BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

IMPROVED MODULAR MULTIPART DNA ASSEMBLY, DEVELOPMENT OF A

DNA PART TOOLKIT FOR *E. coli*, AND APPLICATIONS IN TRADITIONAL

BIOLOGY AND BIOELECTRONIC SYSTEMS

by

SONYA VICTORIA IVERSON

B.S., Montana State University, 2007

M.A., Boston University, 2015

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2016

Approved by

First Reader   _____
                                  Douglas Densmore, Ph.D.
                                  Associate Professor
                                  Electrical & Computer Engineering (ECE)
                                  Affiliated Faculty, Biomedical Engineering (BME)

Second Reader   _____
                                  Ahmad S. Khalil, Ph.D.
                                  Assistant Professor
                                  Biomedical Engineering (BME)

Dedication

This dissertation is dedicated to my parents who encouraged me, to my friends who helped me learn balance, and to the people who taught me the joy of science: Lawrence Fauque, Dr. David Sands, and Dr. Edward Schmidt.

Acknowledgements

IMPROVED MODULAR MULTIPART DNA ASSEMBLY, DEVELOPMENT OF A DNA PART TOOLKIT FOR *E. coli,* AND APPLICATIONS IN TRADITIONAL BIOLOGY AND BIOELECTRONIC SYSTEMS

SONYA VICTORIA IVERSON

Boston University Graduate School of Arts and Sciences, 2016

Major Professor: Douglas Densmore, Associate Professor of Electrical and Computer Engineering

ABSTRACT

DNA assembly and rational design are cornerstones of synthetic biology. While many DNA assembly standards have been published in recent years, only the Modular Cloning standard, or MoClo, has the advantage of publicly available part libraries for use in plant, yeast, and mammalian systems. No multipart modular library has previously been developed for use in prokaryotes. Building upon the existing MoClo assembly framework, we developed a collection of DNA parts and optimized MoClo protocols for use in *E. coli.* We present this assembly standard and library along with part characterization, design strategies,

potential applications, and troubleshooting. Developed as part of the Cross-disciplinary Integration of Design Automation Research (CIDAR) lab collection of tools, the CIDAR MoClo Library is publicly available and contains promoters, ribosomal binding sites, coding sequences, terminators, vectors, and a set of fluorescent control plasmids. Optimized protocols reduce reaction time and cost by >80% from previously published protocols. The CIDAR MoClo Library is the first bacterial DNA part library compatible with a multipart assembly standard.

To demonstrate the utility of the CIDAR MoClo system in a traditional biology context, we used the library and previous expression data to create a series of dual expression plasmids. In this manner, we produced a dual expression plasmid capable of expressing equimolar amounts of two variants of rabbit aldolase, a His-tagged wildtype protein and a single-amino-acid substitution mutant deficient in binding actin. This expression plasmid will enable the production of dimer-of-dimer heterotetramers needed for structural determination of the actin-aldolase interaction by electron microscopy. To employ CIDAR MoClo in a synthetic biology context, we produced a bioelectronic pH-mediated genetic logic gate with DNA circuits built using MoClo and integrated with Raspberry Pi computers, Twitter, and 3D printed components. Logic gates are an increasingly common biological tool with applications in cellular memory and biological computation. MoClo facilitates rapid

iteration of genetic designs, better enabling the development of cellular logic.

The CIDAR MoClo Library and assembly standard enable rapid design-build-test cycles in *E. coli* making this system advantageous for use in many areas of synthetic biology as well as traditional biological research.

Table of Contents

List of Tables

List of Figures

form of the dimer makes defining the actin binding interface difficult. (d) HRA-R42A heterotetramers may enable actin scaffolding while preventing crosslinking with a non-symmetrical aldolase complex. The charge difference between HRA and R42A should allow for the separation of different heterotetramers identifies. The species highlighted inside the black box are needed for actin-decorating. The bottom two-two heterotetramer is not expected to bind actin and will have no influence on downstream experiments. 18

List of Abbreviations

AraC   Arabinose C protein

BFP   Blue fluorescent protein

BIOFAB   BIOFAB: International Open Facility Advancing Biotechnology

CDS   Coding Sequence

CIDAR   Cross-disciplinary Integration of Design Automation Research

CNC   Computer numerical control

CP919   MC4100 derived *E. coli* strain

DNA   Deoxyribonucleic acid

DVA   DVL2 with illegal BsaI site repaired

DVK   Alternate nomenclature for DVL1

DVL0   pSB1C3 vector adapted to include MoClo cloning sites

DVL1   pSB1K3 vector adapted to include MoClo cloning sites

DVL2   pSB1A2 vector adapted to include MoClo cloning sites

eBFP2   Enhanced blue fluorescent protein

FP   Fluorescent protein

GFP   Green fluorescent protein

goi   Gene of interest

ho1   heme oxygenase-1 protein

ho1   heme oxygenase-1 coding sequence

hra   HRA coding sequence

HRA   Rabbit aldolase A protein, wildtype, with a 5' His tag

HSL	Homoserine lactone
ICE	Inventory of Composable Elements
JBEI	Joint BioEnergy Institute
LB	Luria Broth
Level 0	Synonymous with "basic part"
Level 1	Synonymous with "transcription unit"
Level 2	Synonymous with "device"
MAD	Modular allergen detection
MoClo	Modular Cloning Assembly Standard
OD	Optical density
pBAD	Arabinose promoter
PBS	Phosphate buffered saline
PCA	Polycarbonate
pcyA	Phycocyanobilin:ferredoxin oxidoreductase protein
pcyA	Phycocyanobilin:ferredoxin oxidoreductase coding sequence
pOR30	Plasmid acquired from the lab of Christopher Voigt, MIT
pTet	Tetracycline responsive promoter
R42A	Aldolase-R42A mutant coding sequence
R42A	Rabbit aldolase A protein, mutant with an R42A substitution
RBS	Ribosome binding sequence
RFP	Red Fluorescent Protein
TASBE	Tool-Chain to Accelerate Synthetic Biological Engineering

TU	Transcription unit
VF2	Vector Forward 2 – BioBrick standard forward primer
VR	Vector Reverse – BioBrick standard reverse primer
YFP	Yellow fluorescent protein

1 CHAPTER 1 INTRODUCTION

1.1 *Synthetic biology*

Synthetic biology integrates engineering, biology, chemistry, and physics in a multidisciplinary research approach to biological engineering (Cameron et al. 2014, Church et al. 2014). DNA assembly and modification is a cornerstone of the field, with applications ranging from cellular computation and memory devices to metabolic engineering and production of pharmaceuticals (Ro et al. 2006, Ham et al. 2008, Friedland et al. 2009, Bonnet et al. 2012, Keasling 2012). Increased modularity in genetic design in the form of publicly available libraries of characterized DNA parts provide a platform for rapid expression tuning of biological circuits (Casini et al. 2015). Likewise, the advent of affordable table top manufacturing is enabling new research approaches as 3D printing and computer numerical control (CNC) micromilling equipment and software tools become more available (O'Neill et al. 2014, Guckenberger et al. 2015).

Recent examples of computation in living cells include genetic regulatory networks that count (Friedland et al. 2009) and a variety of approaches to creating Boolean logic in living cells (Tamsir et al. 2011, Siuti et al. 2013). Using recombinases enzymes which are able to cut DNA and invert the sequence, researchers created all 16 two-input logic gates in *E. coli* with a single genetic module per gate (Siuti et al. 2013). Other logic gates designs have been produced

using common signaling pathways molecules in *E. coli* (Tamsir et al. 2011) and zinc-finger transcription factors in mammalian cells (Lohmueller et al. 2012).

Synthetic biology has reoriented metabolic engineering, providing rational design tools for the modulation of biosynthesis and isolation of useful compounds (Salis 2011, Keasling 2012). A hallmark publication, the antimalarial medication, Artemisinin, was produced in yeast providing an alternative source of the high-demand drug through a partnership with the Gates Foundation (Ro et al. 2006). In addition to biomedical applications, metabolic engineering is also being employed to produce high value molecules currently produced from petroleum (Lee et al. 2012).

Rapid DNA assembly forms the foundation of the design-build-test paradigm in synthetic biology across all of these applications. Despite improvements in the technologies, high throughput iterative designs and combinatorial methods are still cost prohibitive in a synthesis-based assembly due to the lack of reusable parts (Czar et al. 2009, Kosuri and Church 2014). As a result, a variety of DNA assembly standards based off a variety of methodologies have arisen in the previous 15 years to meet this demand.

Multipart DNA assembly has become a common tool in genetic engineering methodologies since the publication of the Gibson and Golden Gate assembly methods (Engler et al. 2009, Gibson et al. 2009). While complex genetic designs can be built with synthesis-based assembly techniques, these methods produce

no standardized reusable parts and require novel synthesis for every iteration of design. The most efficient assembly method for a given application is largely application dependent.

1.2   *Assembly methodologies and standards*

Commonly used assembly methods are listed in **Figure 1-1**, part of a survey performed to determine the enabling technologies in synthetic biology (Kahl and Endy 2013). According to this survey, synthesis based methods including Gibson assembly (Gibson et al. 2009) are the most commonly used assembly methods followed closely by BioBrick (Knight 2007), Gateway (Katzen 2007), and BglBrick (Anderson 2010) assembly standards. Each of these methods are described further in this section.

Gibson and synthesis based methods are ideal for many applications including genome engineering, but remain cost prohibitive for high throughput cloning and combinatorial assembly. Gateway cloning was one of the earliest standards developed, however it is primarily designed for moving full gene units. BioBrick and BglBrick standards offered the first modular DNA part assembly systems and libraries of basic parts (i.e. promoters, coding sequences), combining two parts at a time in a consuming binary fashion.
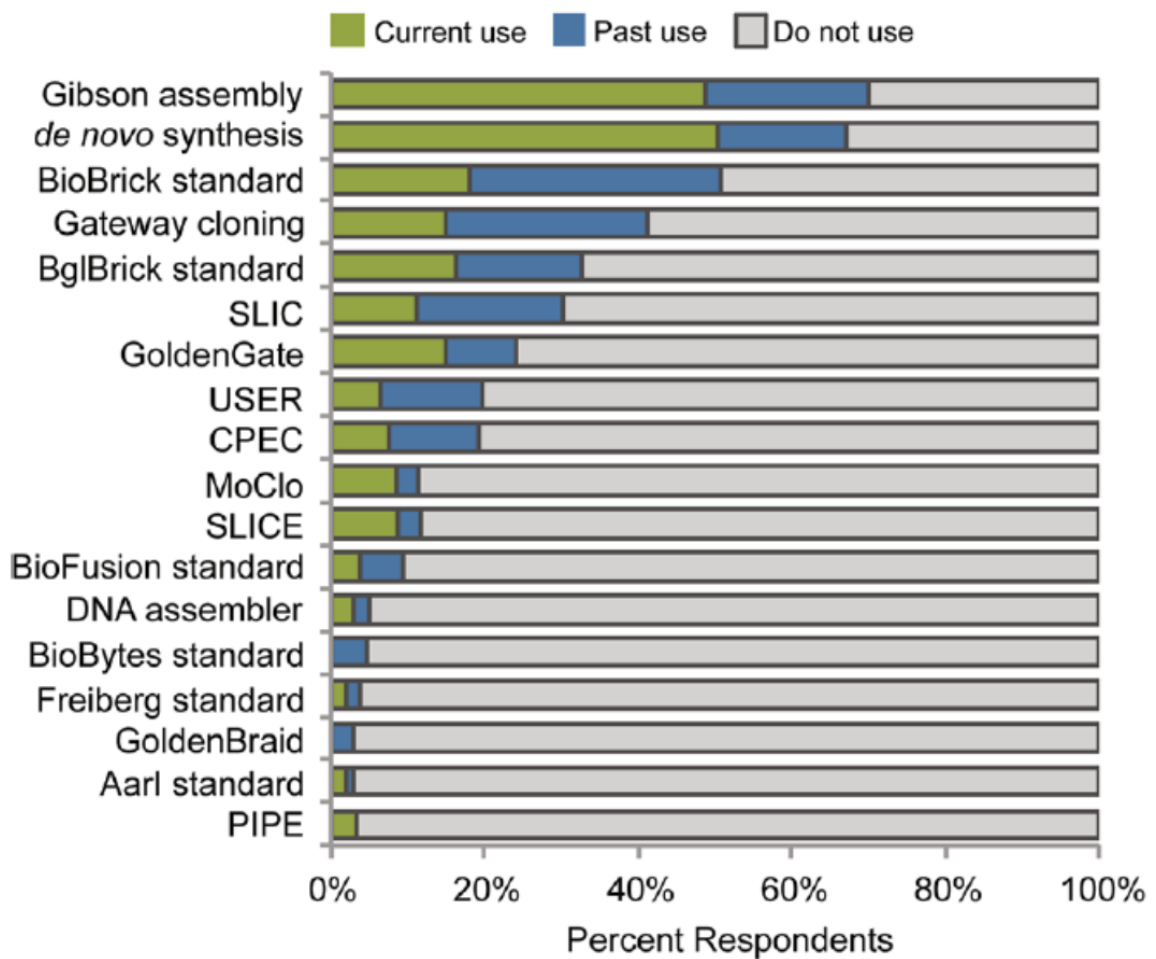
Figure 1-1 Survey of enabling technologies: DNA assembly methods and standards. *List of commonly used DNA assembly methods and standards as published by (Kahl and Endy 2013).*

Golden Gate and its derivatives, including Modular Cloning (MoClo), were the first multipart modular assembly standards available, allowing for the reliable assembly of up to 6 parts at a time (Engler et al. 2009, Weber et al. 2011). At the time this survey was taken, however, DNA part libraries were not yet widely available reducing the utility of these standards. In the past two years, part libraries have become increasingly available for eukaryotic MoClo platforms. This work describes the development of an *E. coli* MoClo assembly standard. With these part

4

libraries available, MoClo and related multipart assembly standards are rapidly becoming more widely used.

DNA assembly standards can largely be broken down by the underlining methodology employed. In general, all assembly methods utilize one or more of the following tools; restriction endonucleases (RE), homing endonucleases (HE), Type IIS restriction endonucleases (Type IIS), or annealing overlapping ends (OE) (Casini et al. 2015). Assembly standards are further categorized by the ability to hierarchically assemble large constructs, to reuse parts with a modular design, and to assemble more than two parts in a single reaction. The most effective practices often utilize multiple assembly methods.

1.2.1 Restriction endonucleases

Restriction endonuclease (RE) digest and ligation techniques, traditional molecular biology methods, have been used to cut and paste DNA since the 1970s (Cohen et al. 1973). RE methods generally only join two DNA parts and are sensitive to illegal restriction sites, recognition sequences which may occur naturally within the sequence of the part and interfere with the cloning reaction. Complex DNA engineering with traditional cloning is an art form onto itself, requiring extensive planning and verification of intermediate steps.

One of the first assembly standards developed, BioBricks (Knight 2007, Smolke 2009), uses standard RE methods with standard prefixes and suffixes each containing two restriction enzyme recognition sequences. BioBrick parts are

cut in a standard fashion resulting in a new part which contains the same prefix and suffix sequences while destroying the connecting restriction site. This allows for hierarchical assembly with a small number of enzymes.

The brick-like assembly format is derived from the use of isocaudomers, restriction enzymes which recognize different sequences yet produce the same overhangs. When ligated, these overhangs produce 'dead' sites, sequences which are no longer recognized by either enzyme. These sites are referred to as 'scars' and are known to influence circuit behavior and stability (Sleight and Sauro 2013). BioBricks provided a foundation for standard, modular parts and gained traction as the method adopted by the International Genetically Engineered Machines competition (iGEM) (Smolke 2009).

Though widely used, the BioBricks standard is limited by the binary nature of RE assembly methods. Parts are susceptible to illegal restriction sites and often require mutation of natural sequences for new parts. The binary nature of RE cloning assembly methods requires sequential step-by-step assembly increasing the time involved in the design-build-test cycle to produce iterative designs.

1.2.2 *Homing endonucleases*

Homing endonucleases (HE) function similarly to traditional restriction endonucleases but recognize longer sequences. This makes illegal sites unlikely. An early form of modular DNA assembly, HE methods first appeared with the Gateway cloning system commercialized by Invitrogen in the 1990s.

Gateway uses recombinases enzymes and long repetitive insertion sequences, *att* sites, to direct cloning. A concerted effort was made to create Gateway compatible plasmids carrying human, mouse, rat, and yeast open reading frames to support the research community (Gelperin et al. 2005, Nakajima et al. 2005, Giuraniuc et al. 2013). However, the *att* sites that grant modularity in Gateway cloning remain in the assembly as large scars between parts which can influence expression behavior (Chee and Chin 2015).

Multiple HE assembly standards have arisen in recent years including an adaptation of the BioBrick standard termed iBrick (Liu et al. 2014) and the HomeRun Vector Assembly System (HVAS) based on a modified Gateway assembly standard (Li et al. 2014). These assembly systems also leave scars between assembled parts (i.e. residual modified or inserted sequences). In HE standards, the scars are generally larger than the BioBrick standard due to the increased size of the recognition sequence.

1.2.3 Overlapping ends

In order to streamline complex DNA assembly beyond the capabilities of traditional RE cloning, various methods rely on the annealing of complementary sequences by using parts with long overlapping ends to assemble intermediate parts (Horton et al. 1989, Bitinaite 2007, Gibson et al. 2009, Quan and Tian 2009, Annaluru 2012). Overlapping ends methods can be seamless or may use linker sequences leaving scars between parts in order to increase modularity. These

assembly methods are often used in concert with computer aided design tools to optimize assembly and design oligos automatically (Hillson et al. 2012, Appleton et al. 2014).

Gibson assembly has become a common method in research laboratories with 48% of researchers surveyed reporting current use of the Gibson method (Kahl and Endy 2013). This scarless method allows for precision cloning of complex devices through synthesis of intermediate parts. However, though the cost of synthesis continues to decrease, Gibson and similar methods largely remain cost prohibitive for combinatorial assemblies and iterative design strategies (Kosuri and Church 2014).

1.2.4 Type IIS restriction endonucleases

An increasingly large collection of assembly standards relies upon Type IIS restriction endonucleases (Casini et al. 2015). These enzymes recognize non-palindromic sequences and cut at a specific distance up or downstream of the recognition sequence. The first method published using these enzymes, Golden Gate (Engler et al. 2009), uses PCR products or stored plasmid DNA and two enzymes in a simultaneous one-pot digestion-ligation reaction. By reusing DNA parts, minimizing synthesis costs, Golden Gate formats tend to be less costly than Gibson assemblies.

A variety of assembly standards have been published which utilize or modify the Golden Gate assembly strategy including Modular Cloning (Weber et al. 2011),

GoldenBraid (Sarrion-Perdigones 2011, Sarrion-Perdigones et al. 2013), BASIC (Storch et al. 2015), and yeast Golden Gate (yGG) (Agmon et al. 2015). Increasingly, combinations of assembly methods are being used to efficiently build complex DNA devices (Werner et al. 2012). This is the case with BASIC which combines overlapping ends and Type IIS enzymes to create an idempotent parts which are attached in a multipart single-tier assembly using linkers (Storch et al. 2015).

1.2.5 Modular cloning (MoClo) – Type IIS standard

Of the assembly methods developed to date, only one, MoClo, has significant part libraries available for eukaryotic (Weber et al. 2011, Duportet et al. 2014), yeast (Lee et al. 2015), plant systems (Engler et al. 2014) and now also for *E. coli* (Iverson et al. 2015). MoClo exploits user-defined overhangs specific to each part type such as a promoter or a coding sequence (CDS), thereby creating interchangeable DNA modules in the form of plasmids. This format allows for simple library propagation and combinatorial assembly from a library of reusable parts with reliable ligation of up to six DNA fragments in a one-pot reaction (Weber et al. 2011).

The original MoClo publication described a complex system using two color and three antibiotic selection in a rotating hierarchical assembly standard (Weber et al. 2011). These protocols called for high concentrations of DNA and long reaction times (>5 hours) and were cost prohibitive (~$10 per reaction). In order to

9

be effective as a high throughput assembly standard in *E. coli*, a more efficient protocol was needed.

Here we have optimized the published protocols, simplified the hierarchical assembly format, and created a library of reliable DNA parts for use in *E. coli*. The CIDAR MoClo Library described in Chapter 3 provides the first *E. coli* Type IIS compatible part library and is now available through Addgene (Iverson et al. 2015). This library enables rapid combinatorial assembly in bacteria and has practical applications in many fields including protein engineering, expression tuning, and library screening.

To demonstrate the utility of the CIDAR MoClo Library outside of the synthetic biology community, we have collaborated with the Tolan lab in the Biology department at Boston University to tune the protein expression of two variants of the rabbit aldolase protein and enable isolation of the heterotetramers formed when these two variants are co-expressed. To employ CIDAR MoClo in a synthetic biology context, in Chapter 5 we produced a bioelectronic pH-mediated genetic logic gate with DNA circuits built using MoClo and integrated with Raspberry Pi computers, Twitter, and 3D printed components. An overview of these three chapters can be seen in **Figure 1-2**.
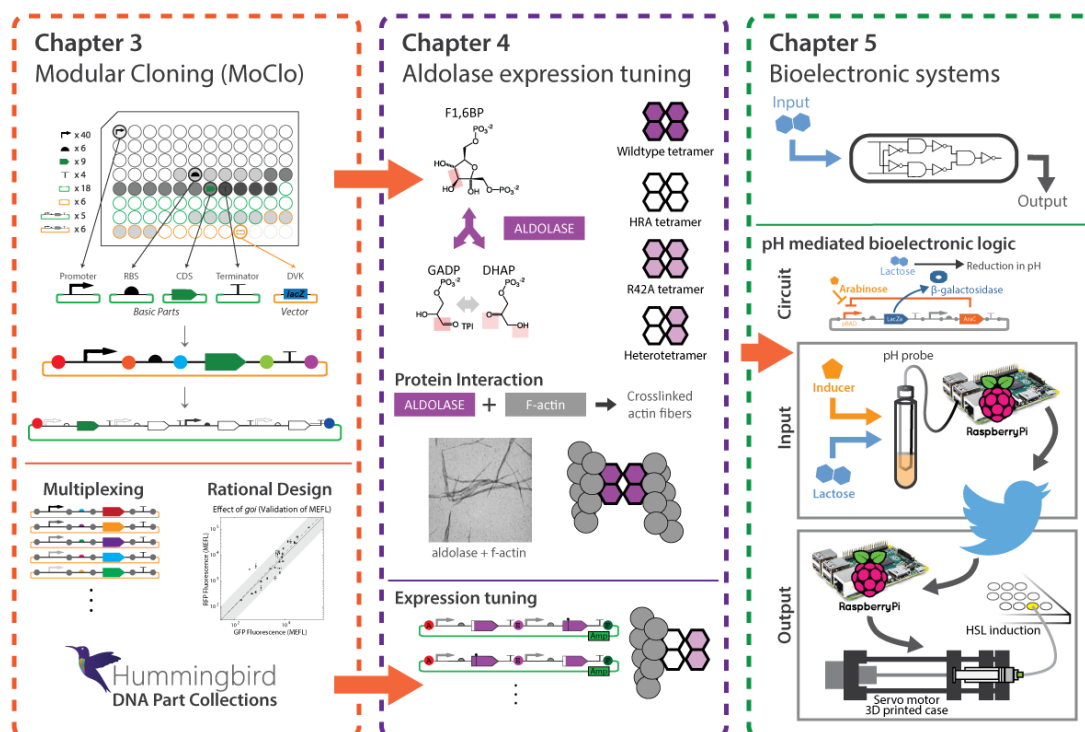
Figure 1-2 Overview of thesis. *Chapter 3 describes the development and optimization of the CIDAR MoClo assembly standard and E. coli part library. Chapter 4 demonstrates the utility of this library and part characterization data to rationally design a equimolar dual expression cassette for production of heterotetrameric alodase proteins. Chapter 5 explores applications of the CIDAR MoClo assembly standard in devleoping bioelectronic cellular logic systems and incorporates 3D printing and computer numerical control (CNC) milling.*

1.3   *Aldolase enzymatic and moonlighting functions*

Aldolase enzymes catalyzes an aldol reaction or its reverse. Two classes of aldolase have been identified with class I enzymes being cofactor-independent in catalyzing aldol reactions while class II enzymes employ a metal ion cofactor (Rutter 1964). Fructose-1,6-bisphosphate (Fru 1,6-P2) aldolase, referred to commonly and in this writing as simply 'aldolase', has a central role in fructose metabolism, glycolysis, and gluconeogenesis.

Aldolase catalyzes a reversible reaction that splits fructose 1,6-bisphosphate into dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (G3P) (Horecker et al. 1980)(**Figure 1-3a**). Aldolase isozymes are further classified by expression in different tissues. Aldolase A, found in muscle, and aldolase C, found in brain, show higher efficiency for Fru 1,6-P2 versus fructose 1-phosphate (Fru 1-P) as these enzymes are primarily involved in gluconeogenesis and glycolysis. Aldolase B however shows equal efficiency for the two substrates corresponding to the importance of fructose metabolism in the liver where it is predominantly expressed (Penhoet and Rutter 1971).

Monomeric and dimeric aldolase possess full catalytic activity (Beernink and Tolan 1994, Beernink and Tolan 1996), yet aldolase is found only as tetramers *in vivo* (Penhoet et al. 1967, Penhoet and Rutter 1971) with the exception of rare mutations which disrupt the dimer interface. Therefore the tetrameric form implies other functions for aldolase beyond the classical catalytic activity of the protein. With a dissociation rate of 10^{-25} M^3 (monomer-tetramer equilibrium) (Tolan et al. 2003), aldolase tetramers demonstrate an unusually stable formation compared to other glycolytic enzymes.

Penhoet and Rutter (1971) dismissed the possibility of allosteric regulation with steady-state kinetic analysis showing no cooperativity. In a study demonstrating that the structure of the dimer and that of one hemisphere of the tetramer do not differ significantly, the researchers conclude that stability of the

12

tetramer suggest aldolase evolved as a 'multimeric scaffold for non-catalytic functions' (Sherawat et al. 2008). Supporting this assertion, class I aldolases display a high degree of amino acid conservation at both the active site and subunit interface residues (Rottmann et al. 1984, Rottmann et al. 1987, Sygusch et al. 1987).

The tetrameric form of aldolase has two distinct subunit interfaces; the A interface is hydrophobic while the B interface is hydrophilic (Sherawat et al. 2008)(**Figure 1-3b**). Disruption of either interface by mutations at key sites leads to the dissociation of the tetramer into dimer form. These mutations do not, however, dramatically alter the tertiary structure of aldolase (Sherawat et al. 2008).

A clinically relevant mutation, D128G, disrupts the B interface and has been associated with a rare nonspherocytic hemolytic anemia (Kishi et al. 1987). Aldolase with D128G or other mutations of this residue maintain catalytic activity with reduced thermostability and are found in dimer form. Mutations of Gln-125 also disrupt the B interface while maintaining actin binding activity (Beernink and Tolan 1994).

Dimer formation can also be seen with the disruption of the A interface by mutations at Glu-224 and Arg-258. Active monomers can be produced by disrupting the remaining interface with a second mutation as seen with the Q125D/E224A double mutant (Beernink and Tolan 1996). However, these monomers are not seen *in vivo*.

13

1.3.1 Moonlighting functions

Consistent with the use of this tetramer as a scaffold, aldolase has been noted to moonlight in a variety of cellular processes. Many of these involve binding to F-actin, such as signal transduction, cell motility, and vesicle trafficking (Wang et al. 1996, Wang et al. 1997, Schindler et al. 2001, Ritterson Lew 2012). Known moonlighting interactions of actin are summarized in **Figure 1-3c**.

Of particular interest, aldolase interacts with the thrombospondin-related Apicomplaxan protein (TRAP) family of transmembrane proteins found in the protozoan parasites responsible for malaria and toxoplasmosis (Sibley 2003). Aldolase forms a bridge between TRAP proteins and the actin cytoskeleton. This bridge allows the parasites to move, enabling infection (Jewett and Sibley 2003).
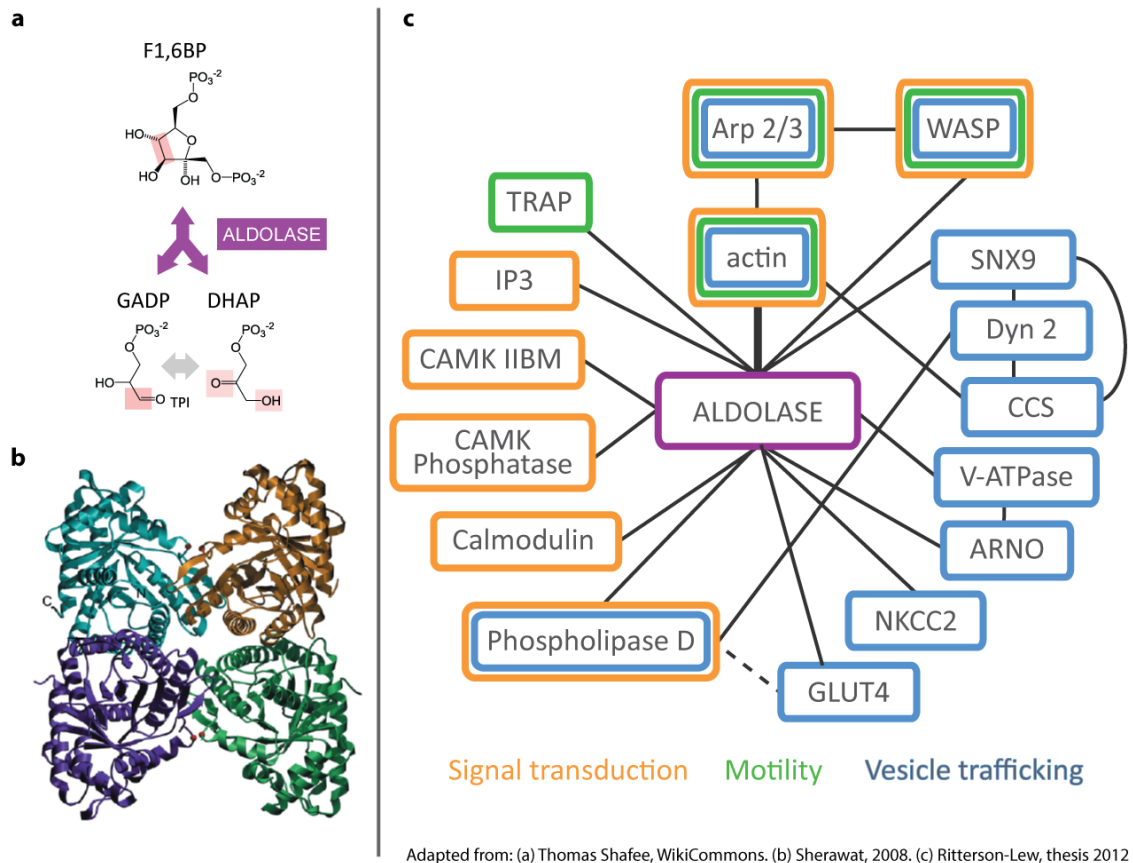
Figure 1-3 Aldolase function and moonlighting. *(a) Aldolase function in glycolysis/gluconeogenesis. (b) Aldolase tertiary structure, tetramer. (c) Known protein interactions involving aldolase suggesting roles in signal transduction, motility, and vesicle trafficking.*

Previous research on the interaction between aldolase and actin proteins has determined that aldolase enzymatic activity and actin binding activity are independent. Mutations that affect enzyme activity do not hinder actin binding and vice versa (Wang et al. 1996). In particular, an arginine residue was substituted to create the R42A mutant aldolase which is catalytically active yet has a 20-fold decrease in affinity for F-actin. Conversely, the D33S variant retains affinity for F-actin while being catalytically inactive (Wang et al. 1996).

1.3.2 Decorating F-actin

Aldolase forms a scaffold for actin filaments, facilitating crosslinking (Schindler et al. 2001, Jewett and Sibley 2003, Pirani et al. 2004, Pirani 2008) (**Figure 1-4a,b**). The predominantly tetrameric aldolase offers two actin binding surfaces, functioning to crosslink actin filaments in non-symmetrical arrays making characterization of the binding interface difficult. An aldolase oligomer which creates symmetrical arrays is needed to decorate actin for structural determination of the binding interface by electron microscopy.

The D128V aldolase mutant forms a dimer and functions as a scaffold for actin filaments without providing a second surface for crosslinking (**Figure 1-4c**). However, the symmetrical structure of the D128V dimer inhibits elucidation of the binding interface. A dimer-of-dimers in which only one dimer is able to crosslink could provide an appropriate aldolase-actin interaction for structure determination.

A His-tagged wildtype rabbit aldolase (HRA) in a dimer-of-dimer formation with the R42A actin-binding-deficient aldolase variant may provide the structure needed for structure determination (**Figure 1-4a,d**). However, the N-terminal His-tag appears to have a significant effect on protein expression such that duel expression of these otherwise nearly identical aldolase variants in the pETDuet expression plasmid produces HRA at a fraction of the level of R42A (Ho and Tolan, personal communication). High levels of protein expression of both variants is

needed to either produce predominately dimer-of-dimer heterotetramers *in vivo* or to create them *in vitro* through hybridizatoin.

With the *in vitro* method for producing heterotetramers, isolated protein of each variant can be denatured to form monomers then mixed to produce all possible heterotetramers compositions. These are then separated by charge with chromatography on a salt gradient. However, loss of protein during renaturation and charge separation steps have thus far prevented significant production of dimer:dimer aldolase tetramers for structure characterization experiments.

In order to produce predominately dimer-of-dimer heterotetramers *in vivo*, equimolar expression of these two variants is required in a single plasmid. Previous attempts at dual expression have resulted in an approximately 1:10 difference in expression of R42A to HRA due, apparently, to the N-terminal His-tag (Ho and Tolan, personal communications). To address this issue, we have used the CIDAR MoClo Library and previously gathered part characterization data to tune the expression of HRA and R42A and assemble a dual expression plasmid with approximately equimolar expression of HRA and R42A.
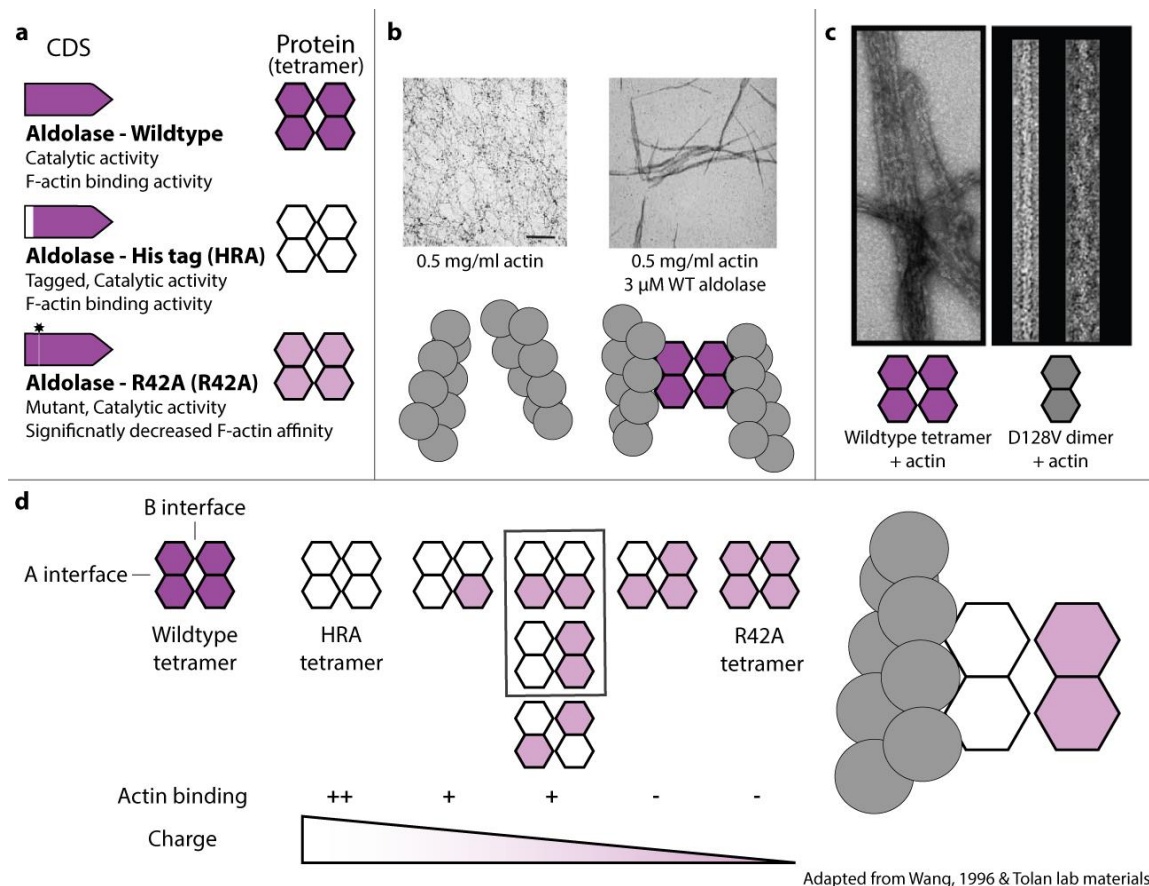
Figure 1-4 Aldolase HRA-R42A heterotetramers for decorating F-actin. *(a) Aldolase CDS and protein monomer icon legend is shown. Wildtype aldolase has both traditional catalytic activity and actin-binding ability. His-tagged wildtype (HRA) retains both catalytic and actin-binding activity. R42A aldolase mutant retains catalytic activity while demonstrating a 20-fold decrease in actin-binding activity. (b) Actin requires a scaffold upon which to structure filaments. Addition of µM wildtype aldolase provides a sufficient scaffold. As only a dimer is needed for scaffolding purposes, the stable wildtype tetramer creates crosslinked structures by providing two actin binding interfaces. (c) In comparison, the D128V disrupts the tetramer interface and this mutant exists predominately in dimer form. D128V dimers allow for actin scaffolding without crosslinking. However the symmetrical form of the dimer makes defining the actin binding interface difficult. (d) HRA-R42A heterotetramers may enable actin scaffolding while preventing crosslinking with a non-symmetrical aldolase complex. The charge difference between HRA and R42A should allow for the separation of different heterotetramers identifies. The species highlighted inside the black box are needed for actin-decorating. The bottom two-two heterotetramer is not expected to bind actin and will have no influence on downstream experiments.*

18

1.4   Rational design of genetic circuits

Rational design of biological devices, as described in 1.3.2, require well characterized parts and is facilitated by utilization of a standardized assembly format. Part libraries which better enable the rapid assembly of DNA devices are increasingly available and contain parts with known behaviors. In many cases while efforts have been made to compile libraries of parts, standard characterization methods and reliably predictably function parts are still lacking (Casini et al. 2015).

