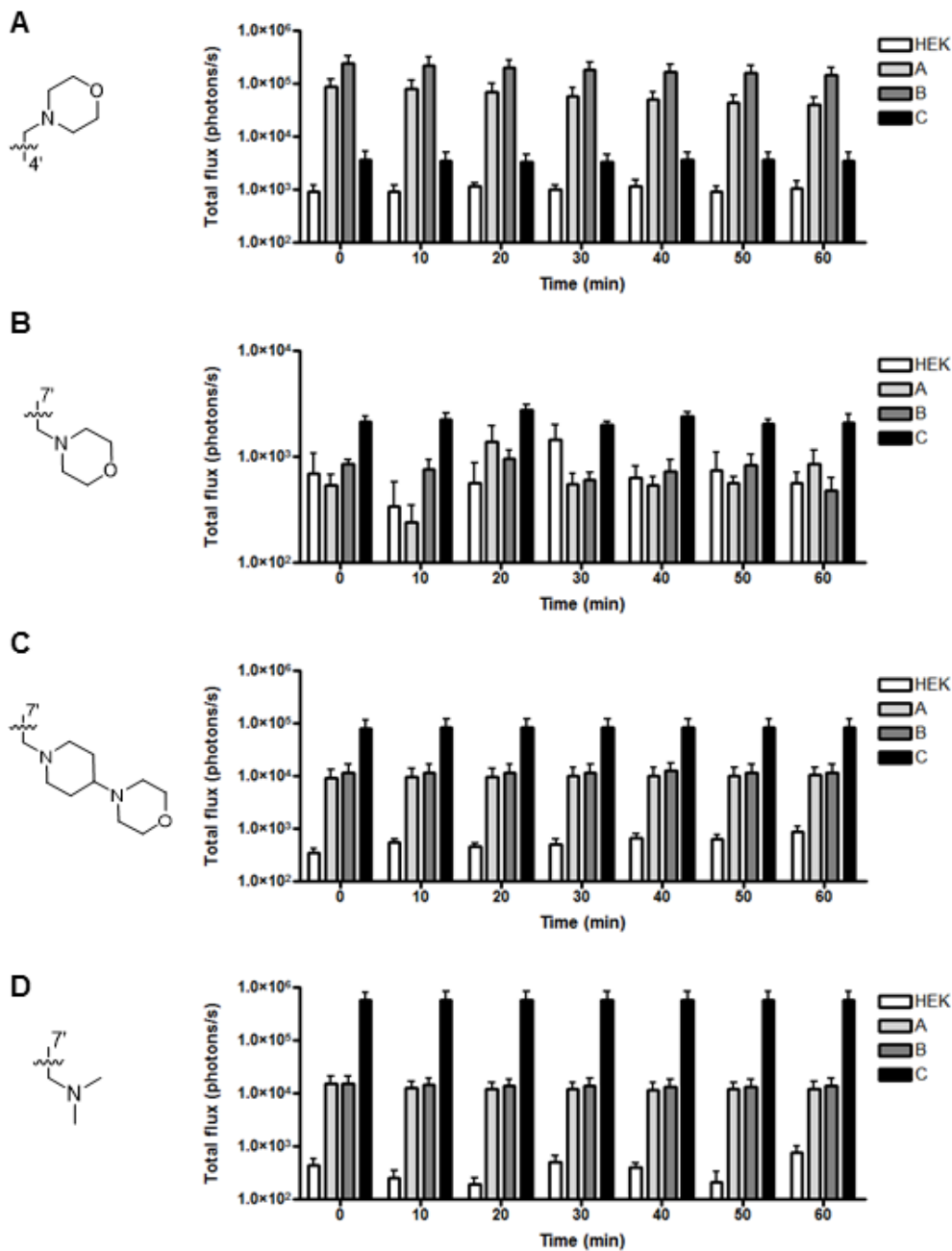
**Figure 5-21.** Luciferase-expressing cells exhibit orthogonality. (A) HEK293 cells ( $2 \times 10^5$ ) expressing wild type or mutant luciferases (A-C) were incubated with 500  $\mu$ M of luciferin analogs and the orthogonality quotient was determined as before. The geometric mean is plotted and the error bars represent the 95% confidence interval for  $n > 4$  experiments. (B) Substrate selectivity between C4' and C7' sterically modified luciferins with lead mutants B and C is maintained in mammalian cells. Error bars represent the standard error of the mean for  $n > 4$  experiments.



**Figure 5-22.** Cellular imaging with mutant luciferases and D-luciferin. (A) HEK293 cells expressing Fluc (WT) or mutant luciferases (A-C) were incubated with D-luciferin (500  $\mu$ M). Representative bioluminescence images are shown. (B) Quantification of the images provided in (A). Error bars represent the standard error of the mean for  $n = 4$  experiments.

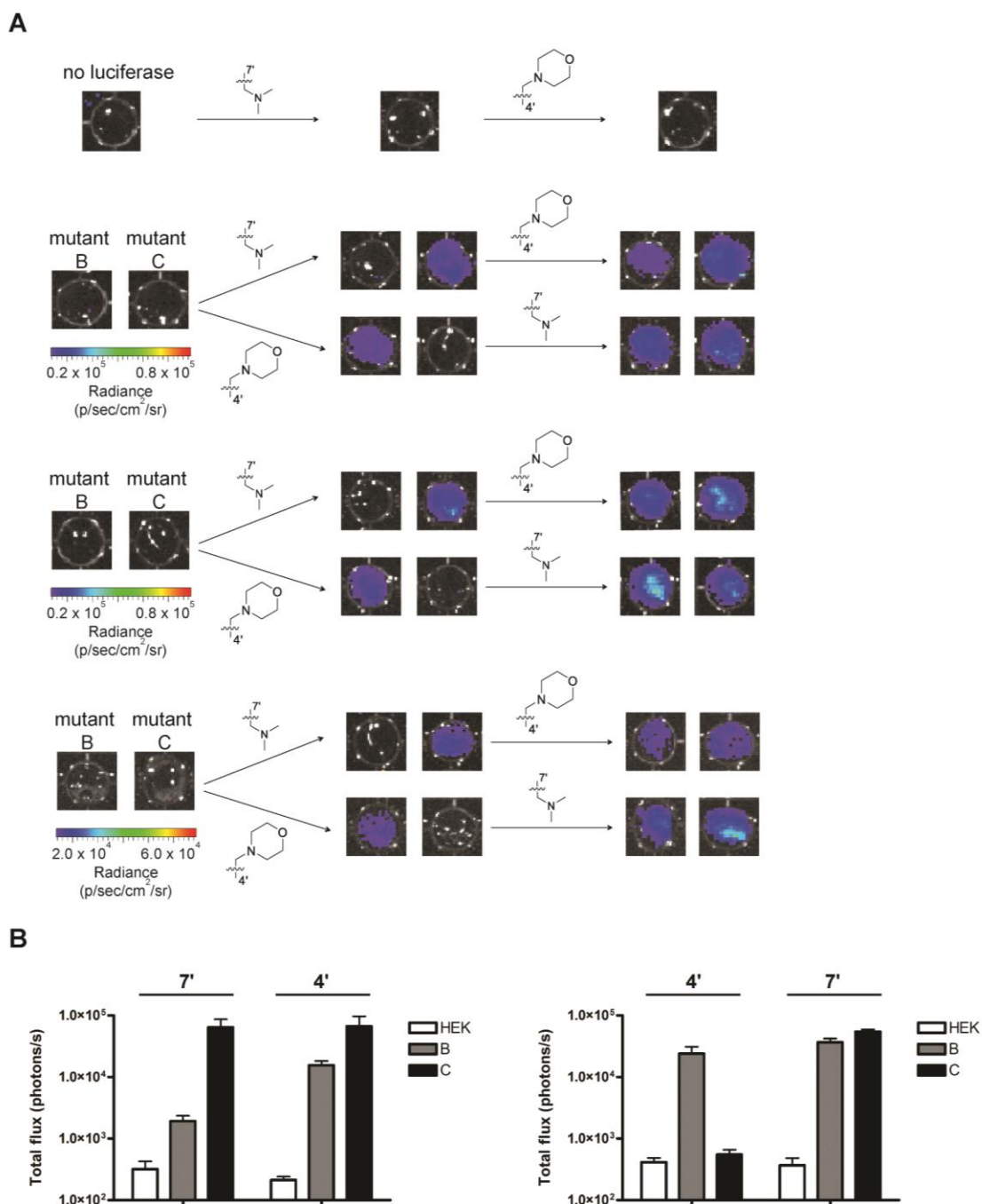


**Figure 5-23.** Cellular imaging with mutant luciferases and luciferin analogs. HEK293 cells expressing Fluc (WT) or mutant luciferases (A)-(C) were incubated with the indicated luciferin analogs (500  $\mu$ M). Representative bioluminescence images are shown.

