DIRECTED EVOLUTION OF THE
TRANSCRIPTIONAL ACTIVATOR LUXR

Thesis by
Cynthia H. Collins

In Partial Fulfillment of the Requirements for the
Degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2006

(Defended December 19, 2005)
ACKNOWLEDGEMENTS

I would like to thank my advisor, Frances Arnold, for the opportunity to work on a truly interesting project from its inception. I thank her for affording me the independence to develop my project and for her enthusiasm and support. Her effort to teach all of the skills required for a successful scientific career has been remarkable. I gratefully acknowledge the wonderful example of how to manage, write, teach, inspire, and truly enjoy science that she has provided throughout my time at Caltech.

I would like to thank Jared Leadbetter, Carl Parker, and Michael Elowitz for serving on my thesis committee. I am grateful to Jared for being an invaluable resource and great mentor. He has played a central role in my research from the beginning, and has remained a constant source of enthusiasm and critical discussion. I thank Carl for allowing me to spend a good portion of the last six months working in his lab. He has patiently answered myriad questions and has helped me become a better biochemist.

I would also like to acknowledge all of the people who have been a part of the Arnold lab and made it such an interesting place to work. Special recognition goes to Yohei Yokobayashi, Joff Silberg, Kim Meyer, Daisuke Umeno, Matthew Peters, Lingchong You, Geethani Bandara, Manish Raizada, Katie Brenner, Chris Otey, and Cara Tracewell. I found the dedication and enthusiasm of my colleagues to be particularly inspiring during my time here.
I do not think I would have survived the past five years without the strong women with whom I have been lucky enough share the struggles and successes of graduate school. I thank Karli, Laura, Nazli and Sarah for being great friends and for sharing “girl’s lunch” with me. Julie B., Betsy and Rachel deserve many thanks for their friendship during my time at Caltech.

My friend Sean has been amazingly supportive and caring. I am truly grateful for his friendship.

Finally, I am most grateful to my parents for their love and support throughout this challenging endeavor.
ABSTRACT

LuxR-type proteins are responsible for density-dependent transcriptional regulation in quorum-sensing systems that employ acyl-homoserine lactones (acyl-HSLs) as signal molecules. The *Vibrio fischeri* *lux* operon, which includes LuxR, has been well studied and shown to function in *E. coli*. The components of the *lux* operon have recently been used to engineer *de novo* genetic circuits because they provide a versatile intercellular communications system. We have used directed evolution to engineer LuxR to expand the “toolbox” of components available for engineering synthetic circuits with a wide range of communications functions and to explore the evolution of LuxR specificity and response.

We developed and validated a screening system to identify mutants of LuxR that activate gene expression with noncognate acyl-HSLs, based on LuxR-mediated gene expression of GFP. We screened libraries of *luxR* mutants for variants exhibiting increased gene activation in response to octanoyl-HSL (C8HSL), with which wild-type LuxR interacts only weakly. Eight LuxR variants were identified that showed a 100-fold increase in sensitivity to C8HSL; these variants also displayed increased sensitivities to a broadened range of acyl-HSLs, including straight-chain acyl-HSLs, while maintaining a wild-type or greater response to LuxR’s cognate acyl-HSL, 3-oxo-hexanoyl-homoserine lactone (3OC6HSL).
LuxR and its homologs respond to their cognate acyl-HSL signals with exquisite specificity. To generate a LuxR with a new signaling specificity, we used a dual selection system to identify LuxR variants that either activate gene expression (ON) or do not (OFF) under desired sets of conditions. The dual selection system was evaluated prior to its use, and a 490-fold enrichment in functional clones per round of ON/OFF selection was observed. We subsequently used the dual selection system to identify a new LuxR variant that retains the response to straight-chain acyl-HSLs attained by screening for activation with C8HSL, but no longer responds to 3OC6HSL. A single mutation in the N-terminal signal-binding domain reduces the response to acyl-HSLs having a carbonyl substituent at the third carbon of the acyl chain. The specificity-enhancing mutation does not affect LuxR response to straight-chain acyl-HSLs and would not have been identified by positive selection alone.

We also used the dual selection system to identify LuxR variants that activate transcription upon binding to a promoter containing a mutated operator site, at which LuxR is unable to activate transcription. An amino acid position critical for modulating the DNA-binding specificity of LuxR for the new DNA target was identified. The dual selection system provides a rapid and reliable method for identifying LuxR variants that have the desired response, or lack thereof, to a given set of acyl-HSL signals or DNA targets. LuxR variants with new specificities identified in this work are potentially useful components for constructing artificial cell-cell communication systems that program population-level behaviors.
Many species of bacteria use the synthesis and recognition of small molecules to assess changes in their local population density and alter the expression of key genes accordingly. The components of these systems have been harnessed for engineering synthetic genetic circuits that include elements of intercellular communication. We have used directed evolution to engineer LuxR, a transcriptional activator from Vibrio fischeri responsible for density-dependent signal recognition and modulation of gene expression. Both the signal binding and signal binding specificities of LuxR have been engineered in order to expand the ‘toolbox’ of components available for engineering synthetic circuits with a wide range of communications functions and to explore the evolution of LuxR specificity and response.

Chapter 1 of this thesis provides an introduction to LuxR and its homologs and to the methods used in the studies described in this thesis. The general mechanism and components required for acyl-homoserine lactone (acyl-HSL)-based quorum sensing (QS) will be presented using the Vibrio fischeri lux regulon as a model system. Similar QS systems have been identified within the Gram-negative phylum Proteobacteria and the synthesis and recognition of diverse acyl-HSLs provides different bacterial species with their own chemical languages. Engineering of the QS transcriptional regulator LuxR is the subject of this work and the studies previously used to characterize this protein will be reviewed in detail. The diversity of LuxR homologs found in species of
α-, β-, and γ-Proteobacteria will also be introduced. Previous efforts to engineer LuxR, its homologs, and other transcription factors will also be discussed.

The lux system has been used to engineer de novo regulatory circuits requiring intercellular communication. LuxR homologs have been investigated, but these can show significant crosstalk and often also behave differently from LuxR in other ways. For applications in synthetic biology, it would be useful to have a set of ‘standardized parts’—for example, a series of LuxR variants that respond to different chemical signals or activate gene expression at different promoters from which new networks can be assembled.

A postdoc in our laboratory, Yohei Yokobayashi, initiated the work towards applying the techniques of directed evolution to the engineering of transcription factors and synthetic genetic circuits. He showed that a detuned regulatory network, wherein the λ repressor CI was ubiquitously preventing expression at a target promoter, could be converted to a functional signal-dependent inverter through directed evolution of the CI protein. In the process, a library of genetic devices, with a range of new behaviors, was generated. I decided to apply a similar approach to the engineering of LuxR and initiated the project by targeting its response to a non-cognate acyl-HSL, octanoyl-HSL (C8HSL). LuxR requires 200-fold higher concentrations of C8HSL to achieve levels of activation observed with LuxR’s cognate signal, 3-oxohexanoyl-homoserine lactone, 3OC6HSL.
Chapter 2 describes the implementation of a LuxR-dependent screen and the identification of LuxR variants with increased sensitivity to C8HSL. A LuxR reporter system in which gene activation leads to the production of green fluorescent protein (GFP) was constructed wherein high-throughput screening of LuxR-expressing colonies is achieved by exciting a plate of colonies with ultraviolet light and visually inspecting colonies for emission of green light. A library of luxR mutants was generated by random mutagenesis and screened for gfp activation with 200 nM C8HSL. Wild-type LuxR requires approximately 2 μM C8HSL to achieve the level of gene activation obtained with 10 nM 3OC6HSL. Six C8HSL-responsive LuxR variants were identified from a library of approximately 20,000 clones. These mutants were recombined and a library of approximately 1200 clones was screened for their ability to activate gene expression with 50 nM C8HSL. Eight LuxR variants were identified that showed a 100-fold increase in sensitivity to C8HSL; these variants also displayed increased sensitivities to pentanoyl-HSL and tetradecanoyl-HSL, while maintaining a wild-type or greater response to 3OC6HSL. The most sensitive variants activated transcription as strongly with C8HSL as the wild type did with 3OC6HSL. Three ‘hypersensitive’ variants were also identified that showed increased activation with 3OC6HSL.

All of the C8HSL-responsive LuxR variants identified showed broadened acyl-HSL specificity profiles and all responded to 3OC6HSL as well, or better, than wild type. We wanted to continue the process of engineering LuxR variants by increasing its specificity for a new acyl-HSL signal. To do this we needed to be able to identify LuxR variants
that both activate gene expression with the new target and do not respond to 3OC6HSL. Chapter 3 describes the implementation and evaluation of a dual selection system capable of identifying LuxR variants that either activate gene expression (positive selection, ON) or do not (negative selection, OFF) under desired sets of conditions. The ON/OFF selection system was adapted from a system developed by Y. Yokobayashi for use with the λ CI repressor. Modification of the dual selection system for use with an activator was required. To test the modified ON/OFF selection system we selected a plasmid encoding wild-type LuxR from a background of nonfunctional clones and showed that a single round of ON/OFF selection yields a 490-fold enrichment of functional clones. This result, which suggests that a functional LuxR variant can be identified from a library of 100 million in only three rounds of ON/OFF selection, demonstrated the potential of our dual selection system and led us to begin selecting libraries for LuxR variants exhibiting new acyl-HSL specificities with confidence that we could identify the desired variant if it existed in the library.

The parent used for the evolution of acyl-HSL specificity was LuxR-G2E, which activates gene expression upon binding a broadened range of acyl-HSLs, including C10HSL. The application of the ON/OFF selection system to identify a LuxR variant that retains the increase in sensitivity to C10HSL achieved by LuxR-G2E, but that has lost its ability to recognize and respond to LuxR’s cognate signal 3OC6HSL is presented in chapter 4. A library was generated by random mutagenesis and LuxR variants were selected for the ability to survive ON rounds of selection with C10HSL and OFF rounds
of selection with 3OC6HSL. A single mutation, arginine 67 to methionine in the N-terminal signal-binding domain, reduces LuxR-G2E’s response to acyl-HSLs having a carbonyl substituent at the third carbon of the acyl chain. This mutation has a similar effect when introduced into a LuxR homolog that responds to a different 3-oxo-HSL (3OC12HSL), LasR from *Pseudomonas aeruginosa*. This specificity-enhancing mutation does not affect LuxR response to straight-chain acyl-HSLs and would not have been identified by positive selection for activation by these signaling molecules. This work demonstrates that accumulation of a small number of point mutations is sufficient to switch specificity from one signal to another, passing through a broadly-specific intermediate. We expect that this dual selection system could be used to further refocus the specificity towards a single acyl-HSL by selecting for the absence of gene activation with multiple acyl-HSLs during OFF rounds of selection.

While LuxR and its homologs specifically recognize very diverse acyl-HSL signal molecules, they can often activate gene expression at each other’s promoters. LuxR variants with new DNA binding specificities would be a powerful addition to the components available to synthetic biologists. The ability to use a single cell type to recognize and specifically respond to different acyl-HSLs by activating gene expression at different promoters would increase the capabilities of current communications systems. Also, by generating the components in the lab from a well-characterized, well-behaved starting point, one can target the desired properties from the beginning. To address this challenge, we applied our dual selection system to modify the DNA-binding specificity
Chapter 5 describes the selection of LuxR variants that activate gene expression at a mutated *lux* promoter at which LuxR does not activate gene expression. We generated a library by random mutagenesis and selected mutants activate gene expression at the new promoter. We also selected for decreased activation at the wild-type *lux* promoter. The two LuxR variants identified contain a mutation at amino acid position 212, from an arginine to either a histidine or a cysteine. The R212C mutation was accompanied by a second mutation, Q232H, and these two mutations are shown to act synergistically to confer the desired change in DNA-binding specificity. The promoter specificity of LuxR-R212C/Q232H was broadened while LuxR-R212H showed a preference for the new promoter. Both variants required approximately 100-fold higher concentrations of 3OC6HSL to achieve levels of activation observed with wild-type LuxR and the *lux* promoter. We have generated the first LuxR variants that respond to a new promoter and shown that random mutagenesis can be used to identify residues critical for DNA-binding specificity.
# Table of Contents

**ACKNOWLEDGEMENTS** ........................................................................................................ iii
**ABSTRACT** ............................................................................................................................... v
**TABLE OF CONTENTS** ........................................................................................................... xiii
**LIST OF TABLES** .................................................................................................................... xvii
**LIST OF FIGURES** ................................................................................................................ xviii
**ABBREVIATIONS** .................................................................................................................. xxii

## CHAPTER

1. Introduction ............................................................................................................................... 1
   
   **A. Genetic circuit engineering and intercellular communication** ................................. 2
   
   **B. Acyl-homoserine lactone dependent transcriptional activation** ................................. 6
      
      B.1. Mechanism of QS activation in *V. fischeri* ............................................................... 6
      B.2. Homologous acyl-HSL dependent QS systems from the α-, β- and γ-Proteobacteria ......................................................................................................................... 8
   
   **C. Elements of acyl-homoserine lactone-dependent quorum sensing systems** ...... 11
      
      C.1. The lux operon ................................................................................................................. 11
      C.2. Acyl-HSL synthesis by LuxI .......................................................................................... 13
      C.3. Acyl-HSL-mediated gene activation by LuxR homologs ........................................... 15
         
         C.3.1. Acyl-HSL response .................................................................................................. 16
         C.3.2. DNA binding and activation ................................................................................... 19
   
   **D. Engineering bacterial transcription factors** ................................................................. 21
      
      D.1. Directed evolution and rational design ...................................................................... 22
      D.2. Screening and selection ............................................................................................... 23
      D.3. Engineering signal binding ........................................................................................ 26
   
   **E. References** ..................................................................................................................... 31

2. Directed Evolution of *Vibrio fischeri* LuxR for Increased Sensitivity to a Broad Spectrum of Acyl-homoserine Lactones ................................................................................. 42
   
   **A. Abstract** ......................................................................................................................... 43
   
   **B. Introduction** ................................................................................................................. 44
   
   **C. Results** .......................................................................................................................... 47
      
      C.1. Cloning vectors and screening tools ......................................................................... 47
      C.2. First-generation laboratory evolution–error-prone PCR ........................................... 47
Appendix

A. Plasmid, gene and protein sequences ........................................... 173

A. Sequences of wild-type LuxR, LuxR homologs, evolved LuxR variants and the pLuxR expression vector .................................................. 174
  A.1. General information ............................................................... 174

B. Sequences of the fluorescent reporters used for screening and characterization of LuxR variants ................................................. 186
  B.1. General information ............................................................... 186

C. Sequences of plasmids, genes and proteins comprising the dual selection system for the directed evolution of LuxR ........................................... 194
  C.1. General information ............................................................... 194
LIST OF TABLES

TABLE PAGE

2.1. Nucleotide and amino acid changes in the recovered luxR/LuxR mutants .......... 50
2.2. Activation of gfpuv gene expression by LuxR variants with four different acyl-HSLs. .................................................................................................................. 51
2.3. Specificity of second-generation LuxR variants. ........................................... 58
2.4. Activation of gfpuv gene expression by luxR single mutants ...................... 60
2.5. Oligonucleotide primers used in Chapter 2 ............................................... 72

3.1. Oligonucleotide primers used in Chapter 3. ............................................. 102

4.1 Nucleotide and amino acid changes in luxR/LuxR and lasR/LasR mutants. ....... 114

A.1. Nucleotide and amino acid changes in the recovered luxR/LuxR mutants described in Chapter 2. ........................................................................................................ 182
A.2. Single-mutants mutants described in Chapter 2. ........................................ 183
A.3. Nucleotide and amino acid changes in luxR/LuxR and lasR/LasR mutants described in Chapter 4. ........................................................................................................ 184
A.4. Nucleotide and amino acid changes in luxR/LuxR variants described in Chapter 5 ...................................................................................................................... 185
## List of Figures

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1. Schematic of the <em>Vibrio fischeri</em> quorum sensing system.</td>
<td>7</td>
</tr>
<tr>
<td>1.2. Diversity of LuxR homologs and acyl-HSLs found in Nature</td>
<td>9</td>
</tr>
<tr>
<td>1.3. Structures, non-IUPAC names and abbreviations of the acyl-HSLs that activate the homologs listed in Fig. 1.2</td>
<td>10</td>
</tr>
<tr>
<td>1.4. <em>Vibrio fischeri lux</em> gene organization</td>
<td>12</td>
</tr>
<tr>
<td>1.5. Schematic of acyl-HSL synthesis by LuxI from acyl-acyl-carrier protein and S-adenosyl-L-methionine</td>
<td>14</td>
</tr>
<tr>
<td>1.6. Chemical structures of tetracycline derivatives used in the directed evolution of TetR</td>
<td>27</td>
</tr>
<tr>
<td>2.1. Amino acid sequence changes to first- and second-generation LuxR variants showing increased gene activation in response to C8HSL</td>
<td>49</td>
</tr>
<tr>
<td>2.2 Detection of C8HSL and 3OC6HSL by wild-type LuxR and eight second-generation LuxR variants</td>
<td>54</td>
</tr>
<tr>
<td>2.3. Protein concentrations of second-generation LuxR variants</td>
<td>56</td>
</tr>
<tr>
<td>2.4. Mutations in LuxR responsible for altering acyl-HSL specificity differ from those which result in a loss of function</td>
<td>63</td>
</tr>
<tr>
<td>2.5. Multiple sequence alignments of 19 members of the LuxR family of transcriptional regulators</td>
<td>65</td>
</tr>
</tbody>
</table>
2.6. Hypothetical positions of residues in TraR corresponding to those found to modulate acyl-HSL specificity in LuxR ......................................................... 68

3.1. Schematic of the dual selection system for the directed evolution of LuxR variants with altered specificity ................................................................. 91

3.2. Plasmid maps of plasmids used to test dual selection system .................. 93

3.3. Agarose gel of colony PCR products used to determine the composition of the plasmid mixture after each round of positive and negative selection .... 96

4.1. Diversity of LuxR homologs and acyl-HSLs ........................................ 109

4.2. Activation of $gfpuv$ transcription with 3OC6HSL, C10HSL, C6HSL, C8HSL, C12HSL and 3OC12HSL by wild-type LuxR, LuxR-G2E and LuxR-G2E-R67M ......................................................................................................................... 116

4.3. Activation of $gfpuv$ transcription with 3OC6HSL and C6HSL by wild-type LuxR and LuxR-R67M .................................................................................. 117

4.4. Activation of $gfpuv$ transcription with 3OC12HSL and C12HSL by wild-type LasR and LasR-R61M ..................................................................................... 118

4.5. Gel mobility shift assays of LuxR, LuxR-G2E and LuxR-C10Sp in the presence of 3OC6HSL and C10HSL ........................................................................... 120

4.6. Protein accumulation is dependent upon acyl-HSL binding .................. 121

4.7. Solid-phase assays show minimal crosstalk between LuxR and LuxR-G2E-R67M .............................................................................................. 122
5.1. List of LuxR homologs and their cognate DNA targets, and the sequences of the wild-type lux box and the mutant, luxG5A, used in this study......................139

5.2. Schematic of the dual selection system for the directed evolution of LuxR variants with altered DNA-binding specificity...................................................141

5.3. Activation of gfpuv transcription from P_{luxI} and P_{luxG5A} by wild-type LuxR, LuxR-R212C/Q232H, LuxR-R212H and LuxR-R212S with 3OC6HSL............147

5.4. Activation of gfpuv transcription from P_{luxI} and P_{luxG5A} by single-mutants LuxR-R212C and LuxR-Q232H with 3OC6HSL................................................148

5.5. Protein accumulation is dependent upon acyl-HSL binding..........................149

5.6. Electromobility shift assays of LuxR and the evolved variants with P_{luxI} and P_{luxG5A}...................................................................................................................151

5.7. RhlR-mediated GFPuv gene expression at P_{luxI} and P_{luxG5A} .......................153

5.8. Alignment of residues in TraR corresponding to residues 212 and 232 in LuxR ....................................................................................................................154

A.1. Nucleotide sequence of full-length, wild-type luxR..............................................176

A.2. Amino acid sequence of full-length, wild-type LuxR..............................................176

A.3. Nucleotide sequence of the pLuxR vector containing wild-type luxR ..............177

A.4. Plasmid map of the pLuxR vector containing wild-type luxR..............................179

A.5. Nucleotide sequences of the luxR homologs used in this work ......................180

A.6. Amino acid sequences of the LuxR homologs used in this work......................181

A.7. Nucleotide sequence of the wild-type P_{luxI} promoter......................................188

A.8. Nucleotide sequences of the wild-type lux box and mutant luxG5A box ............188
A.9. Nucleotide sequences of the fluorescent reporters, gfpuv, gfpmut3 and dsred-express, used in this study .......................................................... 189
A.10. Amino acid sequence of the fluorescent reporters GFPuv, GFPmut3 and DsRed-express, used in this study .......................................................... 191
A.11. Nucleotide sequence of the pluxGFPuv reporter plasmid .......................................................... 192
A.12. Plasmid map of the pluxGFPuv reporter plasmid .......................................................... 193
A.13. Nucleotide sequence of the genes used in the dual selection system .................. 196
A.14. Amino acid sequence of the proteins used in the dual selection system ............ 198
A.15. Nucleotide sequence of the positive (ON) selection plasmid, pluxCAT, containing the wild-type P\textsubscript{luxI} promoter .......................................................... 199
A.16. Plasmid map of the pluxCAT selection plasmid containing the wild-type P\textsubscript{luxI} promoter .......................................................... 201
A.17. Nucleotide sequence of the negative (OFF) selection plasmid, pluxBLIP, containing the wild-type P\textsubscript{luxI} promoter .......................................................... 202
A.18. Plasmid map of the pluxCAT selection plasmid containing the wild-type P\textsubscript{luxI} promoter .......................................................... 203
A.19. Nucleotide sequence of the selection test plasmid pBLIPCAT .......................... 204
A.20. Plasmid map of the selection test plasmid pBLIPCAT .......................... 206
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>QS</td>
<td>Quorum sensing</td>
</tr>
<tr>
<td>Tc</td>
<td>Tetracycline</td>
</tr>
<tr>
<td>Cm</td>
<td>Chloramphenicol</td>
</tr>
<tr>
<td>Cr</td>
<td>Carbenicillin</td>
</tr>
<tr>
<td>bp</td>
<td>Base pairs</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>HTH</td>
<td>$\alpha$-helix-turn-$\alpha$-helix</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-$\beta$-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>Acyl-HSL</td>
<td>Acyl-homoserine lactone</td>
</tr>
<tr>
<td>C4HSL</td>
<td>Butanoyl-homoserine lactone</td>
</tr>
<tr>
<td>C5HSL</td>
<td>Pentanoyl-homoserine lactone</td>
</tr>
<tr>
<td>C6HSL</td>
<td>Hexanoyl-homoserine lactone</td>
</tr>
<tr>
<td>3OC6HSL</td>
<td>3-Oxo-hexanoyl-homoserine lactone</td>
</tr>
<tr>
<td>C8HSL</td>
<td>Octanoyl-homoserine lactone</td>
</tr>
<tr>
<td>C10HSL</td>
<td>Deconoyl-homoserine lactone</td>
</tr>
<tr>
<td>C12HSL</td>
<td>Dodecanoyl-homoserine lactone</td>
</tr>
<tr>
<td>3OC12HSL</td>
<td>3-Oxo-dodecanoyl-homoserine lactone</td>
</tr>
<tr>
<td>C14HSL</td>
<td>Tetradecanoyl-homoserine lactone</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction
A. Genetic circuit engineering and intercellular communication

The ability to sense and respond to chemical stimuli is essential for the survival of all living organisms. Recognition of a chemical signal leads to a response that, hopefully, allows an organism to adapt and thus increase its probability of survival. Often, changes in gene expression occur following recognition of a specific chemical signal. Regulation of gene expression is crucial for interactions with the environment and for complex processes including development, apoptosis, circadian rhythms and biofilm formation. Pieces of the regulatory machinery required for modulating gene expression, including promoters, repressors and activators, have been borrowed from these complex regulatory networks in order to build simple networks commonly called ‘genetic circuits.’ Constructing genetic circuits de novo provides new opportunities to understand the governing principles and design strategies inherent in functional regulatory systems [1, 2]. By combining regulatory components from unrelated systems and generating simple regulatory circuits, researchers can gain an understanding of how more complex natural systems function in a controlled manner where perturbing a single component or condition leads to more easily interpreted results.

Initial work towards engineering synthetic genetic circuits has combined regulatory elements to produce simple circuits including toggle switches and a genetic oscillators [2-5]. Combining regulatory elements and generating a new, functional circuit, however, is not a trivial endeavor. These circuits are assembled from unrelated elements that have
not been optimized to work in the new configuration. An important challenge in engineering genetic circuits is to tune the components to function optimally in the new network. One way to address this problem is to apply what is known about the regulatory components and network design from previous biochemical, genetic and systems analyses to rationally adjust the parameters of the circuit components [6, 7]. A complementary approach is to use directed evolution, the iterative process of generating genetic diversity and screening or selecting for a desired function, to identify functional circuits [8].

As circuit engineers become more competent at generating functional systems, increasingly complex circuits will become more feasible design goals [9]. Engineering increasingly complex behaviors, such as those that mimic developmental processes like pattern formation and tissue development or assembly of microbial communities such as biofilms, become a more attainable goal with the addition of elements of intercellular communication systems [10]. Bacterial ‘quorum sensing’ components can be used to send and receive chemical signals from one cell to another. Quorum sensing (QS) refers to the ability of a bacterium to perceive and respond to fluctuations in cell-population density. In general, bacteria produce and release small chemical signal molecules that increase in concentration as a function of cell density [11, 12]. The detection of these signal molecules leads to changes in gene expression that presumably increase the fitness of the host organism.
QS systems found in bacteria from the Gram-negative phylum *Proteobacteria* use the synthesis and detection of acyl-homoserine lactones (acyl-HSLs) [11, 13, 14]. In these systems an acyl-HSL synthase produces the diffusible signaling molecule and an acyl-HSL-responsive transcription factor activates transcription following signal binding [12]. The levels of acyl-HSL synthesis in these organisms are sufficiently low to prevent accumulation of the signals at low cell densities. At high cell densities, acyl-HSL accumulates in the cell and is bound by the transcriptional activator leading to changes in gene regulation. The potential for these components to be used in synthetic intercellular communications circuits was demonstrated by Weiss and Knight, who expressed the acyl-HSL-responsive activator in *E. coli* to produce ‘receiver’ cells that would respond to signals produced by ‘sender’ cells synthesizing the acyl-HSL signal [15].

The components used by Weiss and coworkers are from the *lux* QS system found in *Vibrio fischeri*. The acyl-HSL synthase found in *V. fischeri*, LuxI, produces 3-oxo-hexanoyl-homoserine lactone (3OC6HSL), which is recognized by the 3OC6HSL-dependent transcriptional activator LuxR. We chose to target the *lux* system as the starting point for engineering novel intercellular signaling components. The *lux* system was first discovered in 1970 and has since become one of the best characterized QS systems [11, 13, 16]. By evolving the quorum-sensing signal response regulator, LuxR, we hoped to explore the evolution of signal specificity and response in acyl-HSL-based QS systems and to develop generally applicable tools for the identification of transcriptional response regulators that are optimized to function in a new context.
The components of the *lux* QS system have been successfully incorporated into synthetic genetic circuits requiring cell-cell communication [15, 17-20]. One of these examples used a LuxR variant identified in this work [17]. Using ‘receiver’ cells containing a LuxR-based band-detect circuit that express a fluorescent protein over a limited range of signaling molecule concentrations, Basu *et al.* [17] demonstrated that an undifferentiated lawn of the ‘receiver’ cells generated a ring of gene expression around a central colony of ‘sender’ cells producing 3OC6HSL. A second band-detect circuit was generated by replacing wild-type LuxR with a ‘hypersensitive’ LuxR variant that requires lower concentrations of 3OC6HSL to activate gene expression. Plating a mixture of two strains containing the two band-detect circuits, each controlling the expression of different fluorescent reporter, led to the production of a bull’s-eye pattern around a colony of sender cells (see appended publication [17]). The band of fluorescence observed with the evolved LuxR variant was further away from the sender cells, and therefore at lower concentrations of 3OC6HSL, than the band of fluorescence from the cells containing wild-type LuxR. This work demonstrated both the feasibility of using bacterial communication components to engineer systems that emulate complex processes such as pattern formation, and that LuxR variants with new behaviors are useful additions to the toolbox of components available for engineering genetic circuits.

In this chapter, I will briefly review the components involved in acyl-HSL-dependent QS and previous work towards engineering these systems and other transcription factors.
Particular emphasis will be placed on LuxR and its homologs, specifically their signal and DNA-binding specificities.

B. Acyl-homoserine lactone dependent transcriptional activation

The lux system from Vibrio fischeri was the acyl-HSL-based QS system discovered [16]. V. fisheri is a bioluminescent marine bacterium that colonizes the light organs of a variety of marine fish and squid, where it is found at very high cell densities (10^{10} cells/mL) and produces light. Nealson et al. were the first to report the role of a chemical signal molecule in the regulation of the light-producing enzyme luciferase [16]. This signal was later determined to be 3-oxo-hexanoyl-L-homoserine lactone (3OC6HSL) by \textsuperscript{1}H NMR spectroscopy [21].

B.1. Mechanism of QS activation in V. fischeri

In V. fischeri, the 3OC6HSL signal is synthesized by the acyl-HSL synthase, LuxI. In environments of low cell density, the amphipathic acyl-HSL diffuses out of the cell through the membrane [22]. As the local environment becomes more crowded, the acyl-HSLs accumulate in the cell, where they bind to the acyl-HSL-dependent transcriptional activator, LuxR [23, 24]. Upon signal binding, LuxR activates transcription at the lux promoter leading to light production through the transcription of the luciferase (lux)
genes. Transcription of the *lux* genes is essential for *V. fisheri*’s symbiotic relationship with marine eukaryotes [25]. In exchange for light production, *Euprymna scolopes*, the Hawaiian bobtail squid, provides *V. fischeri* with branch chain amino acids to support dense populations and luminescence.

**Figure 1.1.** Schematic of the *Vibrio fischeri* quorum sensing system. The left panel depicts the regulatory state at low cell density. Low, basal expression of LuxI produces a small amount of 3OC6HSL, which diffuses out of the cell. The right panel shows that at high cell density accumulation of 3OC6HSL leads to its binding by LuxR. LuxR-3OC6HSL complexes dimerize and activate gene expression at the lux promoter leading to increased expression of the *lux* genes required for light production, *luxCDABE*, as well as *luxI* and *luxR*. 
B.2. Homologous acyl-HSL dependent QS systems from the α-, β- and γ-Proteobacteria

Pairs of LuxI/LuxR homologs have been identified in many species of Gram-negative Proteobacteria (Fig. 1.2) [26, 27] where they regulate genes with diverse functions in response to changes in population density [28-33]. Different species of Proteobacteria synthesize and respond to acyl-HSLs with acyl chains of varying lengths and chemical modifications [34, 35] (Fig. 1.3). Due to the diversity of acyl-HSL signals used by different bacterial species and the high degree of specificity of both acyl-HSL production by LuxI homologs and acyl-HSL response by LuxR homologs, this type of QS has been postulated to be species specific [11].
Figure 1.2. Diversity of LuxR homologs and acyl-HSLs found in Nature. Illustration of the divergence of selected LuxR homologs. Homologs are listed with both species and cognate acyl-HSL signal molecule. The tree was adapted from Gray and Garey, and Lerat and Moran [26, 27].
Figure 1.3. Structures, non-IUPAC names and abbreviations of the acyl-HSLs that activate the homologs listed in Fig. 1.2., as well as the acyl-HSLs used in these studies, are shown.
C. Elements of acyl-homoserine lactone-dependent quorum sensing systems

The essential elements for acyl-HSL-dependent quorum sensing are an acyl-HSL synthase, an acyl-HSL-responsive transcriptional regulator and a target promoter. A LuxR homolog typically responds to an acyl-HSL synthesized by its partner LuxI homolog. The precise arrangement and regulation of these QS systems vary from organism to organism, and there appear to be exceptions to every rule. We will focus on the lux system to introduce these essential elements. Important work regarding homologous systems will also be presented.