1.4.1   Enhancing rational design capabilities

In 2009, BIOFAB: International Open Facility Advancing Biotechnology (BIOFAB) was founded as a biological design-build-test facility. Researchers at the BIOFAB developed libraries of engineered constitutive and controllable promoters and translation elements (Mutalik et al. 2013a) and studied methods for accurately quantifying DNA part behavior (Mutalik et al. 2013b).

These translation elements, termed bicistronic designs (BCDs), decouple transcription from translation by introducing an intermediate cistron between the promoter and the gene of interest. A 30 a.a. peptide is translated under the control of a standard ribosome binding sequence (RBS), and contains a stop codon and secondary RBS overlapping the start codon of the gene of interest. By disrupting the secondary structure between the promoter and second RBS, the use of a BCD was shown reduce variability in fluorescence expression 10-fold, from 16% to 1.6%

19

(Mutalik et al. 2013a). Three of these BCD parts, described in Chapter 3, were included in the CIDAR MoClo library to enable more rational design.

Many methods of predicting protein expression and device behavior rely upon quantification of promoter and RBS sequences. Defining promoter sequences is often complicated as promoter components may be spread across hundreds of base pairs and controllable promoters (inducible/repressible) have an added factor of dose response to be considered. To approach this issue, researchers have proposed various methods for quantifying promoter (Beal et al. 2012, Mutalik et al. 2013b).

Most notable of these is the RBS calculator (Salis 2011) created to provide a simple, accessible tool to aid in synthetic genetic design. The RBS calculator is an online tool which calculates theoretical RBS strength as a factor of secondary structure (https://salislab.net/software/) (Salis et al. 2009, Salis 2011). Other tools exist for more complex aspects of genetic design including Cello (Cellular Logic), a tool for designing cascading cellular logic from abstract specifications (Nielsen et al. 2016).

1.5   Cellular logic; Boolean biology

Cellular logic refers to the use of engineered circuits in living systems which can compute signals. The Voigt Lab at MIT has previously published a variety of transcriptional logic gates using inducible promoters and repressor proteins. One of these designs is shown in **Figure 1-5**. In order to create an XOR gate, cells

carrying specific plasmids were spotted in close proximity on media with or without inducer.
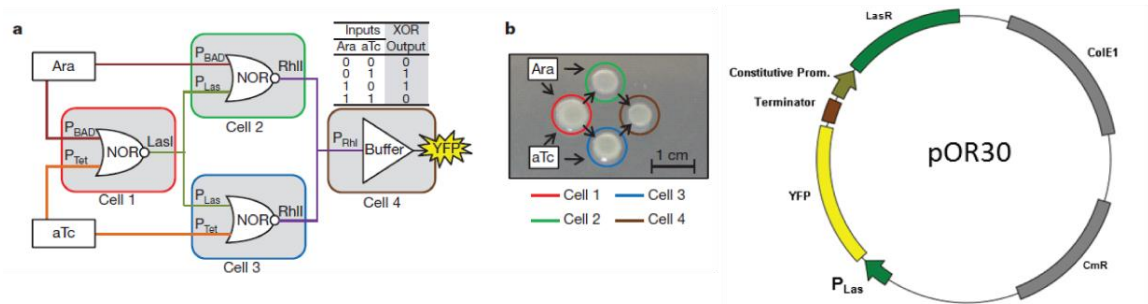
Figure 1-5 XOR gate as an example logic gate. (a) Four different cell strains are required to isolate components of the logic gate. Extruded signaling molecules from the first strain act as inducers for the second level of strains in Cell 2 and Cell 3. Only in the presence of one or the other inducer molecule is the output, YFP, expressed. (b) pOR30 is the plasmid used in Cell 4 and also used in Chapter 5.

These and other designs have demonstrated the potential for computational logic in cells (Friedland et al. 2009, Tamsir et al. 2011, Siuti et al. 2013). However, the need for compartmentalization of partial circuits in individual cells to prevent transcriptional cross talk is limiting. Some attempts to control crosstalk have looked to microfluidics to physically separate components of cellular logic.

If these cellular logic and microfluidic designs could be miniaturized, automated, and integrated with an electronic system, more functional computation could be seen. The Cello software tool aids the design of computational logic within cellular systems by defining cascades built from a library of genetic logic gates, fulfilling the need for automation (In press, Science 2015). Meanwhile, CNC milling, 3D printing and open source electronics systems could serve to provide the

miniaturization and bioelectronic integration aspects and lay the groundwork for more functional cellular computing.

1.6   Bioelectronic synthetic biology; CNC milling, 3D printing and adaptive research tools

Additive and subtractive manufacturing are increasingly affordable tools for synthetic biology. Tabletop 3D printers and computer numerical control (CNC) mills are commercially available in the range of $900-$2200. Free and open source software tools, such as OpenSCAD (www.openscad.org/) and Blender (https://www.blender.org/), facilitate the rapid design and iteration of increasingly complex microfluidics and custom tools (Gutierrez-Arenas 2015, Oxman 2015).

3D printing is particularly useful for designing custom devices and housings for electronic systems, as is the case with the syringe pump designed at MIT (Wijnen et al. 2014). This syringe pump provides an affordable alternative to commercially available injection devices and can be controlled electronically with open source devices. Examples of 3D printing and CNC milling in synthetic biology can be seen in **Figure 1-6**.

Integration of these tools, MoClo, 3D printing, CNC milling, and affordable electronics, opens the door for a wide range of bioengineered tools. Since all components of these designs are easily accessible and affordable, designs can be placed online and replicated anywhere in the world.
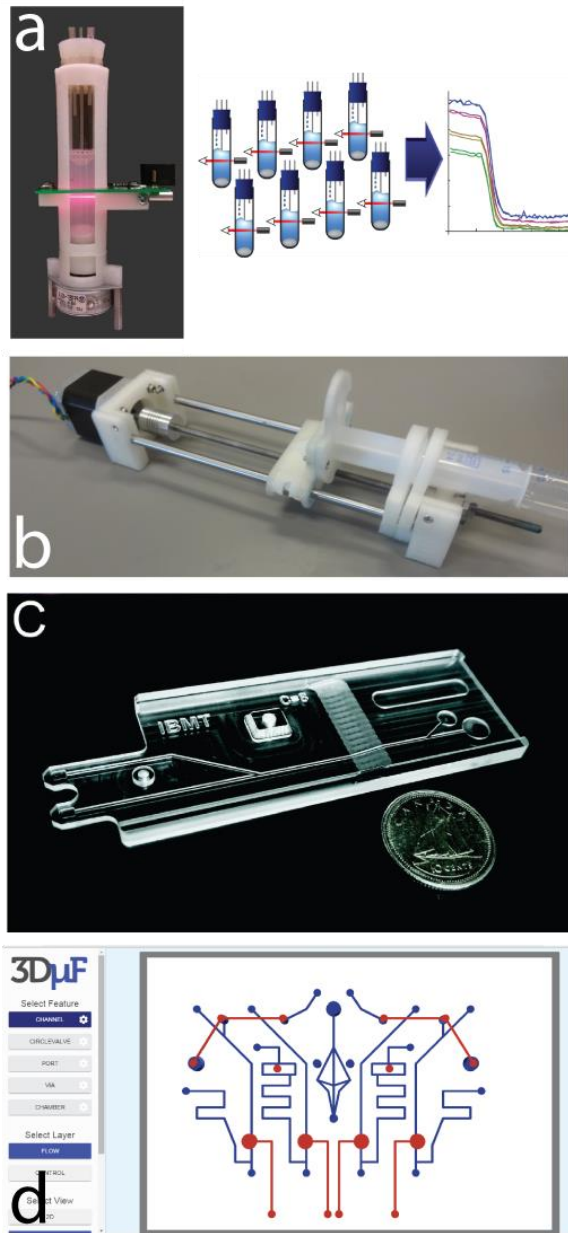
Adapted from Takahashi et. al, 2015; Wijnen et. al, 2014;
Guckenberger et. al, Lab on a Chip 2015; http://cidarlab.github.io/3DuF/

Figure 1-6 CNC and 3D printing in synthetic biology. *Examples of 3D printing and CNC milling in synthetic biology include (a) an open source 3d printed turbidostat, (b) an open source library of 3d printed syringe pumps, and (c) various CNC micromilling methods for microfluidics. Additionally, online design tools like 3DμF (cidarlab.github.io/3DuF/) enable intuitive simple design of 3 dimensional objects with a standard .stl file format which is easily interchanged between tools.*

2 CHAPTER 2 MATERIALS AND METHODS

2.1 General supplies and reagents

PCR performed using Phusion polymerase (NEB Cat #M0530L). Plasmid preparations and PCR cleanup were done using Qiagen Plasmid Mini Prep (Qiagen Cat #12125) and Qiagen Qiaquick PCR Purification Kit (Qiagen Cat #28106) following the manufacturers protocols. Transformations were performed with 6-15 µL of Alpha Select Gold Efficiency *E. coli* cells (Bioline USA Inc., Taunton, MA, USA) with a standard heat shock protocol (incubate on ice with DNA 5', 42°C 90", return to ice and recover in LB or SOC media).

2.2 Bacterial culturing and DNA handling

2.2.1 Bacterial strains, growth and storage conditions

For general cloning, Alpha Select Gold Efficiency DH5α *E. coli* cells (Bioline USA Inc., Taunton, MA, USA) were used in volumes of 5-10 µL per transformation. All plasmids are stored in these cells as glycerol preps and generally these are the cultures used in plasmid preps for sequencing and for use in MoClo reactions.

Cells were plated on appropriate antibiotic selective LB agar, supplemented with 80 µL of 20 mg/µL 5-bromo-4-chloro-3-indolyl ß-D-galactopyranoside (X-GAL) and 100 µL of 0.1M isopropyl-ß-D-thiogalactopyranoside (IPTG) (Zymo Research Corp., Irvine, CA, USA). Blue-white screening was used to select colonies which were grown overnight at 37°C in LB supplemented with the appropriate antibiotic.

In Chapter 5, Top10 (DH10B derived) cells were used in some experiments as noted in specific chapters.

2.2.2 *Plasmid nomenclature*

Plasmid names are abbreviations of the component parts as follows (**Figure 2-1**). pJ02B2Rm_AE(A) or pR404mGm_EF(K). In this format, "p" indicates "plasmid" verses describing the insert alone. The next three characters are generally a capital letter and two numbers identifying the promoter. In these examples, J02 = J23102 and R40 = R0040, both BioBrick parts. The next two or three digits identify the RBS or RBS type part which through this writing will be one of six parts. "2m", "3m", "4m" refer to the Weiss RBS parts from the BioBrick registry, B0032, B0033, and B0034; "m" indicates the parts have been modified and these parts contain the flanking BioBrick scar sequences to maintain the original BioBrick spacing. "B2", "B8" and "B12" refer to three of the BIOFAB bicistronic design elements, BCD2, BCD8 and BCD12 which are cloned as _BC parts and treated as RBS part types for cloning purposes though they actually contain a small coding sequence and second RBS site.

The characters between the RBS designation and the "_" indicate the coding sequence. Commonly these are seen as "Rm" or "Gm" indicating a red fluorescent protein (RFP) or green fluorescent protein (GFP) coding sequence. Here "m" indicates a modification from the original sequence usually pertaining to the removal of one or more illegal restriction endonuclease recognition sequences.

The capital two letter code after the underscore denote the 5' and 3' fusion sites, respectively. The B0015 terminator sequence is used for all plasmids discussed in these works and is not denoted in the plasmid name. The capital letter in parenthesis at the end indicates the antibiotic resistance of the vector used, either ampicillin (A), kanamycin (K) or chloramphenicol (C).
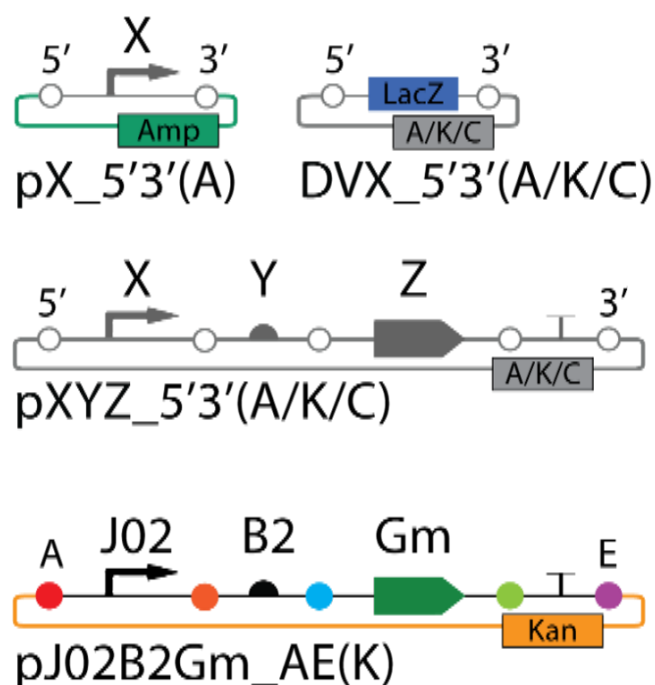
Figure 2-1 CIDAR MoClo plasmid nomenclature. Plasmids are given unique names with descriptive abbreviations to allow for simple identification. Grey part symbols are used to indicate a generic part type rather than a specific part.

When describing a generic part within a plasmid, an X is used as a wildcard to indicate this part is not specifically designated. For example, pJXB2Rm_AE would describe any plasmid or a collection of plasmids which have a J23 series

26

promoter, BCD2, and E1010m (RFP) in an _AE vector. Similarly, in all figures grey symbols indicate an undefined part. Using this nomenclature, pJ02B2Rm_AE(A) can be easily identified as a transcription unit containing J23102, BCD2, E1010m (RFPm), and B0015 in that order cloned into the DVA_AE vector which carries the ampicillin resistance gene, *bla*.

Lists of plasmids with component parts explicitly described are found in **Table A-7**.

2.2.3 Basic part and vector preparation & validation

Plasmid prep cultures were grown up overnight in 3-5 mL of LB with appropriate antibiotic or grown for 6-8 hours during the day in instances where same day sequencing was needed.

2.2.3.1   Basic parts general cloning strategy

Basic parts were either amplified from plasmids obtained from BioBricks (http://partsregistry.org/) or Addgene (Addgene.org, Cambridge, MA, USA) or, if the part was less than 35 bp, were assembled using annealing oligonucleotides containing the appropriate fusion sites and BsaI restriction sequences. PCR reactions were performed as above. Complimentary oligonucleotides were heated to 95°C and cooled at a rate of 0.5°C/minute to a final temperature of 45°C before being diluted for use in a MoClo reaction as above. This PCR product or annealed DNA was then used in a Level 0 or Basic Part MoClo reaction as described below and in **Figure 3-1**.

2.2.3.2   Destination vectors

The *lacZ* alpha fragment was PCR amplified pMJS2AF (donated by Dr.

Michael Smanski) and subsequently cloned into two backbones: DVA vectors used

pSB1A2, DVK vectors used pSB1K3. DNA containing the *lacZ* alpha fragment was

used as template for PCR reactions (20 fmol). PCR reactions with Phusion DNA

polymerase (NEB) following the manufacturer's protocol and were performed as

follows: denaturation at 95°C for 1', 30 extension cycles (98°C 20", 61°C 20", 72°C

20"), 5' extension at 72°C, hold 4°C. PCR product was gel extracted and cleaned

using QIAquick PCR Purification Kits (Qiagen Inc., Valencia, CA, USA).

LacZ PCR products and pSB1K3 and pSB1A2 vectors were digested with

SpeI enzyme (NEB) according to the manufacturer's protocol, cleaned up using

the QIAquick PCR Purification Kit (Qiagen). Ligation reactions used T4 DNA ligase

(NEB). Cells were plated on appropriate antibiotic selective LB agar, supplemented

with 80 µL of 20 mg/µL 5-bromo-4-chloro-3-indolyl ß-D-galactopyranoside (X-GAL)

and 100 µL of 0.1M isopropyl-ß-D-thiogalactopyranoside (IPTG) (Zymo Research

Corp., Irvine, CA, USA). Blue colonies were selected and sequence verified.

DVL2 carried an illegal BsaI site in the *bla* ampicillin resistance gene, which

was fixed with PCR mutagenesis as described in Weber et. al (2011) to create a

new vector, DVA.

2.2.3.3 Transcription Unit (Level 1) and Device (Level 2) assembly

Transcription units, generally consisting of a promoter, RBS part, CDS, and terminator assembled in a DVL1 / DVK vector, and Devices, multiple transcription units combined in a DVL2 / DVA vector, were assembled using the MoClo protocols as described in 2.3.

2.2.3.4 Plasmid isolation and sequence analysis

Plasmids were purified using the QIAprep Spin Miniprep kit (Qiagen) following the manufacturers' protocols. Sequences were validated with the standard BioBrick vector primers, Vector Forward (VF2) (5'-tgccacctgacgtctaagaa-3') and Vector Reverse (VR) (5'-attaccgcctttgagtgagc-3') primers. In larger constructs internal primers were used to sequence verify on a case by case basis. Sequences were analyzed using Benchling (www.benchling.com).

2.2.4 DNA parts construction

In total, ### parts were constructed for use in the CIDAR MoClo assembly standard. The initial CIDAR collection consisted of ## basic parts built in DVL0 (Chloramphenicol) vectors derived from pSB1C3 (BioBrick Part Registry) and ## destination vectors (DVL0, DVL1, and DVL2). These parts were used in optimizing MoClo protocols and building most of the transcription units described in this thesis. The final CIDAR MoClo standard and library replaces the DVL0 vectors with DVA vectors (DVL2 with repaired illegal site) and many basic parts were recloned into DVA vectors for the final published CIDAR library.

2.2.4.1 Initial 3 antibiotic CIDAR part collection

The initial CIDAR parts collection consisted of BioBrick parts converted to MoClo format in a destination vector carrying chloramphenicol resistance, DVL0, derived from the BioBrick vector pSB1C3. The parts were chosen for their utility and verified function within the iGEM Parts Registry.

All promoters and the B0015 terminator were cloned as four variations with appropriate fusion sites as described in **Figure 3-1**. Including both basic parts and destination vectors, 128 plasmids were cloned, verified, and archived as the initial CIDAR MoClo collection. All information about these parts was recorded and maintained in the CIDAR Registry, a detailed excel file. Approximately 18% of these parts were cloned by the 2012 and 2013 BostonU iGEM teams and other members of the CIDAR lab. A list of these parts is found in **Table A-6**. Primer design used in assembly are noted in **Table A-4**.

2.2.4.2 Final CIDAR collection and CIDAR MoClo Library

In order to increase efficiency of cloning and increase plasmid yield, all of the basic parts to be used in the CIDAR MoClo Library were converted from the DVL1 (Chloramphenicol, CAM) backbone to the DVA (Ampicillin, AMP, backbone with repaired illegal site). DVA provided higher plasmid DNA concentrations (average ~100 ng/µL verses ~20 ng/µL) and simplifies the library system by removing CAM plasmids from the rotation. Level 0 (Basic Parts) and Level 2

(Devices) both use DVA vectors minimizing the number of assembly vectors required.

While preparing the CIDAR MoClo Library, 6 coding sequence parts were cloned a second time to replace incorrect sequences or to remove illegal sites. A total of 17 new DVA vectors were created as well as the DVK_AF vector to allow for vector context comparison of two-part Devices. Additionally, empty DVA and DVK vectors were cloned, closed on a SpeI site and lacking the LacZ cassette and fusions sites to enable the production of new vectors.

2.3   CIDAR MoClo assembly methods

2.3.1   Original and Optimized MoClo Protocols

Initial MoClo assembly reactions followed the previously published protocol (Weber et al. 2011), using 45 cycles alternating between 37°C 2' and 16°C 5' followed by 50°C 5' and 80°C 10' with 40 fmol of each DNA part, 10 U of restriction enzyme, 10 U of T4 ligase (Promega) with 1x Promega T4 ligase buffer in a 20 µL volume.

In developing optimal protocols, various reaction conditions were tested. The following components were added to a 0.2 mL tube: 10-60 fmol of each DNA component with up to six components total using equimolar PCR product or previously made MoClo DNA parts, and the appropriate Destination Vector, 10-50 U of BsaI or BbsI (NEB), 5-50 U of T4 DNA ligase (Cat #M1794, Promega, Madison, WI, USA or #M0202 NEB), 1 X T4 DNA Ligase Buffer (Promega or NEB),

31

and sterile, deionized water to at total volume of 10-60 µL. Reactions were performed using the following parameters: 15-40 cycles (37°C 1.5-3', 16°C 3-5'), followed by 50°C for 5' and 80°C for 10' then held at 4°C or -20°C until transformed.

By comparing efficiencies and costs, the optimal protocol was determined as follows: 10 fmol of each DNA component, 10 U of BsaI or BbsI, 20 U of T4 Ligase (NEB or Promega) and 1x T4 DNA Ligase Buffer (Promega only) to a final volume of 10-20 µL depending on concentration of DNA parts. Lower volumes are preferred.

2.3.2 Multiplex MoClo Protocols

MoClo reactions were prepared as above with the following differences: The multiplex part type(s) was added such that the total concentration of that type was equimolar to the other part types. When multiplexing less than 6 of any one part type, samples were pipetted individually. To accommodate accurate measurements, multiplex reactions were performed in 20 µL volumes with 60 fmol of each part type. For example, in a reaction using 6 different promoters, 10 fmol each of these promoters was used along with 60 fmol of each other part. For larger multiplex examples all iterations of a given part type are mixed in equimolar ratio prior and added as one mixed part.

2.4   Fluorescence analysis

2.4.1 Flow cytometry

All fluorescent expression devices were characterized using a BD LSRFortessa SORP flow cytometer. RFP fluorescence was measured using a solid-state Coherent Sapphire 561 nm laser at 100 mw strength with a PE-Texas Red 610/20 filter. GFP fluorescence was measured using a solid-state Coherent Sapphire 488 nm laser at 200 mw strength with a FITC 530/30 filter. Clonal colonies were grown overnight on agar with antibiotic and were used to inoculate 200 µL LB broth (Sigma-Aldrich) with the appropriate antibiotic in sterile 96-deep well plates in triplicate grown for 16 hours,37°C shaking, 300 rpm. Cells were then diluted 100-fold into 200 µL of sterile phosphate buffered saline (PBS) in 96-well round bottom plates before measurement.

2.4.2 Plate reader for optical density and fluorescence measurements

A Tecan SpectraFluor Plus plate reader with Magellen v6.6 software was used for optical density (OD or OD_{600}) measurements and to measure fluorescence in culture. Measurements were performed in 96-well plates with 200 µL volumes. For OD, default settings were used along with blank wells to provide background subtracted data. For fluorescence, top and bottom read options were used as noted with manual gain and filter settings as follows: RFP – ex. 580 / em. 635, gain 87. GFP – ex. 485 / em. 535, gain 89.

2.4.3 Two color control array and color model

In order to create an accurate color model, a series of high and low expression RFP and GFP transcription units, described here as two-color controls were evaluated by flow cytometry as in 2.4.1. These TUs were assembled in _AE and _EF DVK, transferred to DVA, and assembled into dual expression devices in _AF DVA vectors. This series was designed to provide both the _AE and _EF context for each TU as well as both variants of dual expression plasmids for each. A subset of eBFP2 transcription units were also assembled as part of this series for comparison. A complete list of these two-color control plasmids is included in Table 2-1.

The two-color controls were measured by flow cytometry in triplicate, overnight cultures diluted 1:100 in PBS and measured as described in 2.4.1. In order to normalize to fluorescein, one color used in the model must register with the FITC filter set (generally either YFP or GFP). For this color model, designed for use primarily with red and green fluorescence, GFP and RFP constructs were used to build the color model.

To enable normalization, expression of each fluorescent reporter needed to be high enough to register cleanly on all appropriate fluorescence channels. This initial flow cytometry data identified J02B2 as the promoter/RBS combination ideal for use as the color model due to its consistently high level of expression in different

constructs. A small number of eBFP2 transcription units were made to test ability

of a RFP:GFP color model to predict expression in another filter set.

Plasmid ID	2CC#	Construct
pJ02B2Rm_AE(K)	2CC01	J23102:BCD2:RFP:B0015
pJ02B2Gm_EF(K)	2CC02	J23102:BCD2:GFP:B0015
pJ02B12Rm_AE(K)	2CC03	J23102:BCD12:RFP:B0015
pJ02B12Gm_EF(K)	2CC04	J23102:BCD12:GFP:B0015
pJ024mRm_AE(K)	2CC05	J23102:B0034m:RFP:B0015
pJ024mGm_EF(K)	2CC06	J23102:B0034m:GFP:B0015
pJ022mRm_AE(K)	2CC07	J23102:B0032m:RFP:B0015
pJ022mGm_EF(K)	2CC08	J23102:B0032m:GFP:B0015
pJ14B2Rm_AE(K)	2CC09	J23114:BCD2:RFP:B0015
pJ14B2Gm_EF(K)	2CC10	J23114:BCD2:GFP:B0015
pJ14B12Rm_AE(K)	2CC11	J23114:BCD12:RFP:B0015
pJ14B12Gm_EF(K)	2CC12	J23114:BCD12:GFP:B0015
pJ144mRm_AE(K)	2CC13	J23114:B0034m:RFP:B0015
pJ144mGm_EF(K)	2CC14	J23114:B0034m:GFP:B0015
pJ142mRm_AE(K)	2CC15	J23114:B0032m:RFP:B0015
pJ142mGm_EF(K)	2CC16	J23114:B0032m:GFP:B0015
pJ02B2Rm:Gm_AF(A)	2CC18	J02B2Rm:J02B2Gm
pJ02B12Rm:Gm_AF(A)	2CC19	J02B12Rm:J02B12Gm
pJ024mRm:Gm_AF(A)	2CC20	J024mRm:J024mGm
pJ022mRm:Gm_AF(A)	2CC21	J022mRm:J022mGm
pJ14B2Rm:Gm_AF(A)	2CC22	J14B2Rm:J14B2Gm
pJ14B12Rm:Gm_AF(A)	2CC23	J14B12Rm:J14B12Gm
pJ144mRm:Gm_AF(A)	2CC24	J144mRm:J144mGm
pJ142mRm:Gm_AF(A)	2CC25	J142mRm:J142mGm
pJ02B2Gm_AE(K)	2CC38	J23102:BCD2:GFP:B0015
pJ02B2Rm_EF(K)	2CC39	J23102:BCD2:RFP:B0015
pJ024mGm_AE(K)	2CC40	J23102:B0034m:GFP:B0015
pJ024mRm_EF(K)	2CC41	J23102:B0034m:RFP:B0015
pJ02B2Gm:Rm_AF(A)	2CC42	pJ02B2Gm:pJ02B2Rm
pJ024mGm:Rm_AF(A)	2CC43	pJ024mGm:pJ024mGm
pJ02B2B_AE(K)	2CC44	J23102:BCD2:eBFP2:B0015
pJ02B2B_EF(K)	2CC45	J23102:BCD2:eBFP2:B0015
pJ02B2Rm:B_AF(A)	2CC46	pJ02B2Rm:pJ02B2B

Plasmid ID	2CC#	Construct
pJ02B2Gm:B_AF(A)	2CC47	pJ02B2Gm:pJ02B2B
pJ02B2B:Gm_AF(A)	2CC48	pJ02B2B:pJ02B2Gm
pJ02B2B:Rm_AF(A)	2CC49	pJ02B2B:pJ02B2Rm
pJ02B2Rm_AE(A)	2CC01(A)	J23102:BCD2:RFP:B0015
pJ02B2Gm_EF(A)	2CC02(A)	J23102:BCD2:GFP:B0015
pJ02B12Rm_AE(A)	2CC03(A)	J23102:BCD12:RFP:B0015
pJ02B12Gm_EF(A)	2CC04(A)	J23102:BCD12:GFP:B0015
pJ024m1Rm_AE(A)	2CC05(A)	J23102:B0034m:RFP:B0015
pJ024m1Gm_EF(A)	2CC06(A)	J23102:B0034m:GFP:B0015
pJ022mRm_AE(A)	2CC07(A)	J23102:B0032m:RFP:B0015
pJ022mGm_EF(A)	2CC08(A)	J23102:B0032m:GFP:B0015
pJ14B2Rm_AE(A)	2CC09(A)	J23114:BCD2:RFP:B0015
pJ14B2Gm_EF(A)	2CC10(A)	J23114:BCD2:GFP:B0015
pJ14B12Rm_AE(A)	2CC11(A)	J23114:BCD12:RFP:B0015
pJ14B12Gm_EF(A)	2CC12(A)	J23114:BCD12:GFP:B0015
pJ144m1Rm_AE(A)	2CC13(A)	J23114:B0034m:RFP:B0015
pJ144m1Gm_EF(A)	2CC14(A)	J23114:B0034m:GFP:B0015
pJ142mRm_AE(A)	2CC15(A)	J23114:B0032m:RFP:B0015
pJ142mGm_EF(A)	2CC16(A)	J23114:B0032m:GFP:B0015
pJ02B2Gm_AE(A)	2CC38(A)	J23102:BCD2:GFP:B0015
pJ02B2Rm_EF(A)	2CC39(A)	J23102:BCD2:RFP:B0015
pJ024mGm_AE(A)	2CC40(A)	J23102:B0034m:GFP:B0015
pJ024mRm_EF(A)	2CC41(A)	J23102:B0034m:RFP:B0015
pJ02B2Gm:Rm_AF(K)	2CC42(K)	pJ02B2Gm:pJ02B2Rm
pJ024mGm:Rm_AF(K)	2CC43(K)	pJ024mGm:pJ024mGm

Table 2-1 Table of Two Color Control (2CC) plasmids used to develop color model. These plasmids can also be found online in the public Benchling directory.

2.4.4 TASBE tools

Flow cytometry data was converted from arbitrary units to compensated Molecules of Equivalent Fluorescein (MEFL) using the TASBE characterization method (Beal et al. 2012). An affine compensation matrix is computed from single color and blank controls: RFP alone (J23102:BCD2:E1010m:B0015, abbreviated

as pJ02B2Rm, in _AE and EF DVAs), GFP alone (J23102:BCD2:E0040m:B0015 abbreviated as pJ02B2Gm, in _AE and EF DVAs), RFP:GFP (pJ02B2Rm:J02B2Gm in _AF DVAs) as well as the reciprocal GFP:RFP (pJ02B2Gm:J02B2Rm in _AF DVAs) together and untransformed DH5 Alpha Select *E. coli* cells (Bioline), respectively. FITC measurements (for GFP) are calibrated to MEFL using SpheroTech RCP-30-5-A beads (Spherotech 2001). An estimated mapping from RFP measured in the PE-Texas Red channel to equivalent FITC is computed from transformation of constitutive co-expression of RFP and GFP expressed together (RFP:GFP as pJ02B2Rm:J02B2Gm as _AF in DVA, GFP:RFP as pJ02B2Gm:J02B2Rm as _AF in DVA); RFP measurements are translated to MEFL by first mapping to estimated equivalent FITC. Geometric statistics are then computed over data in MEFL units.

2.5 CIDAR MoClo Library preparation

2.5.1 Glycerol storage

All plasmids made for use in CIDAR MoClo and specifically for the published CIDAR MoClo Library were sequence verified and stored at -80°C as individual glycerol preps in 300-1mL of LB with 25% glycerol. In preparation for submitting the CIDAR MoClo Library to Addgene, and to provide backup stocks for internal use, the glycerol stocks were used to inoculate fresh plates and single colonies from theses plates were used to inoculate 96 well deep culture plates (1mL max volumes). These plates were grown overnight in 500 µL of LB with appropriate

antibiotic and frozen after addition of 25% glycerol, mixed by shaking at 900 rpm sealed with PCR sealing film.

2.5.2 Benchling

All plasmids described in this text can be found in public folders on the Benchling lab management tool at https://benchling.com/siverson/.

2.5.3 CIDAR-ICE

All plasmids in the CIDAR MoClo Library are made public in the CIDAR Inventory of Composable Elements (ICE) web based inventory (http://cidar-ice.org). The ICE platform was initially developed by the Joint BioEnergy Institute (JBEI, www.jbei.org) to facilitate the sharing of DNA parts between research groups and facilities (Ham et al. 2012).

2.6   Aldolase Methods

2.6.1 Plasmid design of aldolase expression cassettes

HRA_CD and R42A_CD were cloned as basic parts with the DVA_CD vector as described previously using PCR templates provided by the Tolan lab. Extrapolating from the RFP and GFP data presented in section 3.2.3 (**Figure 3-3** and **Figure 3-4**), three _AB and two _EB plasmid configurations were chosen which were predicted to produce similar high expression of the gene of interest (*goi*). These plasmids were chosen due to their low variability in replicates and similarity between green and red expression cassettes and are described in detail in 4.2.1 and **Table 4-1** and at https://benchling.com/siverson/. Each chosen

combination of promoter, RBS part, vector and aldolase variant (HRA_CD or R42A_CD) were assembled using standard MoClo protocols as described in section 2.3.1.

The HRA transcription units were designated pJ06B12HRA_AE(K), pJ07B2HRA_AE(K), pJ07B12HRA_AE(K), pJ02B2HRA_EF(K), pJ02B12HRA_EF(K) and nicknamed Tolan-1 through Tolan-5 respectively. The HRA transcription units were designated pJ06B12R42A_AE(K), pJ07B2R42A_AE(K), pJ07B12R42A_AE(K), pJ02B2R42A_EF(K), pJ02B12R42A_EF(K) and nicknamed Tolan-6 through Tolan-10 respectively. Tolan-1 and Tolan-6 differ only in by *goi*, and likewise for the rest of this series (2 and 7, 3 and 8, etc.). Sequences were verified using VF2 and VR primers as described previously.

Each combination of pJXBXHRA_AE and pJXBXR42A_EF were combined to create the dual expression cassettes expressing HRA upstream of R42A. Likewise, each combination of pJXBXR42A_AE and pJXBXHRA_EF were also created for a combined total of twelve _AF(A) duel expression plasmids (**Figure 4-1**). These plasmids are designated as follows; pJXBXHRA:pJXBXR42A_AF(A) where pJ6B12HRA:pJ02B2R42A_AF(A) would be described as Tolan-1-9 as it consists of Tolan-1 and Tolan-9 transcription units. These duel expression plasmids were provided to the Tolan lab for protein isolation and quantification.

Due to the nearly perfect sequence homology between HRA and R42A, varying only with the 5' His tag and a single internal amino acid substitution, separation and quantification by antibody methods is difficult. To aid in analyzing the relative expression of each component of the duel expression cassettes, each Tolan-1 through-10 were subcloned into DVA vectors to provide Tolan-1(A), Tolan-2(A), etc.

Upon sequence verification, all of these plasmids were provided to the Tolan lab for protein isolation and quantification.

