2015

ex vivo DNA cloning

Adam B. Fisher

Virginia Commonwealth University

Follow this and additional works at: https://scholarscompass.vcu.edu/etd

Part of the Biological Engineering Commons, Biotechnology Commons, Molecular Biology Commons, and the Molecular, Cellular, and Tissue Engineering Commons

© The Author

Downloaded from

https://scholarscompass.vcu.edu/etd/3962

This Dissertation is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.
ex vivo DNA Cloning
"Now this is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning."
- Sir Winston Churchill, 1947

As I reach the conclusion of a graduate career that seems like an entire lifetime unto itself, there are numerous special people to recognize for making this experience not only possible but also successful. While not nearly adequate the best I can do is give my acknowledgement and thanks where it may be preserved and disseminated.

First I must thank my parents, who both worked full-time so that I should have the freedom to go college and pursue my dreams. From the start you were my inspiration to dedicate my life to that which I am passionate and it has meant everything to me. Thank you for always supporting and encouraging me at every step even when I made that difficult. None of this is possible without you.

To my sisters for always keeping me humble, thank you for taking this difficult task so seriously. You both never let me forget where I come from and this brings me incredible pride. Similarly my friends, many of which I have stayed close with since the start of grade school, have blessed me with guidance, perspective and an unfathomable loyalty.

The long, winding path to where I am today would’ve diverged often without the advice, inspiration and compassion of many teachers, coaches, professors and advisors. To the directors of the Integrative Life Sciences program, Dr. Robert Tombes and Dr. Bill Eggleston, thank you for accepting me and continually being an open resource along with the rest of the program’s staff and faculty. Thank you Dr. Paul Fawcett for teaching me the beautiful and useful art of programming and computation. I owe a large thanks to each of my committee members: Drs. Allison Johnson, Matthew Hartman, Michael Peters and B. Frank Gupton. Each member contributed beyond being active committee members: Dr. Johnson for being the model for bringing science to life in the classroom, Dr. Hartman for an incredible chemical biology course that taught me to see the science beyond the article, Dr. Peters for our collaborative discussions and Dr. Gupton for demonstrating how to be the captain and navigator of not only a lab, but a department.
My fellow labmates, all of whom I’m fortunate to call friends, thank you all for being there as peers, advisors – both professionally and personally – and collaborators. This includes the many incredible undergraduate researchers who have passed under my tutelage. Thank you for remaining patient with me and all your help. In particular, a big thanks is necessary to Zachary Canfield who was integral in doing many of the experiments. I have only the highest of hopes for your future, never settle for anything less. Most of all, I owe more than I can say to my peer and close friend, Dr. George McArthur IV. Whiteboards, 821 café, conferences, posters, papers and projects – we’ve done it all together. It has been an honor and privilege to work and grow alongside you. The future is here for us both and I can only hope that our future will be shared as much as possible. Thank you, George.

Dr. Fong, I owe much of this degree to your belief and continuing support. It is rare to find an advisor more dedicated to his students and their vision. You cultivated my passion for science and biology by allowing me to walk my own path without letting me feel as though I walked it alone. Moreover, daily you exemplify what it means to be a husband and father. I can only hope I am able to come close to balancing my own life so steadily. Thank you again for not only enabling me to be a great scientist, but also showing me how to be a great man. And thank you for always keeping the door open.

Lastly but most of all thank you, Tricia Herbert. Sometimes the universe places your future right in front of you and we in our limited sight are not able to grasp the beautiful significance until much later. The moment you walked into my first collegiate class and sat next to me, I had no idea what the universe had just laid in front of me. I am the luckiest man in the world to have been given these past five years with such a beautiful, kind and intelligent woman. Through two graduate programs; living apart, together or with our families; traveling back-and-forth between Richmond and Lynchburg or traveling around the world; through our successes and failures; health and sickness; I have never once not needed you by my side. Out of the many, many things I’ve learned throughout the past four years this is the most important. What the future holds, I have no idea. All I can do is hope that somehow I am blessed enough to end up an old man holding your hand thinking of how the universe could be so wonderfully serendipitous. I love you Tricia and cannot wait to share the rest of this adventure we call life together.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................. ii
TABLE OF CONTENTS ...................................................................... iv
FIGURES AND TABLES ................................................................... vi
ABBREVIATIONS ......................................................................... viii
ABSTRACT .................................................................................... ix

Chapter 1. Introduction and Background ............................................. 1
  1.1. DNA Assembly and Synthetic Biology ......................................... 1
  1.2. The ex vivo approach ............................................................................. 7
  1.3. Applications of Cell-Free Systems ...................................................... 8
  1.4. Double-strand Break Repair: Natural DNA Assembly ..................... 11
  1.5. The ex vivo Cloning approach ............................................................ 14

Chapter 2. ex vivo DNA ASSEMBLY .................................................. 16
  2.1. Initial 2-way ex vivo DNA Assemblies .............................................. 16
  2.2. Testing the Magnesium:ATP Ratio with 2-way Linear Assemblies ........ 18
  2.3. Optimized Assemblies with Yeast and E. coli Lysates ....................... 21
  2.4. Conclusions and Discussion ............................................................. 24

Chapter 3. Optimization of ex vivo DNA Assembly ..... 26
  3.1. Lysate Preparation ............................................................................. 26
  3.2. Optimizing the ex vivo DNA assembly reaction .............................. 40

Chapter 4. Probing ex vivo DNA Assembly with Enzymatic
           Titrations ................................................................................. 57
  4.1. Enzymatic Titrations ......................................................................... 57
  4.2. RecA Titrations ................................................................................ 62
  4.3. Strain Engineering (nuc4- and autolysis cassette) ............................ 65

Chapter 5. ex vivo PCR and ex vivo Cloning ................................. 70
  5.1. The pUN plasmid and ex vivo PCR ................................................ 70
  5.2. ex vivo Cloning ................................................................................ 79

Chapter 6. Conclusion and Future Directions ................................. 82

Chapter 7. Materials and Methods ..................................................... 89
7.1. Reagents, DNA and Enzymes ........................................... 89
7.2. Generation of DNA Fragments for ex vivo DNA Assembly .... 89
7.3. Preparation of Lysates .................................................. 90
7.4. DNA Assembly Reactions ............................................. 93
7.5. Agarose Gel Analyses and Extraction .............................. 94
7.6. Transformations and Preparation of Competent Cells ........ 95
7.7. Lambda RED Recombineering ........................................ 97

Chapter 8. References ..................................................... 98

Appendix A. Tables ....................................................... 1
Appendix B. Gel Pictures .................................................. 5
Appendix C. DNA Sequences .............................................. 18
   C.1 Primer Sequences .................................................. 18
   C.2 Plasmid Sequences ................................................ 23
FIGURES AND TABLES

Figure 1.1 DNA Assembly by a Chew-back, Anneal and Repair Mechanism 5
Figure 2.1 The 2-way Circular Assembly ................................. 17
Figure 2.2 The 2-way Linear Assembly and Buffer Optimization .... 20
Figure 2.3 The 3-way Assembly and Colony Counts ....................... 21
Table 2.1. Effect of Detergents on Assembly and Transformation .... 24
Figure 3.1 Harvested Growth Stage on Assembly Efficiency ........... 27
Figure 3.2 Rich (TB) versus Minimal (M9) media ...................... 30
Table 3.1 Characteristics of Protein Storage Methods .................. 32
Figure 3.3 Glycerol negatively affects lysate activity .................. 33
Figure 3.4 Detrimental effects of detergents on electroporation ....... 39
Figure 3.5 Concentration of Lysate ......................................... 45
Figure 3.6 Optimization of components of “Base” Buffer ............... 47
Figure 3.7 Titrations of Divalent Metal Cations ......................... 50
Figure 3.8 Titrations of Magnesium and Manganese ..................... 52
Figure 3.9 Lane composition of Manganese to Magnesium Titration .... 53
Figure 3.10 Effect of Manganese on Two-way Circular Assembly ...... 55
Table 4.1 Enzymes used in Titration ........................................ 58
Figure 4.1 Titrations of Various DNA cloning Enzymes ................. 59
Figure 4.2 Lane Composition of Enzymatic Titration .................... 61
Figure 4.3 Titrations of RecA .................................................. 63
Figure 4.4 Lane composition of RecA titration ............................ 64
Figure 4.5 Colony counts of exDA reactions containing RecA ........ 65
Figure 4.6 ex vivo DNA Assembly with nuc4-λR and Lambda RED .... 68
Figure 5.1 Expression of Pfu DNA Polymerase ............................. 76
Figure 5.2 Volumetric titrations of ex vivo PCR .......................... 77
Figure 5.3 ex vivo PCR of 2-way and 3-way Fragments .................. 78
Figure 5.4 ex vivo Cloning ..................................................... 80
Table A.1 Summary of Assemblies ......................................... 1
Table A.2 Strain Genotypes .................................. 2
Table A.3 Description of Relevant Genotypes .................. 3
Table A.4 Unit Definitions of Enzymes used in Titrations .......... 4
Figure B.1 Dra, Sce and Eco Time Course ........................ 5
Figure B.2 Lysate Preparation (Gel 1) .......................... 6
Figure B.3 Lysate Preparation (Gel 2) ............................ 7
Figure B.4 Lysate Dilutions (60 minutes) ........................ 8
Figure B.5 Lysate Dilutions (120 minutes) ...................... 8
Figure B.6 Buffer Optimization (NEB10β) ....................... 9
Figure B.7 Buffer Optimization (XJa Autolysis) .................. 10
Figure B.8 Buffer Titration (NEB10β #1) ....................... 11
Figure B.9 Buffer Titration (NEB10β #2) ....................... 12
Figure B.10 Buffer Titration (XJa) ............................... 13
Figure B.11 Manganese Titration (NEB10β) ...................... 14
Figure B.12 RecBCD, PolI, RecA, Eco Ligase Titration .......... 15
Figure B.13 ssDNA Exonucleases Titration ..................... 16
Figure B.14 RecA Titration (XJa Autolysis) .................... 17
Table C.1 ex vivo Assembly Assay Primers ....................... 18
Table C.2 pUN and UNS Primers ................................ 19
Table C.3 Oligos and Primers used for Strain Engineering ....... 21
ABBREVIATIONS

BCP = Blue Chromogenic Protein (also, amilCP)
BSA = Bovine Serum Albumin
CDS = Protein Coding Sequence
CFS = Cell-Free System
CFU = Colony-Forming Unit
DHAP = dihydroxyacetone phosphate
*Dra* = *Deinococcus radiodurans* lysate
DSB = Double-stranded break
dsDNA = double-stranded DNA
*Eco* = *Escherichia coli* lysate
EE = Early exponential
IDT = Integrated DNA Technologies
ISO = Isothermal DNA assembly (also, Gibson assembly)
LB = Luria-Bertani media
LE = Late exponential
M9 = M9 minimal media
MAGE = Multiplex Automated Genome Engineering
NHEJ = Non-homologous end joining
PCA = Polymerase Chain Assembly
PCR = Polymerase Chain Reaction
*Pfu* = *Pyrococcus furiosus*
pUN = Plasmid Unique Neutral
RBS = Ribosome Binding Site
*Sce* = *Saccharomyces cerevisiae* lysate
SOB = Super Optimal Broth media
SOC = Super Optimal Broth with Catabolite repression Media
SP = Stationary Phase
SSA = Single-strand annealing
ssDNA = single-stranded DNA
Tag = *Thermus aquaticus*
TAR = Transformation-associated recombination
TB = Terrific Broth media
UNS = Unique Neutral Site
Genetic engineering of microbes has developed rapidly along with our ability to synthesize DNA de novo. Yet, even with decreasing DNA synthesis costs there remains a need for inexpensive, rapid and reliable methods for assembling synthetic DNA into larger constructs or combinatorial libraries. While technological advances have resulted in powerful techniques for in vitro and in vivo assembly of DNA, each suffers inherent disadvantages. Here, an ex vivo DNA cloning suite using crude cellular lysates derived from E. coli is demonstrated to amplify and assemble DNA containing small sequence homologies. Further, the advantages of an ex vivo approach are leveraged to rapidly optimize several parameters of the ex vivo DNA assembly methodology testing lysates from different engineered strains of E. coli, with various buffer components and using titrations of purified cloning enzymes. Finally, in order to complete the cloning suite, a vector expressing the Pyrococcus furiosis (Pfu) DNA polymerase was designed, constructed and expressed in E. coli to create a ‘functionalized lysate’ capable of ex vivo PCR. Not only do we demonstrate ex vivo cloning methodology as a complete cloning package capable of replacing the expensive cloning reagents currently required by synthetic biologists, but also establish ex vivo as an overarching approach for conducting molecular biology.
Chapter 1. Introduction and Background

“What I cannot create, I do not understand”
- Richard Phillips Feynman, PhD

1.1. DNA Assembly and Synthetic Biology

Our capacity to (re)engineer living systems is linked to our ability to physically build specific DNA molecules that encode desired functionality and behavior. Thanks to major advances in DNA sequencing and DNA synthesis technologies, a renaissance in genetic engineering known as synthetic biology has sparked incredibly rapid progress in our ability to engineer life. Synthetic biology, metabolic engineering, systems biology and associated disciplines can be leveraged to not only access the estimated ~100 million biological compounds present on our planet, but create sophisticated biological entities as molecular machines capable of replication, catalysis, sensing and computational response\(^1\). The programmable chemistry potentiataed by recombinant DNA technologies can be harnessed to address societal needs such as food scarcity; renewable fuel sources; environmental remediation and protection; medicinal (and veterinary) diagnostics and therapeutics; or defense against bioterrorism\(^2,3\).

Due to difficulties stemming from the inherently ‘analog’ nature and complex interdependence of biochemical systems on top of the staggering permutations of even the smallest of genetic circuits,
synthetic biologists have focused on the characterization, standardization and modularity of biological parts. Yet, to achieve the overarching synthetic biology dream we must continue to optimize the design-build-test engineering loop until first principles of de novo genetic design are ascertained.

While the current bottleneck in the Synthetic biology engineering loop lies in the design portion of the loop (since most initial designs are not optimal and possibly not even functional), the only way to alleviate this design bottleneck is to continue to build and test many genetic designs. This is not just to identify adequate designs for the application at hand, but also to inform subsequent design efforts. The ensuing study not only presents an optimized DNA assembly methodology significantly less expensive and/or time-consuming than current methods, but also presents an overall approach to rapid phenotypic characterization, optimization and alternate route for laboratory molecular biology.

Although our ability to construct novel DNA has come a long way from Nirenberg’s initial synthesis of a nucleic acid homopolymer, it still proves most cost effective to purchase DNA in small fragments or single-stranded oligomers to assemble into larger, gene-length products. This is particularly true when combinatorial libraries are being constructed. Indeed, synthetic biologists often rely on construction and thorough characterization of libraries of modular genetic parts. For instance, characterization of promoter regions and ribosome binding sites (RBS) has generated successful part libraries
capable of tuning transcriptional and translational expression of downstream protein coding sequences\textsuperscript{6-8}. Data collected from these experiments have even produced fairly robust models capable of rationally designing tunable parts, such as the RBS calculator\textsuperscript{9}. Yet the interplay between DNA sequence and phenotype is full of nonlinear, epistatic and stochastic dynamics, minimizing the modularity and reliability of genetic parts. For example, combining a highly expressive promoter with a highly expressive RBS part may result in very low overall protein expression. To address these problems synthetic biologists must often rely on parts tested in a combinatorial manner. Beyond regulatory parts for modulating gene expression, encoding novel functions through protein engineering still remains an undertaking of brute force necessitating generation of extreme sequence diversity through directed evolution (requiring an appropriate screening approach), combinatorial domain-swapping or a synthetic metagenomics approach. In order to address the design bottleneck of the design-build-test loop, the build and test steps must be made extremely rapid.

Traditionally, recombinant DNA has been constructed using restriction cloning (i.e., cutting with an endonuclease and joining with a ligase) and DNA fragments isolated from natural sources by restriction digestion or polymerase chain reaction. In addition to restriction-ligation approaches, other site-specific recombination systems have been employed to assemble DNA fragments with great success\textsuperscript{10}. However, genetic engineering has recently become more flexible with the use of
'sequence-independent' approaches that take advantage of the decreasing cost of DNA synthesis\textsuperscript{11-14}. In a perfect world, synthesis of DNA would be so cheap that all DNA constructs could be synthesized in their entirety, sequence-verified and delivered to researchers for testing. However, current gene synthesis rates are highly limiting often costing several thousand dollars to synthesize full-length genes. Instead, most DNA is ordered in more cost-effective fragments (as of this writing, single-stranded oligomers are $0.35 per base up to 60 bases and $120 for a 750 base pair double-stranded fragment through Integrated DNA Technologies\textsuperscript{\textregistered} [IDT]) and then assembled using overlapping homologous sequences on the ends of the oligomers in a 'sequence-independent' fashion. For example, polymerase chain assembly (PCA), a variation on the general polymerase chain reaction (PCR), uses partially overlapping 60bp single-stranded oligomers alternating strands of an entire gene sequence in equimolarity as a template. These oligos are pooled and allowed to anneal and extend into longer and longer double-stranded sequences until a complete double-stranded gene sequence is generated. These 'synthons' can be quickly inserted into a plasmid backbone by traditional sequence-specific or sequence-independent methods.
Figure 1.1 DNA Assembly by a Chew-back, Anneal and Repair Mechanism

Depicted are the steps employed by Gibson Isothermal assembly (ISO) and analogous putative steps for in vivo single-strand annealing predicted to be the active mechanism in ex vivo DNA assembly. In steps 1-3, an exonuclease (here 5′-to-3′) chews back (resects) the double-stranded ends to expose terminal homologies (colored blue), which can anneal to each other in step 4. If there has been excessive resection as depicted in steps 5-6, a DNA polymerase is used to repair the exposed gaps. Finally, a DNA ligase acts to seal the DNA backbone (7), to yield a fully sealed, accurate assembly/repair. The table describes the tripartite mixture of enzymes in the ISO formulation and some analogous naturally-occurring enzymes from yeast, E. coli and D. radiodurans.

When trying to assess the function of DNA sequences in the cell, it is of great utility to have sequence-independent cloning techniques that do not constrain design (e.g., scars from restriction sites).

Homologous recombination is the longest-tenured sequence-independent cloning method and the most widely utilized for integration onto the genome. Highly active homologous recombination systems native to Bacillus subtilis and S. cerevisiae make effective in vivo recombination systems, while in E. coli, the Lambda Red recombinase
genes can greatly increase the rates of homologous recombination\textsuperscript{15}. In fact, \textit{S. cerevisiae} has been utilized to assemble and propagate entire bacterial genomes from very large overlapping fragments, through transformation-associated recombination (TAR)\textsuperscript{16}. However, the assemblies must be designed to work with the organismal chassis and, unless the construct is to remain in the assembly host, must be designed to shuttle into the final host organism. The TAR cloning process can therefore take up to eight days to get a DNA construct into the terminal host. While monetarily inexpensive, these time expenses are often a disappointing attribute of \textit{in vivo} methodologies.

In contrast to \textit{in vivo} approaches, purified \textit{in vitro} systems offer user-control, efficiency and time savings, but are often expensive to purchase or difficult to establish in-house due to complications in protein purification. Once established there is no denying the potency of these systems. In 2008, a completely synthetic Mycoplasma genome was assembled from chemically synthesized double-stranded DNA (dsDNA) fragments (5-7 kilobase pairs each) using an \textit{in vitro} chew-back assembly method\textsuperscript{16}. More recently, the mouse mitochondrial genome was reconstructed from overlapping synthetic single-stranded oligonucleotides (60 nucleotides each) using a three-enzyme \textit{in vitro} isothermal DNA assembly method (ISO assembly, also known as Gibson assembly)\textsuperscript{17}. Currently the most popular and efficient sequence-independent cloning method, Gibson and colleagues discovered an optimized mixture of a phage 5′-to-3′ DNA exonuclease, DNA ligase, and heat-stable polymerase catalyzing highly efficient DNA end joining.
(Figure 1.1). As the current “gold-standard” for DNA assembly, Gibson assembly suffers from the disadvantages of most purified in vitro systems – it is highly expensive. Distributed as a master mixture from New England Biologicals® (through Synthetic Genomics®) each reaction costs $15 dollars. While very efficient and taking only an hour, if trying to assemble larger libraries of constructs this is not a cost-effective route.

1.2. The ex vivo approach

The terminology ‘ex vivo’ – Latin for “out of the living” – is a common term in organismal biology, where tissue or even full organs are isolated from an organism with minimal perturbation of it’s natural environment. This permits researchers more control over conditions and access to measurements not otherwise possible (or even ethical) in a living organism. In this manner, the “ex vivo” moniker describes an experiment combining the advantages of an in vivo approach and an in vitro approach.

While ex vivo experimentation is traditionally associated with organs and tissues, the approach has a storied history in microbiology and cellular biology. The first ex vivo reaction catalyzed by a microbial cell extract was performed by Eduard Buchner (1897) over 100 years ago, converting sugar to ethanol and carbon dioxide in a yeast extract, eventually winning the Nobel Prize in Chemistry (1907). Using lysate fractions derived through differential centrifugation, microsomes were discovered as catalysts of protein synthesis, leading
to the identification of mRNA as the template for protein synthesis by Nirenberg and Matthaei (1961) in their own extract-based reaction. As history and technology progressed the microbiology and cellular biology systems derived from cellular extracts have become popularly referred to as ‘cell-free systems’ (CFS). However, ‘cell-free system’ is an umbrella term giving very little detail to the composition of the final system. For instance, cell-free systems for protein synthesis range from crude extract to systems reconstituted in vitro with entirely purified components. For this reason, we propose ex vivo be used to describe a CFS which is minimally altered from its’ in vivo origins by only very crude purification or manipulation (e.g., centrifugation of insoluble components).

1.3. Applications of Cell-Free Systems

Since their inception, CFSs have primarily been leveraged to probe the central processes of transcription and translation. Over the past ~25 years, building off work establishing coupled transcription-translation CFS, modern CFS have evolved into robust systems capable of high-level protein production. Recent work with E. coli-based extracts have garnished yields exceeding 2 mg/mL protein at scales exceeding 100 liters. Complex eukaryotic proteins containing high-levels of post-translational modifications can be produced in cell-free extracts from wheat-germ, rabbit reticulocytes and insect cells at high titers. This is important, since CFSs obviate the need to perform in vivo manipulations that can be difficult or impossible in some eukaryotic cell lines. Cell-free protein production is therefore
particularly attractive to industrial production of therapeutic peptides as they are commonly post-translationally modified, exist associated with the membrane or may exhibit toxicity (e.g. antimicrobial peptides).

For synthetic biology, CFSs have been cleverly employed to expedite the prototyping of genetic parts for gene expression. By monitoring the production of fluorescent proteins in a CFS the time investment and noise (i.e., plasmid copy number) associated with in vivo characterization is greatly reduced. Although not an absolute perfect measure of in vivo dynamics, a set of constitutive and inducible promoters tested in a CFS were shown to correlate well with measurements of in vivo gene expression and dose-response²⁰.

Cell-free protein synthesis has benefited greatly from advances in mimicking intracellular conditions, activating integrated biological processes and controlling cell-free metabolism. Control over metabolism in CFSs has motivated the development of CFSs as platforms for the synthesis of metabolites. Although most of these higher-functioning CFSs require further purification of the pathway constituents from the lysate, in a truly excellent demonstration of an ex vivo system, Panke and colleagues optimized the catalysis of dihydroxyacetone phosphate (DHAP) from glucose²¹. The authors were able to reach concentrations far beyond those documented in vivo (12 mM DHAP) by deletion of two genes in the production strain combined with supplementation of butanol and rabbit muscle aldolase to the reaction. DHAP is an unstable product, but the exogenous addition of the
aldolase and butanol allowed conversion to a more stable form. This is a perfect example of the simplicity and flexibility of CFS optimization, not only by modification of the host chassis, but also by circumventing the membrane to directly supply new components (here the small-molecule butanol and a non-bacterial enzyme) impossible in the same in vivo system.