C.1. The lux operon

The lux operon was first characterized following its cloning into E. coli [36]. As shown in Figure 1.4, the lux genes are arranged into two divergently transcribed promoters separated by 155 base pairs (bp). One promoter, P_{luxR}, contains luxR and the other, P_{luxI}, contains luxI and the six other lux genes, luxCDABEG, required for light production. Gene expression from P_{luxI} is activated by LuxR and requires the presence of 3OC6HSL. Expression of luxR was shown to depend on cyclic-AMP receptor protein (CRP) [37, 38], and a CRP-binding site was subsequently identified between luxI and luxR [39]. LuxR also appears to regulate its own expression, both positively and negatively, and depends on cellular LuxR concentrations and a downstream element in the DNA encoding luxD.
The precise mechanisms of luxR autoregulation have not been fully elucidated.

**Figure 1.4.** *Vibrio fischeri* lux gene organization. The luxR gene encoding the acyl-HSL-dependent transcriptional activator is transcribed from P_luxR. P_luxI contains luxI, which encodes the acyl-HSL synthase, followed by the five lux genes required for light production, luxCDABE, and an additional gene of unknown function luxG. The luxA and luxB genes encode the α- and β-components of luciferase. The luxC, luxD and luxE genes encode an acid reductase required for the production of the luciferase substrate, tetradecanal, from tetradecanoic acid.

LuxR activation of P_luxI requires a 20 bp inverted repeat centered 42.5 bp upstream of the P_luxI transcriptional start site called the lux box [43, 44]. Both halves of the lux box’s inverted repeat are required for LuxR-mediated gene activation, and the precise positioning of the lux box relative to the transcriptional start site is also imperative [44]. The lux box was first implicated as the LuxR binding site by the demonstration that point
mutations in the *lux* box could eliminate P_{luxI} activation in *E. coli* [43]. Similar nucleotide sequences have been identified upstream of many promoters regulated by LuxR homologs. TraR, the LuxR homolog from *Agrobacterium tumefaciens*, was the first homolog shown to directly bind to the *tra* box in DNaseI protection assays [45]. Recently, purified LuxR was shown to directly protect the *lux* box *in vitro* [46]. LuxR homologs TraR, LasR, QscR and CepR have also been shown to bind to lux-type boxes *in vitro* [47-50].

C.2. *Acyl-HSL synthesis by LuxI*

LuxI catalyzes acyl-HSL formation from S-adenosyl-L-methionine (SAM) and acyl-acyl-carrier protein (acyl-ACP) [51], both of which are readily available in the cell. SAM is acylated by acyl-ACP and its methionine moiety lactonized to produce acyl-HSLs (Fig. 1.5). LuxI homologs synthesize acyl-HSLs with acyl chains from 4 to 16 carbons in length [35, 52]. The third carbon of the acyl chain is often modified to a carbonyl or, less frequently, a hydroxyl group [52]. Acyl-HSLs containing a single degree of unsaturation have also been identified [53]. The formation of acyl-HSLs with specific acyl chains by LuxI proteins is a result of their ability to specifically recognize only ACPs charged with the “correct” acyl moieties. The crystal structure of EsaI, the LuxI homolog from *Pantoea stewartii* that produces 3OC6HSL, identified a hydrophobic cavity that recognizes the acyl chain of the acyl-ACP [54]. The size of this pocket was hypothesized to confer EsaI’s preference for acyl chains six carbons in length. In contrast, the crystal
structure of LasI, the LuxI homolog from *Pseudomonas aeruginosa* responsible for 3-oxo-dodecanoyl-homoserine lactone (3OC12HSL) production, contains a hydrophobic tunnel for acyl chain recognition that places no apparent restriction on acyl chain length [55]. Molecular modeling and mutagenesis studies indicate that the recognition of the ACP may play a role in recognizing acyl-ACPs with longer acyl chains. Gould *et al.* [55] hypothesize that, because longer acyl chains will have intrinsically higher binding affinities for the hydrophobic tunnel than shorter acyl chains, the required contribution from the ACP for binding by the enzyme is lower. Recognition of acyl-ACPs by LuxI homologs that make acyl-HSLs with shorter acyl chains likely requires the ACP to make a larger contribution to binding affinity. However, binding of acyl-ACPs carrying long acyl chains are sterically precluded from binding.

![Diagram of acyl-HSL synthesis](image_url)

**Figure 1.5.** Schematic of acyl-HSL synthesis by LuxI from acyl-acyl-carrier protein and S-adenosyl-L-methionine.
A LuxI homolog was recently engineered to produce an acyl-HSL with a longer acyl chain [56]. The sequences of closely related LuxI homologs from Erwinia strains producing either 3OC6HSL or 3OC8HSL were examined to identify amino acids that differed between the two groups. Two mutations were shown to be sufficient to change a 3OC6HSL synthase to a 3OC8HSL synthase. An alignment with the crystal structure of EsaI indicates that the two changes may increase the size of the binding pocket, however, the precise molecular mechanisms that control the specificity of acyl-HSL production remain elusive. Furthermore, attempts at switching the specificity from 3OC8HSL to 3OC6HSL by targeting the same residues produced nonfunctional enzymes.

C.3. Acyl-HSL-mediated gene activation by LuxR homologs

LuxR recognizes the acyl-HSL synthesized by LuxI, 3OC6HSL, and activates gene expression by increasing the affinity of RNA polymerase for the lux promoter [57]. Early mutagenesis studies showed that the N-terminal domain of LuxR is responsible for acyl-HSL binding while its C-terminal domain is responsible for DNA binding and activation [24, 58-60]. The mutations critical for LuxR function clustered in two regions, between residues 79 to 127 and residues 184 to 230 [24, 60]. Some of the LuxR variants containing mutations to the N-terminal domain were shown to be capable of activating gene expression with elevated levels of 3OC6HSL and were hypothesized to be involved in acyl-HSL binding. Subsequent studies verified the role of the C-terminal domain in
DNA binding and activation [58, 59]. A region between residues 116 and 160 was implicated in multimerization [61].

C.3.1. Acyl-HSL response

Following the identification of LuxR’s cognate signal molecule as N-3-oxo-hexanoyl-L-homoserine lactone [21], similar chemical structures were used to probe the specificity of the response. Analogs were tested with V. fischeri for their inducing activity and for their ability to inhibit LuxR’s ability to respond to 3OC6HSL [62]. Gene activation by LuxR was greatest in the presence of 3OC6HSL. However, acyl-HSLs with similar structures were also able to stimulate LuxR-mediated activation. The ability to study recombinant LuxR in E. coli led to a second set of studies that examined both binding and activation by LuxR with a set of acyl-HSL analogs [34]. Binding studies measured the amount of tritium-labeled 3OC6HSL ($^3$H-3OC6HSL) that remained associated with the E. coli cells, and presumably bound by LuxR, following incubation with the signal and washing to remove any unbound or loosely bound acyl-HSLs [63]. Acyl-HSL analogs were assessed for their ability to prevent the accumulation of $^3$H-3OC6HSL in LuxR-expressing cells. $^3$H-3OC6HSL accumulation was observed to decrease as the acyl-chain length became increasingly shorter or longer than six carbons [34]. A preference for acyl-HSLs containing a 3-oxo moiety was also observed. Compounds with substitutions in the homoserine lactone ring were not found to bind to LuxR, with the exception of
compounds with a homocysteine thiolactone ring. Compounds observed to bind to LuxR were tested for their ability to stimulate gene transcription by LuxR. None of the analogs tested was as active as 3OC6HSL. Compounds with acyl side chains between five and nine carbons in length have slight activity, while compounds with 10 or more carbon atoms showed little or no activity. Comparisons of gene activation by acyl-HSLs with or without the 3-oxo moiety showed that 3-oxo-containing acyl-HSLs are much more potent inducers. Notably, a few of the acyl-HSLs tested did not stimulate gene expression despite being able to decrease $^{3}$H-3OC6HSL accumulation. The analogs shown to bind to LuxR but do not stimulate gene activation were shown to inhibit LuxR-mediated gene expression in the presence of 3OC6HSL.

The acyl-HSL binding and response specificities of TraR, LasR, CarR and CepR have also been investigated [49, 64-66]. Similar observations were made in each study: acyl chain length, and the presence or absence of the 3-oxo group, are the major determinants of specificity.

The precise mechanism of 3OC6HSL binding and the conformational changes that occur prior to activation are only beginning to be elucidated. It has been consistently observed that LuxR and its homologs accumulate in vivo as functional proteins only in the presence of their target acyl-HSL [45, 46, 48, 49]. Overexpression of these proteins in the absence of acyl-HSL leads to the formation of insoluble inclusion bodies [45, 48, 49, 67]. Acyl-HSL binding by purified LuxR has been shown to be reversible, whereas LasR and TraR
appear to bind the acyl-HSL sufficiently tightly to be considered irreversible. Each of these proteins requires nanomolar concentrations of acyl-HSL to activate gene expression indicating that there may be a difference in the way these proteins bind their cognate signals.

The crystal structure of TraR has been determined and the residues that make contact with the acyl-HSL have been identified [68, 69]. Mutations of the amino acid residues predicted to form hydrogen bonds with 3OC8HSL were shown to be essential for wild-type protein function [70]. The amino acid residues responsible for recognition of specific acyl-HSLs have been more difficult to predict. Residues of TraR predicted to interact with the carbonyl moiety of 3OC8HSL were mutated with the goal of stimulating the response of this protein to an analog, octanoyl-HSL (C8HSL) [70]. Rather than showing improved sensitivity to C8HSL, these variants lost the strong response that wild-type TraR displays towards 3OC8HSL. The attempts of Chai and Winans to shift the specificity of TraR towards 3OC6HSL by introducing bulkier hydrophobic amino acids, and thereby reduce the size of the acyl binding site, also led, in all but one case, to variants with greatly decreased acyl-HSL affinities.
C.3.2. DNA binding and activation

The C-terminal domain of LuxR is responsible for DNA binding and gene activation. When the N-terminal, regulatory domain is absent, it is also capable of activating gene expression in the absence of 3OC6HSL [58]. This positive-control mutant, LuxRΔN, requires RNA polymerase (RNAP) for DNA binding. LuxRΔN and σ^{70}-bound RNAP were subsequently shown to be the only proteins required for transcription from P_{luxI} [71], indicating that LuxR activates gene expression by recruiting RNAP to P_{luxI}. Initiation of transcription by LuxRΔN has since been shown to require the C-terminal domain (CTD) of RNAP in vitro and activation by both LuxRΔN and full-length LuxR was drastically decreased by the overexpression of a truncated RNAP lacking the CTD [72]. Therefore, LuxR activates gene expression by recruiting RNAP to P_{luxI} through direct binding to its CTD.

Amino acid sequence analysis identified a helix-turn-helix (HTH) motif within the C-terminal domain of LuxR [73, 74]. Two alanine scanning mutagenesis experiments have identified residues involved in DNA binding and activation by LuxR [75, 76]. Several mutants unable to bind DNA or activate transcription were identified, including alanine substitutions at positions 191, 193, 212, 217, 225, 229, 230, 238, and 240. DNA binding was assessed indirectly by measuring the ability of each variant to repress a lac promoter containing a lux box [77]. It is impossible to determine from these results which of the mutations yields folded proteins that are unable to bind DNA and which mutations
prevent activation by affecting other properties such as protein stability, folding or
dimerization. The DNA-binding HTH motif in LuxR is predicted to occur between
residues 200 and 224, suggesting that several of the residues identified may be directly
involved in DNA binding. Alanine substitutions at positions 201 and 206 affect
transcriptional activation but not DNA binding. These residues are hypothesized to make
direct contact with RNA polymerase.

The DNA sequence requirements for specific DNA-binding by LuxR have not been
assessed as thoroughly as acyl-HSL specificity. The observation that LuxR and LasR are
capable of activating gene expression at each other’s cognate promoter suggests that the
DNA-binding specificities of these proteins have not significantly diverged [78]. This is
not entirely surprising because the ability to respond to species-specific signals occurs at
the level of acyl-HSL binding. The ability to bind a different DNA target than another
LuxR homolog becomes important when more than one LuxR homolog is present in the
same cell. _P. aeruginosa_, for example, has three LuxR homologs, RhlR, LasR and QscR.
Each of these homologs has been shown to bind to different target sites, although all of
the sites are considered _lux_-type boxes [47, 48, 79]. Substitutions at positions 8 and -8
of the _lux_-type boxes recognized by LasR and QscR are sufficient to switch their
responses between the two sites [47]. LasR is the only LuxR homolog whose DNA-
binding specificity has been investigated with several potential target sites [48]. LasR
was shown to bind several promoters that had previously been identified _las_-responsive.
LasR did not bind to all of the predicted sites and was shown to bind at unexpected sites.
Interestingly, the sites that were bound by LasR did not show a large degree of overall sequence similarity and did not require dyad symmetry.

Prior to this work, LuxR had not been engineered to respond to new acyl-HSLs or to activate gene expression at a new promoter. A lac promoter had been modified to include a lux box, converting LuxR from an activator to a repressor. However, LuxR itself was not modified. As discussed above, an attempt to engineer TraR to recognize a new autoinducer based on structural information yielded proteins that showed broadened specificity but were very poor activators.

D. Engineering bacterial transcription factors

Some of the most successful examples of transcription factor engineering have come from engineering eukaryotic zinc-finger (Zn-finger) proteins [80, 81]. The modularity of Zn-fingers has made these DNA-binding proteins good targets for engineering experiments. Each finger is responsible for binding three consecutive bp of the DNA target and they can be “strung” together to bind longer DNA sequences. The functionality of Zn-fingers is also modular in that an activating domain or repressor domain need only be attached to the protein to confer its function at the target DNA site. Small molecule binding to the effector domain then affects the rate of transcription from the target promoter. In contrast, many bacterial transcription factors, including HTH-
containing proteins like LuxR, are switched ON and OFF by conformational changes that affect the whole protein. This means that engineered proteins must maintain the mechanism of signal transduction from regulatory region to DNA-binding domain and may render these proteins more challenging targets with regards to engineering new functions. Whether targeting a Zn-finger or a bacterial HTH protein, the available strategies are similar and are introduced in the following section. Examples of bacterial transcription factor engineering are then discussed.

D.1. Directed evolution and rational design

Directed evolution is an iterative process of mutation or recombination and selection or screening. It mimics the process of natural evolution by creating genetic diversity and then identifying genes encoding the “fittest” proteins based on the goals of the project. Rational design, on the other hand, uses structure and sequence information, which can be analyzed via computational methods, to identify mutations that are likely to confer the desired change in function.

In general, sequence diversity is generated for directed evolution experiments by random mutagenesis or by recombining multiple parental sequences. Libraries of mutant genes produced by random mutagenesis are especially useful when there is little or no structural information available for a target protein. We chose to use random mutagenesis to target
the acyl-HSL specificity of LuxR because, at the time, there were no structures of LuxR homologs available and, more importantly, because the sequence identity between LuxR and its homologs is very low (<25%), identifying amino acids responsible for modulating acyl-HSL specificity was not feasible from sequence information alone.

A semi-rational approach is commonly used for engineering transcription factors [82-85]. In general, structural or functional information is used to identify residues predicted to modulate the target activity, and libraries that randomize these residues are screened or selected for the desired behavior.

**D.2. Screening and selection**

Identification of proteins that exhibit the desired behavior from a library of variants requires the ability to relate the desired protein function to a measurable output. In a screen, a functional protein yields an easily measured product. Directed enzyme evolution experiments have often relied on the use of surrogate substrates, i.e., substrates that are chemically similar to the substrate of interest and whose products can be easily quantified colorometrically, or the use of additional enzymes that selectively convert the desired product to a readily detectable compound. One of the advantages of working with transcription factors is that protein function can easily be tied to the production of a reporter protein that generates a measurable output. Commonly used reporter proteins
include β-galactosidase, luciferase and fluorescent proteins such as green fluorescent protein (GFP) and red fluorescent protein (RFP).

Selections differ from screens in that they tie the survival of the host cell to the desired protein function. It is often the production of an essential metabolite or an increase in antibiotic resistance that leads to survival. The upregulation of a gene conferring antibiotic resistance is commonly used to identify transcription factors that are on or off under a specific set of conditions. Most antibiotic resistance proteins chemically inactivate the antibiotic through enzymatic cleavage or derivatization. For example, β-lactamase (Bla) catalyzes the opening and hydrolysis of the lactam ring of β-lactam antibiotics [86] and chloramphenicol acetyl transferase (CAT) inactivates chloramphenicol by acetylating the antibiotic [87]. TetA, on the other hand, is a pump and confers tetracycline resistance by exporting the antibiotic from the cell [88].

The major advantage to using a selection instead of a screen is potential library size. Selections can easily cover libraries containing >10^6 variants, while screens are often laborious and can limit library size to ~10^4. Screens, however, remain the more popular method because selections are often plagued by false positives. While false positives occur regardless of the choice between screening and selection, they occur more frequently when cells are placed under selective pressure. Any solution that leads to survival will be selected, even if it is not the solution that has been targeted by the experimenter. As will be discussed in Chapter 3, a successful selection requires an
approach that ensures the surviving bacteria express mutant proteins with the targeted properties. Another advantage to using a screen is that the output is often measured, quantitatively or qualitatively, such that mutants can be easily ranked. The graded output, as opposed to simply being alive or dead, makes it easier to identify the best variant from a pool of variants that perform the desired function.

The use of negative screening or selection also figures prominently into this work. The screens and selections described above are generally used for identifying gain-of-function mutations. As will be discussed in Chapters 3 through 5, identifying variants with loss-of-function mutations is essential for evolving specificity [89]. The absence of reporter production can be used in screens but one runs the risk of identifying nonfunctional proteins. By screening for proteins that confer a desired function prior to screening for the absence of function under a different set of conditions, one can limit the prevalence of the nonfunctional proteins. Selecting for loss-of-function mutations is more complicated because it is impossible to retrieve DNA encoding a protein of interest from a dead bacterium. The most common way to implement negative selections is through the use of a repressor protein as an inverter. In this case, gene activation leads to the production of a repressor, which prevents the production of an antibiotic resistance gene, leading to antibiotic sensitivity. Toxic protein production can be employed in negative selections. In general, the basal level of expression of most toxic proteins leads to a prevalence of false positives and the use of a repressor to invert the signal provides an additional target at which the selection system can be undermined. Yokobayashi and Arnold have
developed a dual selection system that employs an inhibitor of Bla to directly target the protein that confers the ability to survive with β-lactam antibiotics [90].

D.3. Engineering signal binding

The tetracycline repressor (TetR) is a well-studied bacterial transcription factor that has been engineered to both recognize new small molecules and to bind to new DNA sequences [83, 85, 91, 92]. TetR is responsible for repressing the expression of TetA [93]. Because overexpression of TetA is toxic to cells, TetR has evolved a very high affinity for a DNA site located upstream of the tetA gene to prevent its expression. TetR also has a high affinity for tetracycline (tc, Fig. 1.5) and releases its DNA target to allow tetA expression before tc concentrations become toxic.

Variants of TetR that bind to an analog of tc that does not affect wild-type TetR, 4-dedimethylamino-6-demethyl-6-deoxy-tetracycline (cmt3, Fig.1.6), were identified by screening libraries generated by random mutagenesis for cmt3-dependent expression of β-galactosidase [85]. Recombination by DNA shuffling and another round of screening identified a single variant with improved cmt3-inducibility. Almost half of the mutations found in the best clones from the first round were located close to or in the tc-binding region. A new library targeting amino acid positions shown to be close to the sites that differ between tc and cmt3 was generated, using the wild type as parent, and screened for
cmt3-inducibility. A single mutation was identified that was partially inducible with cmt3 but was still showed a preference for tc. Recombination with a mutation identified by random mutagenesis identified a relaxed specificity variant that is induced by cmt3 at levels comparable to wild-type TetR with tc, and that shows decreased induction with tc. A final round screened a library, generated by randomizing the same amino acids as the previous round but with the best double-mutant as parent, for a decrease in response to a potent tc analog. Variants exhibiting the desired improvements in specificity were identified.

![Chemical structures of tetracycline derivatives](image)

**Figure 1.6.** Chemical structures of tetracycline derivatives used in the directed evolution of TetR.

Engineering TetR to respond to cmt3 successfully incorporated both random and semi-rational library design strategies. Mutations critical for modulating TetR’s response to tc and its cmt3 analog were identified. The relaxed specificity variant identified while screening for cmt3 was subsequently used as a parent to identify a TetR mutant inducible
by a different tc analog [92]. Two residues were targeted for randomization due to their proximity to the moiety that differs between tc and the new analog in the TetR. A single mutation was identified that conferred a 200-fold switch in specificity.

A eukaryotic activation domain has been fused to TetR to generate a hybrid tc-dependent transcriptional activator (tTA) that functions in mammalian cells [94]. This system is widely used as an inducible expression system in eukaryotes ranging from yeast to mice [95]. Furthermore, a set of eukaryotic, TetR-based transactivators that respond to cmt3 but not tc were developed from the cmt3-specific TetR variants to allow selective regulation of two genes by the different inducer molecules [96].

The majority of studies for the modification of signal-dependent transcription factors have altered the binding specificity of eukaryotic activation domains [97]. Further examples of engineering signal binding by bacterial transcription factors are unavailable.

D.4. Engineering DNA-binding interactions

The HTH DNA-binding motif, which is found in LuxR, is commonly found in prokaryotic transcriptional regulators. Examples of HTH engineering in the literature are scarce and the examples that are available have relied on structural information regarding the interactions between the protein and its target site [82, 84, 98-106]. Many published HTH engineering experiments have used solutions already present in Nature, including
swapping an entire DNA recognition helix [104] or replacing all of the amino acids that contact a specific DNA target recognized by a close homolog [105]. This strategy is not always successful [98] and does not generate DNA-binding proteins that recognize novel DNA sequences. Attempts at engineering HTH proteins that bind to previously unrecognized sequences have generally relied on libraries that randomize the amino acids shown to make specific contacts with DNA [82-84, 91]. In most cases, modest gains can be accomplished in this manner, however, new specificities are rare. Occasionally, mutations at residues that are not targeted have arisen adventitiously and are critical for binding the new target [91].

One of the successful strategies for engineering HTH proteins is to engineer single-chain proteins. Single-chain proteins that are covalent dimers of the DNA-binding domains from different transcription factors have also been shown to bind to hybrid operators [107-110]. Single-chain proteins that bind to new asymmetric DNA targets have been identified by screening [111] and phage display [102]. More recently, Liang et al. [112] constructed several single-chain heterodimers using engineered 434 repressor DNA-binding domains (identified in previous rational design and selection experiments), which also bound operators made up of half-sites bound by the individual domains. To achieve high DNA-binding affinities ($K_d = 10^{10-11}$ M), the single-chain proteins developed by Liang and colleagues require that the six base pair (bp) sequence between the two half-sites contains mostly A and T nucleotides.
The most successful efforts towards engineering the DNA-binding specificities of prokaryotic proteins have switched the DNA-binding specificity of Cre and Flp recombinases. Work from two independent laboratories used screens incorporating positive and negative selection pressure to identify variants that recombine only at a new target site [113-115]. In one of these experiments, directed evolution of the DNA-binding specificity of a Cre recombinase using positive screening alone and a combination of positive and negative screening showed that both positive and negative selective pressure were required to switch Cre’s DNA-binding specificity [113].
E. References


68. Vannini, A., Volpari, C., Gargioli, C., Muraglia, E., Cortese, R., De Francesco, R., Neddermann, P. & Di Marco, S. The crystal structure of the quorum sensing


CHAPTER 2

Directed Evolution of *Vibrio fischeri* LuxR for Increased Sensitivity to a Broad Spectrum of Acyl-homoserine Lactones

A. Abstract

LuxR-type transcriptional regulators play key roles in quorum-sensing systems that employ acyl-homoserine lactones (acyl-HSLs) as signal molecules. These proteins mediate quorum control by changing their interactions with RNA polymerase and DNA in response to binding their cognate acyl-HSL. The evolutionarily-related LuxR-type proteins exhibit considerable diversity in primary sequence and in their response to acyl-HSLs having acyl groups of differing length and composition. Little is known about which residues determine acyl-HSL specificity, and less about the evolutionary time scales required to forge new ones. To begin to examine such issues, we have focused on the LuxR protein from *Vibrio fischeri*, which activates gene transcription in response to binding its cognate quorum signal, 3-oxohexanoyl-homoserine lactone (3OC6HSL). Libraries of *luxR* mutants were screened for variants exhibiting increased gene activation in response to octanoyl-HSL (C8HSL), with which wild-type LuxR interacts only weakly. Eight LuxR variants were identified that showed a 100-fold increase in sensitivity to C8HSL; these variants also displayed increased sensitivities to pentanoyl-HSL and tetradecanoyl-HSL, while maintaining a wild-type or greater response to 3OC6HSL. The most sensitive variants activated gene transcription as strongly with C8HSL as the wild type did with 3OC6HSL. With one exception, the amino acid residues involved were restricted to the N-terminal, “signal-binding” domain of LuxR. These residue positions differed from critical positions previously identified via “loss-of-function” mutagenesis. We have demonstrated that acyl-HSL-dependent quorum-sensing
systems can evolve rapidly to respond to new acyl-HSLs, suggesting that there may be an evolutionary advantage to maintaining such plasticity.

**B. Introduction**

Several mechanisms have evolved to allow diverse bacterial species to detect changes in their local population density, and to modulate their gene expression accordingly [1, 2]. An opinion that quorum sensing is actually a form of diffusion sensing has recently been articulated [3], however, it is the view of many that the two stances are not necessarily mutually exclusive. The population-density-based control of gene expression in bacteria has been termed “quorum sensing” [4]. A paradigm of the quorum-sensing control of gene regulation within the Gram-negative phylum *Proteobacteria* has been the luminescence (*lux*) operon in *Vibrio fischeri*. In addition to the luciferase genes required for light production, this operon encodes LuxR, an acyl-homoserine lactone (acyl-HSL)-dependent transcriptional activator, and LuxI, an acyl-HSL synthase that catalyzes the production of 3-oxohexanoyl-homoserine lactone (3OC6HSL) [5]. Each bacterium expresses the Lux proteins at low basal levels throughout its entire lifecycle. At low cell densities, the small amounts of the amphipathic 3OC6HSL signal that are produced diffuse away from the cells [6, 7]. However, as a local population increases in density, 3OC6HSL concentrations increase. This results in a shift of the LuxR equilibrium towards its 3OC6HSL-bound, active state. Acyl-HSL binding leads to dimerization of LuxR and binding to the *lux* box, a 20-base pair inverted repeat located in the P<sub>lux</sub>
promoter [8]. There the acyl-HSL-bound LuxR dimer activates expression of the *lux* genes after the recruitment of RNA polymerase [9, 10].

To date, genes encoding more than 50 LuxI/LuxR pairs have been identified in diverse species belonging to the α-, β-, and γ-subdivisions of the *Proteobacteria* [11]. While the basic quorum-sensing mechanism has been fairly well conserved in most species studied, the physiological functions controlled by acyl-HSL-based quorum-sensing systems are varied and include virulence [12, 13], antibiotic synthesis [14, 15], and biofilm formation [16, 17]. A LuxR homolog typically responds to an acyl-HSL synthesized by its partner LuxI homolog. Known quorum-sensing systems are tuned to preferentially synthesize and respond to a specific acyl-HSL having an acyl side group that may range in length from four to 16 carbons, may contain a degree of unsaturation, and may be modified with a carbonyl or hydroxyl functional group at carbon position 3 [18, 19].

Are LuxR homologs evolutionarily pliable, i.e., able to move into new signaling niches on short time scales? We are interested in this issue and in identifying how the specificity of acyl-HSL response is determined. The ability of LuxR to regulate gene expression in response to an acyl-HSL depends on signal recognition and binding; it may also include acyl-HSL-specific dimerization and DNA-binding mechanisms. Because LuxR homologs generally share low amino acid sequence identities (≤25%), however, it is difficult to determine precisely how they discriminate between various acyl-HSL molecules, or how this discrimination evolves and changes as a result of varying selective
pressures. In a recent study on acyl-HSL-binding specificity, residues of TraR (a LuxR homolog) predicted to interact with the carbonyl moiety of its cognate signal, 3-oxooctanoyl-HSL (3OC8HSL) were mutated with the goal of stimulating the response of this protein to an analog, octanoyl-HSL (C8HSL) [20]. Rather than showing improved sensitivity to C8HSL, these variants lost the strong response that wild-type TraR displays towards 3OC8HSL. The attempts of Chai and Winans to shift the specificity of TraR towards 3OC6HSL by introducing bulkier hydrophobic amino acids, and thereby reduce the size of the acyl binding site, also led, in all but one case, to variants with greatly decreased acyl-HSL affinities. Thus, alternative approaches may be more successful at altering the acyl-HSL specificities of LuxR or its homologs. To identify LuxR variants with altered acyl-HSL specificity, and to identify key residues which modulate this property, we have developed and implemented a system to direct the evolution of LuxR-mediated gene activation. Variants of LuxR were generated and screened to identify those that are highly responsive to acyl-HSL signals to which this transcriptional activator does not normally respond.
C. Results

C.1. Cloning vectors and screening tools

Modifications to a commercial two-plasmid system were made to allow for rapid cloning of luxR mutant libraries and screening for gene activation in response to exogenous signal molecules. The first plasmid, pLuxR, contains wild-type luxR under the control of a P_{lac/ara-1} hybrid promoter. The luxR allele is constitutively expressed when placed in E. coli strains lacking lacI and araC. The second plasmid, pLuxGFPuv, contains the gene gfpuv placed under the control of the P_{lux} promoter. The gfpuv gene encodes a green fluorescent protein variant, GFPuv, which becomes visible when excited with standard, long-wavelength ultraviolet light [21]. GFPuv expression from the P_{lux} promoter of pLuxGFPuv is dependent upon the degree of activation that occurs as a result of a functional LuxR variant binding to a specific acyl-HSL signal molecule. The relative levels of gene activation observed with wild-type LuxR and the acyl-HSLs used in this study were similar to those previously reported in V. fischeri [22] and E. coli [23].

C.2. First-generation laboratory evolution—error-prone PCR

We used the two-plasmid system to identify LuxR variants that showed an increased ability to activate gene transcription in response to C8HSL. Past studies have shown that low-micromolar concentrations of C8HSL were required to achieve the same level of gene expression achieved by 5 nM of 3OC6HSL, the cognate signal of LuxR [22]. In the
current system, a C8HSL concentration of approximately 2 μM was required to achieve the level of gene activation obtained with 10 nM 3OC6HSL. Random mutations were introduced into wild-type luxR by error-prone PCR. E. coli hosts containing pLuxGFPuv were transformed with pLuxR vectors containing the mutagenized PCR products. Transformants were initially plated onto LB-agar amended with 50 nM C8HSL and antibiotics as appropriate. However, initial screening with this concentration of acyl-HSL yielded only one mutant, LuxR-G1A. Subsequently, the library was screened for activation with 200 nM C8HSL. Colonies were screened under illumination with 365 nm ultraviolet light. Most colonies appeared dark under such illumination, indicating the expression of LuxR variants having either wild-type or non-functional properties. Of the ~20,000 colonies that were screened, nine fluorescent colonies were identified. These nine colonies were isolated and verified by re-screening, which identified one as a false positive. The remaining eight alleles were amplified and cloned into fresh background plasmids and strains to ensure that the observed phenotype was due to changes within luxR. An additional false positive was identified by doing so.

