

UNIVERSITY OF CALIFORNIA,  
IRVINE

Parallel screening for rapid identification of orthogonal bioluminescent tools

DISSERTATION

submitted in partial satisfaction of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

in Organic Chemistry

by

Colin Michael Rathbun

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

2012

© 2012 Colin Michael Rathbun

## **DEDICATION**

(Optional dedication page) xxx To ...

## TABLE OF CONTENTS

	Page
<b>LIST OF FIGURES</b>	<b>iv</b>
<b>LIST OF TABLES</b>	<b>v</b>
<b>ACKNOWLEDGMENTS</b>	<b>vi</b>
<b>CURRICULUM VITAE</b>	<b>vii</b>
<b>ABSTRACT OF THE DISSERTATION</b>	<b>ix</b>
<b>1 Orthogonal Luciferase[Please insert into preamble]Luciferin Pairs for Bioluminescence Imaging</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Results and Discussion . . . . .	5
1.3 Conclusions . . . . .	14
<b>A Appendix Title</b>	<b>23</b>
A.1 Lorem Ipsum . . . . .	23

## LIST OF FIGURES

	Page
1.1 Expanding the bioluminescence toolkit with unique enzyme–substrate pairs	2
1.2 Measuring luciferin light emission . . . . .	7
1.3 Generating mutant luciferase libraries and screening for orthogonal pairs . .	10
1.4 Analyzing orthogonal enzyme–substrate pairs . . . . .	16
1.5 Imaging cells with orthogonal luciferase–luciferin pairs . . . . .	17

## **LIST OF TABLES**

	Page
1.1 Biochemical analyses of orthogonal enzyme— substrate pairs . . . . .	12

## **ACKNOWLEDGMENTS**

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

You also need to acknowledge any publishers of your previous work who have given you permission to incorporate that work into your dissertation. See Section 3.2 of the UCI Thesis and Dissertation Manual.)

# CURRICULUM VITAE

Colin Michael Rathbun

## EDUCATION

<b>Doctor of Philosophy in Computer Science</b>	<b>2012</b>
University name	<i>City, State</i>
<b>Bachelor of Science in Computational Sciences</b>	<b>2007</b>
Another university name	<i>City, State</i>

## RESEARCH EXPERIENCE

<b>Graduate Research Assistant</b>	<b>2007–2012</b>
University of California, Irvine	<i>Irvine, California</i>

## TEACHING EXPERIENCE

<b>Teaching Assistant</b>	<b>2009–2010</b>
University name	<i>City, State</i>

## REFEREED JOURNAL PUBLICATIONS

**Ground-breaking article** 2012  
Journal name

## REFEREED CONFERENCE PUBLICATIONS

**Awesome paper** Jun 2011  
Conference name

**Another awesome paper** Aug 2012  
Conference name

## SOFTWARE

**Magical tool** <http://your.url.here/>  
*C++ algorithm that solves TSP in polynomial time.*

# **ABSTRACT OF THE DISSERTATION**

Parallel screening for rapid identification of orthogonal bioluminescent tools

By

Colin Michael Rathbun

Doctor of Philosophy in Organic Chemistry

University of California, Irvine, 2012

Associate Professor Jennifer Prescher, Chair

The abstract of your contribution goes here.

# Chapter 1

## Orthogonal Luciferase–Luciferin Pairs for Bioluminescence Imaging

### 1.1 Introduction

Bioluminescence imaging is a popular method for visualizing cells and other biological features *in vivo*.<sup>1</sup> This technology relies on enzymes (luciferases) that catalyze the oxidation of small-molecule substrates (luciferins). The oxidation process is accompanied by the release of light (Figure 1.1A). Since mammalian cells and tissues do not emit substantial numbers of photons, bioluminescent light can facilitate sensitive imaging in these environments.<sup>2</sup> Luciferase-labeled cells can also be imaged repeatedly and noninvasively in a variety of pre-clinical models. This broad dynamic range has enabled numerous studies of fundamental biological processes, including cell homing and differentiation, proliferation, and cell-to-cell communication, in physiologically relevant environments.<sup>3</sup>