The constant struggle between the engineering objective to optimize product yield and the cell’s biological objective to optimize biomass production is remedied by an ex vivo approach. The functionalities of the cell remain intact, while removal of the cell encapsulation enables manipulation of the internal chemistries to activate, integrate and focus cellular resources towards a desired objective. Without compartmentalization there is the opportunity to greatly exaggerate substrate concentrations, remove product, and rapidly sample and monitor reaction conditions. In in vivo systems researchers and engineers must respect the constraints, time-scales and stochasticity inherent to life. On the other hand, purified in vitro systems provide exceptional flexibility and control, resulting in higher reaction efficiencies and better reproducibility. Yet, in vitro systems are low-throughput and orders of magnitude more expensive than their in vivo counterparts. This is clearly demonstrated in the production of cellulosic biofuels, where the primary cost-driver is the expense of purified enzymes for biomass pretreatment. For investigators, in vitro systems sacrifice cellular context necessary to make conclusive inferences about in vivo phenomena. A simple, yet
elegant compromise between in vivo and in vitro approaches, ex vivo systems enhance both engineering and investigatory freedom.

1.4. Double-strand Break Repair: Natural DNA Assembly

The end-to-end joining of free dsDNA fragments is a vitally important process for all cellular life in the repair of double-stranded DNA breakage. DNA double-stranded breaks (DSB) are ubiquitous across all kingdoms of life, occurring during normal housekeeping functions and in response to DNA damaging agents such as chemicals and irradiation. Very simply, a DSB is a break in both strands of duplex DNA. A DSB is a particularly dangerous DNA lesion as these breaks can potentially disrupt gene expression, alter chromosome organization or provide substrates for single- or double-stranded exonucleases, catalyzing the deletion of genetic information. DSB are also notoriously difficult to repair. Where a single-stranded break can be easily repaired using the conserved information on the intact strand, generally DSB have no associated template at the free ends to guide repair. Due to these two reasons DSB repair is widely associated with the development of cancer and has, therefore, been the subject of extensive study. There are two major mechanisms for DSB repair: end-joining and homologous recombination. The major distinction between the two being that the former requires no template to facilitate repair of the DSB. The presence of template DNA is, therefore, essential in the decision and mechanism of DSB repair pathways.
1.4.1. Homologous Recombination

Homologous recombination generally leads to high-fidelity repair, making it the most ubiquitous and dominating DSB repair pathway across organisms. However, unlike end-joining, homologous recombination’s requirements for a template means that a second copy of the region containing the DSB must exist in the cell. For most prokaryotes, this is not a problem as they often exist with multiple or, at least, partial copies of their genome. In eukaryotes, homologous recombination is reserved for the S and G₂ phases of the cell cycle when sister chromatids are present, acting as template for repair and meiotic crossing-over.

1.4.2. Non-homologous End-Joining

End-joining implies the linking of two free duplex DNA termini into a ligated product. Traditionally referred to as non-homologous end-joining (NHEJ), the relatively simple process relies on minor processing of DNA termini to prepare them for direct ligation in the absence of any required homologous template. While this means that repair can take place at any point during the cell cycle regardless of genome copy number, the fidelity of repair is relatively low. Errors are often propagated from loss-of-information during end processing (resection), explaining the relatively low occurrence in prokaryotes in comparison to homologous recombination. In higher eukaryotes, the error-prone nature of NHEJ has evolved into a highly coordinated system to generate sequence diversity in the differentiation of unique
B- and T-lymphocytes. More recent studies have revealed ‘alternate end-joining’ or ‘microhomology-mediated end-joining’ pathways utilizing repeat sequences at or near the terminus of DSBs, such as those that may occur in breaks near or within ribosomal operons of prokaryotes\textsuperscript{28,29}.

1.4.3. Alternative End-Joining

Alternative end-joining pathways appear to exist in a grey-area between the extremes of NHEJ and homologous recombination relying on combinations of cellular machineries from both processes, distinguished by the degree of available homologous template. Outside of rare circumstances, like the ribosomal operon described above, a DSB is very unlikely to form with significant end homologies to guide non-templated repair. Nevertheless, an alternative DSB repair pathway known as single-strand annealing (SSA) exists to join DSB with significant homologies at their termini\textsuperscript{30}. As in the homologous recombination process, SSA involves an extensive end-resection to expose single-strand overhangs\textsuperscript{31}. Yet, in contrast to homologous recombination, these ends subsequently anneal to the newly exposed homologies on the cognate DSB for a final step of ligation (similar to NHEJ). SSA is a mechanistic analog to the Gibson ISO assembly method (Figure 1.1), following the same chew-back (resection), annealing, polymerization and ligation steps, but occurs (albeit infrequently) in completely wild-type microbes. Hopefully, by freeing and concentrating the enzymes necessary for SSA end joining of overlapping DNA fragments will take place. \textit{E. coli}, and \textit{S. cerevisiae} make good chassis not only
due to preliminary research suggesting they are capable of catalyzing end joining in an extract-based reaction, but their relatively easy culturing methods and availability to most labs\textsuperscript{13,32}. \textit{D. radiodurans} makes another excellent chassis for exDA, as it is known to have some of the most active DNA repair systems, making it the world’s (current) most radioresistant organism\textsuperscript{33}.

1.5. The \textit{ex vivo} Cloning approach

In the ensuing study, we endeavored to demonstrate the facilities of the \textit{ex vivo} approach to demonstrate and subsequently investigate and optimize DNA assembly. Using our own assays of assembly we tested lysates prepared from different organisms and strains, under varying growth conditions and with variable buffer/reaction conditions. Titrations of common cloning enzymes were employed to investigate the underlying enzymatic mechanisms of DNA assembly. Applying genome engineering techniques we endeavored to engineer a new strain of \textit{E. coli} lacking competing nucleases and capable of autolysis for \textit{ex vivo} DNA assembly. We designed and constructed pUN (‘plasmid unique neutral’) a novel plasmid featuring a screenable blue chromogenic protein (BCP) insert and ‘unique neutral sites’ for facile cloning via \textit{ex vivo} DNA assembly or other sequence-independent assembly methods. Lastly, using our new pUN plasmid we replaced the BCP insert with a cassette expressing a codon-optimized \textit{Pyrococcus furiosus} (Pfu) DNA polymerase gene. This plasmid yields a lysate with high levels of the thermostable DNA polymerase now ‘functionalized’ to perform a high-fidelity \textit{ex vivo} PCR reaction. With \textit{ex vivo} methods for both DNA
assembly and PCR we report the creation of the first all-encompassing ex vivo Cloning pipeline.
Chapter 2. ex vivo DNA ASSEMBLY

Data and Figures reproduced with permission from:

Author Contributions:
George H. McArthur and Adam B. Fisher designed the experiments, which were carried out in the laboratory of Stephen S. Fong. Adam B. Fisher performed the experiments with Zachary B. Canfield and Laura C. Hayward. George H. McArthur, Adam B. Fisher, and Stephen S. Fong interpreted the data and wrote the manuscript. All authors discussed results and commented on the manuscript.

2.1. Initial 2-way ex vivo DNA Assemblies

To assess the ability of select cellular lysates to join together dsDNA, we first designed two amplicons with appropriate overlapping ends (26 and 30 bp overlaps) to be assembled into a circular plasmid (Figure 1.1A)\(^\text{34}\). Correctly assembled plasmids endow *E. coli* transformants with selective resistance to the antibiotic chloramphenicol and also visually screenable expression of a blue chromogenic protein (BCP) native to the coral *Acropora millepora* (Figure 1.1B)\(^\text{35}\). Template plasmids housing the *bcp* coding sequence (pSB1C3-K592009) and the appropriate antibiotic resistance and replication origin (pSB1C3-J04450) produce white and red/pink colonies, respectively, providing a convenient way to track transformation efficiencies (Figure 1.1B).

In this initial experiment, we used lysates of *S. cerevisiae*, *E. coli*, and *D. radiodurans* (hereafter *Sce*, *Eco*, and *Dra*, respectively) and a simple buffer containing ATP and MgCl\(_2\) to attempt ex vivo DNA assembly.
The two amplicons were incubated for 1 hour with each lysate. Samples of each reaction mixture were subsequently used to transform *E. coli* NEB10β.

Figure 2.1 The 2-way Circular Assembly

(A) 2-way assembly was demonstrated by joining a coding sequence for blue chromogenic protein (BCP; a 698 bp segment of pSB1C3-K592009 colored in dark blue) and the majority of pSB1C3-J04450 (a 2446 bp segment), thereby replacing the RFP coding sequence with BCP (pSB1C3-BCP). The pMB1 origin is colored in green and the chloramphenicol resistance gene is colored orange. (B) Correctly assembled plasmids allow transformants to express BCP (blue colonies) while colonies containing carryover template plasmids appear either red (pSB1C3-J04450) or white (pSB1C3-K592009).

Interestingly, Dra-incubated DNA did not produce transformants, although DNA incubated with Sce or Eco successfully transformed *E. coli*. Further analysis by agarose gel electrophoresis showed significant degradation of the individual amplicons, indicating that a highly active exonuclease system might have prevented the assembly of DNA by Dra (Appendix B Figure B.1). Even expected background transformants resulting from lingering circular PCR template were absent, suggesting that endonuclease activity is also high in Dra. Indeed, after further investigation we found a consensus DrdI site.
(GAC-N₆-GTC) between 1475 and 1486 (GACGCTCAAGTC) in the replication origin of our plasmids.

2.2. Testing the Magnesium:ATP Ratio with 2-way Linear Assemblies

We then explored ways to improve the efficiency of ex vivo DNA assembly with Eco and Sce by varying the relative composition of ATP and MgCl₂ in our buffer, adding NAD to the buffer (to power NAD+-dependent processes such as ligation), modulating the temperature at which the reaction mixture was incubated, increasing the amount of cellular lysate in our reaction, and increasing the duration of incubation. To directly assess the efficacy of ex vivo DNA assembly and avoid variation associated with transformation, we chose to visualize formation of a linear product (Figure 2.2A) from a pair of overlapping amplicons (28 bp overlap) via agarose gel electrophoresis of the ex vivo assembly reaction prior to transformation (Figure 2.2B). In this manner, we were able to compare different reaction conditions on their ability to join two overlapping dsDNA into one linear product.

Buffers were made of 1 mM DTT, 1 mM NAD, and varying concentrations of ATP and Mg²⁺. A wide range of ATP:Mg²⁺ ratios were initially tested with 1-hour Eco reactions at 37°C: 1:5, 1:10, 1:20, 5:5, 5:10, 5:20, 10:5, 10:10, and 10:20 (mM:mM). Only three ATP:Mg²⁺ ratios were chosen for further reaction optimization: 10:5, 5:5, and 1:10 (mM:mM) (buffers 1-
3 in Figure 2.2B, respectively). Linear products resulting from assembly by Sce were not at all visible in the gel, although bands of small pieces of DNA indicate that there is some nuclease activity under most conditions tested. On the other hand, Eco appears to have significant activity at 30°C and 37°C even without the addition of buffer. Previous studies have noted that the nucleolytic behavior of the RecBCD complex of E. coli changes based upon the relative amount of ATP to free Mg\(^{2+}\) \textit{in vitro} \textsuperscript{37}. We observed a similar trend in our lysates; the most efficient assembly reactions (for both 30°C and 37°C) were carried out under conditions of excess magnesium relative to ATP (Buffer 3 in Figure 2.2B), probably because ATP can chelate Mg\(^{2+}\) with high affinity under physiological conditions \textsuperscript{38}. 
Figure 2.2 The 2-way Linear Assembly and Buffer Optimization

(A) Optimized ex vivo reaction conditions were identified by visualizing the joining of a 697 bp segment of pSB1C3-K592009 (BCP) and a 381 bp segment of pSB1C3-J04450 (Promoter) into a larger linear DNA molecule. (B) Reactions were run for 1 and 2 h (left and right in each gel, respectively), at three temperatures and four buffer compositions for each Eco and Sce. All buffers contained 1 mM NAD and 1 mM DTT. ATP and Mg2+ concentrations (mM:mM) varied as follows: (1) 10:5, (2) 5:5, (3) 1:10. The control lanes include a reaction with no supplemented buffer (−) and a negative control ("ctrl") DNA-only lane (no buffer or lysate).

Each ex vivo reaction condition was tested for both 1- and 2-hour incubations. Initial experiments indicated that greater incubation times (3–6 hours) do not improve ex vivo assembly yields (data not shown). For all successful assemblies, 2-hour reactions appear to generate more linear product than 1-hour reactions. The temperatures selected for ex vivo reaction optimization reflect cell culture conditions (30°C for yeast and 37°C for E. coli) and the temperature used for ISO assembly reactions (50°C). The 50°C reactions are considered a negative control since we do not expect the cellular machinery in Eco or Sce to be thermostable, although transient activity may occur initially. Based on the results of these
experiments we selected a buffer composition of 1 mM ATP, 10 mM Mg^{2+}, 1 mM DTT, and 1 mM NAD+ (i.e., Buffer 3 in Figure 2.2B) incubated at 37°C for 2-hour as our optimized conditions for both Eco and Sce.

Figure 2.3 The 3-way Assembly and Colony Counts
(A) Three-way assembly was demonstrated by joining the BCP coding sequence (a 696 bp segment of pSB1C3-K592009), the majority of a plasmid carrying a different antibiotic resistance gene (a 2214 bp segment of pSB1K3-J04450) and a promoter-RBS region used to drive BCP expression (a 387 bp segment of pSB1C3-J04450), resulting in pSB1K3-BCP.
(B,C) The pMB1 replication origin is colored in green, the chloramphenicol resistance gene is colored orange and the kanamycin resistance gene is colored purple. Colony counts for optimized two- and three-way assembly experiments. The negative control assembly reaction is labeled “None” meaning that these colonies result from in vivo assembly in E. coli.

2.3. Optimized Assemblies with Yeast and E. coli Lysates

Under optimized reaction conditions, Eco and Sce were used again to perform the two-way dsDNA assembly. In addition, we designed three overlapping amplicons (30, 29, and 26 bp overlaps) to demonstrate a three-way assembly, which is not only more complex but also more useful for generating combinatorial libraries (Figure 2.3A). For each of these assembly tests, lysate-incubated DNA was allowed to react for 2-hours before transformation of E. coli. Control conditions of zero

<table>
<thead>
<tr>
<th>Lysate</th>
<th>Reaction time</th>
<th># Colonies</th>
<th># Blue</th>
<th># Red</th>
<th># White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eco</td>
<td>0 minutes</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eco</td>
<td>120 minutes</td>
<td>429</td>
<td>423</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Sce</td>
<td>0 minutes</td>
<td>54</td>
<td>51</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Sce</td>
<td>120 minutes</td>
<td>125</td>
<td>124</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>None</td>
<td>0 minutes</td>
<td>63</td>
<td>62</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lysate</th>
<th>Reaction time</th>
<th># Colonies</th>
<th># Blue</th>
<th># Red</th>
<th># White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eco</td>
<td>0 minutes</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eco</td>
<td>120 minutes</td>
<td>498</td>
<td>491</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Sce</td>
<td>0 minutes</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Sce</td>
<td>120 minutes</td>
<td>23</td>
<td>15</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>None</td>
<td>0 minutes</td>
<td>30</td>
<td>21</td>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>
incubation time and reactions with no lysate added were also run to ensure that the lysate was indeed facilitating DNA assembly and not otherwise affecting the transformation process.

Our results (Figure 2.3B and Figure 2.3C) clearly indicate that DNA is indeed being assembled ex vivo. For two- and three-way assemblies, both Eco and Sce, transformation efficiencies significantly increase when the DNA to be assembled is allowed to incubate with the lysate. Unexpectedly, the “no lysate” negative control revealed that for two- and three-way assemblies the overlapping amplicons can be joined together into a circular plasmid in vivo, suggesting that a significant fraction of the ex vivo assembly reactions – and also in vitro reactions (e.g., ISO assembly) – actually occur inside E. coli after transformation (i.e., transformation-associated recombination cloning in E. coli). Although Eco-mediated assembly produced many blue colonies for both two- and three-way assemblies (423 and 491 colonies, respectively), Eco-lysate appeared to be detrimental to the overall transformation efficiency as when transformation reactions were spiked with Eco-lysate (but not allowed to incubate with the DNA amplicons) the presence of colonies was greatly diminished (See, Table 2.1).

2.3.1. Effect of Detergents on Transformants

To test whether or not the detergent-based lysis buffer used to produce Eco was affecting the transformation process, we carried out the assembly reactions and transformations in the presence of the lysis buffer but without cellular lysate. As summarized in Table 2.1,
the bacterial lysis buffer inhibits transformation completely. Therefore, Eco-mediated assembly is likely much more efficient than we have observed and an alternative lysis method would probably increase overall transformation efficiency. The yeast lysis buffer does not appear to inhibit transformation of *E. coli*, which is not surprising because it is designed to lyse yeast cells. Interestingly, the number of colonies produced by DNA incubated with yeast lysis buffer was equivalent to the number of colonies produced by *Sce*-incubated DNA and that produced by the “no lysate” control. These data suggests that *Sce*-mediated DNA assembly was not observed but rather indicates that in vivo end joining in *E. coli* facilitates DNA assembly in this case. Interestingly, there would appear to be an increase in total colonies formed when the DNA is incubated with CelLytic Y. One might speculate that during incubation perhaps the buffer composition influences the DNA to adapt a more tightly supercoiled conformation as is seen in solutions with higher salt concentrations and is observed to greatly increase transformation efficiency\textsuperscript{39,40}. However, without further experimentation and knowledge of the proprietary CelLytic Y formula this remains purely speculative.
<table>
<thead>
<tr>
<th>Lysis Buffer</th>
<th>Reaction time</th>
<th># Colonies</th>
<th># Blue</th>
<th># Red</th>
<th># White</th>
</tr>
</thead>
<tbody>
<tr>
<td>CelLytic B</td>
<td>0 minutes</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CelLytic B</td>
<td>120 minutes</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CelLytic Y</td>
<td>0 minutes</td>
<td>42</td>
<td>41</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CelLytic Y</td>
<td>120 minutes</td>
<td>150</td>
<td>140</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

Tabulation of colony counts for transformations that were carried out in the presence of lysis buffer and the appropriate DNA but without cellular lysate. No transformants were seen when the DNA was incubated with the bacterial lysis buffer (CelLytic B), suggesting that this particular buffer somehow inhibits transformation. On the other hand, the yeast lysis buffer (CelLytic Y) showed no inhibition of transformation, which may not be surprising since the bacterial lysis buffer is designed to act on bacteria such as *E. coli* NEB10β. However, it is also clear from these results that Sce does not offer any advantage over the yeast lysis buffer for DNA assembly and, therefore, it can be deduced that the transformants resulting from Sce incubation are actually a result of *in vivo* DNA assembly in *E. coli*.

### 2.4 Conclusions and Discussion

Through a series of experiments investigating circular and linear DNA assembly, we found that: 1) lysate derived from *E. coli* NEB10β, a RecA-deficient strain, was able to efficiently assemble dsDNA, 2) a fraction of the assembly takes places inside *E. coli* NEB10β post-transformation by *in vivo* DNA assembly, and 3) lysates derived from *D.*
**Radiodurans** and *S. cerevisiae* were not able to join together dsDNA under the conditions we tested. In addition, it is worth noting that we attempted single-stranded DNA (ssDNA) assembly of synthetic 60-mer oligonucleotides (20 bp overlaps) that was accomplished by ISO assembly but was unsuccessful using our ex vivo approach (data not shown). Over the course of all of our experiments we found that the trends in efficiency of assembly held true across organisms independent of batch-to-batch variation.
Chapter 3. Optimization of ex vivo DNA Assembly

3.1. Lysate Preparation

Arguably, the most important step in any ex vivo investigation is the preparation of the lysate. Since this starting material may be used for many subsequent reactions and experiments it is critical to fully investigate the factors contributing to the suitability of resulting cell extracts prior to optimizing downstream reaction parameters. To investigate lysate preparation we tested the effects of varying media composition, harvesting at different points in the growth curve, the addition of cryopreservant (i.e., glycerol) and lysis method.

3.1.1. Harvest Phase

The typical growth curve of E. coli (and other prokaryotic bacteria) show an initial period of non-exponential growth (known as lag phase) in which the cells adjust their intracellular physiology in response to the change in extracellular conditions and the increase in available nutrients. This is followed by the exponential phase in which the cells reproduce at maximal capacity (given the external conditions) until the external environment becomes suboptimal (due to culture density, accumulation of waste/byproducts or depletion of nutrients). At this point the cells transition from exponential phase into a quiescent state known as stationary phase (SP). Reproduction is
halted and metabolism is severely slowed as the cell allocates resources to only the very most essential processes until new resources become available and the whole grow curve can restart.

![Graph showing fold difference for different growth stages](image)

**Figure 3.1 Harvested Growth Stage on Assembly Efficiency**

NEB10β was grown in TB and harvested at various points along the growth curve. Growth phase was judged by optical density at 600 nm, using absorbances of 0.4, 1.0 and >2.0 as markers of early exponential, late exponential and stationary phases, respectively. Lysates were prepared as previously reported using detergents and two-way assemblies performed. Fold difference represents the change in product band intensity of the sample as compared to the intensity of Stationary Phase product band. Band intensities are calculated relative to the BCP band of the negative control lane. Original gel picture are shown in Figure B.2 and Figure B.3.

Initial experimentation with ex vivo DNA assembly grew the *E. coli* cultures to saturation for two reasons. First, this is very convenient for the researcher as the *E. coli* culture can be prepared as an overnight culture and grown to saturation by the next day without the
need to constantly monitor the progress of the culture. Second, during the later stationary phase of growth the cell and genome are bombarded by accumulated waste products (free radicals, high pH)\textsuperscript{41,42}. The damage inflicted upon the cellular DNA may thereby cause the cell to produce DNA repair enzymes useful for DNA end joining. However, the highest levels of DNA replication occur during exponential phase and may in fact contain elevated concentrations of DNA repair enzymes useful to DNA assembly.

Based on previous observations of growth for NEB10β, we defined three different phases of growth, Early Exponential (EE), Late Exponential (LE) and Stationary phase (SP) represented by spectrophotometric measurements of optical density at 600 nm (OD600) of ~0.4, ~1.0 and >2.0, respectively. After harvesting the NEB10β grown to EE (OD600 = 0.46), LE (OD600 = 0.92) and SP (OD600 = 2.75) in previously detailed culturing conditions, lysates were prepared by resuspending in a 3:1 ratio of 2X CelLytic B (Sigma) detergent to pellet mass (e.g., 300 µL CelLytic for 100 mg of cell pellet), pelleting cell debris from lysis and diluting the soluble fraction 50% with glycerol.

The results of 2-way linear assemblies using these lysates displayed in Figure 3.1 shows an improvement in EE and LE harvesting over lysate prepared from SP. However, this improvement is very slight. Since harvesting in the latter stages of growth affords more biomass (and therefore lysate) and is convenient for the researcher we will continue to harvest in the stationary phases of growth.
3.1.2. Rich (TB) versus Minimal (M9)

Cellular physiology is inextricably linked to the extracellular environment. In a laboratory setting the extracellular environment is established by the physical (e.g., temperature and pressure) and chemical conditions of the growth medium. Under optimal conditions, *E. coli* will consume metabolites to increase biomass and replicate as quickly as possible. Conversely, under nutrient-limited or suboptimal conditions *E. coli* will reproduce at a more conservative rate as resources are reallocated to deal with environmental challenges. The rate of growth and extracellular conditions directly impact the production of enzymes involved in DNA reproduction and repair – including those putatively involved in the *ex vivo* DNA assembly process.
Figure 3.2 Rich (TB) versus Minimal (M9) media

Lysates prepared from cells grown in rich and minimal media also show minimal difference in assembly efficiency, although minimal media showed a slightly larger increase in efficiency. Here, late exponential cultures had to be used because cultures grown in M9 media had trouble exceeding optical densities of 2.0 used as the marker of stationary phase. Fold difference represents the change in product band intensity as compared to Stationary Phase. Band intensities are calculated relative to the BCP band of the negative control lane. Original gel pictures are shown in Figure B.2 and Figure B.3.

Previously, we grew our *E. coli* cultures to stationary phase (SP) in Terrific Broth (TB), a rich media very similar to the standard Luria-Bertani (LB) media but containing increased amounts of tryptone and yeast extract, glycerol as an added carbon source and a phosphate-based buffer system. This rich media allows growth to a much higher saturation than typical media such as LB. While we have used other rich media (LB, SOB, SOC; Data not shown) to prepare active lysates, here the activity of lysates prepared in TB were compared to those
prepared in M9 minimal media with glucose. As its name implies, a minimal media contains only the minimum nutrients necessary for bacterial growth, usually salts providing essential elements (Magnesium, Sulfur, Phosphate, etc...), a carbon source (e.g., glucose, sucrose, succinate) and water. In a minimal media environment nothing is “supplemented” to the cell so that all the essential metabolic pathways are active.