Of the seven remaining alleles, two were shown to be identical by DNA sequence analysis. The six LuxR variants from this first generation of laboratory evolution were designated LuxR-G1A through LuxR-G1F (Fig. 2.1). Quantitative characterization of the response of these mutants to 3OC6HSL and C8HSL was performed in liquid-phase assays by monitoring fluorescence levels using a microtiterplate spectrofluorimeter. As shown in Table 2.1, all six variants showed increased gene expression in the presence of
100 nM C8HSL as compared to wild-type LuxR. Amino acid substitutions identified in the first generation are shown in Figure 2.1.

**Figure 2.1.** Amino acid sequence changes to first- and second-generation LuxR variants showing increased gene activation in response to C8HSL. The six first-generation variants were identified from a library of approximately 20,000 clones generated by error-prone PCR. Second-generation variants were generated by DNA shuffling of alleles from the first generation. Eight second-generation variants showing additional increases in gene activation in response to C8HSL were identified after screening 1,200 clones. The positions of amino acid substitutions within the 250 amino acid LuxR protein are indicated as black rectangles (refer to Table 2.1 for nucleotide sequence changes).
Table 2.1. Nucleotide and amino acid changes in the recovered luxR/LuxR mutants

<table>
<thead>
<tr>
<th>luxR plasmid</th>
<th>luxR gene base and substitution</th>
<th>LuxR protein amino acid residue and substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Generation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLuxR-G1A</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>A159→G</td>
<td></td>
</tr>
<tr>
<td>pLuxR-G1B</td>
<td>A133→T</td>
<td>Ile45→Phe</td>
</tr>
<tr>
<td></td>
<td>T501→C</td>
<td>Synonymous</td>
</tr>
<tr>
<td>pLuxR-G1C</td>
<td>A284→G</td>
<td>Asn95→Ser</td>
</tr>
<tr>
<td></td>
<td>A414→C</td>
<td>Synonymous</td>
</tr>
<tr>
<td>pLuxR-G1D</td>
<td>A286→G</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td></td>
</tr>
<tr>
<td>pLuxR-G1E</td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>A683→G</td>
<td></td>
</tr>
<tr>
<td>pLuxR-G1F</td>
<td>T162→A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A403→G</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>A522→G</td>
<td></td>
</tr>
<tr>
<td><strong>Second Generation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLuxR-G2A</td>
<td>A133→T</td>
<td>Ile45→Phe</td>
</tr>
<tr>
<td></td>
<td>A403→G</td>
<td>Met135→Val</td>
</tr>
<tr>
<td>pLuxR-G2B</td>
<td>A133→T</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Ile45→Phe</td>
</tr>
<tr>
<td></td>
<td>A522→G</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td>pLuxR-G2C</td>
<td>T346→G</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td></td>
</tr>
<tr>
<td>pLuxR-G2D</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>A286→G</td>
<td>Ile96→Val</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td></td>
<td>A414→C</td>
<td>Synonymous</td>
</tr>
<tr>
<td>pLuxR-G2E</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td>pLuxR-G2F</td>
<td>A97→G</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A133→T</td>
<td>Ile45→Phe</td>
</tr>
<tr>
<td></td>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td>pLuxR-G2G</td>
<td>A97→G</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A286→G</td>
<td>Ile96→Val</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td>pLuxR-G2H</td>
<td>A97→G</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A284→G</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A522→G</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.2. Activation of *gfpuv* gene expression by LuxR variants with four different acyl-HSLs

<table>
<thead>
<tr>
<th>LuxR type</th>
<th>No acyl-HSL</th>
<th>3OC6HSL</th>
<th>C8HSL</th>
<th>C5HSL</th>
<th>C14HSL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10 nM</td>
<td>100 nM</td>
<td>10 nM</td>
<td>100 nM</td>
<td>500 nM</td>
</tr>
<tr>
<td>LuxR wt</td>
<td>&lt;50</td>
<td>1400</td>
<td>&lt;50</td>
<td>100</td>
<td>&lt;50</td>
</tr>
<tr>
<td><strong>First Generation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LuxR-G1A</td>
<td>&lt;50</td>
<td>2200</td>
<td>&lt;50</td>
<td>800</td>
<td>100</td>
</tr>
<tr>
<td>LuxR-G1B</td>
<td>&lt;50</td>
<td>2600</td>
<td>&lt;50</td>
<td>800</td>
<td>500</td>
</tr>
<tr>
<td>LuxR-G1C</td>
<td>&lt;50</td>
<td>2100</td>
<td>&lt;50</td>
<td>700</td>
<td>&lt;50</td>
</tr>
<tr>
<td>LuxR-G1D</td>
<td>&lt;50</td>
<td>400</td>
<td>100</td>
<td>1500</td>
<td>700</td>
</tr>
<tr>
<td>LuxR-G1E</td>
<td>&lt;50</td>
<td>2200</td>
<td>&lt;50</td>
<td>500</td>
<td>&lt;50</td>
</tr>
<tr>
<td>LuxR-G1F</td>
<td>&lt;50</td>
<td>600</td>
<td>&lt;50</td>
<td>700</td>
<td>600</td>
</tr>
<tr>
<td><strong>Second Generation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LuxR-G2A</td>
<td>&lt;50</td>
<td>1600</td>
<td>200</td>
<td>1700</td>
<td>2200</td>
</tr>
<tr>
<td>LuxR-G2B</td>
<td>&lt;50</td>
<td>2400</td>
<td>900</td>
<td>2900</td>
<td>3000</td>
</tr>
<tr>
<td>LuxR-G2C</td>
<td>&lt;50</td>
<td>500</td>
<td>200</td>
<td>1900</td>
<td>1000</td>
</tr>
<tr>
<td>LuxR-G2D</td>
<td>&lt;50</td>
<td>1300</td>
<td>800</td>
<td>3100</td>
<td>2300</td>
</tr>
<tr>
<td>LuxR-G2E</td>
<td>&lt;50</td>
<td>1400</td>
<td>1100</td>
<td>3600</td>
<td>3200</td>
</tr>
<tr>
<td>LuxR-G2F</td>
<td>&lt;50</td>
<td>3100</td>
<td>800</td>
<td>2900</td>
<td>2600</td>
</tr>
<tr>
<td>LuxR-G2G</td>
<td>&lt;50</td>
<td>3300</td>
<td>500</td>
<td>2800</td>
<td>1400</td>
</tr>
<tr>
<td>LuxR-G2H</td>
<td>&lt;50</td>
<td>3300</td>
<td>900</td>
<td>2500</td>
<td>1200</td>
</tr>
</tbody>
</table>

*a* Units of fluorescence represent the fluorescence due to GFPuv production obtained with a given LuxR variant and p*luxGFPuv*. Fluorescence values were corrected with the background fluorescence obtained without a LuxR expressing plasmid. Both values were normalized to their cell densities prior to correction. All values are the mean of triplicate measurements. Variation was less than ± 15% for all reported values.
C.3. Second-generation laboratory evolution—DNA shuffling of mutant luxR alleles

To investigate whether the mutations could be combined to yield further improvements in C8HSL response, a new library of luxR alleles was generated via DNA shuffling of the six from the first generation, under conditions designed to minimize new point mutations [24]. As with the first generation, transformants were screened for GFPuv production on LB-agar, in this case amended with 50 nM C8HSL. Of 1200 colonies transformed with this library, 102 exhibited fluorescence. To identify those variants most sensitive to C8HSL, these 102 strains were screened using 10 and 20 nM C8HSL. Ten exhibited fluorescence under these conditions and were picked for further characterization. One false positive was identified after re-amplification and re-cloning of the alleles into fresh background materials. Sequence analysis revealed two of the alleles to be identical. Quantitative liquid-phase characterization of the eight remaining mutants showed that all exhibited remarkable increases in fluorescence output in response to not only 100 nM but also 10 nM C8HSL, as compared to either wild-type LuxR or any of the first-generation LuxR variants (Table 2.2). Sequencing revealed that no additional residue positions had been changed; the second-generation sequences were all recombinants of the parent sequences (Fig. 2.1).
C.4. C8HSL and 3OC6HSL dose responses

We compared the influence of a range of 3OC6HSL and C8HSL concentrations on transcriptional activation by wild-type LuxR and each of the second-generation variants (Fig. 2.2). Wild-type LuxR required approximately 2 μM C8HSL to elicit half the maximum GFPuv fluorescence observed using saturating amounts of 3OC6HSL, whereas only 10 nM 3OC6HSL was required for half-maximal activation (Fig. 2.2A). In comparison, all second-generation LuxR variants showed an increased sensitivity to C8HSL, requiring only 15 to 50 nM to achieve the half-maximal 3OC6HSL-induced fluorescence of the wild-type. Thus, C8HSL sensitivity and response by these LuxR variants increased by 20- to 100-fold. Variants LuxR-G2D and LuxR-G2E responded as sensitively and strongly to C8HSL as wild-type LuxR did to 3OC6HSL (Fig. 2.2E, F). The acquisition of this trait had little or no deleterious impact on the sensitivity of these variants to 3OC6HSL. Three variants, LuxR-G2A, LuxR-G2D and LuxR-G2E (Fig. 2.2B, E, F), showed responses to 3OC6HSL indistinguishable from that of wild-type LuxR. Variants LuxR-G2F, LuxR-G2G and LuxR-G2H (Fig. 2.2G–I) showed 5-fold increased sensitivity to 3OC6HSL, exhibiting significant levels of fluorescence with concentrations as low as 1 nM.
Figure 2.2. Detection of C8HSL and 3OC6HSL by wild-type LuxR and eight second-generation LuxR variants. For each panel and variant (A through I), data for 3OC6HSL are represented by closed squares (■), data for C8HSL are represented by open triangles (Δ). Fluorescence data reported are normalized to optical density, and corrected by subtracting background fluorescence from a control strain carrying pPROLar.A122 with pLuxGFPUv. All measurements were performed in triplicate. Error bars indicate either the range of values or the standard deviation, whichever was greater.
**C.5. Increased sensitivity to 3OC6HSL can be attributed to increases in LuxR protein concentration**

To investigate whether increases in expression were responsible for observed increases in gene activation in response to 3OC6HSL, concentrations of the variant LuxR proteins were measured. Fusions of the *c-myc* epitope to the C-termini of wild-type LuxR and each of the second-generation variants were constructed. All *c-myc* LuxR variants functioned well, i.e., exhibited responses to acyl-HSLs which were similar to the variant lacking the *c-myc* epitope. Protein concentrations were monitored under different incubation conditions by Western immunoblot assay and compared.

We tested the accumulation of wild-type LuxR and the second-generation variant proteins after incubation in the absence or presence of saturating amounts of 3OC6HSL. In the absence of acyl-HSL, the amount of protein detected by Western immunoblot for each of the variants differed by less than two-fold from wild-type LuxR; however, in the presence of 1 μM 3OC6HSL, the four variants with increased responses to 3OC6HSL, LuxR-G2B, -G2F, -G2G and -G2H, also showed increased amounts of protein as compared to wild type (Fig. 2.3). Thus mutations leading to higher LuxR levels may lead to the increased sensitivities of these variants to 3OC6HSL and, in part, C8HSL.
Figure 2.3. Protein concentrations of second-generation LuxR variants. The immunoblot was performed on a SDS-PAGE display of the total proteins harvested from strains expressing c-myc tagged second generation LuxR variants incubated in the presence of 1μM 3OC6HSL. Myc-coupled fluorescence for each of the variants varied by less than 10% in the presence and absence of acyl-HSL (latter not presented). In the absence of acyl-HSL, the amount of LuxR protein detected for each variant was within two-fold of that for the wild-type LuxR-myc fusion.

The amount of wild-type LuxR detected in the presence of 3OC6HSL was approximately two-fold more than was detected in the absence of signal. The increase in protein levels observed with the addition of 3OC6HSL is similar to that observed for TraR [20], consistent with the notion that acyl-HSL binding enhances the folding and/or stability of LuxR. If so, acyl-HSL binding and protein stability are inextricably linked.

C.6. Response of LuxR variants to acyl-HSLs with long and short acyl chains

The ability of the C8HSL-responsive LuxR variants to activate gene expression with other acyl-HSLs was also investigated. The response of each variant was examined using 100 nM pentanoyl-homoserine lactone (C5HSL) and 1 μM tetradecanoyl-homoserine lactone (C14HSL), two acyl-HSLs to which wild-type LuxR exhibits essentially no response (Table 2.2). Four first-generation variants (LuxR-G1A, -G1B, -G1D and -G1F)
showed small, yet significant, responses to C5HSL, as did three (LuxR-G1D, -G1E and -G1F) to C14HSL.

The second-generation variants all showed marked improvements in their sensitivity and response to both C5HSL and C14HSL. LuxR-G2A showed a three-fold and LuxR-G2B a >four-fold improvement in sensing C5HSL when compared to the best of the first-generation mutants. LuxR-G2G and LuxR-G2H showed a >seven-fold increase in gene activation in the presence of C14HSL when compared to the best of the first-generation mutants. LuxR-G2D, -G2E, and -G2F all showed broadened acyl-HSL specificity, exhibiting strong responses to both C5HSL and C14HSL. The most responsive mutant, LuxR-G2E, showed an ~five-fold increase in fluorescence with C5HSL and an 11-fold increase with C14HSL, while maintaining a wild-type level of response to 3OC6HSL.

C.7. Acyl-HSL specificity of the second-generation variants

The ratio of the 3OC6HSL and C8HSL concentrations required to attain half-maximal gene expression, termed the specificity, $S$, was used to assess changes in acyl-HSL responses (Table 2.3). While wild-type LuxR has an $S$ of 200, indicating that it requires 200-fold more C8HSL than 3OC6HSL to reach half-maximal fluorescence levels, all of the second-generation mutants have $S < 20$. The variant with the largest increase in response to C8HSL, LuxR-G2E, has $S = 1$: it responds equally to both HSLs. This variant is also sensitive to C5HSL and C14HSL (see above). The variants which showed
increased protein concentration (LuxR-G2F, -G2G and -G2H) also exhibited a more than 10-fold decrease in $S$ as compared to wild type. Whereas increases in protein concentration can increase response to all acyl-HSLs, they can not account for changes in specificity. Acyl-HSL specificity has been significantly altered in the all of the variants.

Table 2.3. Specificity of second-generation LuxR variants

<table>
<thead>
<tr>
<th>LuxR type</th>
<th>$[3\text{OC}6\text{HSL}]_{50}$ (nM)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>$[\text{C8HSL}]_{50}$ (nM)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>$S$&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>LuxR wt</td>
<td>10</td>
<td>2000</td>
<td>200</td>
</tr>
<tr>
<td>LuxR-G2A</td>
<td>10</td>
<td>50</td>
<td>5.0</td>
</tr>
<tr>
<td>LuxR-G2B</td>
<td>5</td>
<td>20</td>
<td>4.0</td>
</tr>
<tr>
<td>LuxR-G2C</td>
<td>20</td>
<td>35</td>
<td>1.8</td>
</tr>
<tr>
<td>LuxR-G2D</td>
<td>10</td>
<td>15</td>
<td>1.5</td>
</tr>
<tr>
<td>LuxR-G2E</td>
<td>20</td>
<td>20</td>
<td>1.0</td>
</tr>
<tr>
<td>LuxR-G2F</td>
<td>1.5</td>
<td>20</td>
<td>13.3</td>
</tr>
<tr>
<td>LuxR-G2G</td>
<td>2.5</td>
<td>30</td>
<td>12.0</td>
</tr>
<tr>
<td>LuxR-G2H</td>
<td>1.5</td>
<td>25</td>
<td>16.7</td>
</tr>
</tbody>
</table>

<sup>a</sup> $[3\text{OC}6\text{HSL}]_{50}$ and $[\text{C8HSL}]_{50}$ are the concentrations of these acyl-HSLs required to reach half-maximal gene activation as measured from the fluorescence output of GFPuv.

<sup>b</sup>The specificity constant, $S$, corresponds to the ratio of the concentrations of $3\text{OC}6\text{HSL}$ and $\text{C8HSL}$ required to achieve half-maximal gene activation, or $[\text{C8HSL}]_{50}/[3\text{OC}6\text{HSL}]_{50}$. 


C.8. Single mutants identify six mutations that increase C8HSL response by LuxR

Because every first-generation luxR mutant contained more than one nucleotide substitution, further information is required to determine which lead to increased response to C8HSL. Single mutants with each of the 13 nucleotide substitutions identified in the six first-generation variants, including four synonymous mutations, were prepared using site-directed mutagenesis. Quantitative liquid-phase characterization showed that amino acid substitutions at positions 33, 45, 95, 116, 135 and 174 confer increases in sensitivity to C8HSL (Table 2.4). Fluorescence levels similar to those obtained with wild-type LuxR were found for substitutions at amino acid position 228 and synonymous mutations at nucleotide positions 162, 414 and 501. Two mutations, one leading to substitution of I96 by valine and the synonymous mutation at nucleotide position 159, are slightly deleterious, causing small decreases in gene activation. The single-mutant containing only the T33 to alanine substitution was more sensitive to C8HSL than LuxR-G1A which contains both the T33A substitution and the synonymous mutation at position 159. However, the response of LuxR-G1D, which contains both the I96V and M135I substitutions, was indistinguishable from the single-mutant containing only the M135I mutation, suggesting that the deleterious effects of the I96V mutation may be masked by the presence of a beneficial substitution.
### Table 2.4. Activation of *gfpuv* gene expression by *luxR* single mutants

<table>
<thead>
<tr>
<th>Nucleotide substitution</th>
<th>Amino acid substitution</th>
<th>Fluorescence (AU)(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No acyl-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HSL</td>
</tr>
<tr>
<td><em>luxR</em> wt</td>
<td><em>LuxR</em> wt</td>
<td>&lt;50</td>
</tr>
<tr>
<td>A97→G</td>
<td>Thr33→Ala</td>
<td>&lt;50</td>
</tr>
<tr>
<td>A133→T</td>
<td>Ile45→Phe</td>
<td>&lt;50</td>
</tr>
<tr>
<td>A159→G</td>
<td>Syn(^b)</td>
<td>&lt;50</td>
</tr>
<tr>
<td>T162→A</td>
<td>Syn</td>
<td>&lt;50</td>
</tr>
<tr>
<td>A284→G</td>
<td>Asn95→Ser</td>
<td>&lt;50</td>
</tr>
<tr>
<td>A286→G</td>
<td>Ile96→Val</td>
<td>&lt;50</td>
</tr>
<tr>
<td>A346→G</td>
<td>Ser116→Ala</td>
<td>&lt;50</td>
</tr>
<tr>
<td>A403→G</td>
<td>Met135→Val</td>
<td>&lt;50</td>
</tr>
<tr>
<td>G405→A</td>
<td>Met135→Ile</td>
<td>&lt;50</td>
</tr>
<tr>
<td>A414→C</td>
<td>Syn</td>
<td>&lt;50</td>
</tr>
<tr>
<td>T501→C</td>
<td>Syn</td>
<td>&lt;50</td>
</tr>
<tr>
<td>A522→G</td>
<td>Ile174→Met</td>
<td>&lt;50</td>
</tr>
<tr>
<td>A683→G</td>
<td>Thr228→Ala</td>
<td>&lt;50</td>
</tr>
</tbody>
</table>

\(^a\)Units of fluorescence represent the fluorescence due to GF Puv production obtained with a given LuxR variant and p*luxGFPuv* as described for Table 2.2.

\(^b\)Syn denotes a synonymous mutation, i.e. a nucleotide change which does not encode an amino acid change.
D. Discussion

The results of this study indicate that LuxR can evolve rapidly to sense and respond to a diversity of acyl-HSL signal molecules. By screening for changes in gene activation in the presence of C8HSL, we were able to identify variants of this signal-dependent transcriptional activator that respond strongly to this molecule and two additional acyl-HSLs. Only two generations and three or fewer residue changes were required to generate LuxR variants that activate gene expression in response to C8HSL with sensitivity equal to its response to the natural signal, 3OC6HSL.

The variants identified in this study may be altered in more than one of a number of properties and processes involved in gene activation by LuxR. LuxR-mediated gene activation involves key events of signal binding, dimerization, conformational changes, DNA binding, and transcriptional activation via interactions with RNA polymerase and other proteins [1]. Changes to LuxR stability (in the presence and absence of a competent acyl-HSL molecule), folding, solubility, and expression could also affect its ability to activate genes. Indeed, the stability of TraR (a LuxR homolog) in the cytosol has been shown to increase dramatically upon binding a competent acyl-HSL [25, 26]. The increase in LuxR concentration observed in the presence of 3OC6HSL, as compared to its concentration without signal, indicates that, like TraR, the stability of LuxR is affected by the presence of acyl-HSL. General improvements in response to various acyl-HSLs could result from changes in many of these properties. Acyl-HSL-specific response, however, can be extricated from these other properties by looking at S, the ratio
of responses, i.e. between the cognate and another acyl-HSL. Using this measure we have shown that the response to C8HSL was enhanced in all of the variants relative to the changes in response to 3OC6HSL, suggesting key changes had occurred beyond general concentration effects, and likely at the level of acyl-HSL binding.

Past studies involving the mutagenesis of luxR have been instrumental in constructing a structural and functional map of LuxR [reviewed in [27]]. Regions of LuxR that are essential for its function were identified by screening random point mutations within luxR [28, 29]. Mutations that led to nonfunctional LuxR variants clustered between residues 79-127 and 184-230. Reversal of the loss-of-function phenotype for several mutations within the N-terminal critical region was achieved by adding high concentrations of 3OC6HSL, suggesting that such mutations weaken the interaction between the sensor protein and its cognate signal [28]. None of the residue positions identified here were shown to be critical for function in the previous studies. Four of the mutations (T33A, I44F, M135I, and M135V) are located outside of the region previously postulated to be involved in acyl-HSL binding, albeit well within the predicted N-terminal domain [27]. Thus, by evolving LuxR to respond to an acyl-HSL molecule to which it is normally insensitive, we have identified residues involved in acyl-HSL specificity and response that are distinct from those revealed during loss-of-function mutant studies (see Fig. 2.4). This was not entirely unexpected, as many laboratory evolution studies have demonstrated that amino acids involved in modulated function are generally distinct from, and more tolerant to change, than those required for folding or function [30].
Figure 2.4. Mutations in LuxR responsible for altering acyl-HSL specificity differ from those which result in a loss of function. The 250-residue LuxR protein is composed of an N-terminal acyl-HSL-binding and regulation domain and a C-terminal DNA-binding and activation domain [27]. Amino acid substitutions identified in this study are shown above the bar, and amino acid substitutions previously shown to lead to a loss of acyl-HSL-mediated LuxR activation are shown below [28, 29, 31, 32].


Analysis of the thirteen single-site mutants showed that a single mutation from each of the first-generation variants was responsible for the observed increase in C8HSL response (Table 2.4), with the exception of LuxR-G1F, in which both the M135V and I174M mutations increase C8HSL sensitivity. Recombining these mutations led to LuxR variants with further increases in their ability to activate gene expression with C8HSL but with varying responses to 3OC6HSL, C5HSL and C14HSL. The absence of N95S in the
second generation may be due to random chance, bias in the construction of the library, or its inability to confer any additional advantage when combined with the other mutations.

D.2. Substitutions at position 135 likely stabilize interactions with straight-chain acyl HSLs

The variants which showed increased response to C8HSL with minimal changes in their response to 3OC6HSL (LuxR-G2A, -G2B, -G2C, -G2D and -G2E) all contain a substitution at position 135 (Fig. 2.2 and Table 2.2). Mutation of M135 to isoleucine or valine appears to stabilize interactions between LuxR and C8HSL without affecting the sensor’s recognition of its cognate signal, 3OC6HSL. Alignments of LuxR with a number of its homologs revealed that proteins that respond to acyl-HSLs containing a carbonyl group at the third carbon of their acyl chain often have a methionine residue at this position, whereas those that recognize straight-chain acyl-HSLs usually contain an isoleucine or a valine residue (see Fig. 2.5). Thus natural and laboratory evolution appear to converge to the same answer. The predominant preference at this position in all LuxR homologs seems to be for a hydrophobic residue (isoleucine, valine, leucine or methionine). Alignment to the published structure of TraR shows that 135 is only two
Figure 2.5. **A.** Multiple sequence alignments of 19 members of the LuxR family of transcriptional regulators. LuxR residue numbering is shown above the alignment. Conserved residues are denoted by shading with black (100% conservation), dark grey with white letters (80-99% conservation), and light grey with black letters (60-79% conservation). Site numbering for residue changes identified in this study are shown above the aligned column bounded by a blue vertical box. **B.** List of LuxR homologs used in the multiple sequence alignment, including the source species and the acyl-HSLs to which they are reported to respond preferentially and strongly.

positions away from an amino acid required for recognition of the 3-oxo moiety of 3OC8HSL in TraR [20], and could potentially be involved inmodulating the response to acyl-HSLs lacking a 3-oxo moiety. It also remains possible that residue 135 does make contact with the acyl-HSL molecule: the structure of LuxR has not yet been determined and may differ significantly from that of its distantly-related homolog, TraR [33].

**D.3. Acyl-HSL specificity is influenced by mutations at LuxR positions 33 and 45**

Each of the second-generation variants which exhibited the greatest increases in gene expression in response to C14HSL (LuxR-G2D to -G2H) contained the T33A mutation (Table 2.2). The T33A mutation may help LuxR to accommodate acyl-HSLs with acyl chains more than six carbons in length. TraR and LasR, which respond to acyl-HSLs having acyl side chains eight and twelve carbons in length, both contain an alanine residue at this position [34, 35]. Substitution of I45 by phenylalanine, on the other hand, led to an increase in GFPuv production in response to C5HSL in each of the three
second-generation variants, LuxR-G2A, -G2B and -G2F, in which it was identified. Therefore, I45F might improve interactions with short-chain acyl-HSLs such as C5HSL.

The identification of four LuxR variants showing increased sensitivity to 3OC6HSL was unexpected, in as much as wild-type already responds strongly to low nanomolar concentrations of this, its cognate signal. Increases in variant protein levels, as observed in the Western blots (Fig. 2.3), due to increases in expression, stability or solubility are the most likely explanation. The crystal structure of TraR was used to map the hypothetical positions of the amino acid substitutions identified in this study (Fig. 2.6). Curiously, the TraR residues which align with LuxR positions I45 and I174 are located at the interface between its signal- and DNA-binding domains. These mutations might act by stabilizing interactions between these two domains required for LuxR’s switch-like behavior, thereby conferring increases in gene activation.

The acyl-HSL sensitivity of LuxR homologs may not have been maximized during natural evolution. LuxR-type proteins may be tunable over short evolutionary periods to meet the physical and chemical nature of the habitats in which the cells encoding them are active. Indeed, the amount of acyl-HSLs that must accumulate for a given quorum-sensing species to exhibit a quorum response can vary widely from low nanomolar to micromolar concentrations [5, 12, 22, 34, 36, 37], and likely reflects the particular physiological ecology of the organism.
Figure 2.6. Hypothetical positions of residues in TraR corresponding to those found to modulate acyl-HSL specificity in LuxR. The crystal structure of the LuxR homolog TraR (PDB 1L3L) has been determined [26]. The two peptides comprising a TraR dimer are shown in red and blue. The double-stranded DNA to which TraR binds is shown in green. The two 3-oxooctanoyl-homoserine lactone (3OC8HSL) molecules are shown using space-filling spherical atoms (yellow represents carbon, red represents nitrogen, and blue represents oxygen). The amino acid residues which align with those identified in the LuxR variants exhibiting altered acyl-HSL responses are shown as sticks with coloring similar to that used for 3OC8HSL. The amino acids identified occur predominantly within the N-terminal domain and do not appear to make direct contact with the signal molecule. Because LuxR and TraR share less than 20% sequence identity at the amino acid level, these assignments are purely hypothetical. However, the overall structures of the two homologs are expected to be in large part similar. Their alignment can provide a useful first approximation of residue location within the three-dimensional structure of any LuxR-type protein.
While this study indicates that LuxR can rapidly evolve to respond to a broad range of acyl-HSL molecules, it did not reveal which or how many evolutionary changes are required to refocus it to a new signal molecule. Our findings indicate that the easiest evolutionary solution to generating a protein that responds to a new signaling molecule is by broadening its specificity. This is consistent with other directed evolution work that has shown that refocusing binding to shift, rather than broaden, specificity often requires additional evolutionary work [38]. Natural evolution of signal and substrate specificity has also been shown to fluctuate between specialized states through broad-specificity intermediates, allowing for the acquisition of new functions by expanding the capabilities of “old” proteins [39-43]. Laboratory evolution experiments which have successfully refocused binding or substrate specificity to a new target have specifically included screening for activity with a new substrate and the absence of activity with the enzyme’s natural substrate [44, 45]. In future laboratory evolution studies with LuxR, we hope to understand the basis for acyl-HSL sensing as well as how diverse acyl-HSL-based intercellular communications systems have evolved and continue to evolve.
E. Materials and Methods

E.1. Bacterial strains, media and growth conditions

The bacterial strains used in this study were: *Escherichia coli* strain DH5α (F φ80dlacZΔM15 Δ(lacZYA-argF) U169 deoR recA1 endA1 hsdR17(π-, m+) phoA supE44 λ- thi-1 gyrA96 relA1) and *E. coli* DH5α containing puxGFPuv (DH5α(puxGFPuv)), a pPROTet.E133-derived LuxR/acyl-HSL-inducible GFPuv expression vector encoding chloramphenicol resistance. Competent DH5α and DH5α(puxGFPuv) cells were prepared by using the Z-competent *Escherichia coli* Transformation Kit (Zymo Research, Orange, CA). *E. coli* strains were cultured at 37 °C in LB medium or on LB agar plates. Antibiotics were added at the following concentrations: 100 μg · mL⁻¹ chloramphenicol, or 20 μg · mL⁻¹ kanamycin. Acyl-HSL stock solutions of 1 and 10 mM were prepared by dissolving appropriate amounts in ethyl acetate acidified with glacial acetic acid (0.01% v/v), and stored at −20 °C. Acyl-HSLs used in these studies were: 3-oxohexanoyl-DL-homoserine lactone (3OC6HSL; Sigma Aldrich, St. Louis, MO), pentanoyl-L-homoserine lactone (C5HSL; B. Hauer), octanoyl-DL-homoserine lactone (C8HSL; Fluka, St. Louis, MO), and tetradecanoyl-DL-homoserine lactone (C14HSL, Fluka). For screening experiments, C8HSL was amended to LB agar media prior to solidification to ensure complete mixing. For liquid-phase experiments, acyl-HSL was dispensed into sterile tubes, the ethyl acetate was evaporated under a stream of air, and sterile medium was added to the dried acyl-HSL. All liquid media containing acyl-HSL were prepared immediately prior to use.
E.2. Plasmid construction

All synthetic oligonucleotide primers used in this chapter are listed in Table 2.5. We constructed the LuxR expression vector, pLuxR, and the signal response screening plasmid, pluxGFPuv, using the two-plasmid pPROLar.A122 and pPROTet.E133 system from BD Biosciences Clontech (Palo Alto, CA) as plasmid backbones. Plasmid pLuxR encodes LuxR under the control of the hybrid $P_{lac/ara-1}$ promoter and was constructed by cloning PCR-amplified $luxR$ from pKE705 [31], using primers 5-LuxR and 3-LuxR, into KpnI and BamHI digested pPROLar.A122. The sequence of pLuxR, including the promoter and the $luxR$ gene, was verified by sequencing using the upstream primer 5-LarSeq2 and internal $luxR$ primer LuxRSeq(int). The pLuxR plasmid contains a kanamycin resistance gene and the p15A replication origin. Plasmid pluxGFPuv encodes a variant of green fluorescent protein (GFPuv from pGFPuv, BD Biosciences Clontech, Palo Alto, CA) under the control of the $P_{luxI}$ promoter. $P_{luxI}$ was PCR-amplified from pKE555 [8] using the primers 5-pluxI and 3-pluxI and $gfpuv$ was PCR-amplified from pGFPuv using 5-GFPuv and 3-GFPuv. These $P_{luxI}$ and $gfpuv$ PCR products were assembled and amplified by PCR with 5-pluxI and 3-GFPuv and cloned into pPROTet.E133 between its AatII and HindIII sites. The sequence of the entire promoter region and $gfpuv$ gene on pluxGFPuv was verified by DNA sequencing using the primers 5-pPROTetSeq and GFPuvSeq(int). This plasmid contains a chloramphenicol resistance gene and the ColE1 replication origin. The plasmid encoding the C-terminal fusions of c-myc with LuxR was cloned using the primers 5-LuxR and 3-LuxR-myc with pLuxR as
template. The fusion was cloned into KpnI and BamHI digested pPROLar.A122 to make pLuxR-myc. Versions of pLuxR-myc encoding the second-generation LuxR variants fused to c-myc were cloned similarly to pLuxR-myc using the appropriate pLuxR plasmid encoding the mutant luxR as template. BamHI, HindIII and KpnI were purchased from Roche Applied Science (Indianapolis, IN) and AatII was purchased from New England Biolabs (Beverly, MA).

