

DEVELOPMENT OF A PROLYL ENDOPEPTIDASE EXPRESSION SYSTEM IN
LACTOBACILLUS REUTERI TO REDUCE THE
MANIFESTATION OF CELIAC DISEASE

A Thesis

presented to

the Faculty of California Polytechnic State University,

San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Biology

by Kara Lynn Jew

July 2019

© 2019

Kara Lynn Jew

ALL RIGHTS RESERVED

COMMITTEE MEMBERSHIP

TITLE: Development of a Prolyl Endopeptidase
Expression System in *Lactobacillus reuteri* to
Reduce the Manifestation of Celiac Disease

AUTHOR: Kara Lynn Jew

DATE SUBMITTED: July 2019

COMMITTEE CHAIR: Michael Black, Ph.D.
Professor of Biology

COMMITTEE MEMBER: Pat Fidopiastis, Ph.D.
Professor of Biology

COMMITTEE MEMBER: Ken Hillers, Ph.D.
Professor of Biology

ABSTRACT

Development of a Prolyl Endopeptidase Expression System in *Lactobacillus reuteri* to Reduce the Manifestation of Celiac Disease

Kara Lynn Jew

Celiac Disease (CD) is an autoimmune disorder that emerges due to the ingestion of gluten, a protein found in a variety of common grains such as wheat, rye, and barley. Approximately 1 in 100 individuals in the US suffer from CD, making it the most commonly diagnosed gastrointestinal disorder (Ciclitira et. al., 2005). These proline-rich gluten peptides are resistant to proteolysis and accumulate in the duodenum of the small intestine. Once in the duodenum, these peptides illicit an autoimmune response resulting in villous atrophy. Current treatment for CD requires a rigorous adherence to a gluten-free diet. Nevertheless, gluten-containing grains are ubiquitous in the western diet, so accidental exposure to gluten remains as a persistent threat.

The approach of this project centers on genetically engineering a strain *Lactobacillus reuteri* to secrete a *Myxococcus xanthus* prolyl endopeptidase (PEP), an enzyme that hydrolyzes a peptide bond adjacent to an internal proline residue. The data from this study revealed that recombinant *M. xanthus* PEP purified from *E. coli* was effective in degrading Suc-Ala-Pro-pNA, a chromogenic substrate containing an internal proline residue. When introduced into a *L. reuteri* expression vector, mutations accumulated in the vector over the course of 5 days. These data suggested that toxicity was possibly associated with *M. xanthus* PEP and the amyE signal peptide.

ACKNOWLEDGMENTS

Dr. Michael Black has impacted my life more than I ever would have expected. On an academic level, he has pushed me to become a more independent researcher and has given me a more eager desire to understand the world around me. He has taught me valuable life lessons that will stick with me for the rest of my life and supported me in my eleventh hour. I hope there is not an alternative universe where our paths do not cross. I would like to further thank my committee members, Dr. Ken Hillers and Pat Fidopiastis, for their support throughout my spent at Cal Poly. The Black lab (Ryan Bitter, Reece Erickson, Jared Hallows, and Joseph Rainaldi) has been the best support group that I am eternally grateful for. Their willingness to provide assistance whenever I was in desperate need has been of immense value to me, and the group dynamic we had will be one that I will always cherish.

TABLE OF CONTENTS

| | Page |
|--|------|
| LIST OF TABLES | ix |
| LIST OF FIGURES | x |
| CHAPTER | |
| 1. CELIAC DISEASE BACKGROUND | 1 |
| 1.1. What is Celiac Disease? | 1 |
| 1.2. The Role of Gluten in CD | 2 |
| 1.3. The Immune Responses Involved in CD | 5 |
| 1.4. Autoimmune Responses Involved in CD | 7 |
| 1.5. Diagnosis of CD | 9 |
| 1.6. Current Treatments for CD | 10 |
| 1.7. Study Goal | 12 |
| 2. THE ASSESSMENT OF THE CLEAVAGE ACTIVITY OF PROLYL AMINOPEPTIDASES AND PROLYL ENDOPEPTIDASES ON PROLINE- CONTAINING SUBSTRATES | 13 |
| 2.1. Introduction | 13 |
| 2.1.1. Serine proteases | 13 |
| 2.1.2. Prolyl Endopeptidases and Prolyl Aminopeptidases | 14 |
| 2.1.3. Study Goal | 15 |
| 2.2. Materials and Methods | 15 |
| 2.2.1. Strains and Growth Conditions | 15 |
| 2.2.2. Molecular Techniques | 16 |

| | |
|--|----|
| 2.2.3. Optimizing Prolyl Endopeptidase (PEP) and Prolyl Aminopeptidase (PAP) fSequences Via Gene Synthesis..... | 17 |
| 2.2.4. Optimizing <i>L. reuteri</i> PAP Sequence Via PCR Stitching..... | 18 |
| 2.2.5. Construction of pET30 Vectors | 23 |
| 2.2.6. Transformation into <i>E. coli</i> Strains TOP10, MC1061, and BL21(DE3) | 25 |
| 2.2.7. Recombinant Protein Expression and Purification | 26 |
| 2.2.8. SDS-PAGE and Western Blots..... | 27 |
| 2.2.9. Enzyme Activity Assay | 28 |
| 2.3. Results..... | 29 |
| 2.3.1. Optimizing PAP and PEP Sequences | 29 |
| 2.3.2. Optimizing <i>L. reuteri</i> PAP Via PCR Stitching..... | 30 |
| 2.3.3. Construction and Transformation of pET30-derived Expression Vectors into <i>E. coli</i> BL21(DE3) | 33 |
| 2.3.4. Expression and Purification of PAPs and PEPs in <i>E. coli</i> Strain BL21(DE3)..... | 34 |
| 2.3.5. Enzyme Activity | 37 |
| 2.4. Discussion | 42 |
| 3. CONSTRUCTION OF A VECTOR-BASED EXPRESSION CASSETTE TO ASSESS THE ACTIVITY OF MXPEP IN LACTOBACILLUS REUTERI | 45 |
| 3.1. Introduction..... | 45 |
| 3.1.1. Probiotics as a Delivery Vehicle..... | 45 |
| 3.1.2. Applications of Probiotics: Food Production and Health Benefits | 46 |

| | |
|--|----|
| 3.1.3. Benefits of Heterologous Protein Expression in Probiotics | 47 |
| 3.1.4. The Use of <i>Lactobacillus reuteri</i> as a Therapy for CD | 48 |
| 3.1.5. Study Goal | 50 |
| 3.2. Materials and Methods..... | 51 |
| 3.2.1. Strains, Plasmids, and Growth Conditions | 51 |
| 3.2.2. Molecular Techniques | 51 |
| 3.2.3. Digestion Independent Cloning | 51 |
| 3.2.4. Construction of pGKMCS-derived Vectors | 52 |
| 3.2.5. Construction of pGKMCS <i>ErmB</i> Fixed Vectors | 56 |
| 3.2.6. Construction of pGKMCS <i>ErmB</i> Δ <i>Lac</i> Promoter Vectors | 58 |
| 3.2.7. Transformation of pGKMCS-derived Vectors into <i>E. coli</i> | 58 |
| 3.2.8. Assessing the Stability of pGKMCS Expression Vectors | 59 |
| 3.2.9. Plasmid Sequencing..... | 59 |
| 3.2.10. Transformation of pGKMCS-derived Vectors into <i>L. reuteri</i> | 60 |
| 3.2.11. <i>L. reuteri</i> Cell Lysis and Protein Purification..... | 61 |
| 3.2.12. SDS-PAGE and Western Blot | 62 |
| 3.3. Results..... | 63 |
| 3.3.1. Expression and Purification of GFP and MxPEP | |
| in <i>E. coli</i> MC1061..... | 63 |
| 3.3.2. Assessing the Stability of pGKMCS Expression Vectors | 66 |
| 3.3.3. Transformation and Protein Expression in <i>L. reuteri</i> 100-23C | 70 |
| 3.3.4. Sequencing of pGKMCS EG, pGKMCS EAG, pGKMCS EAM, and | |
| pGKMCS EG Δ LP | 76 |

| | |
|--|----|
| 3.4. Discussion | 77 |
| BIBLIOGRAPHY | 84 |
| APPENDICES | |
| A. Optimized <i>Aspergillus niger</i> Patent PEP (1581 bp) | 92 |
| B. <i>Lactobacillus acidophilus</i> PAP (918 bp) | 93 |
| C. Optimized <i>Lactobacillus reuteri</i> PAP (928 bp) | 94 |
| D. Optimized <i>Myxococcus xanthus</i> PEP (2070 bp) | 95 |

LIST OF TABLES

| Table | Page |
|---|------|
| 1. Primers Used to Optimize the <i>L. reuteri</i> PAP Sequence | 22 |
| 2. Primers Used to Construct pET30-derived Expression Vectors..... | 24 |
| 3. Codons Used to Rewrite the AnPat PEP and MxPEP Sequences | 30 |
| 4. Summary of the Kinetic Parameters (V_{\max} , K_m , k_{cat} , and k_{cat}/K_m) for LaPAP, LrPAP, and MxPEP..... | 38 |
| 5. Primers Used to Construct pGKMCS-derived Expression Vectors Containing the <i>ErmB</i> Promoter..... | 55 |
| 6. The pGKMCS Vectors and Associated Primers Used for Sequencing | 60 |
| 7. Expected Band Sizes for Each Vector and Set of Primers..... | 72 |

LIST OF FIGURES

| Figure | Page |
|---|------|
| 1. The Six DQ2-specific Epitopes (DQ2- α -I, DQ2- α -II, and DQ2- α -III) Located in the 33-mer | 3 |
| 2. Construction of pCR2.1 Rd1 Mut | 20 |
| 3. Construction of Final Optimized <i>L. reuteri</i> PAP (Rd2 Mut) | 21 |
| 4. <i>RsaI</i> Restriction Digest on Rd1 Mut and LrPAP WT | 31 |
| 5. <i>SalI</i> Restriction Digests on Rd1 Mut and Rd2 Mut | 32 |
| 6. <i>AscI</i> Restriction Digests on pET30 AnPat PEP, LaPAP, LrPAP, MxPEP, and GFP | 33 |
| 7. Lysates and Purified Samples of GFP and LrPAP | 34 |
| 8. Lysates and Purified Samples of GFP, LaPAP, and MxPEP | 35 |
| 9. Pellets, Lysates, and Purified Samples of AnPat PEP | 36 |
| 10. Western Blot of GFP and AnPat PEP | 37 |
| 11. LaPAP Enzyme Kinetics with H-Pro-pNA | 39 |
| 12. LrPAP Enzyme Kinetics with H-Pro-pNA | 40 |
| 13. MxPEP enzyme kinetics with Suc-Ala-Pro-pNA | 41 |
| 14. Digestion Independent Cloning (DIC) | 52 |
| 15. Construction of pGKMCS <i>ErmB</i> AGFP | 54 |
| 16. Construction of pGKMCS <i>ErmB</i> Fixed GFP Vector | 57 |
| 17. Lysates and Purified Samples of GFP from pGKMCS <i>ErmB</i> GFP | 64 |
| 18. Lysates and Purified Samples of MxPEP from pGKMCS <i>ErmB</i> MxPEP | 65 |
| 19. Molecular Evolution of pGKMCS EG and EG Δ LP | 67 |

| | |
|--|----|
| 20. Molecular Evolution of pGKMCS EM and EMΔLP | 68 |
| 21. Molecular Evolution of pGKMCS EAM and EAMΔLP | 69 |
| 22. <i>L. reuteri</i> Colony PCR to Confirm the Presence or Absence of the <i>Lac</i> Promoter in the pGKMCS Expression Vectors | 73 |
| 23. <i>L. reuteri</i> Colony PCR to Confirm the Presence of <i>ErmB</i> , <i>Amyl</i> Signal Sequence, and GFP in pGKMCS Expression Vectors | 74 |
| 24. <i>L. reuteri</i> Colony PCR to Confirm the Presence of <i>ErmB</i> and MxPEP in pGKMCS Expression Vectors | 75 |

1. 0 Celiac Disease Background

1.1 What is Celiac Disease?

Celiac Disease (CD) is an autoimmune disease that is triggered by gluten, a protein found in a variety of common grains such as wheat, rye, and barley. The duodenum, the proximal portion of the small intestine, is the primary site affected by CD, but in some cases it may extend to the ileum (Ruben et al., 1962). When those with CD are exposed to gluten they can experience abdominal pain, diarrhea, and an inflammatory cutaneous disease resulting in lesions, rashes, and blisters on the skin known as dermatitis herpetiformis (Caproni et al., 2009). More severe symptoms lead to the destruction of the small intestine villi, finger-like projections that allow for the absorption of nutrients. These symptoms include an elongation of the grooves located between villi (crypt hyperplasia) and the chronic inflammation of the cells lining the small intestine and destruction of villi (villous atrophy). The Crypts of Lieberkuhn are lined with young intraepithelial cells and are the source of intraepithelial stem cells of the small intestine. Thus, deepening of the Crypts of Lieberkuhn due to crypt hyperplasia indicates an increase in the production of intraepithelial cells. The reduction of the intestinal surface area due to villous atrophy may lead to the malabsorption of nutrients. The prevalence of CD is approximately 1 in 100 among the Caucasian population, making it the most commonly diagnosed chronic gastrointestinal disorder in the United States (Ciclitira et. al., 2005 and García-Manzanares and Lucendo, 2011). However, the prevalence of CD increases to 7.5% for a first-degree relative of an individual with CD (Singh et al. 2015).

1.2 The Role of Gluten in CD

Gluten is a proline- and glutamine-rich protein that can be segregated into two general classifications: glutenins and gliadins. Glutenin is the alcohol-insoluble fraction, which contributes to dough elasticity. Gliadin, the alcohol-soluble peptide, contains high levels of proline and glutamine residues as well as the dominant epitopes involved in inducing the immune response. High levels of proline residues allow gluten peptides to form proteolytic resistant structures that are not easily degraded by gastric, pancreatic, and brush border enzymes found in the human gastrointestinal tract (Kagnoff 2005). Because the structures are resistant to digestion, proline-rich peptides accumulate in the gastrointestinal tract, and the immunogenic epitopes of gluten are preserved. In particular, there are two major categories of immunogenic epitopes derived from α -gliadin: LQLQPFPPQLPYPQPQLPYPQPQLPYPQPQPF (33-mer) and PGQQQPFPPQQPY (p31-p43) (Ozuna et al. 2015). The main immunogenic epitope, 33-mer, is extremely stable and instigates an immune response in virtually all CD patients tested (Shan et al. 2002). As shown in Figure 1-1, the 33-mer contains six overlapping epitopes that have been shown to stimulate three different HLA-DQ2-restricted T cell clones from CD patients (Arentz-Hansen et al. 2000). The p31-p43 peptide has been shown to stimulate the expression of MICA, a molecule associated with cytotoxicity in intestinal epithelial cells (IECs) (Hüe et al. 2004).



DO2- α -I: PFPOPOLPY

DO2- α -II: POPOLPYPO

DO2- α -III: PYPOPOLPY

Figure 1-1. The six DQ2-specific Epitopes (DQ2- α -I, DQ2- α -II, and DQ2- α -III) Located in the 33-mer.

Glutamine residues found within the immunogenic gluten epitopes serve as targets for tissue transglutaminase (tTG), an endogenous enzyme involved in the autoimmune response of CD. This enzyme is ubiquitously expressed throughout the intracellular and extracellular environments of many organs, and it targets glutamine residues in a QXP consensus sequence that will either be transamidated (cross-linked) or deamidated (Dørum et al., 2010). The enzyme primarily catalyzes the irreversible transfer of an acyl group from a glutamine residue to a lysine residue thereby forming an isopeptide bond (Lai et al., 2017). This bond may form between glutamine and lysine residues of the same protein or two different proteins. Under acidic conditions tTG deamidates glutamine residues, an uncharged and polar amino acid, to glutamic acid, a negatively charged amino acid (Di Sabatino et al., 2011). Due to deamidation, gluten develops a stronger binding affinity to the major histocompatibility complex class II (MHC II) molecules that present antigens to pivotal cells of the immune response. The interaction between these gliadin peptides and tTG contributes to the hallmark of CD: anti-tTG antibodies. However, before gluten can invoke the immune and autoimmune responses, the peptides must first cross the intestinal epithelial layer into the lamina propria of the intestinal mucosa.

Under normal conditions macromolecules are inhibited from crossing the intestinal epithelial layer via tight junctions. In CD, the interactions between gliadin, CXCR3 chemokine receptors, adaptor protein MyD88, and zonulin influence the integrity of the small intestine tight junctions and allow the passage of gluten from the intestinal lumen into the lamina propria. CXCR3 is a G-protein-coupled receptor responsible for binding CXC chemokines. These receptors are normally found on lymphocytes, smooth muscle, and epithelial cells, but are overexpressed on the luminal side of IECs in CD (Van Raemdonck et al. 2015; Fasano 2011). Gliadin binds CXCR3 receptors and stimulates the recruitment of the adaptor protein MyD88, which is essential in the signal transduction pathways that culminate in the activation of cytokines and the release of zonulin into the intestine (Thomas et al. 2006; Yu et al. 2006; Lammers et al. 2008). Zonulin is a protein released from IECs after exposure to enteric bacteria and gluten that regulates intercellular tight junctions (Wang 2000; Asmar et al. 2002; Thomas et al. 2006). It has been shown that in the presence of zonulin tight junction permeability increases due to protein kinase C-dependent phosphorylation and displacement of zonula occluden-1 (ZO-1), a scaffold protein that couples integral plasma membrane proteins to other tight junction proteins and the cytoskeleton. Thus, it has been hypothesized that the gluten peptides cross into the intestinal mucosa when the tight junctions of the intestine are weakened due to the release of zonulin. Upon entry into the lamina propria of the intestinal mucosa, gluten peptides will initiate the immune responses.

1.3 The Immune Responses Involved in CD

Villous atrophy in CD is induced by both the innate and adaptive immune systems through the NKG2 family of natural killer (NK) cell receptors and non-classical MHC I chain-related molecule A (MICA). These receptors are expressed on both NK cells and a subset of cytotoxic T lymphocytes (CTLs). NK cells are not antigen-specific but are programmed to kill cells that have been compromised by a virus whereas CTLs target cells that present a specific antigen. Two types of surface NKG2 receptors may be expressed on NK cells and CTLs: an inhibiting (CD94/NKG2A) and activating (NKG2D) receptor. Under normal conditions, classical MHC I proteins, expressed on the surface of all nucleated cells, present self-antigens that are recognized by the inhibitory CD94/NKG2A receptor. MICA is an inducibly expressed ligand that acts as a cellular distress signal to stimulate cytotoxicity, cytokine secretion, and serves as an activating ligand for the activating NKG2D receptor (Baranwal and Mehra 2017; Bauer et al., 2002; Hübner et al., 2004). Inhibition signals produced from the engagement between the CD94/NKG2A receptor and MHC I override activation signals from the NKG2D receptor and MICA interaction. In cells where MHC I expression is greatly reduced (often a consequence of virally infected cells) the inhibition signals are not produced. The unimpeded activation signals lead to the destruction of the compromised cells through NK cell-mediated cytotoxicity or provide a costimulatory signal that enhances T cell receptor (TCR)-dependent cytotoxicity (Groh et al. 2001). Thus, altered expression of MICA and the NKG2 receptors is vital to the development of villous atrophy in CD.

Expression of the NKG2 receptors and MICA is influenced by IL-15, a stress and inflammation induced cytokine critical in the activation of T and NK cells. In CD, the p31-49 peptide of gliadin has been shown to upregulate surface expression of IL-15 on IECs as well as antigen presenting cells (APCs) in the lamina propria (Maiuri et al., 2003 and Bernardo et al., 2016). Overproduction of IL-15 leads to increased expression of NK2GD receptors on CTLs, and the presence of gliadin (α -gliadin and the p31-49 peptide) promotes the overexpression of MICA on IECs (Jabri et al. 2002; Roberts et al. 2001; H  e et al. 2004). The engagement between NK2GD and MICA acts as a costimulatory signal for TCR activation to stimulate cytolytic activity, and IL-15 lowers the activation threshold of the TCR (Groh et al. 2001; Roberts et al. 2001). Altogether, cytolysis of IECs is dependent on both NKG2D and the TCR, but efficient damage can take place at low self-antigen concentrations or low binding affinity for the self-antigen (Liu et al. 2013). Therefore, damage is more pervasive in a given area but restricted to IECs that express both IL-15 and MICA (Abadie and Jabri 2014). In more severe forms of CD, such as refractory CD (symptoms persist despite treatment with a gluten-free diet), cytolysis of IECs is independent of TCR activation and is solely mediated by NKG2D (H  e et al. 2004). Thus, IECs are destroyed through cytotoxic-mediated killing ultimately leading to the development of crypt hyperplasia and villous atrophy. These responses explain how gluten is involved in adaptive and innate immunity of CD; however, the autoimmune response has not been addressed.