Given that the NEB10β lysate appeared to have slight improvements in efficiency when harvested in LE and that when grown in M9 media NEB10β had trouble achieving cellular densities similar to that of our stationary phase cells grown in TB (OD600>2.0), we compared cells grown M9 to LE with cells grown in TB to LE. As in 3.1.1, the results of the 2-way linear assembly are interpreted as fold difference to lysates prepared in TB at SP. The lysates prepared in M9 media at LE show comparability with that of cells grown in TB to LE.
### 3.1.3. Lysate Storage

**Table 3.1 Characteristics of Protein Storage Methods**

<table>
<thead>
<tr>
<th>Storage Condition</th>
<th>4°C</th>
<th>50% Glycerol Solution at -20°C</th>
<th>Frozen at -20°C or -80°C</th>
<th>Lysophilized (usually at -20°C or -80°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical Shelf Life</td>
<td>Days to several weeks</td>
<td>6 months to 2 years</td>
<td>Many years</td>
<td>Many years</td>
</tr>
<tr>
<td>Requires sterility or antimicrobial agent?</td>
<td>Yes</td>
<td>Usually</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Number of times sample may be removed from storage</td>
<td>Many</td>
<td>Many</td>
<td>Once</td>
<td>Once</td>
</tr>
</tbody>
</table>

Proteins are a highly heterogeneous class of biological macromolecules with activities that are extremely dependent upon both their chemical composition and tertiary structure. Some proteins are even dependent upon their association with one or more other peptides. As such, maintaining the stability and activity of proteins outside of their native context can be challenging. If certain conditions are not adhered proteins may lose activity as a result of aggregation, proteolysis/microbial contamination and exposure to suboptimal thermochemistry. While the optimal conditions for storage vary between proteins, Table 3.1 highlights several general guidelines for protein storage and stability.
Figure 3.3 Glycerol negatively affects lysate activity
From the same NEB10β lysates prepared at various growth stages, aliquots were saved prior to glycerol dilution and instead diluted with water. These are compared with their glycerol counterparts and show a significant improvement in assembly efficiency. Fold difference represents the change in product band intensity of the non-glycerol sample as compared to its glycerol-containing counterpart. Band intensities are calculated relative to the BCP band of the negative control lane. Original gel pictures are shown in Figure B.2 and Figure B.3.

For our purposes, once the soluble fraction of our crude lysate is isolated, we dilute with an equal volume of 100% glycerol and store the samples at -20°C. To evaluate the prudence of glycerol addition, Figure 3.3 shows the results of testing freshly prepared NEB10β lysates diluted with either an equal part glycerol or equal part water. Glycerol would appear to have a significant impact on assembly efficiency, decreasing the formation of linear product approximately four-fold.
While such a drastic effect on DNA end-joining might be attributed to enzymatic origins (reduction in exonuclease, polymerase or ligase activities), it is interesting to note that there does not seem to be a difference in the degradation bands formed with or without the presence of glycerol. This means that a roughly equivalent amount of substrate ends are formed by the action of exonucleases, but these compatible substrates are not incorporated into the final product band. Granted, it could be that while the total degradation is equivalent the exonuclease is functioning differently. For example, restriction enzymes are known to have reduced substrate specificity (so called, “star activity”) for particular DNA sites in the presence of glycerol concentrations exceeding 5%\textsuperscript{44}. It would appear that the glycerol might be interfering with the thermodynamics of base pairing between exposed homologous regions. This could make sense given that glycerol is able to decrease the hydration sphere around polynucleotides by competing with water to hydrogen bond with exposed base side groups and, perhaps for this reason, is used in electrophoretic buffers to help prevent DNA entanglement\textsuperscript{45}. Indeed, prior investigation into ligases have shown that the electrostatic reduction of hydrogen bonding interactions experienced with increasing glycerol decrease the activity of the ligase but actually lead to an increase in the accuracy/specificity of the ends ligated together\textsuperscript{46}.

Although omitting the inclusion of glycerol is tempting due to it’s clear detrimental effect on lysate activity, it’s faculty as a cryopreservant provides convenience that – in our opinion – supersedes
this disadvantage. For instance, if glycerol was to be omitted from the lysate there are a few alternatives. Lysate could be aliquoted into tubes, frozen and only thawed immediately prior to use. However, this may be very wasteful as only 2 µL of lysate is needed per reaction so there is either a waste of freezer space as small aliquots are frozen in individual tubes or a waste of lysate as leftover thawed lysate is discarded. Lysates could also be freshly prepared before all assembly reactions; however, this is laborious and requires a culture of E. coli to be prepared ahead of time. For these reasons we have decided to continue to utilize glycerol as a cryopreservant agent and maximize the efficiency of lysis in order to have a very concentrated lysate. This way the lysate may be diluted prior to the reaction and thereby the final concentration of glycerol is diluted, as well.

3.1.4. Detergents versus Autolysis

In the previous section we noted that ultimately, it is desirable to have a concentrated lysate stock, which can be diluted prior to the assembly reaction. In the previous context it was in reference to reducing glycerol contamination, but concentrated protein samples are also more resistant to deactivation and loss due to interactions with the storage vessel. Generating a highly concentrated cell extract is predicated upon an efficient lysis method.

Our current approach relying upon a chemical detergent means of lysis is efficient, compatible with small sample volumes and requires no specialized equipment. The detergents are non-toxic, non-denaturing and available as 1X, 2X and 10X concentrated formulations, so an
increased concentration can be employed to make even more highly concentrated cell extracts. However, the detergents suffer from several major drawbacks including the expense of purchase, a lack of transparency to their proprietary formulation and most of all, extensive toxicity toward the fragile competent *E. coli* utilized for transformation (Table 2.1). To address this major dilemma, we decided to investigate alternative methods to lyse our *E. coli* cultures.

While a number of methods are commonly employed to lyse *E. coli* most are unsuitable. Thermochemical methods such as alkaline lysis and heat-lysis denature the necessary catalytic proteins and present conditions that may handicap downstream processes. On the other hand, physical methods of lysis using a bead basher, French press homogenizer or sonicator offer some of the highest lysis efficiencies with the greatest degree of flexibility in buffer composition, but may have minimum sample volumes, require vigilance to prevent denaturing of the sample and all require the purchase of specialized equipment.

In response to the pitfalls inherent in each of the previous lysis methods, Zymo Research™ has created autolytic strains of *E. coli*. These strains contain an ingenious genetic design drawing upon the lysogenic system of the *E. coli λ* bacteriophage. Briefly, during the lytic phase of the *λ* bacteriophage life cycle the bacteriophage expresses two proteins. One protein (protein ‘R’) is an endolysin/lysozyme capable of cleaving the peptidoglycan bounds found in the cell wall. The other protein is a holin (protein ‘S’) that - as its name implies - punches holes in the cell membrane. Not only
does this disrupt the osmotic balance maintained by the selective permeability of the cell membrane but also releases the endolysin ‘R’ protein into the extracellular space where it can act upon the cell wall. In the XJa design the autolysis cassette only contains the endolysin ‘R’ gene that is driven by the araB promoter from the arabinose degradation operon. By adding arabinose (at ~0.2%) to the growing culture, the endolysin is produced at high level within the intracellular milieu where it is prevent by the cell membrane from acting upon the cell wall. Once ready, the researcher can freeze and thaw the culture (resuspended in buffer of choice) in order to lyse the culture. During the freezing of the cell, intracellular water crystallization damages the cell membrane performing the functions typical of the holin and releasing the endolysin to degrade the cell wall and lyse the cell.

This autolysis cassette presents a unique approach to create a lysate in the buffer of choice, without expensive reagents or equipment and in a highly efficient fashion (Zymo reports lysis efficiencies >90% after one freeze-thaw cycle for the XJa strain)\textsuperscript{49}. As such, we purchased the XJa strain that is derived from the ubiquitous cloning strain JM109. While this strain is genotypically different from NEB10β, from a phenotypic perspective as it relates to DNA, they are very similar. Like most cloning strains, JM109 (XJa) and NEB10β have mutations in endA and recA disrupting their endoplasmic DNA-specific endonuclease and homologous-recombination functions, respectively\textsuperscript{50,51}. The endA mutation is important for protecting the integrity of any
substrate DNA present in the lysate. In its natural context the in 
vivo secretion of EndA prevents the degradation of host episomal or 
genomic DNA, but once the membrane is lysed and removed in an ex vivo 
or in vitro context, any substrate DNA existing or subsequently added, 
assembled or circularized is an active substrate to be cleaved\textsuperscript{52,53}. The 
recA mutation disrupts the cell’s ability to perform homologous 
recombination between repetitive sequences. This way episomal 
(plasmid) DNA is maintained as intended rather than recombined with 
themselves, other plasmids or the genome\textsuperscript{54,55}. In addition they both 
have their restriction systems removed. Strain genotypes are displayed 
in Table A.2 with an explanation in Table A.3.

To test the XJa autolysis for DNA assembly, we performed the growth 
experiments used for NEB10\beta and found that XJa performed similarly in 
2-way linear experiments to NEB10\beta (Figure B.2 and Figure B.3). Based 
on the results of the lysate negative controls from Table 2.1 the 
effects of autolysis lysates on electroporation were studied first.
Using two 3-way assemblies, one performed using a detergent-derived XJa lysate and one using an autolysis-derived XJa lysate, the resulting assemblies were transformed into in-house prepared chemicompetent and electrocompetent NEB10β. The results of these transformations, as well as transformations of pUC19 control DNA, are depicted in Figure 3.4. The first thing to note about this experiment is that the concentration of lysate in the transformation reactions is significantly lower than those performed in Table 2.1 previously and so detergent associated toxicity effects should be expected to be less pronounced in turn. Still, there is clearly a dramatic difference between electroporation efficiencies of lysates. Contrasting the
chemical transformations in Figure 3.4.A, the autolysis reaction appears to have been more efficient than the detergent reaction (either in actual assembly or detrimental effects on transformation) but the overall colony forming units (CFU) per femtomole appear to be comparable. However, when looking at the electroporation results there is a clear effect of detergents which is more succinctly stated in Figure 3.4.B as fold difference between chemical transformation and electroporation. It is also interesting that autolysis seems to have a more pronounced difference in electro- versus chemical transformation than that of the pUC19 positive control however, the pUC19 controls may underestimate the fold improvement of electroporation. Typical efficiencies reported by NEB for their commercially prepared NEB10β chemicompetent and electrocompetent cells would place the improvement at roughly ten-fold, matching the observed improvement seen for autolysis transformation.

3.2. Optimizing the ex vivo DNA assembly reaction

The ‘chew-back, anneal, repair’ end joining mechanism hypothesized to be the biochemical basis of the ex vivo reaction is an interestingly complex optimization problem. While exonuclease activity is imperative, overly active exonuclease activity can destroy the substrate DNA. Additionally, since dsDNA exonuclease activity cannot be constrained a competitive polymerization reaction must be introduced to ‘repair’ uncontrolled exonuclease activity, but not so much as to undo the requisite ‘chew-back’. Several parameters of the
ex vivo reaction were empirically investigated for their impact on this delicate balance.

3.2.1. Reaction Volumes and DNA Concentrations

In its original formulation, ex vivo DNA assembly was performed using 20 µL volumes with 20 ng/µL of the backbone (or larger BCP piece for linear assemblies) and a 6:1 molar ratio of insert fragments to backbone piece. While PCR and gel extraction easily generates enough substrate DNA for a couple reactions, each assembly reaction is consuming over a full microgram of DNA per reaction. Additionally - while able to distinguish background colonies by their red coloration, there was an exceptional amount of red colony formation for gel-purified substrates. To address these issues we made a concerted effort to not only reduce the DNA required for an ex vivo reaction and the amount of red background colonies.

First, the background was addressed by looking at our DNA preparation steps. The bands from our PCR reactions looked very strong and free of side-product formation yet somehow considerable template plasmid was co-purified with our backbone amplicon. Looking at our protocol for PCR we noted that for simplicity we had been adding 1 µL of pure plasmid mini-prep as template for a 50 µL reaction. Concentrations of our mini-preps usually range between 50-100 ng/µL for the high-copy plasmids used by ex vivo DNA assembly. Given that NEB protocols for Q5 DNA polymerase call for a maximum of 1 ng of plasmid template per 50 µL reaction we were adding 50 to over 100 fold excess template DNA necessary for PCR. In response, working dilutions of template plasmid...
at 1 ng/µL were prepared and PCR performed in the same fashion. Resultant gels still showed very bright banding indicative of successful PCR and eluents from gel extractions contained equivalent concentrations of DNA. However, assemblies and subsequent showed a drastic decline in the number of red colonies (~0-2 colonies per plate).

Next, the concentration of DNA used in a typical reaction was decreased along with the total volume of each reaction. Instead of a 20 µL volume, 20 ng/µL backbone and a 6:1 insert to backbone molar ratio, reactions were performed in 10 µL volumes using 5 ng/µL backbone (50 ng per reaction) and a 1.2:1 (2-way assemblies) or a 2:1.2:1 (3-way assemblies) molar ratio. This increased the number of reactions capable of being performed per preparation of DNA fragments by 4-fold (50 versus 200 ng backbone per reaction).

3.2.2. Calculating Protein Concentration

By reducing the DNA substrates used in each reaction and increasing our lysate concentrations (lysis efficiencies), we found many of initial reactions degraded most or all of the substrate DNA. The reduced substrate and increased catalyst concentrations had tipped the delicate balance between chew-back and repair making it clear adjustments to the lysate concentration were needed. Unfortunately, quantifying the “concentration” of the catalysts for ex vivo DNA assembly poses a challenge. First, the cellular extract is a highly heterogeneous environment with many species capable of interfering with quantification. Second, even if the responsible catalysts were
capable of being specifically quantified their identities still remain largely unknown. Even so, measures of total protein concentration were employed as a proxy since the catalysts are almost certainly proteins.

The most common techniques for estimating the concentration of proteins are based on either a colorimetric reaction (e.g., Lowry\textsuperscript{56}, Bradford\textsuperscript{57}) or UV absorbance spectrophotometry (absorbance at 280 nm)\textsuperscript{58}. Deciding on the most appropriate method depends upon multiple factors including the concentration of protein in the sample, the amino acid composition of the peptide and the presence of interfering substances in the sample or buffer. For the ex vivo extracts Bradford colorimetric assays were compared with simple spectrophotometric absorbances of 280 nm for aromatic rings of amino acid side chains (mainly tyrosine and tryptophan), 260 nm for nucleic acids and/or 205 nm for peptide bonds.

Standard curves were generated for Bradford colorimetric assays using either Bovine serum albumin or lysozyme, at a minimum of three different quantities (along with blanks) in either water, the CelLytic detergent or the lysis/storage buffer matching the lysate samples used in the assay. While the Bradford reaction is typically incompatible with detergents, we found that the diluted samples still generated an acceptable standard curve. However, the concentrations of the lysates predicted by the Bradford (as could be expected) was dependent upon the protein used to calibrate the standard curve, the time allowed for color development and even the spectrophotometer used. Additionally,
the assay requires the purchase of reagents and standards and a fairly
time-consuming, fickle protocol to measure samples and standards.

The difficulties posed by the Bradford assay prompted a simplified
approach to concentration estimation. Peptide bonds can be detected in
solution using absorbance at 205 nm, but when tested with lysate the
assay was found to be overly sensitive, inviting error from extensive
dilutions. Protein content can also be estimated using absorbance at
280 nm targeting the aromatic side chains of amino acids such as
tryptophan and tyrosine. The challenge faced in a cellular extract
context is interference from nucleic acids that also absorb in the 280
nm spectra. To correct for the presence of nucleic acids in the lysate
absorbance at 260 nm was measured and a correction applied using the
equation of Warburg and Christian. While Warburg-Christian method is
also susceptible to interference (e.g., NAD/NADP) it was found to be
rapid, requiring minimal sample manipulation and having high
reproducibility (for one sample diluted and measured in triplicate
standard error of the mean = 0.012 mg/mL).
Figure 3.5 Concentration of Lysate

Here XJa lysate prepared by autolysis was measured via spectrophotometry and adjusted by the method of Christian-Warburg. Product Band intensities are calculated relative to the BCP band of the negative control lane. Gel picture for 60 minutes is in Figure B.4 and 120 minutes is in Figure B.5.

Applying the Warburg-Christian method to quantify our lysates a 2-way linear assembly was performed using titrations of autolysis-derived XJa lysate. Samples were removed at 60 minutes and after a full 120 minutes, visualized and interpreted in Figure 3.5 as the intensity of the product band. While in the first 60 minutes there is a product band formation at all tested concentrations, the second 60 minutes distinctly contribute a large portion of the final bands. This could be evidence that most of the “repair” activities do in fact precede in the second half of the 120-minute reaction. However, at higher concentrations of lysate the product bands produced in the initial 60 minutes are consumed, likely by the additional exonucleases present in
the reaction. Looking at the results of the gel it would appear that the optimal final protein concentration of the autolysis lysates is between 10-50 µg/mL. Since our stocks of lysates produced via autolysis were measured at ~7-8 mg/mL, a 10X concentration could be reached by a 1/200 dilution of the lysate.

3.2.3. “Base” Buffer Optimization

In order to function many enzymes rely on the presence of cofactors and various energy sources. In section 2.2 the effects of Magnesium-to-ATP concentration were investigated and the results revealed an important relationship between the relative quantities of magnesium ions and available ATP. Beyond these two components of the reaction buffer, the other components were not varied but may illicit improvements in the overall reaction efficiency.

We tested the formulation of the ex vivo DNA assembly buffer with a few changes to more closely match buffer formulations optimal for purified *E. coli* enzymes, in particular for the *E. coli* DNA ligase (Eco Ligase). For all buffers tested the NAD+ was reduced to 50 µM (typically 26 µM for Eco Ligase) and Bovine Serum Albumin (BSA) at 50 µg/mL. Further the Magnesium and ATP levels were, once again, titrated between 4 mM Magnesium and 50 µM ATP.
Tris-HCl, 50 mM | 50 mM | 50 mM | 50 mM | 50 mM
DTT | 1 mM | 1 mM | 1 mM | 1 mM | 1 mM
MgCl$_2$ | 10 mM | 10 mM | 4 mM | 4 mM | 10 mM
ATP | 1 mM | 1 mM | 1 mM | 50 µM | 50 µM
β-NAD | 1 mM | 50 µM | 50 µM | 50 µM | 50 µM
BSA | 0 µg/mL | 50 µg/mL | 50 µg/mL | 50 µg/mL | 50 µg/mL

| + | 1 | 2 | 3 | 4 |

The components of the base buffer were varied between their original values and values commonly used for the buffers of various molecular biology enzymatic reactions. Fold difference represents the change in product band intensity of the sample as compared to “+” sample of the same lysate type. Band intensities are calculated relative to the BCP band of the negative control lane. Original gel pictures are shown in Figure B.6 and Figure B.7.

The NEB10β prepared by detergent lysis showed decreasing efficiencies with the decrease in NAD+ and addition of BSA when compared to the standard buffer. In contrast, the autolysis XJa showed a slight
increase in assembly efficiency with the addition of BSA and decrease in NAD+. In fact, XJa showed an increase in all tested conditions over the standard buffer. Unfortunately, without additional replicates it is unclear whether the observed trend is due to differences in the XJa autolysis lysate versus the NEB10β detergent lysate or a function of fluctuations in the standard lane.

3.2.4. Divalent Metals Titrations

Throughout biochemistry metal ions play an integral role as enzymatic cofactors. In fact, it has been estimated that metal ion cofactors are needed in nearly one-third of enzyme-catalyzed reactions\textsuperscript{60}. Most often these metals are coordinated within the active site of enzymes where they mediate redox (and sometimes non-redox) reactions by increasing the electrophilicity of the active site. Once a ligand is bound in the active site, the metal can act to remove electron density and cause polarization of reactive bonds in the ligand - usually C-O or P-O in nature\textsuperscript{61}.

As the most abundant divalent metal species in cellular contexts, magnesium is the most frequently utilized metal in active sites. As discussed previously, in physiologic environments magnesium forms stable chelates with molecules containing phosphate such as ATP, RNA, DNA and other nucleic acids. The tendency of magnesium to maintain elevated hydration states is the primary means by which it acts as a cofactor to bind and catalyze nucleic acid substrates\textsuperscript{61}.
Due to its poor spectroscopic and crystallographic properties, researchers have long employed other divalent metal analogues to study nucleic acid enzymology. While analogues such as zinc, calcium and most often manganese are acceptable for structural studies, these metals often display differing biochemistries to magnesium\(^{62}\). In fact, it is often observed that other transition metals, especially manganese, coordinate more tightly with active sites and thereby confer higher levels of activity. In other instances, the alternate metal can change the coordination geometry or structural conformation of the active site altering the substrate specificity, mode of action or completely inhibiting the protein.
Figure 3.7 Titrations of Divalent Metal Cations

This figure shows the titration of Zinc Chloride, Manganese Chloride and Calcium Chloride with Magnesium Chloride and without Magnesium Chloride. Fold difference represents the change in product band intensity of the sample as compared to “+” sample (10 mM Magnesium) of the same lysate type. Band intensities are calculated relative to the BCP band of the negative control lane. Original gel pictures are shown in Figure B.8, Figure B.9 and Figure B.10. NEB10β is the average of duplicate experiments and error bars represent standard error of the mean (N=2).

Previously, we established that the ratio of Magnesium-to-ATP in the ex vivo DNA assembly buffer was crucial for efficient end joining, most likely as a consequence of ATP’s ability to chelate magnesium and the resulting effects on RecBCD activity. Although magnesium dominates as the native cofactor for nucleic acid-directed enzymes, we decided to test the effects of titrating zinc, manganese and calcium on the efficiency of end joining.

Using XJa (autolysis) and NEB10β (detergent) four formulations of each metal were tested with respect to magnesium (mM:mM): 1:10, 10:10, 1:0,
10:0. The results of this 2-way linear titration depicted in Figure 3.7 are compared to a control reaction with the ex vivo DNA assembly “base” buffer. Just as in the “base” buffer optimization in 3.2.3 there are clearly differing trends seen for the NEB10β/detergent and the XJa/autolysis lysates - although there is also some agreement between the two, as well. Both lysates responded negatively to Zinc in the complete absence of magnesium, while both responded positively to the addition of manganese - even in the complete absence of magnesium. XJa actually shows an increase in efficiency across all tested conditions so long as 10 mM of magnesium was present. Perhaps more surprising is that for both manganese and calcium XJa showed equivalent or higher activity even in the absence of magnesium. On the other hand, outside of manganese NEB10β clearly did not tolerate the omission of magnesium. The extremely pronounced response to manganese by both lysates prompted further investigation into the supplementation of manganese to the standard ex vivo DNA assembly buffer.

3.2.5. Manganese

The strong activation of DNA assembly imparted by manganese in the titrations of the previous section (3.2.4) prompted a more detailed study of manganese. Using a similar titration approach between magnesium and manganese, 2-way linear assemblies were performed using NEB10β lysate. Again, Figure 3.8 shows a comparison of the titrated samples relative to the intensity of the product band formed by the basic ex vivo DNA assembly buffer containing 10 mM magnesium and 0 mM
manganese. In the experiment, total divalent species were held at 10-11 mM with the exception of the 0:2 and 0:6 magnesium:manganese (mM:mM) samples. In Figure 3.9, the results are displayed again, but showing the compositions of the lane by band (product band; BCP and Promoter substrate bands and; degradation band of lowest molecular weight) relative to the BCP band of the negative control lane (-).

![Titrations of Magnesium and Manganese](image)

**Figure 3.8 Titrations of Magnesium and Manganese**

Using NEB10β lysate and the 2WL assembly, magnesium and manganese levels were titrated against one another. The ratio labels indicate the concentration of magnesium chloride (mM) to manganese chloride (mM) with the 10:0 ratio representing the original buffer constitution (+). Fold difference represents the change in product band intensity of the sample as compared to “+” sample of the same lysate type. Band intensities are calculated relative to the BCP band of the negative control lane. Original gel picture is shown in Figure B.11.
Figure 3.9 Lane composition of Manganese to Magnesium Titration

Chart shows the relative intensities of the product (Blue), BCP (Red) and Promoter (Green) substrate bands, as well as the low molecular weight band representing degradation products (Purple). Columns are displayed inverted to more closely represent the actual gel and the order in which the bands migrate. Band intensities are calculated relative to the BCP band of the negative control lane. Original gel pictures are shown in Figure B.11.