**Table 2.5.** Oligonucleotide primers used in Chapter 2

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-LuxR</td>
<td>5-CGAACGGGGGTACCCATGAAAAACATAAATGCCGACGACAC-3</td>
</tr>
<tr>
<td>3-LuxR</td>
<td>5-CGTTTCGCGATCCCGTACTTAAATTATAAGTATGGG CAATC-3</td>
</tr>
<tr>
<td>5-LarSeq2</td>
<td>5-CCTGAGCAATCACCTATGAACGTGTC-3</td>
</tr>
<tr>
<td>LuxRSeq(int)</td>
<td>5-CGAAACATCAGGTCTTTATCACACTGGG-3</td>
</tr>
<tr>
<td>5-pluxI</td>
<td>5-CGAACGCAGTCAGTCTTTGATTCTAATAAATGGGA TTTTGTAC-3</td>
</tr>
<tr>
<td>3-pluxI</td>
<td>5-CTTCTCCTTAACCTATGCCGATCCCTTCCTTGC-3</td>
</tr>
<tr>
<td>5-GFPuv</td>
<td>5-GGGAGGTTGTTATGAGTAAAGGAGAAGAAGACTTTTTCAT-3</td>
</tr>
<tr>
<td>3-GFPuv</td>
<td>5-GTACCCAAAGCTTTTATTTGCTAGCTCATCCATGCCATG-3</td>
</tr>
<tr>
<td>5-pPROTetSeq</td>
<td>5-CCTTCTTAACCTGCGATCCCTTCCTTTCG-3</td>
</tr>
<tr>
<td>GFPuvSeq(int)</td>
<td>5-CGAAGGTATGTGAGAGAAGCAGAC-3</td>
</tr>
<tr>
<td>3-LuxR-myc</td>
<td>5-CGTTTCGCGATCTCCATAGATCTCTCTCGTGATCGCTTTCT GTCAGCTCCACCATTAAAGTATGGGCAATCAATTG-3</td>
</tr>
</tbody>
</table>
E.3. Library construction and screening

Error-prone PCR reactions were performed using *AmpliTaq* DNA polymerase (Applied Biosystems, Foster City, CA) and 50 μM MnCl₂ to increase the mutation rate as described [46]. The primers 5-luxR and 3-luxR were used to amplify the luxR gene using pLuxR as the template. The library was constructed by ligating *KpnI* and *BamHI* digested pPROLar.A122 with the products of error-prone PCR using T4 DNA ligase (Invitrogen, Carlsbad, CA). Vent DNA polymerase (New England Biolabs, Beverly, MA) was used to amplify wild-type luxR, which was digested and ligated into pPROLar.A122 for use as a control. The ligation mixtures were transformed into competent DH5α cells harboring pluxGFPuv and plated onto LB agar containing the appropriate antibiotics and 50 or 200 nM C8HSL. The plates were incubated at 37 °C for 18 hours prior to screening. To identify clones of interest, the agar plates were placed over a UV-transilluminator (VWR Scientific, West Chester, PA) at 365 nm and visually inspected for GFPuv fluorescence. The plates were stored at 4 °C and reexamined every 2-3 hours until colonies on a control plate expressing wild-type LuxR became visibly fluorescent.

Second-generation libraries were constructed by DNA shuffling as described [24]. The primers 5-luxR and 3-luxR were used to amplify the mutant luxR genes from six first-generation mutants using *Pfu* Turbo DNA polymerase (Stratagene, La Jolla, CA). After purification and quantification, equal amounts of parent amplification products were mixed and subjected to a DNAsel I digestion. The 100 μL digest contained ca. 4 μg of the parental mix, 10 μL of 0.5 M Tris-HCl (pH 7.4), 5 μL of 0.2 M manganese chloride and
0.14 units of DNAse I. After one minute of digestion at 15°C, the reaction was stopped by adding 5 μL of 1 M EDTA and by placing the mixture immediately on ice. The QIAEXII gel extraction kit (Qiagen, Valencia, CA) was used to purify fragments ranging from 50 to 150 bp in length. Fragments were randomly reassembled in a 50 μL reaction. Full-length luxR genes were synthesized by diluting the reassembly reaction 50- to 500-fold and amplified using Pfu Turbo DNA polymerase and the primers 5-luxR and 3-luxR. Ligations and transformations were performed similarly to the first generation and plated onto LB-agar containing 50 nM C8HSL. All 102 colonies that fluoresced after 18 hours were picked, purified, and inoculated into 1 mL LB containing the appropriate antibiotics. After 24 hours at 37 °C, the cultures were diluted 10-fold into fresh LB, and replica-stamped using a 96-well pin replicator onto solid media containing 10, 20, 50, 100 and 200 nM C8HSL. Ten mutants that fluoresced in response to 10 or 20 nM C8HSL were identified for further characterization.

E.4. Re-cloning and DNA sequencing

All mutants identified during the first- and second- generation screens were re-cloned into fresh background plasmids and strains to eliminate secondary-site effects. For each mutant, the luxR allele was amplified using Pfu Turbo polymerase and treated with DpnI. The PCR products were digested and ligated into pPROLar.A122 (as above) and transformed into competent DH5α cells containing pLuxGFPuv. The promoter and luxR
gene from all mutants of interest were sequenced using the upstream primer 5-LarSeq2 and internal luxR primer LuxRSeq(int).

E.5. Site-directed mutagenesis

DNA fragments containing each single point mutation of interest were generated using standard methods [47]. In the first step, the 5-LuxR primer and a reverse primer containing the desired mutation were used to amplify luxR upstream of the mutation; the 3-LuxR primer and a forward primer also containing the mutation were used to amplify luxR downstream of the mutation. In the second step, the two overlapping fragments were assembled to yield a full length luxR containing a single nucleotide alteration by PCR using the products of step one as template and the 5-LuxR and 3-LuxR primer set. These PCR products were digested and ligated into KpnI and BamHI digested pLuxR to replace the wild-type luxR. Sequences of each site-directed mutant were verified by DNA sequencing.

E.6. Quantitative characterization of LuxR-mediated gene expression of GFPuv

The assay used to measure LuxR-mediated gene activation was adapted from a light-based bioassay protocol [18]. Cells were first grown from single colonies or glycerol stocks in LB overnight, then diluted 200-fold into 100 mL of fresh LB medium
containing 5 mM potassium phosphate buffer, pH 6.5, and the appropriate antibiotics. Such cultures were incubated with shaking at 37 °C until they reached an OD$_{600}$ of 0.5, and then harvested by centrifugation. Cell pellets were washed and resuspended to an OD$_{600}$ of 0.6 in bioassay medium (0.05% w/v tryptone, 0.03% v/v glycerol, 100 mM sodium chloride, 50 mM magnesium sulfate and 5 mM potassium phosphate buffer, pH 6.5, containing antibiotics. The suspension was subsequently transferred into 48-well plates (VWR International) containing 0.5 mL bioassay medium with acyl-HSL, to a total volume of 2.5 mL per well. Thereafter, the 48-well plates were shaken at 37°C for 4 hours. From each well, 200 μL was transferred to wells of a white 96-well microplate with a clear bottom. GFPuv fluorescence (395 nm excitation, 509 nm emission, 495 nm cutoff) was measured using a fluorescence microtiterplate reader (Molecular Devices, SpectraMAX Gemini XS); cell densities were measured using a microtiterplate reader at 600 nm. Fluorescence by cell suspensions was normalized to optical density. The fluorescence output of pPROLar.A122 with pluxGFPuv was used to determine the background fluorescence without LuxR, and this background fluorescence value was subtracted from all fluorescence measurements obtained with wild-type pLuxR and all mutants to determine fluorescence due only to LuxR-dependent gene activation. For a given variant, the same batch of cells was used for comparing responses to (1) different concentrations of a given acyl-HSL, or (2) different acyl-HSLs.

**E.7. Assays of LuxR abundance in vivo by Western immunodetection**
To estimate the relative concentrations of LuxR in uninduced and fully induced second-generation variants, we constructed C-terminal LuxR fusions to the amino acid sequence GGAEQKLISEEDL, i.e. the c-myc epitope tag with an N-terminal GGA linker. pLuxR-myc was transformed into DH5α(pluxGFPuv), the protein it encodes was determined to function similarly to its parent LuxR without the tag. Strains expressing the C-terminal c-myc fusions of wild-type LuxR and the second-generation LuxR variants were grown, harvested, resuspended, and incubated in the absence of acyl-HSL or with 1 μM 3OC6HSL in a similar manner as during the quantitative characterization of LuxR-mediated gene expression of GFPuv (see above). After incubation in the presence or absence of acyl-HSL and after monitoring OD$_{600}$ and fluorescence: 1 mL of each cell suspension of interest was resuspended in 50 μL SDS loading buffer. Samples were boiled for 5 min and centrifuged for 10 min before the proteins in the clarified fluid were size fractionated via standard SDS-PAGE. Proteins were electrotransferred to nitrocellulose membranes (Invitrogen). After an 8 hour blocking step using tris-buffered saline (TBS; 50 mM Tris-HCl, 150 mM NaCl, pH 7.5) buffer containing 5% non-fat dried milk, membranes were incubated for 8 h in TBS buffer containing 1% non-fat dried milk and 4% bovine serum albumin (BSA, Sigma) with anti-myc mouse antisera (Invitrogen). Antibody-protein complexes were visualized using horseradish peroxidase-coupled anti-murine second antibody (Invitrogen) and SuperSignal West Pico Substrate (Pierce, Rockford, IL). Chemiluminescence was detected using Kodak BioMax Light Film (Rochester, NY). The developed film was subsequently scanned using a Hewlett Packard Scanjet 4400c and HP Precisionscan Pro software (version 3.1). Protein
concentrations, relative to wild-type LuxR, were determined from the scanned image via densitometry analysis using Quantity One software (Bio-rad, version 4.1.1).
F. References


CHAPTER 3

Design and Evaluation of a Dual Selection System for
Identifying LuxR Variants with New Specificities
A. Abstract

To continue to our directed evolution of LuxR towards variants with new specificities, we constructed a selection system capable of both positive (ON) and negative (OFF) selection. An ON/OFF selection system is required to identify specific proteins because positive selection alone generally leads to variants with broadened specificities. A dual selection module developed previously was adapted to incorporate the \( P_{luxI} \) promoter and was shown to enrich by 490-fold per round of ON/OFF selection. Additional rounds of selection were necessary to completely eliminate the presence of nonfunctional plasmids. This work both demonstrates the potential of this dual selection system to enrich for functional LuxR variants from very large libraries (>\(10^6\)) and suggests controls that can be used to decrease the probability of identifying false positives.
B. Introduction

The previous chapter described the directed evolution of LuxR variants that respond to a broad range of acyl-homoserine lactones (acyl-HSLs). To identify an evolved LuxR variant that retains (or increases) its newly-acquired ability to respond to a non-natural signal molecule and decreases its ability to respond to LuxR’s cognate signal, 3-oxohexahoyl-homoserine lactone (3OC6HSL), a selection system capable of both positive and negative selection was developed.

Gain-of-function mutants are identified by positive selection. While it is possible for a single amino acid substitution to both increase a protein’s ability to recognize a new target and decrease, or prevent, the recognition of its natural target [1-3], it has been observed that this occurs infrequently in directed evolution experiments targeting new functions [4-6]. Instead, initial rounds of directed evolution often produce variants that exhibit broadened specificity [2, 4-11]. Matsumura and Ellington noted that their β-glucoronidase variants were active with substrates that were very dissimilar to both the native and target substrates [5]. Similar observations led Aharoni et al., to postulate that the observed broadening of specificity is a common feature of protein evolution [4]. They also noted that it is evolutionarily advantageous for a protein to be able to increase its affinity for a new target without affecting its affinity for its native target molecule.

But what happens when it is necessary to refocus specificity? Matsumura and Ellington used repeated rounds of screening for increases in activity with their target substrate and
then characterized several mutants to find the desired inversion of specificity [5]. A more
direct approach uses screening or selecting for both an increased response to a new target
and a loss of response to the native target. The addition of negative rounds of selection or
screening can be used to identify variants with new specificities without having to sift
through a large number of variants identified by positive rounds and hoping that a variant
with decreased affinity for the native molecule is present.

Santoro and Schulz compared positive screening alone to a combination of positive and
negative screening in their directed evolution experiments targeting Cre recombinase
variants capable of recognizing a new DNA target [12]. They showed that positive
screening yielded variants capable of recombining at either the wild-type or mutated
DNA target. The combination of both positive and negative screening led to the
identification of a variant with an increase in recombination at the new target site and a
decrease in its ability to act at its native site. The directed evolution of Cre recombinase
illust rates the importance of including rounds of both positive and negative selective
pressure in order to identify variants with an explicit specificity profile. Positive and
negative screens have also been successfully implemented to identify Cre recombinases
that recognize a different DNA site [13], tRNA-synthetases [14] and the endopeptidase
OmpT [3].

Physical models have also alluded to the importance of including both positive and
negative “pressure” in order to target specificity [1, 15, 16]. In designing specific
protein-protein or protein-DNA interactions, positive design is used to stabilize interactions with a desired target while negative design destabilizes undesired interactions. A physical model of DNA binding by zinc fingers showed that these proteins are optimized for specificity rather than affinity [1]. This result implies a trade-off between specificity and affinity may exist and, more importantly, it reinforces the experimental evidence that looking only for increases in affinity for a new target will infrequently yield variants with refocused specificities.

Before searching libraries for LuxR variants with new specificities, it was imperative that we both design a dual selection system and evaluate its ability to identify functional clones from a pool of nonfunctional “garbage.” We have implemented an ON/OFF selection system for the directed evolution of LuxR based on the dual selection module previously developed by Yokobayashi and Arnold [17] for the identification of a functional repressor. Their selection system uses the expression of the tetracycline resistance gene, \textit{tet}A, to confer survival during ON rounds of selection. OFF rounds of selection use the production of β-lactamase inhibitory protein (Bli) to confer sensitivity to β-lactam antibiotics when unwanted transcription occurs. β-lactamase (bla), which is constitutively expressed from the selection plasmid, is inhibited when bound by Bli. The \textit{bli} gene was extensively modified by Yokobayashi and Arnold [17] to lower the GC content of the gene and to decrease the homology of the periplasmic leader sequences used by \textit{bli} and \textit{bla}. The reduced homology between the two leader sequences lowers the
probability of recombination events that could eliminate bli and undermine the selection process by producing false positives.

The primary advantage of using a selection system, as opposed to a screen, is an increase in potential library size to more than $10^6$. While screens can be cumbersome and limit library size, selections are often limited by the prevalence of false positives that occur when bacteria (or other organisms) are under selective pressure. By employing a dual selection system, which requires proteins to be ON or OFF under different environmental conditions, the frequency of false positives is expected to decrease. We have implemented and evaluated a dual selection system for directed evolution of the transcriptional activator LuxR and its cognate promoter, $P_{luxI}$, and have demonstrated that it can enrich a functional activator approximately 490-fold per ON/OFF round of selection.
C. Results

C.1. Modification of a dual selection system for LuxR-mediated gene activation

Our goal was to adapt the generic selection module developed by Yokobayashi and Arnold [17] to LuxR-mediated transcriptional activation, such that the ability of LuxR to activate gene expression at $P_{\text{lux}}$, under a given set of conditions, would correspond to the ability of a host bacterium to survive in the presence of different antibiotics. While the original dual selection system of Yokobayashi and Arnold [17] encodes both the positive (TetA) and negative (Bli) elements on a single selection plasmid, replacement of the repressed promoter with the $lux$ promoter did not lead to simultaneous LuxR-mediated expression of both elements. We observed that the amount of basal gene expression of the second gene ($tetA$) in the operon always generated an ON output, irrespective of LuxR. Switching the order of the genes and replacing $tetA$ with $cat$ did not alleviate this problem. In general, promoters that require an activator to recruit RNA polymerase have higher levels of basal gene expression than promoters that are repressed by transcription factors blocking RNA polymerase binding [18]. Instead of trying to tune the output levels of the ON and OFF selection genes simultaneously, we separated the components into ON and OFF selection plasmids (Fig. 3.1). This requires that the selection plasmid be inactivated and replaced following each round of selection. However, separate ON and OFF selection plasmids act to lower the probability of propagating mutated selection plasmids that confer antibiotic resistance without LuxR, i.e. false positives, from one round to the next.
Figure 3.1. Schematic of the dual selection system for the directed evolution of LuxR variants with altered specificity. During rounds of ON selection, LuxR variants that are active under the desired conditions are identified by their ability to survive in the presence of chloramphenicol (Cm) due to LuxR-mediated activation of the chloramphenicol acetyl-transferase (cat) gene from the PluxI promoter. During rounds of OFF selection, LuxR variants which are inactive are identified by their ability to survive in the presence of carbenicillin (Cb) due to a lack of LuxR-mediated expression of the β-lactamase inhibitory protein (bli) gene. During OFF rounds of selection, β-lactamase is constitutively expressed in the cells and is inhibited by Bli, rendering the cells sensitive to carbenicillin.

The ON selection plasmid, pluxCAT, confers chloramphenicol resistance in the presence of LuxR-mediated transcriptional activation of the cat gene from the P_{luxI} promoter. The
OFF plasmid, pluxBLIP, renders bacteria in which LuxR is active sensitive to carbenicillin by producing Blip, which inactivates the constitutively expressed Bla. Thus an absence of LuxR-mediated gene activation is detected by survival in the presence of carbenicillin with the OFF selection plasmid.

C.2. Evaluation of the dual selection system

To test our modified selection module, we selected a plasmid expressing luxR (pLuxR, Fig. 3.2B) from a background of nonfunctional plasmids. The nonfunctional plasmids either lack the luxR gene (pPROLar.A122, Fig. 3.2A) or constitutively express both cat and bli (pBLIPCAT, Fig. 3.2C). While pLuxR confers survival with the ON selection plasmid in the presence of 3OC6HSL and survival with the OFF selection plasmid in the absence of 3OC6HSL, pPROLar.A122 should survive only OFF selection rounds while pBLIPCAT should survive only ON rounds of selection.

The three plasmids, pLuxR, pPROLar.A122 and pBLIPCAT, were mixed with a 1:500:500 molar ratio and transformed into DH5α cells harboring the ON selection plasmid luxCAT. The transformed cells were plated on chloramphenicol-containing LB plates with 100 nM 3OC6HSL to select for cells in the ON state (pLuxR or pBLIPCAT). Cells were recovered by overlaying liquid LB medium and the DNA from these cells was isolated by miniprep. To inactivate the ON selection plasmid, the recovered plasmid
DNA was digested with *Apa*LI, which cleaves pluxCAT and pluxBLIP but does not cleave pLuxR, pPROLar.A122 or pBLIPCAT.

---

**Figure 3.2.** Plasmid maps of plasmids used to test dual selection system.  

**A.** pPROLar.A122 is an “empty” vector. It does not encode a functional *luxR* gene and P<sub>luxI</sub> is never activated in cells containing this plasmid.  

**B.** pLuxR encodes the wild-type *luxR* gene under the control of a P<sub>lac/ara-1</sub>. LuxR is constitutively expressed in DH5<sub>α</sub>, which does not contain *lacI* or *araC*. LuxR activates gene expression from P<sub>luxI</sub> only in the presence of 3OC6HSL.  

**C.** pBLIPCAT constitutively expresses both of the selection proteins, CAT and Bli. CAT and Bli are expressed from this plasmid regardless of 3OC6HSL concentration.
The *Apa*LI-digested DNA from Round 1-ON was used to transform DH5α cells harboring the OFF selection plasmid, pluxBLIP. Cells were spread on plates containing carbenicillin without 3OC6HSL to select for cells in the OFF state (pLuxR or pPROLar.A122). Round 1-OFF cells were recovered, and the DNA was isolated and digested as following Round 1-ON. The purified DNA was transformed into DH5α containing pluxCAT to initiate Round 2 of ON/OFF selection.

From each round of selection, 45 colonies were randomly picked and tested for the presence of pLuxR, pPROLar.A122 or pBLIPCAT by colony PCR. Because these plasmids each contain the same backbone, PCR with primers that bind outside of the inserted region amplifies a section of differing length from each of the three plasmids. Amplification of pLuxR via PCR yields a 1100 base pair (bp) fragment, while pPROLar.A122 yields a 400 bp fragment and pBLIPCAT a 1650 bp fragment, allowing the three plasmids to be easily distinguished from one another.

Following Round 1-ON, 45 of the 45 (100%) colonies picked contained the pBLIPCAT plasmid (Fig. 3.3). This was expected due to the 500-fold excess of this plasmid over pLuxR. Twenty-two of the 45 (49%) colonies from Round 1-OFF contained pLuxR (Fig. 3.3). The other 23 colonies contained pPROLar.A122. A 490-fold enrichment of pLuxR was achieved in the first round of ON/OFF selection.
As shown in Fig. 3.3, 44 of 45 colonies contained pLuxR after Round 2-ON. The remaining colony harbored pPROLar.A122. Round 2-OFF returned exactly the same ratio of pLuxR to pPROLar.A122 as the second ON selection because both plasmids confer the same probability of survival OFF rounds. A third ON round yielded 100% of colonies containing pLuxR
Figure 3.3. Agarose gel of colony PCR products used to determine the composition of the plasmid mixture after each round of positive and negative selection. Forty-five colonies were randomly picked and colony PCR with primers that bind to the common plasmid backbone of each of the three test plasmids was used to identify if test colonies contained pPROLar.A122 (Lar), pLuxR (LuxR) or pBLIPCAT (BLIPCAT). Controls of each plasmid are shown on the left three lanes of each gel.
D. Discussion

Implementation of an ON/OFF selection system for identifying LuxR variants required modifying the dual selection system developed by Yokobayashi and Arnold for identifying functional repressors [17]. Using a single selection plasmid for both ON and OFF rounds of selection was undermined by an excessive basal level of protein production from the second gene in the operon. The reason for the observed increase at only the second gene of the operon is unknown, however, an overall increase in the basal level of gene expression is not unexpected at a promoter that is modulated by an activator instead of a repressor [18]. Activators increase RNA polymerase’s affinity for a given promoter and a low level of transcription is expected because the polymerase will occasionally bind to the promoter without being recruited by the activator. On the other hand, repressors physically block RNA polymerase from binding to the promoter so basal levels of transcription from these promoters are significantly lower. The modifications made here are likely applicable to the evolution of other activator proteins using similar dual selection modules.

The use of two separate selection plasmids required their inactivation following each round of selection. This was accomplished by selective digestion with an endonuclease that cleaves the selection plasmids but leaves the circuit plasmids intact. While this step increases the labor required for each ON/OFF round of selection, it ensures that any selection plasmids that have mutated and confer resistance without LuxR are not propagated. Due to this control, the frequency of false positives that arise due to mutated
selection plasmids should be consistent from round to round while the frequency of functional clones should increase exponentially. This reduces the probability of identifying false positives at the end of a directed evolution experiment.

The background level of nonfunctional plasmids in Round 2 indicates that the ON rounds of selection could have been more stringent; however, an extra round of selection was shown to eliminate the nonfunctional plasmids. To ensure that only functional variants are picked for further characterization, the isolated clones could be screened for their ability to activate the expression of a non-lethal reporter protein, such as the GFP variant used in Chapter 2 [7]. A low frequency of nonfunctional clones would not prevent the identification of functional clones with a final screen and could potentially decrease the number of selection rounds required because it would be unnecessary to ensure that all nonfunctional clones had been eliminated. Furthermore, a screen would add the ability rapidly identify the “best” variants from a pool that meet the minimum requirements imposed by the ON/OFF conditions used with dual selection system.

The 490-fold enrichment of functional LuxR variants per round of selection suggests that three rounds of ON/OFF selection should allow a functional LuxR variant to be identified from a library of more than 100 million. The proportional increase in functional variants in each round of selection is limited by the frequency of false positives that arise each round. Decreasing regions of high sequence similarity can be used to decrease the probability that a recombination event will lead to a false positive [17]. If false positives
were plaguing the selection system, it would be worthwhile to investigate other mechanisms that were allowing the bacteria to survive without the desired gene activation, or absence thereof.

With our new selection system for identifying LuxR variantd, the 490-fold enrichment is likely sufficient for identifying LuxR variants from the size of libraries that are possible to generate. The removal of the selection plasmids after each round of selection is important for decreasing the number of false positives identified. The following chapters describe the use of this dual selection system for identifying LuxR variants that respond to a non-cognate acyl-HSL but that have lost their ability to respond to 3OC6HSL and LuxR variants that activate gene expression at a mutated lux promoter.
E. Materials and Methods

E.1. Bacterial strains, media and growth conditions.

The bacterial strains used in this study were: *Escherichia coli* strain DH5α (F-φ80dlacZΔM15 Δ(lacZYA-argF) U169 deoR recA1 endA1 hsdR17(rk-, mk+) phoA supE44 λ- thi-1 gyrA96 relA1) and *E. coli* DH5α containing pluxCAT (DH5α(pluxCAT)) or pluxblip (DH5α(pluxBLIP)). Competent cells were prepared by using the Z-competent *Escherichia coli* Transformation Kit (Zymo Research, Orange, CA). *E. coli* strains were cultured at 37 °C in LB medium or on LB agar plates. Antibiotics were added at the following concentrations: 150 μg/mL chloramphenicol, 100 μg/mL carbenicillin, and 50 μg/mL kanamycin. 3-Oxo-hexanoyl-homoserine lactone (3OC6HSL; Sigma Aldrich, St. Louis, Mo) stock solutions of 1 and 10 mM were prepared by dissolving appropriate amounts in ethyl acetate acidified with glacial acetic acid (0.01% v/v), and stored at -20 °C. For solid-phase experiments, 3OC6HSL was added to LB agar media prior to solidification to ensure complete mixing.

E.2. Plasmid construction

The LuxR expression vector, pLuxR, and the signal response plasmid, pluxGFPuv, have been described previously [7]. Plasmids pluxCAT and pluxBLIP are similar to pluxGFPuv but encode chloramphenicol acetyltransferase (CAT) or β-lactamase inhibitory protein (Bli) under the control of the P_{lux} promoter and contains the β-
lactamase gene (*bla*). The pBPROTet plasmid was constructed by replacing the SacI-AatII *cat* fragment in pPROTet.E133 with the β-lactamase (*bla*) encoding SacI-AatII fragment from pλPRO12Select[17]. *P_{luxI}* was PCR-amplified from pluxGFPuv using the primers 5-pluxI [7] and 3-lux(CAT) or 3-lux(Blip). *cat* was PCR-amplified from pPROTet.E133 using 5-CAT(lux) and 3-CAT(BamHI) (see Table 3.1 for oligonucleotide sequences). *blip* was PCR-amplified from pλPRO12Select using 5-BLIP(lux) and 3-BLIP(HindIII). The *P_{luxI}* and *cat* or *blip* PCR products were assembled and amplified by PCR with 5-pluxI and 3-CAT(BamHI) or 3-BLIP(HindIII) and cloned into pBPROTet between its AatII and BamHI or HindIII sites.

The pBLIPCAT plasmid was constructed by PCR-amplifying *blip* from pλPRO12Select using 5-BLIP(KpnI) and 3-BLIP(HindIII) and *cat* from pPROTet.E133 with 5-CAT(BLIP) and 3-CAT(BamHI). These *blip* and *cat* PCR products were assembled and amplified by PCR with 5-BLIP(KpnI) and 3-CAT(BamHI) and cloned into pPROLar.A122 between its KpnI and BamHI sites.
Table 3.1. Oligonucleotide primers used in Chapter 3

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
</table>
| 5-pluxI         | 5-CGAACGCACGTCACTCCTGTTGAT  
                    TCTAATAATGGATTGTGTACCC-3 |
| 3-pluxI         | 5-CTTTCCTTTTACTCATACCAACCCTCCCTGCGTTTATTC-3                            |
| 5-LarSeq2       | 5-CCTGAGCAATCACCTATGAACTGC-3                                          |
| 3-lux(CAT)      | 5-GATTCTTTTACTCCATACCAACCCTCCCTGCGTTTATTC-3                           |
| 3-lux(BLIP)     | 5-ACCTTATATAATGCATACCAACCCTCCCTGCGTTTATTC-3                           |
| 5-CAT(lux)      | 5-GGGAGGTTGATGAGAAATATACACTGGATAATCC-3                                |
| 3-CAT(BamHI)    | 5-GTTAGCGGACCGAGCGATCGATATCAAATTACGCC-3                               |
| 5-BLIP(lux)     | 5-GGGAGGTTGATGAGAATATACACTGGATAATCC-3                                 |
| 3-BLIP(HindIII) | 5-GTTAGCAAGCTTTTATAACAGGTCCC-3                                        |
| 5-BLIP(KpnI)    | 5-CGAACGGGTATACGTATTATATAATATGAGCAATACC-3                             |
| 5-CAT(BLIP)     | 5-GGGACCTTTGTAATATAAGCTTGCTAAACGATACGCC-3                             |
| 3-LuxR+200      | 5-GTGAGCGAGGAAAGCGGAATATATCC-3                                        |

E.3. Selection protocol

For positive selection, circuit plasmids were transformed into DH5α cells harboring the ON selection plasmid (pluxCAT) and plated on LB agar plates containing 100 nM 3OC6HSL, kanamycin and chloramphenicol. For negative selection, circuit plasmids were transformed into DH5α cells harboring the OFF selection plasmid (pluxBLIP) and plated on LB agar plates containing kanamycin and carbenicillin. Plates were incubated for 18 hours at 37°C. Selected cells were recovered by overlaying liquid LB medium over the agar plates. Plasmid DNA from the cells was recovered by miniprep (QIAGEN). To inactivate the selection plasmids, the DNA was digested with ApaLI and purified. The digested DNA was subsequently used to transform the next round of selection.
E.4. Plasmid identification by colony PCR

Colony PCR was performed to identify each of the functional plasmids after each round of selection. The Colony Fast-Screen Kit (Epicentre) was used according to instructions, with two primers: 5-LarSeq2 and 3-LuxR+200. These primers bind to the backbone common to each of the three plasmids in the starting mix and allow the specific identification of the plasmids present in each colony. Positive controls of each of the three starting plasmids were run simultaneously.
F. References


CHAPTER 4

Dual Selection Creates a New Signaling Specificity in the
Quorum-sensing Transcriptional Activator LuxR
A. Abstract

The LuxR transcription factor from *Vibrio fischeri* activates gene expression in response to binding the signaling molecule 3-oxo-hexanoyl-homoserine lactone (3OC6HSL). LuxR homologs respond to acyl-homoserine lactones (acyl-HSLs) with diverse chemical structures. We previously described a LuxR variant, LuxR-G2E, that activates gene expression upon binding a broadened range of acyl-HSLs, including straight-chain acyl-HSLs. Here, we use an ON/OFF dual selection system to identify a new LuxR variant that retains the response to straight-chain acyl-HSLs, but no longer responds to LuxR’s cognate signal, 3OC6HSL. A single mutation, arginine 67 to methionine in the N-terminal signal-binding domain, reduces LuxR-G2E’s response to acyl-HSLs having a carbonyl substituent at the third carbon of the acyl chain. This mutation has a similar effect when introduced into a LuxR homolog that responds to a different 3-oxo-HSL (3OC12HSL), LasR from *Pseudomonas aeruginosa*. This specificity-enhancing mutation does not affect LuxR response to straight-chain acyl-HSLs and would not have been identified by positive selection for activation by these signaling molecules. The dual selection system provides a rapid and reliable method for identifying LuxR variants that have the desired response, or lack thereof, to a given set of acyl-HSL signals. This work demonstrates that accumulation of a small number of point mutations is sufficient to switch specificity from one signal to another, passing through a broadly-specific intermediate. LuxR variants with new signaling specificities will be useful components
for constructing artificial cell-cell communication systems that program population-level behaviors such as pattern formation.