1.4 Autoimmune Responses Involved in CD

Environmental factors, genetic predispositions, and, most importantly, the failure of immune tolerance contribute to the development of autoimmunity. Two mechanisms, central and peripheral tolerance, regulate B cells to prevent these cells from becoming auto-reactive. Central tolerance takes place in the primary lymphoid organs during lymphocyte repertoire development when B cells are equipped with receptors to identify specific antigens. Immature B cells that bind to self-antigens with high affinity undergo apoptosis, or genes are rearranged so the receptor no longer binds the self-antigen. Peripheral tolerance occurs in the secondary lymphoid organs and acts as a back-up plan for auto-reactive cells that escape central tolerance. This form of immune tolerance is necessary when the self-antigen is solely found outside primary lymphoid organs or when the cells encounter a soluble form of the self-antigen. Recognition of a self-antigen in the absence of a co-stimulatory signal from a helper T cell results in the B cell becoming anergic. The B cell can no longer respond to that antigen even if presented under proper conditions. When either central or peripheral tolerance fails, B cells are incapable of distinguishing between self and non-self proteins, thus an immune response can be triggered by a self-protein. However, unlike most autoimmune diseases, a foreign antigen (gliadin) and self-protein (tTG) are both responsible for initiating autoimmune responses of CD.

Gliadin is a preferred substrate for tTG to act upon; approximately 36% of glutamine residues in gliadin are accessible to tTG modification (Di Sabatino et al. 2012). The deamidation activity of tTG is critical for the presentation of gliadin peptides by MHC II molecules. There is a strong genetic association between CD and the MHC II HLA-DQ

locus. Approximately 90% of CD individuals either have the MHC II HLA-DQ2.5 or DQ8 alleles (Sollid & Thorsby, 1993 and Lucendo, 2011). Moreover, MHC II HLA-DQ2.5 and DQ8 have a binding preference for negatively charged epitopes (Molberg et al., 1998 and Dørum et al., 2010). Therefore, the conversion of uncharged, polar glutamine residues to negatively charged glutamic acid residues bolsters the avidity of MHC II HLA-DQ2.5 or DQ8 for gliadin epitopes. Additionally, the preference for the deamidated immunogenic peptides of gliadin is strong enough to displace antigens that may already be presented by these MHC II HLA-DQ molecules (Xia, Sollid, and Khosla 2005). Furthermore, tTG can cross-link itself to immunodominant peptides of gliadin through transamidation (Dieterich et al., 2006). Auto-reactive self-tTG B cells bind tTG that has been cross-linked with gliadin, and the tTG-gliadin complex is endocytosed, degraded, processed, and the antigens are presented via the MHC II HLA-DQ heterodimer (Di Sabatino et al., 2012). Even though both gliadin peptides and tTG were taken up by the auto-reactive tTG-specific B cell, MHC II DQ2.5 and DQ8 preferentially present the immunogenic peptides of deamidated gliadin (Di Sabatino et al. 2012). Gliadin-specific Th2 cells then recognize the gliadin fragments presented by auto-reactive tTG-specific B cells and stimulate the B cells to produce anti-tTG antibodies. It is significant to note that the activation of auto-reactive tTG-specific B cells does not require stimulation from auto-reactive tTG-specific T cells. Thus, the formation of the cross-linked tTG-gliadin complex and the modification of the negatively charged gliadin peptide are essential for the production of anti-tTG antibodies from auto-reactive tTG B cells. Although these anti-tTG antibodies are characteristic of CD, the exact function of the antibodies in CD pathology remains unclear.

1.5 Diagnosis of CD

The presence of MHC II HLA-DQ2.5 or DQ8 alleles, anti-tTG IgA antibodies, and villous atrophy are distinguishing features of CD. If there is a risk of acquiring CD individuals will undergo preliminary genetic and serological tests to screen for the disorder. Although 20-30% of the general population carries the MHC II HLA-DQ2.5 or DQ8 alleles, the presence of these alleles may increase the risk of developing CD (Bodis, Toth, and Schwarting 2018). Individuals that were homozygous for HLA-DQ2.5 had the highest risk for developing CD; however, the absence of the MHC II HLA-DQ2.5 or DQ8 alleles was a strong indicator that an individual would not develop CD (Sollid and Lie 2005). Moreover, other genetic factors may contribute to the development of CD. Studies conducted on twins and the occurrence of CD have indicated a ~0.80 concordance rate (the probability a twin is affected provided that the co-twin is affected) compared to a concordance rate of <0.20 in dizygotic twins (Nistico 2006; Greco 2002). Thus, the presence of the MHC II HLA-DQ2.5 or DQ8 alleles is necessary for development of CD, but these alleles are not sufficient to confirm the risk of developing CD. Prior to the serological test, the patient is subjected to a gluten challenge (GC) where the individual will include gluten in their diet to promote the immune and autoimmune responses associated with CD. A study has shown that a 14-day GC with ≥ 3 grams of gluten consumed per day was sufficient to induce the serological hallmarks of CD (Leffler et al. 2013). Following the GC, a serological test is performed to detect the presence of anti-tTG IgA antibodies, an indicator that the autoimmune response of CD has been triggered. Approximately 98% of adult patients with CD test positive for tTG-IgA while on a gluten-containing diet, and 98% of individuals without CD test negative for this antibody (Lebwohl et al. 2012). Genetic and serological

tests may eliminate the prospect of CD or narrow down the array of potential disorders; however, they cannot confirm if an individual has CD. If individuals test positive for these genetic and serological tests, then the diagnosis must be confirmed with a duodenal biopsy to detect villous atrophy, crypt hyperplasia, and intraepithelial lymphocytosis (increased production of intraepithelial lymphocytes) consistent with CD.

1.6 Current Treatments for CD

No cure for CD has been discovered, and the only known method to manage the disorder is a strict and lifelong adherence to a gluten-free diet (GFD): elimination of wheat, rye, barley, and any food products that were derived from those grains. Most individuals experience a substantial improvement in symptoms after maintaining a GFD. However, symptoms may persist for others due to accidental gluten consumption, because those gluten-containing grains are thoroughly abundant in the western diet. For example, there is no current law that mandates manufacturers disclose if medications contain gluten. Moreover, sensitivity to gluten and the degree of the immune response may differ amongst individuals with CD. Those with extreme gluten sensitivity may suffer a severe immune response to a trace amount of gluten, therefore, maintaining an absolute GFD is challenging. In addition, individuals are often dissatisfied with a GFD due to the high cost of food and restrictive nature of this lifestyle, so interest in alternative strategies to manage CD is growing (Aziz et al., 2010). Long-term consequences of ingesting gluten may arise in the future even if the immediate symptoms of CD subside. A minority of individuals may develop more severe disorders such as ulcerative jejunitis (characterized by villous atrophy and chronic ulcers in the small intestine), small bowel adenocarcinoma (a type of

cancer that emerges in mucus-secreting glands), and enteropathy-associated T-cell lymphomas (EATL) due to unmanaged CD (Elsing et al., 2005; Basha et al., 2014; and Al-Bawardy et al. 2017). For individuals with CD, the estimated relative risk in developing small bowel adenocarcinoma and EATL increases 80-fold and 30-fold, respectively (Meijer et al. 2004). Thus, a therapy to defend the small intestine from trace amounts of gluten would be beneficial for individuals with CD.

Prolyl endopeptidases (PEPs) and prolyl aminopeptidases (PAPs) are proteases with the unique ability to hydrolyze peptide bonds adjacent to an internal and a terminal proline residue, respectively. These peptidases may be used to break down the immunogenic epitopes of gluten, thereby assuaging the autoimmune response of CD. Previous studies have shown the use of PEPs as a potential treatment for CD (Shan et al. 2004); however, the enzyme therapy requires that an individual anticipates consuming gluten-containing foods. This would not protect an individual from accidental gluten consumption. Moreover, orally administered enzymes must pass through the low pH of the stomach and proteolytic enzymes: pepsin, trypsin, and chymotrypsin. Probiotics such as Lactic Acid Bacteria (LAB) are acid-tolerant, therefore, some probiotics are capable of surviving the acidic environment of the stomach to colonize the gut. Thus, the use of a probiotic to deliver the enzyme to the site affected by CD may provide an alternate defensive strategy against CD.

1.7 Study Goal

The goal of this study was to develop a probiotic expression system that allowed for the production PAPs and PEPs capable of cleaving proline-rich immunogenic peptides. A pET30 expression system in *E. coli* strain BL21(DE3) was employed to evaluate the expression and cleavage activities of the *Myxococcus xanthus* PEP (MxPEP), *Aspergillus niger* PEP (AnPEP), *Lactobacillus acidophilus* PAP (LaPAP), *Lactobacillus reuteri* PAP (LrPAP), and patented *Aspergillus niger* PEP (AnPat PEP). The expression of these proteins was analyzed through western blot, and the activity of the enzymes was evaluated through a chromogenic substrate assay. The enzyme with the most efficient activity was incorporated into a *L. reuteri* expression cassette. Expression of the enzyme was assessed through SDS-PAGE and western blot.

2.0 The Assessment of the Cleavage Activity of Prolyl Aminopeptidases and Prolyl Endopeptidases on Proline-containing Substrates

2.1 Introduction

2.1.1 Serine Proteases

Serine proteases are ubiquitously found in eukaryotes, prokaryotes, archaea as well as viruses. Approximately one third of known proteolytic enzymes are serine proteases which are characterized by a catalytic triad of asparagine, histidine, and serine residues. The name is derived from the nucleophilic serine residue in the active site which donates electrons to the carbonyl of a peptide to form an acyl-enzyme intermediate (Hedstrom 2002). Because of their naturally destructive nature, these enzymes are generally stored as zymogens (an inactive form of the protein) and are activated by location-specific cleavage. In humans, the pancreas secretes the zymogen trypsinogen which localizes to the duodenum where it is cleaved by enteropeptidases to form active trypsin. Trypsin along with chymotrypsin and elastase promote the digestion of proteins in the GI tract; however, these serine proteases remain relatively ineffective against gluten. For example, trypsin acts upon the C-terminus of arginine and lysine residues, but these residues are scarce throughout gliadin peptides (Ferranti et al. 2007). Furthermore, proline residues located at the carboxyl end of the aforementioned residues interfere with the proteolytic activity of trypsin. Alternative serine proteases such as prolyl aminopeptidase (PAP) and prolyl endopeptidase (PEP) are capable of hydrolyzing proline-containing peptides, thus these enzymes may be a valuable asset for gluten hydrolysis.

2.1.2 Prolyl Endopeptidases and Prolyl Aminopeptidases

PAPs cleave at the N-terminus of a proline residue whereas PEPs cleave internal peptide bonds at the C-terminus of an internal proline residue, and these proteases target smaller peptides (~30 amino acids) as their substrates. Because these are features of the immunogenic 33-mer of gliadin, PAPs and PEPs may be a promising therapeutic for CD. Previous studies have successfully isolated and characterized PEPs and PAPs from fungi (*Aspergillus niger*) and bacteria (*Myxococcus xanthus*, *Flavobacterium meningosepticum*, and *Sphingomonas capsulate*). The *A. niger* PEP (AnPEP) displayed enzymatic activity between pH 2-8, but was most effective at hydrolyzing gluten at pH 4 (Montserrat et al. 2015). Another study further assessed the immunoreactivity of AnPEP hydrolyzed wheat flour (AnPEP HWF) on the development of an immune response in Caco-2 cells (Mohan Kumar et al. 2019). The study reported a reduction in the immune response when these Caco-2 cells did not exhibit an increase in zonulin. This study reported that Caco-2 cells stimulated with AnPEP HWF did not exhibit a significant increase in zonulin compared to that of cells exposed to gliadin. Moreover, the spleenocytes of gluten-sensitive mice treated with AnPEP hydrolyzed wheat flour did not yield a marked increase in IL-15 compared to mice challenged with gluten. In a comparative analysis of PEPs derived from *M. xanthus*, *F. meningosepticum*, and *S. capsulate*, all PEPs exhibited optimal catalytic activity at pH 7; however, the *M. xanthus* PEP displayed the highest affinity for the chromogenic substrate, Suc-Ala-Pro-pNA (Shan et al. 2004). Furthermore, it has been reported that the *F. meningosepticum* PEP was effective at reducing immunogenicity associated with gliadin epitopes (Marti et al. 2004)

2.1.3 Study Goal

The purpose of this study was to assess the enzymatic activity of a PEP from *M. xanthus* and PAPs from *A. niger*, *Lactobacillus acidophilus*, and *Lactobacillus. reuteri*. These genes were introduced into the inducible pET30 expression vectors and transformed into *E. coli* BL21(DE3). Recombinant protein expression was evaluated by SDS-PAGE and western blot, and enzyme activity was assessed through chromogenic substrate assays with Suc-Ala-Pro-pNA or H-Pro-pNA. The PEP or PAP that exhibited the highest enzymatic activity was used in the construction of *L. reuteri* expression vectors.

2.2 Materials and Methods

2.2.1 Strains and growth conditions

E. coli strains TOP10 , MC1061, and BL21(DE3) were utilized in this study. TOP10 was used to harbor pCR2.1-derived vectors. Newly constructed pET30 expression vectors were first cloned into MC1061, and BL21(DE3) was used to express recombinant proteins from pET30-based expression vectors. Cultures were grown at 37°C under aerobic conditions in Luria Bertani (LB) broth or plates. Either ampicillin 100 µg/ml (Amp 100) or kanamycin 30 µg/ml (Kan 30) were added to the media for strains containing pCR2.1-derived vectors. Cultures containing pET30 expression vectors were grown in media supplemented with Kan 30.

2.2.2 Molecular techniques

Polymerase chain reaction (PCR) was performed using Taq or Phusion DNA polymerase in 20 µl volumes. All reactions contained a final primer concentration of 0.1 µM and approximately 0.2 ng of template DNA. For reactions using GoTaq 1X Master Mix (Promega, Madison, WI, USA), DNA was initially denatured at 95°C for 2 minutes. Each reaction was then carried out for 30 cycles of 30 seconds at 95°C, 30 seconds at 55°C, and 1 minute per kb of expected product at 72°C. Following the 30 cycles, a final 5-minute extension at 72°C completed the reaction. When using Phusion Flash High-Fidelity PCR 1X Master Mix (Thermo Scientific, Rockford, IL, USA), DNA was initially denatured for 10 seconds at 98°C. Subsequently, reactions were executed for 30 cycles of 1 second at 98°C, 30 seconds at 55°C, and 15 seconds per kb of expected product at 72°C. The reaction ended with a final 5-minute extension step at 72°C and incubation at 4°C.

Colony PCR was used to evaluate 11-22 CFUs to determine if the colonies contained the recombinant plasmid of interest. Briefly, colonies were suspended in separate aliquots from a PCR master mix (primers and 1X GoTaq master mix) and spot inoculated on LB Kan 30 from each PCR reaction mixture. The primer pairs contained in the PCR master mix were chosen to identify the presence and orientation of the insert; one primer was specific to the insert and another annealed solely to the vector. For example, a forward primer specific to the 5' end of the insert and a reverse primer that annealed to the vector downstream from the insert, or a forward primer within the vector upstream from the insert and a reverse primer specific to the 3' end of the insert.

All restriction digests using *AscI*, *BamHI*, *EcoRI*, and *SaII* (Thermo Scientific, Rockford, IL, USA) were carried out in 20 µl reaction volumes that contained 1X FastDigest Buffer and 300 – 800 ng of DNA. Restriction enzymes were incubated at 37°C for 1 hour and heat inactivated at 80°C for 5 minutes. Following digests, vectors were dephosphorylated with 1X OPTIZYME Alkaline Phosphatase (Thermo Scientific, Rockford, IL, USA) to prevent self-ligation. Reactions contained 10X OPTIZYME AP buffer and 600-800 ng of linear DNA. Dephosphorylation reactions were performed at 37°C for 1 hour and heat inactivated 75°C for 5 minutes.

Ligations were performed with T4 DNA Ligase (Thermo Scientific, Rockford, IL, USA) in a 10 µl reaction mixture containing a 3:1 insert:vector molar ratio (300 ng vector) and 1X T4 DNA Ligase Buffer. The reactions were ligated at 22°C for 60 minutes and heat inactivated at 70°C for 5 minutes. Ligated vectors were transformed into *E. coli* strains MC1061 and BL21(DE3).

2.2.3 Optimizing prolyl endopeptidase (PEP) and prolyl aminopeptidase (PAP) sequences via gene synthesis

E. coli and *L. reuteri* codon usage tables (<http://www.kazusa.or.jp/codon/>) were used in a relative adaptiveness analysis (<http://gcu.schoedl.de/>) to optimize the *M. xanthus* PEP and *A. niger* patented PAP sequences for expression in *E. coli* and *L. reuteri*. For this analysis, codon usage tables for *E. coli* and *L. reuteri* were compared with the *M. xanthus* PEP and *A. niger* patented PAP sequences. In the relative adaptiveness analysis, codon frequencies were converted to relative adaptiveness values; the most frequent codon for a particular

amino acid was set to 100% relative adaptiveness, and the remaining codons for that amino acid were scaled as follows:

$$\text{Relative adaptiveness value} = \text{Codon frequency} \left(\frac{100}{\text{Highest codon frequency}} \right)$$

The relative adaptiveness analysis provided the capability to compare codon frequencies between different amino acids. Both the *M. xanthus* PEP and *A. niger* patented PAP sequences were rewritten to contain the codons that yielded the highest relative adaptiveness values between *E. coli* and *L. reuteri*. The analyses were performed by Graphical Codon Usage Analyser. The optimized *M. xanthus* and *A. niger* sequences were synthesized by Life Technologies (Carlsbad, CA, USA).

2.2.4 Optimizing *L. reuteri* PAP sequence via PCR stitching

For use in this work and a related study, the *L. reuteri* PAP sequence was optimized for expression in *Saccharomyces cerevisiae*, *E. coli*, and *L. reuteri* by site-specific mutagenesis through PCR stitching (Figure 2-1 and Figure 2-2.). The Reuteri-PAP-EcoRI-F and Reuteri-PAP-RsaI-mut-R primers were used to amplify piece A. These primers introduced a new *RsaI* site into the sequence and mutated two CGG codons to CGT. In addition, Reuteri-PAP-EcoRI-F and Reuteri-PAP-Arg-GFP-R were used to create piece B that mutated one CGG codon to CGT and incorporated a *BglIII* restriction site and a 3' end of GFP. Piece B was digested with *RsaI* and treated with terminal deoxynucleotidyl transferase (Promega, Madison, WI, USA) to prevent its amplification. Pieces A and B were PCR stitched with Reuteri-PAP-EcoRI-F and GFP-EcoRI-R to generate piece AB,

the round 1 mutant (Rd1 mut). This amplicon was digested with *Bgl*II and *Eco*RI ligated with pCR 2.1 that was digested with *Eco*RI and *Bam*HI. Using pCR 2.1 Rd1 mut as a template, primer pairs Reuteri-PAP-*Eco*RI-F/Lr-PAP-SalI-mut-R and Lr-PAP-SalI-mut-F/Reuteri-PAP-Arg-GFP-R generated pieces C and D, respectively. These pieces were PCR stitched with Reuteri-PAP-*Eco*RI-F and GFP-*Eco*RI-R to create piece CD, the round 2 mutant (Rd 2 mut). Rd2 mut was digested with *Eco*RI and *Bgl*II and inserted into the THA expression cassette in pRS416. The expression cassette contained the *triosephosphate isomerase* promoter (T), histidine tag (H), and *alcohol dehydrogenase* terminator (A). Table 2-1 contains the primers used to optimize the *L. reuteri* PAP sequence.