The results of the 2-way assembly - albeit less pronounced - confirm the observed increase in assembly efficiencies previously observed in the metal titration experiment. The ratio of magnesium to manganese appeared to have minimal impact. In fact, even the 0:10 sample was able to catalyze assembly with at least equal efficiency to that of the positive control. Visualizing the lane compositions in Figure 3.9 provides more details into the possible mechanisms by which manganese may be enhancing assembly. For instance, the largest consumption of substrate DNA did not directly correlate with product formation as there was more remaining substrate DNA in the higher manganese ratio.
lanes containing the highest level of product. Further, both lanes containing 10 mM manganese (1:10 and 0:10) appear to have the most total DNA in the lane overall. This phenomenon is more than likely attributed to increased DNA polymerase activity in the presence of manganese. Indeed, the literature is ripe with evidence showing that in the presence of manganese DNA polymerases (including the E. coli PolA polymerase) experience a decrease in specificity of base incorporation as well as 3’-5’ proofreading functions. Such a decrease in specificity may increase the polymerase activities not only to gap-fill the product fragment, but also the chew-backed, partially degraded substrate bands thereby preserving them.
Figure 3.10 Effect of Manganese on Two-way Circular Assembly
Using XJa lysed with either CelLytic or using freeze-thaw autolysis, circular assemblies were performed in the original buffer (-Mn) or with the addition of 1 mM Manganese Chloride (10:1 ratio). Percentages represent Blue colonies to all colonies (blue and white). Error bars represent standard error of the mean for reactions performed in duplicate (N=2).

If the presence of the manganese is mutagenizing the recombined product then its addition - regardless of the improvements in apparent end joining - are undesirable for cloning. However, the 2-way linear assay has no way of revealing incorporated mutations. To reveal any increases in mutation rates as a result of manganese supplementation shows the results of 2-way circular assemblies with and without 1 mM manganese. The results still show an increase in total colonies formed across all lysates, but an increase in mutation rate is not immediately apparent. This does not necessarily mean there is not an increase in mutation rate from the addition of magnesium, because the number of white colonies falls outside the countable threshold (<30 CFU per plate) for XJa autolysis and NEB10β. For the XJa detergent
sample that did have a countable number of white colonies there was a decrease in percentage of correct colonies.
Chapter 4. Probing ex vivo DNA Assembly with Enzymatic Titrations

The absence of the cell membrane allowed rapid prototyping of buffer compositions, substrate and catalyst concentrations, and titrations of small molecules. To uncover biochemical mechanisms typically enzymes are purified and titrated together in an in vitro study or knocked-out/overexpressed in an in vivo study. Both of these approaches can be leveraged in ex vivo experimentation to uncover underlying phenomena of the larger system.

4.1. Enzymatic Titrations

Tapping into the wealth of commercially purified cloning enzymes already available through molecular biology suppliers, the underlying biochemistry of ex vivo DNA assembly was probed using enzymatic titrations. Building on the premise that the ex vivo DNA assembly reaction proceeds via a chew-back, anneal and repair pathway similar to Gibson assembly, we purchased available exonucleases, polymerases and ligases (as well as RecA, a homologous recombination protein) native to E. coli. Table 3.1 gives details about each enzyme used in the titration experiment.
<table>
<thead>
<tr>
<th>Enzymes used in Titration</th>
<th>E. coli gene</th>
<th>Function</th>
<th>Product(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exonuclease III</td>
<td>xthA</td>
<td>3' to 5' dsDNA exonuclease</td>
<td>ssDNA and dNMP</td>
</tr>
<tr>
<td>RecJf</td>
<td>recJ</td>
<td>5' to 3' ssDNA exonuclease</td>
<td>dNMP</td>
</tr>
<tr>
<td>Exonuclease I</td>
<td>xonA/sbcB</td>
<td>3' to 5' ssDNA exonuclease</td>
<td>dNMP and dinucleotide</td>
</tr>
<tr>
<td>Exonuclease T</td>
<td>exoT</td>
<td>3' to 5' ssDNA and RNA exonuclease</td>
<td>dNMP</td>
</tr>
<tr>
<td>Exonuclease V (RecBCD)</td>
<td>recBCD</td>
<td>dsDNA processive endonuclease</td>
<td>short oligos</td>
</tr>
<tr>
<td>E. coli DNA Ligase</td>
<td>ligA</td>
<td>DNA ligase</td>
<td>phosphodiester bonded DNA</td>
</tr>
<tr>
<td>DNA PolI</td>
<td>polA</td>
<td>DNA polymerase</td>
<td>DNA polymer</td>
</tr>
<tr>
<td>RecA</td>
<td>recA</td>
<td>ssDNA binding protein</td>
<td>N/A</td>
</tr>
</tbody>
</table>

The addition of exonucleases can help elucidate the possible overhangs generated from chew-back. For example, if homologous overhangs were generated as 5' overhangs, it would be expected that supplementing RecJf would greatly decrease assembly efficiencies, as the homologous regions are substrates for degradation by RecJ. The three non-nucleases - E. coli Ligase, Polymerase I and RecA - are supplemented to see if they address a bottleneck or inefficiency.
Figure 4.1 Titrations of Various DNA cloning Enzymes
The results of titrations of various purified enzymes in XJa lysates are shown. The bar colors reflect concentration of enzyme with blue, red and green reflecting the concentrations lowest to highest. Red represents units necessary to catalyze the formation of 10 nmol of product in a 20 µL reaction except for RecA which had no unit definition. Unit definitions are available in Table A.4. Units were added as follows: RecBCD (.05, .2, .4); PolI (.1, .4, 1); RecA (2, 8, 20 µg); Eco Ligase (.1, .4, 1); ExoI (.1, .4, .8); ExoIII (1, 4, 10); ExoT (5, 20, 50); RecJ (30, 120, 300). Fold difference represents the change in product band intensity of the sample as compared to “+” sample (no enzyme). Band intensities are calculated relative to the BCP band of the negative control lane. Original gel pictures are shown in Figure B.12 and Figure B.13.

To determine relevant concentrations for each enzyme, the unit definitions provided by NEB were employed to calculate the number of units necessary to catalyze the formation of 10 nmol of product in a 20 µL reaction volume (5 ng/µL BCP is approximately 11 nmol). These values were used as a baseline to select three concentrations of each enzyme to test in conjunction with a 2-way linear assembly.
Results of the assembly shown in Figure 4.1 were unexpected for several reasons. The most glaring outcome is the stimulating effect of ExoI and ExoT on assembly. It was not expected that any of the exonucleases would improve assembly as this has not presented as a rate-limiting step and can be a substrate sink if homologous ends are consumed. Clearly, for the 3'→5' ssDNA exonucleases ExoI and ExoT (at low quantities) the homologous overhangs are not being consumed. This is made clearer by Figure 4.2, which shows the lane composition of the titration. By looking at the lane composition and status of the substrate DNA, the activities of the various enzymes become clearer. For instance, it is very clear that the lack of product formation in the dsDNA exonuclease lanes is a direct result of consuming the substrate DNA. In contrast, RecJ appears inefficient at assembly not because it is consuming the substrate DNA, but for some other reason. It is very possible that the homologous overhangs produced during the ex vivo DNA assembly are 5' extensions. These extensions would be compatible substrates for degradation by RecJ resulting in a loss of the necessary homology and the creation of inert substrate DNA. However, it is also important to realize that these graphs are still numerical interpretations of the gel results. In the case of ExoT, without looking at the gel picture (Figure B.13) it would be nearly impossible to observe the decreasing resolution and molecular weight of the product band due to extensive degradation.
Figure 4.2 Lane Composition of Enzymatic Titration

Chart shows the relative intensities of the product (Blue), BCP (Red) and Promoter (Green) substrate bands, as well as the low molecular weight band representing degradation products (Purple). Columns are displayed inverted to more closely represent the actual gel and the order in which the bands migrate. Band intensities are calculated relative to the BCP band of the negative control lane. Original gel pictures are shown in Figure B.12 and Figure B.13.

In the non-exonuclease category, improvements were observed using E. coli DNA ligase and RecA while addition of the DNA polymerase I inhibited assembly. Looking at the composition of the PolI lane the substrate fragments are maintained. The PolI protein has dual exonuclease activities: a 3′→5′ proofreading activity typical of DNA polymerases and a 5′→3′ exonuclease activity most likely used for nick translation. Similar to RecJ, the presence of substrate DNA and known 5′→3′ activities suggests that PolI may be consuming the homology of the overhangs.
4.2. RecA Titrations

Improvements in end joining in the presence of RecA was surprising as previous investigation revealed that recA+ strains of E. coli are suboptimal for ex vivo DNA assembly. RecA is a highly multifunctional protein in vivo, acting as an ssDNA binding protein\textsuperscript{63}, catalyzing strand exchange\textsuperscript{64} and even acting as a highly specific protease (of the LexA protein) to signal the host SOS response to DNA damage\textsuperscript{65,66}.

Mechanistically, RecA performs its DNA recombination functions by binding tightly to ssDNA in long nucleoprotein filaments. Once bound to a ssDNA substrate (and ATP) RecA’s secondary DNA binding site will bind to a dsDNA polymer and begin a “homology search” scanning across the dsDNA substrate for sequence matching the ssDNA. Once found, the RecA nucleoprotein filament can catalyze strand exchange and the formation of an intermediary triplex DNA\textsuperscript{64}. Given the unique contradiction seen between in vivo produced RecA and ex vivo supplemented RecA further titration experiments were performed. Using the 2-way linear assembly, Figure 4.3 and Figure 4.4 show the results of RecA supplementation in a 20 μL reaction using XJa autolysis lysate.
Figure 4.3 Titrations of RecA

RecA protein was added to an XJα lysate in a 2-way linear assembly. Fold difference represents the change in product band intensity of the sample as compared to “+” sample (no enzyme). Band intensities are calculated relative to the BCP band of the negative control lane. Original gel picture is shown in Figure B.14.

The repeated RecA titration showed a similar trend to that observed in the initial titration as additional RecA facilitates end joining in the two-way linear assembly. Although, it would appear that at a certain point the concentration of RecA becomes saturating – possibly even reducing assembly efficiency (as in Figure 4.1). The normal function of RecA to recombine ssDNA with a dsDNA substrate is not the same single-strand annealing method proposed earlier, but is clearly aiding in assembly. Again, looking at the lane compositions in Figure 4.4 it would appear that RecA helps to protect the substrate DNA from degradation as the total band intensity of the RecA lanes exceed that of the positive control lane.
Figure 4.4 Lane composition of RecA titration

Chart shows the relative intensities of the product (Blue), BCP (Red) and Promoter (Green) substrate bands, as well as the low molecular weight band representing degradation products (Purple). Columns are displayed inverted to more closely represent the actual gel and the order in which the bands migrate. Band intensities are calculated relative to the BCP band of the negative control lane. Original gel pictures are shown in Figure B.12 and Figure B.13.

Finally, RecA was supplemented to 2-way circular assemblies, transformed and the resulting colony counts compiled into Figure 4.5. It would appear that there is another dynamic happening here – either during circularization or transformation – which decreases the efficiency of assembly for the circular assembly. Moreover, with “homology search” functions of RecA it was disappointing to find that RecA contributed nothing toward improving the mutation rate of assembly either.
Figure 4.5 Colony counts of exDA reactions containing RecA
Assemblies were performed using basic buffer or supplemented 200 ng RecA using autolysed XJa or XJa lysed with detergents. Percentages represent Blue colonies to all colonies (blue and white). Error bars represent standard error of the mean for reactions performed in duplicate (N=2).

4.3. Strain Engineering (nuc4- and autolysis cassette)

Just as previously purified enzymes can be supplemented to the ex vivo reaction to synthetically test the effects of increased protein concentration, the genetic background of the strain can be altered to knockout the expression of native enzymes. Strain engineering or ‘recombineering’ (short for recombination-mediated genetic engineering) uses transformed DNA with homology to the host genome in order to target and integrate changes into the chromosome\(^67\). In E. coli this can be accomplished by the Lambda Red operon of bacteriophage
lambda\textsuperscript{68}. This three-protein operon encodes a 5’-to-3’ exonuclease (alpha), a single-strand annealing protein (beta) and a protein, which binds and inhibits RecBCD activity (gamma). For both the native bacteriophage’s linear dsDNA genome or transformed ds/ssDNA the Lambda Red system works by inhibiting RecBCD (preventing degradation of substrate DNA), converting the dsDNA to a fully single-stranded intermediate and annealing the subsequent ssDNA to homologous regions of the genome exposed as ssDNA during DNA replication\textsuperscript{69,70}. Genomic changes can be selected if the targeting DNA integrates a selectable marker cassette (i.e., antibiotic resistance), otherwise colonies must be plated and screened via a high-throughput genotyping method like allele-specific PCR\textsuperscript{71,72}.

Given the promising applications of genome engineering considerable effort has been invested in optimizing Lambda Red and other methodologies (e.g., homologous recombination, RecET, CRISPR-Cas)\textsuperscript{15,67,73-76}. The Church group pioneered a Lambda Red recombiner method to make many changes across the genome in an automated fashion known as ‘Multiplex Automated Genome Engineering’ (MAGE) using a closed-loop device to grow/recover cells, induce Lambda Red from a heat-inducible operon on the E. coli genome, and transform a solution of editing oligonucleotides\textsuperscript{77-79}. In an effort to optimize MAGE this same group employed MAGE to knockout the most active E. coli ssDNA exonucleases (RecJ, ExoI, ExoVII and ExoX) creating a strain known as nuc4- (Genotype in Table A.2 and Table A.3), which they subsequently
deposited into the Addgene repository (Addgene bacterial strain
#40803)\textsuperscript{80}.

We acquired the nuc4- strain from Addgene in order to test its
suitability as a strain for ex vivo DNA assembly with the underlying
assumption being that these exonucleases may compete to degrade
exposed homologies. Since this strain contained the Lambda Red system
we decided to integrate the lysis cassette from the XJa strain to
create and autolysis version of the nuc4-, which we dubbed nuc4-\(\lambda\)R
(for Genotype and description see Table A.2 and Table A.3; for oligo
sequences see Table C.3).
Figure 4.6 ex vivo DNA Assembly with nuc4-λR and Lambda RED

Assemblies were performed using original buffer, supplemented with 1 mM Manganese or 200 ng RecA using autolysed nuc4-, nuc4- with λRed or XJa. Percentages represent Blue colonies to all colonies (blue and white). Error bars represent standard error of the mean for reactions performed in duplicate (N=2).

Once isolated, the nuc4-λR strain was used to create two lysates. One lysate was prepared by the same methods as XJa (-λRED), while the other was heat-shocked to induce production of the Lambda Red proteins (+λRED) as other studies have shown the Lambda RED system capable of facilitating DNA assembly. The resulting lysates were used to perform 2-way assemblies with the standard buffer, with Manganese and with RecA (Figure 4.6). We were disappointed to find that our nuc4-λR showed no improvement over the standard XJa reaction under any tested conditions with or without the inclusion of the Lambda RED system.
Unfortunately, it is impossible to infer the direct relationship between the genotypic changes in nuc4-\(\lambda\)R and their impact on DNA assembly due to the absence of proper control strains. To clarify, we saw in the enzyme titration (Figure 4.1) that ssDNA exonucleases can have strong stimulating or inhibitory impact on DNA assembly. nuc4-\(\lambda\)R has four ssDNA exonucleases knockout, including one shown to stimulate DNA assembly (ExoI) and one shown to inhibit it (RecJ) in the enzyme titration experiments. Moreover, the parental strain (\textit{E. coli} MG1655) that nuc4-\(\lambda\)R is derived from does not have the mutations in \textit{endA} and \textit{recA} like the cloning strains used throughout this study. While MAGE oligonucleotides were purchased to disrupt \textit{endA} and \textit{recA} via Lambda RED recombineering (sequences in Table C.3), we were never able to identify clones containing the mutant alleles. Without progressive knockouts of the ssDNA exonucleases with background mutations in the \textit{recA} and \textit{endA} loci to use as controls it is impossible to make direct conclusions about the effect of each knockout as it pertains to DNA assembly.
Chapter 5. \textit{ex vivo} PCR and \textit{ex vivo} Cloning

5.1. The pUN plasmid and \textit{ex vivo} PCR

Using lysates derived from completely “wild-type” strains it is possible to perform sophisticated \textit{in vitro} reactions such as protein expression or DNA assembly. As demonstrated in the previous section, the background genotype of the \textit{E. coli} can be easily manipulated using modern genome editing tools to investigate genotype-phenotype relationships or create strains with phenotypic enhancements (i.e., autolysis). In a similar but perhaps simpler manner, expressional plasmids can be constructed and transformed into existing strains to produce one or more proteins (or even mRNA). Plasmid-based expression allows rapid introduction into the various strain backgrounds without the challenges required to integrate cassettes into the genome. We refer to lysates containing an (over)expressed protein(s) as a ‘functionalized lysate’, as a novel function has been imparted to the lysate. The following chapter demonstrates a functionalized lysate containing the \textit{Pyrococcus furiosus} (Pfu) DNA polymerase that has the newly imparted ability to perform PCR. This functionalized lysate is then applied to amplify the fragments for 2- and 3-way assemblies, which are subsequently assembled and transformed using our \textit{ex vivo} DNA assembly approach.
5.1.1. Design and construction of the pUN plasmid

Many plasmids have been developed and are available through numerous for- and not-for-profit distributors. These plasmids usually have features associated with specialized purposes such as clonal amplification, protein expression, interspecies transfer or phage production. Traditionally features are included which facilitate cloning/DNA assembly such as multiple cloning sites, integrase sites (Gateway cloning) or counter-selectable markers.

Up to this point, ex vivo assembly and its assays employed the commonly used high-copy number Biobrick plasmids, pSB1C3 and pSB1K3 to conduct assembly assays. These plasmids are freely distributed under a creative commons license and are fully compatible with the ‘Biobrick Assembly’ cloning methodology. However, we decided to design and develop our own vector useful for expression and cloning via ex vivo DNA assembly and other sequence-independent cloning methods.

The plasmid backbone was derived from the pJ251-GERC vector available in the Addgene repository (Addgene plasmid #47441). The backbone fragment (amplified by pGERC_BB-f/-r; see Table C.2) contains the Kanamycin resistance selectable marker and the low copy number p15a origin of replication (~5-10 copies per cell). The second-half, ‘insert’ portion was synthesized as two linear fragments by Integrated DNA Technologies (IDT) and contains all the designed cloning and expressional features. The ‘insert’ design draws heavily upon the “Unique Neutral Sites” (UNS) developed by Torella, et al. A UNS is a computationally derived 40 base pair site that meets several design
criteria including: approximately 50% GC-content without homopolymeric runs; Unfavorable intramolecular (hairpins) and intermolecular (bonding between different UNSs) interactions; absence of common restriction/cloning sites; absence of start codons, promoter-like sequences or sequence homology to the E. coli MG1655 genome. The authors reported 10 UNSs that were subsequently demonstrated to assemble multiple repeated sequences via Gibson assembly and even provided functional insulation of promoter sequences from genetic context. In our design, some of these UNSs are incorporated throughout the insert separating each of the following features (in order from 5’ to 3’): an upstream terminator; a T5 promoter and 5’ untranslated region with a ribozyme; a newly optimized BCP cassette driven by a strong constitutive promoter and synthetic ribosome-binding site (RBS) and; two strong downstream terminators (sequence available in C.2.6).

The design of the insert was such that the new BCP cassette can be replaced with a desired RBS and protein coding sequence (CDS) and be driven by the strong upstream T5 promoter. Since the original BCP insert causes a very strong (and quicker developing) blue pigmentation, assemblies can be screened for the presence of background colonies. Unfortunately, we have found that the upstream promoter does not transcribe through the downstream UNS site so currently a promoter needs to be included with expression assemblies. However, it is from the UNS sites that our plasmid derives its namesake, pUN (for “plasmid unique neutral”). The embedded UNS sites play a critical role not just for simplifying assembly and insulating
parts as demonstrated by the original authors, but as ideal priming sites for PCR. For all the reasons the UNS sites make excellent termini for assembly reactions, these sites are also ideal primer sequences for amplification or as 5' embedded overhangs. Indeed, we designed and synthesized forward and reverse primers for each UNS site (sequences available in Table C.2) and have performed PCR from all incorporated UNS sites without any preliminary optimization or troubleshooting.

5.1.2. Construction of pUN-PrhaBAD-Pfu

In Addgene there are two DNA polymerase-expressing vectors available: pAKTaq (Addgene plasmid #25712) that expresses the *Thermus aquaticus* (Tag) DNA polymerase and pET16b.Pfu (Addgene plasmid #12509) that expresses the *Pfu* DNA polymerase. While we were able to successful express the Tag protein and use it to perform ex vivo PCR, Tag is a less than ideal cloning polymerase. Tag does not create flush termini (leaves one 3' adenine residue overhang) and also lacks the proofreading activities typical of high-fidelity polymerases. On the other hand, Pfu has proofreading abilities and creates blunt-ended products ideal for cloning purposes, but the pET16b.Pfu construct is a nightmare for *E. coli* expression. The difficulty stems from the poor codon compatibility of the Pfu coding sequence with *E. coli* as the host. To circumvent the poor codon adaption a special strain of *E. coli* (BL21(DE3) Rosetta) is needed which hosts a vector overexpressing some of the less prevalent tRNAs along with the DE3 prophage (for T7 polymerase expression).
In the spirit of open-access ex vivo DNA assembly we decided it was prudent to have the coding sequence codon optimized and synthesized so that the Pfu gene can be expressed in any E. coli. To perform codon-optimization the COOL codon optimization tool (http://bioinfo.bti.a-star.edu.sg/COOL/) was employed to simultaneously maximize codon context and codon adaptability for E. coli\textsuperscript{85}. With the Ribosome Binding Site Calculator tool a synthetic ribosome-binding site was designed based on the codon-optimized Pfu gene and an upstream wild-type rhamnose-inducible promoter (PrhaBAD)\textsuperscript{9,86}. The resulting cassette plus the flanking UNS sequences was synthesized by IDT as two overlapping fragments of 1144 and 1483 base pairs and subsequently assembled along with the backbone using ex vivo DNA assembly. Of the approximately 200 white transformants, 24 colonies were lifted (along with two blue colony controls) and used as template in colony PCR. The PCR amplified from UNSX across the entirety of the Pfu cassette to UNS9, and of the 24 clones all 24 showed amplicons at the expected size (~2800 bp; Blue colonies were ~1500) although the bands were faint. Of these 24, 3 clones were miniprepped the next day and sent for sequencing (Eurofins MWG). Of these three, one clone was identified as completely correct and used for subsequent procedures (pUN-PrhaBAD-Pfu sequence available C.2.7).

5.1.3. Expression and Preparation of ex vivo PCR lysate

As the original plasmid was transformed into NEB10β (and XJa/autolysis strains are a rarity in labs), expression and lysis was performed in NEB10β for all subsequent steps. Expression was performed by
inoculating 24 mL of pre-warmed SOB media (in a 125 mL Erlenmeyer) with 200 µL of an overnight culture, allowing growth for approximately 2 hours, then adding filter sterilized 20% w/v rhamnose to a final concentration of 0.2% (~250 µL).

To assay the culture for the production of the Pfu DNA polymerase, aliquots were removed at 6, 9 and 24 hours post-induction and frozen until needed. The remainder of the culture was collected at 24 hours post-induction. These aliquots were lysed using CelLytic B, diluted and protein content quantified by Bradford assay. Each protein sample was prepared with 25 µg of total protein (along with a NEB10β negative control) and visualized using SDS-PAGE under denaturing conditions. In the resultant gel (Figure 5.1) there is clearly a strong band in all the sample lanes at roughly 90 kD that is absent in the negative control lanes. This size correlates with the predicted molecular weight of the Pfu DNA polymerase (90.11 kD). Interestingly, there appears to be little variation in the concentration of Pfu across the different harvesting times.
Figure 5.1 Expression of Pfu DNA Polymerase
Soluble lysate of NEB10β expressing Pfu was harvested at different time points and separated via denaturing SDS-PAGE. “C” is NEB10β without pUN plasmid, “L” is low-molecular weight ladder.

Given the encouraging results of the protein gel in Figure 5.1, the remaining pellets of the culture harvested at 24 hours were thawed and used to prepare lysates for ex vivo PCR. The lysis reaction was performed using CelLytic B in a 1.5:1 ratio and incubated at 30°C for 15 minutes and subsequently diluted with lysis buffer to a final volume-to-pellet mass ratio of 5:1. While most of the lysate was centrifuged to remove the cellular debris an aliquot was first removed which retained the total cellular extract. The soluble fraction and the total cellular extract were next incubated at 80°C for 30 minutes to kill any host proteins (such as nucleases). The lysates are given one last extended spin in the centrifuge to pellet any denatured insoluble host proteins and the supernatant aspirated and diluted by
half with 100% glycerol. As with the ex vivo DNA assembly lysates, these lysates are stored at -20°C and can be safely removed and replaced into the freezer many times.