B. Introduction

Quorum sensing (QS) is the process by which bacteria sense their local population density and modify the expression of key genes accordingly [1, 2]. While many QS signals have been identified (reviewed by Taga and Bassler [1]), one of the most prevalent and well-studied QS systems employs acyl-homoserine lactones (acyl-HSLs) [3]. In *V. fischeri*, the 3-oxo-hexanoyl-homoserine lactone (3OC6HSL, Fig. 4.1) signal is synthesized by LuxI [4]. In environments of low cell density, the amphipathic acyl-HSL diffuses out of the cell through the membrane [5]. As the local environment becomes more crowded, the acyl-HSLs accumulate in the cell, where they bind to the acyl-HSL-dependent transcriptional activator, LuxR [6, 7]. Upon signal binding, LuxR dimerizes [8] and activates transcription at the *lux* promoter by recruiting RNA polymerase [9]. Pairs of LuxI/LuxR homologs have been identified in many species of Gram-negative *Proteobacteria* [10, 11], where they regulate genes with diverse functions in response to changes in population density [1, 2]. LuxR homologs respond to acyl-HSLs with acyl chains of varying lengths and chemical modifications [12, 13] (Fig. 4.1).
Figure 4.1. Diversity of LuxR homologs and acyl-HSLs. a. Illustration of the divergence of selected LuxR homologs. Homologs are listed with both species and cognate acyl-HSL signal molecule. Homologs discussed in this work are highlighted and the structure of their cognate acyl-HSL signal molecules shown. The tree was adapted from Gray and Garey [10], and Lerat and Moran [11]. b. Structures, names and abbreviations of non-cognate acyl-HSLs used in this study.
LuxR/LuxI and their homologs have been adopted by biological engineers to effect programmed intercellular communication and coordinate the behaviors of cell populations [14-18]. In an early demonstration of engineered cell-cell communication, Weiss and coworkers expressed LuxR in *E. coli* to produce ‘receiver’ cells that would respond to signals produced by ‘sender’ cells that expressed LuxI [17]. More recently, You et al. described a circuit that employs Lux QS components to sense and control the density of a bacterial population by expressing a killer gene, *ccdB*, from the *lux* promoter [18]. Kobayashi *et al.* described input/output modules based on the Lux system that coupled to a genetic ‘toggle switch’ and programmed density-dependent protein production in *E. coli* [16]. Lux QS components have also been used to build a variety of functional modules, including a chemical pulse generator [15] and chemical band-detect network [14]. Synthetic systems exhibiting complex behaviors have exploited these modules, using variants of LuxR that respond to different concentrations of the acyl-HSL signal. For example, using ‘receiver’ cells containing LuxR-based band-detect circuits that express fluorescent proteins over different concentration ranges of the signaling molecule, Basu *et al.* [14] demonstrated that an undifferentiated lawn of the ‘receiver’ cells generated a bulls-eye pattern of gene expression around a central colony of ‘sender’ cells producing 3OC6HSL.

LuxR/I and their homologs can serve a wide range of communication functions in synthetic networks. Applications at the next level of complexity, for example, will implement two-way communications between cellular populations [19, 20]. These
applications, however, require a high degree of specificity in the LuxR receiver functions, since crosstalk among chemical signals can undermine the function of the entire network. LuxR homologs have been investigated, but these can show significant crosstalk [21] and often also behave differently from LuxR in other ways. For applications in synthetic biology, it would be useful to have a set of ‘standardized parts’—for example, a series of LuxR variants which respond to different chemical signals at different levels and specificities—from which new networks can be assembled.

Engineering protein specificity is challenging because it requires both stabilizing a desired interaction (positive design) and destabilizing undesired interactions (negative design) [22]. Directed evolution experiments have also shown that specificity rarely arises as a result of positive selection alone [23]. We previously used directed evolution to create LuxR variants exhibiting broadened acyl-HSL specificity and increased response sensitivity [24]. Random-mutant libraries were screened for increased response to octanoyl-homoserine lactone (C8HSL). All the C8HSL-responsive LuxR variants showed broad response to acyl-HSLs having acyl chains five to fourteen carbons in length; none were specific for the new signal. To generate a LuxR with a new signaling specificity, we modified the dual selection system recently described by Yokobayashi and Arnold [25] to select for LuxR variants that either activate gene expression (ON) or do not (OFF) under desired sets of conditions.
C. Results

C.1. A dual selection system for evolution of LuxR signaling specificity

To identify LuxR variants that respond specifically to a new signal, we adapted a dual selection system developed previously to select for functional genetic networks that properly transition gene expression from ON to OFF under desired conditions[25]. The selection scheme was outlined in Figure 3.1. To select for LuxR variants that activate gene expression under a given set of conditions, i.e., in the presence of a new acyl-HSL, a library of mutagenized luxR genes is selected with the ON selection plasmid pluxCAT, which yields chloramphenicol resistance with LuxR-dependent transcriptional activation. Identification of variants that show a lack of gene activation (i.e., with the wild-type cognate signal) is achieved by transformation with the OFF selection plasmid pluxBLIP, which encodes β-lactamase inhibitory protein (Bli) under the control of the lux promoter and β-lactamase (Bla) under the control of a constitutive promoter. Bli inactivates Bla and renders cells with active LuxR alleles sensitive to β-lactam antibiotics.

C.2. Directed evolution of LuxR-G2E to respond to C10HSL but not 3OC6HSL

To obtain the new LuxR signaling specificity, we started with a broad-specificity variant identified previously, LuxR-G2E (Table 4.1) [24]. To compare gene activation by LuxR and our evolved variants with exogenously added acyl-HSLs, we used a reporter plasmid, pluxGFPuv, that contains gfpuv under the control of the lux promoter. Liquid-phase
**gfpuv** bioassays quantitatively measure fluorescence output due to LuxR-mediated GFPuv production [24]. LuxR-G2E responds to DL-decanoyl-homoserine lactone (DL-C10HSL), while wild-type LuxR does not respond at concentrations to 10 μM. Wild-type LuxR and LuxR-G2E respond similarly to 3OC6HSL (Fig. 4.2).

Using error-prone PCR, we created a library of *luxR-G2E* genes encoding, on average, one to two amino acid substitutions in the N-terminal 160 amino acids of LuxR, shown to be responsible for acyl-HSL binding [26]. Two rounds of ON/OFF selection were used to enrich for LuxR variants that activate gene expression with C10HSL but not 3OC6HSL.

The LuxR expression vectors obtained from colonies surviving Round 2 ON/OFF selection were used as templates to amplify the mutant *luxR* genes by PCR with a high-fidelity polymerase. These genes were ligated into LuxR expression vectors and used for a final OFF/ON round of selection. The LuxR expression plasmids obtained from colonies surviving Round 3 selection were transformed with pluxGFPuv and screened for their ability to activate gene expression with 100 nM C10HSL. Eighty of the most fluorescent clones were transferred to a 96-well plate and grown overnight. These clones were further screened for the inability to activate GFPuv expression with 3OC6HSL on solid media.
We sequenced 12 of the 80 fluorescent clones and found four unique LuxR variants that exhibited the desired retention of C10HSL response and loss of 3OC6HSL response. All had the substitution of an arginine at position 67 for methionine, along with up to five additional new mutations. One, LuxR-G2E-R67M, has only the R67M mutation in addition to the three amino acid substitutions of its parent, LuxR-G2E (Table 4.1). All four unique variants exhibited identical responses with 3OC6HSL, C10HSL, C6HSL and C8HSL, indicating that the substitutions other than R67M are likely neutral (data not shown). We therefore focused on LuxR-G2E-R67M for all further characterization.

Table 4.1. Nucleotide and amino acid changes in luxR/LuxR and lasR/LasR mutants

<table>
<thead>
<tr>
<th>luxR plasmid</th>
<th>luxR gene base substitution</th>
<th>LuxR protein amino acid substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parent (2nd generation C8HSL-responder)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLuxR-G2E</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>T162→A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td><strong>Specificity mutant from selection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLuxR-G2E-R67M</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>T162→A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G200→T</td>
<td>Arg67→Met</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td></td>
<td>T498→C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T510→C</td>
<td></td>
</tr>
<tr>
<td><strong>R67M mutation in wild-type LuxR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLuxR-R67M</td>
<td>G200→T</td>
<td>Arg67→Met</td>
</tr>
<tr>
<td><strong>R61M mutation in wild-type LasR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLasR-R61M</td>
<td>GCT181→ATG</td>
<td>Arg61→Met</td>
</tr>
</tbody>
</table>

The gfpuv gene expression bioassay was used to quantitatively compare the influence of 3OC6HSL and C10HSL on transcriptional activation by wild-type LuxR, LuxR-G2E and LuxR-G2E-R67M (Fig. 4.2). LuxR and LuxR-G2E exhibited similar responses to 3OC6HSL; gene activation by LuxR-G2E-R67M, however, was barely above background even at concentrations as high as 10 μM 3OC6HSL. Responses to C10HSL showed the opposite trend: wild-type LuxR showed no response, while LuxR-G2E and LuxR-G2E-R67M showed similarly high levels of fluorescent protein production.

We then characterized the ability of LuxR-G2E-R67M to activate gene expression with four acyl-HSLs not used during directed evolution: C6HSL, C8HSL, dodecanoyl-homoserine lactone (C12HSL) and 3-oxo-dodecanoyl-homoserine lactone (3OC12HSL) (Fig. 4.2). LuxR-G2E-R67M has retained the increased ability to respond to straight-chained acyl-HSLs C6HSL, C8HSL and C12HSL shown by LuxR-G2E; its response to 3OC12HSL, however, was more than 50-fold weaker than LuxR-G2E’s. Thus, it appears that the R67M mutation disrupts interaction with the 3-oxo group.
Figure 4.2. Activation of *gfpuv* transcription with 3OC6HSL, C10HSL, C6HSL, C8HSL, C12HSL and 3OC12HSL by wild-type LuxR (▲), LuxR-G2E (■) and LuxR-G2E-R67M (◊). Shown are units of fluorescence due to GFPuv production in *E. coli* containing pluxGFPuv and pLuxR, pLuxR-G2E or pLuxR-G2E-R67M. The data are normalized to optical density and corrected by subtracting background fluorescence from a control strain containing pluxGFPuv and pPROLar.A122.
C.4. Effect of the R67M substitution in wild-type LuxR and LuxR homolog LasR

To further assess the role of the R67M mutation in 3-oxo recognition, we introduced it into wild-type LuxR and compared LuxR-R67M’s ability to activate gene expression with 3OC6HSL and its straight-chain counterpart, C6HSL. Fig. 4.4 shows that introducing a methionine at position 67 of LuxR eliminates its ability to respond to 3OC6HSL but has little effect on response to C6HSL. Thus, the R67M substitution appears not to significantly affect binding to straight-chain acyl-HSLs.

**Figure 4.3.** Activation of *gfpuv* transcription with 3OC6HSL and C6HSL by wild-type LuxR (▲) and LuxR-R67M (□). Shown are units of fluorescence due to GFPuv production in *E. coli* containing pluxGFPuv and pLuxR or pLuxR-R67M. The data are normalized to optical density and corrected by subtracting background fluorescence from a control strain containing pluxGFPuv and pPROLar.A122.
LasR is a LuxR homolog from *Pseudomonas aeruginosa* that responds to 3OC12HSL. It has only 20% identity to LuxR in the N-terminal signal-binding domain. LasR contains an arginine at residue 61, equivalent to R67 in LuxR. Because LasR also activates gene expression at the *lux* promoter [27], we used the same *gfpuv* gene expression bioassay to characterize its ability to activate gene expression in the presence of acyl-HSLs. If R67M in LuxR eliminates an interaction with the 3-oxo group, then the R61M substitution in LasR might reduce LasR’s ability to respond to 3OC12HSL. Figure 4.4 shows that LasR-R61M in fact has decreased response to 3OC12HSL, while its ability to respond to C12HSL is largely unaffected.

![Figure 4.4](image.png)

**Figure 4.4.** Activation of *gfpuv* transcription with 3OC12HSL and C12HSL by wild-type LasR (▲) and LasR-R61M (□). Shown are units of fluorescence due to GFPuv production in *E. coli* containing pluxGFPuv and pLasR or pLasR-R61M. The data are normalized to optical density and corrected by subtracting background fluorescence from a control strain containing pluxGFPuv and pPROLa.R.A122.
C.5. Gel mobility shift assays of acyl-HSL stimulated DNA binding

We used gel mobility shift assays to measure the ability of different acyl-HSLs to stimulate DNA binding by purified LuxR, LuxR-G2E and LuxR-G2E-R67M (Fig. 4.5). In the presence of 3OC6HSL, LuxR binds a 20 base pair (bp) pseudopalindromic DNA target known as the lux box, which is centered 42.5 bp upstream of the transcriptional start site in the lux promoter [28, 29]. Wild-type LuxR was able to bind a DNA probe containing the lux promoter in the presence of 3OC6HSL but not C10HSL, while LuxR-G2E could bind the same probe with 3OC6HSL or C10HSL (Fig. 4.5a). LuxR-G2E-R67M bound the DNA only in the presence of C10HSL. A titration of 3OC6HSL in the presence of 3.5 nM LuxR revealed that 100 nM of L-3OC6HSL was required to shift 50% of the specific probe (Fig. 4.5b), in agreement with published results [29]. Approximately 2 μM of L-3OC6HSL was required to shift 50% of the DNA with LuxR-G2E under similar conditions. LuxR-G2E-R67M was not observed to shift the DNA with up to 12.5 μM L-3OC6HSL. Wild-type LuxR, on the other hand, did not shift with the highest concentration of C10HSL, 12.5 μM. LuxR-G2E required approximately 1 μM C10HSL to shift 50% of the probe and LuxR-G2E-R67M achieved 50% binding with approximately 10 μM C10HSL. Hill coefficients determined for each interaction where binding was detected range from 0.7 to 1.0 (data not shown), similar to the value of 0.9 for wild-type LuxR with 3OC6HSL published by Urbanowski et al. [29].
Figure 4.5. Gel mobility shift assays of LuxR, LuxR-G2E and LuxR-G2E-R67M in the presence of 3OC6HSL and C10HSL. a. Gel shift assay of binding to lux promoter DNA. All lanes contain approximately 1 fmol of DNA and 3.5 nM LuxR, LuxR-G2E or LuxR-G2E-R67M. Acyl-HSL was added to a final concentration of 0.01 μM, 0.05 μM, 0.10μM, 0.64 μM, 1.26 μM, 6.26μM and 12.5 μM to each lane from left to right. b. Data for 3OC6HSL are represented by closed diamonds (♦), and C10HSL by open squares (□). The amount of the lux promoter probe DNA shifted was determined at each concentration of acyl-HSL and averaged. Error bars represent standard deviation of the mean based on at least three measurements.

Western immunoblots further demonstrate the new acyl-HSL specificity of LuxR-G2E-R67M. While wild-type LuxR accumulates only in the presence of its cognate signal, 3OC6HSL, and LuxR-G2E accumulates with either 3OC6HSL or C10HSL, an increase in LuxR-G2E-R67M concentration was detected with the addition of C10HSL, but not 3OC6HSL, reflecting the new signal-binding specificity (Fig. 4.6).
Figure 4.6. Protein accumulation is dependent upon acyl-HSL binding. Shown is a representative Western immunoblot of c-myc tagged LuxR, LuxR-G2E and LuxR-G2E-R67M in the absence of acyl-HSL or with 1 μM 3OC6HSL or C10HSL. Western immunoblots were performed as described [24].


A solid-phase assay with two fluorescent cell strains was used to determine whether any crosstalk occurred between LuxR and LuxR-G2E-R67M when cells were spread on plates containing 50 nM 3OC6HSL, 200 nM C10HSL or both. Two strains, one containing gfpuv under the control of LuxR and the other gfpuv under the control of LuxR-G2E-R67M, were plated on separate sides of a Petri dish. In the presence of 3OC6HSL, GFPuv production was observed as a result of LuxR-mediated gene activation (Fig. 4.7a). No fluorescence was observed from the second strain containing LuxR-G2E-R67M. Only LuxR-G2E-R67M-mediated GFPuv expression was observed with C10HSL (Fig. 4.7b). In the presence of 3OC6HSL and C10HSL, both sides of the
plates fluoresced (Fig. 4.7c). This finding indicates that LuxR and LuxR-G2E-R67M are sufficiently specific for their respective acyl-HSL signals to be used in synthetic genetic circuits requiring more than one acyl-HSL response system.

**Figure 4.7.** Solid-phase assays show minimal crosstalk between LuxR and LuxR-G2E-R67M. Cells on the left contain pLuxR and pluxGFPuv; cells on the right contain pLuxR-G2E-R67M and pluxGFPuv. a. Plate containing 3OC6HSL shows fluorescence only in LuxR-containing cells. b. Plate containing C10HSL shows fluorescence only in LuxR-G2E-R67M-containing cells. c. Plate containing 3OC6HSL and C10HSL shows fluorescence from both LuxR and LuxR-G2E-R67M.
D. Discussion

To evolve a new signaling specificity, the selection scheme must identify variants that respond to the new chemical signal (positive selection) but also have decreased response to the cognate signal(s) (negative selection). Variants of the Tet repressor with altered specificity were identified after multiple rounds of positive screening produced several variants capable of recognizing a new effector molecule [30]. Subsequent testing of the variants by screening with different effectors identified a mutant that no longer responded to tetracycline or several tetracycline analogs. Screening and selection systems that directly incorporate both positive and negative selection have been used to modify the specificities of enzymes, including Cre recombinase [31, 32], tRNA synthetases [33] and the endopeptidase OmpT [34]. Both the tRNA synthetase and Cre recombinase experiments used production of a reporter protein (GFP, RFP or β-galactosidase) for rounds of positive selection. The tRNA synthetase experiments relied on a lack of reporter expression during rounds of negative screening. In contrast, the absence of Cre activity was determined by a second reporter gene that was disrupted by site-specific recombination, in a negative screening step. The dual selection system used here also requires a specific output (as opposed to a lack thereof) for both positive and negative rounds of selection.

Using the dual selection system of Fig. 4.2, we discovered a single arginine to methionine substitution at residue 67 that decreased LuxR-G2E’s response to the LuxR cognate signal, 3OC6HSL while maintaining its acquired response to C10HSL. Changes in
specificity can be estimated by comparing the ratio of 3OC6HSL and C10HSL concentrations required to attain half-maximal levels of gene activation for the different LuxR variants. (In cases where no activation was observed, the highest concentration of acyl-HSL measured was used to obtain a lower bound for the specificity change.) A 100-fold change in specificity was attained by LuxR-G2E through its increase in C10HSL sensitivity. An additional change in specificity of at least 500-fold was acquired upon going from LuxR-G2E to LuxR-G2E-R67M. Overall, a more than 50,000-fold change in specificity of the acyl-HSL response was obtained by evolving wild-type LuxR into LuxR-G2E-R67M. Introducing the R67M mutation into wild-type LuxR generated a 1000-fold change in the ratio of 3OC6HSL to C6HSL response, while the R61M mutation in LasR yielded a more conservative 50-fold change in the ratio of 3OC12HSL to C12HSL response.

In LuxR, LuxR-G2E and LasR, this R→M substitution decreases binding to acyl-HSLs with the 3-oxo group, without affecting the recognition of straight-chain acyl-HSLs. Positive selection for response to the new straight-chain C10HSL signal would not have uncovered this specificity-enhancing mutation. A conserved methionine occurs at this position in a group of C8HSL-responsive homologs, including CepR, from the β-Proteobacteria. Thus this particular amino acid substitution, identified during laboratory evolution, was previously discovered during the natural evolution of these transcriptional activators to respond to different signaling molecules. Due to the low sequence identity between LuxR and its homologs, however, identification of this residue
as a key determinant of acyl-HSL specificity would have been difficult from sequence comparisons alone. The location of this residue in the crystal structure of TraR, the LuxR homolog from *Agrobacterium tumefaciens* and the only LuxR homolog for which a structure is available, indicates that it makes van der Waals contacts with the acyl-HSL but does not make direct contact with the 3-oxo moiety [35]. Furthermore, an attempt to target 3-oxo binding by TraR based on structural information did not yield any variants with new specificities[36]. A recent homology model of LuxR that suggests that the acyl side chain is flipped in the LuxR/acyl-HSL complex, as compared to the TraR/acyl-HSL complex, also failed to implicate residue 67 in 3-oxo recognition [37].

Gel mobility shift assays demonstrated that LuxR and the evolved variants bind DNA only in the presence of the specific acyl-HSLs that activate gene expression. A 250-fold change in the ratio of 3OC6HSL and C10HSL concentrations required to shift 50% of the DNA probe was observed with LuxR and LuxR-G2E. While LuxR and LuxR-G2E have similar dose-response curves with 3OC6HSL, LuxR-G2E required approximately ten-fold more signal to achieve 50% binding. LuxR-G2E-R67M also has a lower affinity for C10HSL than its parent, LuxR-G2E, although the activation by these two proteins is very similar. These results show that a decrease in signal binding *in vitro* does not necessarily yield a decrease in gene activation *in vivo* and supports the use of selection systems that directly measure signal-dependent gene expression, rather than DNA or signal binding, to generate variants with the precise output required for a given application.
Switching from a protein capable of recognizing 3OC6HSL but not C10HSL (LuxR) to one that recognizes C10HSL but no longer responds to 3OC6HSL (LuxR-G2E-R67M) was achieved via a series of single mutations. The three amino acid substitutions in LuxR-G2E each increase LuxR’s response to C8HSL and have additive effects when recombined by DNA shuffling [24]. Thus, these three mutations could accumulate in any order under selective pressure for increased response to the straight-chain acyl-HSL. The R67M mutation, on the other hand, could only be detected as beneficial after significant increases in straight-chain acyl-HSL response had been achieved, because LuxR does not respond to C10HSL.

The directed evolution of a LuxR variant with a new specificity therefore involved the acquisition of ‘promiscuous’ functions [24] and subsequent ‘respecialization.’ It has been postulated that broad-specificity proteins are evolutionary intermediates between more specialized states and that they are important for the evolution of proteins with new functions [23, 38]. These new functions can provide selective advantages to organisms as they adapt to changing environments. Results from many directed evolution experiments indicate that a single mutation is much more likely to confer a new function (e.g., the ability to recognize a new signal or substrate) without abolishing the old function [32, 39-41]. If a promiscuous interaction adversely affects the fitness of the organism, negative selective pressure will lead to the accumulation of mutations that minimize the deleterious interaction and lead to an increasingly specialized protein. We have emulated this process with LuxR and its interactions with 3OC6HSL.
This work has also established the use of the dual selection system for the directed evolution of a transcriptional activator and suggests it can serve generally for engineering transcriptional regulators with specific ON/OFF responses. The use of evolved LuxR variants in designer genetic circuits has been established [14], and will likely increase as new variants are engineered. Figure 4.7 illustrates the lack of crosstalk between wild-type LuxR and LuxR-G2E-R67M. This mutant will be a useful component for intercellular communication circuits that require multiple chemical ‘wires’ between cell strains. Its evolution demonstrates that the acyl-HSL binding specificity of LuxR can be tuned using the ON/OFF selection to suit a particular application. Engineers targeting applications that involve engineered cell-cell communication, including biofabrication, tissue engineering and targeted gene therapies, may find these proteins to be valuable additions to their ‘toolbox’ of circuit components.

E. Materials and Methods

E.1. Bacterial strains, media and growth conditions

The bacterial strains used in this study are Escherichia coli strains DH5α (F- φ80lacZΔM15 Δ(lacZYA-argF) U169 deoR recA1 endA1 hsdR17(rK-, mK+) phoA supE44 λ- thi-1 gyrA96 relA1 and BL21 (F- dcm+ Hte ompT hsdS(λB- mB-) gal endA Tet'). E. coli strains were cultured at 37°C in LB medium or on LB agar plates. Acyl-
HSLs used in these studies were: 3-oxohexanoyl-DL-homoserine lactone (3OC6HSL; Sigma Aldrich, St. Louis, MO), hexanoyl-DL-homoserine lactone (C6HSL; Fluka, St. Louis, MO), octanoyl-DL-homoserine lactone (C8HSL; Fluka), decanoyl-DL-homoserine lactone (C10HSL, Fluka), dodecanoyl-DL-homoserine lactone (C12HSL; Fluka) and 3-oxododecanoyl-L-homoserine lactone (3OC12HSL; Quorum Science, Corville, IA). Antibiotics were added at the following concentrations to maintain the plasmids: 100 μg/mL carbenicillin, 20 μg/mL kanamycin, and 100 μg/mL chloramphenicol. During selection experiments antibiotics were added as follows: 100 μg/mL carbenicillin, 50 μg/mL kanamycin, and 150 μg/mL chloramphenicol.

E.2. Plasmid construction

The LuxR expression vector, pLuxR, and the signal response plasmid, pluxGFPuv, have been described previously[24]. Plasmids pluxCAT and pluxBLIP are described in Chapter 3.

pLuxR(SalI) is similar to pLuxR except that the SalI endonuclease site upstream of the P_{lac/ara-1} promoter was removed and new a SalI endonuclease site was engineered into luxR at nucleotide positions 493-498 (amino acid positions 164-165) such that no amino acid changes were introduced. pluxRSalI was constructed by synthetic overlap extension PCR.
E.3. Library construction and selection

Error-prone PCR reactions were performed using AmpliTaq DNA polymerase (Applied Biosystems, Foster City, CA) and 300 μM MnCl₂ to increase the mutation rate as described [42]. The primers 5-luxRv2 (5-AGAGGAGAAAGGTACCCATGAAAAACA-3) and LuxR(SalI)-r2 (5-GCTATATTTATTTTCGATAATTGTCGACTAGAGAAGGAAC-3) were used to amplify the first 500 bp of the luxR-G2E gene using pLuxR-G2E as the template. The library was constructed by ligating KpnI and SalI digested pLuxRSalI with the products of error-prone PCR using T4 DNA ligase (Invitrogen, Carlsbad, CA). For the first ON round of selection (Round 1-ON), the ligation mixtures were transformed into competent DH5α cells harboring pluxCAT (DH5α(pluxCAT)) and plated onto LB agar plates containing 50 μg/mL kanamycin, 150 μg/mL chloramphenicol and 100 nM C10HSL. Similar plates were used for subsequent ON rounds. Plates used for OFF rounds contained 50 μg/mL kanamycin, 100 μg/mL carbenicillin and 100 nM 3OC6HSL. For each round, cells were also plated with kanamycin and carbenicillin to estimate the library size by quantifying the number of colonies that survive under non-selective conditions. For both ON and OFF selections, LB agar plates were incubated at 37 °C for 14 hours and then harvested with LB. The plasmid DNA was recovered via miniprep (QIAGEN, Valencia, CA), digested with Scal to inactivate pluxCAT or pluxBLIP, purified by spin column (Zymo Research, Orange, CA), and used to transform the competent cells for the next round. The plasmid DNA recovered from Round 2-ON was
used as the template for recloning to remove any false positives in which changes had occurred outside *lux-R-G2E*. The *lux-R-G2E* alleles were amplified using Pfu Turbo polymerase (Stratagene) and treated with *DpnI*. The PCR products were digested and ligated into pLuxRSalI (as above) and transformed into competent DH5α cells containing pluxBLIP and subjected to a third round of OFF selection (Round 3-OFF). The cells were harvested and a final ON round of selection (Round 3-ON) was performed as described above. The DNA purified from Round 3-ON was digested with *SacI* and used with pluxGFPuv to transform DH5α. The transformation was plated onto LB agar containing kanamycin, chloramphenicol and 100 nM C10HSL. Fluorescent colonies were identified as described previously [24] and grown in 1 mL cultures in deep-well microplates overnight at 37 °C.

### E.4. Characterization of LuxR variants

Quantitative characterization of LuxR-mediated gene expression of GFPuv was performed as reported previously [24].

LuxR, LuxR-G2E and LuxR-G2E-R67M were purified as described [29] with slight modification: all buffers contained a lower concentration of 0.05% Tween-20. LuxR-G2E-R67M was purified in the presence of 5 μM DL-C10HSL. Gel shift experiments were also performed similarly to those described by Urbanowski *et al*. [29]. The DNA probes were PCR products 173 bp in length from amplification of pluxGFPuv with 5-
pluxI and 3-pluxI [24]. Probes were generated by end-labeling the PCR products using $^{\gamma}$-32-P-ATP plus T4 polynucreotide kinase. Protein-DNA binding reactions contained approximately 1 fmol of the DNA probe in a final volume of 16 μL of DNA binding buffer (20 mM Tris-HCL, [pH 7.4], 50 mM KCl, 1 mM EDTA, 1 mM dithiothreitol, 100 μg bovine serum albumin/mL, and 5% ficoll 400) with 100 μg/mL poly(dI-dC)·poly(dI-dC) (GE Healthcare) competitor DNA. Purified protein and acyl-HSL were added as indicated and incubated for 25 min at 25 °C. The reaction mixtures were run on a native 5% Tris-glycine-EDTA gel at 4 °C. Following electrophoresis, the probes were detected using Kodak BioMax XAR Film (Rochester, NY) and quantified from the scanned image via densitometry analysis using AlphaEaseFc image analysis software (Alpha Innotech).