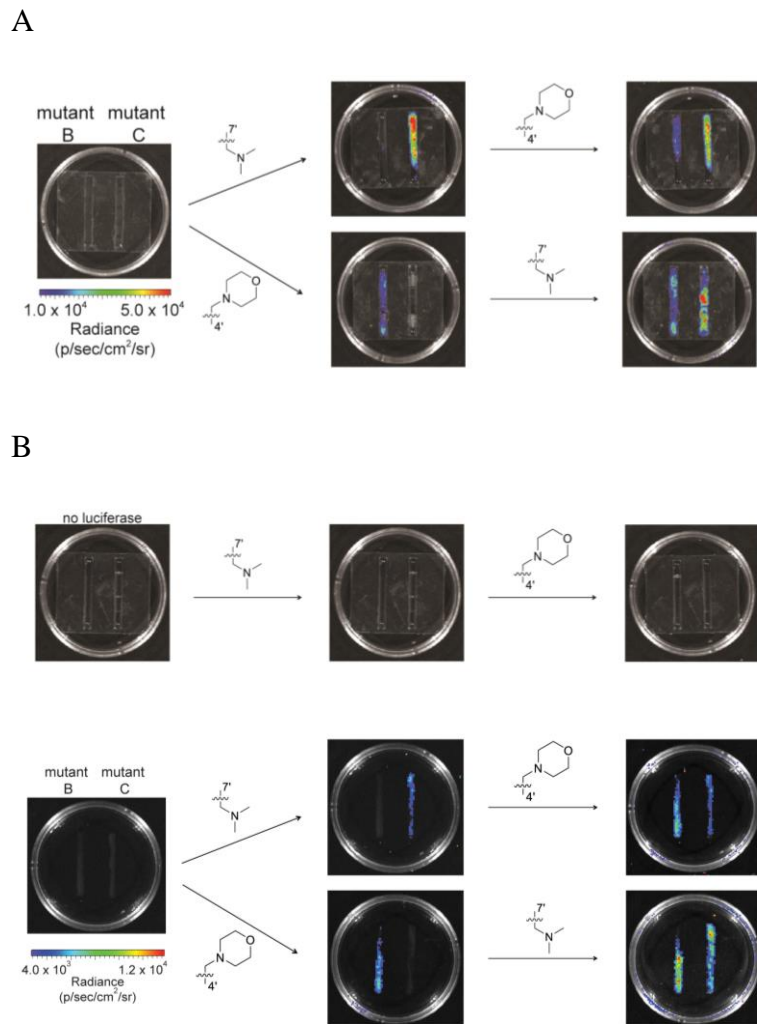


**Figure 5-24.** Cellular bioluminescent photon production is sustained. (A)-(D) HEK293 cells (200,000 cells/well) expressing mutant luciferases (A)-(C) were incubated with 500  $\mu$ M of the indicated luciferin analog. Images were acquired (0-60 min), and light emission was quantified. Error bars represent the standard error of the mean for  $n = 3$  experiments.





**Figure 5-25.** Cellular imaging with orthogonal pairs. (A) HEK293 cells expressing mutant B or C (or no luciferase) were imaged with the indicated analog over 10 min. The same cells were then immediately treated with the second orthogonal luciferin and re-imaged. Sample bioluminescent images from 3 independent experiments are shown. (B) Quantification of the images provided from (A). Error bars represent the standard error of the mean for  $n = 3$  experiments.

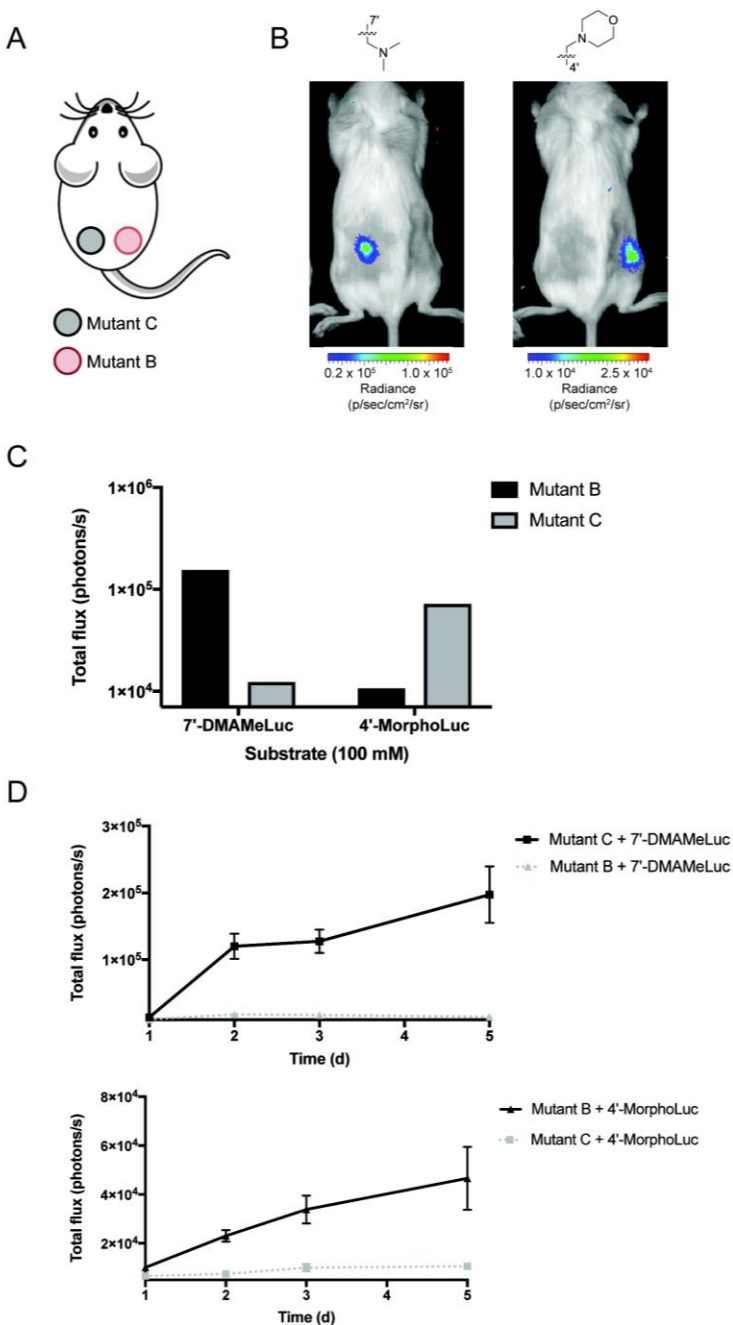


**Figure 5-26.** Cellular imaging with orthogonal pairs patterned with biocompatible stencils. (A) HEK293 or (B) DB7 cells expressing mutant B or C (or no luciferase) were plated with biocompatible stencils and imaged with the first analog (as shown). The same cells were then immediately treated with the second analog and re-imaged. Sample bioluminescent images from (A) = 3 and (B) = 2 independent experiments are shown.

### 5.9 *In vivo* imaging with orthogonal pairs

The level of orthogonality achieved in our cultured cell models suggested that orthogonal luciferase-luciferin pairs could be distinguished *in vivo*. To evaluate mutant B and C for multi-component imaging we transitioned the orthogonal pairs into mouse models. DB7 cells ( $5 \times 10^5$  cells per graft) expressing mutant B or C were implanted at opposite lower flanks of FVB/NJ

mice. Mice were imaged after intraperitoneal injection of analog 4'-MorphoLuc. Photon production was only observed at the site of mutant C. After 5 h, the mice were imaged to confirm signal had returned to background levels before administering analog 7'-DMAMeLuc. After injection of 7'-DMAMeLuc, signal was only detected at the site of mutant B. Based on the number of cells implanted, we estimate that approximately  $1 \times 10^6$  cells are necessary for detection of the orthogonal pairs. Orthogonality was maintained over the five day imaging period, an important criteria for imaging biologically relevant processes.



**Figure 5-27.** *In vivo* imaging with orthogonal pairs. (A) Diagram for mice inoculations with DB7 cells expressing mutant B (red circles) or mutant C (gray circles). (B) Mice were imaged immediately after intraperitoneal injection of 7'-DMAMeLuc (100mM, left). After 5 h, mice were imaged to confirm signal had returned to background levels before administration of 4'-MorphoLuc (100 mM, right). Representative images are shown for imaging at day 5 post-implantation. (C) Quantification of signal from (B). (D) Orthogonality is maintained over time. After each compound (top: 7'-DMAMeLuc, bottom: 4'-MorphoLuc) was administered, each cell implantation site was analyzed for total flux. Error bars represent the standard error of the mean for  $n = 3$  mice.

## 5.10 Conclusions and future directions

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

## **5.11 Materials and methods**

### **5.11a General biological methods**

#### *Fluorescent spectra and assays*

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

#### *Bioluminescence emission spectra*

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

#### *UV-Vis spectra*

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

#### *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.

#### *General chemiluminescence procedure*

Carried out as described previously (50).

### **5-11b Mutant luciferase library design**

Table S1: Summary of site-directed library identities, sizes, and relative phylogenetic conservation at each residue. Codon sets defined as in Reetz, *et al.* and Kille, *et al.* (54-56). Conservation designations based on 13 aligned insect luciferase sequences (11).

Library	Residue	Conserved	Codon set	Nucleic Acid Diversity	Amino Acid Diversity
SD249	M249	moderate	Hydrophobic	3388	2800
	F250	low	22c-trick		
	T251	moderate	22c-trick		
SD314	S314	moderate	22c-trick	5808	4800
	G315	high	NDT		

	G316	mod-high	22c-trick		
SD218	R218	high	All besides WT	19	19

Table S2: Definition of codon sets from Table S8.