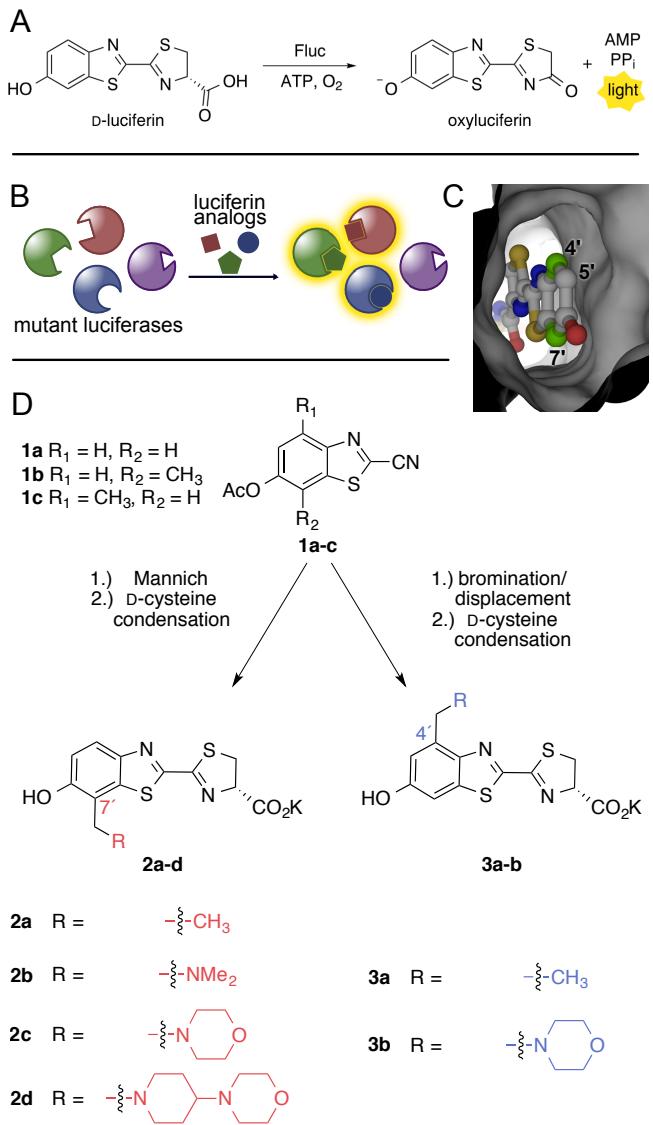


Figure 1.1: Expanding the bioluminescence toolkit with unique enzyme–substrate pairs. (A) Luciferase-mediated light production proceeds via an adenylation–oxidation sequence. (B) Strategy to develop orthogonal luciferase–luciferin pairs via substrate resolution. Genetically engineered luciferases and 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.

While versatile, bioluminescence to date has been largely limited to monitoring one cell type or biological feature at a time. This is due, in part, to a lack of distinguishable luciferase–luciferin pairs for *in vivo* use. The optimal luciferases (from the insect family) use the same substrate, D-luciferin.<sup>1,4</sup> Thus, they cannot easily discriminate multiple cell types in a single subject. Additionally, unlike fluorescent protein technologies, a diverse suite of accessible bioluminescent probes does not yet exist. To address this void, D-luciferin analogues have been engineered to emit different colors of light.<sup>5–7</sup> However, these substrates are still utilized by the same luciferases, precluding the distinct genetic tagging of individual cell types. Insect luciferases have also been engineered to emit different colors of light with D-luciferin.<sup>8–10</sup> The observed emission spectra are not sufficiently resolved, though, for routine use in complex tissues or animals. Discriminating among different wavelengths in bioluminescence (and whole-body optical imaging, in general) is exceedingly difficult.

Contrasting with these attempts to achieve spectral resolution, we aimed to obtain distinguishable bioluminescent probes via substrate resolution. Substrate-resolved bioluminescence is well precedented in nature, as structurally distinct luciferase–luciferin pairs have been identified across diverse phyla.<sup>11–13</sup> Some of these pairs, including those from the firefly and *Renilla reniformis*, have been used extensively.<sup>1,3,14,15</sup> Firefly (Fluc) and *Renilla* luciferase employ chemically unique substrates (D-luciferin and coelenterazine, respectively), enabling their tandem application *in vivo*.<sup>16,17</sup> Coelenterazine is less ideal for use in these environments, though, owing to its suboptimal bioavailability and stability.<sup>1,18</sup> Other naturally occurring luciferases and luciferins can be used in combination with Fluc/D-luciferin or other bioluminescent systems.<sup>17,19</sup> However, most of these native pairs remain poorly characterized or ill-suited for routine use.

Artificial (i.e., mutant) luciferases can exhibit altered bioluminescent properties, including tolerance for chemically modified substrates. Fluc itself has been manipulated to process analogues of D-luciferin.<sup>20</sup> In elegant work along these lines, Miller and co-workers prepared

a class of non-natural aminoluciferins that were found to be robust light emitters with Fluc, but the products inhibited the enzymatic reaction.<sup>21</sup> Product inhibition was relieved using mutated versions of the enzyme.<sup>22</sup> These same mutations also resulted in sharply reduced emission with D-luciferin, providing key precedent for the development and utilization of orthogonal pairs.<sup>23</sup> The mutant enzymes from these studies, though, were less selective for one analogue over another perhaps due to the structural similarities between the luciferin scaffolds. Simultaneous enzyme–substrate manipulation has also been applied to aequorin (a marine photoprotein) and the luciferase from the deep-sea shrimp *Oplophorus gracilirostris*.<sup>24</sup>? In both cases, altered bioluminescent outputs (e.g., colors and stabilities) were achieved, but orthogonal substrate usage was not realized.

Here, we report a strategy for the de novo production of orthogonal luciferase–luciferin pairs. We synthesized a series of sterically modified luciferins that were poor emitters with Fluc but intrinsically capable of robust light production. We then iteratively screened these analogues with libraries of mutant luciferases and identified substrate-selective enzymes. The “hits” were also biochemically characterized. Importantly, when the mutants and analogues were combined, robust light production was observed when complementary enzyme– substrate partners interacted. Sequential administration of substrates enabled unique luciferases to be illuminated (and thus resolved) within cultured cell models. These tools promise to enable a variety of multicellular imaging applications. Importantly, our approach to identifying orthogonal bioluminescence pairs is also general and should enable rapid diversification of the bioluminescence toolkit.