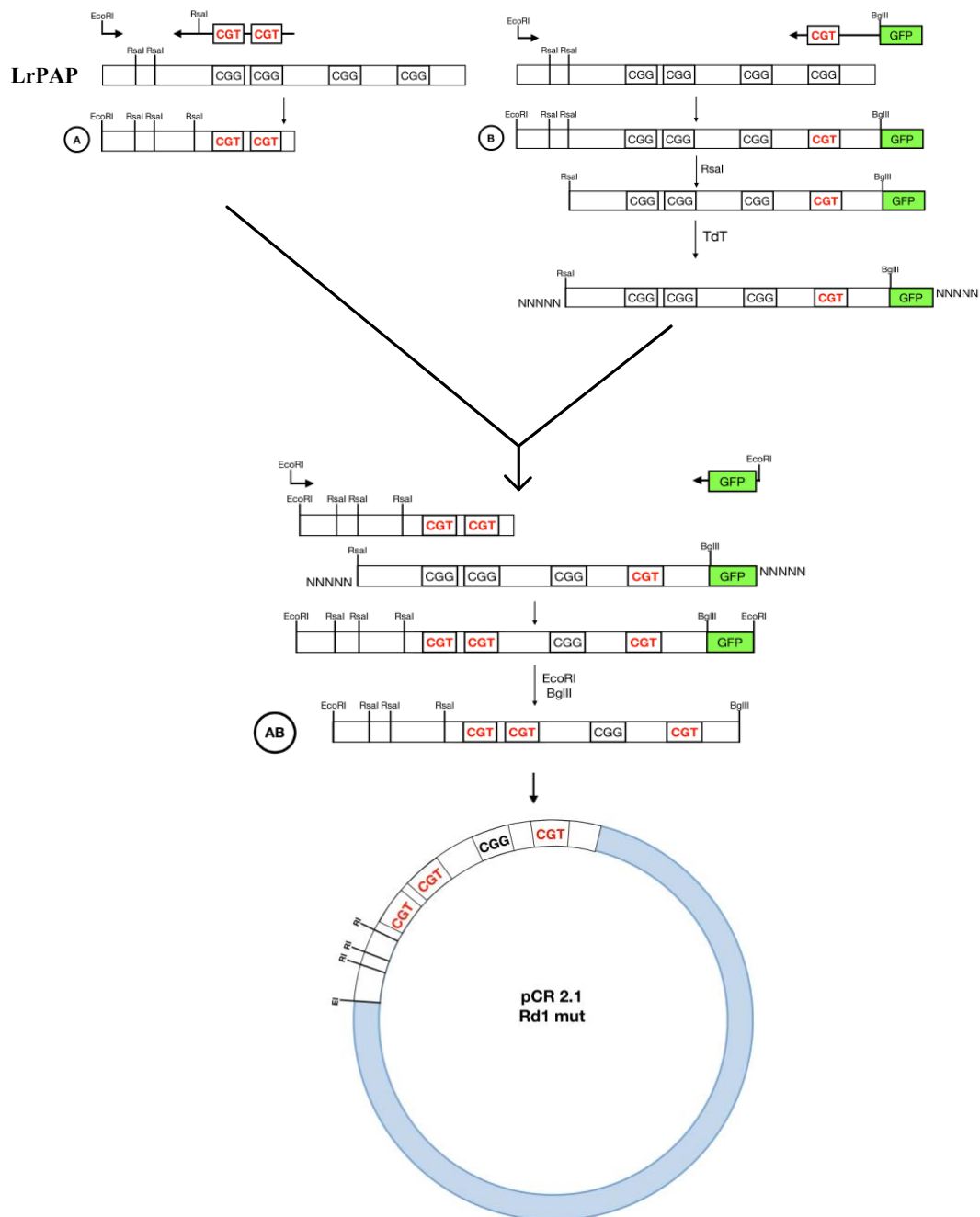


Figure 2-1. Construction of pCR 2.1 Rd1 Mut. Two CGG codons were mutated to CGT and an *RsaI* site was introduced into the 5' end of LrPAP (A). LrPAP was amplified to contain at the 3' end of the sequence (B). The resulting pieces were PCR stitched to generate RdI mut that contained 3 mutated arginine codons (AB).

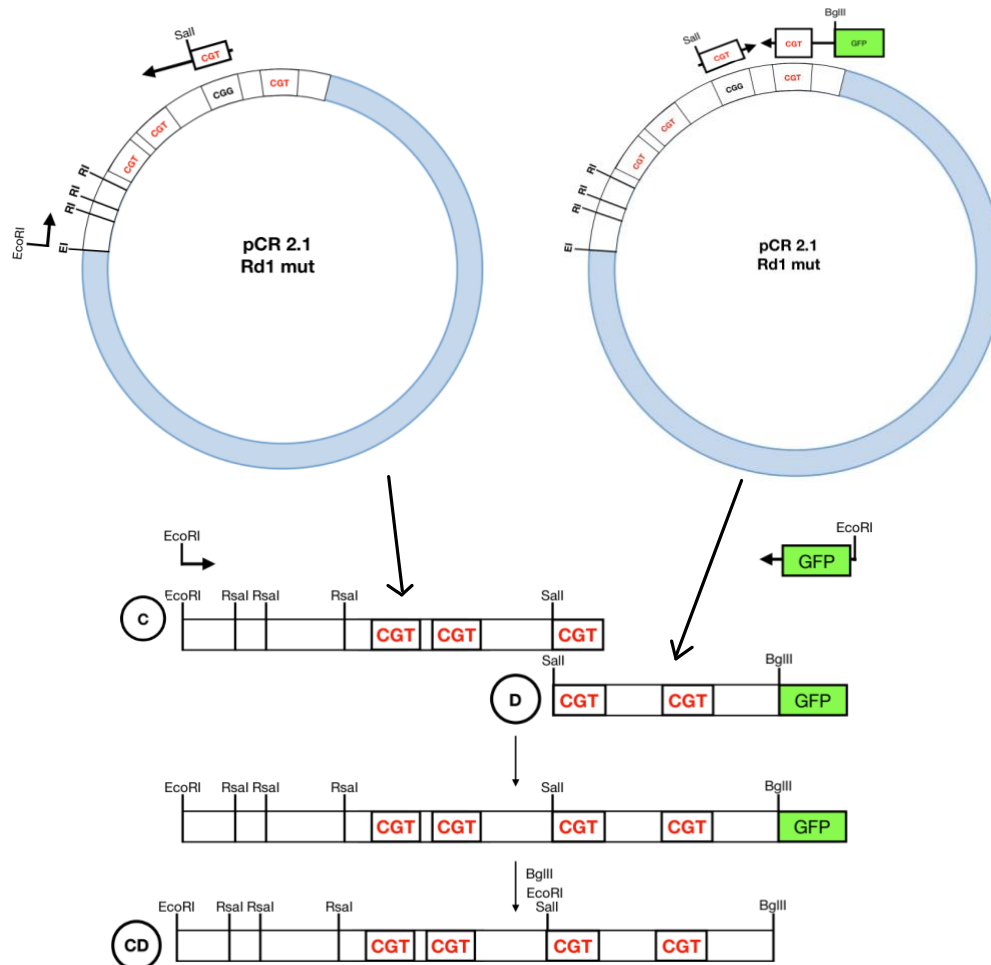


Figure 2-2. Construction of Final Optimized *L. reuteri* PAP (Rd2 Mut). A 5' portion of RdI mut was amplified to incorporate a *SalI* site and mutate the final CGG codon to CGT (**C**). The 3' end of RdI mut was amplified to mutate the final CGG codon to CGT and add the 5' end of GFP as well as a *SalI* site (**D**). Both pieces were PCR stitched to generate Rd2 mut with 4 mutated arginine codons (**CD**).

Table 2-1. Primers Used to Optimize the *L. reuteri* PAP Sequence. Site mutations identified in bold and underline.

| Name | Sequence (5' → 3') | Features | T _m | Target |
|-------------------------------|--|--------------|----------------|--|
| GFP-EcoRI-R | aaa GAATTC agattgtatagttccatcgcca | EcoRI | 71°C | 3' end of GFP flanked by EcoRI |
| Reuteri-PAP-EcoRI-F | cGAATTC ATGAAACAAGGCACCTAAAAATTATTACCC | EcoRI | 72°C | 5' end of <i>L. reuteri</i> PAP flanked by EcoRI |
| Lr-PAP-Sall-mut-F | GAACAATA CGTcGACc gtAAGCAGCCATCCAAAGCTT TATC | Sall | 80°C | Introduces Sall site and mutates CGG to CGT |
| Lr-PAP-Sall-mut-R | CTT acgGTCgAC CGTATTGTTCAATTCATTACTTGAACG | Sall | 75°C | Introduces Sall site and mutates CGG to CGT |
| Reuteri-PAP-Rsal-mut-R | GAGAAAGAGTCTTT Cacg TAATTTCATTAA Cacg GTC GAC gTAc TCATCAATTTCGTCAAcc | Rsal | 84°C | Introduces Rsal site and mutates CGG to CGT |
| Reuteri-PAP-Arg-GFP-R | CTCGCTGACTTTA Tcg tAATGTTGAAAATAATAC GTTTAATAATTA Aagatct GGCATGGATGA AACTATA CAA | BglII | 80°C | Mutates CGG to CGT and flanked by BglII site and the 3' end of GFP |

2.2.5 Construction of *pET30* vectors

The *M. xanthus* PEP, *L. acidophilus* PAP, *L. reuteri* PAP, and *A. niger* patented PAP (Kang, Yu, and Xu 2013) sequences were cloned into pET30 using an *AscI* restriction site located upstream from a histidine tag. Table 2-2 contains the primers that were used in the construction of the pET30 expression vectors.

Table 2-2. Primers Used to Construct pET30-derived Expression Vectors.

| Name | Sequence (5' → 3') | Features | T _m | Target |
|---------------|--|--------------------|----------------|------------------------------|
| An-Patent-F | cggcgcgcgc ATGCGTGCCTTTTTCAGCTGTTGC | <i>Ascl</i> | 88.7°C | pET30 <i>A. niger</i> patent |
| An-Patent-R | tggcgcgcgc AGCATATATTTCTTC | <i>Ascl</i> | 75.2°C | pET30 <i>A. niger</i> patent |
| MxPepOpt-Asc- | ga ggcgcgcgc ATGTCATATCCAGCTACTCG | <i>Ascl</i> | 79.8°C | pET30 <i>M. xanthus</i> |
| MxPepOpt-Asc- | gc ggcgcgcgc gACGACCTTGAGCAGCAACACC | <i>Ascl</i> | 88.1°C | pET30 <i>M. xanthus</i> |
| PepNCFM-F | aggcgcgcATGAAAACCTGGTACTAAAAATTATCA | <i>Ascl</i> | 57.1°C | pET30 <i>L. acidophilus</i> |
| PepNCFM-R | tggcgcgcgc GATTTTGGCCTTTAAAGGTG | <i>Ascl</i> | 60.4°C | pET30 <i>L. acidophilus</i> |
| AscI-PepPN-F | aggcgcgcgc ATGAACAAGGCACTAAAAATTATTACCC | <i>Ascl</i> | 64.8°C | pET30 <i>L. reuteri</i> |
| AscI-PepPN-R | aggcgcgcgc ATTATTAACGTATTATTTCACATT | <i>Ascl</i> | 60.7°C | pET30 <i>L. reuteri</i> |

2.2.6 Transformation into *E. coli* strains TOP10, MC1061, and BL21(DE3)

Three strains of *E. coli*, TOP10, MC1061, and BL21(DE3), were utilized for the transformations performed in this study. For transformations into *E. coli* TOP10, 2 µl of the vector was incubated with 40 µl of TOP10 at 42 °C for 30 seconds, and transformants were immediately incubated with 250 µl SOC recovery medium for 45 minutes at 37°C. After recovery incubation, 100 µl of the cells were plated on LB Amp 100 and grown for 18 hours at 37 °C. The newly constructed pCR 2.1 Rd1 mut was transformed into *E. coli* TOP10.

Electrocompetent *E. coli* BL21(DE3) were prepared by diluting an overnight culture 1:100 in 250 ml SOB. The culture was grown to OD₆₀₀= 0.5 – 0.7 and centrifuged at 3,000xg for 10 minutes at 4°C. Cells were washed twice with 250 ml nanopure water and once with 250 ml 10% glycerol. The supernatant was poured off and the pellet was resuspended in residual 10% glycerol. Electrocompetent cells were divided into 140 µl aliquots and flash frozen in liquid nitrogen and stored at -80°C. All electroporations were performed with 40 µl electrocompetent cells and 200-500 ng of DNA using the BTX Electro Cell Manipulator 600 (Harvard Apparatus Inc; 2.45 kV, 129 Ω). Transformants were incubated at 37°C for 1 hour in 500 µl SOC recovery medium. After recovery incubation, 100 µl of cells were plated on LB Kan 30 agar and grown for 18 hours at 37°C. All pET30-derived vectors were first transformed into MC1061 to obtain purified plasmid and subsequently electroporated into BL21(DE3) to evaluate expression of the proteins of interest.

2.2.7 Recombinant protein expression and purification

BL21(DE3) *E. coli* cultures harboring pET30 expression vectors were grown overnight at 37°C in 3 ml LB Kan 30 broth. Overnight cultures were diluted 1:100 in 250 ml LB Kan 30 broth, grown to OD₆₀₀ = 0.5 – 0.7, and induced for 18 hours with 0.1 mM IPTG at 22°C. Following induction, cells were pelleted at 8,000xg for 4 minutes at 4°C. Pellets were resuspended in 25 ml TKE (50 mM Tris pH 7.5, 2.5 mM KCl, and 0.5 mM EDTA) and sonicated for 30 seconds at power setting 5 (~14 RMS) and incubated on ice for 60 seconds for a total of 3 times. Sonicated samples were spiked with 125 µl of Triton X-100 (0.5% final concentration) and incubated for 15 minutes at room temperature. After the incubation period, samples were centrifuged at 10,000xg for 10 minutes at 4°C to separate unbroken cells, cell debris, and insoluble proteins (pellet) from soluble proteins (lysate). Recombinant proteins containing a C-terminal histidine tag were purified with HisPur™ Ni-NTA resin (Thermo Scientific, Rockford, IL, USA). Resin was prepared by washing twice with 500 µl water and once with 500 µl TKE. Lysates were incubated on ice with ~250 µl bed volume of Ni-NTA resin for 30 minutes. Ni-NTA bound proteins were separated from unbound proteins by centrifugation at 3,000xg for 5 minutes at 4°C. Pellets were resuspended in 4 ml TKE, transferred to a polypropylene column, washed twice with 5 ml TKE, and Ni-NTA bound proteins were eluted with 2 ml 100 mM imidazole. Eluates were transferred into 1 inch dialysis tubing with a molecular weight cutoff of 12 to 14 kDa (Carolina Biological Supply Company, Burlington, NC, USA). Samples were dialyzed in 330 ml storage buffer (50 mM Tris-HCl, pH 8.2, 1 mM DTT, 0.1 mM EDTA, 0.1% Tween-20, and 57% glycerol) for three days, and used storage buffer was replaced twice after 24 hours.

2.2.8 SDS-PAGE and Western Blots

Protein samples were prepared with a 4X SDS-PAGE sample buffer (250 mM Tris pH 6.8, 8% SDS, 40% glycerol, 0.02% bromophenol blue, and 5% β -mercaptoethanol), boiled at 99°C for 3 minutes, and placed on ice. From each sample, 15 μ l were electrophoresed for 90 minutes at 100 V on a 12% polyacrylamide gel. Following electrophoresis, SDS-PAGE gels used to evaluate cell lysates and purified protein samples were then stained with GelCode™ Blue Safe Protein Stain (Thermo Scientific, Rockford, IL, USA) for 60 minutes and destained overnight in water.

SDS-PAGE gels used to identify proteins of interest were subsequently analyzed by western blot. Proteins electrophoresed on SDS-PAGE were immediately transferred to a nitrocellulose membrane using a submersible transfer (Bio-Rad Laboratories, Inc., Hercules, CA, USA) at 150 mA for 30 minutes after electrophoresis. Following the transfer, membranes were incubated with Ponceau S to visualize and mark the molecular weight standard bands. Membranes were blocked overnight in 3% fat-free milk powder dissolved in TBST (25 mM Tris, 0.15 M NaCl, and 0.05% Tween-20; pH 7.2) at 4°C. The next day membranes were incubated in a hybridization oven with either 10 ml 1:2000 anti-GFP-HRP or 1:5000 Ni-HRP for 45 minutes at 25.5°C and subsequently washed 4 times with 5 ml TBST. Blots were incubated with 7 ml Supersignal West Pico Substrate Working Solution (Thermo Scientific, Rockford, IL, USA) to detect HRP activity. Western blots were visualized by chemiluminescent detection using the ChemiDoc XRS+ (Bio-Rad Laboratories, Inc., Hercules, CA, USA).

2.2.9 Enzyme Activity Assay

PEP and PAP enzyme activity was measured with the chromogenic substrates Succinyl-Alanyl-Prolyl-*p*-nitroanilide (Suc-Ala-Pro-pNA) or H-Prolyl-*p*-nitroanilide (H-Pro-pNA) (Bachem, Torrance, CA, USA), respectively. Enzyme concentrations were quantified with a BCA protein assay (Thermo Scientific, Rockford, IL, USA). Reaction mixtures (200 μ l) contained 100 mM potassium phosphate buffer pH 7.0, 0.1–2 mM chromogenic substrate, and 0.6 μ M LaPAP, 0.3 μ M purified LrPAP, or 0.015 μ M MxPEP. Stocks of Suc-Ala-Pro-pNA (dissolved in water) and H-Pro-pNA (dissolved in 20% methanol) were prepared at a concentration of 5 mM. The release of pNA was kinetically measured at a wavelength of 410 nm using a SpectraMax Plus 384 microplate reader (Molecular Devices, Sunnyvale, CA, USA). Enzymatic assays with the PEP were performed for 10 minutes, and absorbance measurements were taken at 10-second intervals. Assays containing a PAP were carried out for 60 minutes with absorbance measurements taken every 30 seconds. All reactions were incubated at 37°C.

Because a vertical beam of light was used to measure the absorbance of pNA, the pathlength was dependent on the height of the reaction mixtures. The pathlength for a 200 μ l reaction mixture was determined by measuring the absorbance of *p*-nitrophenol (PNP). PNP was diluted from 2.06×10^{-6} M – 4.12×10^{-5} M in 0.1M Na₂CO₃, and the absorbances were measured in triplicate at 400 nm. The pathlengths (*b*) were calculated according to Beer's Law:

$$A = \epsilon_{PNP} bc$$

Where A is the absorbance of the PNP, ϵ_{PNP} is the molar absorptivity coefficient of PNP ($1.7 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$), and c is the concentration of PNP. The average pathlength was used to adjust the molar absorptivity value of pNA (ϵ_{pNA}).

Velocities of the reactions were used to calculate international units (IU), the amount of enzyme required to convert 1 μmole of substrate per minute, according the following formula:

$$IU = \frac{(ABS_{410 \text{ nm}})(\text{reaction mixture volume, l})}{(\text{minute})(\mu\text{M value of } \epsilon_{\text{pNA}})}$$

The kinetic data was analyzed through Michaelis-Menten and Lineweaver-Burk plots to determine the V_{max} and K_{m} values for each enzyme.

2.3 Results

2.3.1 Optimizing PAP and PEP sequences

Because the genetic code has redundancies, certain amino acids are encoded by more than one codon. Since the PEP and PAP were obtained from different originating microorganisms, codon bias was an issue when incorporating the sequences into *E. coli* and *L. reuteri*. Table 2-1 shows the codons with the highest relative adaptiveness values between *E. coli* and *L. reuteri*. These codons were used to rewrite the patented *A. niger* PAP (AnPat PAP) and *M. xanthus* PEP (MxPEP) sequences that were synthesized by Life Technologies. See appendix for complete AnPat PAP and MxPEP sequences.

Table 2-3. Codons Used to Rewrite the AnPat PEP and MxPEP Sequences. These codons had the highest relative adaptiveness values between *E. coli* and *L. reuteri*.

| | | | | | |
|--------------------------|-----|--------------------------|-----|-----------------------|-----|
| Alanine (A) | GCT | Glycine (G) | GGT | Proline (P) | CCA |
| Arginine (R) | CGT | Histidine (H) | CAT | Serine (S) | TCA |
| Asparagine (N) | AAT | Isoleucine (I) | ATT | Threonine (T) | ACT |
| Aspartic acid (D) | GAT | Leucine (L) | TTA | Tryptophan (W) | TGG |
| Cysteine (C) | TGT | Lysine (K) | AAA | Tyrosine (Y) | TAT |
| Glutamic Acid (E) | GAA | Methionine (M) | ATG | Valine (V) | GTT |
| Glutamine (Q) | CAA | Phenylalanine (F) | TTT | Stop | TAA |

2.3.2 Optimizing *L. reuteri* PAP via PCR stitching

For expression in *S. cerevisiae* (used in a related study), *E. coli*, and *L. reuteri*, four arginine codons were mutated from CGG to CGT in the *L. reuteri* PAP (LrPAP) sequence. Three of the four arginine codons in LrPAP WT were mutated via PCR stitching to create Rd1 mut. To confirm successful mutation of the 3 arginine codons, piece A, piece B, Rd1 mut, and LrPAP WT were amplified with Lr-PEP-EI-F/GFP-EI-R, and the resulting amplicons were digested with *RsaI* (Figure 2-3). Amplification and subsequent digest patterns were not expected from piece A or B; however, the *RsaI* digest on the piece B PCR product resulted in a pattern similar to that of Rd1 mut (511 + 253 bp). Digests on Rd1 mut and LrPAP WT (651 + 253 bp) yielded expected fragments. Although the 140 and 40 bp fragments were too small to detect on the agarose gel, the shift from the 651 bp band in LrPAP WT to the 511 bp fragment in Rd1 mut indicated the three arginine codons were successfully mutated. Rd1 mut was cloned into pCR2.1, and the resulting vector, pCR2.1 Rd1 mut, transformed into *E. coli* TOP10.