2.6.2 *Plasmid design of aldolase fluorescent protein fusion plasmids*

Concurrently, fluorescent fusion protein (FP) coding sequences were constructed by linking the first 36 nucleotides of each aldolase variant to the full CDS of E0040m (GFP) and E1010m (RFP) as shown in **Figure 4-4** and annotated online at https://benchling.com/siverson/. These fusion protein coding sequences, HRA:GFP_CD, HRA:RFP_CD, R42A:GFP_CD, and R42A:RFP_CD were used to assemble the fluorescent of expression plasmids described in 2.6.1 in order to use fluorescence as a proxy for aldolase protein expression.

2.6.3 *Fluorescence measurement of fusion proteins*

Fluorescence expression of aldolase:FP fusion expression plasmids was recorded using a Tecan plate reader with Magellen software as described in 2.4.2.

2.6.4 Protein expression analysis

Western blots using α-aldolase and α-His antibodies were performed on lysates from each of these dual expression clones and compared to isolated protein standards for HRA and R42A (Tolan lab).

2.7 Inducible pH sensor methods

2.7.1 Bacterial strains, growth conditions

Initial experiments were done using DH5α cells (Bioline, Alpha Select Gold Efficiency *E. coli*). Later experiments used TOP10 cells as noted in the text. Plasmids were built using MoClo as described previously and are described in Chapter 5. Cultures used for single time point pH assays were grown under standard conditions in 14mm plastic culture tubes, 3 mL LB with appropriate antibiotic, 37°C 300 rpm, and pH was tested at defined time points.

Final experiments and logic gate tests were grown as described with the addition of continuous monitoring. Briefly, the pH probe was placed in the culture (1.5 mL) while shaking, sealed with parafilm, with wires placed carefully along the seam of the incubator to connect to the Raspberry Pi placed nearby. pH measurement method are described below. Culture for inducible YFP expression with the pOR30 plasmid was grown overnight under standard conditions. For YFP induction, 100 µL of overnight culture was aliquoted to a well in a 96 well plate along with control wells containing 100 µL of overnight growth of appropriate cultures and 100 µL of fresh LB and sealed with breathable film. Upon induction

as described below, 100 μL of inducer diluted in LB was injected into the sealed well. The plate was incubated at 37°C, 900 rpm for 2 hours.

2.7.2 Plasmid designs

DVK_AE was used as a control for LacZ expression. pR10LacZDV_AE was built to mimic DVK with modular parts and inducible Lac promoter (BBa_R0010). pBADLacZDV:J024mC80_AF contains two transcription units which express LacZ and AraC to allow for inducible expression with arabinose. pTetLacZDV:J024mC40_AF was built to constitutively express the tetracycline repressor, TetR derived from the BioBrick part C0040, to allow for induction of LacZ expression via the pTet promoter and the inducer molecule anhydrous tetracycline (aTc).

pOR30 was a gift from the Voigt lab at the Massachusetts Institute of Technology (MIT). This plasmid expresses YFP under the control of the pLas promoter to enable inducible expression with N-3-oxo-dodecanoyl-L-Homoserine lactone (HSL) (Cayman Chemicals, Cat #10007895, CAS 168982-69-2).

2.7.3 Evaluation of pOR30

Induction curves were performed with concentrations of HSL ranging from 1 mM to 100 mM over a period of two hours with a 1:2 dilution of overnight culture and fresh LB. A final concentration of 10 mM was determined to be optimal with a 1-2 hour incubation time at 37°C, 900 rpm.

2.7.4 pH measurements with Raspberry Pi

A Raspberry Pi processor was programmed language to interpret the signal from an Atlas Scientific pH probe (Cat #ENV-40-pH) with an Ezo pH Circuit (Cat #EZO-pH)after calibration with standard pH calibration solutions (pH 4, 7, and 10) (Cat #KIT-101P). PuTTY terminal (www.putty.org) was used to control the Raspberry Pi interface.

2.7.5 pH mediated induction of YFP

Upon sensing a target pH value the Raspberry Pi connected to the pH probe was programmed to submit a Twitter post to the @CIDARlab Twitter feed using the Twitter API. This tweet functioned as the input signal for a second Raspberry Pi connected to a servo motor and a 3D printed syringe pump calibrated to dispense 100 µL (+/- 3 µL) through 7/32" medical grade tubing and 14G needle point. Upon detecting the Tweet from Raspberry Pi 1, Raspberry Pi 2 activated the servo motor to inject 100 µL of 20 µM homoserine lactone in LB into a sealed well of a 96 well plate to induce the expression of YFP in cells carrying the pOR30 plasmid (Voigt Lab, MIT).

2.8 CNC milling and 3D printing

All 3D printing, done with an Ultimaker 2 printer and stock materials, and CNC milling using an Othermill CNC mill and polycarbonate (PCA) stock, was performed by Ryan Silva, at Boston University. An open source .stl design for a syringe pump was modified for use in Chapter 5. The original designs were

developed at MIT (Wijnen et al. 2014) and files can be found online (http://matter.media.mit.edu/tools/details/3d-printed-syringe-pump).

A 5x5 array of wells, styled after a 96-well plate, was CNC milled in PCA designed to fit into the D8-H12 region of a 3D printed 96-well plate frame. The 5x5 well array was sterilized with alcohol and bleach before inoculation of cultures in LB with appropriate antibiotic. The array was sealed with breathable culture film, placed inside the 96-well plate frame and incubated at 37°C, 900 RPM in a shaking plate incubator.

3 CHAPTER 3 IMPROVED MULTIPART, MODULAR, DNA ASSEMBLY (CIDAR MOCLO) AND NEW *E. COLI* COMMON PART LIBRARY

3.1 *Introduction*

MoClo, which was developed by Weber et. al in 2011, relies upon Type IIS restriction enzymes (BbsI and BsaI). Each recognize a 6 base pair non-palindromic sequence and cut at a specified distance from that recognition sequence resulting in overhanging 4 bp fusion sites (**Figure 3-1**). Due to the nature of these cutting patterns, the 4 bp fusion sites can be any four nucleotides and can thus be defined as standard sites in this assembly method. For example, CDS parts are flanked with '-AATG-' at the 5' end and '-AGGT-' on the 3' end making all parts interchangeable. The restriction recognition sites are placed and oriented such that the digest product ends in these specific four base pair overhangs and no longer contains the restriction enzyme recognition sequence. Once ligated it cannot be recut allowing for a hierarchical multipart assembly.

The original MoClo protocol allowed for the reliable assembly of up to six parts in a five hour digestion-ligation reaction with large reagent volumes and rotating use of three antibiotics and two color selection modules. Although this provided a substantial improvement in modularity compared to Golden Gate, the reaction conditions were still time and cost-prohibitive for most applications. Here we introduce a modified protocol (**Table A-2**) that, while maintaining the same capacity for assembling multiple modules and a >95% cloning efficiency, also

reduces reaction time from five hours to 90 minutes and lowers reaction costs by 85% while simplifying the hierarchical assembly standard.

Two variations of the CIDAR MoClo system are described in this chapter. The first, the 3 antibiotic selection (3Ab MoClo) system was initially used in the CIDAR lab to build many of the plasmids discussed throughout this chapter. In order to streamline and optimize the CIDAR MoClo standard, this system was later converted to a two antibiotic selection system which and published as the CIDAR MoClo assembly standard and library.

3.2   Results

3.2.1   3Ab MoClo parts collection, 3 antibiotic assembly standard

To create destination vectors, pSB1C3, pSB1K2 and pSB1A2 BioBrick vectors were modified by inserting a LacZ alpha fragment expression cassette flanked with BbsI and BsaI recognition sequences and MoClo fusion sites into a SpeI site. Flanking sequences are shown in **Table A-5**. Adapted from the original published MoClo protocols, fusion sites were chosen as shown in **Figure 3-1**. An array of commonly used BioBrick parts were adapted for use in a MoClo standard by cloning into the pSB1C3 derived DVL0 vectors.

The 3Ab MoClo parts collection and standard used an alternating series of vectors (**Figure 3-1**). Basic parts were provided in a Destination Vector Level 0 (DVL0) derived from BBa_pSB1A2, with BsaI sites flanking the part and fusion sites. These basic parts combine to create a simple transcription unit (TU)

46

consisting of promoter:RBS:CDS:terminator in a Level 1 kanamycin vector (DVL1) derived from BBa_pSB1K3 with BbsI sites flanking the newly constructed TU. These units can be further combined into multi-TU devices in the same manner using a Level 2 ampicillin vector (DVL2).

We created a standard four-part structure that allows users to easily assemble a genetic device consisting of a promoter, RBS, CDS, and terminator. This basic format provides a solid foundation for high throughput assembly while remaining adaptable to addition of new parts (i.e. new fusion site combinations) and part types. This 3Ab system (Haddock et al. 2015) was later adapted to become the CIDAR MoClo assembly standard.
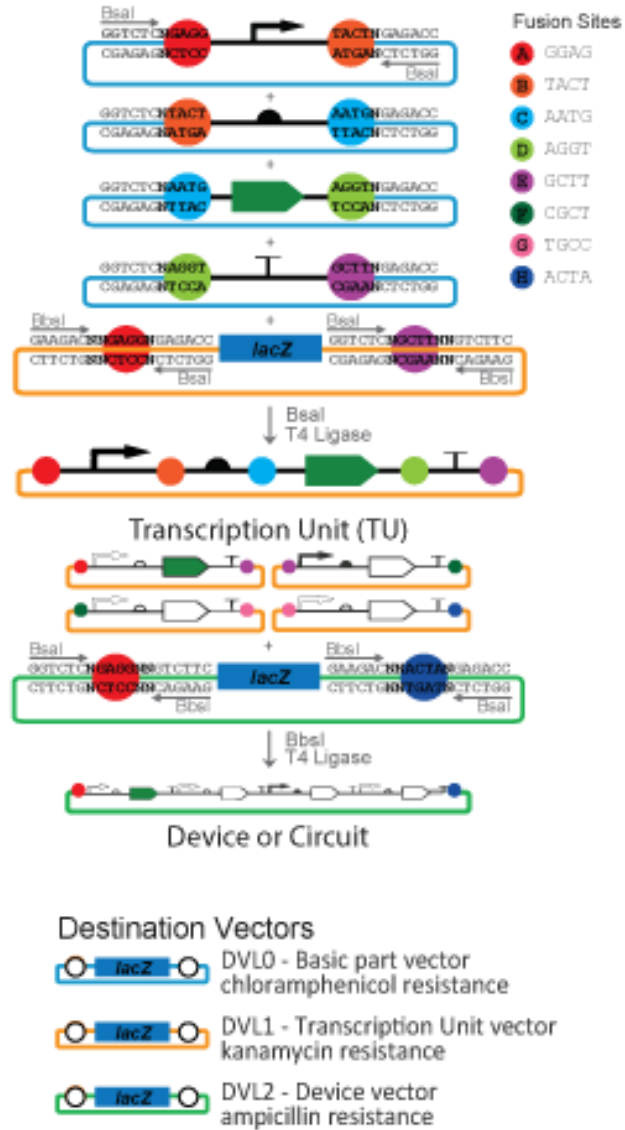
Figure 3-1 Initial three antibiotic selection assembly standard. *The initial CIDAR MoClo part collection used a three-tier antibiotic resistance standard with basic parts (Level 0) cloned into DVL0 chloramphenicol resistance plasmids, transcription units (Level 1) in DVL1 kanamycin resistant plasmids and devices (Level 2) in DVL2 ampicillin resistant plasmids.*

3.2.2 *Optimized MoClo protocols*

The originally published MoClo protocols used 40 fmol of each DNA part and a five hour incubation at 37°C with final 5' 50°C and 10' 80°C steps. In the largest assembly published in this study the authors used a cycling protocol to increase efficiency which consisted of 45 cycles of alternating 2' 37°C and 5' 16°C with the same final steps. Both of these protocols provided high efficiency cloning and both required approximately 5.5 hours of incubations (Weber et al. 2011).

In order to optimize the MoClo protocols both time- and cost-sensitive parameters were evaluated. Identical reactions were performed (Level 1, fluorescent expression transcription units) using a cycling MoClo protocol with 15, 20, 25, 30 or 40 cycles (37°C 1.5 or 3', 16°C 3 or 5'), followed by 50°C for 5' and 80°C for 10' then held at 4°C or -20°C until transformed in either Bioline Alpha Gold Select DH5α cells or the same strain prepped with the Zymo chemically competent cell preparation kit.

BsaI (NEB Cat #R0535S) and BbsI (NEB Cat #R0539S) were used in all assemblies as appropriate. NEB T4 Ligase (Cat #M0202L) and Promega T4 Ligase (Cat # M1794) were compared, each tested with NEB T4 Ligase Buffer (Cat #B0216L) and Promega T4 Ligase Buffer (Cat #C1263). Number of cycles, length of steps, reaction volume and reagent concentration were varied.

- Weber Protocol
  - Time: 5 hours
  - Expense: $10 / reaction
    - Volume: 30 uL
    - Enzymes: 20 Units
    - DNA: 40-60 fmol
    - Transformation: 50 uL cells
  - Efficiency: >95% Correct clones (>100 CFU)
  - Library: Non-existent

- CIDAR Protocol
  - Time: 45-90 minutes
  - Expense: $ 1.5 / reaction
    - Volume: 10 uL
    - Enzymes: 5 Units
    - DNA: 10 fmol
    - Transformation: 7 uL cells
  - Efficiency: >95% Correct clones (>100 CFU)
  - Library: 93 part kit, >200 basic parts created

Figure 3-2 Optimization of Weber protocol to develop CIDAR MoClo protocols. Optimization of protocols decreased cost from approximately $10 to $1.50 per reaction, reduced time from 5 hours to less than 90 minutes and maintained >95% cloning efficiency. Volumes, concentrations, cycling times, and number of cycles were each evaluated.

3.2.3 Two color controls, color model, and part characterization

We used the library and protocols to construct the 2 Color Controls (2CC) series of plasmids described in 2.4.3. FACS analysis was run as described in 2.4.1 and the data was used to define the ideal color controls for future FACS experiments. Additionally, these plasmids were used to create a color model by normalizing flow cytometry fluorescence data to SpheroTech RCP-30-5-A beads as a fluorescein standard (**Figure 3-3a**). This color model allows us to directly correlate RFP and GFP expression as molecules of equivalent fluorescein (MEFL) counts rather than arbitrary units (**Figure 3-3b**).

This color model was validated by comparing MEFL counts of 28 pairs of plasmids with various promoter, RBS, and fusion site combinations (mean square

error = 1.80 fold) in which each pair differed only in the coding sequence (E0040m_CD GFP or E1010m_CD RFP) (**Figure 3-3c**).
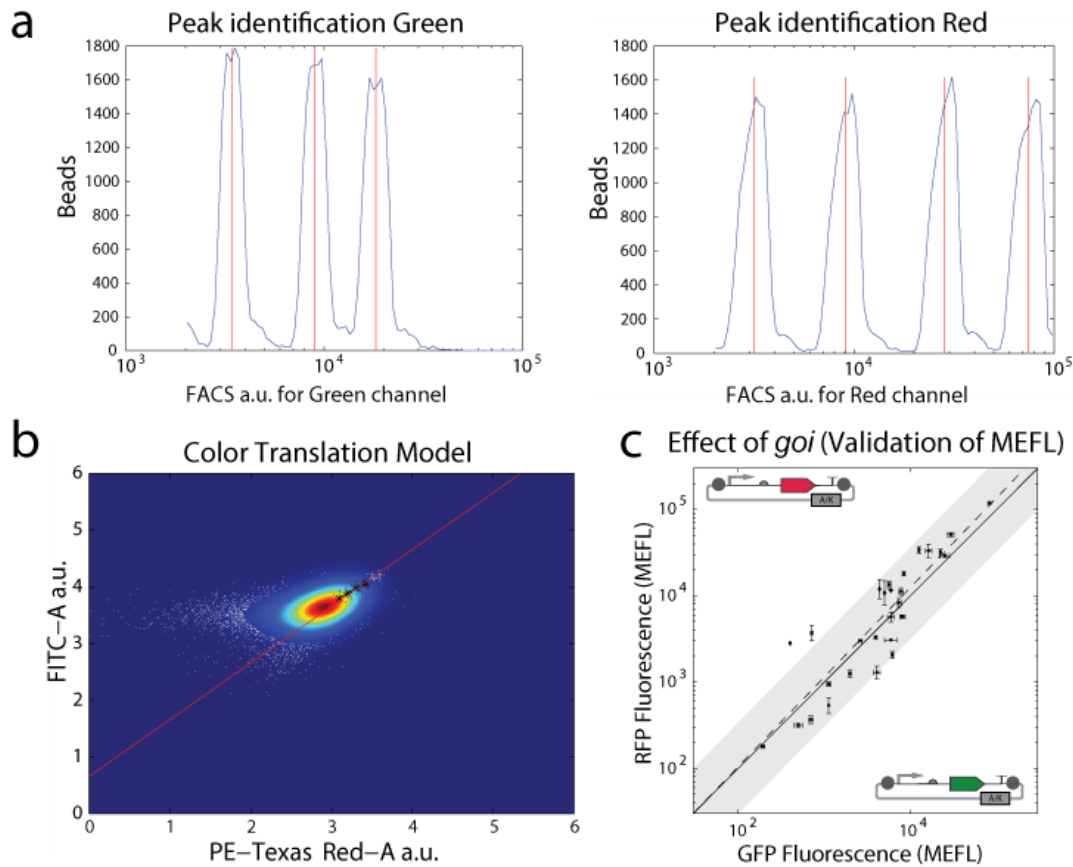
Figure 3-3 Molecules of Equivalent Fluorescein (MEFL) normalization of flow cytometry data.
SpheroTech 8-peak beads provide a basis for normalizing across fluorescence channels using a
physical standard. (a) Peaks visible with selected laser voltages and filters for green fluorescent
protein (left) and red fluorescent protein (right). RFP measurements are translated to MEFL by first
mapping to estimated equivalent FITC. Geometric statistics are then computed over data in MEFL
units. (c) To validate MEFL normalization, 28 pairs of transcription units were compared in which
only the CDS varied (RFP on the y-axis verses GFP counterparts on the x-axis). An approximate
1:1 ratio is observed with a mean square error of 1.80 fold.

An array of GFP reporters composed of 16 constitutive promoters (J23100 Anderson series) and 6 RBS type parts (3 Weiss RBS, 3 BCD) constructed to characterize the transcription and translation elements to explore the potential for rational design using these parts. The resulting 96 iterations of a simple GFP expression plasmid in a DVK_AE vector were analyzed using flow cytometry (**Figure 3-4a,b**). Two-way ANOVA analysis identified 36.3% of expression variation as being due to the promoter, 43.9% due to RBS part, and 19.3% due to the interaction of the two factors.
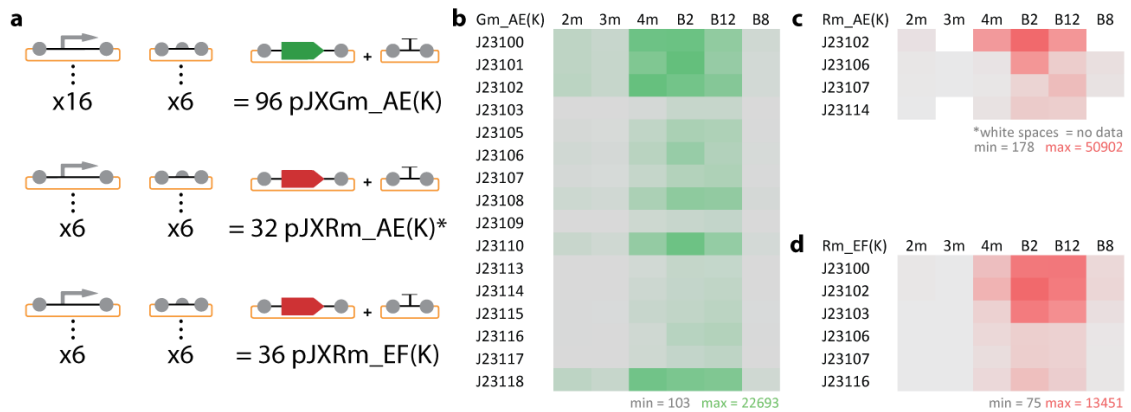
Figure 3-4 Array of expression vectors constructed from a collection of 16 consitutive promoters and 6 RBS type parts (3 Weiss RBS, 3 BCDs). (a) Construction of arrays with multiple promoters and RBS parts, one CDS part, one terminator, vector. (b) pJXGm_AE(K) expression array consists of 16 promoters and six RBS parts. TUs containing RBS parts B0034m, BCD2, and BCD12 and promoters J23100, J23101, J23102, and J23118 have the highest GFP fluorescence. (c) Subset of the contexts demonstrated in (b) with RFP in place of the GFP CDS. (d) Subset of the contexts demonstrated in (b) with RFP in place of GFP and flanked with _EF fusion sites in place of _AE. In both c and d a similar expression pattern can be seen with B0034m, BCD2, and BCD12, J23100, and J23102 demonstrating high levels of expression. Max and min for each array (color coded) are noted in MEFL units.

To compare the expression patterns with a different CDS, a subset of this array was made using RFP in place of GFP (**Figure 3-4a,c**). The overall pattern appears similar, though the range of expression is approximately 2-fold larger than with the corresponding GFP array (max MEFL value of 50902 as compared to 22693). A third small array was analyzed using the RFP transcription units in a DVK_EF vector as compared to the DVK_AE vector in **Figure 3-4** parts b and c. This subset shows a much smaller range and lower max signal compared to the RFP _AE array in part c offering the first indication that the flanking fusion sites influence expression.
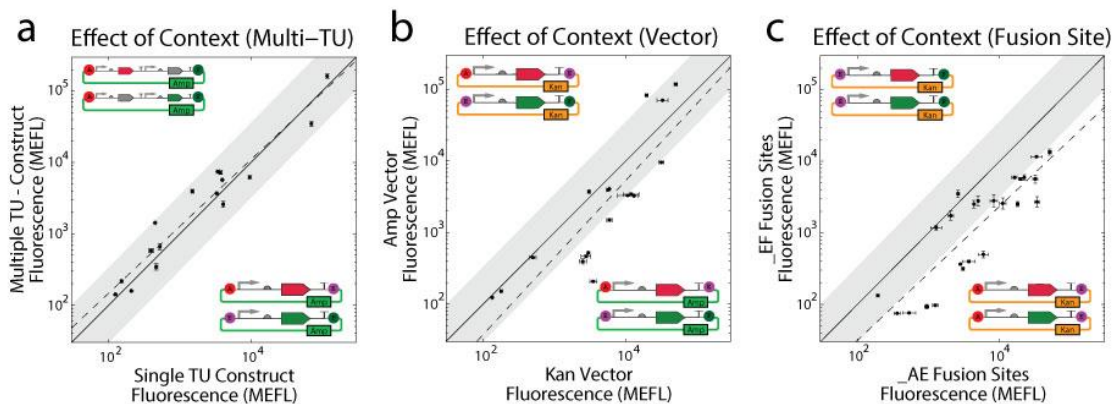
Figure 3-5 Pairwise comparisons of context effects. *(a) Single TU expression compared to the same TU when expressed in the same plasmid with another TU. Expression of a single TU is consistent when assembled into a larger device. (b) Changing vector from DVA to DVK shows a high level of variability, suggesting all expression tuning should be done in the same vector. (c) Changing the 4 base pair fusion sites flanking a given TU influences expression likely due to the proximity of the 5' fusion site to the minimal 35 bp promoter used in this study.*

To further assess the capability for rational design under commonly varied genetic contexts, other pairwise comparisons were performed to evaluate the effects of gene order, variation in vector, and use of specific fusion sites flanking

the transcription unit. Expression of a single TU was shown to remain constant when expressed in a plasmid with a second TU up- or downstream (**Figure 3-5a**) (mean sq. error = 1.53 fold). Changing the vector from DVA to DVK while maintaining the same transcription unit showed a higher variability yet retained a nearly linear relationship (mean sq. error = 2.61 fold) (**Figure 3-5b**).

Modifying the four base fusion sites did have an effect on expression (p<0.001) in a paired analysis in the _EF transcription unit when compared to the _AE paired clone (mean sq. error = 2.02 fold) (**Figure 3-5c**). This difference may be due to the proximity of the 5' fusion site to the simple promoter. Including an insulator upstream of the promoter sequences may mitigate this effect.

In order to select the constitutive promoters to be included in the CIDAR MoClo Library, the data from **Figure 3-4b** was graphed as seen in **Figure 3-6,** and **Figure 3-7.** Six constitutive promoters were chosen which would provide a full range of expression capabilities (as defined by GFP expression ranging from 10^2 – 10^5 MEFLs) when paired with these six RBS parts, also included in the library. The MEFL data used to create **Figure 3-3c, Figure 3-4, Figure 3-5, Figure 3-6,** and **Figure 3-7** can be found in Appendix A along with a concise histogram of the GFP expression of promoter:RBS combinations included in the library (**Figure A-2**).
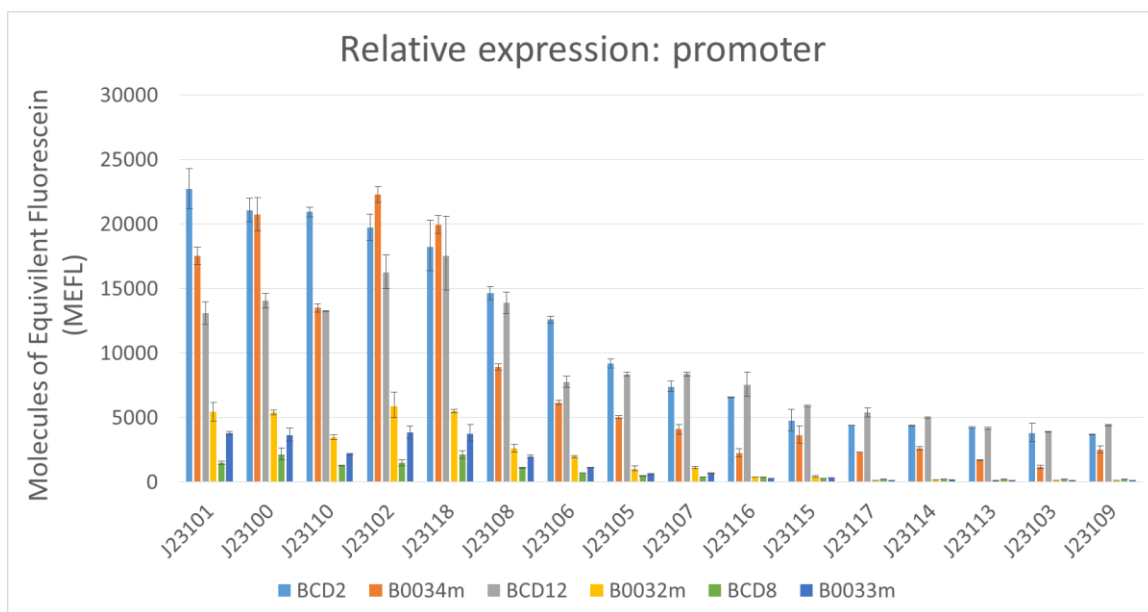
Figure 3-6 GFP expression under the control of variable promoters and RBS. *Data here sorted by promoter to guide selection of promoter parts in future constructs. Same data as presented in* *Figure 3-4b*. *Subset of data for parts included in the CIDAR Library are in Appendix A.*
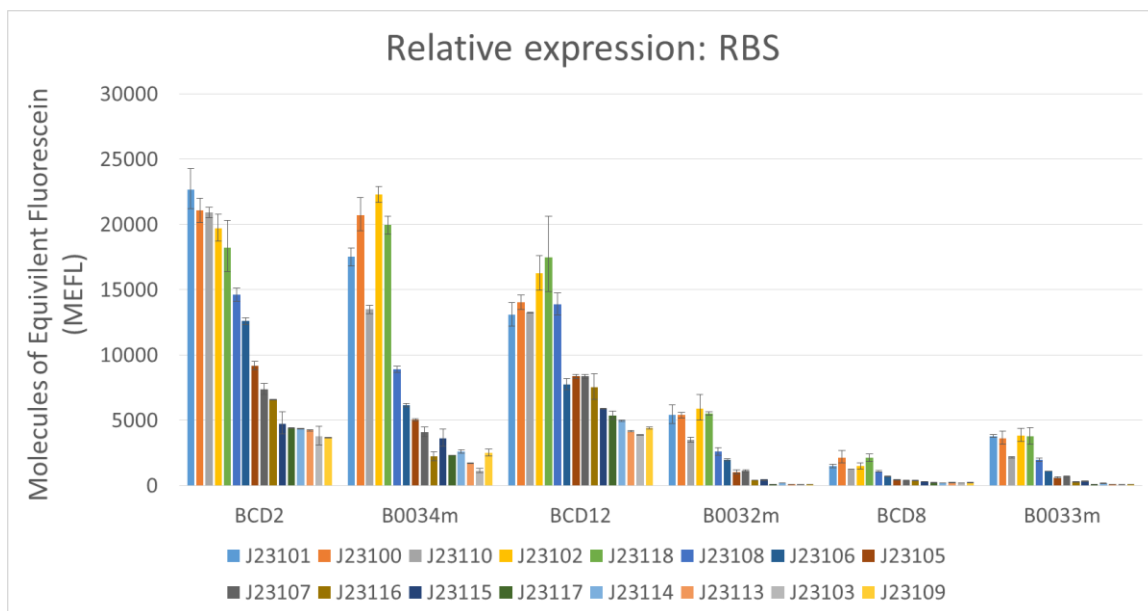
Figure 3-7 GFP expression under the control of variable promoters and RBS. *Data here sorted by RBS to guide selection of RBS parts in future constructs. Same data as presented in* *Figure 3-4b*. *Subset of data for parts included in the CIDAR Library are in Appendix A.*

3.2.4 Publicly available CIDAR MoClo E. coli part library and assembly standard

Most of the parts contained in the CIDAR *E. coli* MoClo Library are derived from the BioBricks Registry (http://parts.igem.org/) and were selected for their functional reliability and utility in synthetic biological designs (**Table A-7**). To enable rational design, three of the basic parts were selected from the BIOFAB collection of BCD translational elements which have been shown to enable more rational design in terms of protein expression (Mutalik et al. 2013a). These BCD parts contain a leader peptide followed by a secondary RBS in order to physically separate transcriptional and translational regulation.

E. coli Modular Parts and Vectors
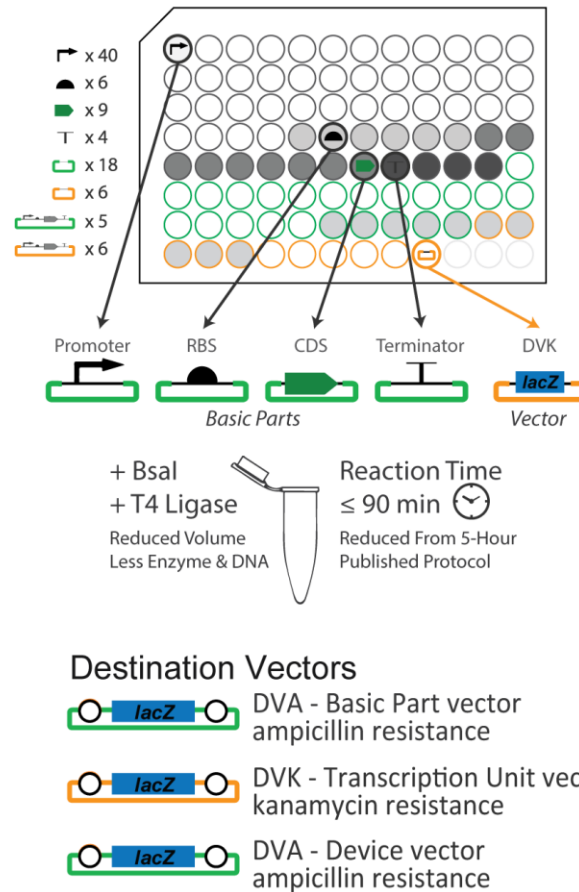
Figure 3-8 CIDAR MoClo Library, part and vector structure. *The CIDAR MoClo Library provided in a 96-well plate (Addgene, #1000000059). CIDAR MoClo assembly standard is based on four part transcription units comprised of a promoter, ribosome binding site, coding sequence and terminator assembled into a DVK (Kanamycin-resistance destination vector). Destination vectors alternate antibiotic resistance at each level and use lacZα blue-white selection.*

Destination vectors are included in the library to allow for simple cloning of new parts with any of the standard fusion site pairs. From the initial 3Ab assembly standard, the DVL0 vectors have been removed from the rotation and replaced

57

with the DVA vector (**Figure 3-8**). DVA is derived from DVL2 which was a modified

pSB1A2 carrying an illegal site in the *bla* ampicillin antibiotic resistance gene. This

illegal site was repaired with a synonymous single codon mutation to produce DVA.

All basic parts are provided in DVA vectors. These are then assembled into

transcription units in DVK vectors (previously designated as DVL1) (**Figure 3-9**).

The circuit or device level of assembly combines two or more TUs into a DVA

vector. Further hierarchical assembly is possible by continuing the DVA – DVK

rotation.

In addition to the parts and vectors, a set of fluorescent expression plasmids

is also included to be used as standards with the TASBE flow cytometry analysis

tools (Beal et al. 2012) (**Table A-7**). Sequence and part information for all plasmids

is available in the CIDAR Inventory of Composible Elements (ICE) registry

(http://cidar-ice.org). Other plasmids for use in *E. coli* not included in this library

are available upon request.
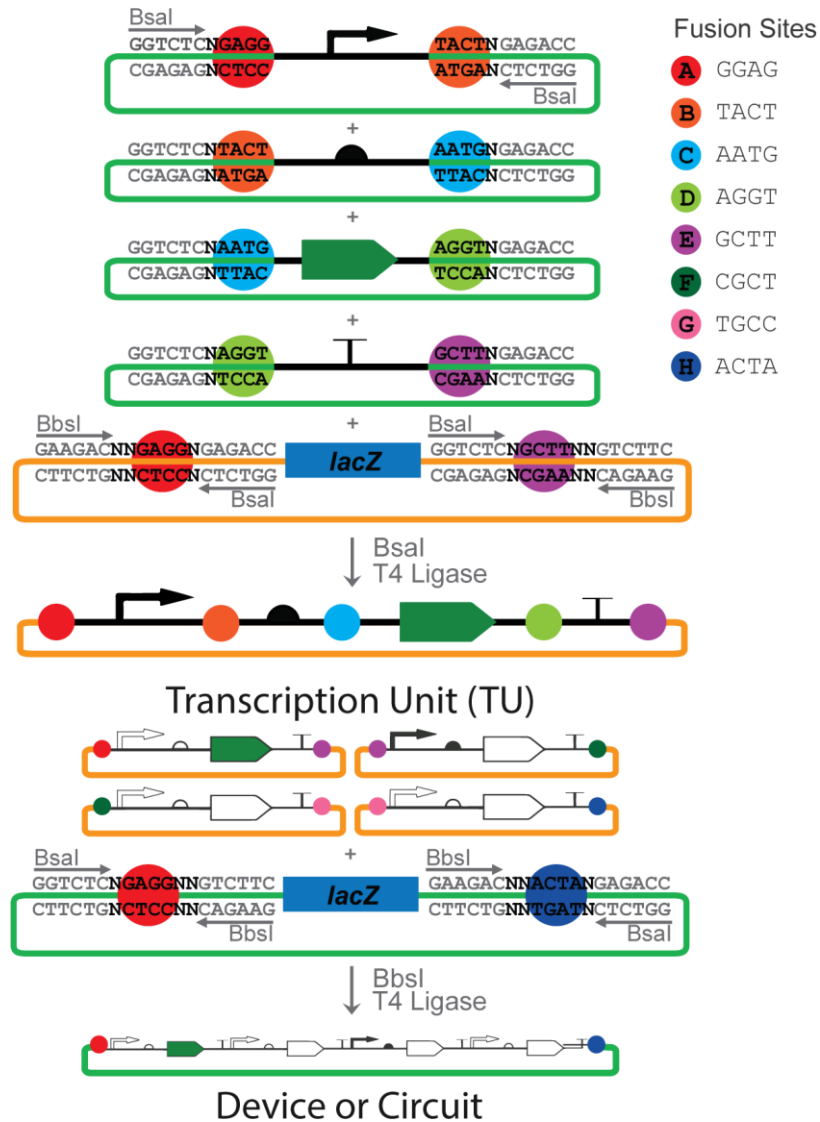
Figure 3-9 CIDAR MoClo Assembly Standard. *Basic parts shown with green backbones are prepared in DVAs with part specific fusion sites, indicated with single capital letters. The 3' fusion site of the upstream part must match the 5' fusion site of the following part in order to be correctly assembled. Digesting parts and a DVK with the BsaI and simultaneous ligation with T4 DNA ligase results in a transcription unit (TU) shown here with the orange backbone. Multiple TUs can be combined into a complex device using BbsI in place of BsaI with the appropriate DVA.*

3.2.5 Multiplex Modular Cloning (MMC)

As a MoClo reaction is dependent upon equimolar ratios of each part type, a logical advance on the methodology is to multiplex reactions by adding various plasmids of the same part type at 1/n the concentration of each other part type where n equals the total number of iterations. Multiplex MoClo (MMC) allows for an expanded utility including simultaneously screening a large number of iterations in parallel while retaining >95% cloning efficiency. Examples of this methodology include screening of variant sequences, or modulating expression levels of a single transcription unit by multiplexing the promoter and/or RBS part (**Figure 3-10**).