## 1.2 Results and Discussion

### 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 1.1B). We used the Fluc/D-luciferin pair as a starting point for several reasons. First, this duo is the most widely used in biomedical imaging applications owing to the nontoxicity of the reagents and bioavailability of the substrate.<sup>25,26</sup> Second, the Fluc/D-luciferin reaction releases the highest percentage of tissue-penetrating light among known bioluminescent families.<sup>27</sup> Thus, new enzymes and substrates based on the firefly pair would be more applicable to *in vivo* studies. Third, a wealth of structural and biochemical information on Fluc could guide our engineering efforts.<sup>13,28?</sup> Finally, D-luciferin derivatives are arguably the most synthetically tractable luciferin architectures.<sup>30?</sup>

Generating an expanded set of bioluminescent tools required access to diverse luciferin scaffolds. A variety of D-luciferin analogues have been synthesized over the past four decades,<sup>31? ? ? ? ,32</sup> and those capable of robust emission with Fluc harbor common features: an electron-donating group at the 6' position, a carboxylate appendage (for adenylation), and an abstractable proton  $\alpha$  to the carboxylate.<sup>33?</sup> Beyond these requirements, Fluc can tolerate a surprisingly large variety of modified luciferins,<sup>30? ?</sup> including 6'-amino substituents,<sup>20? ,21</sup> alkylated<sup>34? ?</sup> and acylated<sup>35</sup> scaffolds, and even luciferins with non-natural chromophores.<sup>6,35</sup> Crystallographic analyses have also corroborated these experimental results, indicating flexibility within the luciferase active site and “space” to accommodate luciferins with appendages at or near the 6' position.<sup>? ?</sup>

Unlike most efforts to produce luciferin analogues reported to date, we were attracted to the 4' and 7' positions of the luciferin core. These positions lie in close proximity to the Fluc

backbone (Figure 1.1C). Substrates with additional steric bulk at these sites would likely be occluded from the Fluc active site and thus good targets for orthogonal probe development: while poor emitters with the native enzyme, the molecules could potentially give off light with designer mutants. Indeed, preliminary docking studies suggested that only analogues with small (e.g., 2–3 atoms) substituents at C4' and C7' could effectively access the active site.

Generating 4'- and 7'-modified luciferins presented an early challenge. These positions have been rarely exploited for analogue development, and no prior syntheses were amenable to preparing libraries or large quantities of these probes. Rapid, high-yielding syntheses were essential, as large quantities of luciferins are required for light emission assays. Fortunately, the core benzothiazole unit (**1a–c**) of the desired analogues could be accessed from a common route (Figure 1.1D) and in multigram quantities.<sup>??</sup> From this single intermediate, we envisioned installing functional handles at C4' and C7' to rapidly assemble a variety of luciferins. We were initially drawn to an aldehyde group, owing to its ease of diversification under mild conditions (e.g., reductive amination) and broad compatibility. Aldehyde installation on **1a** was problematic, though, due to formation of a hydrated hemiacetal.<sup>36</sup> To circumvent this issue, we turned to more reactive iminium ions. These electrophiles can be readily trapped by electron-rich aromatics in a Mannich-type reaction.<sup>37</sup> Toward this end, benzothiazole **1a** was modified with a series of tertiary benzyl amines via in situ iminium formation and coupling (Figure 1.1D and Schemes S2 and S3). The amino appendages were selected to enhance the water solubility of the luciferin core. Importantly, this synthetic approach was modular and amenable to large-scale (1–10 g) syntheses. “Matched” probes with steric modifications at C4' were also prepared (**3a, b**). A different synthetic approach was necessary, though, as the 4' position cannot be selectively targeted with electrophiles (Figure 1.1D and Scheme S4).

## Analyzing bioluminescent light emission with modified luciferins

With the modified luciferins in hand, we first evaluated their optical properties with Fluc. All analogues were competent light emitters and could be processed by the enzyme (Figures 1.2A). However, the emission intensities were much weaker than those observed with