Figure 5.2 Volumetric titrations of ex vivo PCR
Functionalized lysates of Pfu were titrated to amplify the BCP band used in 2-way linear assemblies. Bands to the right were prepared with total lysate in the heat-kill process. The values below the bands indicate µL of lysate used per 50 µL reaction; “C” represents negative control of purified BCP fragment.

5.1.4. ex vivo PCR
Equipped with lysates containing the Pfu DNA polymerase, initial PCR reactions amplifying the 2-way circular BCP fragment were formulated using volumetric titrations of the extracts in a standard 50 µL Pfu PCR reaction. These titrations shown in Figure 5.2 exhibit a very nice trend with a peak at 1-2 µL of cell extract per 50 µL PCR. Comparing the bands on the left (heat-killed soluble fraction only) to the bands on the right (heat-killed total extract) it would appear that including the total cellular extract in the heat-kill step of lysate preparation increases the yield of Pfu in the final extract. This may be due to additional lysis of cells surviving the detergent treatment or may occur by freeing Pfu trapped in the insoluble fraction of the
extract prior to centrifugation. Second, it is clear that using excessive extract (or too little) causes an observable decrease in PCR efficiency. It is therefore advisable to perform a small titration experiment each time new lysates are prepared. While a slightly inconvenient consumption of materials and time, the long-term savings from the ex vivo approach are more than worth the investment.

Figure 5.3 ex vivo PCR of 2-way and 3-way Fragments
Fragments for 2-way circular (2WC), 2-way linear (2WL) and 3-way circular (3W) were amplified using ex vivo PCR. In addition control PCRs were performed to detect the contribution of background genomic (BG 16S) and pUN-PrhaBAD-Pfu plasmid (pUN) DNA. The positive background (+ pUN) control had purified plasmid template added. There is a clear presence of background amplification and side amplification (see BCP lanes).

After constructing our Pfu-functionalized lysate and prototyping the ex vivo PCR reaction with volumetric titrations, ex vivo PCR was applied to generate the substrate DNA for 2-way linear, 2-way circular and 3-way circular assemblies. The analytical gels shown in Figure 5.3 demonstrate that Pfu-functionalized lysate was capable of synthesizing all the substrate DNA used in our assembly assays we decent efficiency. However, there is a fair amount of side-product formed in some of the reactions. The presence of high levels of contaminating host DNA may be the critical factor in the side-product formed. This is clearly evidenced by the strong amplicons formed in the background
control lanes (“BG”). These lanes consisted of ex vivo PCR reactions using primers targeting the highly conserved 16S genomic loci (“16S”) or the pUN-PrhaBAD-Pfu (“pUN”) plasmid but without the addition of template. In the future the steps to remove host DNA contaminations would bolster ex vivo PCR fidelity. Fortunately, since the Pfu enzyme is thermostable, non-thermostable nucleases – such as DNase I or RecBCD – could be used during lysate preparation to destroy host DNA and then semi-selectively denature by a heat-kill before desired DNA substrates or products are introduced to the lysate.

5.2. ex vivo Cloning

Up to this point ex vivo DNA assembly has been demonstrated, investigated and optimized; tools such as clever assays for DNA assembly, self-lysing strains and a novel cloning plasmid were designed and developed; and the Pfu DNA polymerase was optimized, cloned and expressed to functionalize lysates with the ability to perform a high-fidelity polymerase chain reaction. Yet, the ultimate objective of the project was to merge these ex vivo tools and techniques into an encompassing ex vivo-based cloning suite. As with all cloning, this began with the generation and isolation of substrate DNA as the remainder of the 2-way and 3-way ex vivo PCRs from Figure 5.3 were purified by gel extraction. Subsequent ex vivo DNA assembly reactions were performed using our NEB10β detergent-based, XJa detergent-based and XJa autolysis-based lysates, with the basic ex vivo DNA assembly reaction buffer or with 1 mM Manganese supplemented. In-house formulated Gibson mastermix was used as a positive control.
and transformation of substrate DNA (‘TAR’) was used as a negative control for background assembly.

**Figure 5.4 ex vivo Cloning**

Substrate DNA amplified using ex vivo PCR was gel purified and used in ex vivo DNA reactions. 2 µL of the 10 µL reactions were transformed using in-house prepared chemicompetent NEB10β. Gibson lane represents a positive control using Gibson isothermal assembly formulated in-house. TAR is a negative control consisting of DNA fragments transformed directly into NEB10β without assembly reaction. Percentages represent Blue colonies to all colonies (blue and white). Error bars represent standard error of the mean for reactions performed in triplicate (N=3).

The most striking aspect of the ex vivo Cloning pipeline is the elevated mutagenic rate in comparison with assemblies using Q5 DNA polymerase (Figure 3.4). While the total (Blue and White) colony forming units per femtomole are very comparable, the percentage of white colonies-to-blue colonies changes by over 20% in some circumstances. Even the Gibson assembly control shows a greatly increased error rate over typical efficiency. As the only difference
between these two tests are the manner by which the substrate DNA was prepared, it is convincing that the mutations are introduced during the PCR amplification of the substrates. However, this may not be indicative of significantly lowered fidelity of Pfu in comparison with Q5 polymerase, as misamplification products were observed in the ex vivo PCR reaction (Figure 5.3). Since the primers encode all the homology necessary for faithful assembly efficiencies of joining correct amplicons versus misamplified products would be approximately equivalent. If indeed this is the case, a simple post-lysis digestion using a non-thermostable nuclease (e.g., DNase I) followed by the typical 80°C heat-treatment should alleviate the observed mutational rate of ex vivo Cloning.
Chapter 6. Conclusion and Future Directions

Responding to the needs of the synthetic biology community for inexpensive, rapid and efficient DNA manipulation we were able to uncover a unique approach to DNA assembly. Working on the observation that all of the enzymes used in prototypical in vitro manipulations of DNA are highly conserved, we incorporated an ex vivo approach to screen several candidate organisms for DNA assembly functions. To accomplish this we designed a very clever assay of DNA assembly.

Our circular assemblies provided a colorimetric output that allowed us to interpret the origins of each resulting transformant. In our initial pilot studies, we used two chloramphenicol plasmids in our two-way assembly and observed the growth of blue (correct), white (background) and red (colonies) – with many background colonies observable (Figure 1.1.B). While we realized that this would allow us to track transformation efficiencies by observing fluctuations in background, it was not until subsequent studies did we realize that if the donor plasmids were maintained on different antibiotics than the destination vector (like in the original 3-way assembly; see Figure 2.3), we could track mutational rates. Because our assemblies join the BCP coding sequence to the destination vector at a crucial junction for proper expression of the BCP protein (across Ribosome Binding Site and Start Codon), there is a low tolerance for error during end joining. Due to the low throughput of performing transformation to assess DNA assembly
(~28 hours) we also designed a linear assembly assessed by gel electrophoresis (~3 hours) that greatly expanded our investigative capacities.

We leveraged this linear assembly to optimize and investigate ex vivo DNA further. During our titration of the assembly reaction with divalent metals we found contradictory results between our autolysis strain and detergent lysed strain. This could be a function of the composition of the detergent, as it could very well contain the strong chelators EDTA or EGTA in order to improve lysis efficiency. The presence of these chelators would alter the apparent divalent cation ratios between detergent and autolysis treated extracts. However, there was good agreement between both strains that manganese had a stimulatory effect on end joining, which we confirmed in follow up titrations. Similarly, we investigated the biochemical origins of the ex vivo DNA assembly by titrating putative enzymes.

In E. coli, DSB repair, dominated by homologous recombination, relies on the RecBCD complex to expose ssDNA ends and load RecA, forming the nucleofilament that subsequently directs a homology search and strand invasion of dsDNA\(^6\) – but this is different phenomena than directly joining two ends of DNA together directly. Previously, it was shown that recA- strains of E. coli join linear dsDNA with homologous overhangs in lysate more efficiently than their RecA-expressing counterparts\(^1\). However, we observed an interesting dichotomy surrounding RecA. During our titrations using the linear assay, RecA showed consistent enhanced assembly properties, but when applied to
circular assemblies and transformations it showed a consistently negative effect. It could be that RecA plays a role preserving the ends of the linear DNA products, while the nuclease in recA- lysates are free to degrade the product (and substrate) DNAs\textsuperscript{63}. Either way, as recBCD- strains have also been reported to direct accurate DNA end joining, ex vivo assembly reactions may involve mechanisms independent of homologous recombination machinery, perhaps by an alternative end-joining pathway\textsuperscript{29,87}. We suspect that both ex vivo and in vivo DNA assembly is likely facilitated by multiple competing DNA repair mechanisms, but largely by RecBCD.

Exonuclease activity is required for DNA end resection, but is a balancing act. We saw with yeast very little exonuclease activity, which preserves the regions of homology, but may not expose complementary overhangs. Conversely, extreme exonuclease activities, like that of Deinococcus, may be resecting bi-directionally, deleting the necessary homologous regions. Again, knockouts of exonucleases implicated in deletion of exposed overhangs may greatly improve the efficiency of end joining\textsuperscript{88}. In our own titration experiments we saw this effect with the addition of dsDNA exonucleases (RecBCD and ExoIII). Certain ssDNA exonucleases also showed detrimental effects on assembly (DNA PolI and RecJ), but there was significant substrate DNA remaining in the lanes supporting the hypothesis that they possess the ability to recess homology arms. In the cell ssDNA activity is actually beneficial as a sort of check-and-balance to ensure the cell can abort non-productive recombination events\textsuperscript{88}. Previous studies have
shown ExoI and RecJ knockouts have elevated mutational rates due to the loss of this abortive activity\textsuperscript{88}. These hypotheses would be very convincing if the complete opposite effect weren’t observed in the exact same study. During the titration ExoI and - at its lowest levels - ExoT showed the ability to greatly enhance assembly, almost 8-fold in some circumstances. These both have opposite polarities to the inhibitory RecJ ssDNA exonuclease so there is clearly a preference for degrading the 3’ extensions, rather than the 5’ extensions. While we lacked the full library of strains needed to draw exquisite conclusions, the equivalency between nuc4-\lambda R and XJa highlights the complexity and redundancy of exonuclease activity and DNA repair pathways\textsuperscript{89}.

The final task at hand was to complete our ex vivo-based cloning pipeline. To do so we designed a new plasmid that incorporated features designed by the community (UNSs, riboJ), de novo designed parts (RBS, Codon-optimized gene) and the screenable BCP cassette from our own works. Using ex vivo assembly we recombined the plasmid backbone together with two larger synthetic fragments in a very effective manner. Based on 24 tested clones, all 24 contained the correct size inserts. Out of 3 sent for subsequent sequencing only one had no errors in the Pf\textsubscript{u} gene. This represents a clear win for ex vivo DNA assembly as a very practical cloning tool. The Pf\textsubscript{u} polymerase was at one time the highest-fidelity polymerase, though it has now been replaced by more engineered, more expensive polymerases like Q5. The assemblies showed a greatly elevated error rate based on an increased
white:blue colony ratio, but this might be corrected simply by removing the contaminating host DNA or performing PCR optimization (e.g., annealing temperature, buffer composition, etc...).

With the successful transformation of assembly reactions derived from both an optimized ex vivo DNA assembly and an ex vivo PCR, the project has successfully completed it’s ultimate objective of establishing an ex vivo cloning suite. Yet, given the outcomes of the experimentation it is easy to instead focus on perceived shortcomings of the project. There were many questions that remained unanswered, conflicting results and components not carried to ‘full completion’ (i.e., generation of all nuclease-knockouts, optimization of ex vivo PCR). Even more egregious for some is the lack of data backed by extensive replicates. While understandable, this is simply a function of the biological systems being studied, the objectives of the study and even the ex vivo approach to the study.

Biological systems are inherently noisy and complex, meaning establishing definitive causalities is a difficult task often not accomplished in even the very deepest of investigations. Further, the particular subcellular system being probed – DNA maintenance and repair – is conserved amongst every living organism. Such conservation is indicative of the “do-or-die” nature of this pathway and is reflected in the degree of redundancy and orchestration. By using an approach like ex vivo that combines the throughput of an in vitro system with endogenous complexity of an in vivo system, investigations into these mechanisms can generate results that quickly
deviate the project from the original objective into a very large study all of it’s own.

Continued development of faster and cheaper DNA assembly and genetic characterization methods, such as the ex vivo DNA cloning pipeline described here, will further accessibility and success of DNA-based studies and applications. As mentioned in detail, there were a number of surprising findings throughout the study that could be pursued. The first is the interesting case of ExoI. ExoI could be purchased and titrated into lysates with many different host background genotypes to investigate its contribution towards recombination/end joining. Alternatively, modifications could be employed to try and modulate native degradation. For example, using primers with non-standard phosphodiester analogues – such as the phosphorothioate bond – could be employed to halt degradation. This could be very helpful in the linear assembly when the product is formed and subsequently degraded from exposed termini before visualization. Building off the success of the Pfu functionalized lysate, a multienzymatic functionalized lysate featuring a ligase, polymerase and exonuclease could be constructed as an ex vivo mimic of the tripartite Gibson assembly. Moreover, functionalized lysates hold the potential to address some problems unsuitable to living cultures. Biocontainment has become a major point of contention for the employ of genetically modified organisms, particularly in an environmental context\textsuperscript{90,91}. By substituting a functionalized lysate, the same systems can be applied as bioremediation treatment without the risk of spreading throughout the
environment, and without the tremendous costs associated with bulk enzyme purification. While ex vivo engineering and ex vivo molecular biology may not be the “silver bullet” to solve every societal problem, there are many academic and industrial opportunities addressable by ex vivo and synthetic biology.
Chapter 7. Materials and Methods

7.1. Reagents, DNA and Enzymes

All chemical reagents were purchased from Sigma-Aldrich. Divalent metals used in titration experiments were all complexed with chloride (i.e., MgCl₂, MnCl₂, CaCl₂, ZnCl₂) and formulated as 1 Molar Stock solutions. All enzymes used in the titration experiment were purchased from NEB, along with the Q5 polymerase used for PCR. All DNA synthesis of primers and synthetic fragments was performed by IDT. The plasmids used for 2-way and 3-way assembly experiments were obtained from the Registry of Standard Biological Parts. Sequences for all DNA used are provided in (DNA Sequences).

7.2. Generation of DNA Fragments for ex vivo DNA Assembly

The templates, primers and products used for the assembly assays are summarized in Appendix C.

7.2.1. PCR with Q5 DNA Polymerase

The DNA fragments used to demonstrate ex vivo DNA assembly were generated using standard PCR of parts of the following plasmids from the BioBricks registry: pSB1C3-J04450, pSB1C3-K592009, and pSB1K3-J04450. Primers used in this study (Table C.1) were generated using the j5 automated DNA assembly software⁹²-⁹⁴. Amplicons were generated by 100 µL PCR reactions with Q5 polymerase (NEB) under standard reaction
conditions. These reactions were cycled at 98°C for 30 s; 98°C for 10 s, 50°C for 15 s, 72°C for 25 s (repeated for 25 cycles); 72°C for 2:00. Resultant PCRs were subsequently purified using agarose gel electrophoresis and extraction (Section 7.5).

7.2.2. ex vivo PCR with Pfu DNA polymerase

Pfu PCRs were formulated as follows (final concentration): 200 μM each dNTP, 1X Q5 reaction buffer (2 mM MgCl2), 500 μM each primer and ~2 ng template DNA per 50 μL reaction. These reactions were cycled differently for small fragments (<1 kb) and large fragments (>1 kb). For smaller pieces (i.e., BCP and Promoters): 95°C for 90 s; 95°C for 30 s, 55°C for 30 s, 73°C for 90 s (repeated for 28 cycles); 74°C for 3:00. For larger fragments (i.e., backbones, background controls): 95°C for 90 s; 95°C for 45 s, 55°C for 30 s, 73°C for 4 min (repeated for 30 cycles); 74°C for 6:00. Resultant PCRs were subsequently analyzed or purified using agarose gel (See, 7.5).

7.3. Preparation of Lysates

7.3.1. Initial Study

Cellular lysates were prepared from the following strains: E. coli NEB10β (NEB™ #C3019), S. cerevisiae BY4741 and D. radiodurans R1 (ATCC® 13939). E. coli was grown in Terrific Broth with glycerol (Sigma®) at 37°C with shaking at 250 rpm. D. radiodurans was grown in 123 TGY medium (5% Tryptone, 5% Yeast extract, 1% Glucose, 1% Potassium monophosphate) at 30°C and S. cerevisiae was grown with YPD media (Sigma™) at 30°C, both shaking at 250 rpm. The preparation of
the bacterial (E. coli and D. radiodurans) and the S. cerevisiae lysates varied slightly: bacterial cultures were pelleted once OD600nm= 6.00-6.50, while the yeast cultures were pelleted once OD600nm= 4.00-4.50. Volumes of 4-6 mL of culture were centrifuged at 13,200 rpm at 4°C for 2 min, washed with 1 mL of Milli-Q H2O, centrifuged again and the wet pellet massed. 2X CelLytic B Lysis Reagent (Sigma™) was added to the bacterial cell pellets at 3 µL/mg of cells. 1X CelLytic Y Lysis Reagent (Sigma™) was added to the yeast pellets in the same ratio. After the addition of the lysis reagents, the cells were incubated at 37°C for 10 min shaking at 300 rpm. The lysed cells were centrifuged at 13,200 rpm for 15 min and a 20 µL sample of the supernatant (lysate) was mixed with 20 µL of 100% glycerol to yield 40 µL aliquots. All lysate aliquots were stored at -20°C.

7.3.2. Optimized Methods

Cultures of E. coli were grown under similar conditions to those reported in 7.3.1, with the exception that the cultures were inoculated from starter cultures. For cultures using the autolysis cassette, the media was supplemented with 3 mM L-arabinose and 10 mM magnesium chloride from a 500X stock. Cultures were grown into saturation (OD600nm= 2.0-4.0 after about 18 hours growth) and then pelleted at 4200RPM at 4°C for 20 minutes. From this point on cells were kept on ice. Media was aspirated, making sure to remove as much as possible using a pipette tip and the pellets were massed (~75 milligrams per 5 mL) in pre-massed tubes. We found that tubes can vary
significantly and accuracy is best when each tube is massed initially. Cells were washed once with ice-cold 10% glycerol solution in order to prevent premature lysis in a hypotonic solution by the autolysis strains. For detergent lysis, cell pellets are resuspended in 1.5 volumes of CellLytic B 2X per milligram of pellet and incubated at 30°C with shaking for 15 minutes. The resulting lysate is then diluted with an additional 3.5 volumes of Lysis buffer (50 mM Tris-HCl, pH8; 0.2 mM EDTA; 2 mM DTT; 200 mM NaCl; 0.2% Triton X-100; 0.1 mM PMSF). Autolysis pellets are resuspended in 5 volumes of Lysis buffer and frozen at -80°C or in liquid nitrogen. Thawing the tube in a room temperature water bath lyses the autolysis strain. All lysates are then spun at the highest speed (20,000xg) for 20 minutes at 4°C. The resulting supernatant are very carefully aspirated into an equal volume of glycerol, mixed thoroughly (with pipette or vortex) and stored at -20°C.

7.3.3. ex vivo PCR Lysate

From an overnight culture, a SOB media (it is critical to use a medium without glucose) culture is inoculated (at a 10⁻² dilution) and cultured at 37°C rotating at 250 RPM for about 2 hours. At this point the L-Rhamnose inducer is added to 0.2% (from a filter sterilized 20% solution). After another 3-24 hours the culture can be harvested at 8000xg and washed with 10% glycerol. The culture is lysed in 1.5 volumes of CellLytic B 2X and diluted to 5 volumes using lysis buffer (see above). The entire lysate is then incubated at 75°C for 30 minutes, centrifuged at the highest speed (16,000xg) for 20 minutes at
4°C and the supernatant aspirated into an equal volume of glycerol, mixed thoroughly and stored at -20°C.

7.4. DNA Assembly Reactions

7.4.1. Initial Study

Buffers for assembly reactions were prepared from 100× stock solutions. Stock solutions were as follows: 100 mM NADβ, 100 mM ATP, 100 mM DTT, 1 M MgCl₂. Tris-HCl was added to the buffer to 500 mM from a 1 M stock solution. A typical assembly reaction would include 6 µL of cellular lysate, 2 µL of 10× buffer, 2 µL of nuclease-free water and 10 µL of a DNA master mix. DNA master mixes contained all fragments needed for assembly and were formulated with 20 ng/µL of the backbone and 6:1 molar ratio of insert fragments to the backbone.

7.4.2. Optimized and Supplemented Conditions

Buffers were also formulated from 100X stocks with the exception of the divalent cation titrations. For these experiments a “base” buffer with all components except for magnesium and water was formulated and then aliquoted, then these aliquots were fully constituted with the various cations and water to a 10X concentration. A typical assembly optimized assembly was formulated as follows: 1 µL of 10X Buffer, 5 µL of a 2X DNA Mastermix, 1 µL of 10X Lysate (at ~200 µg/mL) and 3 µL of nuclease-free water. The DNA master mixes were formulated with 5 ng/µL of the backbone fragment (BCP band for 2-way linear) and molar ratios of 1.2:1 and 2:1.2:1 for 2-way (BCP:BB) and 3-way (Promoter:BCP:BB),
respectively. These reactions were transferred immediately to the thermocycler and incubated at 37°C for 2 hours. Upon completion reactions were either used immediately for transformations or gel electrophoresis, or frozen at -20°C until used.

7.5. Agarose Gel Analyses and Extraction

7.5.1. Initial Study

Gel analysis was performed using 1.0–1.2% agarose-TAE gels containing GelRed (Biotium™) as the staining agent. All analytical gels were run at 100 V for 30–50 min and visualized with a Molecular Imager® Gel Doc™ XR+ Imaging System with Image Lab™ v4.0 software (Bio-Rad).

To purify PCR reactions, samples were run on 1.0% agarose-TAE gels stained with GelGreen (Biotium™). The gels were run at 100 V for 30 min and subsequently visualized under blue light excitation. The gels showed no side product formation and bands were excised and isolated using ZymoClean™ Gel DNA Recovery Kit (Zymo Research™).

7.5.2. Optimized Analytical Gels

In the optimized format, analytical gels for 2-way linear reactions were formulated at 1.8% agarose in TAE buffer. 3.5 µL of GelRed (Biotium™; 50,000X in Water) was added to the molten gel and once poured the gels were allowed to solidify for at least one hour. To each 10 µL assembly reaction, 3 µL of 5X Loading Dye was added and 5 µL loaded. For Ladders (NEB 100-bp, 1kb or 2-log Quick-load) 3.8 µL
were added to each lane. The gel was run at 180V for 1 minute, the run at 90V for 35 minutes.

7.5.3. Gel Analysis

To analyze the gel they were visualized with a Molecular Imager® Gel Doc™ XR+ Imaging System (Bio-Rad). Using Image Lab™ v4.0 software (Bio-Rad) the lanes and bands were identified. The lane profile was used to manually adjust bands with exceptional tailing or false peaks. Using the Relative Quantity analysis tool, the BCP band of the negative control is selected as the internal reference. Band information was exported into a Comma Separated Values file for subsequent analysis. Pictures of all gels not included in the main text are provided in (Appendix B Gel Pictures).

7.6. Transformations and Preparation of Competent Cells

7.6.1. Initial Study

Transformations were performed using chemically competent E. coli NEB10β (NEB) according to manufacturer’s recommendations. Briefly, 2 µL of reaction mixture (from ex vivo assemblies) or diluted DNA master mix (for in vivo assembly) were added to each transformation, these were incubated for 30 min on ice, heat shocked at 42°C, recovered in 950 µL SOC at 37°C for 60 min and 50 µL of culture was spread onto agar plates containing the appropriate antibiotic.
7.6.2. Optimized Competence and Transformation

To prepare chemicompetent *E. coli* for chemical transformation the “Ultracompetent” method of Inoue, et al was followed closely. As there is not a refrigerated incubator available in our lab to culture at 18°C the cultures were incubated at room temperature (~21°C).