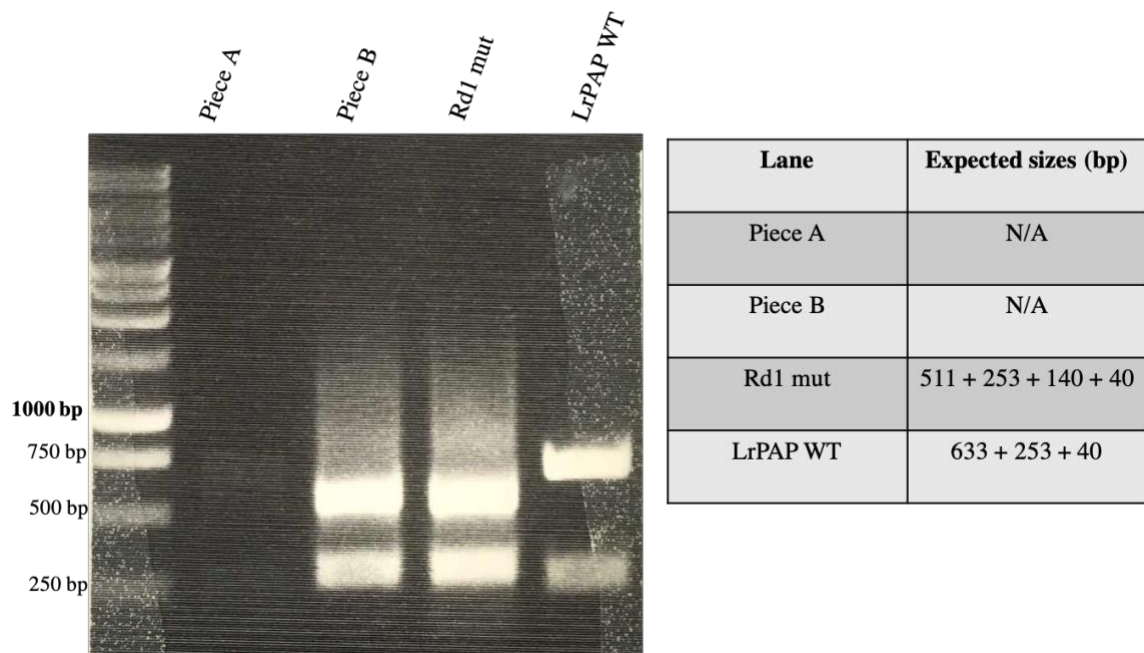


Figure 2-3. *RsaI* Restriction Digest on Rd1 Mut and LrPAP WT. Digests were analyzed by 2% agarose gel. RdI mut (511 + 253 bp) and LrPAP WT (633 + 253 bp) yielded the expected band sizes. The band shift from 633 bp in LrPAP WT to 511 bp in RdI mut demonstrate successful mutation of the 3 arginine codons.

Subsequently, pCR2.1 Rd1 mut was used as a template to generate Rd2 mut which contained all four mutated arginine codons. To confirm successful mutation of the final arginine codon, Rd1 mut, amplified with Lr-PAP-EI-F/Lr-GFP-Arg-R from pCR2.1 Rd1 mut and the PCR stitched Rd2 mut amplicon were digested with *SalI* (Figure 2-4). The expected digest patterns were observed from Rd1 mut (503+437 bp) and Rd2 mut (437+372+135 bp). Although the 135 bp fragment was not seen from the digest on Rd2 mut, the shift from the 503 bp fragment in Rd1 mut to the 372 bp band in Rd2 mut indicated that an additional *SalI* was incorporated and the arginine codon was mutated. Rd2 mut was the optimized version of LrPAP used to clone into pET30.

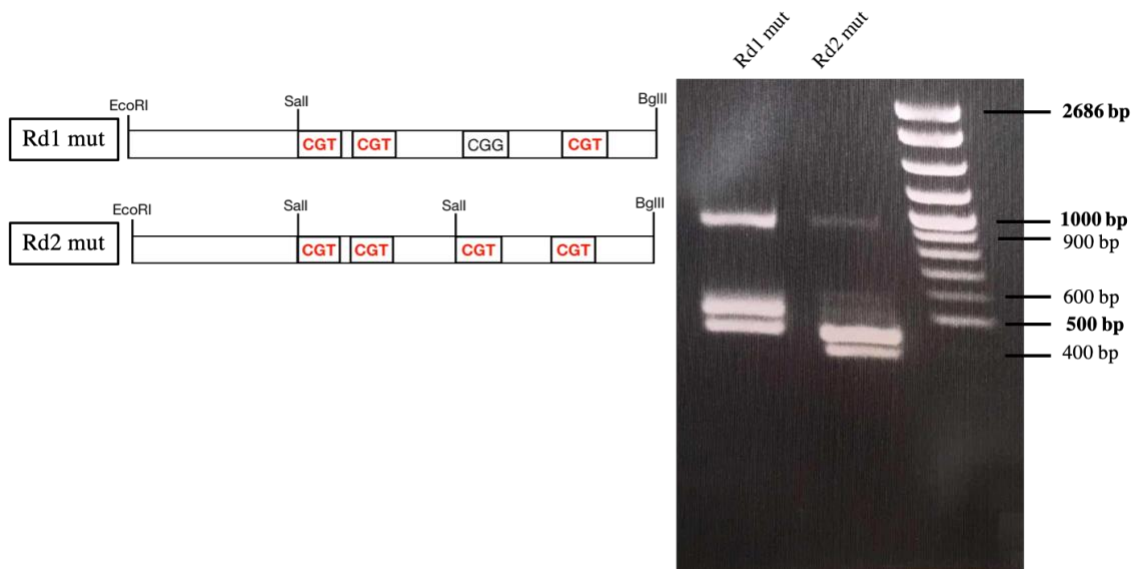


Figure 2-4. *SalI* Restriction Digests on Rd1 Mut and Rd2 Mut. Digests were analyzed by 2% agarose gel. Both Rd1 mut (503+437 bp) and Rd2 mut (437+372+135 bp) yielded expected band sizes which indicated the addition of a *SalI* restriction site and successful mutation of the fourth arginine codon in Rd2 mut.

2.3.3 Construction and transformation of pET30-derived expression vectors into *E. coli* BL21(DE3)

The pET30 vectors containing the AnPat PAP, MxPEP, *L. acidophilus* PAP (LaPAP), and LrPAP were digested with *Asc*I to determine if the vectors were transformed into *E. coli* BL21(DE3). With the exception of pET30 GFP, all *Asc*I restriction digests on the pET30-derived vectors yielded the expected band sizes (Figure 2-5). These pET30 vectors were successfully transformed into *E. coli* BL21(DE3).

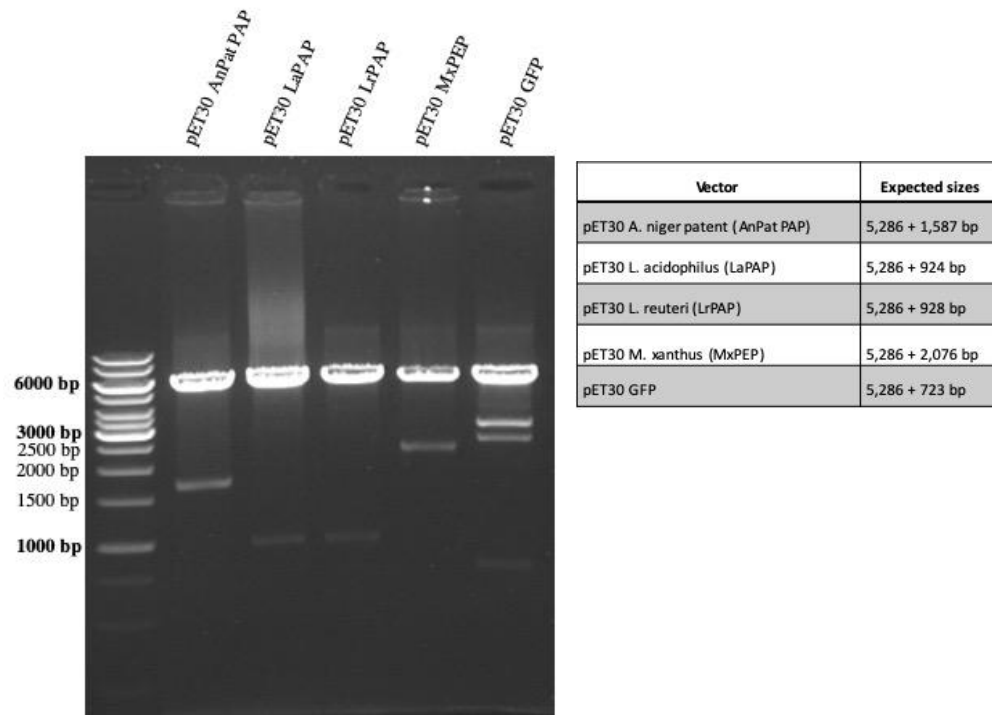


Figure 2-5. *Asc*I Restriction Digests on pET30 AnPat PAP, LaPAP, LrPAP, MxPEP, and GFP. Digests were analyzed on a 1% agarose gel. The *Asc*I digests resulted in the expected band sizes for pET30 AnPEP, pET30 AnPat PAP, pET30 LaPAP, pET30 LrPAP, and pET30 MxPEP confirming successful construction of

2.3.4 Expression and purification of PAPs and PEPs in *E. coli* strain BL21(DE3)

The expression of each PEP and PAP in BL21(DE3) was analyzed via SDS-PAGE. LaPAP (38.13 kDa), MxPEP (79.83 kDa), and LrPAP (37.81 kDa) were purified at the expected sizes from the soluble cell lysate fractions (Figure 2-6A and Figure 2-6B). According to figure 2-6C, AnPat PAP (61.31 kDa) was not detected in any of the soluble samples; however, three bands (~55, 40, and 15 kDa) become more apparent in both insoluble samples (AnPat PAP 90 pellet and AnPat PAP 90 pellet purified). Additionally, a ~60 kDa band was further purified from the insoluble fraction of the induced samples (AnPat PAP 90 pellet purified). The ~55, 40, and 15 kDa fragments from the purified AnPat PAP pellet were detected with Ni-HRP; however, the ~60 kDa fragment was not observed (Figure 2-7). GFP (29.90 kDa) was used to control for the induction with IPTG and purification with Ni-NTA resin.

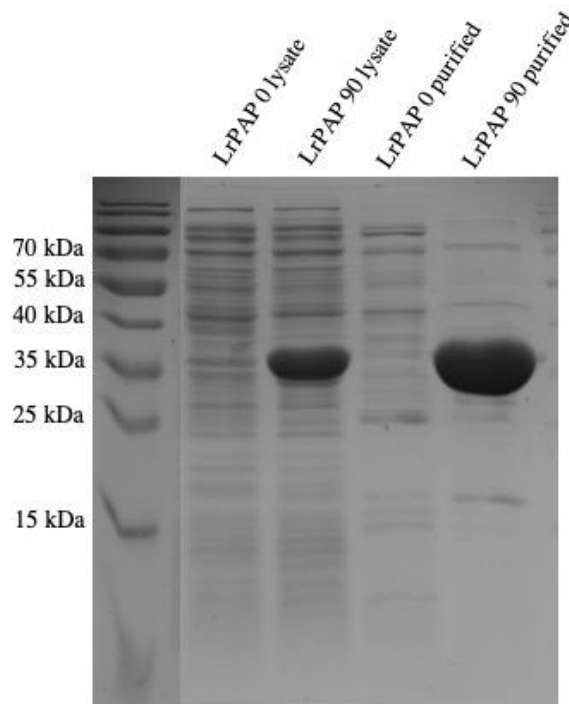


Figure 2-6B. Lysates and Purified Samples of GFP and LrPAP. Samples were analyzed by SDS-PAGE (12%). The expected sizes for GFP and LrPAP were 29.90 kDa and 37.81 kDa, respectively. The gel contains lysates, pellets, and purified samples that were not induced (0) and induced with IPTG for 90 minutes (90).

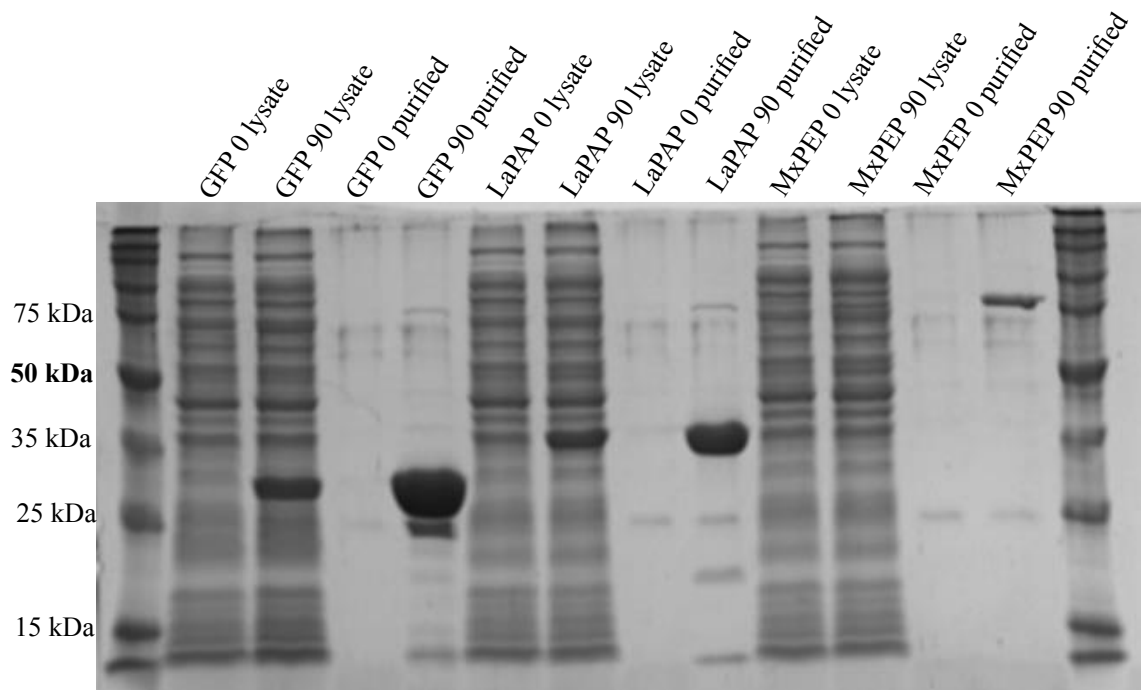


Figure 2-6A. Lysates and Purified Samples of GFP, LaPAP, and MxPEP. Samples were analyzed by SDS-PAGE (12%). The expected sizes for GFP, LaPAP, and MxPEP were 29.90 kDa, 38.13 kDa, and 79.83 kDa, respectively. The gel contains lysates and purified samples that were not induced (0) and induced with IPTG for 90 minutes (90).

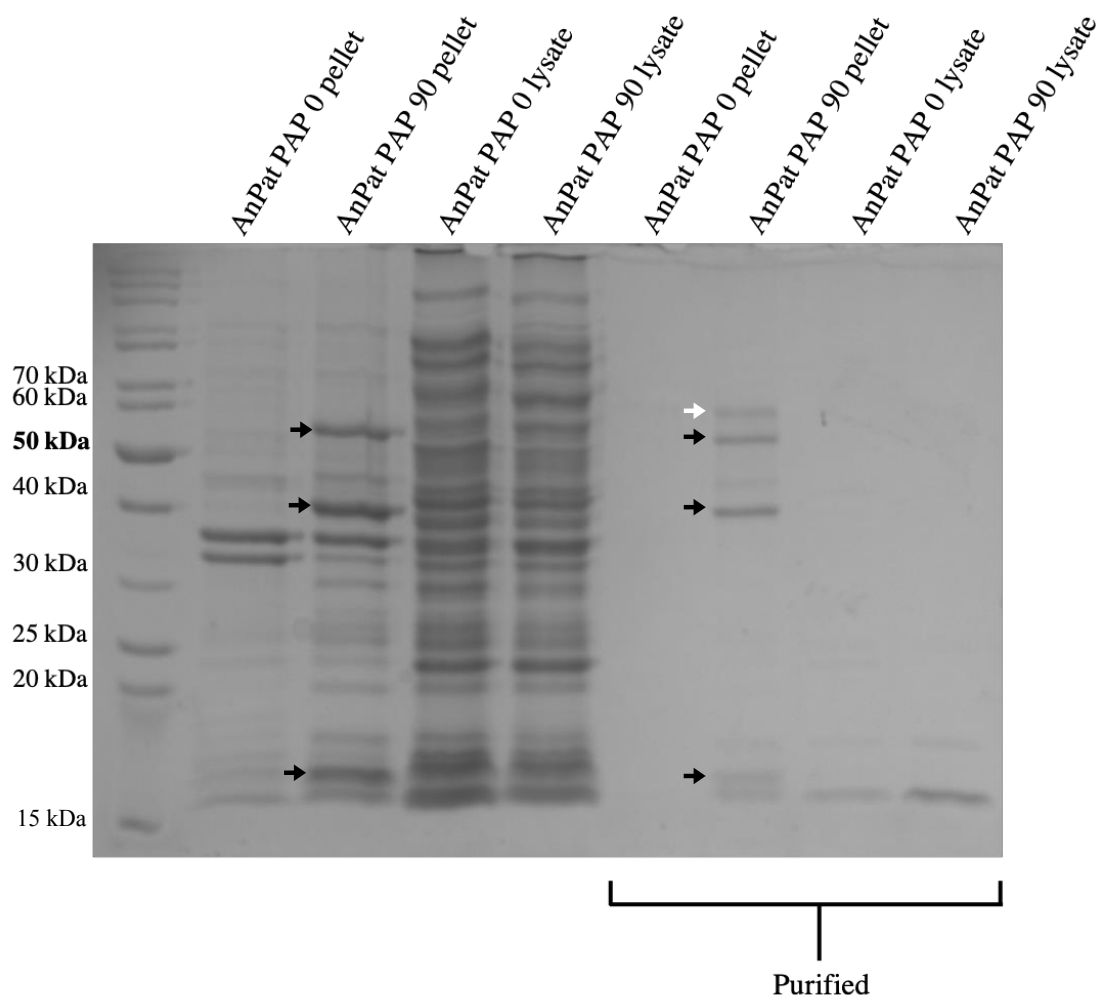


Figure 2-6C. Pellets, Lysates, and Purified Samples of AnPat PAP. The expected sizes for GFP and AnPat PAP were 29.90 kDa and 61.31 kDa, respectively. SDS-PAGE contains pellets, lysates, and purified samples that were not induced (0) and induced with IPTG for 90 minutes (90). Both pellets from the induced samples contained three unique bands (~55, 40, and 15 kDa) that were not apparent in the other AnPat PAP samples (black arrows). Additionally, a ~60 kDa band was detected in the purified AnPat PAP 90 pellet (white arrow).

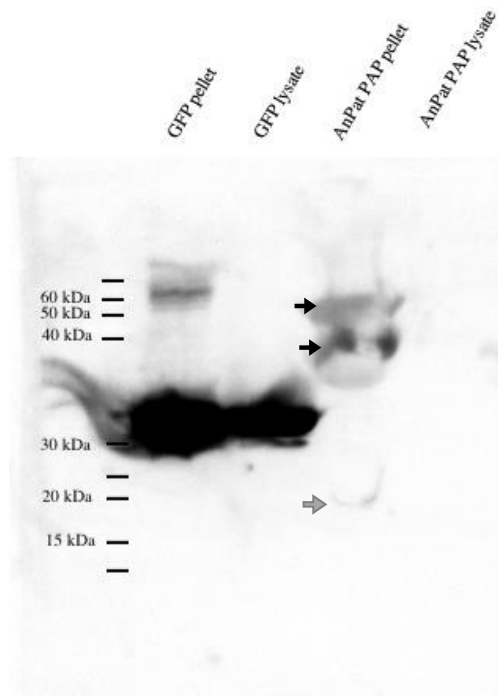


Figure 2-7. Western Blot of GFP and AnPat PEP. All samples were purified with Ni-NTA resin and probed with 1:500 Ni-HRP. As expected, GFP was detected 29.90 kDa. The ~55 and 40 kDa fragments (black arrows) were strongly detected, and the ~15 kDa fragment (gray arrow) was faintly detected from the AnPat PAP pellet.

2.3.5 Enzyme activity

AnPatPEP was not expressed as expected in BL21(DE3), therefore, only LaPAP, LrPAP, and MxPEP were further evaluated for cleavage activity. The protein concentrations of LaPAP, LrPAP, and MxPEP were determined to be 26.63 μ M, 112.13 μ M, and 21.65 μ M, respectively.

According to Beer's law ($A=\epsilon bc$), the absorbance of a sample is dependent on the concentration and molar absorptivity of the substance in question as well as the pathlength of light. Because a vertical beam of light was used in this assay, the pathlength was

dependent on the volume of the reaction. The absorbances of 200 μ l volumes of diluted PNP samples yielded an average pathlength of 0.675 cm. From this pathlength, ϵ_{pNA} was calculated to be 13037.037 $\text{M}^{-1} \text{cm}^{-1}$. This value was used to determine V_{max} , K_{m} , and k_{cat} for LaPAP, LrPAP, and MxPEP (Table 2-4).

In triplicate reactions that assessed the cleavage activity of LaPAP or LrPAP on H-Pro-pNA, the K_{m} values of LaPAP and LrPAP were calculated to be 0.501 mM (Figure 2-8A) and 0.625 mM (Figure 2-8B), respectively. LaPAP yielded a $V_{\text{max}}=1.43 \times 10^{-5}$ IU and LrPAP exhibited $V_{\text{max}}= 6.49 \times 10^{-5}$ IU. Neither LaPAP nor LrPAP exhibited cleavage activity on the Suc-Ala-Pro-pNA substrate (data not shown).