Name	Codons used	Total codons	Total amino acids	Reference
22c-trick	NDT, VHG, TGG	22	20	Kille <i>et al.</i> (3)
NDT	NDT	12	12	Reetz <i>et al.</i> (7)
Hydrophobic*	GBC, WTK	7	7	Pines <i>et al.</i> (8)
All – Phe*	TGK, VBC, ATG, NAW	20	19	Pines <i>et al.</i> (8)
All – Arg*	KGG, WKT, VHG, NAT	19	19	Pines <i>et al.</i> (8)

\*Identified using algorithm provided by Pines *et al.* (59)

#### *Mutant luciferase library generation*

Two sections of the luciferase gene (*pgl4-luc2*), denoted R1 and R2, were targeted for gene assembly. The R1 region comprises amino acids 199-275 and was assembled with primers R1-F0 to R1-F-235 (forward primers, Table S1) and R1-R0 to R1-R119 (reverse primers, Table S2). The R2 region comprises amino acids 275-346 and was assembled with primers R1-F0 to F264 (forward) and R1-R0 to R228 (reverse). The gene assembly primers were designed using gene2oligo to identify desirable  $T_m$  values (see Tables S3 and S4) (2).

To assemble mutant libraries, primers containing the codon(s) of interest (primer tables S5-7) were used in place of the primer coding for the wild-type sequence. Libraries and primers are named “SD” (site directed), followed by the number of the first residue mutated in the library (e.g., SD218 is the library based on R218). Primers with degenerate codons were mixed in proportion to the number of amino acids encoded by the primer (See Kille, *et al.* for a detailed



protocol) (55). Libraries were assembled and all PCR reactions were run using Q5 Hot-start DNA polymerase (New England BioLabs).

#### *Circular polymerase extension cloning (CPEC)*

The libraries generated above were ultimately inserted into a vector lacking the R1 and R2 regions. This deletion vector was generated from pET28a-luc2 using overlapping PCR. Primers Nco-Luc2-For and R1-deletion Rev (PCR1), along with primers R1-deletion-For and Luc2-Not-Rev (PCR2) were used to remove R1. The R2 region was similarly removed using primers Nco-Luc2-For and R2-deletion-Rev (PCR1) and primers R2-deletion-For and Luc2-Not-Rev (PCR2). The PCR1 and PCR2 products were assembled using Nco-Luc2-For and Luc2-Not-Rev (Table S8). The resulting gene products were digested with *NcoI-HF* and *NotI-HF* (New England BioLabs) and inserted into pET28a with T4 ligase (New England BioLabs). The resulting vectors were labeled pET28-R1del and pET28-R2del, referring to the R1 and R2 deleted regions, respectively.

Library DNA was ultimately inserted into pET28-R1del or pET28-R2del (5). The pET28-R1del or pET28-R2del were linearized by amplification using primers R1-vector-Fwd and R1-vector-Rev (for pET28-R1del) or R2-vector-Fwd and R2-vector-Rev for pET28-R2del (Table S9). Q5 Hot start polymerase (New England BioLabs) was used in the amplifications. The DNA was then digested with *DpnI* for 1 h at 37 °C, and the products were purified *via* gel electrophoresis to remove any remaining circular template DNA. The resulting linear vector (500 ng) was combined in a 1:1 or 2:1 molar ratio with library insert DNA (targeting region R1 or R2, described above). The insert and vector were combined via CPEC (6) using the following

conditions: 1x Q5 Hot start DNA polymerase reaction buffer, dNTPs (0.8 mM ), and Q5 Hot start DNA polymerase (1 U) in a total volume of 50  $\mu$ L. The following thermal cycling conditions were used: initial denaturation at 95  $^{\circ}$ C for 60 s; 5-10 cycles of denaturation (95  $^{\circ}$ C, 30 s), slow ramp anneal (72 – 55  $^{\circ}$ C over 2.5 min), and extension (72  $^{\circ}$ C, 132 s). The final step involved extension at 72  $^{\circ}$ C for 5 min. The PCR products were then analyzed using 1% agarose gels and ethidium bromide staining (10  $\mu$ L of the reaction). A portion of each product (2.5  $\mu$ L) was utilized directly for transformation into T7 Express *lysY* competent *E. coli* (New England Biolabs) cells. Transformants were plated according to the primary screening protocol described below.

#### *Primer lists*

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

Table S3: Region 1 (R1, wild type *luc2* primers). Lower case letters denote non-luciferase sequences included for maintaining proper  $T_m$ 's and do not appear in the final assembled gene.

Forward primers	
F0	GGATCCACCGGATTGCCCAAGGGCGTAGCCCTACC
F35	GCACCGCACCGCTTGTGTCCGATTCAGTCATGCCC
F70	GCGACCCCATCTTCGGCAACCAGATCATCCCCGACA
F106	CCGCTATCCTCAGCGTGGTGCCATTTACCACGGC
F141	TTCGGCATGTTACCACGCTGGGCTACTTGATCTGCG
F178	GCTTTCGGGTCGTGCTCATGTACCGCTTCGAGGAG

F213	GAGCTATTCTTGCGAAGCTTcggccctgtacagtgcgtg
F235	cagcgactgtacagggc
Reverse primers	
R0	GGGCAATCCGGTGGATCC
R18	ACAAGCGGTGCGGTGCGGTAGGGCTACGCCCTT
R51	GCCGAAGATGGGGTCGCGGGCATGACTGAATCGGAC
R87	CCACGCTGAGGATAGCGGTGTCGGGGATGATCTGGTT
R124	CGTGGTGAACATGCCGAAGCCGTGGTCAAATGGCA
R159	AGCACGACCCGAAAGCCGCAGATCAAGTAGCCCAG
R194	cgAAGCTTCGCAAGAATAGCTCCTCCTCGAAGCGGTACAT G

Table S4: Region 2 (R2, wild-type *luc2* primers) Lower case letters denote non-luciferase sequences included for maintaining proper T<sub>m</sub>'s and do not appear in the final assembled gene.

Forward primers	
F0	acacatCGAGGAGGAGCTATTCTTGCGAAGCTTGCA
F36	AGACTATAAGATTCAATCTGCCCTGCTGGTGCCCACAC
F74	TATTTAGCTTCTTCGCTAAGAGCACTCTCATCGACAAGT ACGAC
F117	CTAAGCAACTTGCACGAGATCGCCAGCGGCGGG
F150	GCGCCGCTCAGCAAGGAGGTAGGTGAGGCC
F181	GTGGCCAAACGCTTCCACCTACCAGGCATCCG
F213	CCAGGGCTACGGCCTGACAGAAACAAGTAGTGCCA
F248	TTCTGATCACCCCCgtcgcctccggaggaga
F264	tctcctccggaggcg
Reverse primers	
R0	TAGCTCCTCCTCGatgtgt
R19	GGGCAGATTGAATCTTATAGTCTTGCAAGCTTCGCAAGA A
R59	GCTCTTAGCGAAGAAGCTAAATAGTGTGGGCACCAGCA
R97	TCGTGCAAGTTGCTTAGGTCGTAATTGTCGATGAGAGT
R135	TTGCTGAGCGGCGCCCCGCCGCTGGCGATC
R165	GGAAGCGTTTGGCCACGGCCTCACCTACCTCC
R197	AGGCCGTAGCCCTGGCGGATGCCTGGTAGGT
R228	acGGGGGTGATCAGAATGGCACTAGTTGTTTCTGTC

Table S5-S13: Primers used to construct site-directed libraries. The bases highlighted in red denote sites targeted for saturation mutagenesis.

Table S5: SD218 primers	
Forward primers	
SD218-F1	GCACCGCACCGCTTGTGTCKGGTTCAGTCATGCCC
SD218-F2	GCACCGCACCGCTTGTGTCKWKTTCAGTCATGCCC
SD218-F3	GCACCGCACCGCTTGTGTCKVHGTTCAGTCATGCCC
SD218-F4	GCACCGCACCGCTTGTGTCKNATTCAGTCATGCCC
Reverse primers	
SD218-R1	GCCGAAGATGGGGTCGCGGGCATGACTGAAACCMGAC
SD218-R2	GCCGAAGATGGGGTCGCGGGCATGACTGAAAMWGAC
SD218-R3	GCCGAAGATGGGGTCGCGGGCATGACTGAAACDBGAC
SD218-R4	GCCGAAGATGGGGTCGCGGGCATGACTGAAATNGAC

Table S6: SD227 primers	
Forward primers	
SD227-F1	GCGACCCCATCTGKGGCAACCAGATCATCCCCGACA
SD227-F2	GCGACCCCATCVBCGGCAACCAGATCATCCCCGACA
SD227-F3	GCGACCCCATCATGGGCAACCAGATCATCCCCGACA
SD227-F4	GCGACCCCATCNAWGGCAACCAGATCATCCCCGACA
Reverse primers	
SD227-R1	GCCMCAGATGGGGTCGCGGGCATGACTGAATCGGAC
SD227-R2	GCCGVBGATGGGGTCGCGGGCATGACTGAATCGGAC
SD227-R3	GCCCATGATGGGGTCGCGGGCATGACTGAATCGGAC
SD227-R4	GCCWTNGATGGGGTCGCGGGCATGACTGAATCGGAC