Multiplex MoClo reactions have been performed using up to 6 different part types with as many as three part types multiplexed without decreasing efficiency. Library creation using multiplex MoClo could facilitate the development of fusion proteins with additional fusion sites used to build a CDS part from intermediates. The DVL0_CX, DVL0_CY, DVL0_XY, DVL0_XD and DVL0_YD vectors were created for this purpose and used to assemble fusion proteins comprised of three domains (estrogen receptor - cre recombinase - estrogen receptor, in one example).

By creating a library of _CX protein coding sequences and a _XD encoded tag (His-tag for example) a library of tagged protein sequences can be created in a single assembly step. This same approach could be applied to protein engineering to determine the optimal residues at which to create a fusion or to

experiment with linker designs using the _XY part type. Mass production of fluorescent reporters could be assembled with the same methods.
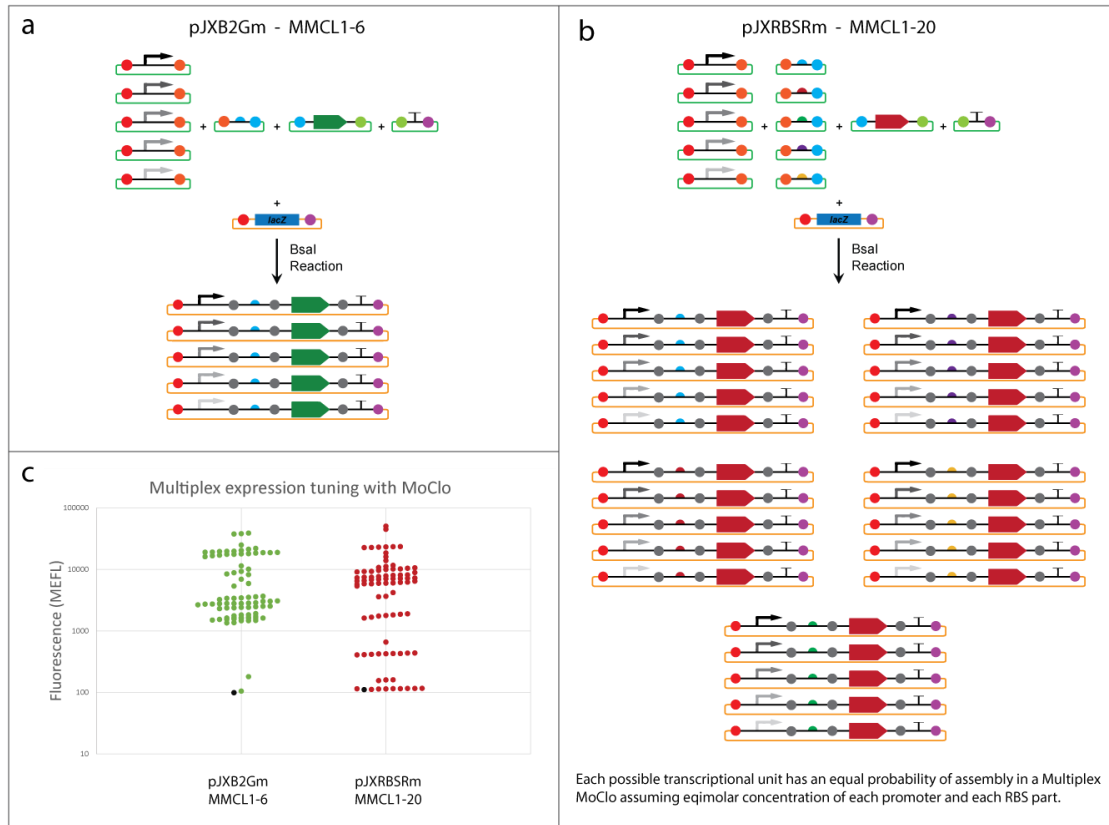
Figure 3-10 Multiplex MoClo. The modular format of MoClo allows for simple multiplexing of one or more part types by the addition of 1/n of the molar concentration of each part where n = the total number of parts of that type. (a) Basic structure of Multiplexed MoClo (MMC). (b) Assembly of a 5 x 5 multiplex reaction in which both the promoter and RBS type part are multiplexed as seen in the red coded section of (c). (c) MMC of five promoters with BCD2, GFP, a standard terminator and DVK_AE (pJXB2Gm) provides five distinct populations of fluorescent cells. Likewise, multiplexing both the promoter and RBS type part results in more populations than can be accurately identified from fluorescence expresssion alone (pJXRBSRm).

One of the most powerful uses of multiplex MoClo may be found in the form of expression tuning of complex devices. By multiplexing the promoter and RBS

parts of one or more TU, then using the mixed population of all TUs assembled in those reactions to assemble a multi-TU device, a vast array of iterative designs can be assembled in a single tube and screened for the desired function. When combined with fluorescent output signals, this expression tuning can be used in conjunction with fluorescence activated cell sorting to measure and isolate only those clones which demonstrate the desired behavior.

3.2.6   *Insulating recombinase sites within BiCistronic Designs (BCDs)*

In the interest of developing *cre* recombinase based memory devices, preliminary investigations into the ability to use BCD parts as insulators was investigated. A commonly used recombinase, cre and its cognate recognition sequence, LoxP, are well studied as genome modifying tools in transgenic mouse research and recombinase memory devices in synthetic biology (Orban et al. 1992, Friedland et al. 2009). The placement of LoxP, the *cre* recognition sequence, between promoter and RBS parts proved problematic in early designs with BioBrick assembly, ameliorating fluorescence expression (data not shown). The LoxP sequence (36 nt) was placed in frame in the short encoded peptide sequence of BCD parts where it would be translated, disrupting any secondary structure which might otherwise influence expression. New RBS type parts were made (BCD2lox_BC, BCD12lox_BC, etc.) which included a potentially insulated LoxP site (**Figure 3-11**).

An accurate measure of insulation would have required a related construct containing LoxP between the promoter and CDS without the use of an insulator. However, these designs would also invariably offer other explanations for sequence variance such as the sequence context surrounding the promoter or 5' region of the gene. In order to accurately measure insulating effect, a wide variety of constructs would need to be tested to identify factors influencing expression, making true comparison of insulation effect on a small scale difficult.

A small preliminary experiment was performed to assess whether transcription units carrying an embedded LoxP sequence expressed a fluorescent goi similarly to the counterpart transcription unit lacking the LoxP sequence. In five of the eight instances tested, transcription units with embedded LoxP sequences showed expression similar to that of counterpart TUs lacking LoxP.

While this data is not conclusive, it does suggest that embedding repetitive or A/T rich sequences within the translated portion of the BCD may prevent secondary structure and minimize detrimental effects of the given sequence on protein expression. Further experimentation is needed to determine the full potential of this technique.
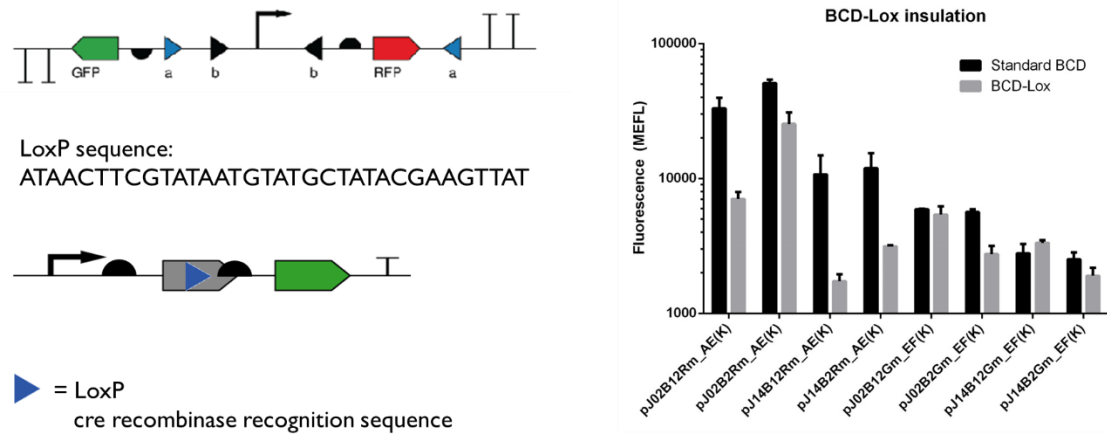
Figure 3-11 BCD parts as insulators. *Devices like those shown in the top left are difficult to design rationally due to the inclusion of hairpin-forming recognition sequences in the transcription and translation control regions. LoxP was placed inside the peptide coding sequence of the BCD part to insulate and disrupt the hairpin structure as illustrated (middle left). (right) In five of the eight instances tested, inclusion of the LoxP sequence within the BCD part resulted in similar levels of fluorescent protein expression similar to that of counterpart TUs lacking LoxP.*

3.2.7 CIDAR Inventory of Composable Elements (CIDAR ICE)

The Inventory of Composable Elements (ICE) is an open source registry designed to assist in management of DNA part and strain information (Ham et al. 2012). ICE was built to facilitate the creation of a web of registries, with support for distributed interconnected use. Currently, ICE registries have been implemented at multiple research facilities including Stanford University, University of California Berkeley, Joint Genome Institute (JGI), Harvard Medical School, Synberc (NSF funded synthetic biology consortium), and now Boston University (Paige 2014).
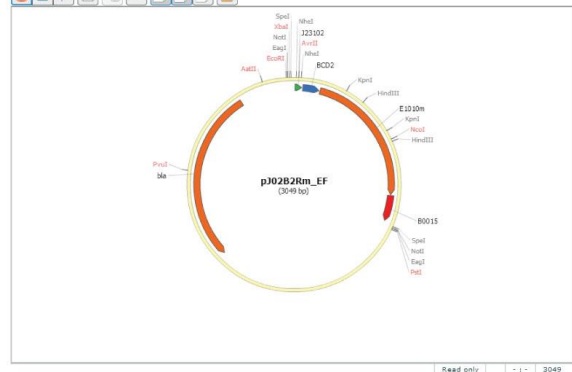
CIDAR-ICE is a publicly available database displaying sequence and part information for the CIDAR MoClo Library (https://cidar-ice.org) (**Figure 3-12**).

Figure 3-12 CIDAR-ICE, publicly available parts registry. *Part of the Web of Registries (JBEI-ICE), the CIDAR-ICE registry can be found at http://cidar-ice.org. Top left: Sign in page. Top right: Plasmid map example. Bottom: Home screen view of plasmids available.*

3.3   Discussion

3.3.1 CIDAR MoClo E. coli part library

Four MoClo part libraries were previously available providing reusable parts and vectors for plant transformation constructs (Addgene, #1000000047), general eukaryotic multigene construct assembly (Addgene, #1000000044), mammals (mMoClo, (Duportet et al. 2014)) and yeast (Addgene, #1000000061). Recent

publications have also detailed the development of yeast Golden Gate (yGG) and BASIC assembly methods both of which employ a similar digestion and ligation reaction, though neither offer a part library to accompany the methods. To address the lack of publically available standardize parts for bacterial systems, we have constructed and characterized a library of commonly used genetic parts and the necessary vectors in MoClo format for use in E. coli (Addgene, #1000000059) (**Figure 3-8** and **Table A-7**).

The CIDAR MoClo part library is the first Type-IIS compatible modular part library available for use in bacteria. It has already been employed by four other research labs by directly sharing the part library. Additionally, it is available publicly through Addgene along with other MoClo libraries enhancing the universality of this assembly standard.

Though the chart shown in **Figure 1-1** shows only sparse adoption of the MoClo methodology, this survey was done prior to the availability of any of these part libraries. With MoClo assembly methods and libraries more available and adapted for use in a wide variety of organisms, it is to be expected that adoption of MoClo methods will increase rapidly. Efforts to educate and promote the use of these systems would help to increase their adoption rates.

One approach to increasing adoption of the MoClo assembly methods is to incorporate these methods into high school and university level biology labs, familiarizing more students with Type IIS assembly methods at an early point in

66

their research careers. MoClo is an ideal teaching tool for demonstrating the concepts of synthetic biology and modular engineering. As such, Appendix C in this work includes basic teaching lab materials which cover the assembly of a single transcription unit, multiplex MoClo assembly and analysis, and the construction of a two color fluorescent plasmid.

3.3.2 MoClo assembly standard variations

Each library uses a variant of the same format to adapt to the target organisms. The original Weber protocol uses five basic parts (promoter, RBS, signal peptide, CDS, and terminator) and a series of vectors. The CIDAR MoClo system was adapted from this standard with a four part format and a series of destination vectors. The yeast system, MoClo-YTK, uses eight primary part types with individual parts making up the origin and marker portion of the backbone (Lee et al. 2015). The mammalian MoClo (mMoClo) standard uses six basic parts (insulators, promoters, 5'UTR, genes, 3'UTR, and polyA) to assemble transcription units or gene trap plasmids containing *att* sites (Duportet et al. 2014).

Two fusion sites flank transcription units and in most MoClo formats end linker parts are used to enhance modularity (Weber et al. 2011). The Weber MoClo format rotates three antibiotics and two color selection (LacZ and CRed) whereas the CIDAR MoClo standard simplifies this rotation to two antibiotics and one color selction (LacZ). Between the Weber and CIDAR assembly standards, the _AB promoters and all RBS part types (_BC) are directly compatible.

67

	Weber	CIDAR	Plant	mMoClo	MoClo-YTK
# of part types	4-6	4-6	5-8	6	8-10
Linkers	yes	no	yes	yes	yes[1]
Enzymes	BsaI, BpiI	BsaI, BbsI	BsaI, BpiI	BsaI, BpiI	BsaI, BsmBI
Library	#1000000044	#1000000059	#1000000047	Submitted	#1000000061
Part Fusion Sites	GGAG	GGAG[2]	GGAG	GGAG	CCCT
	TACT	TACT	TACT	TACT	AACG
	AATG	AATG[3]	CCAT	AATG	TATG
	AGGT	AGGT	AATG	AGGT	TTCT
	GCTT	GCTT[4]	AGGT	GCTT	ATCC
	CGCT		TTCG	CAAC	TGGC
			GCTT	CGCT	GCTG
			GGTA		TACA
			CGCT		GAGT
					CCGA
					CAAT
					CCCT

[1] *Linkers are termed "connectors" and flank every TU.*

[2] *Promoters may contain one of the following 5' fusion sites (GGAG, GCTT, CGCT, TGCC)*

[3] *Fusion proteins can be assembled with the addition of X (CGTT) and Y (TGTG) fusion sites.*

[4] *Terminators may contain one of the following 5' fusion sites (GCTT, CGCT, TGCC, ACTA)*

Table 3-1 Comparison of published MoClo standards. *Though four of the published standards use many of the same fusion sites (highlighted), the part type associated with those sites varies making many of these parts incompatible.*

3.3.3 MoClo & CIDAR Workflow

To better enable a specify-design-build-test-share workflow, the CIDAR workflow includes tools for each stage as described in **Figure 3-13**. All of these tools are open source and publicly available to facilitate the adoption of this workflow by other researchers.

CIDAR Workflow

Specification	Design	Build	Test	Share
Programming language for describing synthetic biology devices	Software tool for creating DNA assembly plans in common assembly formats	Collection of DNA part libraries based on CIDAR MoClo assembly.	Software tool to convert arbitrary fluorescent units to Methyl Equivilant Fluorescein (MEFL) physical units.	Inventory of Composable Elements (ICE), online database of sequences and characterization data.
Eugene	Raven	Hummingbird	TASBE Tools	CIDAR
www.eugenecad.org	www.ravencad.org	www.cidarlab.org/moclo	synbiotools.bbn.com	www.cidar-ice.org

Figure 3-13 CIDAR Workflow: Specification, design, build, test, share. In pursuit of automation, the CIDAR lab has developed a collection of tools. These allow for the specification of genetic circuits in human-readable computer language, automate assembly planning, provide libraries of DNA parts, analyze fluorescence data in MEFL units, and release all DNA parts and data publicly through the ICE database and other resources.
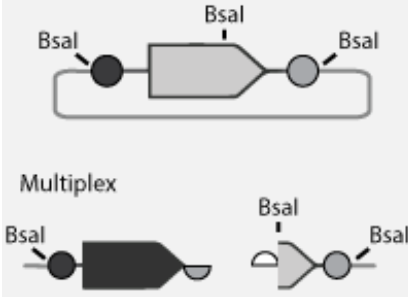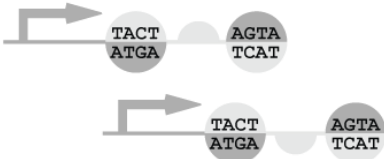
Eugene is an ecosystem of human-readable languages for synthetic biology which facilitates the specification of biological devices (Bilitchenko et al. 2011, Oberortner et al. 2014, Oberortner and Densmore 2015). A methodology agnostic DNA assembly planner, Raven, provides human-readable and machine-executable instructions for constructing complex DNA devices with Golden Gate, MoClo, Gibson, and BioBricks formats (Appleton et al. 2014). The CIDAR MoClo Library is part of the Hummingbird project, a collection composed of this and other MoClo libraries currently in development. The TASBE tools for normalizing fluorescence enable the comparison of experimental data between experiments and researchers (Beal et al. 2012).

3.3.4 Troubleshooting and known error modes

In developing the CIDAR MoClo Library, alternative formats and fusion sites were attempted with varying degrees of success. **Table 3-2** describes known failure modes.

In particular, a series of fusion sites were designed to enable multi-cistronic assembly in the form of promoter:rbs:cds:rbs:cds: terminator and promoter:rbs:cds:rbs:cds:rbs:cds:terminator. Previous experiments had shown 6-part assembly to maintain high efficiency, however, these bicistronic designs were unsuccessful, apparently due to the choice of fusion site. In order to continue the practice of containing the "ATG" start site within the 5' CDS fusion site, the C1 and C2 fusion sites were designated as GATG and TATG, respectively. All attempts to build with these parts failed, providing few white colonies all of which were sequenced as artifacts. These artifacts were predominately with respect to the 5'CDS fusion sites where the _C1D1 CDS part would fuse to the _BC part rather than the _DC1 part, skipping the first transcription unit and other similar incorrect assemblies.

This series of experiments indicated the need for greater variance in the fusion site sequence in a given MoClo reaction. In general, the fusion sites used in the CIDAR MoClo assembly standard will differ in at least 2 of the 4 base pair positions with respect to all other fusion sites in a given reaction.

FAILED ASSEMBLY STRATEGY	NAME	EXPLANATION
	Multicistronic design	C (AATG) and C1 (GATG) are too similar, varying by only 1 bp and allow for mismatch ligations. The A:B:C1:D1:E transcription unit was the most common result. Use of different fusion sites should enable this design.
	Palindromic fusion sites	Palindromic fusion sites result in overhangs which are able to fuse to inverted copies of the same part rather than the intended downstream part, inhibiting correct assembly.
	Illegal sites	Illegal sites in parts or vectors decrease the efficiency of a given reaction. The ad-hoc 'fusion site' made by an illegal restriction site may be palindromic, a reverse compliment of existing sites, or too similar to existing sites to allow for optimal conditions. Illegal sites are particularly problematic in Multiplex reactions when one or more individual parts of a multiplexed part type contains an illegal site.
	Reverse compliment of existing fusion site	Fusion sites which are the reverse complement of an existing fusion site create a backward assembly step. In the case demonstrated here, with one part flanked by reverse compliment sequences, this creates an artifact whereby repeated RBS parts could assemble in a chainlink fashion.

Table 3-2 Known failure modes. *Choice of fusion sites is key for efficient and accurate MoClo assembly. Greater than 1 bp variance is required between sites in a given reaction.*

3.3.5 *Rational design and expression prediction with CIDAR MoClo*

Using the MEFL normalization method, plasmids made with the CIDAR MoClo part library have a rational expression capability with a mean square error of less than 2-fold. A given genetic context will express the same physical amount of protein regardless of the coding sequence. The BCDs in particular enable this rational design by reducing variation in expression due to an interaction between the transcription and translation elements.

We hypothesized that the GFP and RFP data described above could be used to predict protein expression of a given context with any gene of interest. To evaluate this hypothesis, CIDAR MoClo was used to predict and build an equimolar dual expression plasmid, described in Chapter 4. Of the twelve candidate plasmids evaluated, four provided near-equal expression of two variants of a rabbit aldolase protein, one of which contains a 5' His-tag. Of these, the highest co-expression plasmid is being further analyzed and used in downstream experiments.

While the CIDAR MoClo standard does enable rational design, improvements could be made to increase modularity and consistent expression. The Weber and MoClo-YTK (yeast) protocols use linker or connector parts to flank the transcription unit and increase modularity. This has a benefit of reducing the number of part types required in the library for the same number of permutations. Additionally, this modularity further constrains expression variation by standardizing the sequence 5' of the promoter region. To improve modularity in

this fashion with the CIDAR MoClo assembly standard, a 5th and 6th part type could be created which would flank the 5' and 3' ends to remove the need for variable promoter and terminator parts. Instead, all promoters would use A as a 5' fusion site, further controlling expression variations, and all terminators would end in a 3' E site.

3.3.6   Applications of MoClo in traditional biological research and synthetic biology

With interchangeable parts and the ability to multiplex assembly, MoClo has applications as a time and cost efficient means of library creation and propagation, mutagenesis screening, transcription factor characterization, protein engineering, and genomic engineering. In order to demonstrate this utility, in Chapter 4 we collaborated with a biology lab and employed CIDAR MoClo and experimental data from **Figure 3-3** and **Figure 3-4** to balance protein levels with a dual expression plasmid. In Chapter 5, MoClo was used to build modular sensors and cellular logic gates with plasmids designed to induce a change in pH upon induction. This pH signal can be read with affordable open-source electronic devices to create bioelectronic sensors and logic devices.

3.3.7   CIDAR MoClo as an educational tool

The simplified concept of CIDAR MoClo (two antibiotic, one color, shorter protocols) lends itself to an educational setting with the potential for biology practice lab protocols being developed to use the CIDAR parts and protocols.

These materials are being developed and will be available at http://cidarlab.org/moclo.

3.4   Conclusions

This work describes the construction of the first bacterial multipart modular DNA part library for use in bacteria and the development of optimized MoClo assembly protocols and standards. This library and standard lays the groundwork for a wide range of synthetic biology applications, greatly increasing efficiency and modularity in bacterial engineering. The CIDAR MoClo Library, especially when combined with a related design tool (Appleton et al. 2014), allows for rapid assembly of synthetic gene networks and cost efficient combinatorial assembly.

The publicly available CIDAR MoClo Library is intended to provide a starting point for research labs to adapt for use in specific fields through the inclusion of new DNA parts. It is a flexible assembly standard, allowing for multiplexing, and the addition of new part types, new fusion sites, and the streamlined assembly of fusion proteins. It is also ideal for academic research and educational uses in teaching laboratory settings (Appendix C).

Though formatted for cloning in *E. coli*, iterative design of eukaryotic circuits is also possible through the introduction of species specific parts. Bio-design automation activities are further promoted with the available Eugene design files that capture not only the parts in the library but also initial design guidelines and data (Appendix D).

The CIDAR MoClo Library is available through Addgene (www.addgene.org/cloning/moclo/densmore) and additional individual plasmids along with functional information and DNA sequence are publicly available through the CIDAR ICE Public Registry (www.cidar-ice.org).

4 CHAPTER 4 RATIONAL DESIGN WITH CIDAR MOCLO EQUALIZES PROTEIN EXPRESSION OF TWO ALDOLASE VARIANTS IN *E. COLI* TO ENABLE ISOLATION OF HETEROTETRAMERS

4.1   *Introduction*

To demonstrate the utility of the CIDAR MoClo Library, we have collaborated with the Tolan lab (Biology, Boston University) to tune expression of two variants of rabbit aldolase protein, HRA (His-tagged wildtype) and R42A (actin-binding-deficient mutant). Aldolase, well studied for its role in glycolysis, is only found in nature in an unusually stable homotetramer and is known to bind F-actin providing a scaffold for actin crosslinking. This crosslinking of actin fibers creates a disordered matrix unsuitable for crystalizing, preventing characterization of the actin-aldolase interface. HRA:R42A tetramers in which only HRA is able to bind actin may enable crystallization by preventing crosslinking (**Figure 1-4**). Attempts at isolating these heterotetramers from dual expression strains have not been successful, due to low expression of HRA caused by the 5' His-tag. Using the CIDAR MoClo library and previous expression data, a series of dual expression plasmids were created and evaluated. Of these, two show equal expression of each aldolase variant by western blot analysis and are being applied to crystallization studies.

We tested the predictive engineering capability of the CIDAR MoClo system by producing a dual expression plasmid with two variants of the rabbit aldolase

protein expressed constitutively in equal quantities to enable the isolation of heterotetramers. These heterotetramers, dimers of dimers, will be used for electron microscopy to characterize actin-aldolase binding.

4.2   Results

4.2.1 Predictive design of aldolase Transcription Units

Using the data gathered in Chapter 3 on GFP and RFP expression (**Figure 3-3c, Figure 3-4, Figure 3-5** and **Appendix A**) as a reference, a small number of promoter:RBS combinations were chosen which were predicted to express a *goi* in similar molar quantities. Single expression units for each aldolase variant were created, sequence verified, and assembled into dual expression plasmids for the isolation of heterotetramers (**Figure 4-1**).

Five contexts were chosen for high equimolar expression of R42A and HRA in *E. coli*, J06B12X_AE, J07B2X_AE, J07B12X_AE, J02B2X_EF, and J02B12X_EF. The _EF vector context tends to have lower expression of a given transcription unit compared to the same construct in an _AE vector. As expression in the _EF transcription unit is the limiting factor for _AF dual expression devices, the highest expression contexts in _EF which showed consistency in previous data were chosen to enable the highest overall expression levels. Three _AE constructs were chosen to match these _EF plasmids in expected expression.
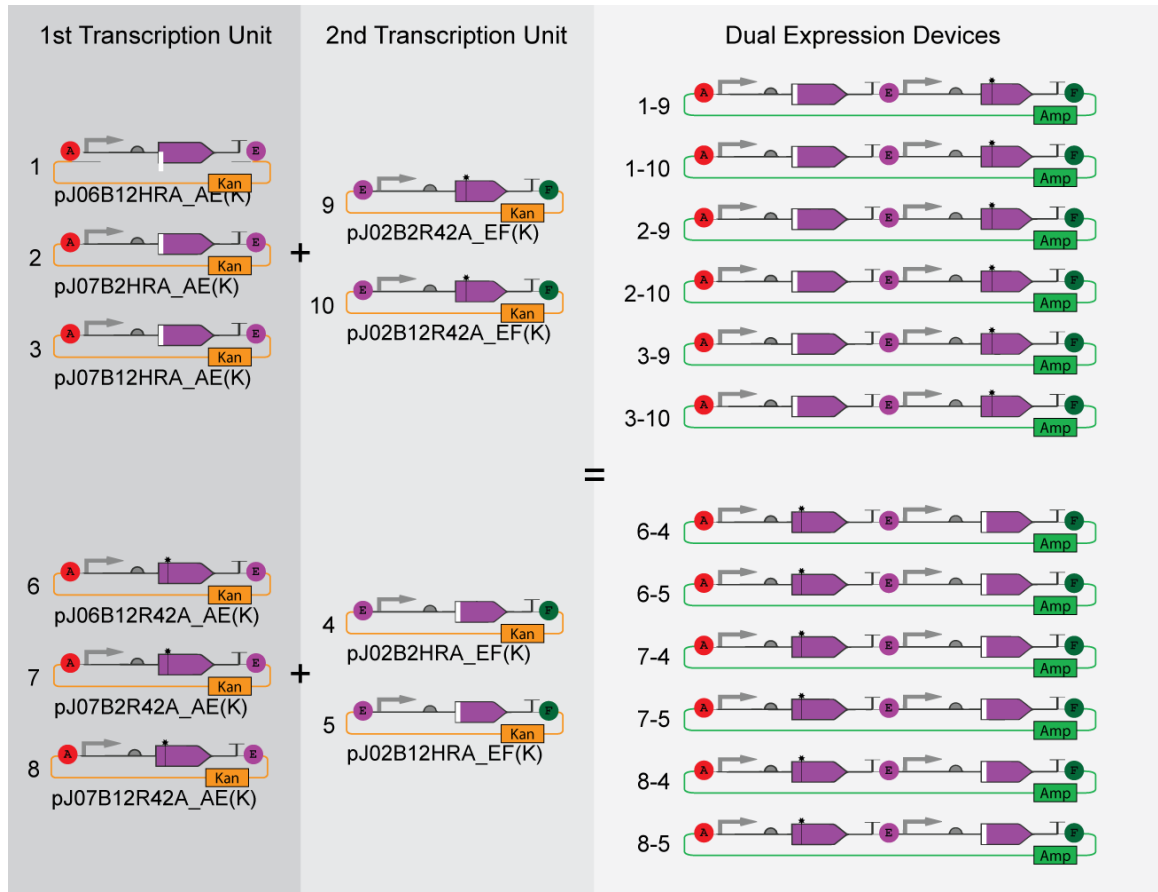
Figure 4-1 Plasmid design for aldolase dual expression devices. *HRA transcription units are pJ06B12HRA_AE(K), pJ07B2HRA_AE(K), pJ07B12HRA_AE(K), pJ02B2HRA_EF(K), pJ02B12HRA_EF(K) and designated Tolan-1 through Tolan-5 respectively. The HRA transcription units are pJ06B12R42A_AE(K), pJ07B2R42A_AE(K), pJ07B12R42A_AE(K), pJ02B2R42A_EF(K), pJ02B12R42A_EF(K) and nicknamed Tolan-6 through Tolan-10 respectively. The _AE(K) plasmids will be the first transcription unit in each dual expression cassette while the _EF(K) plasmids fill the second position. Each combination of pJXBXHRA_AE and pJXBXR42A_EF were combined to create the dual expression cassettes expressing HRA upstream of R42A. Likewise, each combination of pJXBXR42A_AE and pJXBXHRA_EF were also created for a combined total of twelve _AF(A) duel expression plasmids.*

J23102 is consistently the highest strength promoter in the CIDAR MoClo library. Likewise, BCD2 and BCD12 have been shown to be the strongest RBS parts in the library with BCD2 the stronger of the two by a small margin. The

78

pJ02B2X_EF and pJ02B12X_EF fluorescent reporters produced RFP and GFP in
similar quantities (Table 4-1).

Context	GFP	RFP
pJ06B12X_AE(K)	7764 ± 420	11165 ± 1112
pJ07B2X_AE(K)	7400 ± 404	8331 ± 981
pJ07B12X_AE(K)	8363 ± 159	17987 ± 837
pJ02B2X_EF(K)	5656 ± 248	13402 ± 1284
pJ02B12X_EF(K)	5924 ± 55	11530 ± 73
Average for goi	7021 ± 1179	12483 ± 3572
pJ02B2X_AE(K)	29470 ± 2559	50955 ± 2861
pJ142mX_AE(K)	194 ± 10	178 ± 7

All values noted as MEFL units of the geometric mean
± standard deviation measured by flow cytometry.

pJ02B2X_AE(K) and pJ142mX_AE(K) are shown here
to demonstrate the range of expression

**Table 4-1 Fluorescent protein expression data for each genetic context used in aldolase
tuning.** *This data was used to model expression of aldolase expression plasmids.*

4.2.2 Protein expression analysis of dual expression plasmids

HRA_CD and R42A_CD parts were cloned by PCR using template provided
by the Tolan lab. Transcription units composed as described above were
assembled with standard CIDAR MoClo protocols in DVK_AE and DVK_EF
vectors. Upon sequence verification these were further compiled into dual
expression plasmids (Tolan 1-9, Tolan 1-10, etc. as described in **Figure 4-1**) in
DVA_AF vectors. Sequence verified dual expression plasmids were provided to
the Tolan lab for expression analysis.

Due to the high degree of similarity between HRA and R42A, accurate
quantification of each variant in a dual expression system is difficult. Previous data
(**Figure 3-5a**) of fluorescent protein reporters suggests expression of a given

transcription unit when expressed alone or in a dual expression vector is unchanged as long as the same backbone is used. To provide an option for a proxy measurement of protein expression in single gene cassettes, individual transcription units were subcloned from the standard DVK_AE and DVK_EF vectors into DVA_AE and DVA_EF vectors respectively to produce Tolan-1(A) through Tolan-10(A).

Western blots with α-His antibody which recognize HRA but not R42A show high levels of HRA expression in all dual expression plasmids (**Figure 4-2a**) when compared to the control. Increasing HRA expression is the first step to creating equal dual expression plasmids. Importantly, expression of HRA appears fairly consistent between pairs of plasmids containing the same HRA transcription units (2-9 and 2-10, 6-5 and 7-5 for example). This same antibody used with single expression DVA plasmids showed less consistent expression, contrary to expectation, with Tolan-4(A) demonstrating the highest level of expression (**Figure 4-2b**). This inconsistency may be due in part to signal saturation in **Figure 4-2a**.
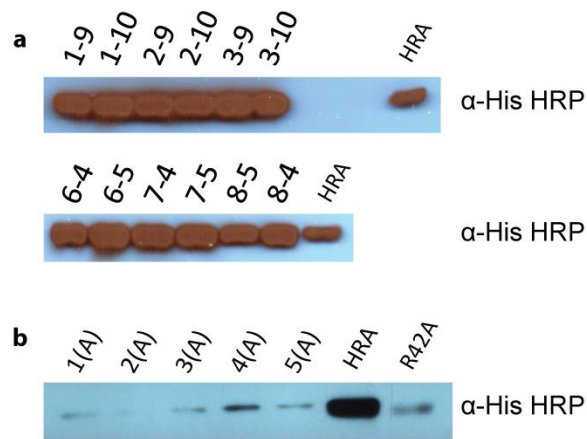
Figure 4-2 Western blots of aldolase protein expression. *(a) α-His antibody, expression of HRA in each of the dual expression clones is higher than the control demonstrating an improvement upon the previous plasmid design. Additionally, pairs of plasmids containing identical HRA TUs show similar expression (2-9 and 2-10, 3-9 and 3-10, 6-5 and 7-5), consistent with the fluorescent data for these genetic contexts. (b) Expression of HRA in the single TU DVA plasmids show less consistent expression of individual TUs with Tolan-4A showing the highest expression of HRA. Western blots were performed by Quinn Ho, Tolan Lab.*

Westerns performed on dual expression plasmids using anti-aldolase antibodies identify both aldolase variants resulting in doublet bands. Though faint, multiple samples appear to have approximately equal expression of HRA and R42A when probed with α-aldolase antibody, namely 3-9, 6-4, and 6-5 (Figure 4-3a). A second western performed with α-aldolase A antibody replicates these results with 6-4 and 6-5 showing the highest levels of expression for each variant in approximately equal quantity (Figure 4-3b).
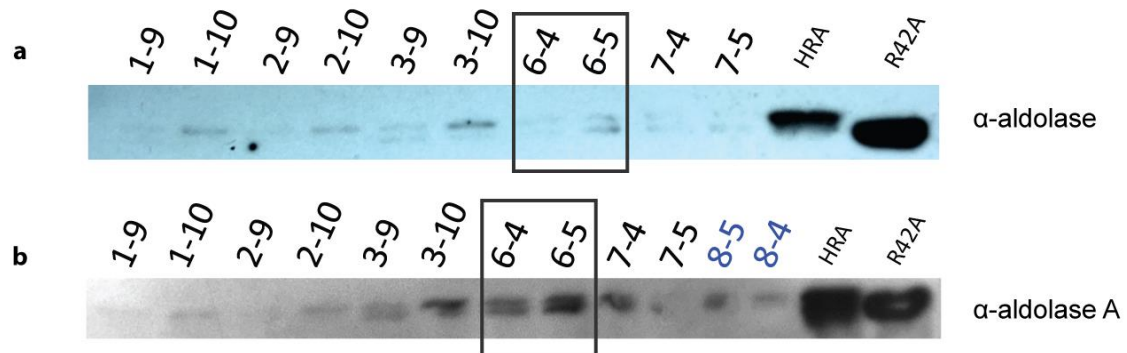
Figure 4-3 Expression analysis of aldolase HRA and R42A. *(a) α-aldolase antibody binds to both HRA and R42A presenting a doublet. Bands are faint, yet a few samples appear to contain approximately equal quantities of HRA and R42A (3-9, 6-4, 6-5, 7-4). (b) A second α-aldolase antibody, α-aldolase A, shows similar results with clearer bands. Of these, 6-4 and 6-5 appear to have the highest expression while presenting approximately equal signals for both variants. Western blots were performed by Quinn Ho, Tolan Lab.*

4.2.3 GFP and RFP fusion proteins as proxy expression measurements

A recent publication (Mutalik et al. 2013a) demonstrated the effect of various 5' coding sequences on protein expression by creating fusion proteins carrying the 5' 36-nt of a given gene fused to the 5' end of GFP or RFP under the control of a standard promoter and a series of ribosome binding sites. From this we hypothesized that fusion proteins carrying the 5' 36-nt of each aldolase variant could substitute as proxy measurements for expression levels of HRA and R42A in non-fusion plasmids. This information could inform on the expression patterns of the dual expression plasmids previously created as **Figure 3-5** demonstrated the consistency of expression in a given context when assembled into a multi-TU device.