In four replicate reactions containing MxPEP and Suc-Ala-Pro-pNA, MxPEP exhibited a $K_{\text{m}}=0.685$ mM and $V_{\text{max}}=2.87 \times 10^{-3}$ IU (Figure 2-8C). There was no cleavage activity from MxPEP on the H-Pro-pNA substrate (data not shown).

Table 2-4. Summary of the Kinetic Parameters (V_{max} , K_{m} , k_{cat} , and $k_{\text{cat}}/K_{\text{m}}$) for LaPAP, LrPAP, and MxPEP.

| Enzyme | V_{max} (IU) | K_{m} (mM) | k_{cat} (sec ⁻¹) | $k_{\text{cat}}/K_{\text{m}}$ (sec ⁻¹ mM ⁻¹) |
|--------------|-----------------------|---------------------|---------------------------------------|---|
| LaPAP | 1.43×10^{-5} | 0.501 | 1.99×10^{-3} | 3.98×10^{-3} |
| LrPAP | 6.49×10^{-5} | 0.625 | 1.80×10^{-2} | 2.88×10^{-2} |
| MxPEP | 2.87×10^{-3} | 0.685 | 15.937 | 23.26 |

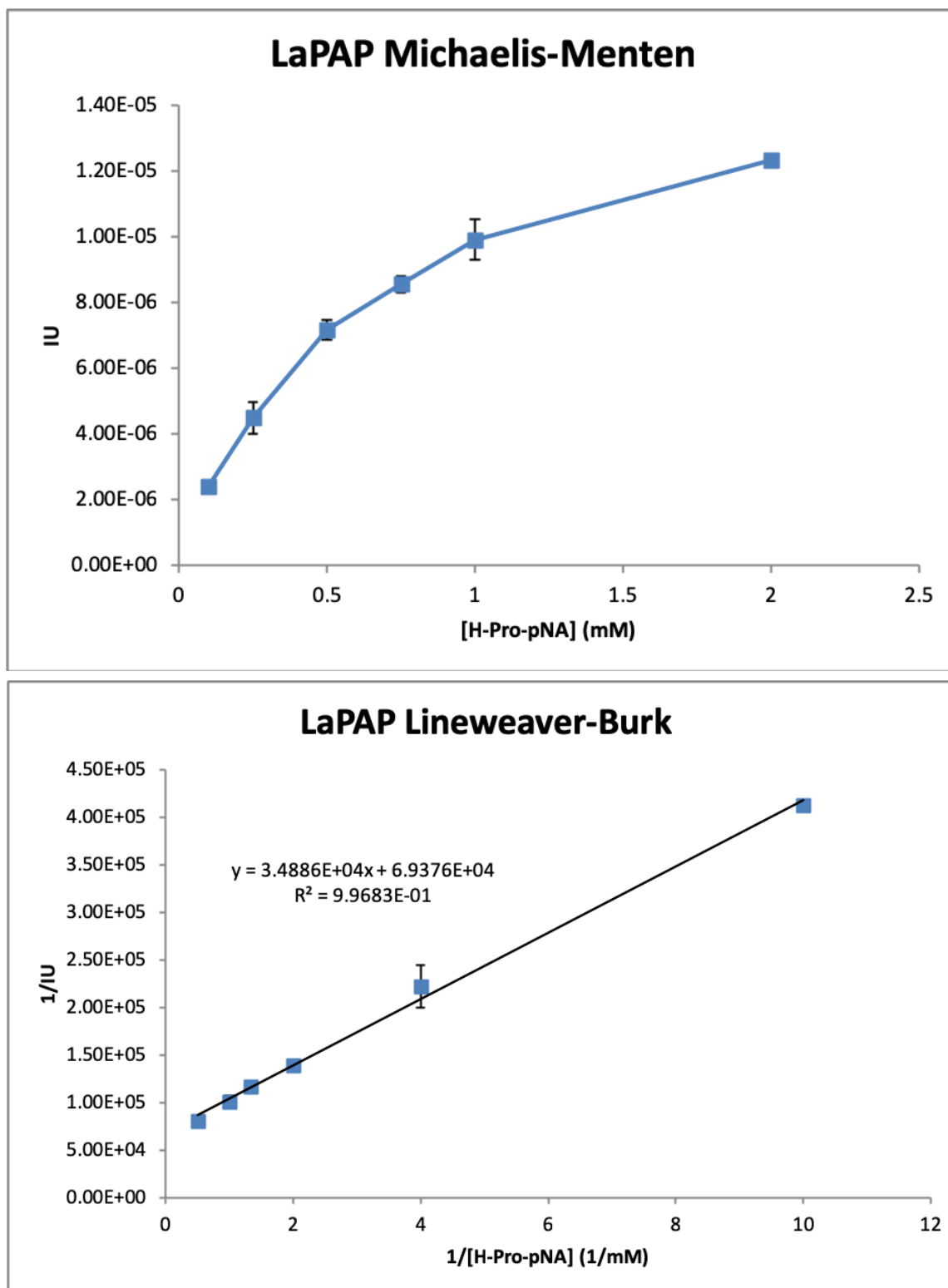


Figure 2-8A. LaPAP Enzyme Kinetics with H-Pro-pNA. (a) Michaelis-Menten plot. (b) Lineweaver-Burk plot that yielded a linear regression line with a slope of 3.49×10^4 and a y-intercept of 6.94×10^4 . LaPAP exhibited $V_{\max} = 1.43 \times 10^{-5}$ IU and $K_m = 0.501$ mM. All values represent the mean \pm SD.

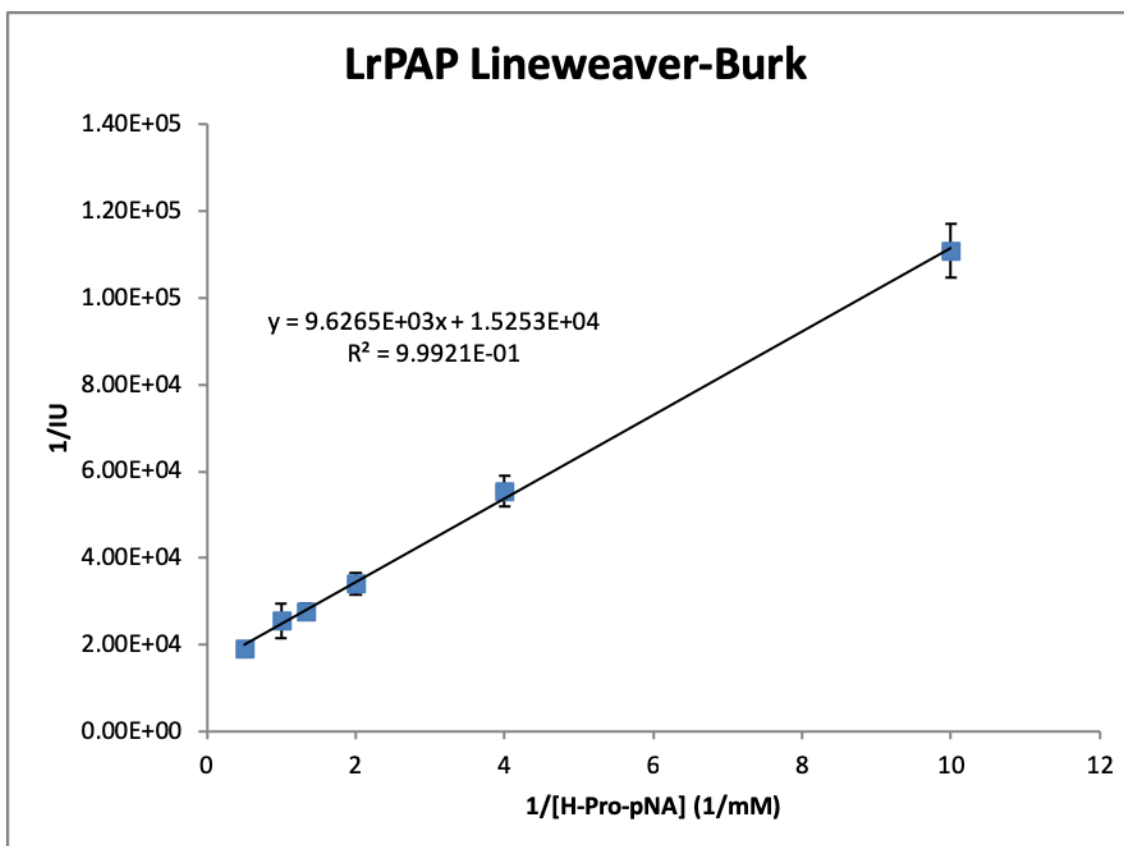
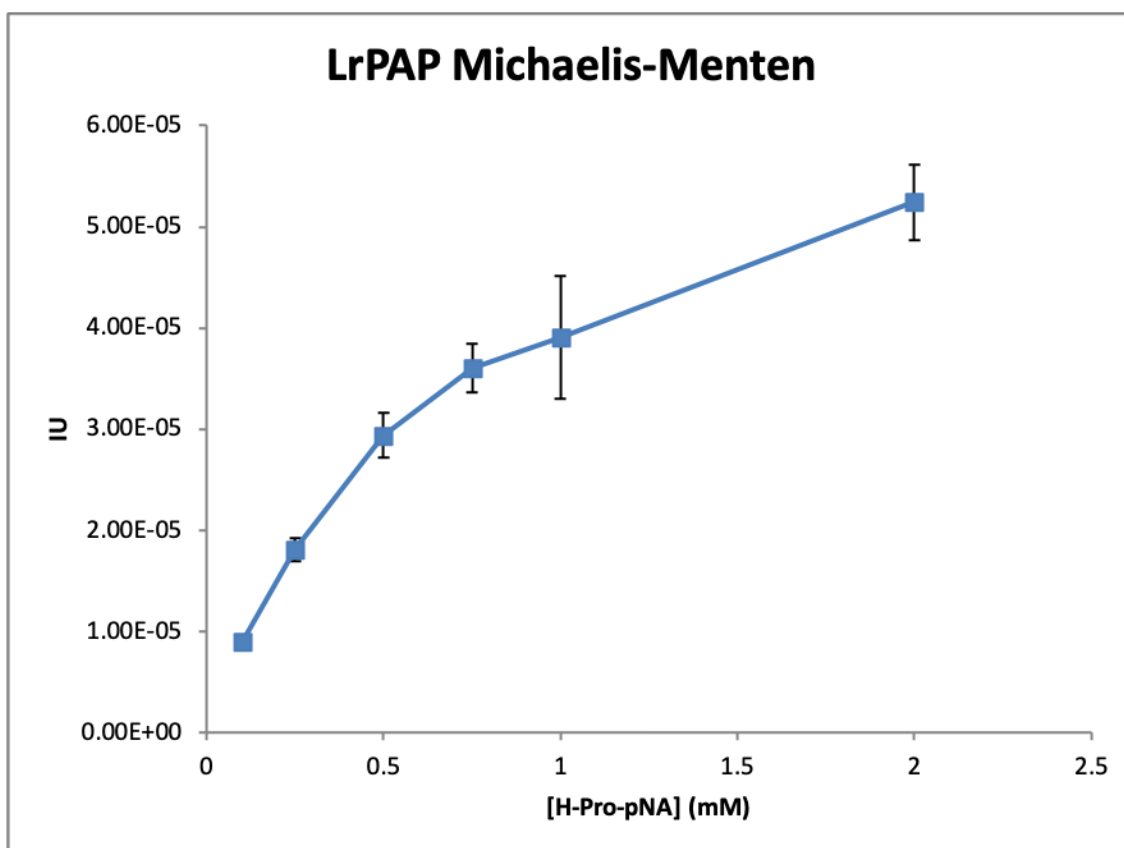


Figure 2-8B. LrPAP Enzyme Kinetics with H-Pro-pNA. (a) Michaelis-Menten plot. (b) Lineweaver-Burk plot that yielded a linear regression line with a slope of 9.63×10^3 and a y-intercept of 1.52×10^4 . LrPAP exhibited $V_{\max} = 6.49 \times 10^{-5}$ IU and $K_m = 0.625$ mM. All values represent mean \pm SD.

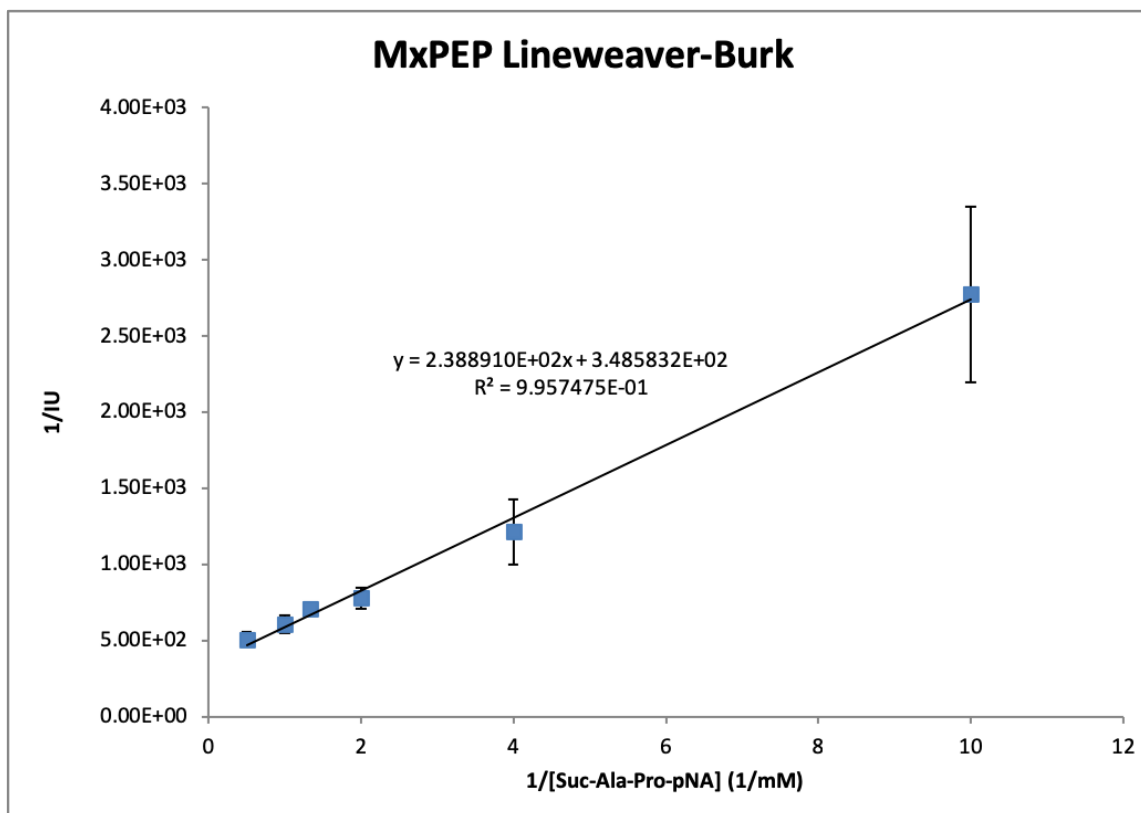
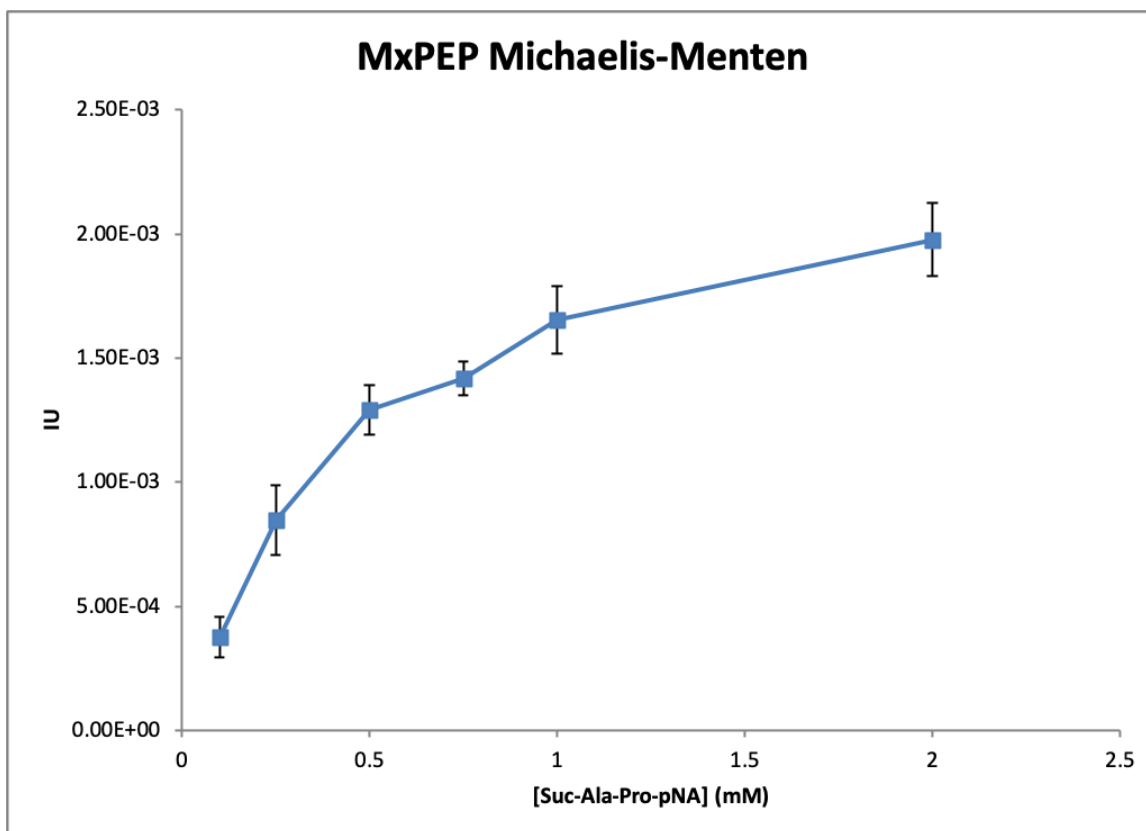


Figure 2-8C. MxPEP Enzyme Kinetics with Suc-Ala-Pro-pNA. (a) Michaelis-Menten plot. (b) Lineweaver-Burk plot that yielded a linear regression line with a slope of 2.39×10^2 and a y-intercept of 3.49×10^2 . MxPEP exhibited $V_{\max} = 2.87 \times 10^{-3}$ IU and $K_m = 0.685$ mM. Each value represents the mean \pm SD.

2.4 Discussion

The AnPat PAP, MxPEP, and LrPAP sequences were successfully optimized for expression in *E. coli* and *L. reuteri*. Following optimization of these sequences, all pET30 expression vectors were successfully cloned into BL21(DE3) (Figure 2-5), but only LaPAP, LrPAP, and MxPEP were expressed as soluble proteins that could be purified (Figure 2-6A and Figure 2-6B). Because all pET30 vectors contained a histidine tag at the C-terminus, the 3 AnPat PAP fragments (~55, 40, and 15 kDa) purified from the insoluble fraction of the cell lysate suggested that AnPat PAP was most likely cleaved at 3 possible sites on the N-terminus. Furthermore, these fragments were detected by western blot. Altogether, these data suggest that AnPat PAP may have been cleaved prior to protein purification. Cleavage that occurred post purification would result in noticeable N-terminal fragments on SDS-PAGE that were absent on the western blot. Moreover, a previous study reported a decrease in protein solubility with the introduction of synonymous codon substitutions (Cortazzo et al. 2002). It has been proposed that an increase in the rate of translation due the elimination of codon bias has an adverse effect on heterologous protein solubility (Rosano and Ceccarelli 2009). Although this may not be the definitive reason for the insoluble AnPat PAP fragments, alternate sequences of AnPat PAP should be considered if this enzyme is of interest for future use. Even though AnPat PAP was not successfully purified, these proteases would not be beneficial in the context of this study. AnPEP exhibits optimal activity at pH 2.5-4, but the duodenal pH ranges from 6-8 (Tsiatsiani et al. 2017). Thus, AnPEP activity would be low in this area of the small intestine. However, the optimal pH environment for AnPEP along with the report that AnPEP is resistant to pepsin degradation make this enzyme suitable for activity in the

stomach (Kubota, Tanokura, and Takahashi 2005; Stepniak 2006). Therefore, AnPEP would be a promising candidate as an oral enzyme therapy since it could digest gluten in stomach before it entered the small intestine. AnPEP, marketed as Tolerase G® by DSM, has been advertised as a dietary supplement that is active under gastric conditions to degrade the immunogenic epitopes of gluten (Salden et al. 2015). Because MxPEP, LaPAP, and LrPAP were successfully expressed as soluble proteins, the enzymatic activities of these proteases were evaluated to determine which would be incorporated into the *L. reuteri* expression cassette.