Table S7: SD240 primers	
Forward primers	
SD240-F1	CCGCTATCCTCAGCGBCGBCCCA GBC CACCACGGC
SD240-F2	CCGCTATCCTCAGCGBCGBCCCA WTK CACCACGGC
SD240-F3	CCGCTATCCTCAGCGBCWTKCCA GBC CACCACGGC
SD240-F4	CCGCTATCCTCAGCWTKGBCCCA GBC CACCACGGC
SD240-F5	CCGCTATCCTCAGCWTKWTKCCA GBC CACCACGGC
SD240-F6	CCGCTATCCTCAGCWTKGBCCCA WTK CACCACGGC
SD240-F7	CCGCTATCCTCAGCGBCWTKCCA WTK CACCACGGC
SD240-F8	CCGCTATCCTCAGCWTKWTKCCA WTK CACCACGGC

SD240-F9	NDTGGCATGTTACACCACGCTGGGCTACTTGATCTGCG
SD240-F10	VHGGGCATGTTACACCACGCTGGGCTACTTGATCTGCG
SD240-F11	TGGGGCATGTTACACCACGCTGGGCTACTTGATCTGCG
Reverse primers	
SD240-R1	GVCVCGCTGAGGATAGCGGTGTCGGGGATGATCTGGT T
SD240-R2	MAWMAWGCTGAGGATAGCGGTGTCGGGGATGATCTG GTT
SD240-R3	MAWGVCCTGAGGATAGCGGTGTCGGGGATGATCTGG TT
SD240-R4	GVCMAWGCTGAGGATAGCGGTGTCGGGGATGATCTGG TT
SD240-R5	CGTGGTGAACATGCCAHNGCCGTGGTGGVCTGG
SD240-R6	CGTGGTGAACATGCCCDBGCCGTGGTGGVCTGG
SD240-R7	CGTGGTGAACATGCCCCAGCCGTGGTGGVCTGG
SD240-R8	CGTGGTGAACATGCCAHNGCCGTGGTGMAWTGG
SD240-R9	CGTGGTGAACATGCCCDBGCCGTGGTGMAWTGG
SD240-R10	CGTGGTGAACATGCCCCAGCCGTGGTGMAWTGG

Table S8: SD249 primers	
Forward primers	
SD249-F1	CACGGCTTCGGCGBCNDTNDTACGCTGGGCTACTTGAT CTGCGG
SD249-F2	CACGGCTTCGGCGBCNDTVHACGCTGGGCTACTTGAT CTGCGG
SD249-F3	CACGGCTTCGGCGBCNDTTGGACGCTGGGCTACTTGAT CTGCGG
SD249-F4	CACGGCTTCGGCGBCVHGNDTACGCTGGGCTACTTGAT CTGCGG
SD249-F5	CACGGCTTCGGCGBCVHGVHACGCTGGGCTACTTGAT CTGCGG
SD249-F6	CACGGCTTCGGCGBCVHGTGGACGCTGGGCTACTTGAT CTGCGG
SD249-F7	CACGGCTTCGGCGBCTGGNDTACGCTGGGCTACTTGAT CTGCGG
SD249-F8	CACGGCTTCGGCGBCTGGVHACGCTGGGCTACTTGAT CTGCGG
SD249-F9	CACGGCTTCGGCGBCTGGTGGACGCTGGGCTACTTGAT CTGCGG
SD249-F10	CACGGCTTCGGCWTKNDTNDTACGCTGGGCTACTTGAT CTGCGG
SD249-F11	CACGGCTTCGGCWTKNDTVHACGCTGGGCTACTTGAT

	CTGCGG
SD249-F12	CACGGCTTCGGC <b>WTKNDTTGG</b> ACGCTGGGCTACTTGAT CTGCGG
SD249-F13	CACGGCTTCGGC <b>WTKVHGNDT</b> ACGCTGGGCTACTTGAT CTGCGG
SD249-F14	CACGGCTTCGGC <b>WTKVHGVHG</b> ACGCTGGGCTACTTGAT CTGCGG
SD249-F15	CACGGCTTCGGC <b>WTKVHGTGG</b> ACGCTGGGCTACTTGAT CTGCGG
SD249-F16	CACGGCTTCGGC <b>WTKTGGNDT</b> ACGCTGGGCTACTTGAT CTGCGG
SD249-F17	CACGGCTTCGGC <b>WTKTGGVHG</b> ACGCTGGGCTACTTGAT CTGCGG
SD249-F18	CACGGCTTCGGC <b>WTKTGGTGG</b> ACGCTGGGCTACTTGAT CTGCGG
Reverse primers	
SD249-R1	<b>GVCAHNAHN</b> GCCGAAGCCGTGGTCAAATGG
SD249-R2	<b>GVCAHNCDB</b> GCCGAAGCCGTGGTCAAATGG
SD249-R3	<b>GVCAHNCCA</b> GCCGAAGCCGTGGTCAAATGG
SD249-R4	<b>GVCCDBAHN</b> GCCGAAGCCGTGGTCAAATGG
SD249-R5	<b>GVCCDBCDB</b> GCCGAAGCCGTGGTCAAATGG
SD249-R6	<b>GVCCDBCCA</b> GCCGAAGCCGTGGTCAAATGG
SD249-R7	<b>GVCCCAAHN</b> GCCGAAGCCGTGGTCAAATGG
SD249-R8	<b>GVCCACADB</b> GCCGAAGCCGTGGTCAAATGG
SD249-R9	<b>GVCCACCA</b> GCCGAAGCCGTGGTCAAATGG
SD249-R10	<b>MAWAHNNDT</b> GCCGAAGCCGTGGTCAAATGG
SD249-R11	<b>MAWAHNCDB</b> GCCGAAGCCGTGGTCAAATGG
SD249-R12	<b>MAWAHNCCA</b> GCCGAAGCCGTGGTCAAATGG
SD249-R13	<b>MAWCDBAHN</b> GCCGAAGCCGTGGTCAAATGG
SD249-R14	<b>MAWCDBCDB</b> GCCGAAGCCGTGGTCAAATGG
SD249-R15	<b>MAWCDBCCA</b> GCCGAAGCCGTGGTCAAATGG
SD249-R16	<b>MAWCCAAHN</b> GCCGAAGCCGTGGTCAAATGG
SD249-R17	<b>MAWCCACDB</b> GCCGAAGCCGTGGTCAAATGG
SD249-R18	<b>MAWCCACCA</b> GCCGAAGCCGTGGTCAAATGG

Table S9: SD255 primers	
Forward primers	
SD255-F1	TTCGGCATGTTCAACACGCTGGGC <b>NDTNDTNDT</b> TGCG
SD255-F2	TTCGGCATGTTCAACACGCTGGGC <b>NDTNDTVHGT</b> TGCG
SD255-F3	TTCGGCATGTTCAACACGCTGGGC <b>NDTNDTTGG</b> TGCG
SD255-F4	TTCGGCATGTTCAACACGCTGGGC <b>NDTVHGNDT</b> TGCG

SD255-F5	TTCGGCATGTTTACCACGCTGGGCNDTVHGVHGTGCG
SD255-F6	TTCGGCATGTTTACCACGCTGGGCNDTVHGTGGTGCG
SD255-F7	TTCGGCATGTTTACCACGCTGGGCNDTTGGNDTTGCG
SD255-F8	TTCGGCATGTTTACCACGCTGGGCNDTTGGVHGTGCG
SD255-F9	TTCGGCATGTTTACCACGCTGGGCNDTTGGTGGTGCG
Reverse primers	
SD255-R1	AGCACGACCCGAAAGCCGCAAHNAHNAHNGCCCAG
SD255-R2	AGCACGACCCGAAAGCCGCA CDBAHNAHNGCCCAG
SD255-R3	AGCACGACCCGAAAGCCGCA CCAAHNAHNGCCCAG
SD255-R4	AGCACGACCCGAAAGCCGCA AHNCDBAHNGCCCAG
SD255-R5	AGCACGACCCGAAAGCCGCA CDBCDBAHNGCCCAG
SD255-R6	AGCACGACCCGAAAGCCGCA CCACDBAHNGCCCAG
SD255-R7	AGCACGACCCGAAAGCCGCA AHNCCAAHNGCCCAG
SD255-R8	AGCACGACCCGAAAGCCGCA CDBCCAAHNGCCCAG
SD255-R9	AGCACGACCCGAAAGCCGCA CCACCAAHNGCCCAG

Table S10: SD286 primers	
Forward primers	
SD286-F1	AGACTATAAGATTCAATCTGCCGBCGBCGBCCCCACAC
SD286-F2	AGACTATAAGATTCAATCTGCCGBCWTKWTKCCCACAC
SD286-F3	AGACTATAAGATTCAATCTGCCGBCWTKGBCCCCACAC
SD286-F4	AGACTATAAGATTCAATCTGCCGBCGBCWTKCCCACAC
SD286-F5	AGACTATAAGATTCAATCTGCCWTKWTKWTKCCCACA C
SD286-F6	AGACTATAAGATTCAATCTGCCWTKGBCGBCCCCACAC
SD286-F7	AGACTATAAGATTCAATCTGCCWTKGBCWTKCCCACAC
SD286-F8	AGACTATAAGATTCAATCTGCCWTKWTKGBCCCCACAC
Reverse primers	
SD286-R1	GCTCTTAGCGAAGAAGCTAAATAGTGTGGG GVC GVC GVC C
SD286-R2	GCTCTTAGCGAAGAAGCTAAATAGTGTGGG MAW MAW GVC
SD286-R3	GCTCTTAGCGAAGAAGCTAAATAGTGTGGG GVC MAW G VC
SD286-R4	GCTCTTAGCGAAGAAGCTAAATAGTGTGGG MAW GVC G VC
SD286-R5	GCTCTTAGCGAAGAAGCTAAATAGTGTGGG MAW MAW MAW
SD286-R6	GCTCTTAGCGAAGAAGCTAAATAGTGTGGG GVC GVC M AW
SD286-R7	GCTCTTAGCGAAGAAGCTAAATAGTGTGGG MAW GVC M