To perform transformations, we incubated 48 µL of cells with 2 µL of DNA sample on ice for 30 minutes. The cells were heat-shocked in a 42°C water bath for 45 seconds and immediately transferred and swirled in an ice-water bath for 90 seconds. 950 µL of room temperature SOC was added to the tube and it was moved to the 37°C incubator with shaking at 250 RPM for 1 hour. Transformants were serially-diluted into sterile PBS and 100 µL plated onto LB with the appropriate antibiotic.

To prepare electrocompetent cells, we followed the protocol “Transformation of *E. coli* by Electroporation” in the Molecular Cloning Manual 4th ed. Electroporation reactions were carried out in a Bio-Rad Micropulser using the same volumes as used for chemical transformation above.

CFU per femtomole values were calculated by taking the raw counts of countable plates (30-300), multiplying by the dilution factor and dividing by the number of femtomoles of DNA transformed. Given the backbone of *ex vivo* assemblies as the limiting substrate at 5 ng/µL, 2 µL equates to approximately 6.6 femtomoles of 2-way and 7.3 femtomoles
2 µL of a 50 pg/µL solution of pUC19 equates to 0.06 femtomoles.

7.7. Lambda RED Recombineering

Genomic manipulations of the nuc4- strains were performed using the protocols laid out by Wang, et al. Summarizing, a single cycle of recombineering consisted of growing a culture of nuc4- in a 3 mL culture of LB + Ampicillin until OD600 = ~0.4, inducing lambda expression at 42°C for 15 minutes, washing 1 mL of the culture twice with ultrapure water, concentrating at ~50 µL and transforming 2 µL of a 10 µM solution of MAGE oligo(s)/λR cassette. In order to modify nuc4-, an already existing chloramphenicol cassette in the host genome was deactivated using the cat_fwd_stop oligo (see Appendix C) and the chloramphenicol mutant (nuc4-.Cm-) was isolated by replica plating. The λR autolysis cassette (including chloramphenicol resistance) was amplified out of XJa, gel extracted and transformed into nuc4-.Cm-. Unfortunately, there was so much homology between the chloramphenicol cassette in the λR autolysis cassette and the deactivated chloramphenicol loci that many, many false-positive chloramphenicol transformants were obtained (as confirmed by allele-specific PCR; Table C.3). Instead to identify the correct integrants, a screen approach utilizing Phenol Red Agar + 2% Arabinose was employed to identify colonies unable to utilize arabinose. A positive clone was isolated and the integration of the lysogen gene was confirmed by allele-specific genotyping and a lytic phenotype. Sequences of all oligos used for recombineering are provided in (Table C.3).
Chapter 8. References

101
Appendix A. Tables

Table A.1 Summary of Assemblies

<table>
<thead>
<tr>
<th>Part</th>
<th>Template</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2-way Circular (2WC)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCP</td>
<td>pSB1C3-K592009</td>
<td>exda1</td>
<td>exda2</td>
<td>698</td>
</tr>
<tr>
<td>Backbone</td>
<td>pSB1K3-J04450</td>
<td>exda3</td>
<td>exda4</td>
<td>2446</td>
</tr>
<tr>
<td><strong>2-way Linear (2WL)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Promoter</td>
<td>pSB1C3-J00450</td>
<td>exda22</td>
<td>exda23</td>
<td>381</td>
</tr>
<tr>
<td>BCP</td>
<td>pSB1C3-K592009</td>
<td>exda12</td>
<td>exda13</td>
<td>697</td>
</tr>
<tr>
<td><strong>3-way Circular (3WC)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backbone</td>
<td>pSB1K3-J04450</td>
<td>exda3</td>
<td>exda19</td>
<td>2214</td>
</tr>
<tr>
<td>Promoter</td>
<td>pSB1C3-J00450</td>
<td>exda20</td>
<td>exda21</td>
<td>387</td>
</tr>
<tr>
<td>BCP</td>
<td>pSB1C3-K592009</td>
<td>exda12</td>
<td>exda2</td>
<td>696</td>
</tr>
</tbody>
</table>

Description of the DNA fragments used in each assembly assay. DNA sequences for primers, template plasmids and products of assemblies can be found in Appendix C.
### Table A.2 Strain Genotypes

<table>
<thead>
<tr>
<th>Strain</th>
<th>Parent</th>
<th>Synonymous Strain</th>
<th>Growth (Celsius; Resistance)</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEB10β</td>
<td>K-12</td>
<td>DH10Beta</td>
<td>37; Str</td>
<td>Δ(ara-leu) 7697 araD139 FhuA ΔlacX74 galK16 galE15 el14-&lt;br&gt;Δ(lac-proAB) glnV44 (supE44) el14-&lt;br&gt;Δ(mrr-hsdRMS-mcrBC) $\phi$80dlacZΔM15 recAl relAl endAl nupG rpsL (StrR) rph spoT1 &lt;br&gt;hsdR17(rK- mK+) relAl recAl araB::λR, cat (CmR)</td>
</tr>
<tr>
<td>XJa</td>
<td>K-12</td>
<td>JM109</td>
<td>37; Cm</td>
<td>F'[traD36 proA+B+ lacIq Δ(lacZ)M15] (McrA-)&lt;br&gt;thi gyrA96 (NalR) endAl hsdsR17(rK- mK+) relAl recAl araB::λR, cat (CmR)</td>
</tr>
<tr>
<td>nuc4-</td>
<td>EcNR2</td>
<td></td>
<td>32; Amp Cm</td>
<td>F- λ- ilvG- rfb-50 rph-1 ΔmutS::cat&lt;br&gt;Δ(ybhB-bioAB)::[lcI857 N(cro-ea59)::tetR-bla] xonA- recJ- xseA- exoX-</td>
</tr>
<tr>
<td>nuc4-.Cm-</td>
<td>nuc4-</td>
<td></td>
<td>32; Amp</td>
<td>F- λ- ilvG- rfb-50 rph-1 ΔmutS::cat-&lt;br&gt;Δ(ybhB-bioAB)::[lcI857 N(cro-ea59)::tetR-bla] xonA- recJ- xseA- exoX-</td>
</tr>
<tr>
<td>nuc4-..λR</td>
<td>nuc4-</td>
<td></td>
<td>32; Amp Cm</td>
<td>F- λ- ilvG- rfb-50 rph-1 ΔmutS::cat&lt;br&gt;Δ(ybhB-bioAB)::[lcI857 N(cro-ea59)::tetR-bla] xonA- recJ- xseA- exoX- araB::λR, cat (CmR)</td>
</tr>
</tbody>
</table>

Genotypes of Strains used in the current study are presented above. Growth indicates temperatures used to culture strain and resistances available to strain. Relevant genotypes are indicated in bold and described below in Table A.3.
<table>
<thead>
<tr>
<th>Genotype</th>
<th>Phenotypic Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>recA1</td>
<td>Enables homologous recombination pathways for reduced occurrence of unwanted recombination in cloned DNA; cells UV sensitive</td>
</tr>
<tr>
<td>endA1</td>
<td>For cleaner preparations of DNA and better results in downstream applications due to the elimination of non-specific digestion by Endonuclease I</td>
</tr>
<tr>
<td>rpsL</td>
<td>Mutation in ribosomal protein S12 conveying streptomycin resistance</td>
</tr>
<tr>
<td>ΔmutS::cat</td>
<td>Deletion of mutS used for mismatch repair for better Lambda Red recombination; Conveys Chloramphenicol Resistance</td>
</tr>
<tr>
<td>ΔmutS::cat-</td>
<td>Inactivation of Chloramphenicol cassette by insertion of stop codon</td>
</tr>
<tr>
<td>lcI857 N(cro-ea59)::tetR-bla</td>
<td>Lambda RED operon of alpha, beta and gamma proteins induced by heat-shock; Cells must be cultured at lower temperatures; Conveys Ampicillin resistance</td>
</tr>
<tr>
<td>araB::λR, cat (CmR)</td>
<td>Insertion of Lambda phage endolysin inducible by arabinose; Disrupts arabinose metabolism; Conveys Chloramphenicol resistance</td>
</tr>
<tr>
<td>xonA- recJ- xseA- exoX-</td>
<td>Inactivation of exonucleases (ExoI, RecJ, ExoVII, ExoX) by premature stop codon insertion</td>
</tr>
</tbody>
</table>

Descriptions of relevant genotypes found in Table A.2
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Concentration (U/μL)</th>
<th>Temperature (Celsius)</th>
<th>Time</th>
<th>Buffer</th>
<th>[Substrate]</th>
<th>[Product]</th>
</tr>
</thead>
<tbody>
<tr>
<td>RecBCD</td>
<td>10</td>
<td>37</td>
<td>30 min</td>
<td>NEB 4</td>
<td>Linear dsDNA (no concentration)</td>
<td>10 nmol (acid-soluble nucleotides)</td>
</tr>
<tr>
<td>ExoI</td>
<td>20</td>
<td>37</td>
<td>30 min</td>
<td>1X ExoI</td>
<td>170 ng/μL (ssDNA)</td>
<td>10 nmol</td>
</tr>
<tr>
<td>ExoIII</td>
<td>100</td>
<td>37</td>
<td>30 min</td>
<td>NEB 1</td>
<td>.15 mM (sonicated dsDNA)</td>
<td>1 nmol</td>
</tr>
<tr>
<td>ExoT</td>
<td>5</td>
<td>25</td>
<td>30 min</td>
<td>NEB 4</td>
<td>1 nmol</td>
<td>.1 nmol (TCA soluble (polythymidine) nucleotides)</td>
</tr>
<tr>
<td>RecJf</td>
<td>30</td>
<td>37</td>
<td>30 min</td>
<td>NEB 2</td>
<td>30 ng/μL (sonicated ssDNA)</td>
<td>.05 nmol (TCA soluble nucleotides)</td>
</tr>
<tr>
<td>DNA PolI</td>
<td>10</td>
<td>37</td>
<td>30 min</td>
<td>NEB 2</td>
<td>70 ng/μL SalmonSperm DNA; 33 μM dNTPs</td>
<td>10 nmol (dNTPs -&gt; acid insoluble material)</td>
</tr>
<tr>
<td>Eco Ligase</td>
<td>10</td>
<td>16</td>
<td>30 min</td>
<td>1X Eco</td>
<td>0.12 μM DNA termini (300 ng/μL)</td>
<td>.06 μM ligated ends</td>
</tr>
</tbody>
</table>

Unit definitions of enzymes supplied by NEB and used in titration experiments. For exonucleases, exonuclease activity is measured as formation of acid-soluble nucleotides. For DNA PolI polymerization activity is defined by formation of acid insoluble material from dNTPs. E. coli Ligase activity is measured by the formation of ligated ends from digested material. All reported unit definitions are for a 50 μL reaction volume.
Appendix B. Gel Pictures

Below are the original agarose gel pictures used in the analysis of ex vivo assembly. The higher molecular weight band (BCP band) in the control lane (“C”) was used as the relative reference for band intensities.

B.1.1 Time Course of Deinococcus, Saccharomyces, and E. coli lysates

Figure B.1 Dra, Sce and Eco Time Course
B.1.2 Lysate Preparation (Growth Phase, Media and Glycerol)

Figure B.2 Lysate Preparation (Gel 1)
Figure B.3 Lysate Preparation (Gel 2)
B.1.3 Concentration of Lysates

Figure B.4 Lysate Dilutions (60 minutes)

Figure B.5 Lysate Dilutions (120 minutes)
B.1.4 Buffer Optimization

Figure B.6 Buffer Optimization (NEB10β)
Figure B.7 Buffer Optimization (XJa Autolysis)
B.1.5 Buffer Titration

Figure B.8 Buffer Titration (NEB10β #1)
Figure B.9 Buffer Titration (NEB10β #2)
Figure B.10 Buffer Titration (XJa)
B.1.6 Manganese Titration

Figure B.11 Manganese Titration (NEB10β)
B.1.7 Enzymatic Titrations

Figure B.12 RecBCD, PolI, RecA, Eco Ligase Titration
Figure B.13 ssDNA Exonucleases Titration
B.1.8 RecA Titration

Figure B.14 RecA Titration (XJa Autolysis)
## Appendix C. DNA Sequences

### C.1 Primer Sequences

**Table C.1 ex vivo Assembly Assay Primers**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>exda1</td>
<td>AGAGAAAGAGAAGAAATGAGTGTGATCGCTAAACAAATGACCTACAGTGGG</td>
<td>50</td>
</tr>
<tr>
<td>exda2</td>
<td>ATTTGATGCCTGTTATTAGGGCACCAGGTTTTGCGTGC</td>
<td>40</td>
</tr>
<tr>
<td>exda3</td>
<td>GGTCGCTTAATAACCAGGCATCAAATAAAACGAAAGCT</td>
<td>39</td>
</tr>
<tr>
<td>exda4</td>
<td>GCGATCACACTCATTTCCTCACCTCTTTCTCTGATGTA</td>
<td>41</td>
</tr>
<tr>
<td>exda12</td>
<td>AGAGAAATACTAGATGAGTGTGTACGCTAAACAAATGACCTACAGTGGG</td>
<td>48</td>
</tr>
<tr>
<td>exda13</td>
<td>GCCTGGCTCTAGTATTATTAGGCGACCACAGGATTCTGCGTGC</td>
<td>41</td>
</tr>
<tr>
<td>exda14</td>
<td>GGTCGCTTAATAACTAGAGCCAGGCATCAAATAAAACG</td>
<td>40</td>
</tr>
<tr>
<td>exda19</td>
<td>TGGTTTCTTAAGAGCTGATCCTTCAACTCAGCA</td>
<td>33</td>
</tr>
<tr>
<td>exda20</td>
<td>TGAGTTGAAGAGTCAGCTTTAAGAGAAACCATTATTATCATGACATTAACC</td>
<td>50</td>
</tr>
<tr>
<td>exda21</td>
<td>AGCGATCACACTCATTCTCTATTTTTCTCTCTCTCTCTAGATGTA</td>
<td>45</td>
</tr>
<tr>
<td>exda22</td>
<td>TGAAGGATCAGCCTCTAAGAAACCATTATTATCATGACATTAACC</td>
<td>45</td>
</tr>
<tr>
<td>exda23</td>
<td>GCGATCACACTCATTCTCTTTCTCTCTCTCTAGATGTA</td>
<td>44</td>
</tr>
<tr>
<td>Primer</td>
<td>Sequence</td>
<td>Length</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>UNS1F</td>
<td>CATTACTCGCATCCATTCTCAG</td>
<td>22</td>
</tr>
<tr>
<td>UNS1R</td>
<td>GAGACGAGACGAGACGC</td>
<td>18</td>
</tr>
<tr>
<td>UNS2F</td>
<td>GCTGGGAGTTCGTAGACG</td>
<td>18</td>
</tr>
<tr>
<td>UNS2R</td>
<td>GCTTGGATTCTCGTTTGT</td>
<td>19</td>
</tr>
<tr>
<td>UNS3F</td>
<td>GCACTGAAGGTCTCAATC</td>
<td>19</td>
</tr>
<tr>
<td>UNS3R</td>
<td>CGACCTTGATTTTCCAGTG</td>
<td>20</td>
</tr>
<tr>
<td>UNS4F</td>
<td>CTGACCTCTGCCAGC</td>
<td>16</td>
</tr>
<tr>
<td>UNS4R</td>
<td>GACTTTGCCTTTGTACTTACT</td>
<td>21</td>
</tr>
<tr>
<td>UNS5F</td>
<td>GAGCCAACCTCCCTTACAAC</td>
<td>20</td>
</tr>
<tr>
<td>UNS5R</td>
<td>CTCTAACGGACTTGAGTGAGG</td>
<td>21</td>
</tr>
<tr>
<td>UNS6F</td>
<td>CTCGTTCGCTGCCACC</td>
<td>16</td>
</tr>
<tr>
<td>UNS6R</td>
<td>GTATGTGACCCTAGATTTCTTAG</td>
<td>25</td>
</tr>
<tr>
<td>UNS7F</td>
<td>CAAGACGCTGGCTCTGA</td>
<td>17</td>
</tr>
<tr>
<td>UNS7R</td>
<td>CGAGTAGTTCAGTAGCGGA</td>
<td>19</td>
</tr>
<tr>
<td>UNS8F</td>
<td>CCTCGTCTCAACAAAGC</td>
<td>18</td>
</tr>
<tr>
<td>UNS8R</td>
<td>CCAGGGTGTTGAGGGT</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Sequence</td>
<td>Length</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>UNS9F</td>
<td>GTTCCTATCATCTGGCGAATC</td>
<td>22</td>
</tr>
<tr>
<td>UNS9R</td>
<td>CAGTGCTCTTGTGGGTTC</td>
<td>17</td>
</tr>
<tr>
<td>UNSXF</td>
<td>CCAGGATACATAGATTACCACAAC</td>
<td>24</td>
</tr>
<tr>
<td>UNSXR</td>
<td>GGTGGAAGGGCTCGG</td>
<td>15</td>
</tr>
<tr>
<td>pGERC-BB_F</td>
<td>TCGAGAAGGACACGGTGA</td>
<td>18</td>
</tr>
<tr>
<td>pGERC-BB_R</td>
<td>AAAGCCTTGTATGCTTCTTT</td>
<td>20</td>
</tr>
<tr>
<td>Oligo Name</td>
<td>Sequence</td>
<td>Product Length</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>cat_fwd_stop</td>
<td>G<em>C</em>A<em>T</em>CGTAAGAACATTTTGA&lt;br&gt;GGCATTTTCAGTCAGTTGCTCAATGA&lt;br&gt;ACCTATAACCAGACCGTTCAGCTGG&lt;br&gt;ATATTACGGCCTTTTTAAA</td>
<td>NA</td>
</tr>
<tr>
<td>recA.KO</td>
<td>C<em>A</em>A<em>A</em>TTGTTTCTCAATCTGGC&lt;br&gt;CCAGTGGCCTGCCCAACGCTcaCTa&lt;br&gt;TTaGTTTTCTGCTAGAgCCATTTTT&lt;br&gt;ACTCCTGTCAATGCAGGTA</td>
<td>NA</td>
</tr>
<tr>
<td>endA.KO</td>
<td>T<em>C</em>G<em>T</em>TTTAACACCGAGTAAGT&lt;br&gt;GATGTAACGTGATTTTTGCTATTTGCT&lt;br&gt;GCTGaGTTGGTACTGAGCGCAGCATT&lt;br&gt;TTCCGGCCCGCGGCTGGCC</td>
<td>NA</td>
</tr>
<tr>
<td>araB_lambdaR-f</td>
<td>CTTCGCCCCGCAAAGCCGTAA</td>
<td>2066</td>
</tr>
<tr>
<td>araB_lambdaR-r</td>
<td>ACAGGTGCTGAAATGCGGC</td>
<td>2066</td>
</tr>
<tr>
<td><strong>asLambdaR-r</strong></td>
<td>GGTCAGTTCGAGCATAAGGC</td>
<td>20</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------------</td>
<td>----</td>
</tr>
<tr>
<td><strong>endA.KO.v2_wt-f</strong></td>
<td>CGTTATTTGTCTATTTGCTGG</td>
<td>191</td>
</tr>
<tr>
<td><strong>endA.KO*_mut-f</strong></td>
<td>CGTTATTTGTCTATTTGCTGA</td>
<td>191</td>
</tr>
<tr>
<td><em><em>endA.KO</em>-r</em>*</td>
<td>GCACGATTCAGATCAAAACG</td>
<td>191</td>
</tr>
<tr>
<td><strong>recA.KO.v2_wt-f</strong></td>
<td>TGCCGCAACGCTTT</td>
<td>474</td>
</tr>
<tr>
<td><strong>recA.KO.v2_mut-f</strong></td>
<td>TGCCGCAACGCTCA</td>
<td>474</td>
</tr>
<tr>
<td><strong>recA.KO.v3-r</strong></td>
<td>GTGTTGACGCGGATTGTCAG</td>
<td>474</td>
</tr>
</tbody>
</table>
C.2 Plasmid Sequences

Below are Genbank formatted files of the various plasmids used throughout the study, complete with annotations.

C.2.1 pSB1C3-J00450 (RFP expression)

<table>
<thead>
<tr>
<th>LOCUS</th>
<th>pSB1C3-J00450</th>
<th>3139 bp ds-DNA</th>
<th>circular</th>
<th>22-APR-2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFINITION</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACCESSION</td>
<td>pSB1C3-J00450</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VERSION</td>
<td>pSB1C3-J00450</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FEATURES             Location/Qualifiers

terminator         3099..3139

        /label="B0012"

terminator         3011..3139

        /label="BBa_B0015"

rep_origin         261..875

        /label="pMB1 rep origin"

terminator         1042..1147

        /label="T0 terminator"

RBS              2279..2290
misc_feature 3106..3125

/label="B0012 Stem-loop"

misc_binding 2049..2070

/label="BioBrick prefix"

primer_bind 12..33

/label="SB-prep-3P-1 primer site"

primer_bind 2031..2054

/label="SB-prep_2Ea primer site"

CDS 1160..1829

/label="CmR"

CDS 2297..3002

/label="E1010"

misc_feature 3114..3117

/label="B0012 Loop"

misc_binding 1..21

/label="BioBrick suffix"

misc_feature 3022..3065
ORIGIN

1 tactagtagc ggccgctgca gtccggcaa a aaaggcaag gtgtcaccac cctgcccttt

61 ttcttttaaa cccaaaagat tacttcgctg tatgcaggct tcctcgctca cctgcActcgct

121 ggcgtcggtc gtggctgc ggcaagcggt atcagctcag tccaaagcgg gaatacgggt

181 atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaaggc agcaaaaggc
241 caggaaccgt aaaaagccg cgttgctggc gtttttccac aggtctccgcc ccccctgacga
301 gcatacaca aaatgcagct caagtcagag gatgccgaaac ccgacaggac tataagata
361 ccaagcgttt ccccttgaaa gctccctcgt gcgccttcct ctmccgacce ggggcctgac
421 cgataacctg tccgcttttc tcccttcgga aagcgtgggc gttttgctgca gtcagctg
481 tagtatctc agttcgggtg aggtgttgtc ctccaaagct ggtgtgttcg acgaaimccc
541 cgtcagccc gacgcctgcg ccctccgtga gctccctcct gttccgaccc tgccg
601 acacgaactta tcgcactcgg cagccggcc acgtaacagg attgccagaag ccagctgatgt
661 aagcggcgct acagagtctt tagaagtttg gcctcaactac gctacacta gaagaacagt
721 atttgtatcc tgctggcttc tgaagccggt gctccctcct gttccgaccc tgccg
781 atccggcaca caaaccaccc ctggttcggt gttttttttt gttgcaagcc acagattac
841 gcgcagaaaa aaggattatt tcgcctcggg tcggctgtttt ggcacgctca
901 gtgggagaaa aactcaactt aagggatta gttgccttcct ttataaaaa gatcctccac
961 ctagatcctt ttataatta aatgaagttt ttaatcact aataatgtat atgatgtaac
1021 ttggtctgac agctcaggtg ggtggatttc accaaattaa aacgcgcggcc gcacacggag
1081 ccttctgac agctcaggtg ggtggatttc accaaattaa aacgcgcggcc gcacacggag
1141 aagcgaagctc gatatakaat tggccttgc tggcgacttc atgcgcagtc tgggtgaatt
1201 cattagcatc tctgacgccg aagactaat ctccaaacgc actgaggaacc tgaatcgcca
1261 gcgcagtcag cactgctgct cttgcgttat aatatttgc ccatggtaaa acggggggcga
agaagttgtc catattgcc acgttaaat caaaactgcc cagggattgg
cagagacgaa aaacatattc tcaataaaacc cttaagggaa ataggccagg ttttcaccgt
aacacgcacc atcttgccaa tatatgtgta gaaactgcgg gaatactgctg tggtatccac
tccagagccga tgaaaacgct tcaataaacc ctttagggaa atag
gccagg ttttcaccgt
agaagttgtc catattgcc acgttaaat caaaactgcc cagggattgg
cagagacgaa aaacatattc tcaataaacc ctttagggaa atag
tatcccatat caccagctca cgcttttca ttgccatagc aaaaatgcgg tggatatccac
tcagagccgac aagaagttgtg ataagggccg gataaaactt gttcttatatttt tctttaacgg
gccagg ttttcaccgt
dagagctca ccaataagt tctttacgat gccatgggaa tatatctgat gttcctttca ttgccatagc aaaaatgcgg tggatatccac
tcagagccgac aagaagttgtg ataagggccg gataaaactt gttcttatatttt tctttaacgg
2401 tcgtccgtac gaaggtaccc agacgcctaa actgaaagtt accaaaggtg gtccgctgcc
2461 gttcgccttg gacatcctgt ccgccagatt ccagtacggt tccaaagctt acgtaaaca
2521 ccgggctgac atccgggactacctgaaact gtccttcccg gaaggtttca aatgggaacg
2581 tgttatgaca ttgcaagacg gtgtggtttgt taccgttacc cagacgtccct cccatgaaga
2641 cgggtggccc atcttcaaac tttaactgcg tggtaaaccg ttcgcctcgg acgcggccgt
2701 tatgcagaaa aaaaacctgg gttggaagc ttcaccgaaa cgataggtac ccggaagacgg
2761 tgctctgaaa ggtgaaatca aatgcgtctt gaaactgaaa gacgatgggt actacgacgc
2821 tgaagttaaa accacctaca tggctaaaaa accgggtcag cctgccggggcttttacac
2881 cgacatcaaa cttgcggatca cctcccaca aaaccgacacg ccttggggtttg aacagatcag
2941 acgtgcctga ggtctggcact ccacgggtgc tttaaaaaagatgtagtactc tgaagttaaa
3001 gctactagag ccaggcatca aataaaaacga aaggtcagct cgaagactgt ggcctttcgt
3061 tttatctgtt tgttctcggt gaacgtcttc tactagagtc aacagtgcct acctggcggtg
3121 ggcctttttc gcgtttata

//
C.2.2 pSB1C3-K592009 (Contains BCP Coding Sequence)

DEFINITION .