Fluorescent fusion proteins of HRA:GFP, HRA:RFP, R42A:GFP and R42A:RFP were created to match Tolan-1 through Tolan-10 in terms of promoter, RBS and vector context. These were grown overnight along with the same contexts with GFP and RFP coding sequences for comparison (pJ06B12X_AE(K), etc. as described in **Figure 4-1**). Fluorescence of each clone was measured in triplicate with a Tecan plate reader.

As the BCD elements are designed to minimize variation in expression due to RBS:CDS interaction, it was expected that expression of a given context would remain fairly constant with the change of *goi* between GFP, HRA:GFP and R42A:GFP or between RFP, HRA:RFP, and R42A:RFP. However, it appears the 5' His-tag in HRA:GFP and HRA:RFP has more influence on gene expression than can be ameliorated with the BCDs.
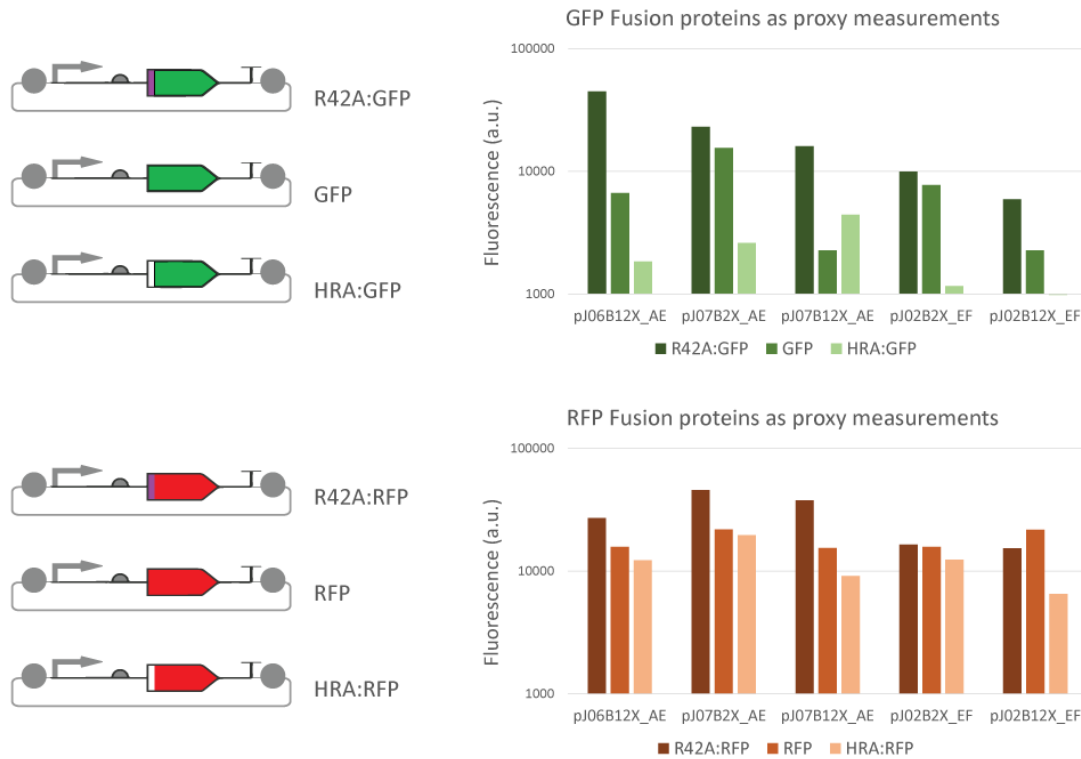
Figure 4-4 Fluorescence data of aldolase proxy plasmids. *Evaluating fluorescent protein fusion proteins as a proxy measurement of HRA and R42A expression. Error ranges are too narrow to be visible at this scale.*

This fluorescence analysis does appear to correlate with the western data presented in **Figure 4-2** and **Figure 4-3**, with HRA fusions fluorescing at notably lower levels (variance of 11.83-fold between R42A:GFP and HRA:GFP units). RFP comparisons are more closely related with an average variance of 2.47-fold between R42A:RFP and HRA:RFP. Overall, while some information can be discerned from the fusion proteins, the reliability of this method as a means of evaluation expression of a native protein is uncertain.

Considering that the BCDs have been shown previously to reduce variance 10-fold (Mutalik et al. 2013a), this suggests that expression of HRA without the influence of a BCD could be as much as 40-fold less than the R42A equivilant. With a BCD present the average variance between HRA:GFP and GFP TUs = 4.40-fold; variance between HRA:RFP and RFP TUs = 1.74-fold.

4.3   Discussion

4.3.1 Balanced expression of both aldolase variants in a single plasmid

While a more exhaustive approach could be employed in tuning expression in a system like the one demonstrated in this chapter, it does appear that we have succeeded at engineering equimolar expression within approximately a 2-fold error (estimated from western data) in 4 of the 12 dual expression plasmids created. Further quantification is pending optimization of protein separation methods. The strongest of these (Tolan-6-4 and Tolan-6-5) are currently being applied to structural analysis of the aldolase-actin binding interface.

The concept behind using fusion proteins as proxy measurements has some merit, however more analysis is needed. For example, data shown in **Figure 4-4** was gathered with a plate reader and represents population readings. Flow cytometry data of these cells may indicate multiple distinct populations within a single culture that could skew mean fluorescence values in a combined population. Additionally, more analysis could be performed to determine the length of the 5'

sequence that provides the most accurate representation of expression with the native CDS.

4.3.2 Limiting factors

The influence of the fusion site choice on expression created an upper limit of expression available for dual expression plasmids as even the highest strength combination of promoter and RBS in an _EF(K) vector would be approximately 3-fold weaker than the counterpart _AE(K) vector. Adaptation of the CIDAR MoClo assembly standard to include outside linkers or connectors as described for other MoClo standards could solve this limitation. Addition of linkers would allow for the standardization of the fusion sites directly flanking the transcription unit and facilitate the use of the same strong fusion site context with both TUs.

4.3.3 Alternative design strategies

The His-tag is a necessary component of downstream experiments regarding the isolated aldolase heterotetramers. However, it is likely unnecessary for the His-tag to be fused to the 5' end of the coding sequence rather than the 3' end where it would have less influence on gene expression (Goodman et al. 2013). Other tags may provide similar functions and have less influence on protein expression.

Inducible, rather than constitutive, expression also offers an alternative approach by which an inducer molecule could be used to control the expression of each variant. This approach would minimize the number of genetic designs

required and may reduce stress upon the cells caused by constitutive high expression of foreign proteins. However, careful control of induction protocols would be required and results may not be as consistent as with constitutive expression plasmids.

Another alternative may be to engineer a dimer-of-dimer fusion protein in which R42A and HRA are encoded as a single coding sequence, a fusion protein, to facilitate 2:2 ratios in heterotetramer formation. The choice of linker domains between the subunits would be critical, and other stochastic or structural issues may arise from the forced ratio. Such alternatives should be considered, however, if this dual expression plasmid and isolation of the correct heterotetramers proves to be unfruitful.

4.4   *Conclusions*

In summary, in Chapter 4 we have successfully tuned the expression of two variants of a rabbit aldolase protein by selecting genetic contexts based on previous expression data. Of the twelve dual expression plasmids examined, four appear to have approximately equal expression of both variants for a success rate of 33% in terms of balancing expression. Of these, two demonstrate both equal and high levels of protein expression (Tolan-6-4 and Tolan-6-5). These are undergoing further quantitative analysis and being applied to protein electron microscopy experiments to decorate F-actin and determine the actin-aldolase binding interface.

5 CHAPTER 5 DEVELOPMENT OF BACTERIAL BIOELECTRONIC LOGIC GATES WITH pH AND FLUORESCENCE AS OUTPUT SIGNALS

5.1 Introduction

A modular DNA assembly system lays the groundwork for a wide variety of applications from cell-free expression of combinatorial designs to rapid prototyping of bioelectronic sensors. Here we used CIDAR MoClo assembly to create plasmids containing DNA circuits which utilize a sugar input and a pH change as an output signal. Hydrolysis of lactose by beta-galactosidase reduces pH through the release of a hydrogen ion when breaking the bond between a galactose and glucose molecule.

Induction of LacZα expression decreases pH as measured by an Atlas Scientific pH probe. pH data is recorded by a Raspberry Pi processor. A change in pH can be used as a signaling mechanism to integrate biological inputs into digital outputs. Multiple pH probes attached to Raspberry Pi processors facilitate the development of bioelectronic Boolean logic gates based on previously published genetic designs (Tamsir et al. 2011).
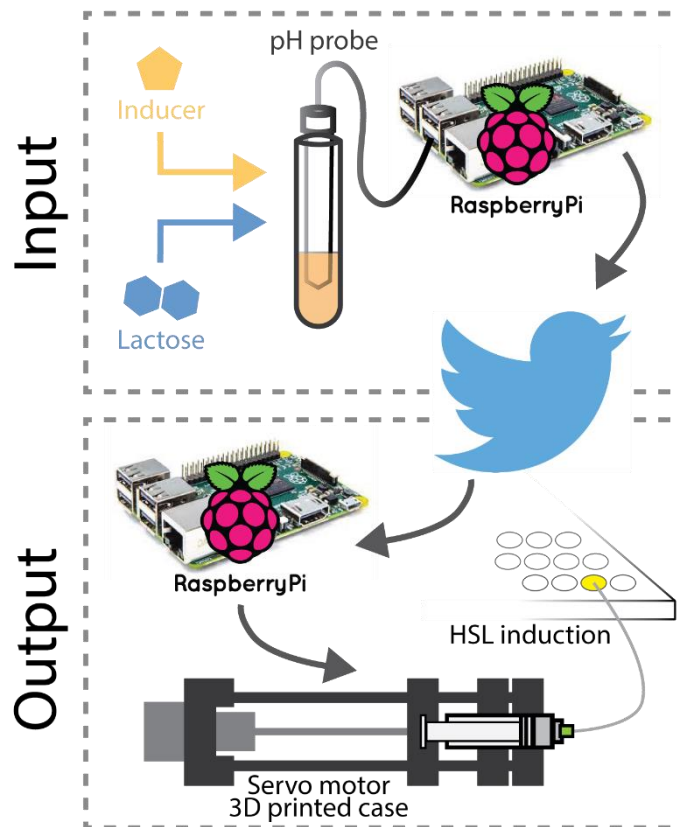
Figure 5-1 pH Experiment Schematic. *Inducer in the media turns on LacZα, expressing beta-galactosidase enzyme which metabolizes lactose resulting in an increase of hydrogen ions. This reduction in pH is measured with an Atlas Scientific pH probe attached to a Raspberry Pi. The Raspberry Pi reports pH every three seconds until a predetermined threshold is reached. Upon reaching threshold, a signal is sent via the Twitter API and a tweet is posted online. This tweet triggers the next stage when the signal is received by another Raspberry Pi which controls a 3D printed syringe pump calibrated to inject 100 μL of a second inducer molecule into inducible YFP expression cells.*

5.2   Results

5.2.1 Induction of LacZ mediated pH change

Two plasmids were evaluated as inducible pH circuits, pBADLacZDV:J024mAraC_AF(A) and pR10mLacZDV_AE(K) (**Figure 5-2**). To

89

evaluate the ability to measure a distinct change in pH with lactose induction, a dosage response curve was done using pR10LacZDV_AE(K), a modularly assembled form of the DVK_AE destination vector. This plasmid contains an interchangeable promoter (R0010_AB, pLac) and combined RBS, CDS, and terminator LacZDV_BE. This modularity allows for the simple creation of multiple LacZ induction plasmids by varying the promoter part.
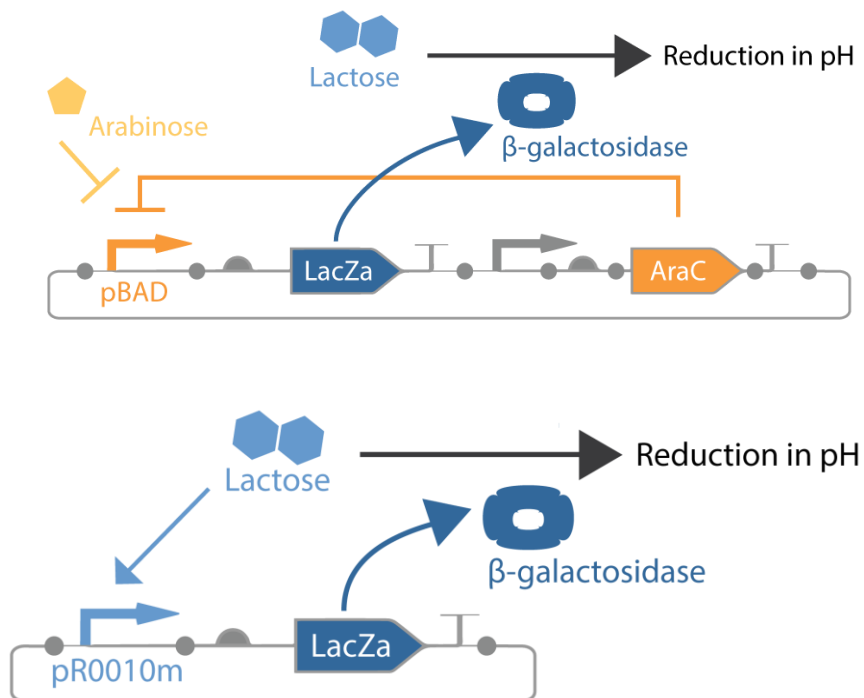
Figure 5-2 LacZ under the control of pBAD and AraC (above) and R0010m (below). Top: pBADLacZDV:J024mAraC_AF(A). Bottom: pR10mLacZDV_AE(K) These constructs allow for inducible pH reduction.

In **Figure 5-3**, DH5α cells were grown overnight in the presence of lactose and pH was measured at 14 hours. Under these conditions, a significant change in pH was seen with induction of 50 µM lactose (**Figure 5-3**) with

pR10LacZDV_AE(K) in DH5α cells. As this appears to be a saturating concentration, an induction curve was performed to determine whether a lower dose would suffice. The resulting curved showing saturation between 25 and 50 mM (data not shown); therefore a 50 mM induction concentration was used for the next round of experiments.

In order to mimic the logic gates presented in by the Voigt lab (Tamsir et al. 2011), we transformed the plasmids into TOP10 cells, a DH10B strain. The previous experiment was then replicated with pR10m controlling LacZ expression as a control and pBAD with arabinose induction as a new circuit (**Figure 5-2**). This data failed to replicate the results (**Figure 5-4**), with 50 mM of lactose failing to decrease pH with pR10m as drastically as seen previously. Additionally, expression of LacZα under the control of pBAD and AraC demonstrates a dramatic change in pH in the presence of arabinose alone.

INDUCTION OF PH CHANGE IN PR10LACZDV_AE(K)

	0mM	50mM	100mM
pR10mLacZDV_AE(K)	8.7	5.0	4.9
pR10LacZDV_AE(K)	7.5	5.0	4.9
pJ02B2GFP_AE(K)	6.5	6.5	6.8
DVK_AE	7.6	4.9	5.0
DH5a	8.2	8.4	8.5
LB+Kan	6.9	6.8	6.7
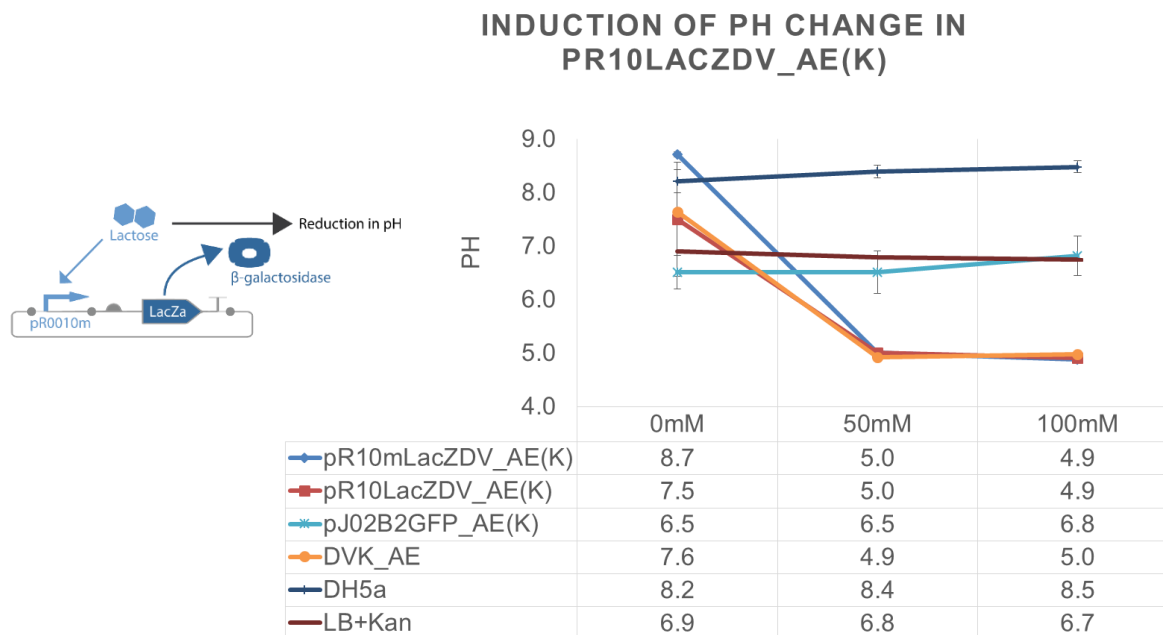
Figure 5-3 Induction of pH change using lactose and pLac promoter driving LacZa in DH5a cells. pR10mLacZDV_AE(K) differs only slightly from pR10LacZDV_AE(K), containing additional basepairs to match the DVK_AE sequence. Cells grown overnight (14h) in the presence of lactose.
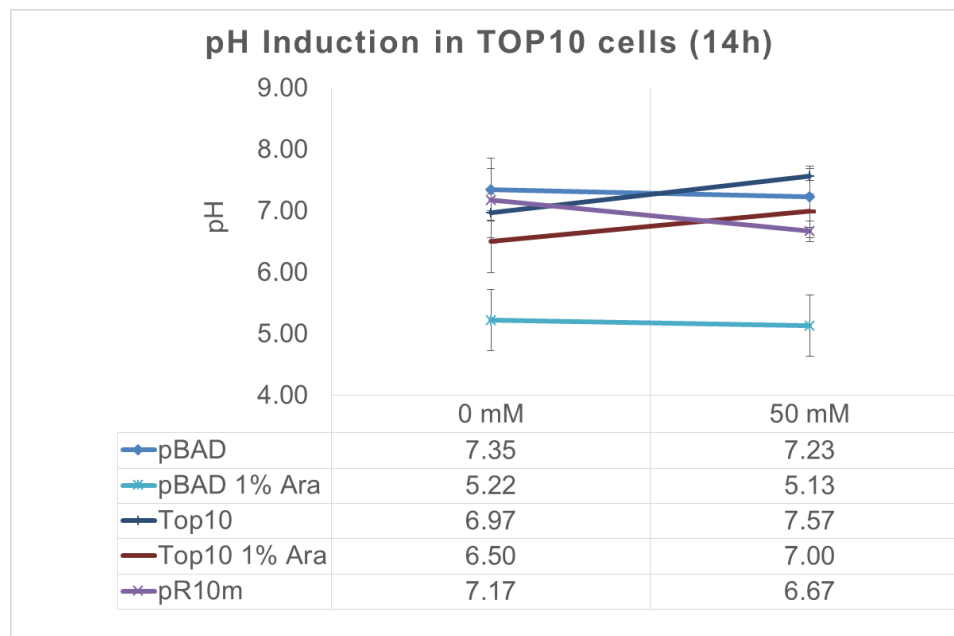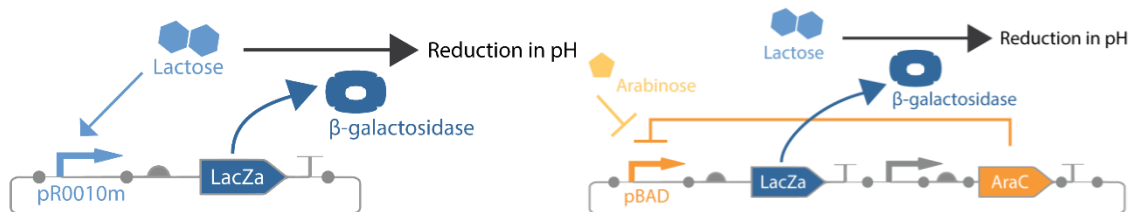
Figure 5-4 pH response with pBAD controlled LacZ expression. DH10B cells were grown overnight (14h) in the presence of lactose. pBAD controlled LacZ expression shows pH change in response to arabinose alone. Additionally, R10m controlled LacZ expression fails to induce a substantial shift in pH with a 50 mM lactose concentration as seen in DH5α cells.

pH change of the R0010m control is much less substantial in **Figure 5-4** than is seen in DH5α cells. In an attempt to recreate the dramatic pH difference between controls and experimental cells, a second induction curve was performed with 0, 50, and 150 mM lactose (**Figure 5-5**). pR10m cells displayed a significant change in pH upon induction with 150 mM lactose ($p < 0.001$). Having recalibrated induction for DH10B cells, 150 mM or 200 mM was used in experiments going forward.
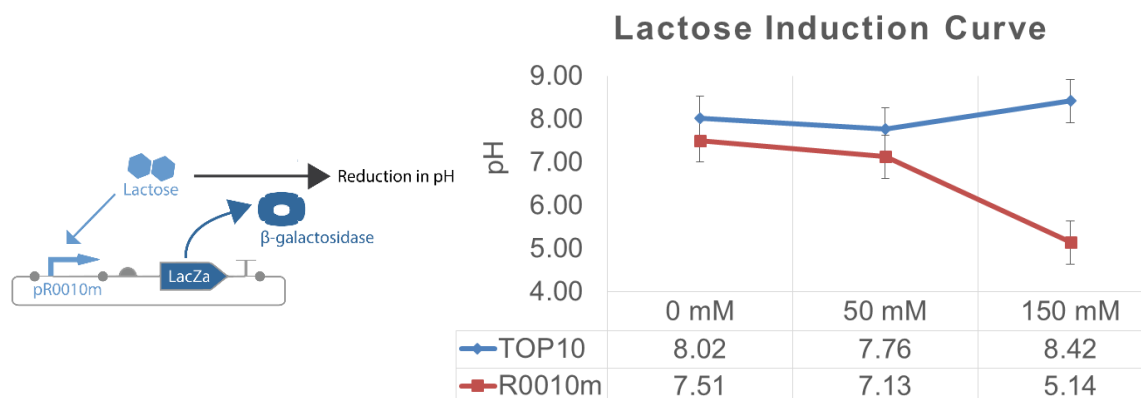
Figure 5-5 Lactose induction curve in DH10B cells. *Consistent with Figure 5-4, lactose induction appears to be less efficient in DH10B cells, requiring doses of 150 mM or 200 mM (not shown here) for substantial decrease in pH.*

All pH experiments to this point were performed as endpoint analyses. To determine the effects of incubation time, further induction experiments were conducted with pH measured at intervals (**Figure 5-6**). Cells expressing LacZ under the control of pBAD were incubated with or without 0.1% arabinose and 150 mM lactose. A reduced concentration of arabinose was used in an attempt to prevent arabinose mediated pH change. At 7.5 hours, a difference in pH can be seen between cells treated with arabinose (pH = 6.67) and those treated with arabinose and lactose (pH = 6.00). By tuning LacZ expression in this time course, we were able to define the parameters under which pBAD induced LacZα could be used as a logic gate component.
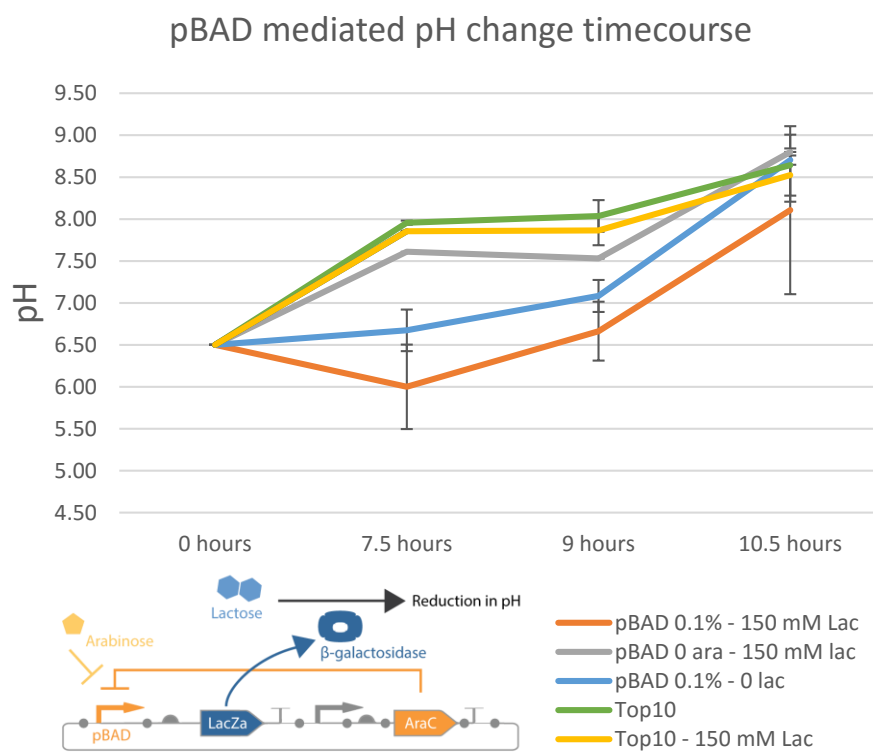
Figure 5-6 pBAD mediated pH change time course. *In order to optimize pBAD mediated pH change a time course was performed to observe pH at various intervals. Shorter incubations appear to provide a window in which pH is substantially different in in the presence of lactose and arabinose when compared to either arabinose alone or no sugar inputs.*

To optimize the induction and pH component of a bioelectronic logic gate, we wished to reduce the time required for signal acquisition. To do this, cells were grown overnight without inducer or lactose and diluted 1:2 into fresh culture tubes. These tubes were then induced and grown for three hours. This three hour protocol was sufficient to generate a pH change in R0010m controlled LacZ expression cells ($p < 0.01$) and in pBAD controlled cells ($p < 0.01$) despite the background caused by high arabinose concentration.
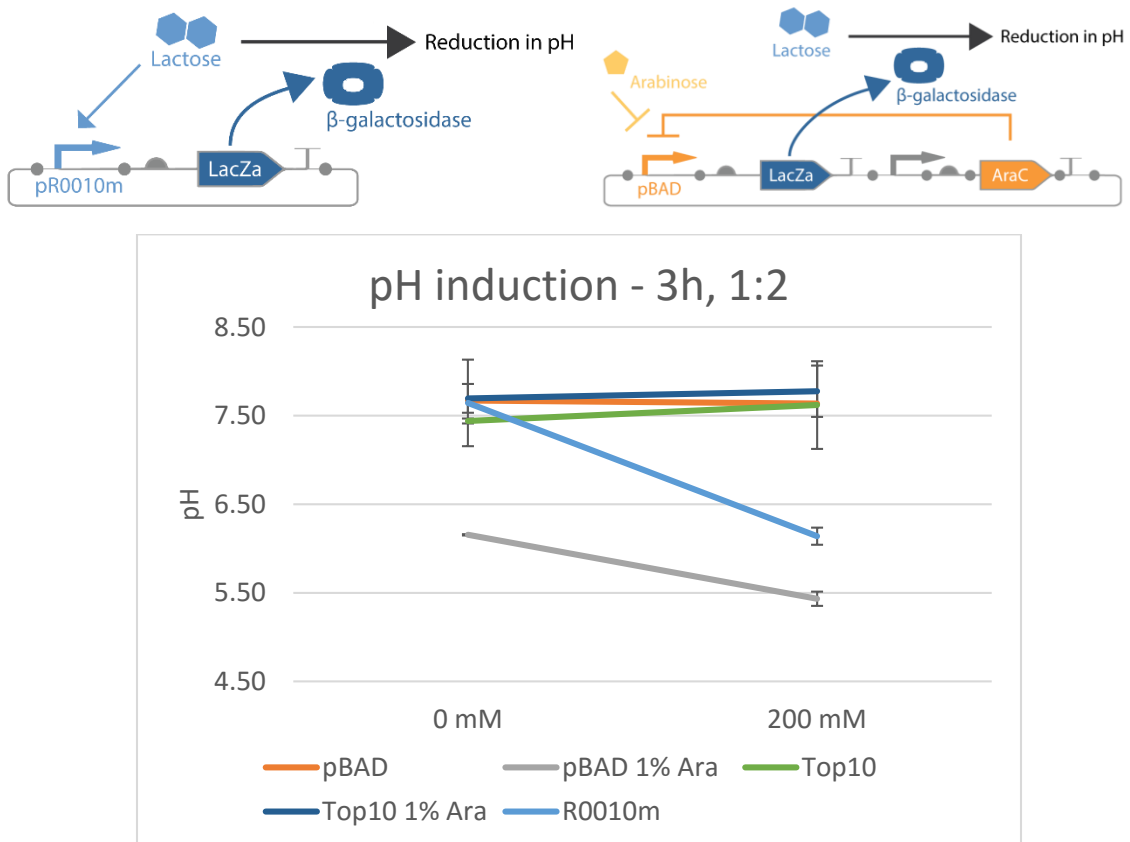
Figure 5-7 Three hour pH induction protocol. *This rapid protocol decreases the time required for signal acquisition by decreasing pH in R0010m controlled cells from 7.65 to 6.14 (p < 0.05). This experiment was run concurrently with Figure 5-6, using a 1% arabinose induction which is seen here to decrease pH in the absence of lactose. Lower concentration is used in the final experiment.*

5.2.2 pH mediated induction of Yellow fluorescent protein (YFP)

A YFP expression plasmid was acquired from the Voigt lab in order to closely replicate the logic gates described in Tamir et. al (2011). pOR30, shown in **Figure 1-5**, expresses YFP in the presence of homoserine lactone (HSL) under the control of the pLas promoter. Having tuned the expression of the pR10m and

pBAD controlled LacZ expression plasmids in the previous experiments, these two plasmids were used as the basis of a bioelectronic OR gate.

An induction curve was performed to determine the concentration of HSL necessary to rapidly induce YFP expression in DH10B cells carrying the pOR30 plasmid. Doses of 10 and 100 mM were analyzed with flow cytometry at 30' intervals from t=0 to t=180. A concentration of __ was selected as an optimal induction solution.

5.2.3 Cellular logic with sugar input, pH intermediate and fluorescent output

A culture tube containing 200 mM lactose in LB (1 mL) was inoculated with 500 µL of either pR10mLacZ and or pBADLacZ from an overnight culture. The pBADLacZ sample was also induced with 0.1% arabinose. pH probes wired to independent Raspberry Pi processors were inserted into each tube. The tubes were placed into a shaking incubator and pH was reported every three seconds.

A 96 well plate containing only controls and one sample well (uninduced pOR30 cells in LB) was prepared and the wells covered with breathable film. HSL was loaded into the syringe which was placed into the 3D printed syringe pump calibrated to inject 100 µL (± 4 µL) at a time. This syringe pump was attached to a third Raspberry Pi programmed to monitor the @Tweecoli Twitter account for a post reading "True" for a pH change in either culture when pH reached a predefined threshold (pH of 5.5).

Upon either culture reaching a pH of 5.5, a tweet would be posted live online stating either "Lac True at <Date> <Time>" or "Ara True at <Date> <Time>", respectively. Either of these posts would be accepted by the third Raspberry Pi as a signal to activate the servo motor in the syringe pump and inject 100 μL of HSL into a well containing uninduced cells.
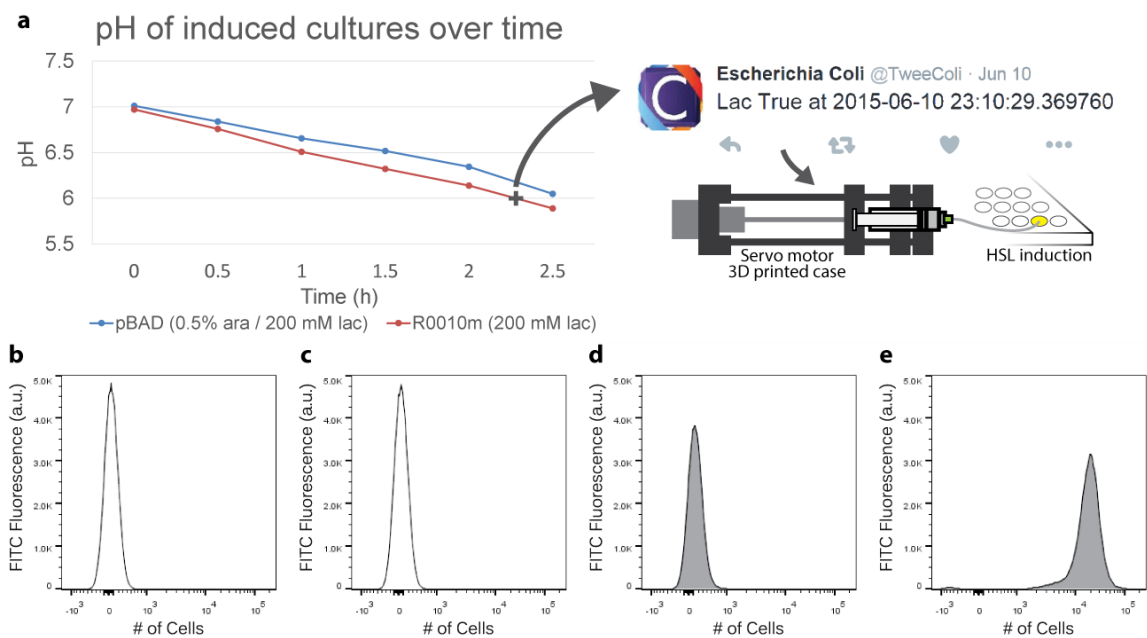
Figure 5-8 pH mediated logic gate with YFP expression output. (a) pH at 30 minute intervals of cultures containing DH10B cells with pBADLacZ:J06B2AraC_AF(A) plasmid, induced with arabinose (described in Figure 5-2) in media containing 200 mM lactose (blue) and DH10B cells with pR10mLacZ_AE(K) plasmid induced with 200 mM lactose alone (red). When pH of either sample falls below 5.5, a Tweet is posted via the Twitter API which functions as a trigger for a second Raspberry Pi which mediates induction of a fluorescent reporter by controlling the motor embedded in a 3D printed syringe pump containing inducer. (b-e) Fluorescence measurement of control and experimental samples from 96 well plate with HSL induction pump demonstrating induction of specified well with 2 h. incubation at 37C 900 rpm – (b) DH10B cells, no plasmid, - HSL. (c) DH10B cells, no plasmid, + HSL. (d) DH10B cells, pOR30 plasmid, - HSL. (e) DH10B cells, pOR30 plasmid, + HSL.

Figure 5-8 demonstrates a functioning bioelectronic logic gate in which a biological input (sugar) is measured via a pH change and transmitted as a digital output (Twitter) which in turn triggers a mechanical input mechanism (HSL injection) inducing YFP expression as a biological output. This experiment was designed to mimic a previously published logic gate (Tamsir et al. 2011) to demonstrate the utility of MoClo in the development of bioelectronic cellular logic. While this proof of concept experiment is relatively simple, it lays the groundwork for more complex forays into bioelectronic engineering.