	AW
SD286-R8	GCTCTTAGCGAAGAAGCTAAATAGTGTGGGGVCM <del>AWM</del> AW

Table S11: SD314 primers	
Forward primers	
SD314-F1	CTAAGCAACTTGCACGAGATCGCC <del>NDTNDTNDT</del>
SD314-F2	CTAAGCAACTTGCACGAGATCGCC <del>NDTNDTVHG</del>
SD314-F3	CTAAGCAACTTGCACGAGATCGCC <del>NDTNDTTGG</del>
SD314-F4	CTAAGCAACTTGCACGAGATCGCC <del>VHGNDTNDT</del>
SD314-F5	CTAAGCAACTTGCACGAGATCGCC <del>VHGNDTVHG</del>
SD314-F6	CTAAGCAACTTGCACGAGATCGCC <del>VHGNDTTGG</del>
SD314-F7	CTAAGCAACTTGCACGAGATCGCC <del>TGGNDTNDT</del>
SD314-F8	CTAAGCAACTTGCACGAGATCGCC <del>TGGNDTVHG</del>
SD314-F9	CTAAGCAACTTGCACGAGATCGCC <del>TGGNDTTGG</del>
Reverse primers	
SD314-R1	TTGCTGAGCGGCGC <del>AHNAHNAHN</del> GGCGATC
SD314-R2	TTGCTGAGCGGCGC <del>CDBAHNAHN</del> GGCGATC
SD314-R3	TTGCTGAGCGGCGC <del>CCAAHNAHN</del> GGCGATC
SD314-R4	TTGCTGAGCGGCGC <del>AHNAHNCDB</del> GGCGATC
SD314-R5	TTGCTGAGCGGCGC <del>CDBAHNCDB</del> GGCGATC
SD314-R6	TTGCTGAGCGGCGC <del>CCAAHNCDB</del> GGCGATC
SD314-R7	TTGCTGAGCGGCGC <del>AHNAHNCCA</del> GGCGATC
SD314-R8	TTGCTGAGCGGCGC <del>CDBAHNCCA</del> GGCGATC
SD314-R9	TTGCTGAGCGGCGC <del>CCAAHNCCA</del> GGCGATC

Table S12: SD337 primers	
Forward primers	
SD337-F1	<del>NDTNDT</del> GGC <del>NDT</del> GGCCTGACAGAAACA <del>ACTAGTGCCA</del>
SD337-F2	<del>NDTVHG</del> GGC <del>NDT</del> GGCCTGACAGAAACA <del>ACTAGTGCCA</del>
SD337-F3	<del>NDTTGG</del> GGC <del>NDT</del> GGCCTGACAGAAACA <del>ACTAGTGCCA</del>
SD337-F4	<del>VHGNDT</del> GGC <del>NDT</del> GGCCTGACAGAAACA <del>ACTAGTGCCA</del>
SD337-F5	<del>VHGVHG</del> GGC <del>NDT</del> GGCCTGACAGAAACA <del>ACTAGTGCCA</del>
SD337-F6	<del>VHGTGG</del> GGC <del>NDT</del> GGCCTGACAGAAACA <del>ACTAGTGCCA</del>
SD337-F7	<del>TGGNDT</del> GGC <del>NDT</del> GGCCTGACAGAAACA <del>ACTAGTGCCA</del>
SD337-F8	<del>TGGVHG</del> GGC <del>NDT</del> GGCCTGACAGAAACA <del>ACTAGTGCCA</del>
SD337-F9	<del>TGGTGG</del> GGC <del>NDT</del> GGCCTGACAGAAACA <del>ACTAGTGCCA</del>
Reverse primers	



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

Table S13: SD347 primers	
Reverse primers	
SD347-R1	CTTCGGGGGTGATCAGAATAHNAHNAAGTTGTTTCTGTCA AGGCCG
SD347-R2	CTTCGGGGGTGATCAGAATAHNCDBGAGTTGTTTCTGTCA GGCCG
SD347-R3	CTTCGGGGGTGATCAGAATAHNCCAAGTTGTTTCTGTCA GGCCG
SD347-R4	CTTCGGGGGTGATCAGAATCDBAHNAAGTTGTTTCTGTCA GGCCG
SD347-R5	CTTCGGGGGTGATCAGAATCDBCDBGAGTTGTTTCTGTCA GGCCG
SD347-R6	CTTCGGGGGTGATCAGAATCDBCCAAGTTGTTTCTGTCA GGCCG
SD347-R7	CTTCGGGGGTGATCAGAATCCAAHNAAGTTGTTTCTGTCA GGCCG
SD347-R8	CTTCGGGGGTGATCAGAATCCACDBGAGTTGTTTCTGTCA GGCCG
SD347-R9	CTTCGGGGGTGATCAGAATCCACCAAGTTGTTTCTGTCA GGCCG

Table S14: Primers used to create R1 and R2 deletions in pET Luc2

Outside primers	
Nco-Luc2-For	ATACCATGGATGGAAGATGCCAAAACATTAAGAAG GG
Luc2-Not-Rev	ATAGGCCGCCACGGCGATC
R1 deletion	
R1-deletion-For	GAACAGTAGTGGATCCAAGCTTGCAAGACTATAAGA TTCAATCTGCCC

R1-deletion-Rev	TAGTCTTGCAAGCTTGGATCCACTACTGTTCATGATC AGGGC
R2 deletion	
R2-deletion-For	TTCTTGCGAAGCTTACTAGTGCCATTCTGATCACCCC CG
R2-deletion-Rev	CAGAATGGCACTAGTAAGCTTCGCAAGAATAGCTCC TCCTCG

Table S15: Primers used to amplify inserts for the R1 or R2 regions, along with the pET vector backbone.

R1 insert amplification primers	
R1-insert-Fwd	GACAAAACCATCGCCCTGATCATGAACAGTAGTGG ATCCACCGGATTGCCCAA
R1-insert-Rev	CACCAGCAGGGCAGATTGAATCTTATAGTCTTGCA AGCTTCGCAAGAATAGCTCCTCCTC
R1- vector amplification primers	
R1-vector-Fwd	ACTACTGTTCATGATCAGGGCGATGGTTTTGTC
R1-vector-Rev	GCAAGACTATAAGATTCAATCTGCCCTGCTGGTG
R2 insert amplification primers	
R2-insert-Fwd	CGAGGAGGAGCTATTCTTGCG
R2-insert-Rev	GGGGGTGATCAGAATGGCACTA
R2- vector amplification primers	
R2-vector-Fwd	TAGTGCCATTCTGATCACCCCC
R2-vector-Rev	CGCAAGAATAGCTCCTCCTCG

### 5-11c. Screening protocols

#### *Primary screening protocol*

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

#### *Secondary screening of library members*

Light-emitting colonies from the agar plates (primary screen) were picked and added to 5 mL of LB broth (containing 40  $\mu$ g/mL Kan). Colonies expressing wild-type Fluc (or another luciferase from an earlier round) were also picked for growth and comparative analysis. At least two colonies (per thirty library members) were picked and grown up simultaneously, as controls. All culture tubes were shaken and incubated at 37  $^{\circ}$ C for 2-5 h until the OD<sub>600</sub> reached ~0.8. Rapidly growing cultures were stored at 4  $^{\circ}$ C until all cultures reached the desired OD<sub>600</sub>. A portion of each culture (0.5 mL) was reserved and stored at 4  $^{\circ}$ C for sequencing analysis. IPTG (0.5 mM) was added to the remainder of the cultures to induce protein expression for 2 h at 30  $^{\circ}$ C or for 12 h at 25  $^{\circ}$ C.

After induction, the cultures were centrifuged for 10 min at 3400 x g (4 °C). The supernatant was removed, and the pellet was resuspended in 0.4 – 0.6 mL of Fluc lysis buffer (50 mM Tris•HCl, 500 NaCl, 0.5% (v/v) Tween, 5 mM MgCl<sub>2</sub>, pH = 7.4). Lysed bacterial suspensions (90 µL/well) were added to a 96-well black plate (in triplicate). The lysed cultures were imaged using a stock solution of 10X luciferin and 10X ATP in PBS (1-5 mM luciferin, 10 mM ATP). The appropriate luciferin stock solution (10 µL) was added to each well, 6 wells at a time (final [luciferin] = 100-500 µM and [ATP] = 1 mM). Images were acquired over 10 min as described above. Bioluminescence readings acquired 5 min post-luciferin addition were used for most analyses. Library members emitting light equivalent to or greater than wild-type Fluc (first round of screening) or the “hits” from previous rounds (subsequent rounds of screening) were further analyzed. The reserved cultures for these mutants were added to 3.5 mL of LB-Kan and grown in a shaking incubator at 37 °C overnight. DNA from the colonies of interest was harvested and sequenced.