The activities of LaPAP and LrPAP were assessed with H-Pro-pNA because PAPs are known to have terminal cleavage activity. This was supported when neither the LaPAP nor LrPAP exhibited cleavage activity with Suc-Ala-Pro-pNA (data not shown). Analysis on LaPAP and LrPAP revealed that these enzymes have a K_m of 0.501 mM and 0.625 mM, respectively. Thus, these data suggest that LaPAP had a higher affinity for H-Pro-pNA than LrPAP. Because MxPEP acts upon internal proline residues, the activity of this enzyme was assessed with Suc-Ala-Pro-pNA. The MxPEP analyzed in this study had $K_m = 0.685$ mM; however, previous studies have shown MxPEP to have $K_m = 0.2-0.4$ mM when used to cleave Suc-Ala-Pro-pNA (Shan et al. 2004; Shan, Mathews, and Khosla 2005; Kocadag Kocazorbaz and Zihnioglu 2017). K_m values show that the MxPEP used in this study had a lower affinity towards Suc-AlaPro-pNA than the MxPEP used in the other studies. Moreover, the k_{cat}/K_m , indicator for catalytic efficiency, revealed a discrepancy from previous reports on MxPEP. In this study the k_{cat}/K_m for MxPEP was 23.26 sec⁻¹ mM⁻¹ whereas Shan et al. and Kocazorbaz and Zihnioglu reported 97 and 20.33 sec⁻¹ mM⁻¹,

respectively (Shan et al. 2004; Kocadag Kocazorbaz and Zihnioglu 2017). Overall, the MxPEP characterized in this study showed sufficient cleavage activity against Suc-Ala-Pro-pNA; however, it appeared to exhibit similar or lower effectiveness than previous investigations. Both aforementioned groups employed a C-terminal histidine tag to isolate MxPEP from whole cell lysates, therefore, the C-terminal histidine tag utilized in this study should not have any notable interference on MxPEP activity. It cannot be elucidated why there is a marked difference in catalytic efficiency of MxPEP between this study and Shan et al..

Although LrPAP, LaPAP, and MxPEP all cleaved their respective substrates, PEPs would be more effective than PAPs to cleave the immunogenic epitopes of gluten in the context of CD. The 33-mer (LQLQPFPPQLPYPQPQLPYPQPQLPYPQPQPF) that bestows the immunogenic properties on gliadin contains a handful of internal proline residues that would not be accessible by a PAP. Thus, both LaPAP and LrPAP would be ineffective against this peptide as a whole. If the 33-mer was broken down into smaller components, then the PAPs might be able access any terminal prolines to further detoxify the epitopes. An enzyme such as MxPEP that can access the internal proline residues is more valuable to this study. Moreover, MxPEP was reported to exhibit optimal enzymatic activity at a pH of 7, thus it would be a strong candidate for gluten degradation in the duodenum which ranges from pH 6-8. Because MxPEP exhibited features of interest to this study, it was used in the construction of the *L. reuteri* expression cassette in pGKMCS.

3.0 Construction of a Vector-based Expression Cassette to Assess the Activity of MxPEP in *Lactobacillus reuteri*

3.1 Introduction

3.1.1 Probiotics as a delivery vehicle

As mentioned in chapter 2, PEPs that cleave proline-containing oligopeptides may be effective in degrading the immunogenic epitopes of gluten to assuage the CD autoimmune response. Unfortunately, PEPs are notably absent from the gastrointestinal (GI) tract, therefore, a mechanism must be developed to introduce the enzymes into this environment. Orally administered enzymes are subjected to harsh conditions of the stomach, pancreatic juices, bile, and cleavage by brush border proteases that may render the enzymes inactive. An enzyme such as AnPEP would be valuable in the setting of the stomach since it maintains optimal activity at pH 2.5-4 and is resistant to pepsin degradation. As of 2015, the biotechnology company DSM launched Tolerase® G, a commercially available dietary supplement of AnPEP advertised to break down residual gluten. One drawback is that Tolerase® G must be consumed before each meal for the effects of the enzyme to take place. If an individual with CD forgets to administer the enzyme, then they are susceptible to the symptoms of accidental gluten exposure. Another mechanism of interest would be to maintain the PEP in duodenum, so an individual would not have to actively administer the enzyme prior to eating. PEPs delivered to the small intestine would need to be transported by a vehicle that can survive the environments of the GI tract and maintained at this site. Probiotic bacteria are well-suited for this function because they survive exposure to gastric acid in the stomach, colonize the small intestine, and are resistant to the bile salts and pancreatic juices that enter into the duodenum.

3.1.2 Applications of probiotics: food production and health benefits

According to the Food and Agriculture Organization of the United Nations and the World Health Organization, probiotics are live microbial organisms that promote healthy digestion and influence the gut microbial community of a host when administered in sufficient quantities. These bacteria are endogenous to the human GI and urogenital tracts as well as the oral cavity; however, probiotics may also be supplemented through lyophilized forms (e.g. tablets and capsules) or foods (e.g. yogurt, sauerkraut, kimchi, etc.). These bacteria are often used to modify flavors and textures of food products, inhibit the growth of bacteria that lead to food spoilage, and protect against food-borne pathogens in humans. During the production of yogurt, probiotics produce various organic acids (e.g. lactic acid, acetic acid, pyruvic acid) and volatile compounds (e.g. acetone, diacetyl, acetoin) that influence the overall flavor profile of yogurt. Nisin, an anti-microbial peptide produced by probiotic *Lactococcus lactis*, was licensed by the US Food and Drug Administration (USFDA) as a food additive for preservation and is commercially available under the name Nisaplin®. As an alternative to the use of antibiotics, probiotics have been used as feed additives to reduce infection of chickens and hens by *Salmonella*, thereby decreasing the risk of transmission of *Salmonella* through the human food supply chain. In addition to the contributions made within the food industry, supplementing probiotics into one's diet may directly confer great health benefits to the host.

There has been growing interest in the use of probiotics as orally delivered therapeutics for various GI disorders and health issues. In the murine gut, the presence of *Lactobacillus reuteri* corresponded to a reduction of fecal bacteria in the stomach when compared to that

of a *L. reuteri*-free murine gut and a reduction in urease activity, a virulence factor that promotes the survival of pathogenic bacteria in the stomach (Wilson et al. 2014). Other studies have shown that probiotics may be beneficial as an adjunct treatment for enteric pathogens such as *Helicobacter pylori*, which is associated with the development of gastric ulcers. Some *Lactobacilli* probiotics may reduce the severity of an infection by impairing the capability of *H. pylori* to colonize the gastric mucosa (Ljungh and Wadström, 2006; de Klerk et al. 2016). These probiotics competitively adhere to the gastric epithelium or inhibit adhesion genes to impede *H. pylori* colonization. Moreover, *Lactobacilli* have also been shown to aid in the prevention of irritable bowel syndrome and the reduction of treatment-related side effects such as antibiotic-associated diarrhea (Wilkins 2017). As described here, probiotics have intrinsically beneficial effects on a host, but these benefits may be further enhanced through genetic modification of the probiotics to express heterologous proteins.

3.1.3 Benefits of heterologous protein expression in probiotics

Lactic acid bacteria (LAB) such as *Lactobacillus*, *Pediococcus*, *Streptococcus*, *Oenococcus*, and *Lactococcus* are the most widely used probiotics. These bacteria are a group of Gram-positive, non-spore forming bacteria that primarily produce lactic acid from carbohydrate fermentation. Many LAB have been genetically engineered to express heterologous proteins for various applications. Recombinant *Lactococcus lactis* has been used to deliver ovalbumin (OVA), a protein found in egg whites, to the GI tract of OVA-immunized mice and induced peripheral tolerance to OVA (Huibregtse et al. 2007). Another study reported that mice orally immunized with recombinant *L. acidophilus*

expressing Hp0410 (*H. pylori* adhesion protein) yielded an increase in Hp0410-specific IgG and IgA in the serum and intestinal mucosa, respectively (Hongying et al. 2014). These mice were further challenged with *H. pylori* and exhibited lower levels of colonization by the pathogen in the stomach. In another study, post-weaning Balb/c mice, an ideal animal model for studying lactose intolerance, were orally administered recombinant *L. lactis* expressing heterologous β -galactosidase and challenged with lactose (Li et al. 2012). The presence of the recombinant *L. lactis* strain alleviated symptoms, such as diarrhea, commonly associated with lactose intolerance in these mice. As can be seen from these studies, recombinant probiotics expressing heterologous proteins confer propitious effects on a host. These bacteria can be used as vaccines to deliver antigens or therapies to supplement the host with enzymes that alleviate disorder-related symptoms. In this study, we propose the use of the probiotic *L. reuteri* to deliver an enzyme therapy to assuage the immune response following accidental gluten exposure.

3.1.4 The use of Lactobacillus reuteri as a therapy for CD

L. reuteri has been recognized as GRAS (generally regarded as safe) by the USFDA and is commonly used in the production of sourdough bread and other fermented cereals. This bacterium has been isolated from the GI tract of humans, pigs, rats, chickens, and guinea pigs (Hou et al. 2015). In humans, *L. reuteri* was found to colonize the stomach as well as the duodenum, jejunum, and ileum of the small intestine (Reuter 2001; Valeur et al. 2004). Because this bacterium is indigenous to the small intestine, *L. reuteri* commonly outcompete foreign probiotics, such as *L. lactis*, for colonization in the host (Walter, Britton, and Roos 2011). Strains of human *L. reuteri* are known to synthesize reuterin, a

unique broad-spectrum antimicrobial peptide produced during glycerol fermentation. Enteric pathogens (e.g., *E. coli*, *Shigella*, *Salmonella*, and *Vibrio*) and intestinal commensal bacteria (e.g. *Bifidobacterium*) have been shown to be susceptible to reuterin produced by *L. reuteri* (Park, Park, and Song 2008; Spinler et al. 2008). Thus, it has been hypothesized that the production of reuterin helps establish *L. reuteri* as a prominent member of small intestine microbiota. Moreover, *L. reuteri* has been shown to reverse a “leaky” gut by promoting the expression of tight junction proteins in lupus-prone mice (Mu et al. 2017). These characteristics of *L. reuteri* make it an ideal vehicle to deliver PEPs at the duodenum for the management of CD.

Several strains of *Lactobacilli* encode functional PEPs (Degraeve et. al., 2003; Sanz et. al., 2001; Rollan et. al., 2001; Xu et. al. 2001). Although *L. reuteri* expressed this peptidase activity and localizes to the site of pathogenesis, previous work in our lab revealed that it does not secrete the enzyme and only effectively digests the proline-containing substrates when lysed by sonication (Shurtleff 2009). As seen in chapter 2, MxPEP efficiently cleaved the chromogenic substrate Suc-Ala-Pro-pNA, and it is optimally active at a pH similar to that of the duodenum. Delivering MxPEP to the site of CD pathogenesis may be an effective therapy to assuage the immune response to gluten. The goal of this study was to genetically engineer *L. reuteri* to produce and secrete recombinant MxPEP.

3.1.5 Study Goal

In this study, a vector-based expression system was employed for the production of recombinant proteins in *L. reuteri*. The system required a promoter and terminator that could be utilized by *L. reuteri*. Protein expression was driven by the following: the erythromycin ribosomal methylase gene (*ermB*) promoter derived from the broad-host vector pAM β 1 (Kim, Baek, and Pack 1991). This constitutive promoter was used successfully to drive recombinant protein expression in *Lactobacillus spp.* (Lizier et al. 2010). The strong Rho-independent terminator from the r50 ribosomal L7/L12 gene (*rpIL*) of *E. coli* was used to terminate transcription of the cassette (Morita et al. 2015). A signal sequence from the *Bacillus licheniformis* amylase gene (*amyl*) was incorporated into the cassette to generate a secreted form of the recombinant protein. This signal sequence was successfully used to secrete heterologous α -amylase in *E. coli* as well as *L. reuteri* (Malik et al. 2013; Wu et al. 2006). As seen in chapter 2, MxPEP was determined to be more effective than the LaPAP and LrPAP in the chromogenic substrate assay. Thus, MxPEP was incorporated into an expression cassette that was cloned into the *L. reuteri* expression vector, pGKMCS. The cassette contained the *ermB* promoter, *amyl* signal sequence, MxPEP, and *rpIL* terminator for the expression of a secreted form of MxPEP. Another variant of the expression cassette was constructed without the signal sequence to generate a cytosolic version of MxPEP. Additional vectors were constructed to produce cytosolic and secreted GFP as controls. After transformation into *E. coli* with the cytosolic versions of the vector, transformants exhibited stronger GFP fluorescence when the sequence was oriented in the opposite direction relative to the *ermB* promoter. It was hypothesized that *lac* promoter contained in the MCS was interfering with expression from the *ermB*

promoter. The *lac* promoter was removed from the expression vector; however, the absence of this promoter resulted in vector instability.

3.2 Materials and Methods

3.2.1 Strains, plasmids, and growth conditions

E. coli strain MC1061 and *L. reuteri* strain 100-23C were utilized for this study. *E. coli* and *L. reuteri* cultures were grown in Luria Bertani (LB) and Man-Rogosa-Sharpe (MRS), respectively, broth or plates. To select for *E. coli* and *L. reuteri* containing pGKMCS-derived vectors 250 µl/mg erythromycin (Erm 250) was added to LB and 5 µg/ml erythromycin (Erm 5) was added to MRS, respectively.

3.2.2 Molecular techniques

All molecular techniques were the same as in chapter 2.

3.2.3 Digestion independent cloning

Digestion independent cloning (DIC) reactions contained 1X Phusion master mix and a 3:1 molar ratio of insert to vector (Figure 3-1). Thermocycling parameters for the DIC reactions were as follows: an initial denaturation at 98°C for 10 second; 10 cycles of 98°C for 1 second, 65°C for 5 seconds, and 72°C for 15 seconds per kb of product; and a final extension at 72°C for 1 minute. The thermocycling parameters for the construction of the pGKMCS *ermB fixed* vectors were identical as previously stated; however, the annealing temperature of 65°C was omitted.

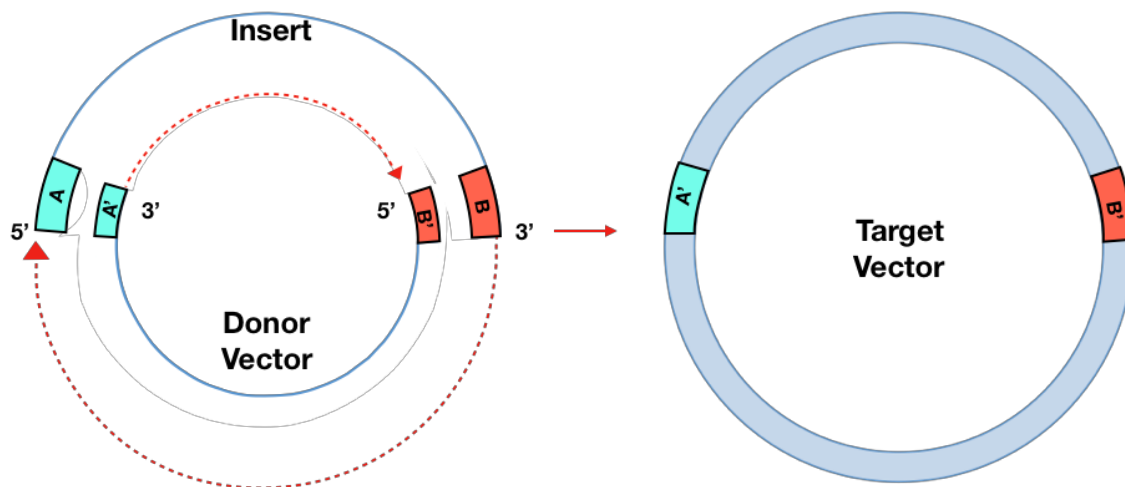


Figure 3-1. Digestion Independent Cloning (DIC). The single stranded insert (A and B) anneal to complementary regions (A' and B') within the single stranded donor vector. The donor vector acts as a template to extend the 3' end of the insert and form the double stranded target vector.

3.2.4 Construction of pGKMCS-derived vectors

The pGKMCS vectors containing the *Enterococcus faecalis* erythromycin ribosomal methylase gene (*ermB*) promoter and the signal sequence from the *Bacillus licheniformis* α -amylase gene (*amyl*) were constructed as follows. According to figure 3-2, the *ermB* promoter containing a 5' portion of the *amyl* signal sequence (Figure 3-2A) was amplified from a DNA string containing the *ermB* promoter as well as a *Lactobacillus acidophilus* lactate dehydrogenase gene (*ldh*) promoter, synthesized by Life Technologies (Carlsbad, CA, USA), with primers ErmB-AmyL-DIC-F/Erm-AmyL-DIC-R. An intermediate vector, pGKMCS 373 *GFP Link*, was used as a template to generate a DNA segment with the *amyl* signal sequence, GFP, his tag, and *rplL* terminator (Figure 3-2B) with the AmyL-F/rpIL-R primers. Subsequently, piece A and B were PCR stitched with the ErmB-AmyL-DIC

F/rpIL-R primers to create an amplicon containing the *ermB* promoter, *amyl* signal sequence, GFP, his tag, and *rplL* terminator (Figure 3-2AB). DIC was used to clone piece AB into pGKMCS 373 *GFP Link* that was digested with *EcoRI* and *AscI* to create pGKMCS *ermB amyl GFP*. The reaction contained 1X Phusion master mix (Thermo Scientific, Rockford, IL, USA) and a 3:1 molar ratio of insert to vector. Thermocycling parameters for DIC were as follows: an initial denaturation at 98°C for 10 seconds, 10 cycles of 98°C for 1 second, 65°C for 5 seconds, and 72°C for 23 seconds, and a final extension at 72°C for 1 minute. pGKMCS *ermB amyl GFP* was digested with *AscI* to remove GFP and ligated with MxPEP digested with *AscI* to produce pGKMCS *ermB amyl MxPEP*. Table 3-1 contains the primers used to construct these vectors.

In addition, pGKMCS vectors were created without the *amyl* signal sequence to produce a cytosolic version of the recombinant proteins. The *ermB* promoter was amplified from the *ermB/ldhL* DNA string with ErmB-AmyL-DIC-F/Erm-AscI-R. This generated an *ermB* promoter flanked by *EcoRI* on the 5' end and *AscI* on the 3' end. The amplified *ermB* promoter and pGKMCS *ermB amyl GFP* were digested with *EcoRI* and *AscI*, and these two fragments were ligated together to produce pGKMCS *ermB*. Subsequently, pGKMCS *ermB* and GFP or MxPEP were digested with *AscI* and ligated together to produce pGKMCS *ermB GFP* or pGKMCS *ermB MxPEP*.

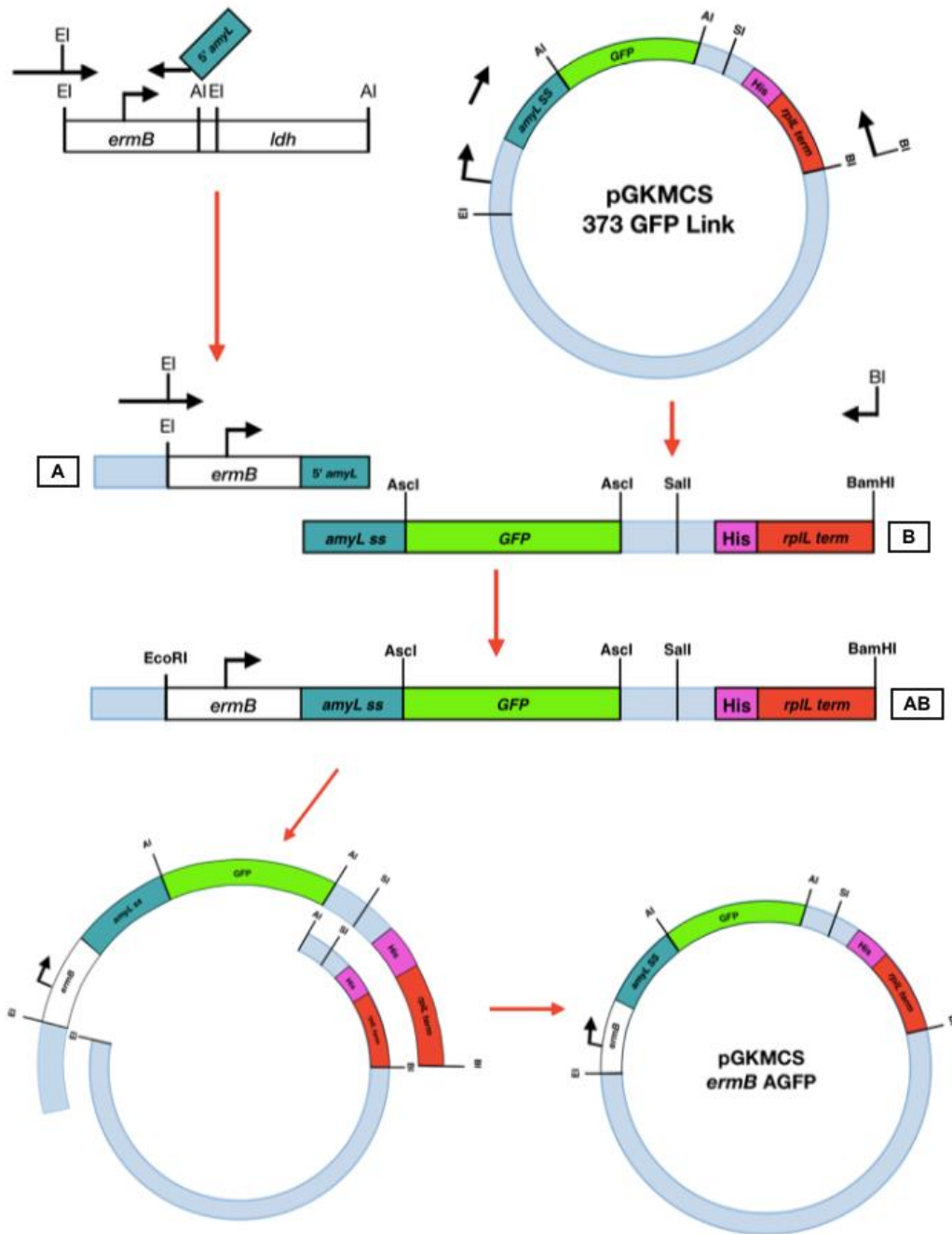


Figure 3-2. Construction of pGKMCS *ErmB* AGFP. The *ermB* promoter and a 5' end of the *amyL* signal sequence was amplified (**A**). An amplicon containing *amyL* signal sequence, GFP, histidine tag, and *rplL* terminator was generated from pGKMCS 373 GFP Link (**B**). Pieces A and B were PCR stitched to generate a *ermB* promoter, *amyL* signal sequence, GFP, his tag, and *rplL* terminator (**AB**). Following amplification, piece AB was incorporated into the vector via DIC. See text for details.