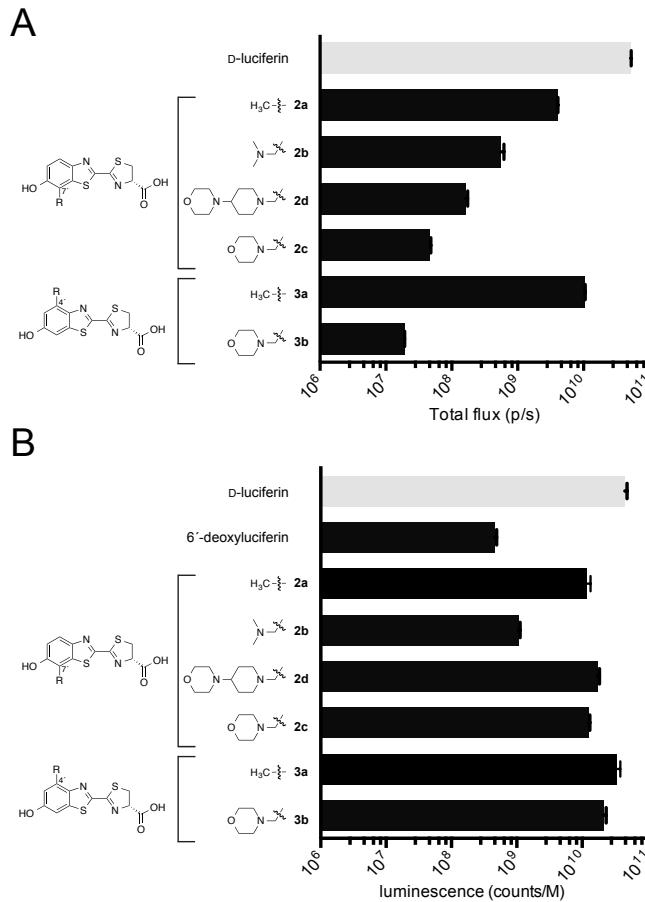


Figure 1.2: Measuring luciferin light emission. (A) Bioluminescence from luciferin analogues (100  $\mu$ M) incubated with 1  $\mu$ g of Fluc. Emission intensities are plotted as total photon flux values on a log scale. Error bars represent the standard deviation of the mean for  $n \geq 3$  experiments. (B) Chemiluminescence with luciferin analogues. Emission intensities are plotted as counts per molar luciferin on a log scale. Error bars represent the standard error of the mean for  $n \geq 3$  experiments.

D-luciferin, the native substrate. Interestingly, the largest analogue (**2d**) was not the weakest emitter, suggesting that steric modification alone does not dictate luciferin utilization. Similar trends in light emission were observed across a range of physiological pH values

(Scheme S3). Consistent with the observed light outputs, the measured kinetic constants for all analogues showed reduced performance relative to D-luciferin (Table S10). For example, the measured Km values were ~100-fold larger than the native substrate, with the largest analogues (**2c** and **2d**) exhibiting the lowest relative binding affinities. Despite their large Km values, **2b–d** exhibited emission spectra similar to that of D-luciferin (Figures S4 and S5). Only the C4'-modified analogue **3b** emitted noticeably red-shifted bioluminescent light, likely due to poor Fluc binding in the excited state<sup>7</sup> or the luminophore being forced into a more polar environment.<sup>38?</sup>

## Measuring the light-emitting potential of luciferin analogues

We attributed the weak bioluminescence of the analogues to poor utilization by Fluc. It was possible, though, that the luciferins were simply not capable of photon production upon activation and oxidation in the active site. For productive bioluminescence, an analogue must be able to reach an electronic excited state (S1) and relax back to the ground state with concomitant photon release.<sup>7</sup> If an analogue 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 analogues were intrinsically capable of light emission, we utilized a previously described chemiluminescence assay.<sup>39</sup> This process mimics the enzymatic reaction itself via formation of an activated ester intermediate, followed by proton abstraction and subsequent reaction with molecular oxygen.<sup>40?</sup> When analogues **2a–d** and **3a, b** were subjected to the assay, robust light emission was observed (Figure 1.2B). In fact, photon outputs for some of the weakest bioluminescent emitters (including **2c** and **3b**) were on par with D-luciferin. A control compound (**6'-deoxyluciferin**) lacking an electron-dense residue on the aromatic ring (a key feature of luciferins) exhibited only weak levels of emission. These results provided assurance that while luciferin scaffolds

may be poor substrates for Fluc, they are still capable of photon production and thus good candidates for orthogonal tool development.

## Evolving substrate-specific luciferases

Having prepared candidate orthogonal luciferins, we set out to identify mutant luciferases that could selectively process the molecules. Predicting enzyme mutations that confer substrate selectivity or otherwise beneficial properties is challenging. Fluc is a highly dynamic enzyme,<sup>41</sup>? complicating the selection of residues from static structural or sequence data. Moreover, amino acids known to play key roles in enzyme function have been identified far from the luciferin binding site;<sup>23</sup> such critical residues are often revealed only by random mutagenesis approaches.<sup>42,43</sup> Screening libraries of completely random mutants was impractical in our case, though, owing to the large library sizes needed to achieve adequate enzyme coverage.<sup>43</sup> Screening in bulk is also difficult as bioluminescent light emission is too weak to detect on conventional cell sorters or other high-throughput instruments. Thus, each enzyme—substrate combination must be physically segregated (to a certain extent) and interrogated for light emission with a sensitive camera.

Recognizing that manual screening necessitated the use of smaller libraries, we developed focused, semirational libraries where the mutations were confined to regions known to modulate substrate binding.<sup>44</sup> “Hits” from these smaller individual libraries could then be easily combined and assayed in subsequent library generations for improved function. We initially targeted residues 218, 249–251, and 314–316 for mutagenesis (Figure 1.3A). These selections were partially based on phylogenetic data gathered from across the insect luciferase family,<sup>13,45</sup> along with previous biochemical assays: Arg218 is known to interact with D-luciferin and influence the local structure of the binding pocket;<sup>28</sup> F250 lies in close proximity ( $\sim 3 \text{ \AA}$ ) to the benzothiazole ring of D-luciferin; T251 has been shown to potentiate substrate