#### **5-11d. Identification of luciferase mutant B selectivity**

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

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

5'-CTTCGGGGGTGATCAGAATGGC-3' for PCR 1

5'-GACAAAACCATCGCCCTGATCATGAAC-3' and

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

Bases highlighted in red denote residue 218 mutated to alanine. The vector backbone was linearized and amplified from pET28a *luc2* encoding wild-type Fluc or a mutant (F250M, S314T, and/or G316T) sequence using primers R1-vector-Fwd and R2-vector-Rev (Table S9). The insert and vector were combined by CPEC (as above). A portion of the reaction (2.5  $\mu$ L) was then utilized directly for transformation into XL1 competent *E. coli*. Sequencing analysis confirmed correct mutation and assembly.

#### **5-11e. Mammalian plasmid construction**

To express the mutant luciferases in mammalian cells, the R1-R2 region of the mutant luciferases was amplified and inserted into pcDNA-Luc2-IRES-GFP (9).

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

5'- GACAAAACCATCGCCCTGATCATGAAC-3' and

5'-CTTCGGGGGTGATCAGAATGGC-3'

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

#### **5-11f. Recombinant protein expression and purification**

The pET-luciferase plasmids (WT, mutants A, B, and C) were transformed into chemically competent *E. coli* BL21 cells. Cultures (1 L) were grown in 2 L flasks at 37 °C in LB broth (containing 40  $\mu$ g/mL Kan) to mid-log phase ( $OD_{600}$  = 0.8 - 1.0), induced with 0.5 mM IPTG,

and incubated at 22 °C for 16 – 18 h. The cells were harvested by centrifugation at 4 °C and resuspended in 20 mL of a solution of 50 mM NaPO<sub>4</sub>, 300 mM NaCl, 1 mM DTT, and 1 mM PMSF at pH 7.4. Lysozyme (1 mg) was added, and the cells were sonicated and centrifuged at 10000 x g for 1 h at 4 °C. WT Fluc and mutant luciferases were purified from clarified supernatants using nickel affinity chromatography (BioLogic Duo Flow Chromatography System, Bio-Rad). Proteins were dialyzed into 25 mM Tris-acetate (pH 7.8) buffer containing 1 mM EDTA and 0.2 mM ammonium sulfate (4 °C). DTT (1 mM) and 15% glycerol were added to the dialyzed samples prior to storage at -20 °C. Final protein concentrations were determined using a standard BCA assay or UV spectroscopy. SDS-PAGE was also performed, and gels were stained with Coomassie R-250.

### **5-11g Light emission assays with recombinant luciferase**

Bioluminescence assays with all luciferin compounds were carried out in triplicate, using solid black, flat-bottom, 96-well plates (BD Bioscience). Assay wells contained purified Fluc (1-2 µg), luciferin substrate (0-1 mM), ATP (Sigma, 0-1 mM), and reaction buffer (20 mM Tris-HCl, 0.5 mg/mL BSA, 0.1 mM EDTA, 1 mM TCEP, 2 mM MgSO<sub>4</sub>, pH 7.6), totaling 100 µL. For pH studies, the buffer comprised 20 mM BIS-TRIS propane (with 100 µM MgCl<sub>2</sub>, 1 mM ATP, 500 µM Coenzyme A). All non-enzyme assay components were premixed in the wells prior to wild-type or mutant luciferase addition. Images for all assays were acquired as described above and analyzed with Microsoft Excel or GraphPad Prism (version 6.0f for Macintosh, GraphPad Software).

### **5-11h 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<sub>m</sub> and relative k<sub>cat</sub> values were determined using nonlinear regression analyses and robust fit outlier removal in GraphPad Prism (version 6.0f for Macintosh, GraphPad Software).

### **5-11i Mammalian cell culture and imaging**

HEK293 (ATCC) and DB7 cells (courtesy of the Contag laboratory, Stanford) were cultured in DMEM (Corning) supplemented with 10% (vol/vol) fetal bovine serum (FBS, Life Technologies), penicillin (100 U/mL), and streptomycin (100 μg/mL). Cells were maintained in a 5% CO<sub>2</sub> water-saturated incubator at 37 °C. Transient transfection of mutant or WT Fluc DNA were performed using cationic lipid formulations (Lipofectamine 2000; Invitrogen). Cells were analyzed for expression or used for imaging analysis 24 – 36 h post transfection.

#### *Flow cytometry*

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

#### *Bioluminescence imaging with mammalian cells*

HEK293 cells transiently expressing Fluc or mutant luciferases were added to black 96-well plates ( $2 \times 10^5$  cells per well). A stock solution of luciferin (5 mM in PBS) was added to each well (500  $\mu$ M final concentration). Sequential imaging in black 96-well plates was performed as above. In some cases, though, a second luciferin solution (500  $\mu$ M final concentration) was added after the initial imaging session. Bioluminescent-images were then acquired again.

#### *Sequential bioluminescence imaging with patterned cells*

Biocompatible stencils were generated and placed in tissue culture dishes as previously described (10). HEK293 or DB7 cells expressing mutant B or C were plated ( $2 \times 10^5$  (HEK293) or  $3 \times 10^5$  (DB7) per stencil chamber). For HEK293 cells, luciferins were added sequentially to each stencil chamber immediately after plating. For the DB7 cells, the cells were allowed to adhere for 24 h prior to stencil removal and sequential luciferin addition. Bioluminescent images were acquired as described above except with 10 min acquisitions.

### **5-11j Molecular docking studies**

Native Fluc (PDB:4G36) was prepared for docking using the protein prep wizard in Maestro (version 2014-3). The OPLS2005 force field was used for minimization. A Glide-grid was generated using this minimized structure and DSLA was used to provide the coordinates for ligand binding. The luciferin analogs were prepared as the AMP conjugates; their geometries were prepared through the LigPrep module prior to docking. SP docking was used to obtain the



input poses for high-level calculations. XP docking was then performed using flexible ligand sampling, sampling of nitrogen inversions and ring conformations. Epik state penalties were used to exclude non-physiologically relevant protonation states and non-planar amide bonds were also penalized. An XP descriptor file was written in order to facilitate post-docking analysis. The docked ligands were evaluated manually via pose-viewer to choose the most relevant poses as well with XP visualizer to quantify the relative contributions of different ligand interactions to the assessed GlideScore.

## References

1. Adams ST & Miller SC (2014) Beyond D-luciferin: expanding the scope of bioluminescence imaging in vivo. *Curr. Opin. Chem. Biol.* 21:112-120
2. Paley MA & Prescher JA (2014) Bioluminescence: a versatile technique for imaging cellular and molecular features. *MedChemComm* 5(3):255-267
3. Branchini BR, Hayward MM, Bamford S, Brennan PM, & Lajiness EJ (1989) Naphthyl- and quinolyl-luciferin: green and red light emitting Firefly luciferin analogs. *Photochem. Photobiol.* 49(5):689-695
4. Jathoul AP, Grounds H, Anderson JC, & Pule MA (2014) A dual-color far-red to near-infrared Firefly luciferin analogue designed for multiparametric bioluminescence imaging. *Angew. Chem. Int. Ed.* 53(48):13059-13063
5. Mofford DM, Reddy GR, & Miller SC (2014) Aminoluciferins extend Firefly luciferase bioluminescence into the near-infrared and can be preferred substrates over D-Luciferin. *J. Am. Chem. Soc.* 136(38):13277-13282
6. Branchini BR, *et al.* (2007) Thermostable red and green light-producing Firefly luciferase mutants for bioluminescent reporter applications. *Anal. Biochem.* 361(2):253-262
7. Branchini BR, *et al.* (2007) Synergistic mutations produce blue-shifted bioluminescence in firefly luciferase. *Biochemistry* 46:13847-13855
8. Mezzanotte L, *et al.* (2011) Sensitive dual color *in vivo* bioluminescence imaging using a new red codon optimized Firefly luciferase and a Green Click Beetle luciferase. *Plos One* 6(4):e19277

9. Haddock SHD, Moline MA, & Case JF (2010) Bioluminescence in the sea. *Ann. Rev. Mar. Sci.* 2:443-493
10. Oba Y & Schultz DT (2014) Eco-evo bioluminescence on land and in the sea. *Adv. Biochem. Eng. Biotechnol.* 144:3-36.
11. Viviani VR (2002) The origin, diversity, and structure function relationships of insect luciferases. *Cell. Mol. Life Sci.* 59(11):1833-1850
12. Badr CE & Tannous BA (2011) Bioluminescence imaging: progress and applications. *Trends Biotechnol.* 29(12):624-633
13. Massoud TF, Paulmurugan R, De A, Ray P, & Gambhir SS (2007) Reporter gene imaging of protein-protein interactions in living subjects. *Curr. Opin. Biotechnol.* 18(1):31-37
14. Porterfield WB & Prescher JA (2015) Tools for visualizing cell-cell 'interactomes'. *Curr. Opin. Chem. Biol.* 24:121-130
15. Bhaumik S & Gambhir SS (2002) Optical imaging of Renilla luciferase reporter gene expression in living mice. *Proc. Natl. Acad. Sci. U. S. A.* 99(1):377-382
16. Maguire CA, *et al.* (2013) Triple bioluminescence imaging for *in vivo* monitoring of cellular processes. *Mol. Ther. Nucleic Acids* 2:e99
17. Pichler A, Prior JL, & Piwnicka-Worms D (2004) Imaging reversal of multidrug resistance in living mice with bioluminescence: MDR1 P-glycoprotein transports coelenterazine. *Proc. Natl. Acad. Sci. U. S. A.* 101(6):1702-1707
18. Dubinnyi MA, *et al.* (2015) Novel mechanism of bioluminescence: oxidative decarboxylation of a moiety adjacent to the light emitter of Fridericia luciferin. *Angew. Chem. Int. Ed.* 54(24):7065-7067
19. Evans MS, *et al.* (2014) A synthetic luciferin improves bioluminescence imaging in live mice. *Nat. Methods* 11(4):393-395
20. Reddy GR, Thompson WC, & Miller SC (2010) Robust light emission from cyclic alkylaminoluciferin substrates for Firefly luciferase. *J. Am. Chem. Soc.* 132(39):13586-13587
21. Harwood KR, Mofford DM, Reddy GR, & Miller SC (2011) Identification of mutant Firefly luciferases that efficiently utilize aminoluciferins. *Chem. Biol.* 18(12):1649-1657.
22. Adams ST, Mofford DM, Reddy GSKK, & Miller SC (2016) Firefly luciferase mutants allow substrate-selective bioluminescence imaging in the mouse brain. *Angew. Chem. Int. Ed.* 55(16):4943-4946