Table 3-1. Primers Used to Construct pGKMCS-derived Expression Vectors Containing the *ErmB* Promoter. Restriction sites are highlighted and identified in Features. The targets are the sections of DNA that each primer was designed to bind.

| Name | Sequence (5' →3') | Features | Target |
|------------------------|---|--------------|--|
| AmyL-F | TGAACCAACAAACGTTATATGC | | 5' end of <i>amyL</i> signal sequence |
| ErmB-AmyL-DIC-F | cggcagtgtagtagtatctgca gaa ttc aqtctagaatcg atac | EcoRI | 5' end of <i>ermB</i> promoter containing a section identical to pGKMCS |
| ErmB-AmyL-DIC-R | CGAGCATATAAACGTTTTTGTGTTCA t CCTTCTtaa ttac | | 3' end of <i>ermB</i> promoter containing a section identical to the 5' end of <i>amyL</i> signal sequence |
| ErmB-Ascl-R | gt ggcgcgcc actCCTTC T taattacaattttagc | Ascl | 3' end of <i>ermB</i> promoter |
| ErmB-Fix-F | cccttttatcaagaagcgca caaaa aag | | Used to incorporate 3 missing nucleotides (in bold and underlined) into the <i>ermB</i> promoter. |
| ErmB-Fix-R | caatttcgtttttcttt ttg ttgcgc | | Used to incorporate 3 missing nucleotides (in bold and underlined) into the <i>ermB</i> promoter. |
| rplL-R | ttgga ttcc aaaaaagc tg tgtagc | BamHI | 3' end of rplL terminator |

3.2.5 Construction of pGKMCS *ermB* fixed vectors

After initial construction of the pGKMCS vectors, it was discovered that the *ermB* promoter was missing three nucleotides that led to a lack of protein expression in *E. coli*. These vectors were rebuilt to incorporate the three missing nucleotides to create a functional *ermB* promoter. The 5' end of the *ermB* promoter was amplified with the ErmB-AmyL-DIC-F/Erm-Fix-R primers (Fig. 3-3C). Using pGKMCS *ermB GFP* as a template, the 3' end of the *ermB* promoter, GFP, his tag, and *rplL* terminator were amplified with the Erm-Fix-F/rplL-BI-R primers (Fig. 3-3D). The reaction was then treated with the restriction enzyme *DpnI* to remove pGKMCS *ermB GFP* from downstream reactions. Following the digest, piece C and D were PCR stitched with ErmB-AmyL-DIC F/rplL-BI-R to generate a DNA sequence containing the fixed *ermB* promoter, GFP, histidine tag, and *rplL* terminator (Fig. 3-3CD). DIC was performed to create pGKMCS *ermB fixed GFP*. Thermocycling parameters were as follows: an initial denaturation at 98 °C for 10 seconds, 10 cycles of 98 °C for 1 second and 72 °C for 23 seconds, and a final extension at 72 °C for 1 minute. For the construction of pGKMCS *ermB fixed MxPEP*, pGKMCS *ermB fixed GFP* was digested with *AscI* to remove GFP and MxPEP was cloned into the vector through ligation. Two additional vectors, pGKMCS *ermB fixed amyl GFP* and pGKMCS *ermB fixed amyl MxPEP*, were built to contain the *amyl* signal sequence. The procedure to create pGKMCS *ermB fixed amyl GFP* was identical to the construction of pGKMCS *ermB fixed GFP*; however, pGKMCS *ermB amyl GFP* was used as a template instead of pGKMCS *ermB GFP*. Table 3-1 contains the primers that were utilized in the construction of these fixed vectors. From here on any mention of a pGKMCS *ermB* vector will refer to these vectors with the three nucleotides added to the *ermB* promoter.

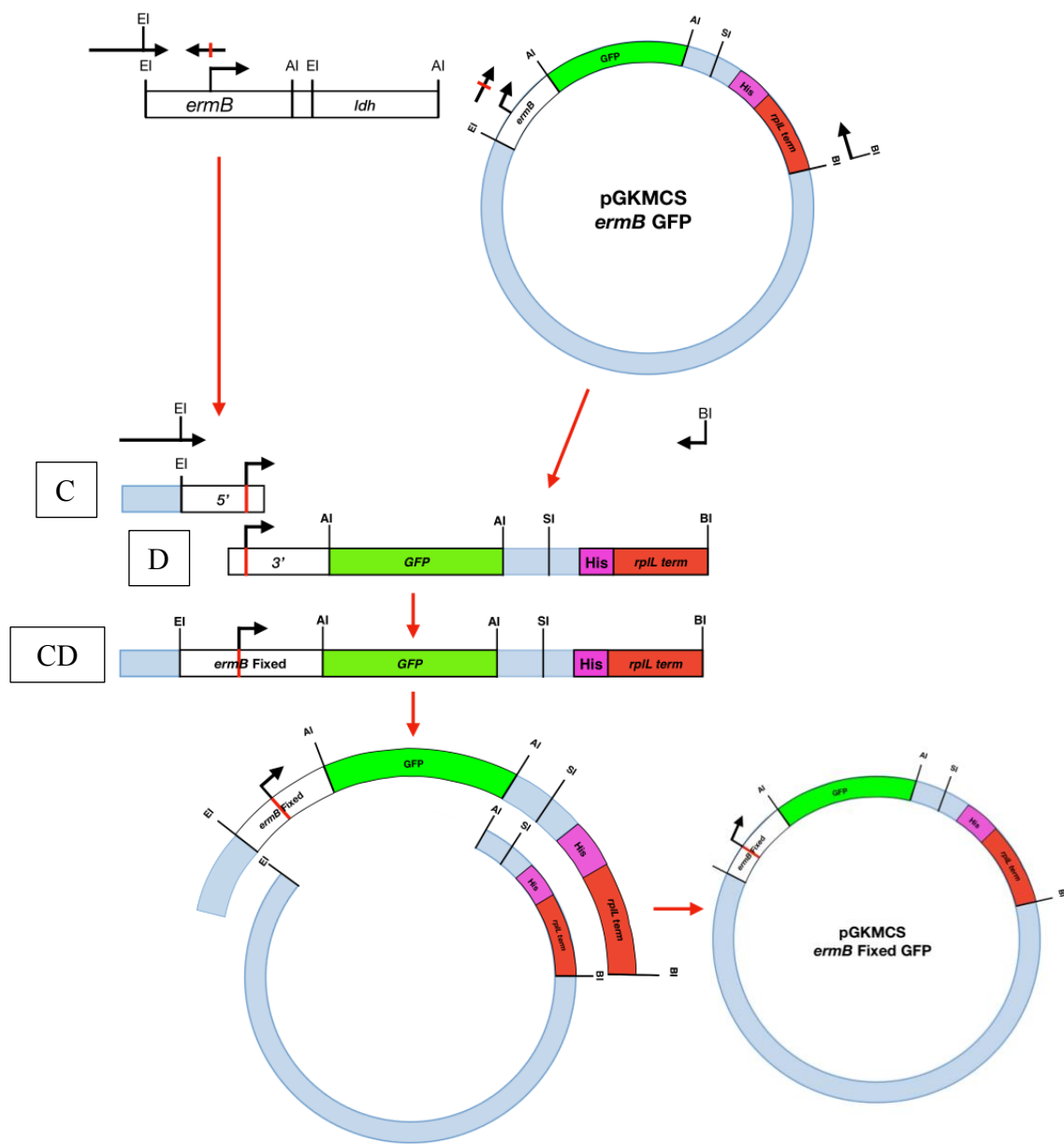


Figure 3-3. Construction of pGKMCS *ErmB* Fixed GFP Vector. The pGKMCS *ermB* fixed GFP vector contains 3 nucleotides in the *ermB* promoter that were omitted from the pGKMCS *ermB* GFP vector. The 5' end of the *ermB* promoter containing the 3 nucleotides was amplified from the *ermB/ldh* DNA string (C). The 3' end of the *ermB* promoter, GFP, His tag, and *rplL* terminator were amplified from pGKMCS *ermB* GFP (D). Pieces C and D were stitched to generate an amplicon containing the fixed *ermB* promoter, GFP, his tag, and *rplL* terminator (CD). See text for details.

3.2.6 Construction of pGKMCS *ermB* Δ *lac* promoter vectors

The *lac* promoter was included in the initial construction of the pGKMCS vector as part of the multiple cloning sequence (MCS) obtained from pCR2.1; however, it was removed from the pGKMCS-derived vectors in this study. The *lac* promoter was removed by amplifying pGKMCS *ermB GFP* with Del-Lac-Pro-BI-F and rplL-R. Amplification resulted in an amplicon flanked by *Bam*HI sites on both the 5' and 3' ends. The amplicon was digested with *Bam*HI and ligated to create pGKMCS *ermB GFP* Δ *lac* promoter. This process was repeated with pGKMCS *ermB MxPEP* and pGKMCS *ermB amyl MxPEP* to construct these vectors without the *lac* promoter.

3.2.7 Transformation of pGKMCS-derived vectors into *E. coli*

All pGKMCS-derived vectors were first transformed into *E. coli* strain MC1061 to obtain purified plasmids for subsequent transformation into *L. reuteri*. The preparation of electrocompetent MC1061 cells was the same as in chapter 2.

All electroporations were performed using the BTX Electro Cell Manipulator 600 (Harvard Apparatus Inc; 2.45 kV, 129 Ω , 25 μ F). For each electroporation, 40 μ l of cells transformed with 200 – 800 ng/ μ l pGKMCS-derived vectors were incubated at 37°C for 2 hour in 500 μ l SOC recovery medium supplemented with 100 ng/mL erythromycin. This sub-inhibitory concentration of erythromycin was required to induce expression of the *ermC* promoter. After recovery incubation, 100 μ l of cells were grown in LB Erm 250 broth and 100 μ l plated on LB Erm 250 agar and grown for 18 hours. For colony PCR (see materials

and methods in chapter 2), 11-22 CFUs were evaluated to determine if the colonies contained the vector of interest.

3.2.8 Assessing the stability of pGKMCS expression vectors

Cultures of *E. coli* MC1061 carrying pGKMCS *ermB GFP*, pGKMCS *ermB GFP Δ lac promoter*, pGKMCS *ermB MxPEP*, pGKMCS *ermB MxPEP Δ lac promoter*, pGKMCS *ermB amyL MxPEP* and pGKMCS *ermB amyL MxPEP Δ lac promoter* were each grown overnight in 4 ml LB Erm 250 at 37 °C. Plasmids were isolated from 3 ml of each broth culture, and 10 µl of the remaining broth of each culture was inoculated into 4 ml LB Erm 250 and grown overnight at 37°C. This process was repeated for 5 consecutive days. All purified vectors (200 ng – 330 ng) were digested with *XhoI* and *NcoI* to analyze the stability of the vectors in *E. coli*.

3.2.9 Plasmid sequencing

The pGKMCS *ermB GFP*, pGKMCS *ermB amyL GFP* pGKMCS *ermB GFP Δ lac promoter*, pGKMCS *ermB amyL MxPEP*, and pGKMCS *ermB amyL MxPEP Δ lac promoter* were sequenced by GENEWIZ (South Plainfield, NJ, USA). Table 3-2 indicates which primers were used for each vector to obtain the sequence data. The MCS-R primer targeted the region upstream from the *ermB* promoter in all pGKMCS vectors whereas ErmB-AmyL-DIC-R annealed to the 3' end of the *ermB* promoter. GFP-AscI-New-R and Mx-Opt-AscI-R bound to the 3' end of GFP and MxPEP, respectively. Chromatograms were assessed with Sequencher (Gene Codes Corporation, Ann Arbor, MI, USA).

Table 3-2. The pGKMCS Vectors and Associated Primers Used for Sequencing.

| Vector | Forward Primer | Reverse Primer |
|---------------------------------------|---------------------------|-----------------------|
| pGKMCS ermB GFP | MCS-R | ErmB-AmyL-DIC-R |
| pGKMCS ermB amyl GFP | MCS-R | GFP-AscI-New-R |
| pGKMCS ermB GFP Δ lac promoter | MCS-R | ErmB-AmyL-DIC-R |
| pGKMCS ermB amyl MxPEP | MCS-R | Mx-Opt-AscI-R |

3.2.10 Transformation of pGKMCS-derived vectors into L. reuteri

L. reuteri was grown in 3 ml MRS overnight and the following day it was diluted 1:100 in 10 ml MRS and grown at 37°C in a candle jar until OD₆₀₀ 0.6. The culture was centrifuged at 4,000xg for 10 minutes at 4°C and washed twice with ice-cold, sterile nanopure water. After washing the cells with water, the pellet was washed a final time in 1/10 volume of EHRG buffer (1 mM HEPES, pH 7.0, 0.5 M raffinose, and 10% glycerol), resuspended in 1/100 volume of EHR buffer, stored in 100 μ L aliquots, frozen in liquid at – 80°C.

All electroporations into *L. reuteri* were performed using the BTX Electro Cell Manipulator 600 (Harvard Apparatus Inc; 1.5 kV, 480 Ω , 25 μ F), and each 100 μ L aliquot was transformed with 200 – 800 ng of plasmid DNA. Immediately after transformation, cells were recovered in 1 mL pre-warmed MRS containing 100 ng/mL erythromycin and incubated in a candle jar at 37°C for 3 hours. After recovery, 250 μ L of transformed cells

were inoculated 4 ml MRS Erm 5 broth and grown overnight at 37°C in a candle jar for the enrichment of cells. The following day 3 ml of MRS Erm 5 was inoculated with 50 µL of enriched cells and grown overnight. Overnight cultures were streaked on MRS Erm 5 plates for the selection of colonies. To detect the presence of the transformed vector, 5-10 colonies were inoculated into 1 ml MRS Erm 5 and allowed to grow overnight. The following day 100 µl aliquots were pelleted, a pipet tip was used to obtain a small portion of each pellet, and the cells were resuspended in separate aliquots from a PCR master mix (primers and 1X GoTaq master mix).

3.2.11 L. reuteri cell lysis and protein purification

L. reuteri cultures harboring pGKMCS-derived vectors were grown overnight at 37°C in 4 ml MRS Erm 5 broth. Overnight cultures were diluted 1:50 in 10 ml MRS Erm 5 broth and grown to OD₆₀₀ = 0.5 – 0.7. After reaching log phase, cells were pelleted at 4,000xg for 10 minutes at 4°C. Pellets were washed twice in 1/10 volume of nanopure water and washed once in 1/100 volume TET buffer (20 mM Tris-HCl, pH 8.0, 2 mM EDTA, and 1.2% Triton X-100). After removing TET buffer, the pellets were resuspended in 1/10 volume lysozyme buffer (20 mM Tris-HCl, pH 8.0, 2 mM EDTA, and 1.2% Triton X-100) supplemented with 10 mg/ml lysozyme and 10 µg/ml mutanolysin. The cells were incubated at 37°C for 3 hours at 80 RPM. After the incubation period, samples were centrifuged at 4,000xg for 10 minutes at 4°C and the supernatant was discarded to remove the lysozyme and mutanolysin. The pellets were resuspended in 500 µl urea lysis buffer (20 mM KPO₄, pH 7.4, 8 M urea, and 0.5M NaCl) to lyse the cells. Samples were sonicated twice at ~14 RMS for 10 seconds to shear DNA molecules. Recombinant proteins containing a C-terminal

histidine tag were purified with HisPur™ Ni-NTA resin (Thermo Scientific, Rockford, IL, USA). Resin was prepared by washing twice with 500 µl nanopure water, once with 500 µl urea lysis buffer, and resuspended in 300 µl urea lysis buffer. Lysates were incubated on ice with ~5 µl bed volume of Ni-NTA resin for 30 minutes. Ni-NTA bound proteins were separated from unbound proteins by centrifugation at 14,000xg for 1 minute at 4°C, and the resin was washed twice with 500 µl urea lysis buffer.

3.2.12 SDS-PAGE and western blot

SDS-PAGE and western blot techniques were the same as in chapter 2.

3.3 Results

3.3.1 Expression and purification of GFP and MxPEP in *E. coli* MC1061

SDS-PAGE and western blots were used to analyze expression of GFP and MxPEP in *E. coli* MC1061 from pGKMCS *ermB GFP* and pGKMCS *ermB MxPEP*, respectively. GFP was not identified by SDS-PAGE; however, purified GFP (29.90 kDa) and a ~25 kDa band unique to the cell lysate appeared on the western blot when probed with anti-GFP conjugated to HRP (Figure 3-4A). Because the pGKMCS expression vectors contained a C-terminal histidine tag, GFP may have been cleaved at the C-terminus prior to purification with Ni-NTA to generate the ~25 kDa band. MxPEP (79.83 kDa) expression was confirmed via SDS-PAGE and western blot using Ni-HRP (Figure 3-4B). As seen in from SDS-PAGE and western blot, MxPEP was purified as a soluble protein from the cell lysate. A similar sized band was detected by Ni-HRP in the MxPEP purified pellet indicating that some MxPEP was insoluble. Additionally, a ~35 kDa band was identified by Ni-HRP in both cell lysates of MxPEP. As seen with GFP, cleavage within MxPEP may have occurred. Evaluation of secreted GFP and MxPEP was not carried out in *E. coli* since the *amyl* signal peptide (SP) from *B. licheniformis* is used to export proteins in Gram positive bacteria. Although expression of cytosolic GFP and MxPEP was detected, *E. coli* harboring an intermediate pGKMCS *GFP* vector exhibited stronger GFP fluorescence with the reverse complement of the gene relative to the promoter of interest (data not shown). This observation led to a deeper investigation into the backbone of the pGKMCS vector.

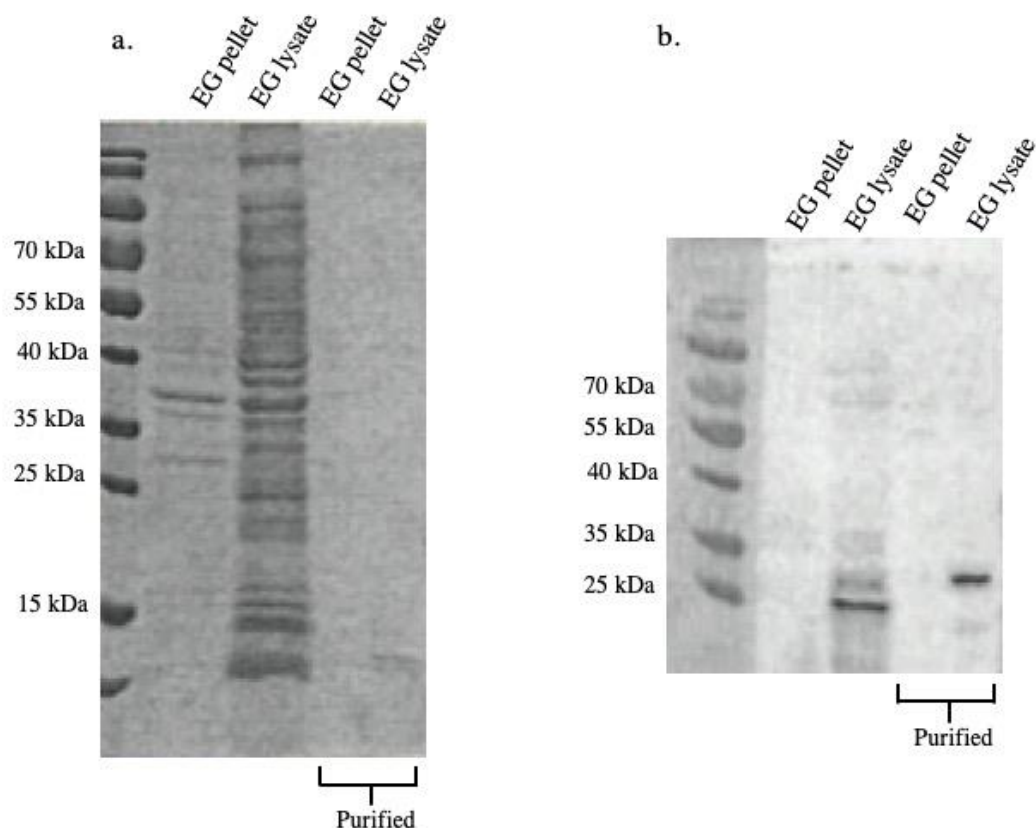


Figure 3-4A. Lysates and Purified Samples of GFP from pGKMCS *ErmB* GFP. Samples were analyzed by 12% SDS-PAGE and western blot probed with anti-GFP conjugated to HRP. **(a)** GFP was not observed via SDS-PAGE **(b)** GFP was detected at the expected size (29.90 kDa) by western blot from the cell lysate purified with Ni-NTA. An additional band ~25 kDa was detected by anti-GFP in the cell lysate. The pET30 GFP lane was used as a control for probing with anti-GFP.