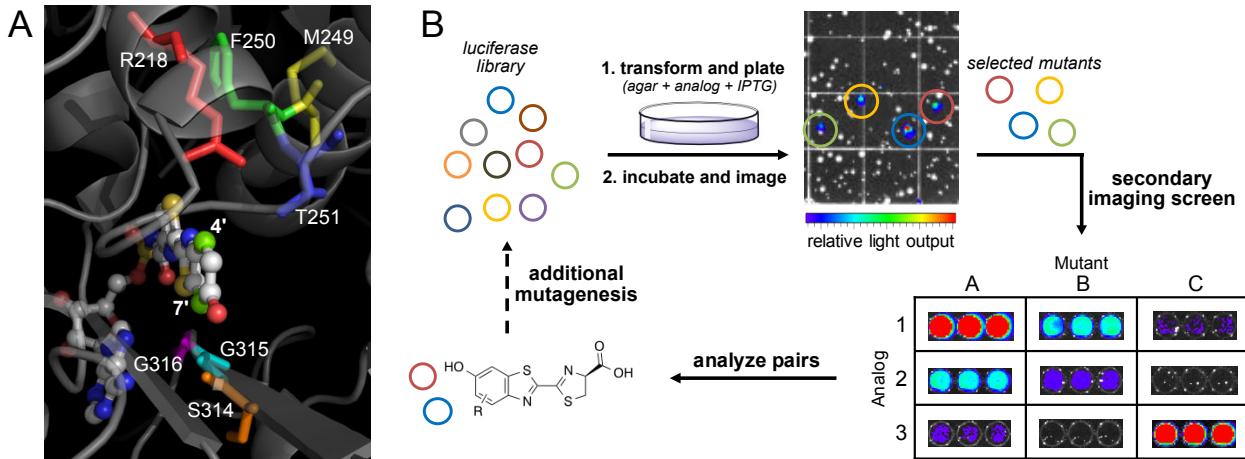


Figure 1.3: Generating mutant luciferase libraries and screening for orthogonal pairs. (A) Amino acids targeted for mutagenesis. These residues were selected based on their proximity to the 4' and 7' positions of luciferin. (B) Library screening strategy. An initial on-plate screen identified functional mutants. These “hits” were subjected to a secondary screen for orthogonality with other mutants and luciferin analogues.

binding;<sup>29</sup> residues 314–316 line a critical edge near the luciferin phenolate and C7' position. Mutations at all of these target sites have been shown to perturb D-luciferin binding (and thus light emission), while preserving the overall structural integrity of the enzyme.<sup>9,29,46</sup>

Saturation mutagenesis was used to prepare the desired libraries. The degree of mutation applied at each residue was based on the following considerations: sequence conservation among the insect luciferase family, the identity of the native residue, and the location of the residue. For example, nonconserved residues were mutated to a higher degree compared to conserved residues in the active site. Codon compression methods were further used to eliminate redundancies and reduce the number of transformants (Tables S1 and S2).<sup>47</sup> The final libraries ranged from 19 to 4800 members in size and were constructed using synthetic gene assembly<sup>48</sup> in combination with circular polymerase extension cloning (CPEC) (Tables S3–S9).Quan:2011fa

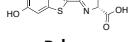
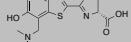
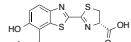
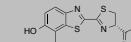
The libraries were screened for orthogonal substrate usage using a two-tiered approach. Library DNA was first introduced into bacteria, and the transformants were arrayed across agar plates containing embedded luciferins (Figure 1.3B). Light-emitting colonies were easily

identified, and in some cases, the light emission values were on par with native Fluc and D-luciferin. A handful of the corresponding mutants were sequenced. Some mutations were observed for multiple analogues, suggesting that they might be selective for bulky luciferins. Other mutations were unique to each compound, which is notable, given the subtle structural differences between some of the analogues. The number of colonies screened was  $\sim 3\times$  the calculated diversity for each library.

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

To mine the entire collection of imaging data for substrateselective pairs, we first developed a measure of orthogonality (equation shown in Figure 1.4A). Favorable values are obtained when two mutants (e.g., A and B) react robustly with unique substrates (e.g., compounds **1** and **2**, respectively) in a mutually exclusive manner. Thus, the more selective a pair of enzymes for their cognate substrates, the larger the orthogonality rating. Since the number of potential pairings exceeded 3000 in our data set, we wrote a computer script to rapidly examine all pairs in an unbiased fashion. The program iterated through each possible combination, calculating the corresponding orthogonality rating. The script ultimately returned a list of pairs ranked by their potential for orthogonality (and thus utility for multicomponent imaging).

Table 1.1: Biochemical analyses of orthogonal enzyme– substrate pairs

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

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

The top pairs identified by the script exhibited selectivity for analogues **3b** (mutants A and B) and **2b–d** (mutant C). The magnitude of each mutant’s preference—defined as the orthogonality quotient—was analyzed. As shown in Figure 1.4B and C, mutants A and B exhibited nearly a 100-fold preference for **3b** over other analogues, while mutant C strongly favored C7'-modified analogues. Similar trends in orthogonal substrate usage were observed using bacterial lysates (Figure S9) and across a range of luciferin concentrations (Figures S10 and S11). Biochemical analyses further indicated that the “brightest” mutant enzymes were those capable of most efficient substrate turnover (Table 1.1).