**E.5. Solid-phase fluorescence assays**

For the solid-phase experiment, separate cultures of *E. coli* DH5α containing either pLuxR and pluxGFPuv or pLuxR-G2E-R67M and pluxGFPuv were grown to stationary phase, diluted four-fold and blotted on top of LB-agar plates with a sterile sponge. Plates were incubated at 37 °C. Fluorescence images were captured with a digital camera while the plates were illuminated by a 360 nm ultraviolet light source.
F. References


CHAPTER 5

ON/OFF Selection for LuxR Variants with

Altered DNA-binding Specificities
A. Abstract

Variants of the acyl-homoserine lactone-dependent transcriptional activator LuxR with altered DNA-binding specificities were identified using a directed evolution scheme incorporating both positive (ON) and negative (OFF) selection. LuxR variants that activate gene expression at new promoters are powerful additions to the regulatory components available to genetic circuit engineers. Currently, the use of more than one LuxR homolog in a single cell is limited by the crosstalk between these systems at the level of DNA-binding specificity. To identify LuxR variants that activate gene expression at a mutant promoter, \( \text{P}_{\text{luxG5A}} \), a library of luxR variants was selected for antibiotic resistance due to gene activation. Surviving clones were selected for decreased activation at the wild-type lux promoter. Two variants, LuxR-R212H and R212C/Q232H, were identified and in vivo bioassays showed that the promoter specificity of LuxR-R212C/Q232H was broadened to include \( \text{P}_{\text{luxG5A}} \) as well as \( \text{P}_{\text{luxI}} \) while LuxR-R212H showed a 10-fold preference for \( \text{P}_{\text{luxG5A}} \), a 50,000-fold change in promoter specificity was achieved by the R212H mutation. The R212C and Q232H mutations were shown to be synergistic; the R212C mutation likely confers the desired change in DNA-binding affinity but destabilizes the protein, and the Q232H mutation appears to compensate for this defect. Selection of a second-generation library did not identify mutants with increased specificity for \( \text{P}_{\text{luxG5A}} \), indicating new library design strategies are required to engineer LuxR variants with inverted promoter specificity.
B. Introduction

LuxR is a 28 kDa acyl-homoserine lactone (acyl-HSL)-dependent transcriptional activator from *Vibrio fischeri* [1-3]. LuxR recognizes a diffusible, quorum-sensing signal molecule, 3-oxo-hexanoyl-homoserine lactone (3OC6HSL) [4, 5], and activates gene expression at the *lux* promoter (P\textsubscript{luxI}) by recruiting RNA polymerase [6, 7]. Gene activation by LuxR requires a 20 base pair (bp) operator site within P\textsubscript{luxI} known as the *lux* box (Fig 5.1a,b) [6, 8]. The *lux* box is centered 42.5 bp upstream of the P\textsubscript{luxI} transcriptional start site and its position has been shown to be essential for gene activation [6]. Sequence analyses have predicted that LuxR binds to its DNA target via an α-helix-turn-α-helix (HTH) motif between amino acid residues 200 and 224 [9]. The HTH motif is a common DNA-binding fold found in a myriad of bacterial activators and repressors [10, 11].

Alanine scanning mutagenesis of the C-terminal domain of LuxR was used to identify residues involved in DNA binding and activation [9, 12]. Several mutants unable to bind DNA or activate transcription were identified, including alanine substitutions at positions 191, 193, 212, 217, 225, 229, 230, 238 and 240. Alanine substitutions at positions 201 and 206 affect transcriptional activation but not DNA binding.
To date, more than 50 LuxR homologs have been identified within the Gram-negative phylum *Proteobacteria* [13]. Each LuxR homolog modulates gene expression following acyl-homoserine lactone (acyl-HSL) binding. All acyl-HSL signals contain a homoserine lactone ring; varied acyl side chain moieties allow species-specific signal recognition. While LuxR homologs have evolved to specifically respond to different acyl-HSL
signals, homologs from different species have been shown to interchangeably activate gene expression at other lux-type promoters. Both LuxR and LasR, a homolog from *Pseudomonas aeruginosa*, can activate gene expression at the other’s target promoter in *Escherichia coli* [14]. The lux-type boxes identified from a range of species show significant sequence diversity (Fig. 5.1a). However, elements within these sites, including highly conserved nucleotides at positions 3, 4, 5, -3, -4, and -5, appear to be sufficient for DNA binding by most LuxR homologs. As LuxR and its homologs are purified, their specific DNA-binding specificities are being investigated [15, 16]. Precisely which sequences are recognized by LuxR and its homologs is not known.

LuxR and its homologs have been used in several synthetic genetic circuits to confer intercellular communication capabilities [17-20]. As the desired circuits become increasingly complex, the ability to use of more than one acyl-HSL response regulator in a single cell type will be advantageous. While it may be possible to identify LuxR homologs that do not show any crosstalk with respect to promoter specificity, an alternative approach is to use directed evolution to generate variants of LuxR that activate gene expression at a mutant promoter. A benefit of this approach is the production of a set of ‘standardized’ parts for engineering intercellular communications circuits.

We have previously identified a LuxR variant with a new acyl-HSL binding specificity (Chapter 4). LuxR variants with switched acyl-HSL specificities were targeted using a dual selection system that allows for ON/OFF selection. We sought to continue this work
by targeting the DNA-binding specificity of LuxR to identify variants that recognize a new DNA target and that have decreased affinities for the wild-type lux box. Directed evolution of the DNA-binding specificity of a Cre recombinase using positive screening alone and a combination of positive and negative screening showed that both positive and negative selective pressure were required to switch Cre’s DNA-binding specificity. We hypothesized that a similar approach, using our dual selection system, would allow us to identify LuxR variants with novel DNA-binding specificities. LuxR variants that show increased gene activation at a mutated lux box and decreased activation at the wild-type lux box were identified using our ON/OFF selection system. Both variants with broadened DNA-binding specificity and variants with a preference for the new target were identified.

C. Results

C.1. Dual selection for evolution of LuxR promoter specificity

We targeted the 5 and -5 positions of the lux box for generating a mutated DNA target. This position is highly conserved in lux-type boxes. In constructing the luxG5A box, we chose to maintain the palindrome and replace these residues in a conservative manner, purine for purine and pyrimidine for pyrimidine (Fig. 5.1b). A transition is less likely to cause large structural differences than either of the possible transversions. Liquid-phase bioassays measuring fluorescence due to LuxR-mediated GFPuv production showed
wild-type LuxR was unable to activate gene expression at a *lux* promoter containing the *luxG5A* box (*P_{luxG5A}* with up to 50 μM 3OC6HSL (Fig. 5.3).

To select for LuxR variants capable of recognizing the *luxG5A* box and activating gene expression at *P_{luxG5A}* we used a gain-of-function (ON) selection plasmid encoding chloramphenicol acetyl-transferase (CAT) under the control of *P_{luxG5A}* (Fig. 5.2). Selection of functional clones relies on LuxR-dependent activation of *cat* leading to chloramphenicol resistance. To select for mutants with decreased affinities for wild-type *P_{luxG5I}* we used a loss-of-function (OFF) selection plasmid encoding the β-lactamase inhibitory protein (Bli) under the control of *P_{luxI}* and a constitutively expressed β-lactamase (*bla*) gene (Fig. 5.2). Selection of clones which are unable to activate gene expression relies on the absence of LuxR-dependent activation of Bli, which inactivates Bla and confers carbenicillin sensitivity.

A library generated by random mutagenesis of the DNA encoding the C-terminal third of LuxR, amino acid residues 166-250, was generated by error-prone PCR. The C-terminal domain was targeted because it has been shown to contain the DNA-binding and activation domains [21-23]. Clones transformed with this library were selected for chloramphenicol resistance due to gene activation at *P_{luxG5A}* with 5 μM 3OC6HSL. The plasmid DNA was recovered from the surviving cells and subjected to digestion with *ScaI* to selectively inactivate the selection plasmid, *pluxG5ACAT*. In a second ON round of selection, several clones survived with a 10-fold lower concentration of 3OC6HSL
(500 nM) and reduced incubation time of 14 hours. All subsequent rounds of ON selection used this set of more stringent conditions.

**Figure 5.2.** Schematic of the dual selection system for the directed evolution of LuxR variants with altered DNA-binding specificity. During positive rounds of selection, LuxR-mediated activation at a lux promoter containing the mutated luxG5A box, PluxG5A, leads to chloramphenicol acetyl-transferase (CAT) expression and chloramphenicol resistance. Variants with decreased activation at the wild-type lux promoter are identified during rounds of negative selection by their ability to survive with carbenicillin. Unwanted gene activation by LuxR leads to carbenicillin sensitivity due to β-lactamase inhibitory protein (Blip) production and its inhibition of constitutively expressed β-lactamase (Bla) protein.
We transformed the *luxR* mutant-encoding plasmids recovered after Round 2-ON with the OFF selection plasmid, *pluxBLIP*, and selected colonies surviving with carbenicillin and 10 nM 3OC6HSL. Cells expressing wild-type LuxR with this concentration of acyl-HSL are not viable. DNA recovered from the surviving Round 2-ON cells was used for a third ON round of selection. At this point, approximately 25% of the library survived with chloramphenicol and 500 nM 3OC6HSL.

To ensure that the observed resistance was a result of changes within *luxR*, we used the DNA recovered from Round 3-ON and PCR-amplified the *luxR* genes with a high-fidelity polymerase. The amplified *luxR* genes were subjected to a final round of ON selection. At this point, more than two-thirds of the cells survived in the presence of chloramphenicol and 500 nM 3OC6HSL. To identify the best P$_{luxG5A}$-activating LuxR variants, *luxR* expression plasmids isolated from colonies surviving the final round of selection were transformed with a plasmid encoding GFPuv under the control of P$_{luxG5A}$, *pluxG5AGFPuv*. More than 60% of the colonies fluoresced with 500 nM 3OC6HSL. Fifteen strains that fluoresced with 100 nM 3OC6HSL were isolated.

DNA sequencing of the 15 clones identified two unique *luxR* mutants. All of the mutants contained a mutation at residue 212 of LuxR. Three of the genes identified contained two mutations: R212→C and Q232→H. The other 12 genes identified all
encode a single R212→H substitution. This result suggests that the 212 position is critical for recognition of the luxG5 box.

Because error-prone PCR cannot access all of the possible amino acids at a given position, we targeted position 212 for saturation mutagenesis. The NNN saturation library was transformed with pluxG5AGFPuv and plated with 100 nM 3OC6HSL. Ten of the most fluorescent clones were isolated and their luxR genes sequenced. Of the ten, seven encoded R212→H and three encoded R212→S.

C.2. Second-generation library selection

A second-generation library was generated by error-prone PCR of the C-terminal domain encoding region of luxR-R212H. The library was determined to have an average mutation rate of approximately 4.5 nucleotide substitutions per luxR-R212H gene. Selections were carried out as described above with concentrations of 100 nM 3OC6HSL added for ON rounds of selection and 50 nM 3OC6HSL for OFF rounds. More than 200,000 clones were selected under these conditions. No positive clones were identified.
C.3. Characterization of luxR mutants

To compare transcriptional activation at P_luxI and P_luxG5A by wild-type LuxR and the evolved LuxR variants, we used a liquid-phase bioassay and measured fluorescence due to GFPuv production with a range of concentrations of 3OC6HSL (Fig. 5.3). The two variants identified via ON/OFF selection, LuxR-R212C/Q232H and LuxR-R212H, showed similar dose-responses with P_luxG5A. Levels of LuxR-R212C/Q232H-mediated fluorescence at P_luxI and P_luxG5A were indistinguishable. LuxR-R212H showed a preference for P_luxG5A, requiring approximately ten-fold more 3OC6HSL to achieve similar levels of activation at P_luxI. LuxR-R212S also showed a subtle preference for P_luxG5A (~five-fold) but required a two-fold higher 3OC6HSL concentration to activate gene expression at both promoters compared to LuxR-R212H.
Figure 5.3. Activation of gfpuv transcription from $P_{luxI}$ (□) and $P_{luxG5A}$ (▲) by wild-type LuxR, LuxR-R212C/Q232H, LuxR-R212H and LuxR-R212S with 3OC6HSL. Shown are units of fluorescence due to GFPuv production in E. coli containing pluxGFPuv and pLuxR, pLuxR-R212C/Q232H, LuxR-R212H or pLuxR-R212S. The data are normalized to optical density and corrected by subtracting background fluorescence from a control strain containing pluxGFPuv and pPROLar.A122.
C.4. LuxR-R212C/Q232H activation at P\textsubscript{luxG5A} requires both mutations

To assess the role of the individual mutations in LuxR-R212C/Q232H, the single mutants R212C and Q232H were constructed by site-directed mutagenesis. Quantitative liquid-phase characterization showed that the LuxR-R212C single mutant does not activate gene expression at P\textsubscript{luxI} or P\textsubscript{luxG5A} with up to 50 μM 3OC6HSL (Fig. 5.4). LuxR-Q232H showed increased levels of activation at both the wild-type and the mutated lux promoter, however, activation at P\textsubscript{luxG5A} required more than 100-fold more 3OC6HSL than the double mutant, LuxR-R212C/Q232H. These two mutations must work synergistically to confer the increased response at P\textsubscript{luxG5A}.

**Figure 5.4.** Activation of gfp\textsubscript{uv} transcription from P\textsubscript{luxI} (□) and P\textsubscript{luxG5A} (▲) by single mutants LuxR-R212C and LuxR-Q232H with 3OC6HSL. Shown are units of fluorescence due to GFP\textsubscript{uv} production in E. coli containing pLuxR-R212C or pLuxR-Q232H. The data are normalized to optical density and corrected by subtracting background fluorescence from a control strain containing pluxGFP\textsubscript{uv} and pPROLar.A122.
C.5. Accumulation of LuxR variants in vivo

Each of the LuxR variants was tagged with a c-myc epitope and assayed for its abundance in vivo using Western immunoblots. In the absence of 3OC6HSL, minimal levels of protein were observed for LuxR and each of the variants (data not shown). Accumulation of wild-type LuxR, LuxR-R212H, LuxR-R212C/Q232H, LuxR-R212S and LuxR-Q232H increased with the addition of 3OC6HSL (Fig. 5.5). LuxR-R212C did not increase in abundance with the addition of 3OC6HSL. The concentrations of LuxR-R212H, LuxR-R212C/Q232H and LuxR-R212S were lower than wild-type LuxR. LuxR-R212H showed the largest decrease in protein expression, to $<10\%$ of wild-type levels. LuxR-R212C/Q232H and LuxR-R212S accumulate to approximately 50\% of wild-type levels. In contrast, the amount of LuxR-Q232H detected with 3OC6HSL was at more than five-fold higher than wild-type LuxR, indicating this mutation stabilizes LuxR.

![Figure 5.5](image)

**Figure 5.5.** Protein accumulation is dependent upon acyl-HSL binding. Shown is a representative Western immunoblot of c-myc tagged LuxR (WT), evolved DNA-binding variants and single-mutants with 10 μM 3OC6HSL. Minimal protein accumulation, similar to the amount shown for LuxR-R212C with 3OC6HSL, was observed for each of the proteins in the absence of acyl-HSL.
C.6. DNA-binding experiments

Wild-type LuxR, LuxR-R212C/Q232H, LuxR-R212H and LuxR-R212S were purified, and their ability to bind DNA in vitro was assayed using electromobility shift assays (EMSAs). As expected from the in vivo bioassay data, all of the variants showed lower affinities for DNA probes containing either the \( P_{\text{luxI}} \) or \( P_{\text{luxG5A}} \) promoter than wild-type LuxR with \( P_{\text{luxI}} \). To observe DNA binding by the evolved variants, micromolar protein concentrations were required. Due to the elevated protein concentrations it was necessary to increase the amount of nonspecific DNA used in these studies to ensure the observed shifts were due to specific DNA-protein interactions. Under these stringent conditions, wild-type LuxR has a \( K_D \) of approximately 75 nM (Fig. 5.6), approximately 150-fold higher than the previously reported value of 0.5 nM [3]. LuxR did not bind to a DNA target containing the \( \text{luxG5A} \) box at concentrations up to 2 \( \mu \)M. The three DNA-binding variants showed similar affinities for the \( \text{luxG5A} \) box. Approximately 15% of the probe was bound with 2 \( \mu \)M protein (Fig. 5.6). Further increases in protein concentration led to excessive levels of non-specific binding and precipitation of the protein. The \( K_D \) of the mutants for the new site was estimated to be more than 1000-fold higher than LuxR’s dissociation constant with its cognate DNA target. All three variants also showed a preference for the \( \text{luxG5A} \) box. LuxR-R212C/Q232H did not detectably bind the wild-type \( \text{lux} \) box and LuxR-R212H reproducibly bound 1-2% of the wild-type probe at the highest protein concentration used in this study. Approximately 6% of the wild-type DNA probe
Figure 5.6. Electromobility shift assays of LuxR and the evolved variants with $P_{\text{luxI}}$ and $P_{\text{luxG5A}}$. Gel shift assays are shown on the left and quantitative plots of the fraction of shifted DNA versus protein concentration are shown on the right. All lanes contained approximately 5 fmol of a 176 bp DNA fragment containing either $P_{\text{luxI}}$ or $P_{\text{luxG5A}}$, the indicated concentration of LuxR, LuxR-R212C/Q232H, LuxR-R212H and LuxR-R212S and 25 μM DL-3OC6HSL.
was bound by 2 μM LuxR-R212S. With only a two-fold preference for the mutant \textit{luxG5A} box, LuxR-R212S has the most broadened DNA-binding specificity.

\textbf{C.7. \textit{RhlR}-mediated gene activation at \textit{P}lux\textit{I} and \textit{P}lux\textit{G5A}}

An alignment of LuxR homologs showed that a LuxR homolog from \textit{P. aeruginosa}, \textit{RhlR}, has a serine residue at the amino acid position that aligns with 212 in LuxR. \textit{RhlR}-mediated GFPuv production from \textit{P}lux\textit{I} and \textit{P}lux\textit{G5A} was measured using the same bioassay as LuxR. As shown in Figure 5.7, \textit{RhlR} activated gene expression at both promoters. Approximately five-fold higher concentrations of \textit{RhlR}’s cognate acyl-HSL, butanoyl-homoserine lactone (C4HSL), were required to achieve similar levels of gene activation at the wild-type \textit{lux} promoter compared to \textit{P}lux\textit{G5A}. This result indicates that a serine residue at this position may confer a broadening of DNA-binding specificity in both \textit{RhlR} and LuxR.
Figure 5.7. RhlR-mediated GFPuv gene expression at $P_{\text{luxI}}$ and $P_{\text{luxG5A}}$. Shown are units of fluorescence due to GFPuv production in *E. coli* containing pRhlR with pluxGFPuv or plux$_{\text{G5A}}$GFPuv. The data are normalized to optical density and corrected by subtracting background fluorescence from a control strain containing pluxGFPuv and pPROLar.A122.
D. Discussion

Examples of HTH engineering for new DNA-binding specificity in the literature are scarce, and the examples that are available have relied on structural information regarding the interactions between the protein and its target DNA site [24-34]. A crystal structure of LuxR is not available, and although LuxR and its homologs share higher sequence identity in their DNA-binding domains than their acyl-HSL binding domains, the sequence similarity between the C-terminal domains of LuxR and TraR, the only LuxR homolog for which a complete crystal structure is available, is very low (<20%). This is likely due to the fact that tra box contains an internal deletion on each side of the inverted repeat, as compared to the lux box (Fig. 5.1a) [14].

We generated a library of luxR mutants via error-prone PCR and used an ON/OFF selection system to identify LuxR variants that both activate gene expression at a promoter containing a mutated lux box, P_{luxG5A}, and that show decreased gene activation at the wild-type lux promoter. The dual selection system was employed to prevent the identification of LuxR variants with increased gene activation at both P_{luxI} and P_{luxG5A}. Our selection identified two variants, LuxR-R212H and LuxR-R212C/Q232H. The conditions used for OFF rounds of selection were not sufficiently stringent to require the variants identified to be specific for the P_{luxG5A} promoter, and both LuxR variants exhibit broadened promoter specificity. However, a 50-fold increase in 3OC6HSL concentration required to reach half-maximal gene expression at P_{luxI} was observed with LuxR-R212C/Q232H and a 500-fold increase was observed with LuxR-R212H. Therefore, the
dual selection system successfully identified variants with decreased gene activation at $P_{\text{luxI}}$. A third variant, LuxR-R212S, was identified by screening a library that randomized codon 212. These results indicate that residue 212 is a modulator of specific interactions with positions 5 and -5 of the $lux$ box. Mutating this residue to an alanine was previously shown to inactivate LuxR [9, 12] and an alignment of this residue with the published structure of TraR shows that the residue may make direct contact with the phosphate group between nucleotides 4 and 5 of the $lux$ box (Fig. 5.8). This residue has not been previously implicated in affecting the DNA-binding specificity of any LuxR homolog.

Lux-type boxes containing the A/T bp targeted in this work have only been identified in *Pseudomonas* strains. LasR has been shown to bind and activate gene expression at promoters containing either an A/T or G/C bp at this position, although it contains an alanine residue at position 212 [16, 35]. The presence of a serine residue in RhlR at the amino acid position that aligns with residue 212 in LuxR prompted us to test its ability to activate gene expression at both $P_{\text{luxI}}$ and $P_{\text{luxG5A}}$. As with LuxR-R212S, RhlR showed a preference for $P_{\text{luxG5A}}$, requiring approximately five-fold more 3OC6HSL to achieve similar levels of activation at $P_{\text{luxI}}$. The finding that RhlR and LasR have relaxed DNA-binding specificities with regards to positions 5, -5 is interesting in light of the fact that the presence of three acyl-HSL response regulators coexist in this bacterial species. Intuitively, one might expect three regulators to have higher DNA-binding specificity to avoid activating gene expression at the ‘wrong’ promoter. The ability to distinguish the different target promoters of the three LuxR homologs must arise from differences at
other nucleotide positions. A recent study showed that two of the LuxR homologs found in *P. aeruginosa*, LasR and QscR, identify their respective promoters by recognizing bp 8 and -8 [36].

**Figure 5.8.** Positions of residues in TraR corresponding to residues 212 and 232 in LuxR. The crystal structure of the LuxR homologue TraR (PDB 1L3L) has been determined [11]. The *tra* box DNA target and the C-terminal domains of two TraR peptides are shown. The two peptides comprising a TraR dimer are shown in green and blue. The *tra* box is shown as sticks (grey represents carbon, red represents nitrogen, blue represents oxygen and orange represents phosphorus). The amino acid residues which align with R212 and Q232 are shown space-filling spheres with coloring similar to that used for DNA. The phosphate group between positions 4 and 5 of the lux box is shown as space filling spheres. The glutamic acid (E196) residue located proximally to Q232 is also shown.
Changes in specificity can be estimated by comparing the ratio of 3OC6HSL concentrations required to attain half-maximal levels of gene activation at the $P_{\text{luxI}}$ and $P_{\text{luxG5A}}$ promoters for the different LuxR variants. A 5,000-fold change in specificity was attained by LuxR-R212C/Q232H. Both LuxR-R212H and LuxR-R212S showed a more than 50,000-fold change in promoter specificity.

The observation from the gel shift experiments that each of the evolved variants has a higher DNA-binding affinity for $P_{\text{luxG5A}}$ than $P_{\text{luxI}}$ was unexpected. Our bioassays showed that LuxR-R212C/Q232H activates gene expression equally well at either promoter but it did not bind $P_{\text{luxI}}$ in vitro. This discrepancy indicates that the conformational changes or interactions with RNA polymerase required for activation may also affect DNA binding. The finding that a truncated LuxR consisting only of the C-terminal domain required RNA polymerase to bind DNA demonstrated that LuxR’s interactions with RNA polymerase can influence its ability to bind the $lux$ box [23]. This finding suggests that selection or screening systems that directly measure the desired output, i.e., gene activation, may be better for optimizing genetic regulatory networks than assays targeting DNA binding directly.

The finding that the R212C single-mutant is nonfunctional and that the Q232H mutation does not confer altered promoter specificity on its own indicates that the R→C mutation promotes recognition of the $lux_{G5A}$ box, but that it is also deleterious with regards to
activating gene expression. The observation that LuxR-R212C does not accumulate in vivo indicates that its inability to activate gene expression is due to decreased protein stability or misfolding. LuxR-R212C/Q232H, LuxR-R212H, and LuxR-R212S were also shown to accumulate to lower protein concentrations. In general, mutations at position 212 negatively affect protein stability and/or folding, and positively affect P luxG5A recognition. The Q232H mutation likely compensates for the defect caused by the R212C mutation. The Q232H mutation could confer the observed increase in activation in one or more ways including increasing protein stability or solubility, stabilizing interactions with RNA polymerase, and increasing DNA affinity. The five-fold higher protein concentration observed with LuxR-Q232H, compared to wild type, suggests that the increase in gene activation observed with LuxR-Q232H is a result of increased protein stability. Alignment to the published structure of TraR suggests that the Q232H mutation may stabilize the dimeric form of the protein through an electrostatic interaction with glutamic acid 196 on the adjacent polypeptide chain (Fig. 5.8).

The majority of published HTH engineering experiments have used solutions already present in Nature, including swapping an entire DNA recognition helix or replacing all of the amino acids that contact a specific DNA target recognized by a close homolog. This strategy is not always successful [24], and does not generate DNA-binding proteins that recognize novel DNA sequences. Attempts at engineering HTH proteins that bind to previously unrecognized sequences have previously relied solely on libraries that randomize the amino acids shown to make contacts with the DNA. In most cases,
modest gains can be accomplished in this manner. New specificities, however, are rare. Our results suggest that DNA-binding proteins that recognize new DNA targets can be identified from libraries generated by random mutagenesis. This strategy is especially useful when structural information is unavailable. Furthermore, the identification of mutations at residues not targeted in site-saturation libraries, which have arisen adventitiously [37], are critical for binding the new target suggests that screening or selecting a library generated by random mutagenesis is a viable complementary approach for identifying DNA-binding mutants even when a crystal structure is available.

The identification of a LuxR variant with two synergistic mutations that increase activation at \( P_{luxG5A} \) indicates that multiple mutations may be required to switch DNA-binding specificity. We used our best first-generation mutant as parent and selected more than 200,000 clones transformed with a luxR-R212H library for additional improvements in promoter specificity. We did not identify any positive clones from these experiments. This suggests that either switching LuxR’s specificity from the G/C bp to an A/T bp at positions 5 and -5 is not chemically possible, or that a LuxR variant meeting the criteria determined by the selection conditions was not present in the second-generation library. The second, more plausible explanation indicates that a new library design is necessary. Alanine scanning mutagenesis of a 29 residue “miniature” DNA-binding protein demonstrated the importance of a well-packed hydrophobic core for DNA binding and specificity [38]. Therefore, making further gains in engineering the DNA-binding specificity with HTH motifs may require the generation of libraries that target a large
number of residues, including both those that interact directly with the DNA and those that form the core of the protein. The dual selection system is ideal for this use because libraries of more than $10^8$ unique clones could be selected for variants with new DNA-binding specificities. If, like the R212C mutation, residues that alter the DNA-binding specificity frequently destabilize the protein, an alternative library design strategy is to start with a stabilized parent [39]. We have previously identified LuxR variants with increased sensitivities to 3OC6HSL that exhibit much higher protein concentrations than wild-type LuxR in the presence of acyl-HSL. One of these stabilized variants may provide the necessary advantage required to identify new LuxR variants with switched DNA-binding specificities.
E. Materials and Methods

E.1. Bacterial strains, media and growth conditions

The bacterial strains used in this study are *Escherichia coli* strains DH5α (F φ80d lacZΔM15 Δ(lacZYA-argF) U169 deoR recA1 endA1 hsdR17(ris, mK+) phoA supE44 λ- thi-1 gyrA96 relA1 and BL21 (F dcm+ Hte ompT hsdS rB- mB-) gal endA Tet'). *E. coli* strains were cultured at 37 °C in LB medium or on LB agar plates. Antibiotics were added at the following concentrations to maintain the plasmids: 100 μg/mL carbenicillin, 20 μg/mL kanamycin, and 100 μg/mL chloramphenicol. During selection experiments antibiotics were added as follows: 100 μg/mL carbenicillin, 50 μg/mL kanamycin, and 150 μg/mL chloramphenicol. 3-Oxohexanoyl-DL-homoserine lactone (3OC6HSL; Sigma Aldrich, St. Louis, MO) stock solutions were prepared by dissolving appropriate amounts in ethyl acetate acidified with glacial acetic acid (0.01% v/v), and stored at -20 °C.

E.2. Plasmid construction

The positive selection and GFPuv reporter plasmids containing P_{luxG5A}, pluxG5ACAT pluxG5AGFPuv, were constructed from pluxGFPuv [40] and pluxCAT (described in Chapter 3) using standard site-directed mutagenesis techniques. The oligonucleotides used were: luxBoxB-f (5'-CTTAACATAAGCACCCTATAGGATCGGTATAGGTTTAC GCAAAGAAAATGG-3) and its reverse complement, luxBoxB-r.
E.3. Mutagenesis of luxR

Mutagenesis of the region of luxR encoding the C-terminal third of the protein was carried out using error-prone PCR [41] to yield a mutation frequency of approximately 2 nucleotide substitutions per luxR gene. The LuxRSalI-f (see above) and 3-LuxR [40] primers were used to amplify the desired region. The pool of mutagenized luxR genes was cloned into pLuxRSalI between its SalI and BamHI sites.

The forward and reverse primers used for saturation mutagenesis at position 212 were:
5-LuxRSat212 (5'-GGGATATTTCAAAAAATATTAGGCTGCAGTGAGNNNACTGTCACTTTCC-3), and 3-LuxRSat(int) (5'-CTGCAGCCTAATATTTTTGAAATATCCC-3). The 3' end of luxR containing the randomized 212 codon was amplified from pLuxSalI with LuxRSat212 and 3-LuxR+200 (5'-GTGAGCGAGGAAGCGGAATATCC-3).

The 5' end of the gene was amplified with 5-LuxR [40] and 3-LuxRSat(int). Overlap extension PCR using the PCR products from the two reactions described above, 5-LuxR and 3-LuxR was used to construct the entire luxR gene containing the randomized 212 codon. The library was ligated into pLuxRSalI following digestion with KpnI and BamHI.

LuxR single-mutants containing the R212C and Q232H mutations were constructed using standard site-directed methods.
E.4. Selection of LuxR variants with new DNA-binding properties

In Round 1-ON, competent cells harboring pluxG5ACAT were transformed with a pool of mutagenized pLuxRSalI and plated on LB-agar amended with 5 μM 3OC6HSL, 150 μg/mL chloramphenicol and 50 μg/mL kanamycin. Surviving colonies were harvested after 20 h at 37°C by overlaying LB medium. pLuxSalI plasmids were recovered from the selected cells via miniprep (QIAGEN) and endonuclease digestion of pluxG5ACAT with ScaI. The recovered DNA was transformed into DH5α cells harboring the pluxG5ACAT plasmid for Round 2-ON. In Round 2-ON (and all subsequent ON selections), the LB-agar contained 500 nM 3OC6HSL. Surviving colonies were harvested after 14 h at 37°C. The plasmid DNA recovered and pluxG5ACAT inactivated via digestion. The luxR expression plasmids recovered after Round 2-ON were transformed into DH5α cells harboring pluxWTBlip and plated with 10 nM 3OC6HSL, 100 μg/mL carbenicillin and 50 μg/mL kanamycin (Round 2-OFF). Cells surviving after 14 h at 37°C were harvested and treated similarly to after Round 1-ON and Round 2-ON. Re-cloning of the library following Round 3-ON was carried out with a proofreading polymerase (Pfu, Stratagene), 5-LuxR and 3-LuxR. PCR products were ligated into pLuxRSalI following digestion with KpnI and BamHI. pLuxRSalI plasmids recovered after the Round 4-ON were transformed into DH5α with pluxG5AGFPuv and plated with 100 nM 3OC6HSL, 100 μg/mL chloramphenicol and 50 μg/mL kanamycin. Eighty
colonies were picked and plated with a titration of 3OC6HSL. The 15 best colonies, based on fluorescence, were chosen for subsequent characterization.

Plasmids encoding the selected LuxR variants were isolated by transforming miniprep DNA into plasmid-free DH5α and plating onto LB-plates supplemented with kanamycin. The plasmids encoding the LuxR variants of interest were also transformed with pluxGFPuv so that their ability to activate gene expression at the wild-type lux promoter could be assessed.