5.2.4 CNC milled 96-well plate adaptions for customized cellular growth and analysis experimentation

Computer numerical control (CNC) milling offers a wide range of applications in synthetic biology. In order to lay basic groundwork for ongoing development of CNC milled devices in the CIDAR lab, we have performed preliminary analysis of growth and fluorescence measurement capabilities in polycarbonate devices (**Figure 5-9**). Cells grown overnight in 100 µL of LB in this 5x5 PCA well array had an OD600 of 0.7 and showed normal fluorescence patterns with only pJ02B2Rm_AE(K) registering on the red channel and both pJ02B2Rm_AE(K) and pJ02B2Gm_AE(K) registering on the green channel with the red expression cells showing normal autofluorescence.
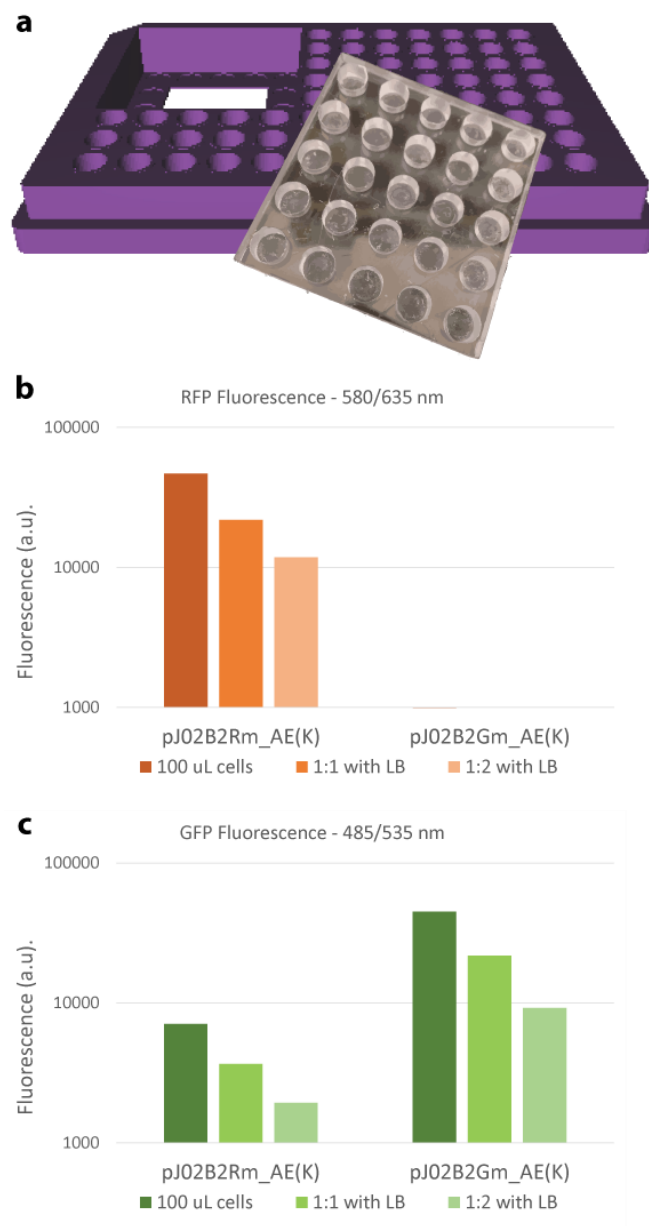
Figure 5-9 CNC milled 96-well plate adaptation for customized cellular growth and analysis experimentation. *Polycarbonate plastic (PCA) can be milled to specification for a wide range of purposes. (a) Proof of concept 3D printed 96-well plate frame designed to hold a CNC milled 5x5 well array to evaluate growth and fluorescence measurements in PCA. OD600 ~ 0.80 after 12 hours growth at 37°C 900 rpm. (b) RFP fluorescence was measured on a plate reader with excitation at 580 nm and emission at 635 nm. (c) GFP fluorescence was measured at 485/535 nm and shows some background fluorescence as is common with FITC channel measurements of live cells.*

5.3   Discussion

We have successfully built and tested a bioelectronic OR gate with pBAD and R0010m mediated LacZ expression vectors. This demonstration provides the basis for the development of more complex and more physically compact devices. Additionally, preliminary experiments show normal cellular growth and fluorescence measurement ability in PCA milled devices.

A pLac promoter, R0010m, reliably controls pH of DH5α and DH10B cells through expression of LacZα and induction with lactose (50 mM in DH5α, 150-200 mM in DH10B). pBAD mediated expression of LacZα is more intricate, requiring careful moderation of arabinose concentration to avoid triggering a pH change with arabinose alone. Induction of cells carrying the pBADLacZDV:J024mC80_AF(A) to create a substantial pH change was performed with a 0.1% concentration of arabinose and 200 mM concentration of lactose with a peak variance in pH seen at or before 7.5 hours.

pTetLacZDV:J024mC40_AF(A) was also evaluated as an inducible pH circuit. This design proved to be even less responsive to inducer than the pBAD mediated device. A schematic of this circuit (**Figure B-3**) and time course induction data (**Figure B-4**) can be found in Appendix B.2.

Though a previously published biosensor has demonstrated this LacZ-pH signaling mechanism (de Mora et al. 2011) the pH induction plasmids evaluated here indicate a high amount of noise in biological systems with relation to pH.

Growth, optical density, and other metabolic pathways influence pH as seen with the data in this chapter in which pH of uninduced cultures shifts over time and arabinose alone triggers a change in pH. Though pH may be a viable bioelectronic interface signal, other interface methods should also be considered moving forward.

Polycarbonate plastic (PCA) can be milled to specification for a wide range of purposes and was shown here to be a viable option for containment of cells during growth and fluorescence analysis. Thickness of the bottom surface through which OD and fluorescence measurements were made was kept standard at 3 mm. Under the conditions reporter here, growth and fluorescence measurement appear normal compared to using standard 96 well plates.

RFP fluorescence was measured on a plate reader with excitation at 580 nm and emission at 635 nm. GFP fluorescence was measured at 485/535 nm and shows some background fluorescence as is common with FITC channel measurements of live cells. In both cases readings were taken from the bottom of the plate and a gain of 87 (red) and 89 (green) was used as determined by the "optimal gain" option on the Tecan plate reader.

The applications of 3D printed and CNC milled custom devices in biological sciences is infinite, with potential uses including small tools like the syringe pump and milled microfluidics as well as larger scale applications in bioproduction and metabolic engineering. By producing micromilled devices designed to fit within a

96 well plate frame as shown in **Figure 5-9** many measurement and analysis tools are readily adaptable to custom formats. For example, centrifugal fluidic devices which work by spinning a flat plastic disc to move liquid through channels, can be cut to fit within a 96 well plate with "output" wells aligned for use in a plate reader or other 96-well compatible equipment. Further research into such practices is ongoing in the CIDAR lab.

5.4   Conclusions

Though a previously published biosensor has demonstrated this LacZ-pH signaling mechanism (de Mora et al. 2011) the pH induction plasmids evaluated here indicate a high amount of noise in biological systems with relation to pH. Growth, optical density, and other metabolic pathways influence pH as seen with the data in this chapter in which pH of uninduced cultures shifts over time and arabinose alone triggers a change in pH. Though pH may be a viable bioelectronic interface signal, other interface methods should also be considered moving forward.

CNC milled devices have the potential to be infinitely advantageous in the development of custom bioelectronic circuits, micromilled microfluidics, and other areas of research under the umbrella of synthetic biology. Preliminary analysis here demonstrates the ability to use PCA as a chamber for growing cells and measuring fluorescence on red and green channels. Further analysis is needed to determine the effects of polishing rough surfaces, movement of cell cultures

103

through milled channels, and the ability to measure other fluorescence in other wavelengths.

These bioelectronic sensor experiments, paired with the aldolase expression tuning experiments described in Chapter 4 demonstrate the utility of CIDAR MoClo in both synthetic and traditional biology research. The CIDAR MoClo Library is the first bacterial multipart modular DNA library available. It is the hope of the author that these materials allow for the wider adoption of MoClo as a research and educational tool.

Appendix A – Additional MoClo data

A.1 MoClo protocol optimization

In order to determine the optimal buffer and ligase for use with MoClo, an array of components was compared. This data represents one transformation reaction per line and was used as a preliminary assessment. Zymo cells are chemically competent cells produced in house from the Bioline Alpha Select DH5α strain. While the Zymo protocol is intended to produce extremely high competent cells and negate the need for heat shocking and recovery, this rapid transformation protocol works only for isolated plasmids. Both Bioline and Zymo cells were used with a standard heat shock protocol as described in Chapter 2.

Construct	Buffer	Ligase	Cells	White	Blue	% White
pJ00B2Rm_EF	Promega	Promega	Bioline	118	24	83%
4 part	NEB	NEB	Bioline	260	50	84%
	Promega	NEB	Bioline	740	40	95%
	NEB+BSA	NEB	Bioline	374	0	100%
	Promega	Promega	Zymo	458	58	89%
	NEB	NEB	Zymo	230	42	85%
	Promega	NEB	Zymo	164	10	94%
	NEB+BSA	NEB	Zymo	208	110	65%
pJ00B2CAM_EF	Promega	Promega	Zymo	172	2	99%
4 part	NEB	NEB	Zymo	224	38	85%
	Promega	NEB	Zymo	88	28	76%
	NEB+BSA	NEB	Zymo	30	5	86%
pJ00ERCreER_EF	Promega	Promega	Zymo	166	28	86%
6 part	NEB	NEB	Zymo	66	14	83%
	Promega	NEB	Zymo	60	20	75%
	NEB+BSA	NEB	Zymo	40	46	47%

Table A-1 Preliminary optimization of MoClo protocols. *Initial steps compared NEB and Promega ligase reagents and competent cell type.*
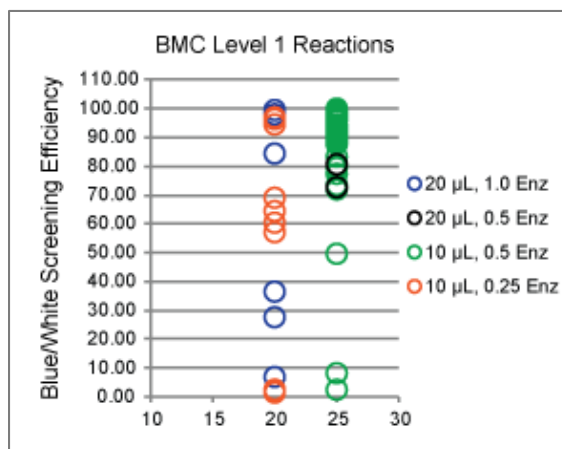
Figure A-1 MoClo cycle and volume optimizations. BMC = basic MoClo, standard assembly protocol, 4 part transcription unit. 10 µL reactions with 0.5 µL of restriction enzyme and 25 cycles resulted in the highest level of consistent assembly efficiency. Lower enzyme concentration and fewer cycles still results in sufficient correct clones despite lowered efficiency.

CIDAR MoClo Protocols

	Step	Temp	Time (min)
Standard	Step 1	37°C	1.5
	Step 2	16°C	3
	Cycle 1-2	x15	
	Step 3	50°C	5
	Step 4	80°C	10
	Total time		82.5

	Step	Temp	Time (min)
Troubleshooting	Step 1	37°C	1.5
	Step 2	16°C	3
	Cycle 1-2	x25	
	Step 3	50°C	5
	Step 4	80°C	10
	Total time		127.5

	Step	Temp	Time (min)
Rapid	Step 1	37°C	20
	Step 2	37°C	1.5
For new basic parts	Step 3	16°C	3
	Cycle 2-3	x5-10	
	Step 4	50°C	5
	Step 5	80°C	10
	Total time		37.5-60

Reaction Conditions

Basic Part or Device (DVA Reactions)

10 fmol each part

1x Promega Ligase buffer

20 U/rxn T4 Ligase

10 U/rxn BbsI

Total Volume: 10 -20uL

Transcriptional Unit (DVK Reactions)

10 fmol each part

1x Promega Ligase buffer

20 U/rxn T4 Ligase

10 U/rxn BsaI

Total Volume: 10 -20uL

Transform 2-5 uL per reaction

Either NEB or Promega Ligase

Table A-2 CIDAR MoClo Protocols as published in ACS Synth Biol 2015.

Main matrix — each cell is a 4-mer; column header gives nucleotides 1·3, row header gives nucleotides 2·4. Parenthetical codes are printed cell labels (green = assigned, orange = RC).

2·4 \ 1·3	G·G	G·T	G·C	G·A	T·G	T·T	T·C	T·A	C·G	C·T	C·C	C·A	A·G	A·T	A·C	A·A
A·A	GAGA	GATA	GACA	GAAA	TAGA (I*)	TATA	TACA	TAAA	CAGA	CATA (C2*)	CACA (Y*)	CAAA	AAGA	AATA	AACA	AAAA
A·C	GAGC	GATC	GACC	GAAC (D1*)	TAGC	TATC	TACC	TAAC	CAGC	CATC (C1*)	CACC	CAAC	AAGC (E*)	AATC	AACC	AAAC
A·T	GAGT	GATT	GACT	GAAT	TAGT (H*)	TATT	TACT (B)	TAAT	CAGT	CATT (C*)	CACT	CAAT	AAGT	AATT	AACT	AAAT
A·G	GAGG	GATG (C1)	GACG	GAAG	TAGG	TATG (C2)	TACG	TAAG	CAGG (D2*)	CATG	CACG	CAAG	AAGG	AATG (C)	AACG (X*)	AAAG
C·A	GCGA	GCTA	GCCA	GCAA	TCGA	TCTA (I)	TCCA	TCAA	CCGA	CCTA	CCCA	CCAA	ACGA	ACTA (H)	ACCA	ACAA
C·C	GCGC	GCTC	GCCC	GCAC	TCGC	TCTC	TCCC	TCAC	CCGC	CCTC	CCCC	CCAC	ACGC	ACTC	ACCC	ACAC
C·T	GCGT	GCTT (E)	GCCT	GCAT	TCGT	TCTT	TCCT	TCAT	CCGT	CCTT	CCCT	CCAT	ACGT	ACTT	ACCT (D*)	ACAT
C·G	GCGG	GCTG	GCCG	GCAG	TCGG	TCTG	TCCG	TCAG	CCGG	CCTG (D2)	CCCG	CCAG	ACGG	ACTG	ACCG	ACAG
T·A	GTGA	GTTA	GTCA	GTAA	TTGA	TTTA	TTCA	TTAA	CTGA	CTTA	CTCA	CTAA	ATGA	ATTA	ATCA	ATAA
T·C	GTGC	GTTC (D1)	GTCC	GTAC	TTGC	TTTC	TTCC	TTAC	CTGC	CTTC	CTCC (A*)	CTAC	ATGC	ATTC	ATCC	ATAC
T·T	GTGT	GTTT	GTCT	GTAT	TTGT	TTTT	TTCT	TTAT	CTGT	CTTT	CTCT	CTAT	ATGT	ATTT	ATCT	ATAT
T·G	GTGG	GTTG	GTCG (J*)	GTAG	TTGG	TTTG	TTCG	TTAG	CTGG	CTTG	CTCG	CTAG	ATGG	ATTG	ATCG	ATAG
G·A	GGGA	GGTA	GGCA (G*)	GGAA	TGGA	TGTA	TGCA	TGAA	CGGA	CGTA	CGCA	CGAA	AGGA	AGTA (B*)	AGCA	AGAA
G·C	GGGC	GGTC	GGCC	GGAC	TGGC	TGTC	TGCC (G)	TGAC	CGGC	CGTC	CGCC	CGAC (J)	AGGC	AGTC	AGCC	AGAC
G·T	GGGT	GGTT	GGCT	GGAT	TGGT	TGTT	TGCT	TGAT	CGGT	CGTT (X)	CGCT (F)	CGAT	AGGT (D)	AGTT	AGCT	AGAT
G·G	GGGG	GGTG	GGCG	GGAG (A)	TGGG	TGTG (Y)	TGCG	TGAG	CGGG	CGTG	CGCG	CGAG	AGGG	AGTG	AGCG (F*)	AGAG

Summary assignment rows:

	A	B	C	D	E	F	G	H	I	J	X	Y	C1	C2	D1	D2
FS	GGAG (A)	TACT (B)	AATG (C)	AGGT (D)	GCTT (E)	CGCT (F)	TGCC (G)	ACTA (H)	TCTA (I)	CGAC (J)	CGTT (X)	TGTG (Y)	GATG (C1)	TATG (C2)	GTTC (D1)	CCTG (D2)
RC	CTCC (A*)	AGTA (B*)	CATT (C*)	ACCT (D*)	AAGC (E*)	AGCG (F*)	GGCA (G*)	TAGT (H*)	TAGA (I*)	GTCG (J*)	AACG (X*)	CACA (Y*)	CATC (C1*)	CATA (C2*)	GAAC (D1*)	CAGG (D2*)

Fusion site choice is crucial for maintaining high efficiency with MoClo. Table A-3 acts as a guide for designating new sites. Assigned sites are indicated in green with the reverse complement of these sites highlighted orange. Pink sequences have only a 1 base pair variation from one or more previously designated fusion sites and may cause a decrease in efficiency if used with these sites. Red sequences are either repetitive or palindromic and should be avoided at all times.

A.2   Primers for CIDAR MoClo parts and vectors

Primer	Sequence
Part Forward	5'ATGAAGACGTnnnn-20nt-
Part Reverse	5'ACGAAGACCTnnnn-20nt-
DV0/DV2/DVA Forward	5'GTACTAGTGGGTCTCAnnnnATGTCTTCtgcaccatatgcggt
DV0/DV2/DVA Reverse	5'CTACTAGTAGGTCTCTnnnnACGTCTTCcccgcgcgttggccgat
DV1/DVK Forward	5'GTACTAGTGGAAGACATnnnnAGAGACCtgcaccatatgcggtgtgaaatac
DV1/DVK Reverse	5'CTACTAGTAGAAGACATnnnnAGAGACCcccgcgcgttggccgattcattaa

Table A-4 Primer design strategy for new basic parts and vectors. At least 20 nucleotides of annealing part sequence is suggested for creation of new basic parts. CDS annealing sequence begins with the second codon as the "ATG" codon is included in the C fusion site. Sequences are color coded; red = BbsI, blue = BsaI, grey = SpeI, green 'nnnn' = fusion site in forward orientation, brown 'nnnn' = reverse compliment of fusion site.

A.3   Parts for 3Ab CIDAR MoClo part collection (initial)

Vector	Level	Antibiotic	5' Flanking Region (5' to 3')	3' Flanking Region (5' to 3')
DVL0_AB	0	CAM	GGTCTCAGGAGATGTCTTC	GAAGACGTTACTAGAGACC

Vector	Level	Antibiotic	5' Flanking Region (5' to 3')	3' Flanking Region (5' to 3')
DVL0_BC	0	CAM	GGTCTCATACTATGTCTTC	GAAGACGTAATGAGAGACC
DVL0_CB	0	CAM	GGTCTCAAATGATGTCTTC	GAAGACGTTACTAGAGACC
DVL0_CD	0	CAM	GGTCTCAAATGATGTCTTC	GAAGACGTAGGTAGAGACC
DVL0_CI	0	CAM	GGTCTCAAATGATGTCTTC	GAAGACGTTCTAAGAGACC
DVL0_CX	0	CAM	GGTCTCAAATGATGTCTTC	GAAGACGTCGTTAGAGACC
DVL0_CY	0	CAM	GGTCTCAAATGATGTCTTC	GAAGACGTTGTGAGAGACC
DVL0_D1E	0	CAM	GGTCTCAGTTCATGTCTTC	GAAGACGTGCTTAGAGACC
DVL0_DC1	0	CAM	GGTCTCAAGGTATGTCTTC	GAAGACGTGATGAGAGACC
DVL0_DC2	0	CAM	GGTCTCAAGGTATGTCTTC	GAAGACGTTATGAGAGACC
DVL0_DE	0	CAM	GGTCTCAAGGTATGTCTTC	GAAGACGTGCTTAGAGACC
DVL0_DF	0	CAM	GGTCTCAAGGTATGTCTTC	GAAGACGTCGCTAGAGACC
DVL0_DG	0	CAM	GGTCTCAAGGTATGTCTTC	GAAGACGTTGCCAGAGACC
DVL0_DH	0	CAM	GGTCTCAAGGTATGTCTTC	GAAGACGTACTAAGAGACC
DVL0_EB	0	CAM	GGTCTCAGCTTATGTCTTC	GAAGACGTTACTAGAGACC
DVL0_FB	0	CAM	GGTCTCACGCTATGTCTTC	GAAGACGTTACTAGAGACC
DVL0_GB	0	CAM	GGTCTCATGCCATGTCTTC	GAAGACGTTACTAGAGACC
DVL0_ID	0	CAM	GGTCTCATCTAATGTCTTC	GAAGACGTAGGTAGAGACC
DVL0_XD	0	CAM	GGTCTCACGTTATGTCTTC	GAAGACGTAGGTAGAGACC
DVL0_XY	0	CAM	GGTCTCACGTTATGTCTTC	GAAGACGTTGTGAGAGACC
DVL0_YD	0	CAM	GGTCTCATGTGATGTCTTC	GAAGACGTAGGTAGAGACC
DVL1_AE	1	KAN	GAAGACATGGAGAGAGACC	GGTCTCTGCTTATGTCTTC
DVL1_EF	1	KAN	GAAGACATGCTTAGAGACC	GGTCTCTCGCTATGTCTTC
DVL1_FG	1	KAN	GAAGACATCGCTAGAGACC	GGTCTCTTGCCATGTCTTC
DVL1_GH	1	KAN	GAAGACATTGCCAGAGACC	GGTCTCTACTAATGTCTTC
DVL2_AF	2	AMP	GGTCTCAGGAGATGTCTTC	GAAGACGTCGCTAGAGACC
DVL2_AG	2	AMP	GGTCTCAGGAGATGTCTTC	GAAGACGTTGCCAGAGACC
DVL2_AH	2	AMP	GGTCTCAGGAGATGTCTTC	GAAGACGTACTAAGAGACC
DVL2_EH	2	AMP	GGTCTCAGCTTATGTCTTC	GAAGACGTACTAAGAGACC

Table A-5 Destination vectors from 3Ab CIDAR MoClo part collection. *Fusion sites are shown in purple, BsaI recognition sequences in red, and BbsI recognition sequences in blue. Black nucleotides indicate spacers required for appropriate placement of enzyme recognition sequences. Spacers can be made of any nucleotide, however standard sequences are used to reduce the potential for unintended restriction recognition sequences.*

Part ID	Description	Cloned by
B0015_DE	Double terminator	SI
B0015_DF	Double terminator	SI
B0015_DG	Double terminator	SI
B0015_DH	Double terminator	TH
B0030_BC	Weiss RBS	SI
B0031_BC	Weiss RBS	SI
B0032_BC	Weiss RBS	SI
B0033_BC	Weiss RBS	SI
B0034_BC	Weiss RBS	SI
C0012_CD	LacI - repressor CDS	EA
C0040_CD	TetR - repressor CDS	SI
C0051_CD	cI - repressor CDS	MF
C0062_CD	luxR - repressor CDS	EA
C0071_CD	rhlR - repressor CDS	EA
C0079_CD	lasR - activator CDS	MF
C0080_CD	araC - regulator CDS	EA
CyPet_CD	Flourescent protein CDS	EA
DsRed_Express_CD	Flourescent protein CDS	EA
DVL0_AB	Destination vector	SI
DVL0_BC	Destination vector	SI
DVL0_CD	Destination vector	SI
DVL0_DE	Destination vector	SI
DVL0_DF	Destination vector	SI
DVL0_DG	Destination vector	SI
DVL0_DH	Destination vector	
DVL0_EB	Destination vector	SI
DVL0_FB	Destination vector	
DVL0_GB	Destination vector	SI
DVL1_AE	Destination vector	SI
DVL1_EF	Destination vector	SI
DVL1_FG	Destination vector	SI
DVL1_GH	Destination vector	SI
DVL2_AF	Destination vector	SI
DVL2_AG	Destination vector	SI
DVL2_AH	Destination vector	SI
DVL2_EG	Destination vector	
DVL2_EH	Destination vector	SI
DVL2_FH	Destination vector	
E0030_CD	YFP fluorescent protein CDS	SI
E0040m_CD	GFP flourescent protein CDS	SI
E1010m_CD	RFP fluorescent protein CDS	SI
E2-Crimson_CD	Fluorescent protein CDS	EA
EBFP2_CD	BFP fluorescent protein CDS	TH
I13453_FB	pBAD promoter	

Part ID	Description	Cloned by
I13458_AF	pC promoter and AraC CDS	
iRFP_CD	iRFP fluorescent protein CDS	TH
J23100_AB	Constitutive promoter	MF
J23100_EB	Constitutive promoter	MF
J23101_AB	Constitutive promoter	MF
J23101_EB	Constitutive promoter	MF
J23102_AB	Constitutive promoter	MF
J23102_EB	Constitutive promoter	MF
J23103_AB	Constitutive promoter	MF
J23103_EB	Constitutive promoter	SI
J23104_AB	Constitutive promoter	MF
J23104_EB	Constitutive promoter	MF
J23105_AB	Constitutive promoter	MF
J23105_EB	Constitutive promoter	MF
J23106_AB	Constitutive promoter	MF
J23106_EB	Constitutive promoter	SI
J23107_AB	Constitutive promoter	MF
J23107_EB	Constitutive promoter	SI
J23108_AB	Constitutive promoter	MF
J23108_EB	Constitutive promoter	MF
J23109_AB	Constitutive promoter	MF
J23109_EB	Constitutive promoter	MF
J23110_AB	Constitutive promoter	SI
J23110_EB	Constitutive promoter	MF
J23111_AB	Constitutive promoter	MF
J23111_EB	Constitutive promoter	MF
J23112_AB	Constitutive promoter	MF
J23112_EB	Constitutive promoter	MF
J23113_AB	Constitutive promoter	TH
J23113_EB	Constitutive promoter	MF
J23114_AB	Constitutive promoter	MF
J23114_EB	Constitutive promoter	MF
J23115_AB	Constitutive promoter	SI
J23115_EB	Constitutive promoter	SI
J23116_AB	Constitutive promoter	SI
J23116_EB	Constitutive promoter	SI
J23117_AB	Constitutive promoter	SI
J23117_EB	Constitutive promoter	MF
J23118_AB	Constitutive promoter	SI
J23118_EB	Constitutive promoter	MF
J23119_AB	Constitutive promoter	SI
J23119_EB	Constitutive promoter	SI
LSSmOrange_CD	Fluorescent protein CDS	EA
mAmetrine_CD	Fluorescent protein CDS	EA

Part ID	Description	Cloned by
mCitrine_CD	Fluorescent protein CDS	EA
mNeptune_CD	Fluorescent protein CDS	EA
mOrange_CD	Fluorescent protein CDS	EA
R0010_AB	pLac promoter	EA
R0010_EB	pLac promoter	MF
R0010_FB	pLac promoter	EA
R0010_GB	pLac promoter	EA
R0040_AB	pTet promoter	SI
R0040_EB	pTet promoter	SI
R0040_FB	pTet promoter	SI
R0040_GB	pTet promoter	SI
R0051_AB	cI promoter	
R0051_EB	cI promoter	
R0051_FB	cI promoter	
R0051_GB	cI promoter	
R0062_AB	luxR pR promoter	
R0062_EB	luxR pR promoter	
R0062_FB	luxR pR promoter	
R0062_GB	luxR pR promoter	
R0063_AB	luxR pL promoter	MF
R0063_EB	luxR pL promoter	MF
R0063_FB	luxR pL promoter	EA
R0063_GB	luxR pL promoter	EA
R0071_AB	RhlR promoter	
R0071_EB	RhlR promoter	
R0071_FB	RhlR promoter	
R0071_GB	RhlR promoter	
R0079_AB	LasR promoter	EA
R0079_EB	LasR promoter	MF
R0079_FB	LasR promoter	EA
R0079_GB	LasR promoter	EA
T-Sapphire_CD	Fluorescent protein CDS	EA

Table A-6 Parts created as initial 3Ab CIDAR MoClo collection. All basic parts are in DVL0 backbones (CAM resistance). All DVL vectors are as described in the previous table. Parts with no initials in the "Cloned by" column were not verified and archived by the time the new CIDAR MoClo standard was developed and parts were moved to a DVA backbone. Note: C0012_CD described here contained an illegal BsaI site which was repaired to create C0012m_CD(A) in the new library.

A.4   Parts for CIDAR MoClo Library (Published version)

A set of 93 plasmids were submitted to Addgene (#1000000059) as described in Chapter 3. **Table A-7** describes the parts in the kit along

with their individual Addgene catalog number, CIDAR-ICE ID, and Benchling location. Additional plasmids can be requested directly from the CIDAR lab.

Plasmid	Part Description	Backbone	Addgene Cat#	CIDAR-ICE ID#	Benchling (benchling.com)
J23100_AB	Basic Part - Promoter (Constitutive)	DVA	65980	CIDAR_490	https://benchling.com/s/gS6iXc3X/edit
J23100_EB	Basic Part - Promoter (Constitutive)	DVA	65981	CIDAR_491	https://benchling.com/s/aq8Gxp2O/edit
J23100_FB	Basic Part - Promoter (Constitutive)	DVA	65982	CIDAR_492	https://benchling.com/s/rTzq9ToN/edit
J23100_GB	Basic Part - Promoter (Constitutive)	DVA	65983	CIDAR_493	https://benchling.com/s/5yn5Zy6g/edit
J23102_AB	Basic Part - Promoter (Constitutive)	DVA	65984	CIDAR_494	https://benchling.com/s/aQOLd1xr/edit
J23102_EB	Basic Part - Promoter (Constitutive)	DVA	65985	CIDAR_495	https://benchling.com/s/zOCQC6dM/edit
J23102_FB	Basic Part - Promoter (Constitutive)	DVA	65986	CIDAR_496	https://benchling.com/s/1pfY4NG/edit
J23102_GB	Basic Part - Promoter (Constitutive)	DVA	65987	CIDAR_497	https://benchling.com/s/LmBWThyt/edit
J23103_AB	Basic Part - Promoter (Constitutive)	DVA	65988	CIDAR_498	https://benchling.com/s/9KEsWqta/edit
J23103_EB	Basic Part - Promoter (Constitutive)	DVA	65989	CIDAR_499	https://benchling.com/s/ynovZBP9/edit
J23103_FB	Basic Part - Promoter (Constitutive)	DVA	65990	CIDAR_500	https://benchling.com/s/oZSwbWfi/edit
J23103_GB	Basic Part - Promoter (Constitutive)	DVA	65991	CIDAR_501	https://benchling.com/s/4qwyO9ja/edit
J23106_AB	Basic Part - Promoter (Constitutive)	DVA	65992	CIDAR_502	https://benchling.com/s/riO8LQFW/edit
J23106_EB	Basic Part - Promoter (Constitutive)	DVA	65993	CIDAR_503	https://benchling.com/s/ljtaz48p/edit
J23106_FB	Basic Part - Promoter (Constitutive)	DVA	65994	CIDAR_504	https://benchling.com/s/b8wazW5M/edit
J23106_GB	Basic Part - Promoter (Constitutive)	DVA	65995	CIDAR_505	https://benchling.com/s/UqDZkix7/edit
J23107_AB	Basic Part - Promoter (Constitutive)	DVA	65996	CIDAR_506	https://benchling.com/s/G8ouTfIM/edit
J23107_EB	Basic Part - Promoter (Constitutive)	DVA	65997	CIDAR_507	https://benchling.com/s/SGzIA9lZ/edit
J23107_FB	Basic Part - Promoter (Constitutive)	DVA	65998	CIDAR_508	https://benchling.com/s/0mSZV0LM/edit
J23107_GB	Basic Part - Promoter (Constitutive)	DVA	65999	CIDAR_509	https://benchling.com/s/InUihbMF/edit
J23116_AB	Basic Part - Promoter (Constitutive)	DVA	66000	CIDAR_510	https://benchling.com/s/1mQPZFtM/edit
J23116_EB	Basic Part - Promoter (Constitutive)	DVA	66001	CIDAR_511	https://benchling.com/s/xqAO8tJl/edit
J23116_FB	Basic Part - Promoter (Constitutive)	DVA	66002	CIDAR_512	https://benchling.com/s/o8lwrldJ/edit
J23116_GB	Basic Part - Promoter (Constitutive)	DVA	66003	CIDAR_513	https://benchling.com/s/YYNkGLBV/edit
R0010_AB	Basic Part - Promoter (Controllable)	DVA	66004	CIDAR_514	https://benchling.com/s/etg7OrN6/edit
R0010_EB	Basic Part - Promoter (Controllable)	DVA	66005	CIDAR_515	https://benchling.com/s/FamBps3h/edit
R0010_FB	Basic Part - Promoter (Controllable)	DVA	66006	CIDAR_516	https://benchling.com/s/oi1T00L0/edit
R0010_GB	Basic Part - Promoter (Controllable)	DVA	66007	CIDAR_517	https://benchling.com/s/Yw6bgPjg/edit
R0040_AB	Basic Part - Promoter (Controllable)	DVA	66008	CIDAR_518	https://benchling.com/s/3RYAjD6u/edit
R0040_EB	Basic Part - Promoter (Controllable)	DVA	66009	CIDAR_519	https://benchling.com/s/SNZHRImR/edit
R0040_FB	Basic Part - Promoter (Controllable)	DVA	66010	CIDAR_520	https://benchling.com/s/nyzVqekN/edit
R0040_GB	Basic Part - Promoter (Controllable)	DVA	66011	CIDAR_521	https://benchling.com/s/473M59P8/edit

Table A-7 CIDAR MoClo Library part list. *All basic parts, vectors, and fluorescent control plasmids are listed here with their Addgene and CIDAR-ICE ID numbers and Benchling location. The full collection can be requested from Addgene (#1000000059).*

Plasmid	Part Description	Backbone	Addgene Cat#	CIDAR-ICE ID#	Benchling (benchling.com)
R0063_AB	Basic Part - Promoter (Controllable)	DVA	66012	CIDAR_522	https://benchling.com/s/1EUFuZKHP/edit
R0063_EB	Basic Part - Promoter (Controllable)	DVA	66013	CIDAR_523	https://benchling.com/s/PN1VmcWg/edit
R0063_FB	Basic Part - Promoter (Controllable)	DVA	66014	CIDAR_524	https://benchling.com/s/xnA98Y5U/edit
R0063_GB	Basic Part - Promoter (Controllable)	DVA	66015	CIDAR_525	https://benchling.com/s/jHQHwAaD/edit
I13453_AB	Basic Part - Promoter (Controllable)	DVA	66016	CIDAR_526	https://benchling.com/s/HE4joK6k/edit
I13453_EB	Basic Part - Promoter (Controllable)	DVA	66017	CIDAR_527	https://benchling.com/s/YC8Ney.Ae/edit
I13453_FB	Basic Part - Promoter (Controllable)	DVA	66018	CIDAR_528	https://benchling.com/s/c4zFwYww/edit
I13453_GB	Basic Part - Promoter (Controllable)	DVA	66019	CIDAR_529	https://benchling.com/s/i9YCgIRi/edit
B0032m_BC	Basic Part - RBS (RBS)	DVA	66020	CIDAR_530	https://benchling.com/s/jwisLDg9/edit
B0033m_BC	Basic Part - RBS (RBS)	DVA	66021	CIDAR_531	https://benchling.com/s/FQFaG78R/edit
B0034m_BC	Basic Part - RBS (RBS)	DVA	66022	CIDAR_532	https://benchling.com/s/kBNYDAHo/edit
BCD12_BC	Basic Part - RBS (BiCistronic Design (BCD))	DVA	66023	CIDAR_533	https://benchling.com/s/mUcZUqJo/edit
BCD2_BC	Basic Part - RBS (BiCistronic Design (BCD))	DVA	66024	CIDAR_534	https://benchling.com/s/jM8sZdRg/edit
BCD8_BC	Basic Part - RBS (BiCistronic Design (BCD))	DVA	66025	CIDAR_535	https://benchling.com/s/8UnD5Vaf/edit
C0012m_CD	Basic Part - CDS (Transcription factor)	DVA	66026	CIDAR_536	https://benchling.com/s/1RopAYal/edit
C0040_CD	Basic Part - CDS (Transcription factor)	DVA	66027	CIDAR_537	https://benchling.com/s/M9uZpTOg/edit
C0062_CD	Basic Part - CDS (Transcription factor)	DVA	66028	CIDAR_538	https://benchling.com/s/FFyxWiwT/edit
C0080_CD	Basic Part - CDS (Transcription factor)	DVA	66029	CIDAR_539	https://benchling.com/s/FFyxWiwT/edit
cre_CD	Basic Part - CDS (Recombinase)	DVA	66030	CIDAR_540	https://benchling.com/s/CAXPcaMy/edit
E0030_CD	Basic Part - CDS (Fluorescent reporter - YFP)	DVA	66031	CIDAR_541	https://benchling.com/s/cr12BVOE/edit
E0040m_CD	Basic Part - CDS (Fluorescent reporter - GFP)	DVA	66032	CIDAR_542	https://benchling.com/s/cj91KrVH/edit
E1010m_CD	Basic Part - CDS (Fluorescent reporter - RFP)	DVA	66033	CIDAR_543	https://benchling.com/s/1AbCv00F/edit
eBFP2_CD	Basic Part - CDS (Fluorescent reporter - BFP)	DVA	66034	CIDAR_544	https://benchling.com/s/bbkcfFdr/edit
B0015_DE	Basic Part - Terminator (Double Terminator)	DVA	66035	CIDAR_545	https://benchling.com/s/a1cHND33/edit
B0015_DF	Basic Part - Terminator (Double Terminator)	DVA	66036	CIDAR_546	https://benchling.com/s/f6RaEEG2K/edit
B0015_DG	Basic Part - Terminator (Double Terminator)	DVA	66037	CIDAR_547	https://benchling.com/s/5pFYTkSF/edit
B0015_DH	Basic Part - Terminator (Double Terminator)	DVA	66038	CIDAR_548	https://benchling.com/s/8st8vlJQ/edit
DVA_AB	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66039	CIDAR_549	https://benchling.com/s/slHkWK6p/edit
DVA_AE	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66040	CIDAR_550	https://benchling.com/s/JbqtwMAv/edit
DVA_AF	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66041	CIDAR_551	https://benchling.com/s/NviiNLlp/edit
DVA_AG	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66042	CIDAR_552	https://benchling.com/s/pPe92CO5/edit
DVA_AH	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66043	CIDAR_553	https://benchling.com/s/iaYDmV2/edit