## Analyzing the origins of orthogonality

The identities of the mutant “hits” provided some insights into the origins of substrate orthogonality. Mutant A had a single arginine to alanine mutation at amino acid 218. Mutant B comprised the same R218A mutation but harbored additional mutations at residues 250

(Phe to Met), 314 (Ser to Thr), and 316 (Gly to Thr). These residues are known to play a role in modulating binding and interaction with the luciferin substrate. The R218A mutant is especially interesting, as it is known to greatly reduce light production and red shift emission with D-luciferin.<sup>29</sup> It has been hypothesized that the smaller Ala group allows more water molecules to access the active site, potentially quenching light emission.<sup>29</sup> The bulky morpholino substituent of 3b could fill this active site void to retain photon production. The third mutant (mutant C) was more selective for the C7'-modified luciferins compared to the C4'-modified compound. Mutant C harbored a single mutation, R218K. R218K may slightly enlarge the active site of the luciferase. This mutation has also been shown to boost activity with bulkier cyclic aminoluciferin analogues.<sup>22</sup> The improved selectivity with 2b–d could be the result of active site positioning. The C7' substituents could potentially place the luminophore in a more advantageous spot for light emission.

To delve into the origins of selectivity, we prepared a small library of additional mutants based on enzyme B (R218A, F250M, S314T, G316T). R218A seemed critical for discriminating the regioisomeric compounds, so this residue was held constant across the series. All possible combinations of the remaining mutations (F250M, S314T, G316T, or native Fluc residues) were then allowed. Imaging analyses of these combinatorial mutants indicated that R218A and F250M were critical for luciferin discrimination (Figures S12 and S13). Both mutations should result in a larger active site, but why they preferentially accommodate 3b over other analogues remains unknown. It is possible that the mutations disrupt critical binding interactions with the luciferin core, but that steric appendages (e.g., on the C4' side) retain sufficient contacts for subsequent oxidation. Indeed, when 3b was incubated with R218A/F250M, light emission was maintained (as compared to Fluc, Figure S13C). When D-luciferin and the C7'-modified analogues were incubated with this same mutant, though, light emission was drastically reduced. Interestingly, the R218A/S314T mutant exhibited an opposite trend in analogue selectivity: 2b was preferred to 3b (Figure S13C). 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.

## Cellular imaging with orthogonal pairs

As a step toward multicomponent imaging applications, we evaluated the orthogonal enzymes and probes in cultured cell models. Mammalian cell lines (HEK293 and DB7) were engineered to express orthogonal mutants A–C. Equivalent expression levels were confirmed using flow cytometry (Figure S14). Cells were then incubated with analogues 2b–d and 3b, and photon outputs were measured. As shown in Figure 1.5, 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 analogue 3b, whereas cells with mutant C were only visible upon treatment with analogues 2b–d (Figures S15–S18). Importantly, the orthogonal pairs could also distinguish unique cell types in a single imaging session. For example, DB7 cells stably expressing mutant B or C could be readily detected via sequential administration of the requisite substrates (Figure 5). Similar trends were observed upon imaging HEK293 cells (Figure S19) and cocultures (Figure S20). These data suggest that crossreactivity between mutants B and C and their non-orthogonal substrates is minimal. The orthogonal pairs also exhibit unique emission spectra (Figure S21) that can further enhance some multicomponent imaging applications (Figure S22).

## 1.3 Conclusions

We developed a general strategy to evolve and identify mutant versions of firefly luciferase that accept distinct, chemically modified luciferins. Bioluminescence has been largely limited

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

Future studies will be aimed at generating additional bioluminescent probes with improved brightness and other optical properties. The enzyme–substrate “hits” reported here, while immediately useful, are weaker light emitters than native bioluminescent systems. Improved light outputs can be achieved using additional rounds of mutagenesis and screening. Previous studies have also demonstrated that distant mutations can profoundly influence the architecture of the luciferase active site, and these regions will be incorporated into future libraries. The screening strategy is also broadly applicable to diverse luciferins, including analogues with altered chromophores that could provide drastically different colors of light. Our results suggest that enzymes capable of discriminating even subtle substrate modifications can be readily identified. Such an outcome bodes well for generating additional orthogonal pairs and filling a long-standing void in imaging capabilities. We anticipate that collections of designer luciferins and luciferases will inspire new discoveries in a variety of disciplines, similar to how fluorescent protein technology enabled seminal advancements in numerous fields.