E.5. Characterization of selected LuxR variants

Bioassays to quantitatively characterize LuxR-mediated gene expression of GFPuv from both P_{luxI} and P_{luxG5A} were performed as described [40]. Western blots to assay the abundance of the LuxR variants in vivo as described [40].

LuxR variants were purified as described [3]. Modifications to the published procedure include: the concentration of tween-20 in all buffers was lowered to 0.05% v/v.

Gel mobility shift assays were performed similarly to the methods described as described in Chapter 4. The DNA probes containing either the wt lux box or lux_{G5A} box were 173 bp in length and amplified from pluxGFPuv or plux_{G5A}GFPuv with 5-LuxI and 3-LuxI
Probes were generated by end-labeling the PCR products using $[\gamma^{32}\text{P}]-\text{ATP}$ plus T4 polynucleotide kinase. Protein-DNA binding reactions contained approximately 5 fmol of the DNA probe, 200 $\mu$g/mL poly(dI-dC)$\cdot$poly(dI-dC) and 250 $\mu$g/mL sonicated salmon sperm DNA in a final volume of 10 $\mu$L of DNA binding buffer (20 mM Tris-HCl, [pH 7.4], 50 mM KCl, 1 mM EDTA, 1 mM dithiothreitol, 100 $\mu$g bovine serum albumin/mL, and 5% ficoll 400). Purified protein and acyl-HSL were added as indicated and incubated for 25 min at 25 °C. The reaction mixtures were run on a native 5% Tris-glycine-EDTA gel at 4 °C. Following electrophoresis, the probes were detected using Kodak BioMax XAR Film (Rochester, NY) and quantified from the scanned image via densitometry analysis using AlphaEaseFc image analysis software (Alpha Innotech).
F. References


Conclusions
We have generated variants of the quorum sensing transcriptional regulator LuxR with new signal and DNA-binding specificities for use in synthetic genetic circuits. One of our hypersensitive variants has already been used as a component in a multicellular band-detect network that generates programmed pattern formation on a solid surface. We anticipate that our new LuxR variants will be the cornerstones of future efforts to construct more complex intercellular communications systems. Furthermore, laboratory evolution of specific signal- and DNA-binding interactions has provided insight into how these systems have evolved in Nature and the biochemical mechanisms involved in distinguishing between chemically similar targets.

Through the directed evolution of LuxR, we have identified several amino acid residues that modulate acyl-HSL and DNA-binding specificity that had not been identified previously. It will be interesting to reassess these mutations as more structural details become available. Would we have been able to identify these sites as specificity determinants if a crystal structure was available? While this scenario is unlikely, perhaps the structures of wild-type LuxR and variants from this work will provide some insight into how these mutations confer the observed change in LuxR-mediated gene activation.

The generation of transcriptional activators with new functions in the absence of structural information demonstrates that directed evolution is a valid design strategy for producing new regulatory components. These evolved variants add new functionalities to make a ‘standardized’ set of components that share many biochemical properties but respond to
different signals or regulate gene expression at different promoters with a range of sensitivities. The wide range of circuit parameters attainable with these new components will aid in the generation of increasingly complex synthetic genetic circuits.

The dual selection system, which successfully identified both acyl-HSL and DNA-binding LuxR variants, is a generally applicable technology and can easily be modified for the directed evolution of other transcription factors. The potential to select libraries of $10^8$ or more may enable, for the first time, the directed evolution of an entire genetic circuit by simultaneously randomizing multiple components and selecting for the desired output.
APPENDIX A

Plasmid, gene and protein sequences
A. Sequences of wild-type LuxR, LuxR homologs, evolved LuxR variants and the pLuxR expression vector

A.1. General information

LuxR (Genbank number AAD48473) is composed of 250 amino acids (750 bp). The N-terminal domain of the protein, residues 1 to ~160 residues, is involved in acyl-HSL binding. The C-terminal domain, residues ~180-250, is involved in DNA-binding and activation. The LuxR expression vector, pLuxR, was derived from the commercially available pPROLar.A122 plasmid and contains luxR under the control of a P_{lac/ara-1} hybrid promoter. This plasmid was used for all cloning procedures as well as for protein expression.

Two synonymous mutations, T498→C and T510→C, were used to introduce a SalI restriction site into the luxR gene between the N- and C-terminal domains so that each domain could be targeted independently. The pLuxR plasmid was also modified to remove a SalI site upstream of the P_{lac/ara-1} promoter. All luxR mutants presented in Chapters 4 and 5 contain the internal SalI restriction site. In random mutagenesis experiments targeting the acyl-HSL binding specificity of LuxR, error-prone PCR was used to amplify the region stretching from the KpnI site to the SalI site to include the first 167 amino acids of LuxR. In random mutagenesis experiments targeting the DNA-binding specificity of LuxR, error-prone PCR was used to amplify the region stretching from the SalI site to the BamHI site.
Figure A.1. lists the nucleotide sequence of full-length, wild-type luxR.

Figure A.2. lists the amino acid sequence of full length, wild-type LuxR.

Figure A.3. lists the nucleotide sequence of the pLuxR vector containing wild-type luxR.

Figure A.4. shows the plasmid map of the pLuxR vector.

Figure A.5. lists the nucleotide sequences of the luxR homologs used in this work. The wild-type luxR gene is substituted by the wild-type lasR gene in pLasR. The wild-type rhlR gene replaces luxR in pRhlR. A BamHI site within the rhlR gene prevented the use of this site when cloning the RhlR expression vector. A HindIII site directly upstream of the BamHI site in pPROLar.A122 is the 3’ restriction site in pRhlR.

Figure A.6. lists the amino acid sequence of full length, wild-type LasR and RhlR.

Table A.1. lists the sequences of the luxR mutants generated by random mutagenesis in Chapter 2.

Table A.2. lists the sequences of the luxR single-mutants constructed in Chapter 2.

Table A.3. lists nucleotide and amino acid substitutions of luxR/LuxR and lasR/LasR mutants described in Chapter 4.

Table A.4. lists nucleotide and amino acid substitutions of luxR/LuxR mutants described in Chapter 5.
Figure A.1. Nucleotide sequence of full-length, wild-type luxR.

1   ATGAAAAACA TAAATGCCGA CGACACATAC AGAATAATTA ATAAAATTAA
51  AGCTTGTAGA AGCAATAATG ATATTAATCA ATGCTTATCT GATATGACTA
101 AAATGGTACA TTGTAATAT TATTTACTCG CGATCATTTA TCCTCATTCT
151 ATGTTAAAAT CTGATATTTTC AATTCTAGAT AATTACCTTA AAAATGGGAG
201 GCAATATTAT GATGACGCTA ATTTAATAAA ATATGATCCT ATAGTAGATT
251 ATCTTACCTC CAATCATTCA CCAATTAATT GGAATATATT TGAAACAAAT
301 GCTGTAATAA AAAATCTCC AAATGTAATT AAAGAAGCGA AAACATCAGG
351 TTCTTATCAG GGGTTTAGTT TCCCTATTCA TACCGCTAAC AATGGCTTCG
401 GAATGCTTAG TTTTGCACTG TCAGAAAAG ACAACTATAT AGATAGTTTA
451 TTCTTACATG CTTTTGATGAA CAAATATTATA CTATCCTATTT TTCTAGGTGA
501 TAATTATCGA AAAATAATAA TAGCAAAATAA TAAATCAAAAC AAGCATTTAA
551 CCAAAAGAGA AAAGGAATGT TTAGCTGGG CATGCGAAGG AAAAGCTCT
601 TGGGATATTT CAAAATATTG AATGGCTGAGT AGGCTACTG TTCCTTCTCC
651 TTAAAAAGAT GCGCAATGA AACTCAATAC AACAAACCGC TGCCAAAGTA
701 TTTCTAAAGAC AATTTTAACA GGAGCAATTG ATTGCCCATAT CTTAAAATT

Figure A.2. Amino acid sequence of full-length, wild-type LuxR.

1    MKNINADDTY RIINKIKACR SNNDNQCLS DMTKMVHEY YLLAIYPHS
50   MVKSDISILD NYPKWXQYY DDANLKYDPI IVDYSNSHS PINWNIFENN
100  AVNKKSPNVI KEAKTSGLIT GFSFPIHTAN NGFGMSFAH SEKDNYIDSL
150  FLHACMNIPL IVPSLVNRYR KINIAANNSN NDLTKEKCE LEWACEGKSS
200  WDISKIGCS ERTVTFLHTN AQMKLNTNRR CQSISKAILT GAIDCPYFKN
Figure A.3. Nucleotide sequence of the pLuxR vector containing wild-type luxR.

```
1    GTCGACTCGA GCATAGCATT TTTATCCATA AGATTAGCGG ATCTAACCTT
51   TACAATTTTG AGGCCCTACA ATTATGATAG ATTTCAATTGT GAGCGGATAA
101  CAATTTCCACA CAGAATTCAT TAAAGAGGAG AAAGGTACCC ATGAAAAACA
151  TAAATGCCGA CGACACATAC AGAATAATTT ATAAATTTAA AGCTTGTAGA
201  AGCAATAATG ATATTAATCA ATGCTTATCT GATATGACTA AAATGGTACA
251  TGGTGAATAT TATTTACTCG CGATCATTTA TCCCTATCTG ATGGTTAAAT
301  TGTGACTATG ATATTAATCA ATGCTTATCT GATATGACTA AAATGGTACA
351  GAGATTTCCT GGAAGATGCC AGGAAGATAC TTAACAGGGA AGTGAGAGGG
401  CCGCGGCAAA GCCGTTTTTC CATAGGCTCC GCCCCCCTGA CAAGCATCAC
451  GAAATCTGAC GCTCAAATCA GTGGTGGCGA AACCCGACAG GACTATAAAG
501  ATACCAGGCT ACACCGCTGT TATGGCCGCG TTTGTCTCAT
551  CCGGCAAAAT GTCGACTGCG GGCGAGCGGA AATGGCTTAC GAACGGGGCG
601  GAGATTTCCT GGAAGATGCC AGGAAGATAC TTAACAGGGA AGTGAGAGGG
651  CCGGCAAAAT GTCGACTGCG GGCGAGCGGA AATGGCTTAC GAACGGGGCG
701  TTTTACTCG CGATCATTTA TCCCTATCTG ATGGTTAAAT
751  GAGATTTCCT GGAAGATGCC AGGAAGATAC TTAACAGGGA AGTGAGAGGG
801  CCGGCAAAAT GTCGACTGCG GGCGAGCGGA AATGGCTTAC GAACGGGGCG
851  GAAATCTGAC GCTCAAATCA GTGGTGGCGA AACCCGACAG GACTATAAAG
901  ATACCAGGCT ACACCGCTGT TATGGCCGCG TTTGTCTCAT
951  CCGGCAAAAT GTCGACTGCG GGCGAGCGGA AATGGCTTAC GAACGGGGCG
1001 GTTTTATCTG TTGTTTGTCG GTGAACGCTC TCCTGAGTAG GACAAATCCG
1051 CCGCCTAGA CCTAGGGGAT ATATTCGGCT TCCCTATCTG ATGGTTAAAT
1101 ACGCTCGGTC GTTCCGACTCG GCGGAGCGGA AATGGCTTAC GAACGGGGCG
1151 GAGATTTCCT GGAAGATGCC AGGAAGATAC TTAACAGGGA AGTGAGAGGG
1201 CCGGCAAAAT GTCGACTGCG GGCGAGCGGA AATGGCTTAC GAACGGGGCG
1251 GAAATCTGAC GCTCAAATCA GTGGTGGCGA AACCCGACAG GACTATAAAG
1301 ATACCAGGCT ACACCGCTGT TATGGCCGCG TTTGTCTCAT
1351 CCGGCAAAAT GTCGACTGCG GGCGAGCGGA AATGGCTTAC GAACGGGGCG
1401 GTTTTATCTG TTGTTTGTCG GTGAACGCTC TCCTGAGTAG GACAAATCCG
1451 CCGCCTAGA CCTAGGGGAT ATATTCGGCT TCCCTATCTG ATGGTTAAAT
1501 ACGCTCGGTC GTTCCGACTCG GCGGAGCGGA AATGGCTTAC GAACGGGGCG
1551 GAAATCTGAC GCTCAAATCA GTGGTGGCGA AACCCGACAG GACTATAAAG
1601 ATACCAGGCT ACACCGCTGT TATGGCCGCG TTTGTCTCAT
1651 CCGGCAAAAT GTCGACTGCG GGCGAGCGGA AATGGCTTAC GAACGGGGCG
1701 GAAATCTGAC GCTCAAATCA GTGGTGGCGA AACCCGACAG GACTATAAAG
1751 ACGCTCGGTC GTTCCGACTCG GCGGAGCGGA AATGGCTTAC GAACGGGGCG
1801 TTTTACTCG CGATCATTTA TCCCTATCTG ATGGTTAAAT
1851 GAAATCTGAC GCTCAAATCA GTGGTGGCGA AACCCGACAG GACTATAAAG
1901 CCGGCAAAAT GTCGACTGCG GGCGAGCGGA AATGGCTTAC GAACGGGGCG
1951 TTTTACTCG CGATCATTTA TCCCTATCTC CTGCAACCCG AGAGTCCCGC
2001 TCAGAAAGAC TCGTCAAGAA GCCGATAGAA GCCGATGCGC TGCAGATCGG
```
178

```
2051 GAGCGGCAGT ACCGTAAGGC ACGAGGAAGC GTGCAACCAC TTTGCCGGCA
2101 AGCTCTTCAG CAATATCAGG GGTAGCAGGC CATGGGCTCA GACAGATCCC
2151 CGGCCAACCAC AGTCCGAGGA CTTCGAGGTC ACGCCCACCC AGCCGGCCAC
2201 AGTCGATGAA TCCAGAAAAG CGGCCATTTT CCACCATGAT ATTCGGCAAG
2251 CAGGCATCGC CATGGGTACAC GACGAGATCC TCGCCGTCGG GCATGCGCGC
2301 CTTGAGCCTG GCGAACAGTT CGGCTGGCGC GAGCCCCTGA TGCTCTTCGT
2351 CCAGATCATC CTGATCGACA AGACCGGCTT CCATCCGAGT ACGTGCTCGC
2401 TCGATGCGAT GTTTCGCTTG GTGGTCGAAT GGGCAGGTAG CCGGATCAAG
2451 CCGGACAGTT TGCCAAATGC TGCCAGTCCC TTCCCGCTTC AGTGACAACG
2501 TCGCTGCCTCG TCCTGCAGTT CATTCAGGGC ACCGGACAGG TCGGTCTTGA
2551 CAAAAAGAAC CGGGCCGCCC TGGCGTGGAC GCCGGAACAC GCCGGCATCA
2601 GAGCAGCCGA TTGCTCTGTTG TGCCCAATTC ACGATCCTCA TCCTGTCTCT
2651 TGATCAGATC TTGATCCCCT GCGCCATCAG ATCCTTGGCG GCAAGAAAGC
2701 CATCCAGTTT ACTTTGCAGG GCTTCCCAAC CTTACCAGAG GGCGCCCCAG
2751 CGGAGCTTTT TTGCCCTGCG TGACCAGATC CCGGAGTTGG AAAACAATGA
2801 AAAGGCCCCA CAATTTTCACA CAGGGCCCTC GACACCGAG GAGAATGTCA
2851 ATCCCTGCGG GCAAGAGGAC CATCCAGTTT ACTTTGTGAG GCTTCCCAAC
2901 CCTTACCAGG GGGCCGCCAC CTGGCAATCC ATCTTTGCAAT GATGGATGGT
2951 GAGCGGATAA CAATTTCACA CAGGGCCCTC GACACCGAG GAGAATGTCA
3001 AGGAGGCCAAC ACAAAAGGTC TTGGAGGGCC AGAGAAGGAA CGAGCTAAA
3051 CGGAGCTTTT TTGCGCCTGG TGACCAAGTC CCGGAGTTGG AAAACAATGA
3101 AAAGGGCTCC AAGTGATGTT TCCTTTAAAAA GCGACACGCA TGCATCGTGT
3151 CCGTCCAAGC AGAGGAGCAA AAGCTCATTT CTGAAGGAGG CTGTTGCGGC
3201 AAACGACCTAG AACAGTGAAT ACACAAGCTT GAACAGCTAC GGAACTCTTG
3251 TGGGTAAAAG AAGAAGAAGA AAACGACTCC TTATAACAGA AATGTTCGTA
3301 GCAATCACC CGGACACT
```
Figure A.4. Plasmid map of the pLuxR vector containing wild-type lux.
Figure A.5. Nucleotide sequences of the *luxR* homologs used in this work

**lasR**

1  ATGGCCCTTGG TTGACGGTTT TCTTGAGCTG GAACGCTCAA GTGGAAAATT  
51  GAGTGGAGAC GCCATCTGGC AGAAGATGCC GACCCAGGCTT GTAGTCTCGA  
101  AGATCTCTGTT CGGCCCTGTG CTAAGGACCA GCCAGGACTT CGAGAACGCC  
151  TGTCTACGCG CGGGTCGACC CGACGTTCAG TCACCTGACC CGAGACGAGC  
201  TGCCGATTTT CTGGGAAACC AGGCTTCTTC TAGAGTGGTT CCTGCGCCAC  
251  CTGGGAAACC GCAGAGGACC AAAGGTTTCC TAGAGTGGTT CCTGCGCCAC  
301  CGGAAAACCC GCAGAGGACC AAAGGTTTCC TAGAGTGGTT CCTGCGCCAC  
351  CGAATCTGGA TCGCTACGCG CGAGAAGGAT TGATGGACGA TTCTGCGCCAC  
401  CGAATCTGGA TCGCTACGCG CGAGAAGGAT TGATGGACGA TTCTGCGCCAC  
451  CGAATCTGGA TCGCTACGCG CGAGAAGGAT TGATGGACGA TTCTGCGCCAC  
501  CGAATCTGGA TCGCTACGCG CGAGAAGGAT TGATGGACGA TTCTGCGCCAC  
551  CGAATCTGGA TCGCTACGCG CGAGAAGGAT TGATGGACGA TTCTGCGCCAC  
601  CGAATCTGGA TCGCTACGCG CGAGAAGGAT TGATGGACGA TTCTGCGCCAC  
651  CGAATCTGGA TCGCTACGCG CGAGAAGGAT TGATGGACGA TTCTGCGCCAC  
701  CGAATCTGGA TCGCTACGCG CGAGAAGGAT TGATGGACGA TTCTGCGCCAC  

**rhlR**

1  ATGAGGAATG ACAGAGGCTT TTTGCTGTG TGAGGATGTT TGGGATGCGA  
51  GATGAGAGCG ATCCAGAGCA CAGAGGCTT TGGCATGCTG GAGGAGGATG  
101  TGCTCTTCTCA AGCAGAGGCTT TGGCATGCTG GAGGAGGATG  
151  AGCAGAGGCTT TGGCATGCTG GAGGAGGATG  
201  AGCAGAGGCTT TGGCATGCTG GAGGAGGATG  
251  AGCAGAGGCTT TGGCATGCTG GAGGAGGATG  
301  AGCAGAGGCTT TGGCATGCTG GAGGAGGATG  
351  AGCAGAGGCTT TGGCATGCTG GAGGAGGATG  
401  AGCAGAGGCTT TGGCATGCTG GAGGAGGATG  
451  AGCAGAGGCTT TGGCATGCTG GAGGAGGATG  
501  AGCAGAGGCTT TGGCATGCTG GAGGAGGATG  
551  AGCAGAGGCTT TGGCATGCTG GAGGAGGATG  
601  AGCAGAGGCTT TGGCATGCTG GAGGAGGATG  
651  AGCAGAGGCTT TGGCATGCTG GAGGAGGATG  
701  AGCAGAGGCTT TGGCATGCTG GAGGAGGATG
**Figure A.6.** Amino acid sequences of the LuxR homologs used in this work

**LasR**

<table>
<thead>
<tr>
<th></th>
<th>MALVDGFLEL ERSSGKEWS AILQKMASDL GFSKILFGLL PKDSQDYENA</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>FIVGNYPAAW REHYDRAGYA RVDPTVSHCT QSVPWIEFSEP SIYQTRKQHE</td>
</tr>
<tr>
<td>10</td>
<td>FFEEASAAGL VYGLTMLHG ARGELGALSL SVEAENRAEA NRFIESVLPT</td>
</tr>
<tr>
<td>15</td>
<td>LWMLKDYALQ SGAGLAEEHP VSKPVLTSR EKEVLQWCAI GKTSEISVI</td>
</tr>
<tr>
<td>20</td>
<td>CNCSEANVNF HMGNIKRRKFG VTSRRVAAIM AVNLGLITL</td>
</tr>
</tbody>
</table>

**RhlR**

<table>
<thead>
<tr>
<th></th>
<th>MRNDGGFLLW WDGLRSEMQP IHDSQGVFAV LEKEVRLGF DYYAYGVRHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>IPFTRPKTEV HGTPKAEWLE RYQMNYGAV DPAILNGLRS SEMVVWSDSL</td>
</tr>
<tr>
<td>10</td>
<td>FDQRSMLWNE ARDWGLCVGA TLPIRAPNNL LSVLSVARDQ QNISFEE</td>
</tr>
<tr>
<td>15</td>
<td>IRLRLRCMIE LLQKTLTDLHP MLMSNPGCLSHREERILQ WTADGKSSGE</td>
</tr>
<tr>
<td>20</td>
<td>IAIILSIES TVNFKHKNIQ KKFDPNKTL AAAYAAPALG I</td>
</tr>
</tbody>
</table>
Table A.1. Nucleotide and amino acid changes in the recovered *luxR*/LuxR mutants described in Chapter 2

<table>
<thead>
<tr>
<th>luxR gene</th>
<th>LuxR protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>base and substitution</td>
<td>amino acid residue and substitution</td>
</tr>
</tbody>
</table>

**First Generation**

<table>
<thead>
<tr>
<th>luxR</th>
<th>luxR gene base and substitution</th>
<th>LuxR protein amino acid residue and substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>pLuxR-G1A</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>A159→G</td>
<td>Synonymous</td>
</tr>
<tr>
<td>pLuxR-G1B</td>
<td>A133→T</td>
<td>Ile45→Phe</td>
</tr>
<tr>
<td></td>
<td>T501→C</td>
<td>Synonymous</td>
</tr>
<tr>
<td>pLuxR-G1C</td>
<td>A284→G</td>
<td>Asn95→Ser</td>
</tr>
<tr>
<td></td>
<td>A414→C</td>
<td>Synonymous</td>
</tr>
<tr>
<td>pLuxR-G1D</td>
<td>A286→G</td>
<td>Ile96→Val</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td>pLuxR-G1E</td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>A683→G</td>
<td>Thr228→Ala</td>
</tr>
<tr>
<td>pLuxR-G1F</td>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>A403→G</td>
<td>Met135→Val</td>
</tr>
<tr>
<td></td>
<td>A522→G</td>
<td>Ile174→Met</td>
</tr>
</tbody>
</table>

**Second Generation**

<table>
<thead>
<tr>
<th>luxR</th>
<th>luxR gene base and substitution</th>
<th>LuxR protein amino acid residue and substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>pLuxR-G2A</td>
<td>A133→T</td>
<td>Ile45→Phe</td>
</tr>
<tr>
<td></td>
<td>A403→G</td>
<td>Met135→Val</td>
</tr>
<tr>
<td>pLuxR-G2B</td>
<td>A133→T</td>
<td>Ile45→Phe</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td></td>
<td>A522→G</td>
<td>Ile174→Met</td>
</tr>
<tr>
<td>pLuxR-G2C</td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td>pLuxR-G2D</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>A286→G</td>
<td>Ile96→Val</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td></td>
<td>A414→C</td>
<td>Synonymous</td>
</tr>
<tr>
<td>pLuxR-G2E</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td>pLuxR-G2F</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>A133→T</td>
<td>Ile45→Phe</td>
</tr>
<tr>
<td></td>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td>pLuxR-G2G</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>A286→G</td>
<td>Ile96→Val</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td>pLuxR-G2H</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>A284→G</td>
<td>Asn95→Ser</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>A522→G</td>
<td>Ile174→Met</td>
</tr>
</tbody>
</table>
Table A.2.  Single-mutants mutants described in Chapter 2

<table>
<thead>
<tr>
<th>luxR gene base and substitution</th>
<th>LuxR amino acid residue and substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td>A133→T</td>
<td>Ile45→Phe</td>
</tr>
<tr>
<td>A159→G</td>
<td>Synonymous</td>
</tr>
<tr>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td>A284→G</td>
<td>Asn95→Ser</td>
</tr>
<tr>
<td>A286→G</td>
<td>Ile96→Val</td>
</tr>
<tr>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td>A403→G</td>
<td>Met135→Val</td>
</tr>
<tr>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td>A414→C</td>
<td>Synonymous</td>
</tr>
<tr>
<td>T501→C</td>
<td>Synonymous</td>
</tr>
<tr>
<td>A522→G</td>
<td>Ile174→Met</td>
</tr>
<tr>
<td>A683→G</td>
<td>Thr228→Ala</td>
</tr>
</tbody>
</table>
**Table A.3.** Nucleotide and amino acid changes in luxR/LuxR and lasR/LasR mutants described in Chapter 4

<table>
<thead>
<tr>
<th>luxR</th>
<th>luxR gene base and substitution</th>
<th>LuxR protein amino acid residue and substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parent (2nd generation C8HSL-responder)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLuxR-G2E</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td></td>
<td>G200→T</td>
<td>Arg67→Met</td>
</tr>
<tr>
<td>Specificity mutants from selections</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLuxR-G2E-R67M</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td>(recovered 7/12)</td>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>G200→T</td>
<td>Arg67→Met</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td></td>
<td>A172→G</td>
<td>Ile58→Val</td>
</tr>
<tr>
<td></td>
<td>G200→T</td>
<td>Arg67→Met</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td></td>
<td>A423→G</td>
<td>Synonymous</td>
</tr>
<tr>
<td>pLuxR-C10Sp1</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>T114→C</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>A172→G</td>
<td>Ile58→Val</td>
</tr>
<tr>
<td></td>
<td>G200→T</td>
<td>Arg67→Met</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td></td>
<td>A423→G</td>
<td>Synonymous</td>
</tr>
<tr>
<td>pLuxR-C10Sp2</td>
<td>A34→G</td>
<td>Ile12→Val</td>
</tr>
<tr>
<td></td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>A101→G</td>
<td>Lys34→Arg</td>
</tr>
<tr>
<td></td>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>G200→T</td>
<td>Arg67→Met</td>
</tr>
<tr>
<td></td>
<td>T255→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>A298→C</td>
<td>Asn100→His</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>A355→G</td>
<td>Ile119→Val</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td></td>
<td>A314→G</td>
<td>Synonymous</td>
</tr>
<tr>
<td>pLuxR-C10Sp_8</td>
<td>A97→G</td>
<td>Thr33→Ala</td>
</tr>
<tr>
<td></td>
<td>T162→A</td>
<td>Synonymous</td>
</tr>
<tr>
<td></td>
<td>G200→T</td>
<td>Arg67→Met</td>
</tr>
<tr>
<td></td>
<td>A314→G</td>
<td>Lys105→Arg</td>
</tr>
<tr>
<td></td>
<td>T346→G</td>
<td>Ser116→Ala</td>
</tr>
<tr>
<td></td>
<td>G405→A</td>
<td>Met135→Ile</td>
</tr>
<tr>
<td></td>
<td>A470→G</td>
<td>Asn157→Ser</td>
</tr>
<tr>
<td></td>
<td>T498→C</td>
<td>Synonymous (SalI)</td>
</tr>
<tr>
<td></td>
<td>T510→C</td>
<td>Synonymous (SalI)</td>
</tr>
</tbody>
</table>

**Introduction of R67M mutation into wild-type LuxR**

| pLuxR-R67M | G200→T | Arg67→Met |

**Introduction of R61M (aligns with R67M of LuxR) mutation into wild-type LasR**

| pLasR-R61M | GCT181→ATG | Arg61→Met |
Table A.4. Nucleotide and amino acid changes in *luxR/LuxR* variants described in Chapter 5

<table>
<thead>
<tr>
<th></th>
<th>luxR gene base and substitution</th>
<th>LuxR protein amino acid residue and substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLuxR(SalI)</td>
<td>T498→C</td>
<td>Synonymous (SalI)</td>
</tr>
<tr>
<td></td>
<td>T501→C</td>
<td>Synonymous (SalI)</td>
</tr>
<tr>
<td><strong>DNA-binding mutants from selections</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLuxR-R212C/Q232H</td>
<td>C634→T</td>
<td>R212→C</td>
</tr>
<tr>
<td></td>
<td>A695→T</td>
<td>Q232→H</td>
</tr>
<tr>
<td>pLuxR-R212H</td>
<td>G635→A</td>
<td>R212→H</td>
</tr>
<tr>
<td><strong>DNA-binding mutants from saturation of residue 212</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLuxR-R212S</td>
<td>C634→A</td>
<td>R212→S</td>
</tr>
<tr>
<td><strong>Single-mutants</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLuxR-R212C</td>
<td>C634→T</td>
<td>R212→C</td>
</tr>
<tr>
<td>pLuxR-Q232H</td>
<td>A695→T</td>
<td>Q232→H</td>
</tr>
</tbody>
</table>
B. Sequences of the fluorescent reporters used for screening and characterization of LuxR variants.

B.1. General information

The pluxGFPuv plasmid was constructed for screening LuxR variants that activate gene expression in the presence of non-cognate signal molecules, based on LuxR-mediated expression of a green fluorescent protein (GFP) variant. GFPuv (Clontech) was used because it can be excited by standard long-wavelength ultraviolet light such that a plate of colonies can be easily screened for GFPuv production by illuminating the plate with in a standard gel box and visually inspecting for green colonies.

A variant of pluxGFPuv containing a mutated lux promoter was constructed for the final round of screening used to identify the best variants from the pool of variants selected for their ability to activate gene expression at the new promoter. The promoter differs from P_{lux} at two positions: the 5 and -5 position of the lux box were altered simultaneously to maintain the inverted repeat structure.

A liquid-phase bioassay system was developed that assesses LuxR-mediated gene activation by quantitatively determining the level of fluorescence due to GFPuv.

Reporter proteins containing two other fluorescent reporters were constructed for use with microscopes. The GFPmut3 variant has an excitation maximum at 489. The DsRed
construct was used to show a lack of crosstalk between LuxR and the evolved acyl-HSL specificity mutant, LuxR-G2E-R67M.

Figure A.7. lists the nucleotide sequence of the wild-type \( P_{lux} \) promoter.

Figure A.8. lists the nucleotide sequences of the wild-type lux box and mutant \( luxG5A \) box.

Figure A.9. lists the nucleotide sequences of the fluorescent reporters, \( gfpuv, gfpmut3 \) and \( dsred\)-express.

Figure A.10. lists the amino acid sequences of the fluorescent reporters, GFPuv, GFPmut3 and DsRed-express.

Figure A.11. lists the nucleotide sequence of the pluxGFPuv reporter plasmid.

Figure A.12. shows the plasmid map of pluxGFPuv.
**Figure A.7.** Nucleotide sequence of the wild-type $P_{lux}$ promoter.

1  AGTCCTTTGA TTCTAATAAA TTGGATTTTT GTCACACTAT TGTATCGCTG  
51  GGAATACAAT TACTTAACAT AAGCACCTGT AGGATCGTAC AGGTTTACGC  
101 AAGAAATGG TTTGTTATAG TCGAATAAAC GCAAGGGAGG TTGGT

**Figure A.8.** Nucleotide sequences of the wild-type $lux$ box and mutant $lux_{G5A}$ box.

$Lux$ box

1  ACCTGTAGGA TCGTACAGGT

$Lux_{G5A}$ box

1  ACCTATAGGA TCGTATAGGT
Figure A.9. Nucleotide sequences of the fluorescent reporters, *gfpuv*, *gfpmut3* and *dsred-express*, used in this study.