Plasmid	Part Description	Backbone	Addgene Cat#	CIDAR-ICE ID#	Benchling (benchling.com)
DVA_BC	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66044	CIDAR_554	https://benchling.com/s/o7TOrROp/edit
DVA_CD	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66045	CIDAR_555	https://benchling.com/s/xgdB8Znv/edit
DVA_DE	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66046	CIDAR_556	https://benchling.com/s/mtpMynxT/edit
DVA_DF	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66047	CIDAR_557	https://benchling.com/s/53Whxb2s/edit
DVA_DG	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66048	CIDAR_558	https://benchling.com/s/bJdQOwZP/edit
DVA_DH	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66049	CIDAR_559	https://benchling.com/s/sGGQDAg6/edit
DVA_EB	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66050	CIDAR_560	https://benchling.com/s/i5JEIhUg/edit
DVA_EF	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66051	CIDAR_561	https://benchling.com/s/OISBaZCj/edit
DVA_EG	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66052	CIDAR_562	https://benchling.com/s/xGeYtbZx/edit
DVA_EH	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66053	CIDAR_563	https://benchling.com/s/fnTqha2S/edit
DVA_FB	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66054	CIDAR_564	https://benchling.com/s/3CuYwFwU/edit
DVA_GB	Destination Vector (For Basic Parts and Devices Ampicillin)	DVA	66055	CIDAR_565	https://benchling.com/s/UvuVapbV/edit
DVA	Destination Vector (Backbone for new DVs)	DVA	66056	CIDAR_566	https://benchling.com/s/6hIiqfV7/edit
pJ02B2Gm_AE	Transcriptional Unit (FACS Controls)	DVA	66057	CIDAR_567	https://benchling.com/s/mgYDcL4z/edit
pJ02B2Gm_EF	Transcriptional Unit (FACS Controls)	DVA	66058	CIDAR_568	https://benchling.com/s/3WQbdIzU/edit
pJ02B2Rm_AE	Transcriptional Unit (FACS Controls)	DVA	66059	CIDAR_569	https://benchling.com/s/Kb4AAbox/edit
pJ02B2Rm_EF	Transcriptional Unit (FACS Controls)	DVA	66060	CIDAR_570	https://benchling.com/s/FOM8uxtS/edit
pJ02B2Gm:Rm_AF	Device (FACS Controls)	DVA	66061	CIDAR_571	https://benchling.com/s/D75zi v6K/edit
pJ02B2Rm:Gm_AF	Device (FACS Controls)	DVA	66062	CIDAR_572	https://benchling.com/s/OOkaVw1a/edit
pJ02B2Gm_AF	Transcriptional Unit (FACS Controls)	DVK	66063	CIDAR_573	https://benchling.com/s/ShVJxDdF/edit
pJ02B2Gm_EF	Transcriptional Unit (FACS Controls)	DVK	66064	CIDAR_574	https://benchling.com/s/uUTmefO7/edit
pJ02B2Rm_AE	Transcriptional Unit (FACS Controls)	DVK	66065	CIDAR_575	https://benchling.com/s/syzZcMB4/edit
pJ02B2Rm_EF	Transcriptional Unit (FACS Controls)	DVK	66066	CIDAR_576	https://benchling.com/s/nNICVGuk/edit
DVK_AE	Destination Vector (For Transcriptional Units Kanamycin)	DVK	66067	CIDAR_577	https://benchling.com/s/sueyphOw/edit
DVK_AF	Destination Vector (For Transcriptional Units Kanamycin)	DVK	66068	CIDAR_578	https://benchling.com/s/1i0mX41y/edit
DVK_EF	Destination Vector (For Transcriptional Units Kanamycin)	DVK	66069	CIDAR_579	https://benchling.com/s/7FwJO5tD/edit
DVK_FG	Destination Vector (For Transcriptional Units Kanamycin)	DVK	66070	CIDAR_580	https://benchling.com/s/Cz5hQ9wM/edit
DVK_GH	Destination Vector (For Transcriptional Units Kanamycin)	DVK	66071	CIDAR_581	https://benchling.com/s/MFw9pNDU/edit
DVK	Destination Vector (Backbone for new DVs)	DVK	66072	CIDAR_582	https://benchling.com/s/oIS42cDR/edit

A.5   Constitutive promoter and RBS part type characterization

A total of 6 constitutive promoters and 6 RBS parts were selected for inclusion in the CIDAR MoClo Library. As described previously, an array of GFP expression plasmids was constructed to evaluate the strength of the RBS and constitutive promoter parts. Using this data, J23100, J23102, J23103, J23106, J23107, and J23116 were selected for the CIDAR MoClo Library along with all 6 RBS parts tested (B0032m, B0033m, B0034m, BCD2, BCD8, and BCD12). A subset of the data presented in Figure 3-6 and Figure 3-7 is shown below for all combinations of these 6 promoters and 6 RBS parts. This chart is included to assist in selecting promoters and RBS parts in future designs.
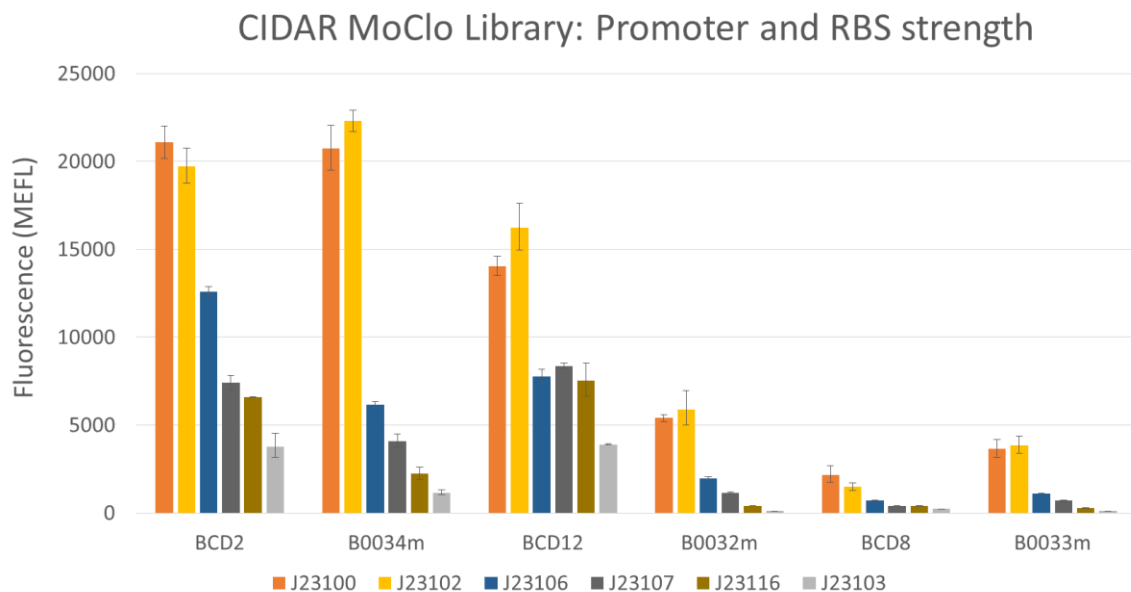
Figure A-2 CIDAR MoClo Library – Constitutive promoters and RBS parts, GFP expression plasmids. A subset of the data from Figure 3-6 and Figure 3-7, all combinations of constitutive promoter and RBS parts which were included in the CIDAR MoClo Library are shown here to aid in selection of parts for future designs.

A.6   Data tables for Chapter 3 figures

	GFP			RFP		
	Mean	**GSD+**	**GSD-**	**Mean**	**GSD+**	**GSD-**
pJ02B2X_AE(A)	82571	83011	82133	118142	121733	114658
pJ02B2X_AE(K)	29396	32065	26948	50902	53843	48121
pJ02B2X_GH(K)	25133	28327	22300	28993	30355	27691
pJ024mX_AE(K)	22299	22918	21698	31844	34567	29335
pJ02B12X_AE(K)	16242	17629	14965	33180	39642	27771
pJ06B2X_AE(K)	12595	12866	12329	33840	36962	30981
pJ07B12X_AE(K)	8362	8523	8205	17974	18831	17156
pJ07B2X_AE(K)	7393	7808	7000	8292	9331	7370
pJ02B2X_FG(K)	8080	8698	7505	5717	5966	5478
pJ06B12X_AE(K)	7756	8187	7348	11128	12295	10072
pJ064mX_AE(K)	6169	6318	6025	2062	2265	1878
pJ024mX_EF(K)	5928	6474	5428	5646	6440	4950
pJ02B12X_EF(K)	5924	5979	5869	11530	11602	11457
pJ022mX_AE(K)	5899	6954	5004	3064	3066	3061
pJ02B2X_EF(K)	5652	5906	5410	13451	14797	12228
pJ14B12X_AE(K)	4982	5037	4927	10721	14867	7731
pJ14B2X_AE(K)	4372	4404	4341	11930	15418	9231
pJ074mX_AE(K)	4089	4475	3737	1284	1530	1077
pJ02B2X_EF(A)	3953	4070	3840	3273	3363	3185
pJ144mX_AE(K)	2617	2755	2486	2993	3029	2957
pJ062mX_AE(K)	1981	2068	1897	1249	1371	1137
pJ072mX_AE(K)	1124	1188	1063	944	1008	883
pJ063mX_AE(K)	1119	1123	1115	534	662	431
pJ06B8X_AE(K)	716	740	692	3696	4538	3010
pJ073mX_AE(K)	698	742	656	364	404	329
pJ022mX_EF(K)	498	555	446	315	331	299
pJ07B8X_AE(K)	401	407	396	2806	2806	2806
pJ142mX_AE(K)	194	204	184	178	185	171

Table A-8 MEFL units - Geometric mean of fluorescence as displayed in Figure 3-3c. All samples were measured in triplicate and data is provided as geometric mean, high, and low.

	2CC GFP L1			2CC GFP L2		
	Mean	**GSD+**	**GSD-**	**Mean**	**GSD+**	**GSD-**
pJ02B2	3953	4070	3840	5682	5750	5615
pJ02B12	4063	4173	3957	2587	2858	2343
pJ024m	1500	1572	1431	3947	4193	3715
pJ022m	449	462	436	1423	1440	1406
pJ14B2	395	429	365	580	617	546
pJ14B12	465	488	443	343	383	308
pJ144m	208	212	204	158	159	156
pJ142m	124	128	121	141	145	136

	2CC RFP L1			2CC RFP L2		
	Mean	**GSD+**	**GSD-**	**Mean**	**GSD+**	**GSD-**
pJ02B2	118142	121733	114658	162270	173731	151564
pJ02B12	70614	70944	70285	34736	37543	32139
pJ024m	9618	9804	9436	6217	6540	5910
pJ022m	3721	3922	3530	7221	7654	6813
pJ14B2	3429	3595	3270	7422	7470	7375
pJ14B12	3303	3396	3213	3678	3751	3606
pJ144m	524	539	508	657	728	594
pJ142m	152	155	149	216	226	206

Table A-9 MEFL units - Geometric mean of fluorescence as displayed in Figure 3-5a. All samples were measured in triplicate and data is provided as geometric mean, high, and low.

	Kan			Amp		
	Mean	**GSD+**	**GSD-**	**Mean**	**GSD+**	**GSD-**
pJ022mRm_AE	3064	3066	3061	3721	3922	3530
pJ024mRm_AE	31844	34567	29335	9618	9804	9436
pJ02B12Rm_AE	33180	39642	27771	70614	70944	70285
pJ02B2Rm_AE	50902	53843	48121	118142	121733	114658
pJ142mRm_AE	178	185	171	152	155	149
pJ144mRm_AE	2993	3029	2957	524	539	508
pJ14B12Rm_AE	10721	14867	7731	3303	3396	3213
pJ14B2Rm_AE	11930	15418	9231	3429	3595	3270
pJ02B2Gm_AE	19733	20775	18743	82571	83011	82133
pJ02B2Gm_EF	5652	5906	5410	3953	4070	3840
pJ02B12Gm_EF	5924	5979	5869	4063	4173	3957
pJ024mGm_EF	5928	6474	5428	1500	1572	1431
pJ022mGm_EF	498	555	446	449	462	436
pJ14B2Gm_EF	2513	2831	2230	395	429	365
pJ14B12Gm_EF	2785	3263	2377	465	488	443
pJ144mGm_EF	3505	3920	3135	208	212	204
pJ142mGm_EF	133	134	131	124	128	121
pJ02B2Rm_EF	13170	13364	12979	3273	3363	3185

Table A-10 MEFL units - Geometric mean of fluorescence as displayed in Figure 3-5b. All samples were measured in triplicate and data is provided as geometric mean, high, and low.

	_AE			_EF		
	Mean	**GSD+**	**GSD-**	**Mean**	**GSD+**	**GSD-**
pJ022mRm(K)	3064	3066	3061	315	331	299
pJ024mRm(K)	31844	34567	29335	5646	6440	4950
pJ02B12Rm(K)	33180	39642	27771	11530	11602	11457
pJ02B2Rm(K)	50902	53843	48121	13451	14797	12228
pJ062mRm(K)	1249	1371	1137	97	102	93
pJ063mRm(K)	534	662	431	76	77	75
pJ064mRm(K)	2062	2265	1878	1732	2013	1489
pJ06B12Rm(K)	11128	12295	10072	2573	3066	2159
pJ06B2Rm(K)	33840	36962	30981	2703	3204	2280
pJ06B8Rm(K)	3696	4538	3010	398	415	381
pJ072mRm(K)	944	1008	883	93	100	86
pJ073mRm(K)	364	404	329	75	75	74
pJ074mRm(K)	1284	1530	1077	1177	1269	1091
pJ07B12Rm(K)	17974	18831	17156	2522	2717	2341
pJ07B2Rm(K)	8292	9331	7370	2797	3403	2300
pJ07B8Rm(K)	2806	2806	2806	364	377	352
pJ02B2Gm(K)	19733	20775	18743	5652	5906	5410
pJ02B12Gm(K)	16242	17629	14965	5924	5979	5869
pJ024mGm(K)	22299	22918	21698	5928	6474	5428
pJ022mGm(K)	5899	6954	5004	498	555	446
pJ14B2Gm(K)	4372	4404	4341	2513	2831	2230
pJ14B12Gm(K)	4982	5037	4927	2785	3263	2377
pJ144mGm(K)	2617	2755	2486	3505	3920	3135
pJ142mGm(K)	194	204	184	133	134	131

*Table A-11 MEFL units - Geometric mean of fluorescence as displayed in Figure 3-5c. All samples were measured in triplicate and data is provided as geometric mean, high, and low.*

Appendix B Additional experimental information

B.1 Chapter 4 – Sequence structure of aldolase fusion proteins.

pJ06B12HRAGFP_AE(1G) (3248 bp)
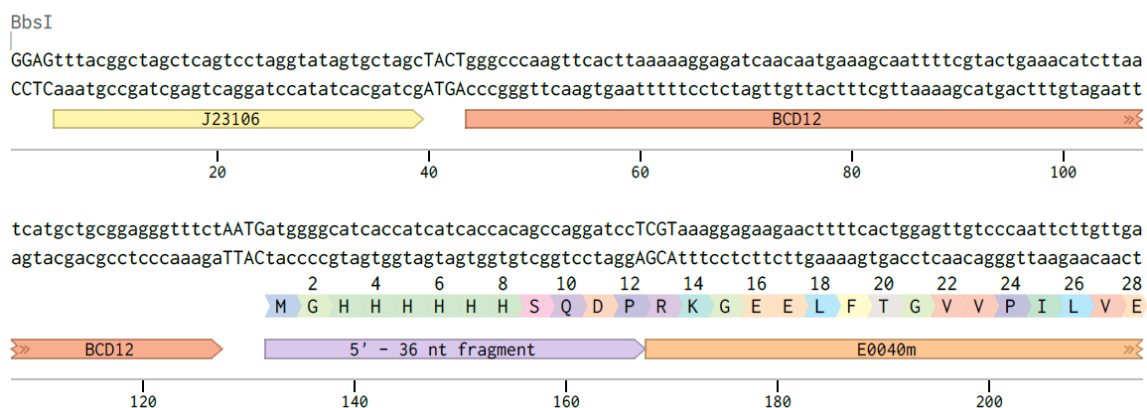
pJ06B12HRARFP_AE(1R) (3209 bp)
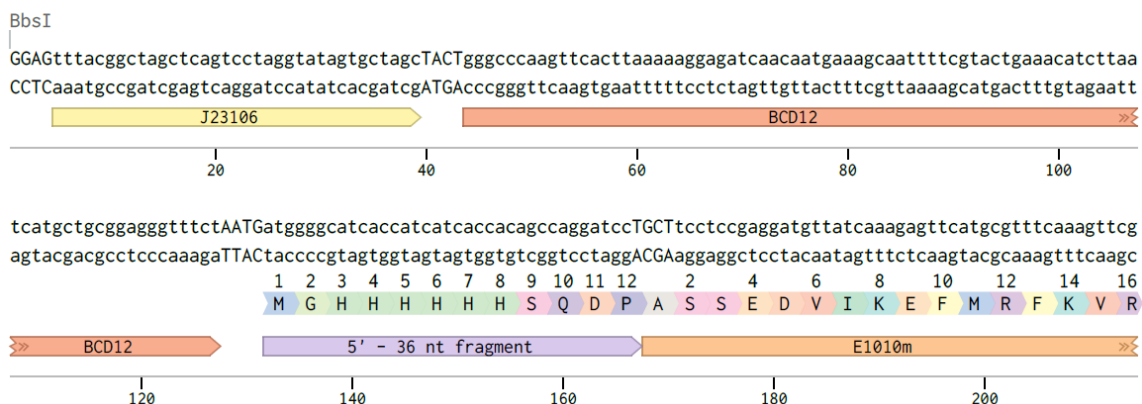
Figure B-1 Sequence annotation of 5' HRA fusion protein sequence. *Full sequence analysis available on Benchling.*

pJ06B12R42AGFP_AE(6G) (3248 bp)

BbsI

GGAGtttacggctagctcagtcctaggtatagtgctagcTACTgggcccaagttcacttaaaaaggagatcaacaatgaaagcaattttcgtactgaaacatcttaa
CCTCaaatgccgatcgagtcaggatccatatcacgatcgATGAcccgggttcaagtgaattttttcctctagttgttactttcgttaaaagcatgactttgtagaatt

J23106   BCD12

20   40   60   80   100

tcatgctgcggagggtttctAATGatgcctcactcccatccagcgttaacccggagcAGCGtaaaggagaagaacttttcactggagttgtcccaattcttgttga
agtacgacgcctcccaaagaTTACtacggagtgagggtaggtcgcaattggggcctcgTCGCatttcctcttcttgaaaagtgacctcaacagggttaagaacaact

2   4   6   8   10   12   14   16   18   20   22   24   26   28
M P H S H P A L T P E Q R K G E E L F T G V V P I L V E

BCD12   5' – 36 nt fragment   E0040m

120   140   160   180   200

pJ06B12R42ARFP_AE(6R) (3209 bp)

BbsI

GGAGtttacggctagctcagtcctaggtatagtgctagcTACTgggcccaagttcacttaaaaaggagatcaacaatgaaagcaattttcgtactgaaacatcttaa
CCTCaaatgccgatcgagtcaggatccatatcacgatcgATGAcccgggttcaagtgaattttttcctctagttgttactttcgttaaaagcatgactttgtagaatt

J23106   BCD12

20   40   60   80   100

tcatgctgcggagggtttctAATGatgcctcactcccatccagcgttaacccggagcaGGCTtcctccgaggatgttatcaaagagttcatgcgtttcaaagttcg
agtacgacgcctcccaaagaTTACtacggagtgagggtaggtcgcaattggggcctcgtCCGAaggaggctcctacaatagtttctcaagtacgcaaagtttcaagc

2   4   6   8   10   12   14   16   18   20   22   24   26   28
M P H S H P A L T P E Q A S S E D V I K E F M R F K V R

BCD12   5' – 36 nt fragment   E1010m

120   140   160   180   200

Figure B-2 Sequence annotation of 5' R42A fusion protein sequence. *Full sequence analysis available on Benchling.*
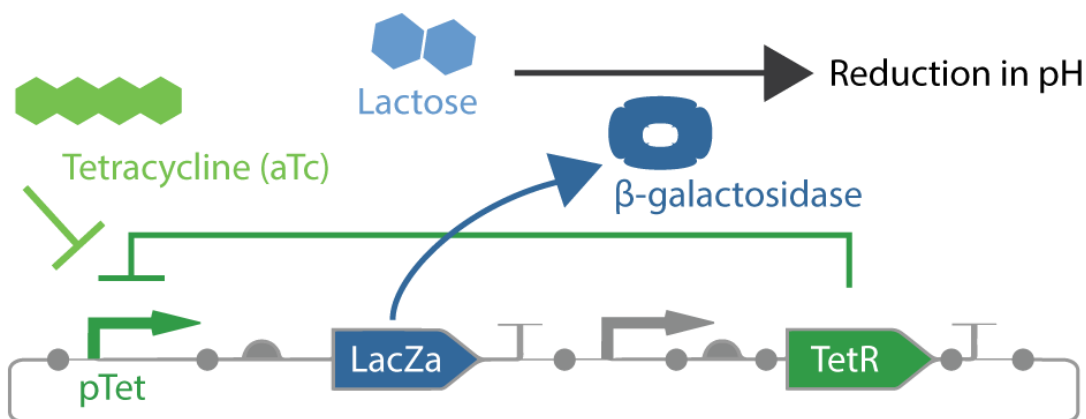
Figure B-3 pTetLacZDV:J024mC40_AF(A) circuit. *This circuit was evaluated for use in the pH biosensor experiments but failed to provide a distinctive pH change upon induction with anhydrous tetracycline (aTc).*
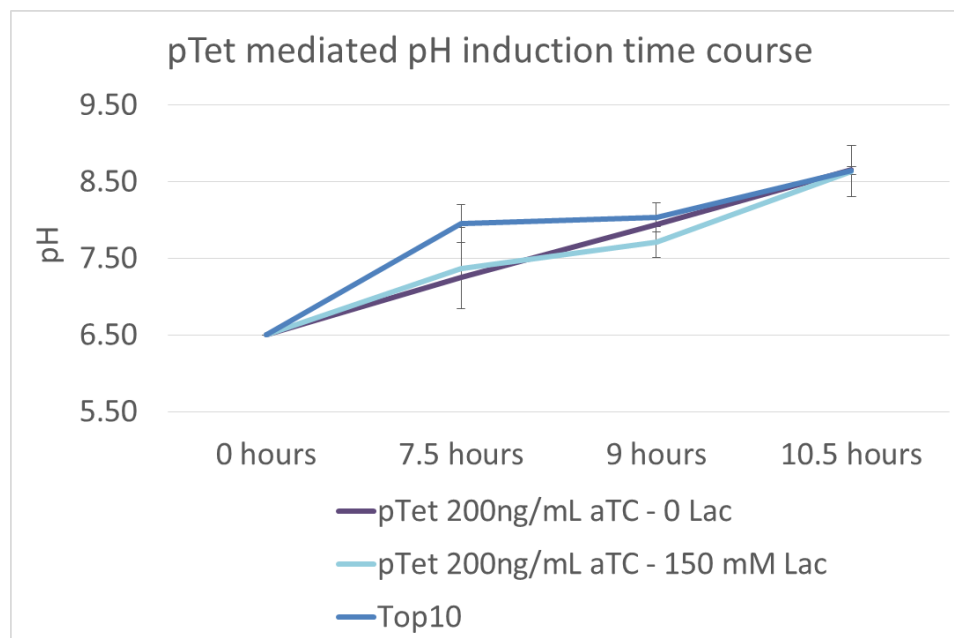
Figure B-4 pTet mediated pH induction time course. *pTet mediated pH induction was less functional as a biosensor than the other circuits tested. Here cells were grown in 3 mL cultures and pH was measured at the defined intervals. Earlier time points may offer more functionality.*

***Appendix C* Education**

CIDAR Educational Labs are found in supplementary files attached to this dissertation along with calculation files and an accompanying Powerpoint presentation.

Also hosted on www.cidarlab.org.

***Appendix D* External files**

Also hosted on www.cidarlab.org.

D.1   MoClo assembly calculation files in Excel linked to part list.

CIDARAssemblyCalculations.xlsx

D.2   Raven file containing part list

raven_CIDARMoCloLibrary.csv

D.3   Eugene file

CIDAR_MoClo_Eugene_final.eug

***Appendix E* Development of modular biosensors for allergen detection**

This appendix describes the background and experimental design for a bioelectronic whole cell gluten biosensor designed in a modular fashion using MoClo and multiplex MoClo protocols. This project was discontinued. Preliminary progress is documented here to provide a resource for further study.

E.1   Two component systems for biosensor development

Recent work using domain exchange methods have utilized histidine kinase two-component systems and focused on the Per-Arnt-Sim (PAS) protein domain

family (Gu et al. 2000). PAS sensor domains are ubiquitous across all forms of life where they are involved in regulating gene expression in response to various inputs including light and oxygen. While the PAS sensor domain of a histidine kinase adopts a common globular structure, the effector domains of these proteins are often highly varied, functioning as kinases, transcription factors, and phosphodiesterases. This allows for a standardized yet versatile signal recognition motif with an expansive repertoire of response capabilities. This apparent modularity suggests the ability to interchange PAS domains to rewrite the natural biosensing pathways towards engineering novel biosensors and create unique gene regulatory systems.

Multiple groups have shown this to be true both in vitro and in vivo primarily with LOV domains (Tar:EnvZ (Utsumi et al. 1989) YtlV:FixL (Moglich et al. 2009) Cph1:EnvZ (Levskaya et al. 2005, Tabor et al. 2009)). Named for their stimuli, light-oxygen-voltage (LOV) domains bind flavin nucleotides, inducing an auto-phosphorylation of the cytoplasmic kinase domain and subsequent transfer of this phosphate to a downstream response regulator (Gu et al. 2000).

Previous research has shown successful domain exchange in histidine kinases, primarily in the development of optogenetic tools (Levskaya et al. 2005, Moglich et al. 2009, Tabor et al. 2009, Moglich and Moffat 2010) and provided insight towards further rearrangements, particularly with respect to the characteristics of the required Jα linker domain (Jin and Inouye 1994, Moglich et

al. 2009). This linker shows little homology among related TCS yet conserves a standard 7(n)-mer or 7(n)+2 pattern. This heptad periodicity is characteristic of α-helical coiled coil and is required for correct function seemingly by controlling the orientation of the sensor and kinase domains as indicated by the lack of function in YtvA:FixL chimeric proteins which disrupted this periodicity (Moglich et al. 2009). Two papers specifically lay out design constraints for the development of future two-component chimeric proteins (Yoshida et al. 2007, Moglich et al. 2009).

Skerker et al. have also shown the ability to rationally rewire two-component systems by mutating specific residues which dictate the interaction of the histidine kinase with its cognate response regulator predicted by analyzing the patterns of coevolution in two-component systems. By mutating as few as three amino acids of the EnvZ kinase domain, T250V, L254Y and A255R, the authors were able to show specific phosphorylation of the non-cognate response regulator RstA and amelioration of the natural phosphorylation of OmpR (Skerker et al. 2008). The independence of protein domains in histidine kinases suggests the possibility of a modular recombination method in rewiring the specificity of known two-component systems by domain exchange with Multiplex Modular Cloning. Additionally, the directed point mutation based rewiring of EnvZ demonstrates the potential for rational modification of phosphorylation targets.

E.2   Design of a Whole-Cell Rapid Bio-electronic Gluten Detector

Glutamine binding protein (GlnBP) is a 26 kDa periplasmic E. coli protein which, in a monomeric form, is responsible for initializing the active transport of L-glutamine into the cell. GlnBP has been shown previously to selectively bind gliadin, the alcohol soluble component of gluten. This induced the authors to design a bioelectronics sensor using purified GlnBP and a nanostructured porous silicone (PSi) waver with the mild success of detecting gluten at levels as low as 2,000 ppm. Regardless of the inadequate sensing capabilities, the cost of purifying GlnBP and production of silicone wavers would have been prohibitive of a commercial detection kit. This research, however, provides a starting point for a synthetic biology approach to developing a gluten biosensor.

By taking advantage of the natural selective binding of GlnBP to gliadin, the simultaneous iterative design methods developed above, and previous research upon the modularity of two-component systems, I propose to develop a rapid, accurate, sensitive and affordable whole-cell gluten biosensor.

E.2.1 Development of a "Slow" Gluten Sensor

The logical first step is the development of a transcription-based gluten-sensing strain of E. coli through fusion of GlnBP to the transmembrane and cytoplasmic domains of EnvZ. All of the following experiments will performed in the EnvZΔ strain(Levskaya et al. 2005) of E. coli reported by Levskaya et al. (RU1012 [MC4100 ara+ φ(OmpC-lacZ) 10-25 ΔenvZ::KanR]) or a derivative strain.

E.2.1.1   Experimental Design: GlnBP:EnvZ fusion

Structural analysis and sequence homology of each family of proteins will be used to inform on the most probable position of splicing between these two proteins. GlnBP is 224 residues in length and contains two globular domains. The cleft formed between the larger domain (a.a. 1-84 and 186-224) and the smaller domain (a.a. 90-180) contains the ligand binding site. As such, it is likely the complete GlnBP coding sequence will be required for a successful chimeric protein.
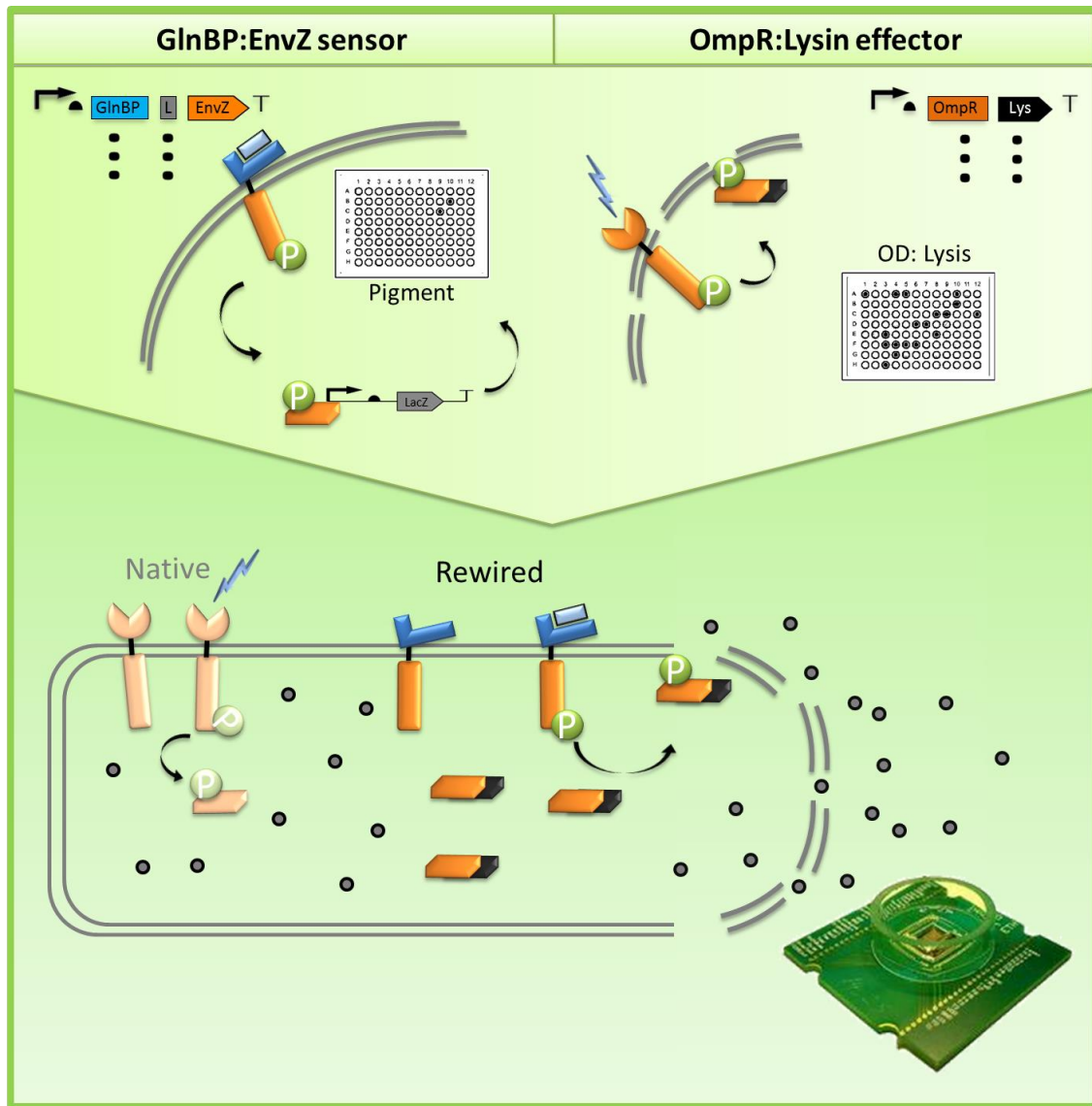
Figure E-1 Whole cell bioelectronic gluten biosensor design. *Gluten sensing and phospho-activated cell lysis systems will be engineered separately using LacZα catalyzed pigmentation and OD as measures of gluten binding and lysis respectively. A selectable marker may also be used in the screening of GlnBP:EnvZ chimeras. Once calibrated, the two systems will be tuned and optimized for use with a bioelectronics interface to provide a digital report.*

The Cph1:EnvZ fusion created by Levskaya et al. used the same C-terminal 229 a.a. fragment of EnvZ as was previously used in one of the first experiments

to demonstrate the modularity of two-component systems by rewiring the aspartate sensing Tar and oxygen responsive EnvZ histidine kinases in a similar manner. Initial attempts at developing this GlnBP:EnvZ fusion will use this same fragment as well as a degenerate linker library based upon the Cph1:EnvZ, Tar:EnvZ and YtlV:FixL chimeric protein sequences and extensive analysis histidine kinase linker domains included in these publications (Utsumi et al. 1989, Zhu and Inouye 2003, Levskaya et al. 2005, Moglich et al. 2009).

In order to efficiently screen this library of chimeric proteins, the previously established reporter system will be used in which activation of OmpR by the cytoplasmic domain of EnvZ leads to the expression of LacZα and the pigmentation. In order for this screening method to function, a modified destination vector which does not use LacZα expression as a screenable marker will be required. Affinity of this chimeric protein for gliadin as a whole and the toxic peptide specifically will be determined.

Potential troubleshooting:

Other species of bacteria may be more appropriate as a biosensor especially in later stages of this development. However, as the MoClo library and protocols are currently oriented towards work in E. coli, initial attempts will be made in this species with a change of chassis as one of the first avenues of adjustment should the need arise. While EnvZ has historically been successful in rewiring

132

endeavors, another histidine kinase may be required for successful development of a GlnBP:Kinase chimera.

E.2.2 Post-translational Control of Lysis

The response regulator portion of the gluten sensor can be developed separately alongside the development of the GlnBP:EnvZ chimera. As phosphorylation is a common and rapid mechanism of controlling enzyme function, engineering phosphorylation (Johnson and Lewis 2001) dependent allosteric control of a lytic enzyme may be a viable method of rapidly inducing lysis as a response to gluten. Previously successful modular rearrangement of the Yersina and T4 phage lysin proteins, effectively rewired specificity (Lukacik et al. 2012) and lends support to the engineering of an allosterically controlled T4 lysin chimeric protein.

Once again using a chimeric protein library, composed of the EnvZ cognate response regulator OmpR and a bacteriophage lysin protein, I propose to develop a phosphosensitive lysin protein which will induce cell lysis upon phosphorylation by GlnBP:EnvZ providing a rapid and amplified signal in response to gluten.