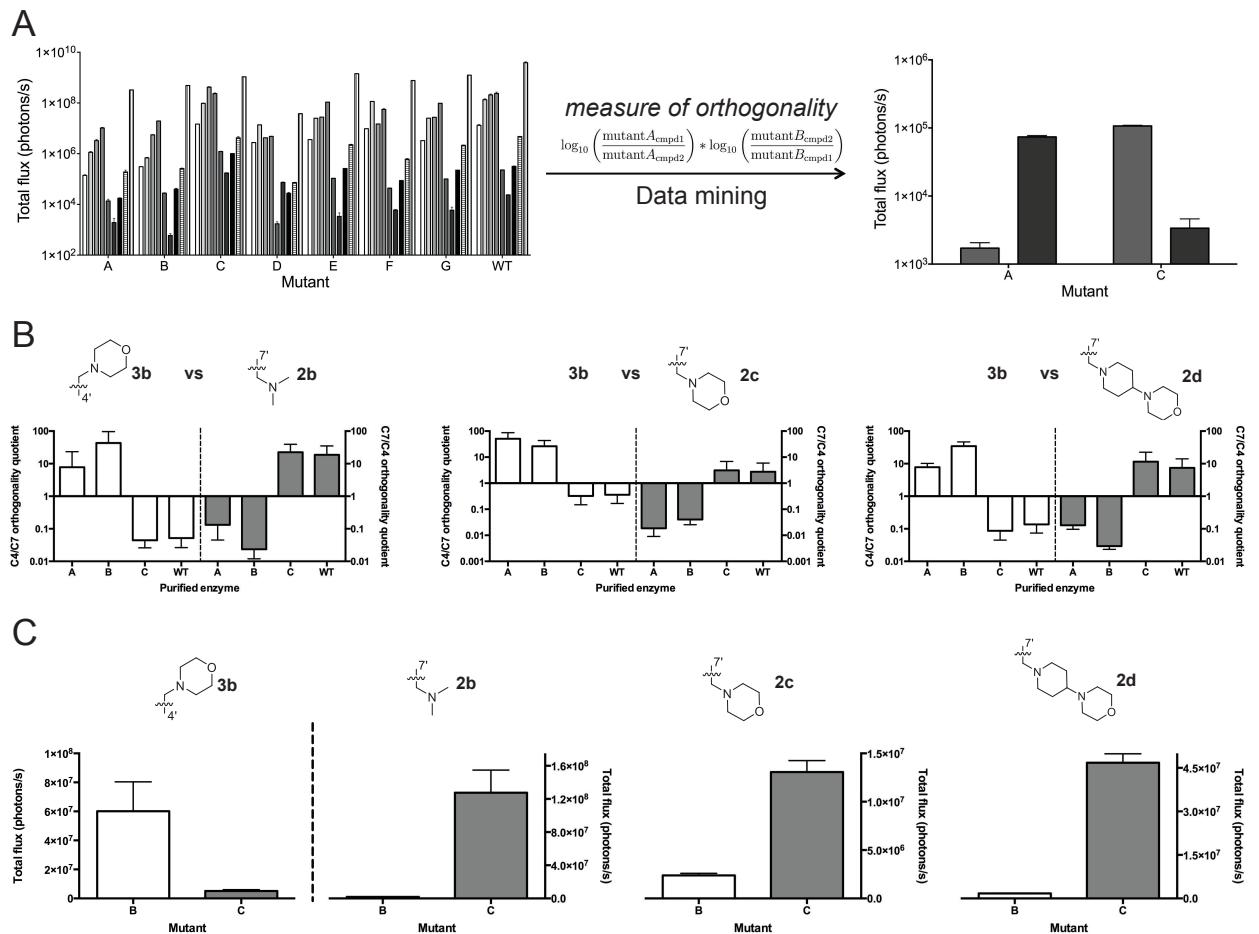


Figure 1.4: Analyzing orthogonal enzyme–substrate pairs. (A) Representative emission of luciferase mutants screened against a panel of luciferin analogues. These data were analyzed with a computer algorithm to determine lead mutants with the strongest orthogonality. (B) Purified mutants exhibit orthogonality. Enzyme ( $1 \mu\text{g}$ ) was incubated with  $100 \mu\text{M}$  of luciferin analogues, and emission intensities were used to determine the orthogonality quotient (the ratio of the total flux for the C4/C7 or C7/C4 pairings). The geometric mean is plotted, and the error bars represent the 95% confidence intervals for  $n > 4$  experiments. (C) Total flux for lead mutants B and C highlights substrate selectivity between C4' and C7' sterically modified luciferins. Error bars represent the standard error of the mean for  $n > 4$  experiments.

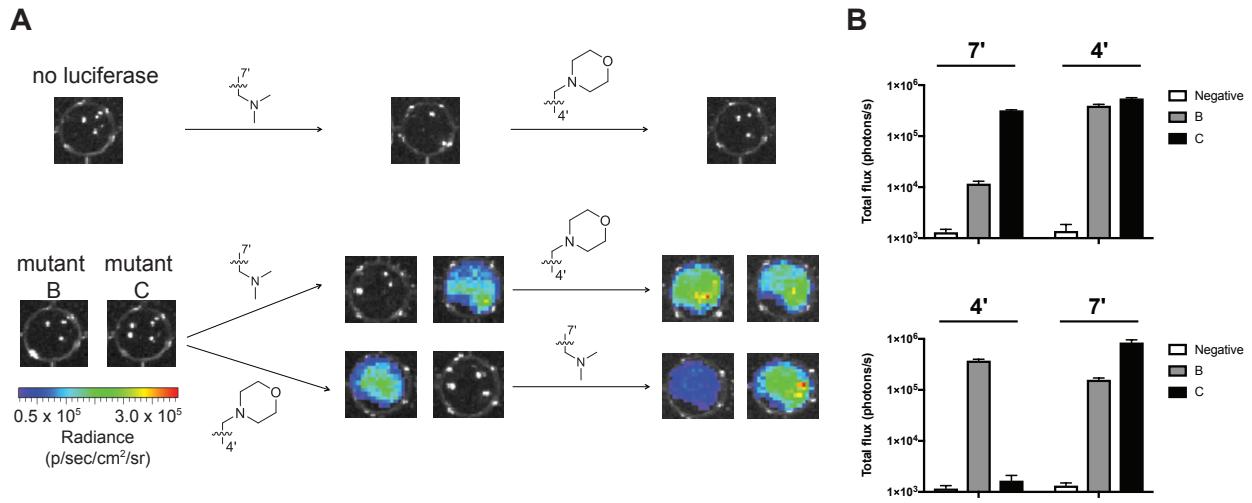


Figure 1.5: Imaging cells with orthogonal luciferase–luciferin pairs. (A) Mutant luciferase-expressing DB7 cells were plated ( $1.5 \times 10^5$  cells/well) in 96-well black plates and sequentially incubated with C4' and C7' sterically modified luciferins ( $750 \mu\text{M}$ ). Representative bioluminescence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