23. Hall MP, *et al.* (2012) Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem. Biol.* 7(11):1848-1857
24. Rowe L, Dikici E, & Daunert S (2009) Engineering bioluminescent proteins: expanding their analytical potential. *Anal. Chem.* 81(21):8662-8668
25. Berger F, Paulmurugan R, Bhaumik S, & Gambhir SS (2008) Uptake kinetics and biodistribution of C-14-D-luciferin-a radiolabeled substrate for the Firefly luciferase catalyzed bioluminescence reaction: impact on bioluminescence based reporter gene imaging. *Eur. J. Nucl. Med. Mol. Imaging* 35(12):2275-2285
26. Contag CH, *et al.* (1997) Visualizing gene expression in living mammals using a bioluminescent reporter. *Photochem. Photobiol.* 66(4):523-531
27. Zhao H, *et al.* (2005) Emission spectra of bioluminescent reporters and interaction with mammalian tissue determine the sensitivity of detection *in vivo*. *J. Biomed. Opt.* 10(4):041210
28. Branchini BR, Magyar RA, Murtiashaw MH, & Portier NC (2001) The role of active site residue arginine 218 in Firefly luciferase bioluminescence. *Biochemistry* 40:2410-2418
29. Branchini BR, Southworth TL, Murtiashaw MH, Boije H, & Fleet SE (2003) A mutagenesis study of the putative luciferin binding site residues of firefly luciferase. *Biochemistry* 42:10429-10436
30. Nakatsu T, *et al.* (2006) Structural basis for the spectral difference in luciferase bioluminescence. *Nature* 440(7082):372-376
31. Sundlov JA, Fontaine DM, Southworth TL, Branchini BR, & Gulick AM (2012) Crystal structure of Firefly luciferase in a second catalytic conformation supports a domain alternation mechanism. *Biochemistry* 51(33):6493-6495
32. McCutcheon DC, Porterfield WB, & Prescher JA (2015) Rapid and scalable assembly of Firefly luciferase substrates. *Org. Biomol. Chem.* 13(7):2117-2121
33. Meroni G, Rajabi M, & Santaniello E (2009) D-Luciferin, derivatives and analogues: synthesis and *in vitro/in vivo* luciferase-catalyzed bioluminescent activity. *Arkivoc*:265-288
34. Branchini BR, *et al.* (2015) Experimental support for a single electron-transfer oxidation mechanism in Firefly bioluminescence. *J. Am. Chem. Soc.* 137(24):7592-7595
35. Conley NR, Dragulescu-Andrasi A, Rao JH, & Moerner WE (2012) A selenium analogue of Firefly D-luciferin with red-shifted bioluminescence emission. *Angew. Chem. Int. Ed.* 51(14):3350-3353

36. McCutcheon DC, Paley MA, Steinhardt RC, & Prescher JA (2012) Expedient synthesis of electronically modified luciferins for bioluminescence imaging. *J. Am. Chem. Soc.* 134(18):7604-7607
37. White EH & Worther H (1966) Analogs of Firefly luciferin. III. *J. Org. Chem.* 31(5):1484-1488
38. White EH, Worther H, Field GF, & Mcelroy WD (1965) Analogs of Firefly luciferin. *J. Org. Chem.* 30(7):2344-2348
39. White EH, Worther H, Seliger HH, & Mcelroy WD (1966) Amino analogs of Firefly luciferin and biological Activity thereof. *J. Am. Chem. Soc.* 88(9):2015-2019
40. Woodrooffe CC, *et al.* (2012) Novel heterocyclic analogues of Firefly luciferin. *Biochemistry* 51(49):9807-9813.
41. Kojima R, *et al.* (2013) Rational design and development of near-infrared-emitting Firefly luciferins available *in vivo*. *Angew. Chem. Int. Ed.* 52(4):1175-1179
42. Kuchimaru T, *et al.* (2016) A luciferin analogue generating near-infrared bioluminescence achieves highly sensitive deep-tissue imaging. *Nat. Commun.* 7:11856.
43. Seliger HH, Mcelroy WD, Field GF, & White EH (1961) Stereospecificity and Firefly bioluminescence, a comparison of natural and synthetic luciferins. *Proc. Natl. Acad. Sci. U. S. A.* 47(7):1129-1134.
44. Shinde R, Perkins J, & Contag CH (2006) Luciferin derivatives for enhanced *in vitro* and *in vivo* bioluminescence assays. *Biochemistry* 45(37):11103-11112
45. Takakura H, *et al.* (2011) Aminoluciferins as functional bioluminogenic substrates of Firefly luciferase. *Chem. Asian J.* 6(7):1800-1810
46. Woodrooffe CC, *et al.* (2008) N-alkylated 6'-aminoluciferins are bioluminescent substrates for Ultra-Glo and QuantiLum luciferase: New potential scaffolds for bioluminescent assays. *Biochemistry* 47(39):10383-10393
47. Viviani VR, *et al.* (2014) Bioluminescence of Beetle Luciferases with 6'-amino-D-luciferin analogues reveals excited keto-oxyluciferin as the emitter and phenolate/luciferin binding site interactions modulate bioluminescence colors. *Biochemistry* 53(32):5208-5220
48. da Silva LP & da Silva JCGE (2012) Firefly chemiluminescence and bioluminescence: efficient generation of excited states. *ChemPhysChem* 13(9):2257-2262

49. Hopkins TA, Seliger HH, White EH, & Cass MW (1967) Chemiluminescence of Firefly luciferin. A model for bioluminescent reaction and identification of product excited state. *J. Am. Chem. Soc.* 89(26):7148-7150
50. Steinhardt RC, *et al.* (2017) Brominated luciferins are versatile bioluminescent probes. *ChemBioChem* 18:96
51. Kato D, *et al.* (2014) A firefly inspired one-pot chemiluminescence system using n-propylphosphonic anhydride (T3P). *Photochem. Photobiol. Sci.* 13(12):1640-1645
52. Mao Y (2011) Dynamics studies of luciferase using elastic network model: how the sequence distribution of luciferase determines its color. *Protein Eng. Des. Sel.* 24(4):341-349
53. Chen MMY, Snow CD, Vizcarra CL, Mayo SL, & Arnold FH (2012) Comparison of random mutagenesis and semi-rational designed libraries for improved cytochrome P450 BM3-catalyzed hydroxylation of small alkanes. *Protein Eng. Des. Sel.* 25(4):171-178
54. Reetz MT, Prasad S, Carballeira JD, Gumulya Y, & Bocola M (2010) Iterative saturation mutagenesis accelerates laboratory evolution of enzyme stereoselectivity: rigorous comparison with traditional methods. *J. Am. Chem. Soc.* 132(26):9144-9152
55. Reetz MT, Kahakeaw D, & Lohmer R (2008) Addressing the numbers problem in directed evolution. *ChemBioChem* 9(11):1797-1804
56. Kille S, *et al.* (2013) Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. *ACS Synth. Biol.* 2(2):83-92
57. Amaral DT, Arnoldi FGC, & Viviani V (2012) Molecular phylogeny of the neotropical bioluminescent beetles. *Luminescence* 27(2):96-96
58. Viviani VR, Amaral DT, Neves DR, Simões A, & Arnoldi FGC (2013) The luciferin binding site residues C/T311 (S314) influence the bioluminescence color of beetle luciferases through main-chain interaction with oxyluciferin phenolate. *Biochemistry* 52(1):19-27.
59. Pines G, *et al.* (2014) Codon compression algorithms for saturation mutagenesis. *ACS Synth. Biol.* 4(5):604-614
60. Ness JE, *et al.* (2002) Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently. *Nat. Biotechnol.* 20(12):1251-1255
61. Quan J & Tian J (2011) Circular polymerase extension cloning. *Nat. Protoc.* 6:242-251

62. Baud MG, *et al.* (2014) Chemical biology. A bump-and-hole approach to engineer controlled selectivity of BET bromodomain chemical probes. *Science* 346(6209):638-641.
63. Bishop AC, *et al.* (2000) A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* 407(6802):395-401
64. Shah K, Liu Y, Deirmengian C, & Shokat KM (1997) Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proc. Natl. Acad. Sci. U. S. A.* 94(8):3565-3570

## **CHAPTER 6: Conclusions and future directions**

Multicellular networks and dynamic cellular movements are crucial for a vast array of biological processes (1-3). Dysregulated cellular crosstalk potentiates numerous pathologies, including infection and metastases. While monitoring individual cells with molecular detail are important to understanding these events, a comprehensive view requires new knowledge at the whole organism level. In this thesis, I developed new sets of bioluminescent tools for monitoring biological processes on the macroscale.