COMMENT ApEinfo:methylated:1

FEATURES Location/Qualifiers

misc_binding 670..690

/primer_bind 825..845

CDS 1..669

terminator 691..762

/primer_bind 2700..2723

rep_origin 930..1544

/primer_bind 2612..2631
primer_bind 2640..2666

CDS 1829..2498

terminator 1711..1816

misc_binding 2718..2739

primer_bind 81..102

primer_bind 681..702

ORIGIN

1 atgagtgtga tcgctaaaca aatgacctac aaggtttata tgtcaggcac ggtcaatgga
61 cactactttg aggtcgaagg cgatggaaaa ggtaagccct acgaggggga gcagacggta
121 aagctcactg tcaccaaggg cggacctctg ccatttgctt gggatatttt atcaccacag
181 tgtcagtacg gaagcatacc attcaccaag taccctgaag acatccctga ctatgtaaag

30
1321 gaggtatgta ggcggtgcta cagagttcctt gaagtgggtga cctaactacg gctacactag
1381 aagaacagta ttttgtatct gcgcctctgct gaagccagtt acccttcggaa aaagagtggg
1441 tagctcttga tccggcaaaaaaaaa aaccaccgcgc tggtagcggt ggtttttttttttgtttaaagca
1501 gcagattact gcagaaaaa aagagatcctca agaagatccttttgatctcttctacgt gcgctctgct gaagccagtt accttcggaa aaagagttgg
1561tgacgctcag tggaacgaaactacggtta agggattttt gtcagagatagatcagagtaaacttggtctgaca gctcgaggct tggattctca ccaataaaaa acgcgcgggcg
1621gatcttccacctagatcttttttaaaataaaaa atgaagttttttaaatcaatcctt aaatatatata
1681tgagtaaacttggtctgacagctcgaggct tggattctca ccaataaaaa acgcgcgggcg
1741gcacccgagctttctgaac aatgccagatg gaggctctgag gcggc
1801 caggagtccag cggagctcggctcagatccttaagttctacgt gcgctctgaac aatccagatg gagttctgaggtcattactg gatcttccacctgatagatcagagtaaacttggtctgaca gctcgaggct tggattctca ccaataaaaa acgcgcgggcg
1861gatcttcacc tagatcttttttaaaataaaaa atgaagttttttaaatcaatcctt aaatatatata
1921gaatcgcctacgcggacagcctataaacttca acgtcctgggatgtgattgcagtaac ccacacacatgcgctctgcctggtgattcact ccagagcgat gaaaacgttt cagtttgctcggtgtaacaagggtgaacact atcccatatcaccagctcac cgtctttcat tgccatacga aattccggat ggcattcatcaggcgggca agaatgtgaa taaaggccgg ataaaacttg tgcttatttttctttacggt ctttaaaaag gccgtataatcagctgaac ggtctggtta taggtacatt
2401 gagcaactga ctgaaatgcc tcaaaatgtt ctttacgatg ccattgggat atatcaacgg
2461 tggatatcc a.gtga.tttt ttctccattt tagcttccctt agctcctgaa aatctcgata
2521 actcaaaaaa tacgcceggt a.gtgaatcta tttcattatg gtgaagttg gaacctctta
2581 cg.tgcceggt caacgtcagt gccacctgac gtctaaagaa ccattattat catgacatta
2641 ac.etaaaaa ataggcgtat cacgaggcag aatircagat aaaaaaact ctagcttttc
2701 gctaaggagtg atttcctgaa ttcgcggccg cttctagag

//
C.2.3 pSB1K3-J00450 (RFP expression with Kanamycin Resistance)

LOCUS pSB1K3-J04450 3273 bp ds-DNA circular 21-APR-2015

DEFINITION .

FEATURES Location/Qualifiers

stem_loop 30..64

/label="Stem loop"

CDS 2431..3136

/label="E1010"

misc_feature 3240..3259

/label="B0012 Stem-loop"

primer_bind complement(157..176)

/label="Verification reverse (VR) primer binding site"

gene 2205..3273

/label="BBa_J04450"

misc_feature 1..21

/label="DNA: BioBrick suffix"

misc_feature 2183..2204
/label="Biobrick Prefix Standard (RFC10)"

terminator 3233..3273

/label="B0012"

cmisc_difference complement(1880..1880)

/label="Silent Mut: G→C mutation to remove XhoI site"

terminator 3145..3224

/label="B0010"

promoter 2205..2404

/label="R0010"

primer_bind 2067..2086

/label="Verification forward (VF2) primer binding site"

terminator 3145..3273

/label="BBa_B0015"

misc_feature 3248..3251

/label="B0012 Loop"

stem_loop 22..93

/label="E. coli his operon terminator"

misc_feature complement(2017..2060)
/label="repeat region misc"

misc_feature 3156..3199

/label="B0010 Stem-loop"

RBS 2413..2424

/label="B0034"

misc_feature complement(261..875)

/label="rep (pMB1) misc"

misc_feature 3176..3179

/label="B1010 Loop"

misc_feature 276..276

/label="ORI misc"

stem_loop complement(2143..2146)

/label="end of terminator"

stem_loop complement(2156..2175)

/label="Stem loop"

misc_feature 1..21

/label="Biobrick Suffix (RFC10)"

CDS complement(1097..1912)
ORIGIN

1 tactagtgc ggccgctgca gtccggcaaa aaagggcaag gtgtcaacac cctgcccttt
61 ttccttttaaa ccgaaagagat tacttcgcgt tatgcagget tctctgctca ctgactcgct
121 gcgctcggtc gttccgctgc ggccagcgggt atcagctcac tcaaaggcc ttaaaggttt
181 atccacagaa tcaggggata acgcagggaa gaacatgtgag caaaaagccc agcaaagggc
241 caggaaccgt aaaaaggccg cgttgctggc gttttttcacc agggctcgcgg cccctgacga
301 gcatacagaa aatcgacgct caagtccagag gtggcgaaaac ccgacaggac tataaagata
361 cccaggctttt cccctggaac gttccctcctgt gcggctcctcttt cttccagacc tcgcgcttac
421 cggatactct agttggtgcgt aggtgcttcgct tcctcaagctc ggtggcattg gcgtgtaagcc
481 taggtatctct agtttctccgt gttccagctgc ttccggcttcgc ccgctctcct gcggctttcctg
gggtgtggtcgt tccgtggtgc cttttgcgg gcgcggttttt tccggcgaac gcgatagttt
541 cgttcagccc gaccgctgag cccttatccgg taactatcgta tctgtgcacac cccgggttaag
601 acacgacctt tcggccactgg cagccagccat tggtaaacag attagcagag cgaggtatgt
661 aggccgggtct acagagttcat tgaagtgggtg gccttaacctgc gctacactac gaagaacagt
721 attttgtatct gcgcgtctgc tgaagccagt taccttcgga aaaaagtttg gtagctcttg
781 atccggcaga caaaccaccc cgggtcaggg gggtttagattttt cttgggcaag agcagattac
841 gcgcagaaaa aaggagctct aagaagatcc tttgcatttt tctacggggt ctgacggtca
901 gtggagacgaa aactacggtt aagggatttt ggtcaagccca tcctaaaaa ggttcacttc
961  ctagatcctt ttaaatta aaatgaagttt taaatcaatac taaagtatat atgagtaaac
1021 ttggtctgac agctcgagtc ccgtaaattc agcgtaatgc tgtgccagtg ttacaaccaan
1081 ttaaccaatt ctgattagaa aaactcatcg agcatcaaag gaaactgcaaa tttattcata
1141 tcaggattat caataccata tttttgaaaa aacgcgtttttct gtaatgaaggg agaaaactcag
1201 cggaggcagt tccataggat ggcaagatcc tgttctggtct cttgcgtat cc acctgttcca
1261 acatcaatac aacctattat ttcccctcgg ctaaataaa gtttatcaag tggaaatcac
1321 ccatgagtga cgactgaatcc cggtaaagat ggcaaaaagct tatgcatttc tttccagact
1381 ttgccacag cggcgccat caagctggtta ctaaaatcac tgcctcaaac ccaaccgtaa
1441 ttcattggtct attggccttg acgagacgaa aatagcggat cgtgtttaaa aggacaattta
1501 caaaccgaa tcgaatgcaaa cgggccgagg aacactgcaac ggcgtcaaac aatattttca
1561 cctgatatcg atatatcggg aatggtgctgt cttccgggat cgcagtgggtcg
1621 agtacagtcatcag cgggtcctgg gtcggtgttc gttgctggaa aggcataaat
1681 tcccctgacct gacctgctct gcgtgctgctg cttgctgtttt tccccgggat cgcagtgggtc
1741 ccatgtttca gaaacaactc tggcgcatcg ggcttcccat acaatcgata gattgtcgcat
1801 cctgattgcc gcgattggtcg ctccgctgcc caacagctctt aacgtctgggtaa atagcgttaa
1861 aaataatc gcgggctggga gcaagacgtt cccgtttgaa tatggctcat aacacccctt
1921 gtattactgt ttatgtaagc agaagctttt atgttctagt atgatatatct tttatctttt
1981 gcaatgtaac atcagagatt ttgagacaca acgtggtcctt gttgaataaa tgcgaactttt

38
gctgagttga aggatcagct cgagtgccac ctgacgtctta agaaaccatt attatcatga

cattaccta taaaaatagg cgtatcagga ggccagaaattt cagataaaaa aaatcctttag

cgttccccgcg cggtggccgga ttcattaatg cagctggcac gacaggtttc ccgactggaa

tctcccccgcg cggtggccga ttcattaatg cagctggcac gacaggtttc ccgactggaa

tttacaccttt atgttcccgg ctcgtatgttt gtgtggttaatt gtgagcggatat caaaatttca

cacatactag agaaagagga gaaatacgag atgggttctct ccgaagaaggt tataaaagag

tctgcctgtt tcaaatgcgtt ttaaacggtt ccgacgttaacg tcaacgagtt ccgaatcga

ggtaagggt aaggtcgtcc gtaacggaaggt acccagacgg ctaaatgga aagttacaaa

ggtggtccgc tgcgttccgc cggccacgtc ctgtccccgc agttcagta ccggtcpara

gcttcgttaa aaccccgagg cgatacttgc gactacctga aactgtactt ccggaaggt

ttaaatggg aacgtgttat gaactttcga gacggttgggt tttaccgt taccagaccac

tccctctgc cagagcggtgag tataacctgca aaaaaggtgt gatgtgatctg taccagac

tcgttcttt cgggttatgc gaaaaacc atgggttgag aacgctttcag cgaacgtatg

tacccgagag acggtcgtct gaaaggtgaa atcaaaatgc gtctgaaact gaaagcgtgt

gtcactacg acgctgaagt taaaaccacc tacatggcta aaaaaaccgggt tcaacggtcg

gttgcaataa aacccgacat ccaactggac atcaacctcc aacaacgaaga ctacaccatc

gttaacagt acgaaacgtgc tgaaggtcgt cactccacgg gtgttaata acgctgatatg
3121 tgcgtgtga gatcgctact agagccaggg atcaaaaa acgaaaggct cagtcgaaag

3181 actgggacct tcgttttattc tgttttttgt cggtgaacgc ttcctactag agtcactctg

3241 gtcacctccc ggggggacct ttttggtttt ata
C.2.4 Circular Assembly Product (in Kanamycin Backbone)

LOCUS pSB1K3-J04450 3273 bp ds-DNA circular 21-APR-2015

DEFINITION .

FEATURES Location/Qualifiers

stem_loop 30..64

/label="Stem loop"

CDS 2431..3136

/label="E1010"

misc_feature 3240..3259

/label="B0012 Stem-loop"

primer_bind complement(157..176)

/label="Verification reverse (VR) primer binding site"

gene 2205..3273

/label="BBa_J04450"

misc_feature 1..21

/label="DNA: BioBrick suffix"

misc_feature 2183..2204

/label="Biobrick Prefix Standard (RFC10)"
terminator      3233..3273

/label="B0012"

misc_difference complement(1880..1880)

/label="Silent Mut: G>C mutation to remove XhoI site"

terminator      3145..3224

/label="B0010"

promoter        2205..2404

/label="R0010"

primer_bind     2067..2086

/label="Verification forward (VF2) primer binding site"

terminator      3145..3273

/label="BBa_B0015"

misc_feature    3248..3251

/label="B0012 Loop"

stem_loop       22..93

/label="E. coli his operon terminator"

misc_feature    complement(2017..2060)

/label="repeat region misc"
misc_feature 3156..3199

/label="B0010 Stem-loop"

RBS 2413..2424

/label="B0034"

misc_feature complement(261..875)

/label="rep (pMB1) misc"

misc_feature 3176..3179

/label="B1010 Loop"

misc_feature 276..276

/label="ORI misc"

stem_loop complement(2143..2146)

/label="end of terminator"

stem_loop complement(2156..2175)

/label="Stem loop"

misc_feature 1..21

/label="Biobrick Suffix (RFC10)"

CDS complement(1097..1912)

/label="Kanamycin resistance marker CDS"
ORIGIN

1 tactagtagc ggccgctgca gtccggc aaaaagggcaag gtgtcaccac cctgcccc tttt
61 ttcttttaaaa ccgaaagat tacttcgagt tatgcaggct tcctcgctca ctgactgctt
121 gcgctcggtc gttegcggtgc ggccagcggt atcagctcac tcaaaagcgg taataggtt
181 atccacagaa tcaggggata acgcagga aaacatgtag gaacaaaggg aagcaaaaggc
241 caggaaccgt gaaaaagcgcc ggttgcgtgc gttttcccc aggcctccgc cccctgcacga
301 gcctacaaata atcagcgcgt caagtccagag gttgagaaac ccgagcttac tataagata
361 ccaggcgttt cccccctgga aacctccctgt gcgctctctct ctgccgaccc tgccgctttac
421 cgctacccgt tcgccttttc ccccttgggg aagcgtttgc gagttccata gcctacgcgt
481 taggtatcgt agtccggtgc ccttcagggt gcggtgggtgc gttttagctta cccagccccc
541 cggctcagccc gacgcgctgcg ccttatacgg taactatcgt tttgagctca acccgtaag
601 acacgactta tggcacttgg cagcagccac tggtaacagg attagcagag cgggtatatgt
661 aggcgggtct acagaaatta tgaagttgttc gcctactta gcctacacta gaagaacagt
721 atcctggtac tcgcgctgc gtaagccagc gtaattctga aaaaaagttg ttagttcgg
781 atccggc aaaaacaccg ctgttaggct gtttgggggg gtttgccagc acgcaggttac
841 gcgcgacaga aaagatctc aagaagatcc tttgacggtc tctacggttc ctagcgtcaca
901 gtggagaaag aacctagcgtt aaggatgttt ggtcatagga ttaaaaaa gatctttcag
961 ctagatccttg taataattaa aatgaagtttt taaatcaata taaatgata atagtagtaacc
ttggctcagc agctcgagtc ccgtaagtc agcgtatact tctgccagtg ttacaaccaaa

ttaaccaatt tggattgagaa aactctagcg agcatcaaat gaaactgcaaa tttattcata

tcaggattat caataccata ttttgaagaa ggccgatatg tgaatgaagg agaaactca

cgaggtcagt tcctatcaggt ggcaagatcc ttggbcgtcct gactcgtcca

acatcaatac acactattaa cttcccctcg tcaaaataaa gtttatcaag tgggaaactca

ccattggtg agcactgaatc cgctgaagat ggccagaagt tattcatttc tttcagact

ttttcaacag gccagcaatt acgctgtaac ctaaaatc acgcatcaac ccaacgttaa

ttcattcgtg atgctgcttg agcgagacga aattacgcat cgtgtttaaa aggacaatta

caaaagagaa tcgaatgcag tggggaagag gcgtcaac aatatcttaa

cctgaatcag gattatttgc taaaacggggt atggctgttt cccggggat cgcagtgggt

agtaacctga tcatcacagg agtaaggata aatgtcttga tttcggagaag aggcataaat

tccgtcagcc agtttagtct gccatctca cattctgtaag cttggcaaac gctacccctt

caggtttcag aaaaactctt tcggccatcc ggcttttcct acaatcgata gttgtgca

cctgattgc ccagctatat ccgagcccc cttatacctg ccacaatcgc ataattgcttt

gaatattaac gcggccctgga gcaagacgttt tcccgttgaa tatgctttt ctttttaaat

gtattactgt tttgtggtgc aggacttttt attttctgtg atgatatatt tttatctttg

gcaatgtaaac atcagagatt ttagacacaag acgtggcttt gttgaaataaa tcgaaccttt

gctgaggtta aggacagcct cgagcctcagc ctcacgtcagc aagaaccatt attatcatga

45
2101 cattaaccta taaaatagg cgtatcacga gccagaattt cagataaaa aaatccttag
2161 ctttcgctaa ggatgatttc tggaattcgc ggccgcttct agagcaatgc gcaaaccgcc
2221 tctcccgccg cgtggccgaa ttcatattaag cagcgtgctc cagagtttac ccgactggaa
2281 agcgggcagt gacgcaacgc caattaatgt tagttagctc actcattagg caccacaggc
2341 ttacacttta atgcgccggg cgcgtatgga cagctggcac gacaggtttc ccgactggaa
2401 ggcgggcagt gagcgcaacgc caattaatgt tagttagctc actcattagg caccacaggc
2461 ttcatgcttt tcaaagtctc tgattaaggt ttgcgtaacg tgcacagagtt cgaaatcgaa
2521 ggtgaaggtg aaggtcgtcc gcacgaaggt acccagacgc ctaaactgaa agttaccaa
2581 ggtggtccgc tgcggtccgc ttggacaccct cggtcccgcc agttccagta cgggtccaa
2641 gcttacgcta aacccggcgc tcacacccgc gctacctgaa gacggttgggt ttggataaag taccagac
2701 ttcctctggc aagcgttggt tatggtggtg tctggtcctc taccagac
tcccctgctc aagcgttggtc aaccggtcgc gatccatgtc cgggtccaa
2761 tccgaccgcc cggttcgtta gaaaaaccct atgggttggg aagccctgac ccggtccaa
2821 tgtcaggtcg aacacccggc tcacacccgcc gcacgttggg tctgtaatgc gagaagcgtt
2881 taccgggaag acgtggtctc gaaggtgagc atcggatcgc ctataatgc gaaagacggt
2941 ggtcactagc acgtggtctc gaaaaaccct acggttggtg aagcgttggg tctgtaatgc gaaagacgtt
3001 ggtgcttaaa aacccgcaaat aacggtgggtc actcctggtc ctaaattaccacaacggtac ccgtgtggggt
tgatcctttc gacagt attaacggt taccggtctg tctgtaatgc gaaagcgtt
3061 tggatcgtcag agccacgctc tgcgttcggt gtaaggtttc ccgctgtaata aacggttggg tctgtaatgc gaaagcgtt
3121 tgtgatgtag aacacccgtc gatcaggtggg tctgtaatgc gaaagcgtt
3181 actgggccct tcgtttttatc tgtgttttgt cggtaacgc tctctactag agtcacactg

3241 gctcaccttc ggtggtggcct ttctgcgttt ata

//
C.2.5 2-way Linear Product

LOCUS  2-Way Linear Product  1050 bp ds-DNA  linear  27-JUN-2015

DEFINITION .

FEATURES Location/Qualifiers

misc_feature  120..141

/label="Biobrick Prefix Standard (RFC10)"

misc_feature  1..13

/label="exda22  5'overhang"

misc_feature  1..381

/label="2WL-Promoter"

misc_feature  354..1049

/label="2WL-BCP"

misc_feature  354..381

/label="2WL-overlap"

promoter  142..341

/label="R0010"

primer_bind  448..469
/label="pSBamil_for"

stem_loop complement(80..83)

/label="end of terminator"

CDS 368..1036

/label="amilCP"

misc_feature 354..365

/label="exda12 5'overhang"

stem_loop complement(93..112)

/label="Stem loop"

misc_feature complement(371..381)

/label="exda23 5'overhang"

RBS 350..361

/label="B0034"

ORIGIN

1 TGAAGGATCA GCTTCTAAGA AACCATTATT ATCATGACAT TAACCtataa aaataggcgt

61 atcacgaggc agaatattcag ataaaaaaaa tccttagctt tcgctaagga tgatttctgg

121 aattcgccggc cgcttctaga gcaatacgca aaccgcctct cccgcgcgt tggccgattc

181 attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa
241 ttaatgtgag ttagctcact cattaggcac cccaggtttt acaactttatg cttcgggttc
301 gtatgttttgtg tggattgtg agcggataac aatttaCAC ATACTAGAGA AAGAGGAGAA
361 ATACTAGATG AGTGTGATCG CTAAACAAAT GACCTACAAG Gtttatgtg caggcacggt
421 caatggcac acattttgagg tccaagggca tggaaaaggt aagccctacg aggggagca
481 gacggttaaag ctcactgtca ccaagggagg acctctggca tttgcttggg atatatttatac
541 accacagtgt cagtacggaa gcatcacttt cacaagttac cctgaagaca tccctgacta
601 tgtaaqgacg tcattccggg agggctatak atgggagagg atcatgaact ttgaagatggg
661 tgcagtggtg actgtagcag atgattccag catccaaagc aactgattttca tctaccatgt
721 caagtctctct gtctttgaaict ttcctccccaa tggacctgtc atgcagaaga agacacaggg
781 cttgggacccc aacactgagc gtctcttttg cagagatgga atgctgtcgag gaaacaactt
841 tatggtctctg aagttgaaag gaggcggtca ctatcttgtg gatatttttaa ctactttaca
901 ggcaagagag cctgtgaaga tgcctaggt ta tcactatgtt gaccgcaaac tggatgtaaac
961 caatcacaac aaggattaca ttcgggttga gcaqtgtaa atttcattG CACGCAAACC
1021 TGTGGTCGCC TAATAATACT AGAGCCAGGC

//
C.2.6 pUN Plasmid

LOCUS      T5-J23100-modBCP-GERC 4348 bp ds-DNA circular 10-FEB-2015

DEFINITION .