# Bibliography

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

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

- [24] Rowe, L.; Dikici, E.; Daunert, S. *Anal. Chem.* **2009**, *81*, 8662–8668.
- [25] Berger, F.; Paulmurugan, R.; Bhaumik, S.; Gambhir, S. S. *Eur. J. Nucl. Med. Mol. Imaging* **2008**, *35*, 2275–2285.
- [26] Contag, C. H.; Spilman, S. D.; Contag, P. R.; Oshiro, M.; Eames, B.; Dennery, P.; Stevenson, D. K.; Benaron, D. A. *Photochem. Photobiol.* **1997**, *66*, 523–531.
- [27] Zhao, H.; Doyle, T. C.; Coquoz, O.; Kalish, F.; Rice, B. W.; Contag, C. H. *J. Biomed. Opt.* **2005**, *10*, 041210.
- [28] BRANCHINI, B. R.; Magyar, R. A.; Murtiashaw, M. H.; Portier, N. C. *Biochemistry* **2001**, *40*, 2410–2418.
- [29] Branchini, B. R.; Southworth, T. L.; Murtiashaw, M. H.; Boije, H.; Fleet, S. E. *Biochemistry* **2003**, *42*, 10429–10436.
- [30] Meroni, G.; Rajabi, M.; Santaniello, E. *Arkivoc* **2009**, 265–288.
- [31] WHITE, E. H.; WORTHER, H.; SELIGER, H. H.; MCELROY, W. D. *J. Am. Chem. Soc.* **1966**, *88*, 2015–&.
- [32] Woodrooffe, C. C.; Meisenheimer, P. L.; Klaubert, D. H.; Kovic, Y.; Rosenberg, J. C.; Behney, C. E.; Southworth, T. L.; Branchini, B. R. *Biochemistry* **2012**, *51*, 9807–9813.
- [33] White, E. H.; Wörther, H. **1966**, *31*, 1484–1488.
- [34] Wang, J.; Lee, T. S.; Zhang, Z.; Tung, C.-H. *Chem Asian J* **2017**, *12*, 397–400.
- [35] Kuchimaru, T.; Iwano, S.; Kiyama, M.; Mitsumata, S.; Kadonosono, T.; Niwa, H.; Maki, S.; Kizaka-Kondoh, S. *Nat. Commun.* **2016**, *7*, 11856.
- [36] Jones, P. R.; Gelinas, R. M. *J. Org. Chem.* **1981**, *46*, 194–196.
- [37] Phillips, J. P.; Barrall, E. M. *J. Org. Chem.* **1956**, *21*, 692–694.

- [38] Hirano, T.; Nagai, H.; Matsuhashi, T.; Hasumi, Y.; Iwano, S.; Ito, K.; Maki, S.; Niwa, H.; Viviani, V. R. *Photochemical & Photobiological Sciences* **2012**, *11*, 1281–1284.
- [39] Steinhardt, R. C.; Rathbun, C. M.; Krull, B. T.; Yu, J. M.; Yang, Y.; Nguyen, B. D.; Kwon, J.; McCutcheon, D. C.; Jones, K. A.; Furche, F.; Prescher, J. A. *ChemBioChem* **2016**,
- [40] Kato, D.-i.; Shirakawa, D.; Polz, R.; Maenaka, M.; Takeo, M.; Negoro, S.; Niwa, K. *Photochem. Photobiol. Sci.* **2014**, *13*, 1640–1645.
- [41] Mao, Y. *Protein Engineering Design and Selection* **2011**, *24*, 341–349.
- [42] Chen, M. M. Y.; Snow, C. D.; Vizcarra, C. L.; Mayo, S. L.; Arnold, F. H. *Protein Engineering Design and Selection* **2012**, *25*, 171–178.
- [43] Reetz, M. T.; Prasad, S.; Carballeira, J. D.; Gumulya, Y.; Bocola, M. *J. Am. Chem. Soc.* **2010**, *132*, 9144–9152.
- [44] Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z.-G.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P. *ACS Synth. Biol.* **2012**, *2*, 83–92.
- [45] Amaral, D. T.; Arnoldi, F.; Rosa, S. P.; Viviani, V. R. *Luminescence* **2014**, *29*, 412–422.
- [46] Viviani, V. R.; Amaral, D. T.; Neves, D. R.; Simoes, A.; Arnoldi, F. G. C. *Biochemistry* **2013**, *52*, 19–27.
- [47] Pines, G.; Pines, A.; Garst, A. D.; Zeitoun, R. I.; Lynch, S. A.; Gill, R. T. *ACS Synth. Biol.* **2014**, 141030112600008.
- [48] Ness, J. E.; Kim, S.; Gottman, A.; Pak, R.; Krebber, A.; Borchert, T. V.; Govindarajan, S.; Mundorff, E. C.; Minshull, J. *Nat. Biotechnol.* **2002**, *20*, 1251–1255.