Fluorescence imaging, with fluorophores and fluorescent proteins, is among the most commonly utilized techniques for visualizing cells with remarkable spatial resolution (4-6). These tools have revolutionized our understanding of fundamental biological processes but require an external light source. This requirement results in only being able to visualize small, previously selected areas (7-8). To study cell interactions and migration over longer length times and distances necessitates new optical instrumentation or the development of new imaging tools.

A complementary imaging technology, bioluminescence, is better suited for macroscale imaging (9-11). Bioluminescent enzymes (luciferases) emit light through the chemical oxidation of small molecule luciferins. The emitted light can penetrate tissues and skin in rodent models and be detected with sensitive cameras. Since an excitation light source is not required and mammalian tissues do not produce large numbers of photons, there is virtually no background signal. This sensitivity has made bioluminescence imaging useful for imaging in a number of areas such tumor cell proliferation, cell homing, and numerous other processes. As with all macroscale techniques, bioluminescence suffers from low resolution. Additionally, bioluminescence imaging has been limited to monitoring one cell type or biological feature at a time.

I aimed to address both of these limitations by developing new bioluminescence tools by engineering luciferases. I developed two distinct genetically encodable probes for monitoring interacting cells. One class comprised secreted split luciferase fragments and the second group comprised cell surface localized fragments. I also evolved luciferase mutants to enable substrate resolved, multi-component imaging.

For monitoring cell-cell contacts, I utilized a split reporter strategy with an ATP-independent marine luciferase from *Gaussia princeps*. In this method, a given cell population will express one split reporter, while the other reporter is expressed by a distinct second cell population. When the two cells associate, a functional luciferase (and thus, light) is generated through the binding of the split reporter fragments. With our completely secreted system, we determined split forms of Gluc can assemble in the extracellular space to enable sensitive imaging of cell proximity (Figure 2-11). While these studies provided proof-of-concept for monitoring cell-cell interactions, the light emission was ~100-fold below the native enzyme (Figure 3-2). While these tools are likely to have immediate utility for imaging interacting cells in excised tissues, *in vivo* imaging would require at least  $2.5 \times 10^5$  total interacting cells. This signal is therefore too weak to monitor processes with rare events, such as earlier metastatic cell-cell interactions.

We therefore set out to improve fragment assembly to boost photon production. We assessed a variety of strategies to optimize complementation including linker lengths, resonance energy transfer, and “designer” luciferases (Figure 3-1). Surprisingly, dramatic design modifications were needed for the most significant enhancements in light emission. The most robust light emission was observed when the split reporters were interfaced with an alternative



luciferase, Nanoluc (Figure 3-11). These improvements are likely to facilitate the visualization of a broad range of cellular interactions.

With split reporters that are exclusively secreted, we recognized that detecting interacting cells *in vivo* would require the complemented Gluc to remain localized to the location of cell-cell interactions. Therefore, while we evaluated the secreted system *in vivo* we also aimed to develop a strategy to localize the reporters to the cell surface. Our initial design tethered a split reporter to the transmembrane domain from the immune cell surface marker CD4. When cells secreting one split reporter were cultured with cells expressing the membrane-anchored split reporter, signal was detected but at much lower levels than the secreted system (Figure 4-4). To improve the sensitivity of this approach, we aimed to improve the accessibility of the split reporter at the cell surface. We investigated split reporter orientation and linker design to push the transmembrane split reporters above receptors or glycan structures at the cell surface. Both longer flexible linkers and rigid linkers enhanced light emission but the tools did not translate to cancer-relevant cell lines. The studies provided important design considerations for this approach.

We designed next generation split reporters interfaced with native ligand-receptor interacting proteins (CD40 ligand and receptor). This resulted in ~20-fold enhancement in signal over our first-generation transmembrane reporters (Figure 4-6). Importantly, the split Gluc-CD40 system localizes signal to the cell surface and provides a distance-dependent readout on cell-cell interactions (Figure 4-8). This system is likely applicable to detecting early liver metastases. The split Gluc-receptor will be tethered to liver cells and the complementary split Gluc-ligand will be secreted by tumor cells. When metastases to the liver occur, the secreted and tethered Gluc fragments should be able to assemble to emit light. These tools are also likely to be applicable to

monitoring many processes involving cell-cell contacts including host-pathogen dynamics and cell migration.

The second set of tools I developed was aimed at engineering new luciferase-luciferin probes for multi-component imaging. Since spectral resolution in whole organism imaging is extremely challenging, we sought to generate distinguishable bioluminescent probes based on substrate resolution. In this strategy, cells would express a mutant luciferase that can only oxidize its unique luciferin. Another cells population would utilize an orthogonal luciferase-luciferin pair thus enabling multi-cellular imaging.

To develop these tools, we required both new luciferin analogs and luciferase enzymes. Based on structural analysis, we generated 4'- and 7'-modified luciferins with additional steric bulk at these sites (Figure 5-1). New luciferase enzymes were developed by generating mutant libraries which targeted residues proximal to the active site (Figure 5-6). With both components in hand, we developed a two-tiered screening approach. The first step involved an on-plate screen which allowed us to quickly remove non-functional enzymes from our libraries. This was followed by a secondary screen in which “hits” could be confirmed to create next-generation libraries. Secondary screens also allowed us to identify substrate selectivity by screening panels of mutant enzymes against panels of luciferin analogs.

This screening approach has generated 5,000 orthogonal pairings with at least 10-fold substrate resolution. Analysis of these large datasets has informed likely substrate-substrate and library-library pairings. This information is important to our approach as it can guide new library and analog design efforts. As such, ongoing work in the lab is providing new luciferin analogs with modified electronic properties and alternative aromatic cores. New libraries based on “hot” spots in the enzyme are also being generated with combinatorial codon mutagenesis techniques.

The latest computational analysis of our imaging dataset suggests orthogonality scores improve with more substrates and enzymes, further validating the need for diverse collections of bioluminescent probes.

We selected a top “hit” to further characterize (mutant B and C). In purified enzyme, mutant B exhibited nearly a 100-fold preference with 4'-MorphoLuc and mutant C favored 7'-DMAMeLuc (Figure 5-13). Additional mutants, based on enzyme B, were analyzed and discerned critical mutations for selectivity with 4'-MorphoLuc. Targeting these positions in future libraries could provide improvements to orthogonality.

The orthogonal enzyme-substrate probes were also evaluated in mammalian cell lines and in mice (Figure 5-20 and 5-27). In both cases, these tools produced light only when complementary enzyme-substrate partners interacted. Based on their selectivity, these designer pairs will bolster multi-component imaging and enable the direct interrogation of cell networks not currently possible with existing tools. Current efforts are aimed at evolving “brighter” luciferase-luciferin pairs for more sensitive *in vivo* detection.

In summary, I was able to expand the bioluminescent toolkit to begin to address the limitations of this imaging technology. I developed new probes relevant to monitoring cell-cell interactions and for multi-cellular imaging. We anticipate these tools will broaden the scope of bioluminescence imaging and that they will have utility in tandem for monitoring a wide-range of biological processes.

## References

1. Batista FD & Dustin ML (2013) Cell:cell interactions in the immune system. *Immunol. Rev.* 251(1):7-12
2. Mayor R & Carmona-Fontaine C (2010) Keeping in touch with contact inhibition of locomotion. *Trends Cell Biol.* 20(6):319-328
3. Ng CT, Snell LM, Brooks DG, & Oldstone MB (2013) Networking at the level of host immunity: immune cell interactions during persistent viral infections. *Cell Host Microbe* 13(6):652-664
4. Troy T, Jekic-McMullen D, Sambucetti L, & Rice B (2004) Quantitative comparison of the sensitivity of detection of fluorescent and bioluminescent reporters in animal models. *Mol. Imaging* 3(1):9-23
5. Ai HW (2015) Fluorescent-protein-based probes: general principles and practices. *Anal. Bioanal. Chem.* 407(1):9-15
6. Giepmans BN, Adams SR, Ellisman MH, & Tsien RY (2006) The fluorescent toolbox for assessing protein location and function. *Science* 312(5771):217-224
7. Alexander S, Weigelin B, Winkler F, & Friedl P (2013) Preclinical intravital microscopy of the tumour-stroma interface: invasion, metastasis, and therapy response. *Curr. Opin. Cell Biol.* 25(5):659-671
8. Hawkins ED, *et al.* (2016) T-cell acute leukaemia exhibits dynamic interactions with bone marrow microenvironments. *Nature* 538(7626):518-522
9. Paley MA & Prescher JA (2014) Bioluminescence: a versatile technique for imaging cellular and molecular features. *MedChemComm* 5(3):255-267
10. Porterfield WB & Prescher JA (2015) Tools for visualizing cell-cell 'interactomes'. *Curr. Opin. Chem. Biol.* 24:121-130
11. Prescher JA & Contag CH (2010) Guided by the light: visualizing biomolecular processes in living animals with bioluminescence. *Curr. Opin. Chem. Biol.* 14(1):80-89