E.2.2.1 Experimental Design: OmpR:Lysin Fusion

Sequence, structure, host specificity and pathogenicity will factor into the selection of a subset of lytic enzymes for use in this design, beginning with the T4 phage lysin. Informed by the rules and constraints generated in Aim 2 and analysis of both structure and sequence conservation in OmpR and the chosen lysin(s),

mixed clones of each will be created in a CXD format as described in 2.2.2. Due to the use of a lytic enzyme, tight inducible control of expression will be required to allow for growth of cells containing this chimeric gene sequence. Additionally, as most lysin proteins require the assistance of a second enzyme, a holin, (Young 1992) to escape the cell membrane, expression of this enzyme could also be used as a means of control. Cultures will be grown to early log phase, induced, and monitored for changed in OD measured by plate reader to identify continued growth versus lysis. As a positive control for lysis, the chosen lysin will be screened as well. Ideally, an inactive form of the chosen lysin will also be available.

Potential troubleshooting:

Depending upon sequence analysis and the success of the initial round of chimeric proteins, different fragments of each protein may be identified and attempted. Additionally, internal insertion either protein may provide more successful fusions as described in 2.2.2. Preventing non-specific cell death will be difficult. Requiring multiple levels of control and may be prohibitive to this design strategy. Other means of triggering lysis may need to be considered as well as other forms of signal output. Some success has been seen in bioelectronics interfaces sensing luciferase (unpublished, Cambridge iGEM 2012) or using the metal reducing *S. oneidensis* to create a measurable charge (unpublished, Edinburgh iGEM 2012), though with low sensitivity.

E.2.3 Phospho-activated Lysin as the Response Regulator of GlnBP:EnvZ

If the above experiments are successful, the two cassettes (or collection of iterations of each) would be combined in the same cell and screened for function. Once completed, lysis becomes an amplified signal for the presence of gluten and can be integrated with a bioelectronics interface to provide a digital readout. Such bioelectronics interfaces have been shown in the form of proteins fixed to CMOS sensors which can detect changes in voltage or pH due to biological stimuli (Ho et al. 2007, Manickam et al. 2010, Welch and Christen 2012). Adapting this system to detect cell lysis and the accompanied change in voltage would likely be a collaborative effort and conversations to this end have been initiated.

E.2.3.1 Experimental Design: Gluten induced lysis

As each half of this sense-response system will be engineered in a multiplex fashion and screened for function individually, it is possible multiple potential iterations of each will be available for further screening in a combined system. Assembly at this point is determined by the outcome of the previous experiments, but will be performed using MoClo in the most efficient manner possible. Adjustments for copy number and adjustments to expression levels may need to be performed and it is likely a gnomically encoded system will be required. Once a laboratory version of the gluten detection whole-cell biosensor is developed the assay will be simplified for use by consumers, ideally through a simple bioelectronic integration on par with a standard glucose blood sugar detection system.

Potential troubleshooting:

While experiments will begin with OmpR and the T4 lysin protein, different molecules may function better in developing ligand-triggered lysis. Success with any of the previously proposed chimeric proteins will shed light upon design constraints and inform more successful designs, particularly in the case of the highly conserved modular histidine kinases. Even if both systems are working independently, the larger device may not function as a whole requiring reiterations upon the design. Once again, any reiterations will be informed by previous attempts and successes.

E.3   Current Status

Components of the EnvZ synthetic genetic light detection system were acquired from Addgene (pCph8 plasmid, #22869, chloramphenicol resistant) and from Dr. J. Clark Lagarias at the University of California – Davis (pPL-PCB plasmid, spectinomycin resistant). pPL-PCB contains pcyA and ho1 coding sequences under the control of a single promoter (Levskaya et al. 2005). The CP919 strain, an MC4100 *E. coli* derivative, carries a genomic pOmpC promoter driving expression of LacZα. CP919 functions as a reporter plasmid for pCph8 (acquired from Voigt Lab) and is resistant to both kanamycin and spectinomycin.

The coding sequences for Cph8, pcyA and ho1 have been converted to MoClo CDS parts, designated as Cph8_CD, pcyA_CD and ho1_CD respectively. Transcription units containing these coding sequences were created and

136

designated pMAD1 and pMAD2 (ho1), pMAD3 and pMAD4 (pcyA), pMAD5 and pMAD6 (Cph8). Co-expression plasmids using pMAD1-4 were assembled and designated pMAD7-8 in _AF vectors. These are archived as mini preps and glycerol stocks alongside a glycerol stock of CP919.

Exploration into the bioelectronic aspects of this sensor design led to the experiments documented in Chapter 5. In light of those results, while pH can function as an electronically integrated signal, there appears to be many contributing factors to adjusting pH of biological cultures. This noise may necessitate more manageable signal integration approaches.

References

Agmon, N., L. A. Mitchell, Y. Cai, S. Ikushima, J. Chuang, A. Zheng, W. J. Choi, J. A. Martin, K. Caravelli, G. Stracquadanio and J. D. Boeke (2015). "Yeast Golden Gate (Ygg) for Efficient Assembly of S. Cerevisiae Transcription Units." <u>ACS Synthetic Biology</u>.

Anderson, J. C. (2010). "Bglbricks: A Flexible Standard for Biological Part Assembly." <u>Journal of Biological Engineering</u> **4**: 1.

Annaluru, N. (2012). "Assembling DNA Fragments by User Fusion." <u>Methods Molecular Biology</u> **852**: 77-95.

Appleton, E., J. Tao, T. Haddock and D. Densmore (2014). "Interactive Assembly Algorithms for Molecular Cloning." <u>Nature Methods</u> **11**(6): 657-662.

Beal, J., R. Weiss, D. Densmore, A. Adler, E. Appleton, J. Babb, S. Bhatia, N. Davidsohn, T. Haddock and J. Loyall (2012). "An End-to-End Workflow for Engineering of Biological Networks from High-Level Specifications." <u>ACS Synthetic Biology</u> **1**(8): 317-331.

Beernink, P. T. and D. R. Tolan (1994). "Subunit Interface Mutants of Rabbit Muscle Aldolase Form Active Dimers." <u>Protein Science</u> **3**(9): 1383-1391.

Beernink, P. T. and D. R. Tolan (1996). "Disruption of the Aldolase a Tetramer into Catalytically Active Monomers." <u>Proceedings of the National Academy of Sciences of the United States of America</u> **93**(11): 5374-5379.

Bilitchenko, L., A. Liu, S. Cheung, E. Weeding, B. Xia, M. Leguia, J. C. Anderson and D. Densmore (2011). "Eugene--a Domain Specific

138

Language for Specifying and Constraining Synthetic Biological Parts, Devices, and Systems." <u>PLoS One</u> **6**(4): e18882.

Bitinaite, J. (2007). "User Friendly DNA Engineering and Cloning Method by Uracil Excision." <u>Nucleic Acids Research</u> **35**: 1992-2002.

Bonnet, J., P. Subsoontorn and D. Endy (2012). "Rewritable Digital Data Storage in Live Cells Via Engineered Control of Recombination Directionality." <u>Proceedings of the National Academy of Sciences of the United States of America</u>.

Cameron, D. E., C. J. Bashor and J. J. Collins (2014). "A Brief History of Synthetic Biology." <u>Nature Reviews Microbiology</u> **12**(5): 381-390.

Casini, A., M. Storch, G. S. Baldwin and T. Ellis (2015). "Bricks and Blueprints: Methods and Standards for DNA Assembly." <u>Nature Reviews Molecular Cell Biology</u> **16**(9): 568-576.

Chee, J. Y. and C. F. Chin (2015). "Gateway Cloning Technology: Advantages and Drawbacks." <u>Cloning & Transgenesis</u> **2015**.

Church, G. M., M. B. Elowitz, C. D. Smolke, C. A. Voigt and R. Weiss (2014). "Realizing the Potential of Synthetic Biology." <u>Nature Reviews Molecular Cell Biology</u> **15**(4): 289-294.

Cohen, S. N., A. C. Chang, H. W. Boyer and R. B. Helling (1973). "Construction of Biologically Functional Bacterial Plasmids in Vitro." <u>Proceedings of the National Academy of Sciences of the United States of America</u> **70**: 3240-3244.

Czar, M. J., J. C. Anderson, J. S. Bader and J. Peccoud (2009). "Gene Synthesis Demystified." <u>Trends Biotechnology</u> **27**: 63-72.

de Mora, K., N. Joshi, B. L. Balint, F. B. Ward, A. Elfick and C. E. French (2011). "A Ph-Based Biosensor for Detection of Arsenic in Drinking Water." <u>Analytical and Bioanalytical Chemistry</u> **400**(4): 1031-1039.

Duportet, X., L. Wroblewska, P. Guye, Y. Li, J. Eyquem, J. Rieders, T. Rimchala, G. Batt and R. Weiss (2014). "A Platform for Rapid Prototyping of Synthetic Gene Networks in Mammalian Cells." <u>Nucleic Acids Research</u> **42**(21): 13440-13451.

Engler, C., R. Gruetzner, R. Kandzia and S. Marillonnet (2009). "Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type Iis Restriction Enzymes." <u>PLoS One</u> **4**(5): e5553.

Engler, C., M. Youles, R. Gruetzner, T. M. Ehnert, S. Werner, J. D. Jones, N. J. Patron and S. Marillonnet (2014). "A Golden Gate Modular Cloning Toolbox for Plants." <u>ACS Synthetic Biology</u>.

Friedland, A. E., T. K. Lu, X. Wang, D. Shi, G. Church and J. J. Collins (2009). "Synthetic Gene Networks That Count." <u>Science</u> **324**(5931): 1199-1202.

Gelperin, D. M., M. A. White, M. L. Wilkinson, Y. Kon, L. A. Kung, K. J. Wise, N. Lopez-Hoyo, L. Jiang, S. Piccirillo, H. Yu, M. Gerstein, M. E. Dumont, E. M. Phizicky, M. Snyder and E. J. Grayhack (2005). "Biochemical and Genetic Analysis of the Yeast Proteome with a Movable Orf Collection." <u>Genes & Development</u> **19**(23): 2816-2826.

Gibson, D. G., L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, 3rd and H. O. Smith (2009). "Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases." <u>Nature Methods</u> **6**(5): 343-345.

Giuraniuc, C. V., M. MacPherson and Y. Saka (2013). "Gateway Vectors for Efficient Artificial Gene Assembly in Vitro and Expression in Yeast Saccharomyces Cerevisiae." <u>PLoS One</u> **8**(5): e64419.

Goodman, D. B., G. M. Church and S. Kosuri (2013). "Causes and Effects of N-Terminal Codon Bias in Bacterial Genes." <u>Science</u> **342**(6157): 475-479.

Gu, Y. Z., J. B. Hogenesch and C. A. Bradfield (2000). "The Pas Superfamily: Sensors of Environmental and Developmental Signals." <u>Annual Review Pharmacology and Toxicology</u> **40**: 519-561.

Guckenberger, D. J., T. E. de Groot, A. M. Wan, D. J. Beebe and E. W. Young (2015). "Micromilling: A Method for Ultra-Rapid Prototyping of Plastic Microfluidic Devices." <u>Lab Chip</u> **15**(11): 2364-2378.

Gutierrez-Arenas, O. (2015). "Handling and Analyzing Meshed Rendering of Segmented Structures from 3d Image Stacks in Blender." <u>Neuroinformatics</u> **13**(2): 151-152.

Haddock, T. L., D. M. Densmore, E. Appleton, S. Carr, S. Iverson, M. De Freitas, S. Jin, J. Awtry, D. Desai and T. Lozanoski (2015). "Bbf Rfc 94: Type Iis Assembly for Bacterial Transcriptional Units: A Standardized Assembly Method for Building Bacterial Transcriptional Units Using the Type Iis Restriction Enzymes Bsai and Bbsi." <u>BioBricks Foundation Request For Proposals</u>.

Ham, T. S., Z. Dmytriv, H. Plahar, J. Chen, N. J. Hillson and J. D. Keasling (2012). "Design, Implementation and Practice of Jbei-Ice: An Open Source Biological Part Registry Platform and Tools." <u>Nucleic Acids Research</u> **40**(18): e141.

Ham, T. S., S. K. Lee, J. D. Keasling and A. P. Arkin (2008). "Design and Construction of a Double Inversion Recombination Switch for Heritable Sequential Genetic Memory." PLoS One **3**(7): e2815.

Hillson, N. J., R. D. Rosengarten and J. D. Keasling (2012). "J5 DNA Assembly Design Automation Software." ACS Synthetic Biology **1**: 14-21.

Ho, W. J., J. S. Chen, M. D. Ker, T. K. Wu, C. Y. Wu, Y. S. Yang, Y. K. Li and C. J. Yuan (2007). "Fabrication of a Miniature Cmos-Based Optical Biosensor." Biosensors and Bioelectronics **22**(12): 3008-3013.

Horecker, B., J. MacGregor, V. Singh, E. Melloni and S. Pontremoli (1980). "Aldolase and Fructose Bisphosphatase: Key Enzymes in the Control of Gluconeogenesis and Glycolysis." Current topics in cellular regulation **18**: 181-197.

Horton, R. M., H. D. Hunt, S. N. Ho, J. K. Pullen and L. R. Pease (1989). "Engineering Hybrid Genes without the Use of Restriction Enzymes: Gene Splicing by Overlap Extension." Gene **77**: 61-68.

Iverson, S., T. L. Haddock, J. Beal and D. Densmore (2015). "Cidar Moclo: Improved Moclo Assembly Standard and New E. Coli Part Library Enables Rapid Combinatorial Design for Synthetic and Traditional Biology." ACS Synthetic Biology.

Jewett, T. J. and L. D. Sibley (2003). "Aldolase Forms a Bridge between Cell Surface Adhesins and the Actin Cytoskeleton in Apicomplexan Parasites." Molecular Cell **11**(4): 885-894.

Jin, T. and M. Inouye (1994). "Transmembrane Signaling. Mutational Analysis of the Cytoplasmic Linker Region of Taz1-1, a Tar-Envz

Chimeric Receptor in Escherichia Coli." <u>Journal of Molecular Biology</u> **244**(5): 477-481.

Johnson, L. N. and R. J. Lewis (2001). "Structural Basis for Control by Phosphorylation." <u>Chem Rev</u> **101**(8): 2209-2242.

Kahl, L. J. and D. Endy (2013). "A Survey of Enabling Technologies in Synthetic Biology." <u>Journal of Biological Engineering</u> **7**(1): 13.

Katzen, F. (2007). "Gateway((R)) Recombinational Cloning: A Biological Operating System." <u>Expert Opin Drug Discov</u> **2**(4): 571-589.

Keasling, J. D. (2012). "Synthetic Biology and the Development of Tools for Metabolic Engineering." <u>Metabolic Engineering</u> **14**(3): 189-195.

Kishi, H., T. Mukai, A. Hirono, H. Fujii, S. Miwa and K. Hori (1987). "Human Aldolase a Deficiency Associated with a Hemolytic Anemia: Thermolabile Aldolase Due to a Single Base Mutation." <u>Proceedings of the National Academy of Sciences of the United States of America</u> **84**(23): 8623-8627.

Knight, T. (2007). "Draft Standard for Biobrick Biological Parts." <u>BioBricks Foundation Request For Proposals</u>.

Kosuri, S. and G. M. Church (2014). "Large-Scale De Novo DNA Synthesis: Technologies and Applications." <u>Nature Methods</u> **11**(5): 499-507.

Lee, J. W., D. Na, J. M. Park, J. Lee, S. Choi and S. Y. Lee (2012). "Systems Metabolic Engineering of Microorganisms for Natural and Non-Natural Chemicals." <u>Natural Chemical Biology</u> **8**(6): 536-546.

Lee, M. E., W. C. DeLoache, B. Cervantes and J. E. Dueber (2015). "A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly." ACS Synthetic Biology.

Levskaya, A., A. A. Chevalier, J. J. Tabor, Z. B. Simpson, L. A. Lavery, M. Levy, E. A. Davidson, A. Scouras, A. D. Ellington, E. M. Marcotte and C. A. Voigt (2005). "Synthetic Biology: Engineering Escherichia Coli to See Light." Nature 438(7067): 441-442.

Li, M. V., D. Shukla, B. H. Rhodes, A. Lall, J. Shu, B. S. Moriarity and D. A. Largaespada (2014). "Homerun Vector Assembly System: A Flexible and Standardized Cloning System for Assembly of Multi-Modular DNA Constructs." PLoS One 9(6): e100948.

Liu, J. K., W. H. Chen, S. X. Ren, G. P. Zhao and J. Wang (2014). "Ibrick: A New Standard for Iterative Assembly of Biological Parts with Homing Endonucleases." PLoS One 9: e110852.

Lohmueller, J. J., T. Z. Armel and P. A. Silver (2012). "A Tunable Zinc Finger-Based Framework for Boolean Logic Computation in Mammalian Cells." Nucleic Acids Research 40(11): 5180-5187.

Lukacik, P., T. J. Barnard, P. W. Keller, K. S. Chaturvedi, N. Seddiki, J. W. Fairman, N. Noinaj, T. L. Kirby, J. P. Henderson, A. C. Steven, B. J. Hinnebusch and S. K. Buchanan (2012). "Structural Engineering of a Phage Lysin That Targets Gram-Negative Pathogens." Proceedings of the National Academy of Sciences 109(25): 9857-9862.

Manickam, A., A. Chevalier, M. McDermott, A. D. Ellington and A. Hassibi (2010). "A Cmos Electrochemical Impedance Spectroscopy (Eis)

Biosensor Array." <u>IEEE Transactions Biomedical Circuits Systems</u> **4**(6): 379-390.

Moglich, A., R. A. Ayers and K. Moffat (2009). "Design and Signaling Mechanism of Light-Regulated Histidine Kinases." <u>Journal of Molecular Biology</u> **385**(5): 1433-1444.

Moglich, A. and K. Moffat (2010). "Engineered Photoreceptors as Novel Optogenetic Tools." <u>Photochemical & Photobiological Sciences</u> **9**(10): 1286-1300.

Mutalik, V. K., J. C. Guimaraes, G. Cambray, C. Lam, M. J. Christoffersen, Q. A. Mai, A. B. Tran, M. Paull, J. D. Keasling, A. P. Arkin and D. Endy (2013a). "Precise and Reliable Gene Expression Via Standard Transcription and Translation Initiation Elements." <u>Nature Methods</u> **10**(4): 354-360.

Mutalik, V. K., J. C. Guimaraes, G. Cambray, Q. A. Mai, M. J. Christoffersen, L. Martin, A. Yu, C. Lam, C. Rodriguez, G. Bennett, J. D. Keasling, D. Endy and A. P. Arkin (2013b). "Quantitative Estimation of Activity and Quality for Collections of Functional Genetic Elements." <u>Nature Methods</u> **10**(4): 347-353.

Nakajima, D., K. Saito, H. Yamakawa, R. F. Kikuno, M. Nakayama, R. Ohara, N. Okazaki, H. Koga, T. Nagase and O. Ohara (2005). "Preparation of a Set of Expression-Ready Clones of Mammalian Long Cdnas Encoding Large Proteins by the Orf Trap Cloning Method." <u>DNA Research</u> **12**(4): 257-267.

Nielsen, A., B. Der, J. Shin, P. Vaidyanathan, V. Paralanov, E. Strychalski, E. Ross, D. Densmore and C. Voigt (2016). "Genetic Circuit Design Automation." Science.

O'Neill, P. F., A. Ben Azouz, M. Vazquez, J. Liu, S. Marczak, Z. Slouka, H. C. Chang, D. Diamond and D. Brabazon (2014). "Advances in Three-Dimensional Rapid Prototyping of Microfluidic Devices for Biological Applications." Biomicrofluidics **8**(5): 052112.

Oberortner, E., S. Bhatia, E. Lindgren and D. Densmore (2014). "A Rule-Based Design Specification Language for Synthetic Biology." Journal of Emerging Technologies in Computing Systems **11**(3): 1-19.

Oberortner, E. and D. Densmore (2015). "Web-Based Software Tool for Constraint-Based Design Specification of Synthetic Biological Systems." ACS Synthetic Biology **4**(6): 757-760.

Orban, P. C., D. Chui and J. D. Marth (1992). "Tissue- and Site-Specific DNA Recombination in Transgenic Mice." Proceedings of the National Academy of Sciences of the United States of America **89**(15): 6861-6865.

Oxman, N. (2015). "Templating Design for Biology and Biology for Design." Architectural Design **85**(5): 100-107.

Paige, S., Iverson, S., Heuckroth, A., Carr, S., Haddock, T., Densmore, D. (2014). The Boston University Center of Synthetic Biology (Cosbi) Ice Repository. International Workshop on Bio-Design Automation.

Penhoet, E., M. Kochman, R. Valentine and W. Rutter (1967). "The Subunit Structure of Mammalian Fructose Diphosphate Aldolase*." Biochemistry **6**(9): 2940-2949.

Penhoet, E. E. and W. J. Rutter (1971). "Catalytic and Immunochemical Properties of Homomeric and Heteromeric Combinations of Aldolase Subunits." Journal of Biological Chemistry **246**(2): 318-323.

Pirani, A. (2008). Electron Microscopy and Single Particle Analysis of Two Actin-Based Macromolecular Assemblies: The Muscle Thin Filament and the Aldolase-Actin Complex, ProQuest.

Pirani, A., K. Allen, D. Tolan, R. Craig, V. Hatch and W. Lehman (2004). Em and 3d Reconstruction of Dimeric Aldolase on F-Actin. BIOPHYSICAL JOURNAL, BIOPHYSICAL SOCIETY 9650 ROCKVILLE PIKE, BETHESDA, MD 20814-3998 USA.

Quan, J. and J. Tian (2009). "Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways." PLoS One **4**: e6441.

Ritterson Lew, C. (2012). Non-Glycolytic Roles for Aldolase in Actin-Dependent Cellular Processes. 3500344 Ph.D., Boston University.

Ro, D. K., E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, M. C. Chang, S. T. Withers, Y. Shiba, R. Sarpong and J. D. Keasling (2006). "Production of the Antimalarial Drug Precursor Artemisinic Acid in Engineered Yeast." Nature **440**(7086): 940-943.

Rottmann, W. H., K. R. Deselms, J. Niclas, T. Camerato, P. S. Holman, C. J. Green and D. R. Tolan (1987). "The Complete Amino Acid Sequence of the Human Aldolase C Isozyme Derived from Genomic Clones." Biochimie **69**(2): 137-145.

Rottmann, W. H., D. R. Tolan and E. E. Penhoet (1984). "Complete Amino Acid Sequence for Human Aldolase B Derived from Cdna and

Genomic Clones." <u>Proceedings of the National Academy of Sciences</u> **81**(9): 2738-2742.

Rutter, W. J. (1964). "Evolution of Aldolase." <u>Fed Proc</u> **23**: 1248-1257.

Salis, H. M. (2011). "The Ribosome Binding Site Calculator." <u>Methods Enzymology</u> **498**: 19-42.

Salis, H. M., E. A. Mirsky and C. A. Voigt (2009). "Automated Design of Synthetic Ribosome Binding Sites to Control Protein Expression." <u>Nature Biotechnology</u> **27**: 946-950.

Sarrion-Perdigones, A. (2011). "Goldenbraid: An Iterative Cloning System for Standardized Assembly of Reusable Genetic Modules." <u>PLoS One</u> **6**: e21622.

Sarrion-Perdigones, A., M. Vazquez-Vilar, J. Palaci, B. Castelijns, J. Forment, P. Ziarsolo, J. Blanca, A. Granell and D. Orzaez (2013). "Goldenbraid 2.0: A Comprehensive DNA Assembly Framework for Plant Synthetic Biology." <u>Plant Physiology</u> **162**(3): 1618-1631.

Schindler, R., E. Weichselsdorfer, O. Wagner and J. Bereiter-Hahn (2001). "Aldolase-Localization in Cultured Cells: Cell-Type and Substrate-Specific Regulation of Cytoskeletal Associations." <u>Biochemistry and Cell Biology</u> **79**(6): 719-728.

Sherawat, M., D. R. Tolan and K. N. Allen (2008). "Structure of a Rabbit Muscle Fructose-1,6-Bisphosphate Aldolase a Dimer Variant." <u>Acta Crystallographica Section D</u> **64**(Pt 5): 543-550.

Sibley, L. D. (2003). "Toxoplasma Gondii: Perfecting an Intracellular Life Style." <u>Traffic</u> **4**(9): 581-586.

Siuti, P., J. Yazbek and T. K. Lu (2013). "Synthetic Circuits Integrating Logic and Memory in Living Cells." <u>Nature Biotechnology</u> **31**(5): 448-452.

Skerker, J. M., B. S. Perchuk, A. Siryaporn, E. A. Lubin, O. Ashenberg, M. Goulian and M. T. Laub (2008). "Rewiring the Specificity of Two-Component Signal Transduction Systems." <u>Cell</u> **133**(6): 1043-1054.

Sleight, S. C. and H. M. Sauro (2013). "Visualization of Evolutionary Stability Dynamics and Competitive Fitness of Escherichia Coli Engineered with Randomized Multi-Gene Circuits." <u>ACS Synthetic Biology</u> **2**: 519-528.

Smolke, C. D. (2009). "Building Outside of the Box: Igem and the Biobricks Foundation." <u>Nature Biotechnology</u> **27**(12): 1099-1102.

Spherotech, I. (2001). "Measuring Molecules of Equivalent Florochrome (Mef) Using Spherotm Rainbow and Ultra Rainbow Calibration Particles." <u>SPHEROTM Technical Note</u> (STN-9 Rev D. 041106).

Storch, M., A. Casini, B. Mackrow, T. Fleming, H. Trewhitt, T. Ellis and G. S. Baldwin (2015). "Basic: A New Biopart Assembly Standard for Idempotent Cloning Provides Accurate, Single-Tier DNA Assembly for Synthetic Biology." <u>ACS Synthetic Biology</u>.

Sygusch, J., D. Beaudry and M. Allaire (1987). "Molecular Architecture of Rabbit Skeletal Muscle Aldolase at 2.7-a Resolution." <u>Proceedings of the National Academy of Sciences</u> **84**(22): 7846-7850.

Tabor, J. J., H. M. Salis, Z. B. Simpson, A. A. Chevalier, A. Levskaya, E. M. Marcotte, C. A. Voigt and A. D. Ellington (2009). "A Synthetic Genetic Edge Detection Program." <u>Cell</u> **137**(7): 1272-1281.

Tamsir, A., J. J. Tabor and C. A. Voigt (2011). "Robust Multicellular Computing Using Genetically Encoded nor Gates and Chemical 'Wires'." <u>Nature</u> **469**(7329): 212-215.

Tolan, D. R., B. Schuler, P. T. Beernink and R. Jaenicke (2003). "Thermodynamic Analysis of the Dissociation of the Aldolase Tetramer Substituted at One or Both of the Subunit Interfaces." <u>Biological Chemistry</u> **384**(10-11): 1463-1471.

Utsumi, R., R. E. Brissette, A. Rampersaud, S. A. Forst, K. Oosawa and M. Inouye (1989). "Activation of Bacterial Porin Gene Expression by a Chimeric Signal Transducer in Response to Aspartate." <u>Science</u> **245**(4923): 1246-1249.

Wang, J., A. J. Morris, D. R. Tolan and L. Pagliaro (1996). "The Molecular Nature of the F-Actin Binding Activity of Aldolase Revealed with Site-Directed Mutants." <u>Journal of Biological Chemistry</u> **271**(12): 6861-6865.

Wang, J., D. R. Tolan and L. Pagliaro (1997). "Metabolic Compartmentation in Living Cells: Structural Association of Aldolase." <u>Experimental Cell Research</u> **237**(2): 445-451.

Weber, E., C. Engler, R. Gruetzner, S. Werner and S. Marillonnet (2011). "A Modular Cloning System for Standardized Assembly of Multigene Constructs." <u>PLoS One</u> **6**(2): e16765.

Welch, D. and J. B. Christen (2012). "Cmos Biosensor System for on-Chip Cell Culture with Read-out Circuitry and Microfluidic Packaging." <u>Conference Proceedings IEEE Engineering in Medicine and Biology Society</u> **2012**: 4990-4993.

Werner, S., C. Engler, E. Weber, R. Gruetzner and S. Marillonnet (2012). "Fast Track Assembly of Multigene Constructs Using Golden Gate Cloning and the Moclo System." Bioengineered Bugs **3**(1): 38-43.

Wijnen, B., E. J. Hunt, G. C. Anzalone and J. M. Pearce (2014). "Open-Source Syringe Pump Library." PLoS One **9**(9): e107216.

Yoshida, T., S. Phadtare and M. Inouye (2007). "The Design and Development of Tar-Envz Chimeric Receptors." Methods Enzymology **423**: 166-183.

Young, R. (1992). "Bacteriophage Lysis: Mechanism and Regulation." Microbiology Reviews **56**(3): 430-481.

Zhu, Y. and M. Inouye (2003). "Analysis of the Role of the Envz Linker Region in Signal Transduction Using a Chimeric Tar/Envz Receptor Protein, Tez1." Journal of Biological Chemistry **278**(25): 22812-22819.

Curriculum Vitae

Sonya Victoria Iverson
(406) 209-9692
siverson@bu.edu
206 Railroad Ave
Sunburst, MT 59482

Education

PhD at Boston University, January 2016

Master of Arts at Boston University, September 2015

Molecular Biology, Cell Biology and Biochemistry

Thesis: Improved modular multipart DNA assembly, development of a DNA part toolkit for *E. coli*, and applications in traditional biology and bioelectronic systems.

Bachelor of Science at Montana State University:

Biotechnology - Animal Systems, May 2007

Minors in Genetics and Psychology

Relevant Work Experience

PhD Candidate – Rational design and standardization in E. coli engineering
September 2011 to January 2016
Dr. Douglas Densmore, PhD (617) 358-6238
Electrical and Biomedical Engineering, Boston University

Research Associate, Lab Manager – Transgenic mouse models, liver cancer
June 2007 to October 2008; March 2009 to August 2011
Dr. Edward Schmidt, PhD (406) 994-6375
Immunology and Infectious Disease, Montana State University

Research Assistant - Transgenic mouse models, developmental biology
August 2004 to May 2007
Dr. Edward Schmidt, PhD (406) 994-6375
Immunology and Infectious Disease, Montana State University

Research Assistant – Evaluation of medicinal plants, novel pharmaceuticals
January 2004 to October 2004
Dr. Gary Strobel, PhD (406) 994-5148
Plant Science Plant Pathology, Montana State University

Research Skills

Molecular Biology
- DNA/RNA isolation, PCR, RT-PCR, qPCR, cloning, mutagenesis
- DNA assembly methods: Gibson, Golden Gate, MoClo (Type IIS)
- Yeast 2-hybrid, other protein-protein experimental methods
- Familiar with antibody based assays, ELISA, western, etc.

Cellular Biology & Biochemistry
- Graduate level Biochemistry (2 semesters)
- Quantitative Biochemistry course
- Flow cytometry and Fluorescence Activated Cell Sorting (FACS)
- Trained on a BD Fortessa Flow Cytometer (advanced user), BD Aria
- Microscopy: Nikon/Olympus compound and confocal microscope systems
- Histology, immunohistochemistry, fluorescence in-situ hybridization (FISH)

Data analysis and tracking and presentation
- Expert skill with Excel, and Powerpoint
- Proficient with GraphPad Prism, various other statistical software
- Bioinformatics tools - ClustalW, MEGA, NCBI BLAST, Gene Ontology
- Adobe Photoshop and Illustrator to make schematics, figures, etc.

Mammalian research biology
- Mouse models of evolutionary biology, liver development and cancer
- Designing transgenic mouse systems, cre-lox and FRT/FLP systems
- Proficient handling mice: injections, IVF, blood sampling, tissue harvest

BioElectronics, Microfluidics, Nanotechnology
- Involved in 3D printed microfluidics bioelectronics research projects
- Familiar with microfluidic fabrication methods
- Class experience with nanotechnology and nanomedicine

Experience with *E. coli, S. cerevisiae, Drosophila sp.,* lambda phage
Literature research and critical analysis
Extensive experience writing, editing, and reviewing papers and grants

Awards and Meetings

- · Invited to speak at Systems and Synthetic Biology Summer School, Taormina Sicily, 2015
- · Presented at Synberc, Boston and Berkeley, multiple years 2012-14
- · Presented at Synthetic Biology 6.0 Workshop in London, UK 2013
- · Attended IWBDA meetings in London and Boston, 2013 and 2014
- · National Conference on Undergraduate Research (NCUR), 2007
- · Research Experience for Undergraduates (REU, NSF), 2006
- · IDeA Networks of Biomedical Research Excellence (INBRE), 2005-06

Additional Experience

Visiting Scientist – International Center for Agriculture in Dryland Areas (ICARDA) - Aleppo, Syria, 2008
Visiting Scientist – Gates Foundation Grand Challenges Project - Maseno, Kenya, 2011
Science Advisor – Montana Gluten Free Processors, ongoing since 2012
Science Mentor – iGEM team, BostonU 2012 and 2013, other mentorships

Publications (Peer Reviewed Journals)

Iverson, S.V., T. Haddock, J. Beal, D. Densmore. CIDAR MoClo: Improved MoClo Assembly Standard and New *E. coli* Part Library Enables Rapid Combinatorial Design for Synthetic and Traditional Biology. ACS Synth Biol (2015)

Iverson, S.V., S. Eriksson, J. Xu, J.P. Prigge, E.A. Talago, T.A. Meade, E.S. Meade, M.R. Capecchi, E.S. Arner, E.E. Schmidt. Altered metabolic profile in livers of mice lacking Thioredoxin Reductase 1 protects against acetaminophen toxicity. Free Radical Biology and Medicine 63:369-380 (2013)

Prigge, J.R., S. Eriksson, S.V. **Iverson**, T.A. Meade, M.R. Capecchi, E.S. Arnér, E.E. Schmidt. Hepatocyte DNA replication in growing liver requires either glutathione or a single allele of txnrd1. Free Radical Biology and Medicine 52:4 803–810 (2012)

Iverson, S.V., K.M. Comstock, C.M. Weisend, J.A. Kundert, E.E. Schmidt. Contributions of new hepatocytes lineages to liver growth, maintenance, and regeneration. Hepatology 54:2 655-663 (2011)

Prigge, J.R., S.V. **Iverson**, A.M. Siders, E.E. Schmidt. Interactome for auxiliary splicing factor U2AF65 suggests diverse roles. Biochimica et Biophysica Acta 1789 487–492 (2009)

Bondareva, A. A., M. R. Capecchi, S. V. **Iverson**, Y. Li, N. I. Lopez, O. Lucas, G. F. Merrill, J. R. Prigge, A. M. Siders, M. Wakamiya, S. L. Wallin, and E. E. Schmidt, 2007. Effects of thioredoxin reductase-1 deletion on embryogenesis and transcriptome. Free Radical Biology & Medicine 43 911-923 (2007)

Publications (Not Peer Reviewed)

Haddock, Traci L., Douglas M. Densmore, Evan Appleton, Swati Carr, Sonya **Iverson**, Monique De Freitas, Shawn Jin et al. "BBF RFC 94: Type IIS Assembly for Bacterial Transcriptional Units: A Standardized Assembly Method for Building Bacterial Transcriptional Units Using the Type IIS Restriction Enzymes BsaI and BbsI." (2015).

Publications (Posters)

Heuckroth, A., Huang, C., Silva, R., Iverson, S., Haddock, T., Pacheco, A., Douglas Densmore. Accessible Microfluidic Mold Fabrication Using 3D Printing. Synberc Spring Retreat, Berkeley CA (2015).

Pacheco, A., Silva, R., Iverson, S., Haddock, T., Douglas Densmore. Multicellular Logic through Conversion of Genetic Circuit Outputs into Electrical Signals. Synberc Spring Retreat, Berkeley CA (2015)

Iverson, S., Beal, J., Haddock, T., Douglas Densmore. E. Coli MoClo assembly & part library: Applications in traditional biological research. Systems and Synthetic Biology Summer School SSBSS, Taormina Sicily (2015)

Iverson, S., Beal, J., Haddock, T., Douglas Densmore. Efficient Modular, Mulitpart and Multiplex DNA Assembly System. International Workshop on Bio-Design Automation IWBDA, Boston MA (2014)

Paige, Stephanie, Sonya Iverson, Aaron Heuckroth, Swati Carr, Traci Haddock, and Douglas Densmore. The Boston University Center of Synthetic Biology (CoSBi) ICE Repository. International Workshop on Bio-Design Automation IWBDA, Boston MA (2014)

Iverson, S., Haddock, T., Douglas Densmore. Multiplex Modular Assembly and Protein Engineering. Synthetic Biology 6.0 London UK (2013)

Iverson, S., Haddock, T., Douglas Densmore. Multiplex Modular Assembly and Protein Engineering. Synberc Fall Retreat Cambridge MA (2013)

Iverson, S., Haddock, T., Douglas Densmore. Development of a Multiplex Modular Cloning System. Synberc Spring Retreat Berkeley CA (2013)

Iverson, S. Douglas Densmore. Context dependent characterization of promoter libraries and site-specific recombinase systems. Synberc Fall Retreat Cambridge MA (2012)