**gfpuv**

1  ATGAGTAAAG GAGAAGAACT TTTCACTGGA GTTGTCCCAA TTCTTGTGA

51  ATTAGATGGT GATGTTAATG GGCACAAATT TTCTGTCAAT GGAGAGGGTG

101  AAGGTGATGC AACATACGGA AAACCTACCC TTAAATTAT TTGCACTACT

151  GGAAGACTAC CTGTTCATGC GCAACACACT TCACTACCT TCTCTATGG

201  TGGTCAATGC TGGTTCCGTT ATCCGGATCA TATGAAACGG CATGACTTTT

251  TCAAGAGTGC CATCGCCGAA GTTTATGTAC AGGAACGCAC TATATCTTTC

301  AAGGATGACG GGAACCTACA AAGGACGCTG TGAATCAGAT TTGAAAGTTA

351  TACTCTGTGG TACGTCATCG AGGTAAAAGG TATTGATTTT AAAGAAGATG

401  GAAACATTCT CGGACACAAA CTGCACTACA ACTACTACCT ACTACTAGGA

451  TAGATTCGAC GCAACACTACA AAGAATGGA ATCAAAGCTA ACTTCAAAAT

501  CGGCCCAACA ATTTAACATC TGGACAGCCT TGGATCGAT CAAGGACAGA

551  TACATCATTG TGGAAATACG TATGAAAACT ATGTTTATTT AAGGACATG

601  CATCGGTGCT TCACTGTGGT TAACTGCTGC TGGGATTACA CATGGAAGAA

651  ATGAGTAAAG GAGAAGAACT TTTCACTGGA GTTGTCCCAA TTCTTGTGA

701  ATGAGTAAAG GAGAAGAACT TTTCACTGGA GTTGTCCCAA TTCTTGTGA

**gfpmut3**

1  ATGCGTAAAG GAGAAGAACT TTTCACTGGA GTTGTCCCAA TTCTTGTGA

51  ATTAGATGGT GATGTTAATG GGCACAAATT TTCTGTCAAT GGAGAGGGTG

101  AAGGTGATGC AACATACGGA AAACCTACCC TTAAATTAT TTGCACTACT

151  GGAAGACTAC CTGTTCATGC GCAACACACT TCACTACCT TCGGTTATGG

201  TGGTCAATGC TGGTTCCGTT ATCCGGATCA TATGAAACAG CATGACTTTT

251  TCAAGAGTGC CATCGCCGAA GTTTATGTAC AGGAAAGAAC TATATTTTTC

301  AAGGATGACG GGAACCTACA AAGGACGCTG TGAATCAGAT TTGAAAGTTA

351  TACTCTGTGG TACGTCATCG AGGTAAAAGG TATTGATTTT AAAGAAGATG

401  GAAACATTCT CGGACACAAA CTGCACTACA ACTACTACCT ACTACTAGGA

451  TAGATTCGAC GCAACACTACA AAGAATGGA ATCAAAGTTA ACTTCAAAAT

501  CGGCCCAACA ATTTAACATC TGGACAGCCT TGGATCGAT CAAGGACAGA

551  TACATCATTG TGGAAATACG TATGAAAACT ATGTTTATTT AAGGACATG

601  CATCGGTGCT TCACTGTGGT TAACTGCTGC TGGGATTACA CATGGAAGAA

651  ATGAGTAAAG GAGAAGAACT TTTCACTGGA GTTGTCCCAA TTCTTGTGA

701  ATGAGTAAAG GAGAAGAACT TTTCACTGGA GTTGTCCCAA TTCTTGTGA
dsred-express

1    ATGGCCTCCT CCGAGGACGT CATCAAGGAG TTCATGCCTGCT TCAAGGTGC
51   CATGGAGGGC TCCGTGAACG GC CACGAGTT CGAGATCGAG GGCGAGGGCG
101  AGGGCGGCC ACCGAGGGC ACCCAGACCG CCAAGCTGAA GGTGACCAAG
151  GCCGCCCCTG GCCTTTGCCTGCTGACATC AGTTCCAGTA
201  CGGCCTCAAG GTGTACGTGA AGCACCAGCG CGACATCCCG GACTACAAGA
251  AGCTGTCCCT CCCCCAGGGC TTCAAGTGGG AGCGCGTGAT GAACCTCGAG
301  GACGGGGCGT TGGTGACCCT GACCCAGGAC TCCCTCCCTGCT AGGACGGCTC
351  CTTCATCTAC AAGGTGAGTT TCTCGGCGCT GAACTTCCCC CTCCAGGGCC
401  CGCCTATGCA GAGAAGACT ATGGGCTGGG AGGCCCTCCAG CAGACGCCTG
451  TACCCCCCGG AGGCGTGCT GAGGCGTGAG ATCCACAAGG CCCCTGACCT
501  GAAGGACGCC GGCACACTTAC TGGTGGAGTT CAAATCCATC TACATGGCCA
551  AGAAGGCCCTG GACGCTGCC GCCCTACTCTA CTGACTCTGCT AGGACCTGAG
601  ATCACCTCCC ACAACGAGGA CTACACACTC GTGAGGCTGT ACGACGGCGC
651  CGAGGGCCGC CACCACCTGT TCCTG
**Figure A.10.** Amino acid sequence of the fluorescent reporters GFPuv, GFPmut3 and DsRed-express, used in this study.

**GFPuv**

1  MSKGEELFTG VVPILVELDG DVNGHKFSVS GEGEDATYG KLTLKFICTT
51  GKLPPVPWPTL VTTFSYGVCQ FSRYPDHMKR HDFFSAMPE GYVQERTISF
101  KDDGNYKTRA EVKFEGDTLV NRIELKGIDF KEDGNILGHK LEYNYSNHNV
151  YITADQKQNG IKANFKIRHN IEDGSVQLAD HYQQNTPIGD GPVLLPDNYH
201  LSTQSALSKD PNEKRDHMVL LEFVTAGAGT HGMDELYK

**GFPmut3**

1  MSKGEELFTG VVPILVELDG DVNGHKFSVS GEGEDATYG KLTLKFICCTT
51  GKLPPVPWPTL VTTFGYGVQC FARYPDHMKQ HDFFSAMPE GYVQERTIFF
101  KDDGNYKTRA EVKFEGDTLV NRIELKGIDF KEDGNILGHK LEYNYSNHNV
151  YIMADQKQNG IKANFKIRHN IEDGSVQLAD HYQQNTPIGD GPVLLPDNYH
201  LSTQSALSKD PNEKRDHMVL LEFVTAGAGT HGMDELYK

**DsRed-express**

1  MASEDVIKE FMRFKVRMEG SVGHEFEIE GEGGRPYEG TQTAKLKVTK
51  GGPLPFAWDI LSPQFYGSK VYVHPADIP DYKKLSFPEG FWKERVNMEF
101  DGGVVTQVD SSLQDNSFIY KVBIQVFVNP SDGPVMQKKT MGWEASTERL
151  YPRDGVLKGE IHKALKLKDG GHLVEFKSI YMAKPVQLP GYYYVDSKLD
201  ITSHNEDYTI VEQYERAEG PHLFL
Figure A.11. Nucleotide sequence of the pluxGFPuv reporter plasmid.

1  GACGTCAGTC CTTTGATTCT AATAAATTGG ATTTTTGTCA CACTATTGTA
51  TCGCTGGGAAA TACAATTAAT TAAACATAGC ATTTGTCAAC ATCGACAGGT
101  TTAGCCAGGA GTAAAGGGGT TTAGTTGCA TCTAAGTCGA TGGAGGGTTG
151  TATGAGTCAA GGAGAAGAAT TTTTACTGTC AATATGTTAC ATTTCTTGTG
201  AATTAGATGG TGATGTGATA ATGGCACAAT TTTCTGTGAG TGAGAGGGTT
251  GAAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC TTTCTTATG
301  GTATGCGCGT CTTTGGATGG AATGAGTCAA ATATGTTAAG ATTTCTTGA
351  GTGTTTGGCTA AATAGATGG TTTTACGCAA TGGAGGTTTC TTTCTTATG
401  CTTAAGATTC CGATGACGCA GGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA
451  TGGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC TTTCTTATG
501  AATTAGATGG TGATGTGATA ATGGCACAAT TTTCTGTGAG TGAGAGGGTT
551  GAAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC TTTCTTATG
601  ATACACATACG GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA
651  TGGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC TTTCTTATG
701  CTTAAGATTC CGATGACGCA GGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA
751  CTTAAGATTC CGATGACGCA GGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA
801  ATACACATACG GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA
851  TGGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC TTTCTTATG
901  AATTAGATGG TGATGTGATA ATGGCACAAT TTTCTGTGAG TGAGAGGGTT
951  GAAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC TTTCTTATG
1001  ATACACATACG GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA
1051  TGGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC TTTCTTATG
1101  AATTAGATGG TGATGTGATA ATGGCACAAT TTTCTGTGAG TGAGAGGGTT
1151  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1201  AATTAGATGG TGATGTGATA ATGGCACAAT TTTCTGTGAG TGAGAGGGTT
1251  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1301  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1351  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1401  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1451  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1501  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1551  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1601  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1651  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1701  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1751  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1801  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1851  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1901  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
1951  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
2001  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC
2051  GCACAGAGGT GAGAGGCTGTCAG AAAGTCAAAT TTTTACGCAA TGGAGGTTTC

192
Figure A.12. Plasmid map of the pluxGFPuv reporter plasmid.
C. Sequences of plasmids, genes and proteins comprising the dual selection system for the directed evolution of LuxR.

C.1. General information

Two separate selection plasmids were constructed for ON and OFF round of selection using the dual selection system. The ON selection plasmid, pluxCAT, yields chloramphenicol resistance with LuxR-dependent transcriptional activation. Chloramphenicol acetyl-transferase (cat) is expressed when LuxR activates gene expression at P_{luxI}.

The OFF selection plasmid, pluxBLIP, encodes β-lactamase inhibitory protein (Bli) under the control of the lux promoter and β-lactamase (Bla) under the control of a constitutive promoter. Bli inactivates Bla and renders cells with active LuxR alleles sensitive to β-lactam antibiotics.

Both pluxCAT and pluxBLIP contain the ColE1 replication origin and the bla antibiotic resistance gene.

The pBLIPCAT plasmid encodes both bli and cat under the control of the P_{lac/ara-1} promoter. In DH5α, both the positive and negative selection elements are constitutively expressed, such that cells are always resistant to chloramphenicol and sensitive to
carbenicillin. This plasmid confers an “always ON” phenotype and was used in the characterization of the dual selection system.

Figure A.13. lists nucleotide sequence of the genes used in the dual selection system.

Figure A.14. lists the amino acid sequences encoded by the cat, bla and bli genes.

Figure A.15. lists the nucleotide sequence of the positive (ON) selection plasmid, pluxCAT, containing the wild-type $P_{luxI}$ promoter.

Figure A.16. shows the plasmid map of the positive (ON) selection plasmid, pluxCAT, containing the wild-type $P_{luxI}$ promoter.

Figure A.17. lists the nucleotide sequence of the negative (OFF) selection plasmid, pluxBLIP, containing the wild-type $P_{luxI}$ promoter.

Figure A.18. shows the plasmid map of the negative (OFF) selection plasmid, pluxBLIP, containing the wild-type $P_{luxI}$ promoter.

Figure A.19. lists the nucleotide sequence of pBLIPCAT.

Figure A.20. shows the plasmid map of pBLIPCAT.
Figure A.13. Nucleotide sequence of the genes used in the dual selection system.

cat

1   ATGGAGAAAA AAATCACTGG ATATACCACC GTTGATATAT CCCAATGGCA
51  TCGTAAGGAA CATTTTGAGG CATTTGCTGA AGTTGCTCAA TGTACCTATA
101 ACCAGACCCT TCAGCTGGAAT ATTACGGCCT TTTTAAAGAC CGTAAAGAAA
151 AATAAGCACA AGTATTATCC GCCTTTTATT CACATTCTTG CCCGCTGAT
201 GAATGCTCAT CCGGAAATTTC GTATGGCAAT GAAAGACGTT GAGCTTGGTA
251 TATGGGATAG TGTTACCGGT TGTTACCGGT TTTCATGAA GCAAATGGCA
301 AGCTTTCATT CGCTCTGGGA TGGCAGTCAA CCGGATTACG CGGTAACTTA
351 ACACATATAT TCGCAAGAAG TGGCGTGTGA CGTAAAGAAA CTGGAATTTC
401 TCCCTAAAGG GTTTTATGGA AATTATTTTT TGGTCATCAG TCACTCTTTG
451 GTGAGTTTCTT CCAATTTTGG TTTAAACGTT CACATATGGA CCACACTCTT
501 CGCCCCCGTG TTAGCCTCAGG CCTATGCTGA ATGTTACCGT TATGAACTAA
551 TGGCGAAGAA AGTGTTTCA TTTCTGCTGC CTTGAGCAAT TGGACATTTT
601 GTCGGCAAGAA TGGTAAATAG TTTGATGGCA ATGTTGAGAA CTGGAATTTC
651 CGGGCGGCTAA

bla

1   ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCAGCTTG
51  TTTGCCCTCT CTTTTGACTC ACCGATTTCA GCTGCGATAA GAAATAGGTA
101 CTTGAAAGAC CGTGAGCCGG AGATGAGCAG ATGCTGCTGT GACATTGTCT
151 AGCGTGAATC TTTGCCCTCT ACATTTTGGTT AGATGACGCT TGGCGCAGCA
201 GAGCAGTCTT CAAAGGCAAA TTATTTACAT TTGCTGATTT CTTGAGAACT
251 CTTCCCGAGA ACCATTTTGC CGCGATTTAT AGTACAGTGC CCGGCGCTCA
301 GGTGAGTTTT ATCCTTGTCT TCATGGAAAT CTTATGCTGA AGTGGATTTG
351 AAGAAAGGCT AATTTTACAT GTGGTACATT ATGTTGAGAA ATGATGAGCT
401 ACTTACTTCT GGGAGACAGA GACAGGCGAT CTTGAGAATT AGGCGAGAAT
451 CACAACATGG TGGATGCCTT AATTTTACAT TTGCTGCTGT TGGCGGCTCA
501 GAATGAAGCC ATACAAACGC AGAGGCTATT AATATTTTAC AGTGCTGCTT
551 TGCCACCAAT CTTGAGCAAA AGTATTATCT ATGCCAATCT CTTGAGGACT
601 TGCAGCCTAC GCATTTATGC ATGGTGCTCA TGTGCTGCTT ATTGGATTTA
651 ACGCTCGTAC CTTGAGCAAT ATGTTGAGAA GTGCGCTGCT GATGAGCTA
701 GAGACCGGTC TGGGAGGCTT ATCCTGCTGA GTGCTGCTT ATGTTGAGAA
751 GCCGATGCCT CTTGAGCAAA ATGTTGAGAA AGAGGAGTAA ATGTTGAGAA
801 TATGATATGC AGAATAGAAG ATGCTGCTGA TGGCGGCTC ATGTTGAGAA
851 AGCATTTGCTT
bli

1  ATGCTTTTAT ATAAAAATGTG TGACAATCAA AATTATGGGG TTACTTACAT
51 GAAGTTTTTA TTGGCATT TT CGCTTTTAAT ACCATCGGTG GTTTTTGCAA
101 GTAGTGCAAG TGTTATGACA GGAGCAAAT ATCAGCGAGATT CAGTTTGGT
151 ATGACACGTC AGCAGGTCTT CGACATACGA GGTGCTGAGA ACATGTGAGAC
201 TGCTGGATCG TTCGCGTACA GCATCCATTG TCGTGGACAT GCAGCAGGAG
251 ACTATTATGC ATACGCAACC TTGGCGTTTC AGAGCTGACG TGCAGACGCA
301 AAGGTAAGTT CTGAAAGGCA GAAADACTG CTTGCACCAA GCAGCAAAAC
351 TCTTACTCTT GCTAAGTTCA ACCAAGTCAC TGTTGGTATG ACTAGAGCAC
401 AAGTACTTTC CACCGTCGGA CGAGGTTCCT GTACCACTTG GAGTGAGTAC
451 TATCCAGCAT ATCCACCGAC GCCAGGAGTG ACTCTCAAGGC TGTCTGCTT
501 CGATGTGGAG GGTACTGCTT CGACGGGTTT CTACCAGGTC TCGGCGCACCC
551 TCTGGTTCAC GGACGGGGTG TCTCAGGCGA AGCGGAGTGA GGACCTTGT
Figure A.14. Amino acid sequence of the proteins used in the dual selection system.

**CAT**

```plaintext
1  MEKKITGYTT VDISQWRHKE HFEAFQSVAQ CTYNQTVQLD ITAFLKTOKK
51  NKHKFYPAPFI HILARLMNAH PEFRMAMKG ELVIWDSVHP CYTVFHEQTE
101  TFSSLWSEYH DDFRQFLHIY SQDVACYGEN LAYFPKGFIE NMFFVSNAPW
151  VSFTSFPLNV ANMDMFFAPV FTMGKYYTQG DKVLMPAIQ VHHAVCDGFH
201  VGAEMNEELQQ YCDEWQGGA
```

**Bla**

```plaintext
1  MSIQHFRVAL IPFFAACLVP VFAHPETLKV VKDAEDQLGA RVGYIELDLN
51  SGKILESFRP EERFMNSTF KVLLCGAVLS RIDAQGEQLG RRHYSQNDL
101  VEYSPVTEKH LTDGNTVREL CSAAITMSDN TAANLiITTI GGPKELTAFL
151  HNMGDVTRVL DRWEPNELA IPNDERDYM TPAMATTLKR LLTGEILTLA
201  SROQLIDITME ADKVAGPLLRL SALPAGFIA DKSGAGERGS RGIAIAALPD
251  GKPSSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW
```

**Bli**

```plaintext
1  MLYKMCDNQ NYGVTYMFLK LAFLSLLIPS VFASSAGVMT GAKFTQIQFG
51  MTRQOVLDIA GAENCETGGSG FDGSLHCRG HAAADYYAYAT FGFTSAAADA
101  KVDKSQOEKL LAPSAPTLTL AKFQVTVMG TRAQVLVTVG QGSCTTWSEY
151  YPAPYSTAGV TLSLSCFVGD GYSSTGFYRG SAHLWFTDGVL LQQGKQWDLV
```
**Figure A.15.** Nucleotide sequence of the positive (ON) selection plasmid, pluxCAT, containing the wild-type $P_{lux}$ promoter.

<table>
<thead>
<tr>
<th></th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GACGTCAATC CTTTGGTATC AATAATTTG ATTTTGTCA CACTATTTGA</td>
</tr>
<tr>
<td>51</td>
<td>TCGCTGGGGA TAACATATCT TAACATTACG ACCTGAGGGA TGATACAGGT</td>
</tr>
<tr>
<td>101</td>
<td>TTACGCAAGGA AAATGGTTTG TTATAGTCTG AAATAAGCCAA GGGAGGTTGG</td>
</tr>
<tr>
<td>151</td>
<td>TATGGGAGTAA AAAATACATTG CAGTATCATA TCCAAATGAGG</td>
</tr>
<tr>
<td>201</td>
<td>ATCGTAAAGA ACATTTTGAG GCATTTCAGT CAGTTGCTCA ATGTACCTAT</td>
</tr>
<tr>
<td>251</td>
<td>AACGAGGCCG TACGCTGGCA TATTACGGCC TTTTTAAAGA CCGCCAGGTT</td>
</tr>
<tr>
<td>301</td>
<td>AACGAGCTAC CAGTTTTATC CCGCTTTATAT TTTCCATAGA AAGATCGTCC</td>
</tr>
<tr>
<td>351</td>
<td>CCATGATATCT ATCGGCTTAC GAGGCTCAGA TGATGCTAAAG AAGGAATGGG</td>
</tr>
<tr>
<td>401</td>
<td>AGCATGAAAA GTGTATTTGA AAATTAAGCA AAATAAATCA AAGGCAGAGG</td>
</tr>
<tr>
<td>451</td>
<td>CCAATTTATC CAGTATCATC CTGGGATGCA GAGGTGTTT GAGGCTGTAAG</td>
</tr>
<tr>
<td>501</td>
<td>CCGCTTCGCG GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>551</td>
<td>TGTCGCGAGA ATGCTTATGC AGAAAATGGA AAATGATGGG AACATGCTAG</td>
</tr>
<tr>
<td>601</td>
<td>GGCAGTGGTT TTTCTCCGAG TTTCCGCTAG GACAGGAAGA AAGATCGTCC</td>
</tr>
<tr>
<td>651</td>
<td>CCGCTTCGCG GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>701</td>
<td>TGTCGCGAGA ATGCTTATGC AGAAAATGGA AAATGATGGG AACATGCTAG</td>
</tr>
<tr>
<td>751</td>
<td>GGCAGTGGTT TTTCTCCGAG TTTCCGCTAG GACAGGAAGA AAGATCGTCC</td>
</tr>
<tr>
<td>801</td>
<td>GGCAGTGGTT TTTCTCCGAG TTTCCGCTAG GACAGGAAGA AAGATCGTCC</td>
</tr>
<tr>
<td>851</td>
<td>GGCAGTGGTT TTTCTCCGAG TTTCCGCTAG GACAGGAAGA AAGATCGTCC</td>
</tr>
<tr>
<td>901</td>
<td>GGCAGTGGTT TTTCTCCGAG TTTCCGCTAG GACAGGAAGA AAGATCGTCC</td>
</tr>
<tr>
<td>951</td>
<td>TTGTTTGTGC GTGAACGCTC TCCTGAGTAG GACAAATGCC CCGCCCTAGA</td>
</tr>
<tr>
<td>1001</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1051</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1101</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1151</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1201</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1251</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1301</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1351</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1401</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1451</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1501</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1551</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1601</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1651</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1701</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1751</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1801</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1851</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1901</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>1951</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>2001</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
<tr>
<td>2051</td>
<td>CCGTGGCGGT GACGTTTATG TTTCCAAATG GATGACTGCT ATGTACCTAT</td>
</tr>
</tbody>
</table>
2101 CACCGGCTCC AGATTTATCA GCAATAAACCC AGCCAGCCGG AAGGGCCGAG
2151 CGCAGAAGTG GTCTGGAACC TTTATCCGCC TCCATCCAGCT ATATTAATTTG
2201 TTGCCGGGAAG GCTAGAGTGA GTAGGTCGCC AGTTAATAGTT TGCGCAACG
2251 TTGGTTGCCAT TGCTACAGGC ATCGTGGTTGT CACGCTCGTC GTTGTGTATG
2301 GCTTCATTGCA GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC
2351 CATGGTTGTGC AAAAAAGCGG TTAGCTCTTT CGGTCTCCGG ATCGTTGTCA
2401 GAAGTAAAGTT GGCCCGAGTT TGATCACTCA TGGTTATGGC AGCACTGCAT
2451 AATTTCTCTTA CTGTGATGCC ATTCGTAAGA TGCTTCTCTG TGACTGCTGA
2501 GTACTCAACC AAGTCACTCT GAGAATAGTG TATGCGGCGA CCGAGTTTCT
2551 CTTGCCCGGC GTCAATACGG GATAATACCAG CGCCACCATG CAGAAGTTTA
2601 AAAAGTGCTCA TCATTGGAAA ACGTCTTCTTG GGGGAAACAC TCTCAAGGAT
2651 CTTACCGCTG TTGAGATCCA GTTGTACGTA ACCACTCGTG GCACCACACT
2701 GATCTTCAGC ATCTTCTACT TTACACGACGG TTTCTGGGTG AGCAAAACAA
2751 GGAAGGCAAA ATGCAGCCAAA AAAGGGAATA AGGGCGACAC GGAAATGTTG
2801 ATAACGTACTA CTCTTCTCATTT TCAATATTTA TGGAGCTATT CATCAGGTT
2851 ATTTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGATC AATAACCAA
2901 ATAGGGGTTTC CGCGCACATT TCCCCGAAAA GTGCCACCT

Note: pluxG5ACAT, a version of this plasmid containing the luxG5A box (see A.2.2 for sequence information), was utilized for the directed evolution of DNA-binding specificity by LuxR.
Figure A.16. Plasmid map of the pluxCAT selection plasmid containing the wild-type $P_{luxI}$ promoter.
Figure A.17. Nucleotide sequence of the negative (OFF) selection plasmid, pluxBLIP, containing the wild-type $P_{luxI}$ promoter.

1  GACGTAGTC CTTTGTATTCT AATAAAATGG ATTTTTGTCA CACTATTGTA
51  TCGCTGGGAA TACAATTACT TAACATAAGC ACCTGTAGGA TCGTACAGGT
101  TTACGCAGAA AAATGGTTTG TTATAGTCGA ATAAACGCAA GGGAGTTTGG
151  TATGACTAGT CAGCAGTCC TCGACATAGC AGGTGCTGAG AACTGTGAGA
201  TTACGCAAGA AAATGGTTTG TTATAGTCGA ATAAACGCAA GGGAGTTTGG
251  TTACGCAAGA AAATGGTTTG TTATAGTCGA ATAAACGCAA GGGAGTTTGG
301  CTCTTACTCT TGCTAAGTTC ATCAACGTGA CACCATCTGG CAGGTACGG
351  CTCTTACTCT TGCTAAGTTC ATCAACGTGA CACCATCTGG CAGGTACGG
401  CTCTTACTCT TGCTAAGTTC ATCAACGTGA CACCATCTGG CAGGTACGG
451  CTCTTACTCT TGCTAAGTTC ATCAACGTGA CACCATCTGG CAGGTACGG
501  CTCTTACTCT TGCTAAGTTC ATCAACGTGA CACCATCTGG CAGGTACGG
551  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
601  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
651  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
701  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
751  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
801  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
851  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
901  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
951  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1001  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1051  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1101  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1151  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1201  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1251  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1301  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1351  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1401  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1451  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1501  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1551  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1601  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1651  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1701  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1751  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1801  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1851  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1901  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
1951  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
2001  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
2051  ATAAAGTACC TGTTTTTCTAG GAGAGTCAGT GTTATACTGA AATAGGAGTA
Figure A.18. Plasmid map of the pluxBLIP selection plasmid containing the wild-type PluxI promoter.
Figure A.19. Nucleotide sequence of the selection test plasmid pBLIPCAT.

1  GTCGACTCGA GCATAGCATT TTTATCCATA AGATTAGGGG ATCTAACTTT
51  TACAATTTGTG AGGCCTCACA ATTATGATAG ATTCATAATTG GAGGGGATAA
101  CAAATTTTCA CAGAATTCAT TAAAGAGGAG AAAGTGACC ATGCTTTTAT
151  ATAAAAATGTG TGGACTAATCA AATTATGGGG ATGCTTTTAT GAGGGGATAA
201  TTGGGCAATTG CTGTTTAAATT ACCATCCGTG GTTGGGGATAA ATGCTTTTAT
251  TGTTATGCAAC AGGAAAAATT TACAGCGGTG CAGGGGATAA ATGCTTTTAT
301  AGTACAGTCTT CAGAAATTGT TTTTGGGATAA ATGCTTTTAT GAGGGGATAA
351  TCGGTCGACT TCGACAGGTC TCAGGGGATAA ATGCTTTTAT GAGGGGATAA
401  ATACGCAACCA TTTGCTTCCA CAGGCGGATAA ATGCTTTTAT GAGGGGATAA
451  CGAAAAAGCCT AGGAAACTTT TCGTCATCTT CAGGGGATAA ATGCTTTTAT
501  GCTAAGTCTTC ACCATTCTTC TCGTGGGCTT CAGGGGATAA ATGCTTTTAT
551  TACCAGGTTTG CACGAGAGAT AGGGTTGGGA ATGCTTTTAT GAGGGGATAA
601  CTAAACTTTT ACCAGCTTCA CAGGCGGATAA ATGCTTTTAT GAGGGGATAA
651  GCTAAGTTCT TCGCTCAGTC TGGGACAGAT CACGACAGAT TATGGCAATG
701  GGTGTCGACT TCGACAGGTC TCAGGGGATAA ATGCTTTTAT GAGGGGATAA
751  TCGGTCGACT TCGACAGGTC TCAGGGGATAA ATGCTTTTAT GAGGGGATAA
801  TGTTATGCAAC AGGAAAAATT TACAGCGGTG CAGGGGATAA ATGCTTTTAT
851  ATACGCAACC TTTGCTTCCA CAGGCGGATAA ATGCTTTTAT GAGGGGATAA
901  CGAAAAAGCCT AGGAAACTTT TCGTCATCTT CAGGGGATAA ATGCTTTTAT
951  GCTAAGTCTTC ACCATTCTTC TCGTGGGCTT CAGGGGATAA ATGCTTTTAT
1001  TACCAGGTTTG CACGAGAGAT AGGGTTGGGA ATGCTTTTAT GAGGGGATAA
1051  CTAAACTTTT ACCAGCTTCA CAGGCGGATAA ATGCTTTTAT GAGGGGATAA
1101  GGTGTCGACT TCGACAGGTC TCAGGGGATAA ATGCTTTTAT GAGGGGATAA
1151  TCGGTCGACT TCGACAGGTC TCAGGGGATAA ATGCTTTTAT GAGGGGATAA
1201  TGTTATGCAAC AGGAAAAATT TACAGCGGTG CAGGGGATAA ATGCTTTTAT
1251  ATACGCAACC TTTGCTTCCA CAGGCGGATAA ATGCTTTTAT GAGGGGATAA
1301  CGAAAAAGCCT AGGAAACTTT TCGTCATCTT CAGGGGATAA ATGCTTTTAT
1351  TACCAGGTTTG CACGAGAGAT AGGGTTGGGA ATGCTTTTAT GAGGGGATAA
1401  CTAAACTTTT ACCAGCTTCA CAGGCGGATAA ATGCTTTTAT GAGGGGATAA
1451  GGTGTCGACT TCGACAGGTC TCAGGGGATAA ATGCTTTTAT GAGGGGATAA
1501  TCGGTCGACT TCGACAGGTC TCAGGGGATAA ATGCTTTTAT GAGGGGATAA
1551  TGTTATGCAAC AGGAAAAATT TACAGCGGTG CAGGGGATAA ATGCTTTTAT
1601  ATACGCAACC TTTGCTTCCA CAGGCGGATAA ATGCTTTTAT GAGGGGATAA
1651  CGAAAAAGCCT AGGAAACTTT TCGTCATCTT CAGGGGATAA ATGCTTTTAT
1701  GGTGTCGACT TCGACAGGTC TCAGGGGATAA ATGCTTTTAT GAGGGGATAA
1751  TCGGTCGACT TCGACAGGTC TCAGGGGATAA ATGCTTTTAT GAGGGGATAA
1801  TGTTATGCAAC AGGAAAAATT TACAGCGGTG CAGGGGATAA ATGCTTTTAT
1851  ATACGCAACC TTTGCTTCCA CAGGCGGATAA ATGCTTTTAT GAGGGGATAA
1901  CTGTTTCTTG CTTGCTGGTT ACCAGTGCTC TGGGACAGAT CACGACAGAT
1951  TGTTATGCAAC AGGAAAAATT TACAGCGGTG CAGGGGATAA ATGCTTTTAT
2001  ATACGCAACC TTTGCTTCCA CAGGCGGATAA ATGCTTTTAT GAGGGGATAA
2051  GGTGTCGACT TCGACAGGTC TCAGGGGATAA ATGCTTTTAT GAGGGGATAA
Figure A.20. Plasmid map of the selection test plasmid pBLIPCAT.