FEATURES Location/Qualifiers

  terminator      complement(1813..1866)
  /label="putative terminator"

  misc_RNA     244..250
  /label="cleaved 5'UTR"

  promoter     complement(3771..3799)
  /label="AmpR promoter"

  misc_feature  1275..1314
  /label="UNS9"

  promoter     383..417
  /label="J23100 (strong promoter)"

  terminator    3910..4067
  /label="rrnB terminator"

  misc_feature  2086..2915
/label="p15A origin"

**terminator** complement(4255..4295)

/label="Sal2 terminator"

**misc_binding** 2754..2765

/label="Bsp24I site"

**misc_binding** 1907..1912

/label=""

**terminator** 4182..4226

/label="RNAI terminator"

**misc_feature** 244..318

/label="RiboJ"

**primer_bind** 119..158

/label="UNS7"

**misc_feature** complement(1125..1164)

/label="UNS5"

**terminator** complement(1441..1483)

/label="L17 terminator"

**CDS** complement(2920..3729)
/label="KanR"

terminator 1328..1381

/label="R2-17 terminator"

terminator 41..118

/label="B0055 Terminator (NCBI:AY643800)"

misc_feature 319..358

/label="UNS6"

enhancer 359..382

/label="Full UP Element"

misc_feature 204..243

/label="UNS4"

misc_feature 1..40

/label="UNSX"

promoter 159..203

/label="T5 promoter"

CDS 456..1124

/label="modified amilCP CDS"

ORIGIN
ccaggataca tagattacca caactcggag cccttcacc aaggaatatt cagcactc
61 cccgtgccga agaaagccc acccgtaag gtgacgagt gatgtgatag ctacgtaaca
121 agacgctgcc tctgacattt ccgctactga actactcgTA ATAAAAAAATT TATTTGCTTT
181 GTGAGCGGAT AAACAATTATA ATActagacct cctgccagca atagtaagac aacacgcaa
241 gtcaagctgc acgggatgtg ctttccggtc tgatgagttc gtggagccga aacagctct
301 acaataatt ttgttttaaCT CGTTCGCTGC CACCTAAGAA TACTCTACGG TCACATACGG
361 AAAATTTTTT TAAAAAAACTTTTACAGC TAGCTCAGTC CTAGGTATTA TGCTAGCTAC
421 ATCCAAATCGC CACTCAATA AAGGACACGT AAAAATGAG CGTTATCGCT AAAAAATAGA
481 CCTACAAGGT TTATATGTCA GGCACCGTCA ATGGACACTA CTTTGAAGTC GAAAGCGATG
541 GAAAGGTAAT GCCCTACGAG GGGGACGAG CGGTAAAGCT CACTGTCACC AAGGGCGGAC
601 CTCTGCCATT TGCTTGGGAT ATTTTTATAC CACAGTGCTCA GTACGGAAGC ATACCATTCA
661 CCAAGTACCC TGAAGACATC CCTGACTATG TAAAGCAGTC ATTCCCGGAG GGCTATACAT
721 GGGAGAGGAT CATGAACCTT GAAGATGGTG CAGTGTGATC TGTCAGCAAT GATTCCAGCA
781 TCCAAGGCAA CTGTTCATAC TACCAGTCTCA AGTCTCTGTT TTTGAACCTT CCTCCCAATG
841 GACCTGTCAT GCAGAAAGAC ACACAGGGCT GGGAGCCCAA CACTGAGCCT CTCTTGACAC
901 GAGATGGAAAT GCTGCTAGGA AAACAATTATA TGGCTCTGGA GTTGAAGGGA GCCGGTCACT
961 ATTTGAGTGA ATTTAAAAC ACTTACAAGGCAAAGAGCC TGTGAAGATG CCAGGGTATC
1021 ACTATGTTGA CCGCAAAGCT GATGTAACCA ATCACAACAA GGATTACACT TCGGTTGAGC
1081 AGTGTGAAAT TTCCATTGCA CGCAAAACCTG TGGTCGCCTA ATAAACTCTAA CGGACTTGAG
1141 TGAGGTTGTA AAGGGAGTTG GCTCCTCGGT ACCAATTCCC AGAAAAAGGG CCTCCCGAAAA
1201 GGGGGGCTTT TTTTCGTTTTT GGTCGGGGGAG ACCAGAAAACA AAAAAAGGCC GCGTTAGCGG
1261 CTTTCATATAA TTGGGTCCCT TATCATCTGG CGAATCGGAC CCACAAGAGC ACTGtgcaga
1321 aggacacggt taatactagg cctgcttgct ggttaatcgc acgaggcctt ttttttggg
1381 ggagagggaa gtcgcaaaa aactaacctt tgaaattcgc tctccaccac atcagctctg
1441 aagcaacgta aaaaaaccgg ccccggcggg tttttttata cccgtagtat ccccaacttat
1501 ctacaatagc tgtccttaat taatctagaa aaatgaaggg aagttctctat acctctctag
1561 gaataggac ttctataggg agtcgaataa gggcgacaca aaggatcct aaatgcata
1621 ataaactctg ataacacttt atagttgta ttatatattt gatcgcgata
1681 attttgatat caaaaactga ttttcccttt attattttcg agatttttt ccttaattct
1741 cttaacataa ctgtagaatg tctatataca aaaaatcata aataatagat gaatagttta
1801 attataggtg ttcatcaatc gaaaaagcaa cgtatctttatatataaaggg ctggcagttt
1861 ctctatataa aggtaataa atctctatat atcaagaaaa ggtgagggct cccttaaata
1921 ttctgacaaa tgctctttcc ctaaactccc cccataaaaa aacccgcgg gctgggattt
1981 tacggtatct ggcggattaac gattactcgt tatacaaaa gcccaggatg cctggcagtt
2041 ccctacttetc ggcgctgcgc tcggctgcttc ggctgcgggga cctcagctcg acgggagtgt
2101 atactggcatt actatgttgg cactgatgag ggtgtcagtg aagttctctca tgtggcagga
2161 gaaaaaaggc tgacccggtg cgtcagcaga atatgtgata caggatatat tccgcttctct
2221 cgctcactga ctcgctacgc tcggtcgttc gactgcggcg agccgaaatg gcttaacgaac
2281 ggggccggaga tttcctggaa gatgccagga agatactttaa cagggaagtg agagggccgc
2341 ggcaagccgc tttttttctata ggtcctgccc ccttgacaag ctctcagcact ccctgagcgc
2401 aatacagtgcc ggcggaaccc cggacagacct ataaagatacggacggtttc cccctggcgc
2461 ctccccctcggt cgcgtctcctgt tctcttccttt cgggtttacc ggttgctatc cgcgtttatg
2521 gccgctcgttg ttcctaccca ggcctcagac tgcggtcctgg gtaggcagta gcgtcagcgc
2581 tggactgtatat ccaccgaaacc ccctgcagt cgcgacctgc gcgccttcatc gccataactgc
2641 gcggcggttct cggcgctcttc aacagcgctg tggactgtat gcacgagcta ggcagcgcctt
2701 gatttagagg agtaacctct gaaagttactgc gcggggttag gccttaactt cagacaagat
2761 tttggtgact gcggctctttc aacagcgctg tggactgtat gcacgagctt ggcctcagcgc
2821 acctctgcgt cgtcctcttc gcggctctgc agttactgcc gcggacggtta cccgaactct gcgggcttctc
2881 agaccaaac gatctcaaga agatcatctt attaagcttt tagaaaaacga gatctcagcgc
2941 caaatgcacttg cactacctg tattactcttt ccatatagttt cagacgcac agtgcagctt
tttacttact gcggcttctct ctcgcttttc aagctcagcgc gggcggccctt gctttatttt
cagaggagag aactccagtc gcgtcagcact gggcggcctt ctcgcttttc aagctcagcgc
gtttggact gcgtctctttc aacagtcttgc gcggcttctct ctcgcttttt cagagcagcag cggcggcctt
gtttggact gcgtctctttc aacagtcttgc gcggcttctct ctcgcttttt cagagcagcag cggcggcctt
gtttggact gcgtctctttc aacagtcttgc gcggcttctct ctcgcttttt cagagcagcag cggcggcctt
gtttggact gcgtctctttc aacagtcttgc gcggcttctct ctcgcttttt cagagcagcag cggcggcctt
3241 atcactgcga tcaaccaaac cgttattcat tcgtgattgc gcctgagc
3301 gcgatcgctg tttaaaggac aattacaaac aggaatcgcag tgcaaccggc gcaggaacac
3361 tgccagcgcga tcaacaatat ttccacctga atcaggatat tcttcataaatctggaacgc
3421 tgttttttcgg ggtctgcgtg ttaaaaaggac aattacaaac aggaatcgcag tgcaaccggc gcaggaacac
3481 cttgatggttc ggaagtgccga taatattcctg tgcacatcga tcttcgtgtgc aagctggaacac
3541 aacatcattg gcaacgctac ctttgccatg ttgcagaaac aactctggcgc cggactgtttt
3601 ccacatacaag ctagatgattc tggcgccgca tttcagaaac aactctggtcgc cgcgatgatg
3661 ccatataaa tcagcatcca tgtagaaatt taatgcgcgc ctgcaggtttt cccgtgatgga
3721 atggctcatat tttctccttt ttcaataatta ttagaagctcat tttcttcgctg gcgttttata
3781 tagccggatac atatgttgaat gttatttagaa aaataaaacaa ataggggtca gttgataac
3841 caattaacca attctgaaca ttatcgccag gcctttataa gttgatatag gctcataaca
3901 ccctgttttt gcctgccgcgc acgtgcgtgga tgggctccacc tgaccccatgtg cggcgttctc
3961 aagttgaacgc cctgacggcc gatgggttagt ggggactcct ccacgccggaa gtaggaactc
4021 gggcgcgctac aaataaacaag aaagggttcg tcggaaagcac ggccttttgc cccgggctaa
4081 cgggctgctg tgcgctttttct gtgctgctttgtgagaaagttgc ctttctctca gaaagtataag
4141 gaaacctgtgca agggggggttc gacgccgcccc aaaaaaaag aatcggccaa acatccgcaac
4201 gtgtttgtgg gcctttgttc gacaatgctc gtaagcttgca tcaataagcag
4261 taaaaaaacc ggccatggcc cggtttttta atgagaattt ttacctgtag cagccggccac
4321 catccggca aagaagcatac aagcttt

//
### C.2.7 Pfu expression plasmid for ex vivo PCR

**LOCUS**  
PUN-rhaBp-Pfu  
5887 bp ds-DNA  
circular  
27-JUN-2015

**DEFINITION**

**FEATURES**  

<table>
<thead>
<tr>
<th>Location/Qualifiers</th>
<th>Start/Stop</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10_signal</td>
<td>280..285</td>
</tr>
<tr>
<td>/label=&quot;-10 region&quot;</td>
<td></td>
</tr>
<tr>
<td>primer_bind</td>
<td>119..158</td>
</tr>
<tr>
<td>/label=&quot;UNS7&quot;</td>
<td></td>
</tr>
<tr>
<td>RBS</td>
<td>302..335</td>
</tr>
<tr>
<td>/label=&quot;synRBS (60K TIR)&quot;</td>
<td></td>
</tr>
<tr>
<td>terminator</td>
<td>complement(2980..3022)</td>
</tr>
<tr>
<td>/label=&quot;L17 terminator&quot;</td>
<td></td>
</tr>
<tr>
<td>CDS</td>
<td>complement(4459..5268)</td>
</tr>
<tr>
<td>/label=&quot;KanR&quot;</td>
<td></td>
</tr>
<tr>
<td>protein_bind</td>
<td>159..180</td>
</tr>
<tr>
<td>/label=&quot;CRP-cAMP binding site&quot;</td>
<td></td>
</tr>
<tr>
<td>misc_binding</td>
<td>4293..4304</td>
</tr>
</tbody>
</table>

59
/label="Bsp24I site"

terminator 2867..2920

/label="R2-17 terminator"

misc_feature 1..40

/label="UNSX"

misc_feature 3625..4454

/label="p15A origin"

terminator complement(5794..5834)

/label="Sal2 terminator"

misc_feature 2814..2853

/label="UNS9"

-35_signal 259..264

/label="-35 region"

terminator 41..118

/label="B0055 Terminator (NCBI:AY643800)"

misc_signal 292..292

/label="TSS"

misc_feature 1221..1262

60
/label="gBlock overlap"

protein_bind 244..260

/label="RhaS binding site"

promoter 159..301

/label="rhaBADp"

misc_feature complement(2664..2703)

/label="UNS5"

CDS 336..2663

/label="PfuDNAPol"

terminator 5721..5765

/label="RNAI terminator"

protein_bind 189..210

/label="CRP-cAMP binding site"

promoter complement(5310..5338)

/label="AmpR promoter"

terminator 5449..5606

/label="rrnB terminator"

misc_binding 3446..3451
misc_signal 264..269

protein_bind 211..227

terminator complement(3352..3405)

ORIGIN

1 ccaggataca tagattacca caactccgag cccttcacc aaggaatatt cagcaatttg
ccccgtgccga agaaaggccc acccgtgaag gtgagccagt gagttgattg ctacgtaaca
121 agacgctggc tctgacattt ccgctactga actactcgcg gtgagcatca catcaccaca
181 attcagcaaa ttgtgaacat catcacgttc atctttccct gttgcccaat ggcccatttt
241 cctgtcagta acgagaaggt cgcttattca ggcgcttttt agactggtcg taatgaaatt
cCGAAGAAGG CAAACCGGTG ATCCGTCTGT TCAAAAAAG AATGGTAAA TTCAAGATAG
301 cAAATAAGCA AGCAAAACCT ATAAAGGAGTC CGCACATGAT CCTGGATGTC GATTACATCA
361 CCGAAAGAAGG CAAACCGGTG ATCCGTCTGT TCAAAAAAG AATGGTAAA TTCAAGATAG
421 AACACGATCG CACCTTTTCG CTTATATTT ATGCACCTGTT ACGCGATGAC AGCAAAAATCG
481 AAGAAGTCAA AAAGATCACC GGCGAACGCC ACGGTAAAAT CGTGCGCATT GTGTGTTTG
541 AAAAAAGTGGGA GAAAAAATTC CTGGGCAAGC CAATTACTGT CTGGAAGCTC TACCTGGGAAC
ATCCGCAGGA TGTGCCGACG ATTCGTGAAA AGGTGCGTGA GCATCCCGCG GTGGTGGATA
TTTTTGAATA TGACATTCCG TTTGCCAAAC GGTATCTGAT TGATAAAGGG CGACATGAA AATGAAGCAA
GGCAATACGT AGAGGAACTG AAAATTCTGG CATTTGATAT CGAAACGCTG TATCACGAAG
GTGAGGAGTT CGGCAAAGGC GCAGTATCATCA TGAGCTGCTA CGGTGATGAA AATGAAGCAA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
AAGTAATTAC CTGGAAAAAC ATCGACTTGC CGTATGTGCA GTTTTTCTTC TGACACCGCG
AAATGATTAA ACGCTTTTTA CGGATTATTC GCGAAAAGGT GGGAGACATC ATTTGTCACCT
ACAACGGGTGA CAGTTCGAC ATCCCATTTCT ACCTTTATCA TGTTATCAGC GCCAGATCA
ACCTGCAAC ACATACCCCTC GAGGCGGTTT ATGAAGCGAT CTTCGGTACT CCGCAAGATA
1681 TTATTCCTTC GCTGCTGGGG CATCTGCTGG AAGAGCGGCA AAAAATTAAA ACGAAAATGA
1741 AAGAAACACA AGATCCGATT GAGAAGATTT TGCTCGATTA TCGCCAGAAA GCAATAAAAAT
1801 TGCTGGCGAAA CAGCTTCTAC GCCTGGGGCC GCAAATATAT TGAACCTGTA TGAAAGAGGC
1861 AGTGCAGCAGA AAGTGTCACT GCCTGGGGCC GCAAATATAT TGAACCTGTA TGAAAGAGGC
1921 TGAAAGAGGA GATAAAGGAA AAGGTAAGA AGAAAGACTG TGAATTTGTC AAGTATATCA
1981 TTCCGAGCAG TGGAGCGAGA TGGCAAGGGA TCGCCAGAAA GAAATTAAGA AGAAAGACTG
2041 TTCCGAGCAG TGGAGCGAGA TGGCAAGGGA TCGCCAGAAA GAAATTAAGA AGAAAGACTG
2101 TTCCGAGCAG TGGAGCGAGA TGGCAAGGGA TCGCCAGAAA GAAATTAAGA AGAAAGACTG
2161 TAGAGATCGT TCCTCGCCGTA TGGAGCGAGA TGGCAAGGGA TCGCCAGAAA GAAATTAAGA
2221 AAACCATCTT AAAACACGGC GACGTTGAGG AAGCCGTTTG CATTGTGAAG GAAGTGATTTC
2281 AAACCATCTT AAAACACGGC GACGTTGAGG AAGCCGTTTG CATTGTGAAG GAAGTGATTTC
2341 GATCCGCTGCA GCAAATCCAAA CCGATGGGGC CGCATGTCGC AGTGGCAAAA AAGCTGGCGG
2401 CAAAAGGTGT GAAAATCAAA CCGGGCATGG TGATCGGTAA TATCGTACTG CGTGGTGATG
2461 GCCCAATCGT AATTCCTGAG ATACTGCGTC AAGAATACGA CCCGAAGAAA CATAAATATG
2521 ATGCGGATAT TTACATTGAA AACCAGGTTC TGCCGGCGGT ATTACGTATT CTGGAGGGCT
2581 TTGGCCTACCG TAAAGAGGAT CTGCGCTATC AAAAAACCCG TCAGGTGGGA CTGACGAGCT
2641 GGCTGAATAT CAAAATACCG TAACTCTAAC GGACTTGAGT GAGGTTGTAA AGGGAGTTGG
2701 CTCCCCGTA CCAAATTCGCA AAAAAAGGCA CCGCAGGCGT TGGGGCTTTTT TTTGGTTTTG
2761 GTCCGGGAGA CCAGAAACAA AAAAAAGCGG CGTTAGCGGC TTCAATAAT TGGGTTCCCTT
2821 ATCATCTGGC GAATCGGACC CACAAGAGCA CTGtctgaga agcaacgttaa aaaacccgc
2881 ctgcttgctg gaatcggacc cacaagagca ctgactgaga ggacacggtt aatactaggc
2941 actaaccttt taaatactga ctccaccaca tcaagctctgga agcaacgttaa aaacccgc
3001 cccgcgcgggt ttttttatac ccgtagtatc cccacttata tcaatagct gcctttaatt
3061 aatctagaaa aatgaagggga agttcctata ctttctagag aataggaact tctatagggag
3121 gtcgaataag ggcgacacaa aaggtattctaat aatgcataa taaatctgta taaatacttta
3181 tagtttgatat tatttttatt tatattcttg acatgtatagac ttgtataatc aaaacactgat
3241 tttccctttta taatttttsga gattttatat tattaatttctattaataaactc tagaatatttt
3301 gttatataaa aaatgataaa ataataagtgg aatagtttaaa ttataggtgtg tcatcaatcg
3361 aaaaagcaac gtatcttttt taaagtgcgt tgttttttttc tcatatttataa gtttaatatt
3421 ttctcatata tcaagcacaag tgacaggcgc cccttaaatatt tctgacatcg gccctttccc
3481 taacctcccc xacataaaaa acccgcgcagc gcggtttttt actgtatatg cggattaacg
3541 attactcgtt atcagaaacgg ccaggatgc ctggcagtct cctactctcg cgctgcgctgt
3601 cggtcggtcg ctgcgcggag atcgacgcta cgggtggtgta tactaattata tcttgttgaac
3661 actgtagtgg atgctccgatg aagaatttcat gtgccaggg agaaaaagct gcaccgggtgc
3721 gtcacgcagaa tatgcctgataagatatgc cctgctcttc gcgccttgcgc tcgctcgcgtct
3781 cggctggttcg actgcgcggga cggaaatgg cttacgaacg gggcggagat ttctggaag
atgccagga gatacttaac aggggaatga gagggccggt gcaaagcggt tttccatag

gctccgcccc cctgacaagc atcacgaaat ctagcctca aatcagttgt ggcgaaaccc
gacaggacta taaagatacc agggcgttcc cccttggccg tccctcgtgc gctctcctgt
tcctgccttt cgggttaccg gtgctcatcc gctgttatgg ccgcgttggt tctattccac
gctgacact cagttcgggg taggcagttc gcttcaagct ggtctgtatg cacgaaaccc
cggttcagtc cgacccgtgc gccttatcct gcgtctatcg tcttgatccc aaccggaaa
gacatgcaaa agcaccactg gcagcagcca ccctggcggc tccctcgtgc gctctcctgt
tcctgccttt cggtttaccg gtgtcattcc gctgttatgg ccgcgttggt tctattccac
gctgacact cagttcgggg taggcagttc gcttcaagct ggtctgtatg cacgaaaccc
cggttcagtc cgacccgtgc gccttatcct gcgtctatcg tcttgatccc aaccggaaa
gacatgcaaa agcaccactg gcagcagcca ccctggcggc tccctcgtgc gctctcctgt
tcctgccttt cggtttaccg gtgtcattcc gctgttatgg ccgcgttggt tctattccac
gctgacact cagttcgggg taggcagttc gcttcaagct ggtctgtatg cacgaaaccc
cggttcagtc cgacccgtgc gccttatcct gcgtctatcg tcttgatccc aaccggaaa
gacatgcaaa agcaccactg gcagcagcca ccctggcggc tccctcgtgc gctctcctgt
tcctgccttt cggtttaccg gtgtcattcc gctgttatgg ccgcgttggt tctattccac
gctgacact cagttcgggg taggcagttc gcttcaagct ggtctgtatg cacgaaaccc
cggttcagtc cgacccgtgc gccttatcct gcgtctatcg tcttgatccc aaccggaaa
gacatgcaaa agcaccactg gcagcagcca ccctggcggc tccctcgtgc gctctcctgt
tcctgccttt cggtttaccg gtgtcattcc gctgttatgg ccgcgttggt tctattccac
gctgacact cagttcgggg taggcagttc gcttcaagct ggtctgtatg cacgaaaccc
cggttcagtc cgacccgtgc gccttatcct gcgtctatcg tcttgatccc aaccggaaa
gacatgcaaa agcaccactg gcagcagcca ccctggcggc tccctcgtgc gctctcctgt
tcctgccttt cggtttaccg gtgtcattcc gctgttatgg ccgcgttggt tctattccac
gctgacact cagttcgggg taggcagttc gcttcaagct ggtctgtatg cacgaaaccc
cggttcagtc cgacccgtgc gccttatcct gcgtctatcg tcttgatccc aaccggaaa
gacatgcaaa agcaccactg gcagcagcca ccctggcggc tccctcgtgc gctctcctgt
tcctgccttt cggtttaccg gtgtcattcc gctgttatgg ccgcgttggt tctattccac
gctgacact cagttcgggg taggcagttc gcttcaagct ggtctgtatg cacgaaaccc
cggttcagtc cgacccgtgc gccttatcct gcgtctatcg tcttgatccc aaccggaaa
gacatgcaaa agcaccactg gcagcagcca ccctggcggc tccctcgtgc gctctcctgt
tcctgccttt cggtttaccg gtgtcattcc gctgttatgg ccgcgttggt tctattccac
gctgacact cagttcgggg taggcagttc gcttcaagct ggtctgtatg cacgaaaccc
cggttcagtc cgacccgtgc gccttatcct gcgtctatcg tcttgatccc aaccggaaa
gacatgcaaa agcaccactg gcagcagcca ccctggcggc tccctcgtgc gctctcctgt
tcctgccttt cggtttaccg gtgtcattcc gctgttatgg ccgcgttggt tctattccac
gctgacact cagttcgggg taggcagttc gcttcaagct ggtctgtatg cacgaaaccc
cggttcagtc cgacccgtgc gccttatcct gcgtctatcg tcttgatccc aaccggaaa
4921 ttcacctgaa tcaggatatt ctctctaatac ctgggaacgct gtttttccgg ggatcgcagt
4981 ggtgagtaac catgcatcat caggagtacg gataaaatgc ttgatggtcg gaagtggcat
5041 aaatccggtc agccagttta gtctgaccat etcatctgta acatcattgg caacgctacc
5101 ttgtgccatgt ttcaagaaaca actctggcgc atcgggcttc ccatacaagc gatagattgt
5161 gcgcacctgat tgcgacgagc catttatac ccataaacaactcatc catccacat
5221 gttggaattt aatcgccgacg tcgactgcttc ccgttaataa tggctgctat tctctctttt
5281 tcaatattat tgaagctatt atcaggtta tttgctctatg agcggatataa tatttaatg
5341 tatttagaaa aataaaacaa taggggtcag tggtaaacc aattaaccaa ttctgaacat
5401 tatcgcgagc ccatttatac ctgaatatgg ctcataacac cccttgtttg cctggcggca
5461 gtagcgaggt ggtcccaacc gaccccatgc cgaactcaga agtgaaacgc cgtagcgccg
5521 atggtagatgt ggggactccc catgcgagag tagggaactg ccaggcatca aataaaacga
5581 aaggctcagt cgaaagactg ggcctttcgc ccgggctaat taggggggtt ccgccctatt
5641 cgactctata gggaagttcc tattctctag aaagtatagg aacttctgaa ggggggctcg
5701 agcggccgca aaaggaagaag atccggcaca caaaccaccg ttggtagcg gggttttttt
5761 gttttggatcg acaatcttcg taagcgctat cattaagctt aaaaaacgg gcgaatgcc
5821 gttttttttaa tgagaatttt tacctgtcgt agcggccacc atccggcaca gaagcatcaca
5881 aggcttt

//