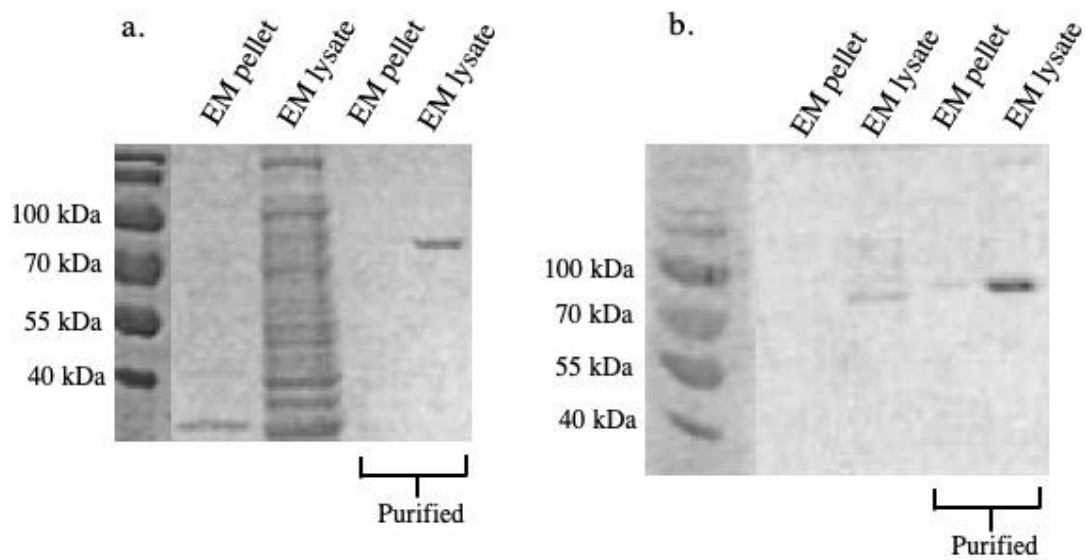


Figure 3-4B. Lysates and Purified Samples of MxPEP from pGKMCS *ErmB* *MxPEP*. Samples were analyzed by 12 % SDS-PAGE and western blot probed with Ni-HRP. **(a)** MxPEP was purified from the cell lysate and appeared at the expected size of 79.83 kDa via SDS-PAGE. **(b)** Western blot analysis shows that MxPEP was detected from the purified pellet and cell lysates. Additionally, a ~35 kDa band was detected by Ni-HRP in the whole cell lysate as well as the purified cell lysate samples.

3.3.2 Assessing the stability of pGKMCS expression vectors

It was discovered that the multiple cloning site of pGKMCS contained the *lac* promoter. *In silico* analysis revealed that the pGKMCS *ermB* expression vectors were constructed with the *lac* promoter located downstream and in the opposite orientation of the *ermB* promoter. Thus, the heightened GFP fluorescence that was associated with the reverse complement gene was likely due to expression from the *lac* promoter. We hypothesized that the *lac* promoter overpowered the *ermB* promoter and removed the *lac* promoter from the pGKMCS vectors in order to obtain unobstructed heterologous protein expression from *L. reuteri*.

After the removal of the *lac* promoter, *E. coli* transformants were screened via diagnostic digest with *Xho*I and *Nco*I; however, inconsistent digestion patterns were observed among transformants that were confirmed by colony PCR (data not shown). This was seen most prominently with cells transformed with pGKMCS *ermB amyI MxPEP ΔLP* (EAM ΔLP). Due to these observations, the stability of all pGKMCS expression vectors was evaluated by subculturing and isolating vectors over a longer period of time (5-6 days). These vectors were assessed through diagnostic digests with *Xho*I and *Nco*I. According to figure 3-5A, all pGKMCS *ermB GFP* (EG) cultures consistently yielded the expected digest patterns (4,466, 750, and 594 bp). Similarly, all pGKMCS *ermB GFP Δ lac promoter* (EG ΔLP) cultures produced the expected 4,308, 750, and 594 bp bands; however, a unique ~3,000 bp band developed by the 3rd day of subculturing (EG ΔLP rd 3-5).

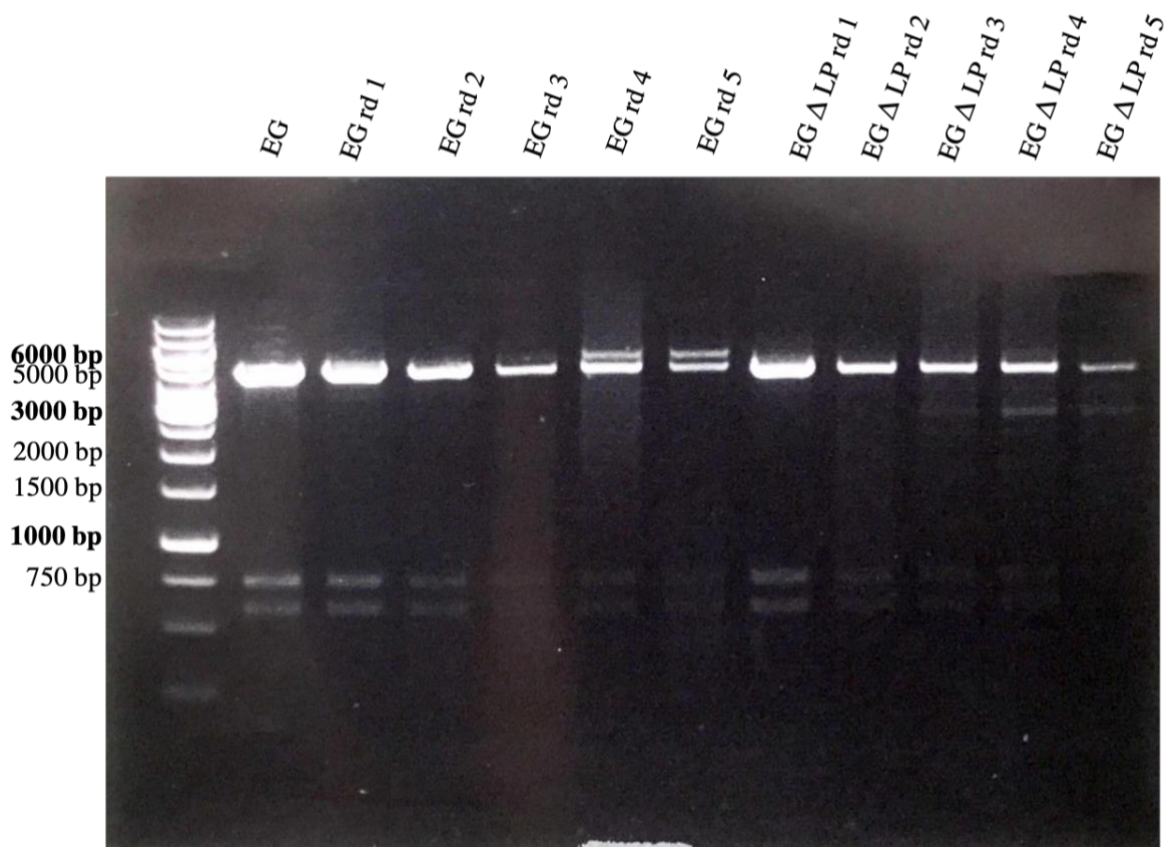


Figure 3-5A. Molecular Evolution of pGKMCS EG and EG Δ LP. Digests were analyzed on 1% agarose gel. EG rd 1-5 (4,466, 750, and 594 bp) and EG Δ LP rd 1-5 (4,308, 750, and 594 bp) vectors display the expected band patterns; however, EG Δ LP rd 3-5 produced a new band at \sim 3,000 bp.

All pGKMCS *ermB MxPEP* (EM) cultures maintained consistent but unexpected digest patterns (Figure 3-5B). The 4,466, 1,992, and 705 bp bands were expected; however, the \sim 3,500 bp band was not anticipated for these *XhoI* and *NcoI* digests (EM rd 1-5). This \sim 3,500 bp band may be attributed to the formation of covalently closed circular DNA which can form during alkaline lysis plasmid purification (Sayers, Evans, and Thomson 1996). Contrastingly, only pGKMCS *ermB MxPEP Δ lac promoter* (EM Δ LP) rd 1-3 maintained the expected 4,308, 1,992, and 705 bp patterns. Subsequent digests revealed a loss of the 1,992 bp band and the introduction of a \sim 2,500 bp band (EM Δ LP rd 4-5).

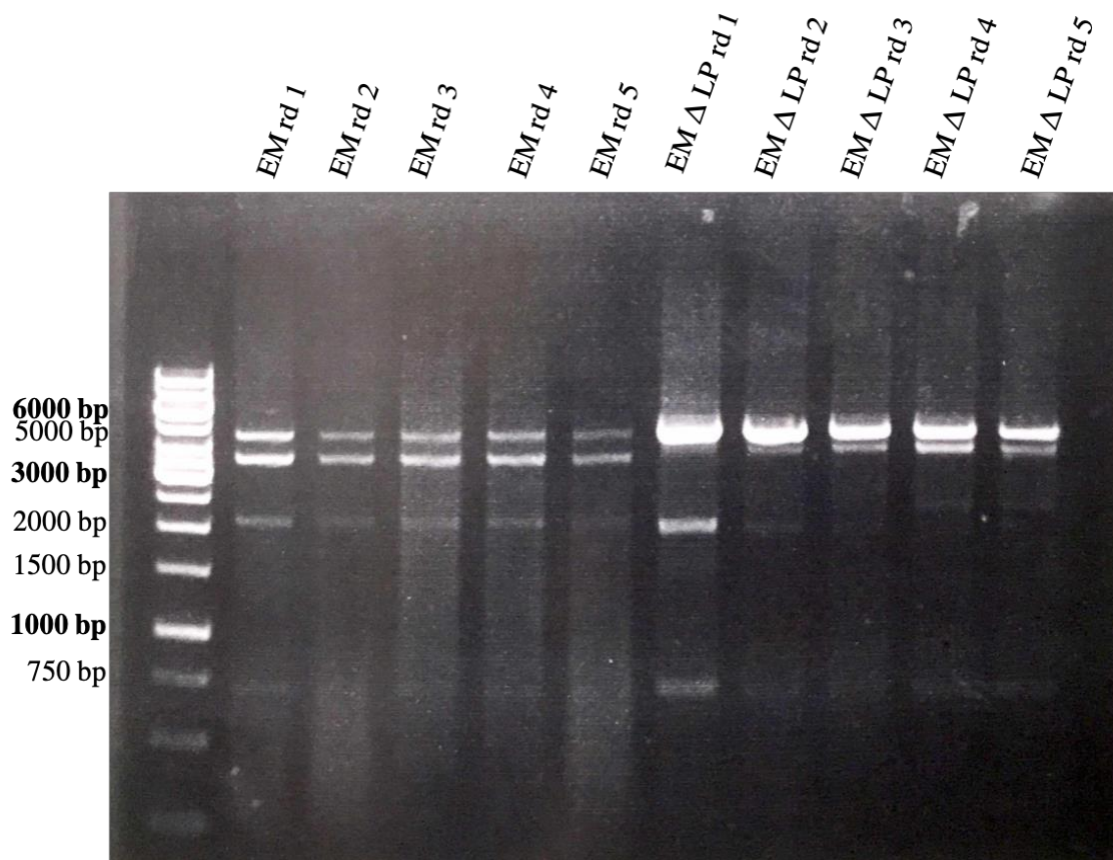


Figure 3-5B. Molecular Evolution of pGKMCS EM and EM Δ LP. Digests were analyzed on 1% agarose gel. All EM vectors produced the expected band patterns (4,466, 1,992, and 705 bp); however, an additional ~3,500 bp band resulted in all digests as well. EM Δ LP rd 1-3 display the expected band patterns (4,308, 1,992, and 705 bp). In both EM Δ LP rd 4-5 the 1,992 bp bands shifted to ~2,500 bp.

As seen in figure 3-5C, all 5 restriction digests on pGKMCS *ermB amyI MxPEP* (EAM) consistently yielded the expected band sizes of 4,466, 2,107, and 705 bp (EAM rd 1-5); however, a faint ~3500 bp band first appeared on the 3rd day of subculturing (EAM rd 3-5). The digest on pGKMCS *ermB amyI MxPEP Δ lac promoter* (EAM Δ LP) culture yielded the expected fragment sizes of 4,308, 2,107, and 705 bp on the 1st – 4th days of growth (EAM Δ LP rd 1-4); however, the vectors obtained on the 2nd - 4th days of growth contained an additional band at ~3,000 bp (EAM Δ LP rd 2-4). A drastic change in the

band patterns occurred between the 4th and 5th days of growth. At the 5th and 6th days of growth the intensity of the ~3,000 bp band became more prominent, another unique band developed ~1,500 bp, the 705 bp band was lost, and inconsistent banding was seen at higher molecular weights (EAM Δ LP rd 5 and 6). Overall, these data indicate that there was selective pressure to alter the vectors without the *lac* promoter, but vectors with the *lac* promoter maintained stability over the course of 5 days.

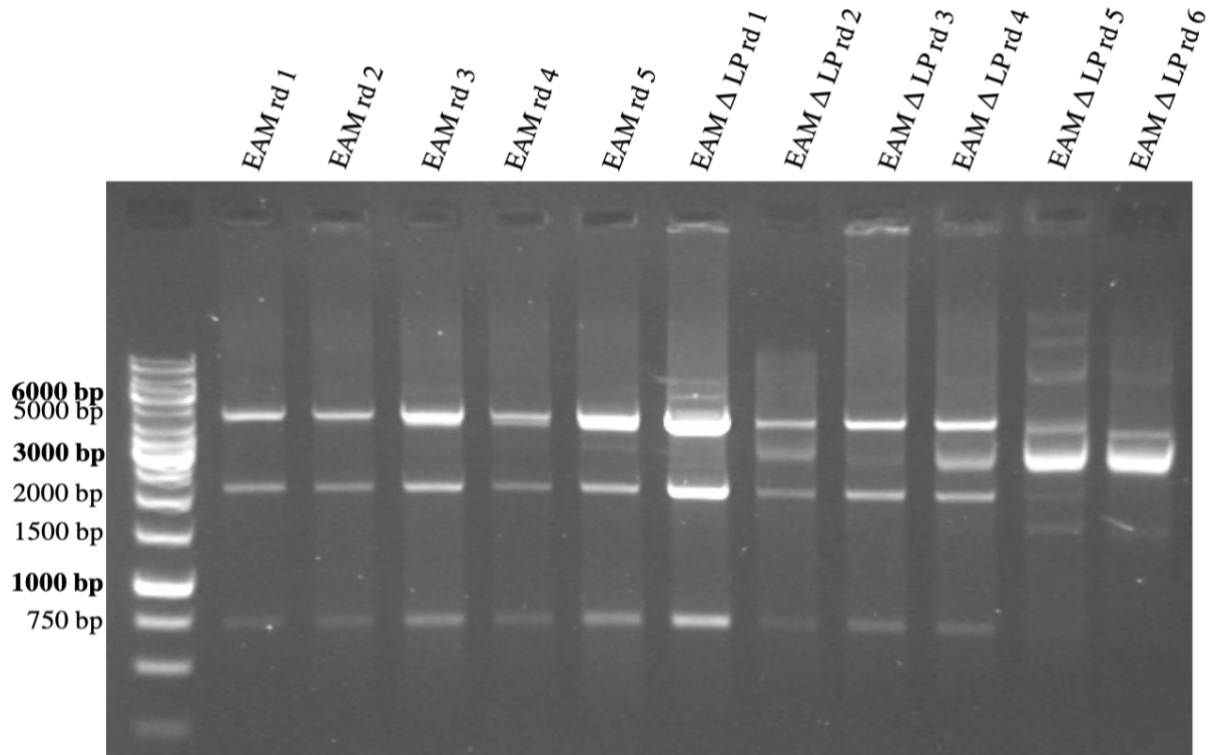


Figure 3-5C. Molecular Evolution of pGKMCS EAM and EAM Δ LP. Digests were analyzed on 1% agarose gel. All digests on EAM rd 1-5 consistently produced the expected band sizes (4,466, 2,107, and 705 bp), but a faint band at ~3,500 bp began to appear in EAM rd 3-5. The expected band patterns were produced in EAM Δ LP rd 1-4 (4,308, 2,107, and 705 bp); however, an unexpected band was observed at ~3,000 bp. The digest patterns for EAM rd 5 -6 did not yield any of the expected band sizes.

Additionally, multiple attempts were conducted to isolate both pGKMCS EM Δ LP and pGKMCS EAM Δ LP from *E. coli* transformants, but there was a noticeable difference in the recovery frequencies between the two. Unaltered pGKMCS EM Δ LP was frequently isolated from the transformants, but only one clone contained pGKMCS EAM Δ LP that generated the expected band patterns (data not shown). Furthermore, mutations accumulated earlier in the cultures containing pGKMCS EAM Δ LP compared to pGKMCS EM Δ LP, and the mutations appeared to be more drastic in pGKMCS EAM Δ LP (Figures 3-5B and 3-5C).

3.3.3 Transformation and protein expression in *L. reuteri* 100-23C

The pGKMCS vectors that yielded the appropriate digestion patterns with *Xho*I and *Nco*I were transformed into *L. reuteri* 100-23C. Successful transformation of the pGKMCS expression vectors in *L. reuteri* was confirmed through PCR. Table 3-3 contains the primer pairs that were used and the expected band sizes for each vector. The pGK-His-Link-F/TS315-R primers were used to detect the absence or presence of the *lac* promoter. As seen in figure 3-6A, PCR on cultures harboring pGKMCS EG, EAG, EM, and EAM resulted in the expected band size of 270 bp indicating the presence of the *lac* promoter. Alternatively, those with pGKMCS EG Δ LP, EM Δ LP, and EAM Δ LP produced an amplicon at 112 bp to confirm the removal of the *lac* promoter. Vectors containing the *ermB* promoter and GFP were assessed with the ErmB-AmyL DIC-F/GFP-AscI-R primers. As expected, the cultures with pGKMCS EG and EG Δ LP produced an amplicon of 1299 bp and EAG 1414 bp, respectively (Figure 3-6B). Additionally, amplification with these primers confirmed that GFP was correctly oriented in relation to the *ermB* promoter. The

vectors containing *ermB* with the MxPEP were not amplified, thereby indicating that these primers were specific to a vector containing GFP. A *L. reuteri* WT control was included in all colony PCRs to show that these primer pairs were specific to the vectors of interest. Lastly, transformation of vectors containing the *ermB* promoter and MxPEP were confirmed with the ErmB-Fix-F/Mx-Opt-AscI-R primers (Figure 3-6C). The PCR on the cultures with pGKMCS EM, EM Δ LP, EAM, and EAM Δ LP resulted in expected band sizes near ~2,700 bp; however no distinction could be made between vectors with (2768 bp) or without (2653 bp) the *amyl* signal sequence. No amplification was observed when these primers were used with cultures containing GFP.

Table 3-3. Expected Band Sizes for Each Vector and Set of Primers. Vectors amplified with pGK-His-Link-F/TS315-R verified the presence (270 bp) or absence (112 bp) of the *lac* promoter. Amplification with GFP-AscI-R or Mx-Opt-AscI-R confirmed the presence

| Vector | pGK-His-link- F TS315-R | ErmB-AmyL-DIC-F GFP-AscI-R | ErmB-Fix-F Mx-Opt-AscI-R |
|---|----------------------------|-------------------------------|-----------------------------|
| pGKMCS <i>ermB GFP</i> (EG) | 270 bp | 1299 bp | - |
| pGKMCS <i>ermB GFP Δ</i> <i>lac promoter</i> (EGΔLP) | 112 bp | 1299 bp | - |
| pGKMCS <i>ermB amyl</i> <i>GFP</i> (EAG) | 270 bp | 1414 bp | - |
| pGKMCS <i>ermB Mx</i> (EM) | 270 bp | - | 2653 bp |
| pGKMCS <i>ermB Mx Δ lac</i> <i>promoter</i> (EMΔLP) | 112 bp | - | 2653 bp |
| pGKMCS <i>ermB Mx amyl</i> (EAM) | 270 bp | - | 2768 bp |
| pGKMCS <i>ermB amyl Mx</i> <i>Δ lac promoter</i> (EAMΔLP) | 112 bp | - | 2768 bp |

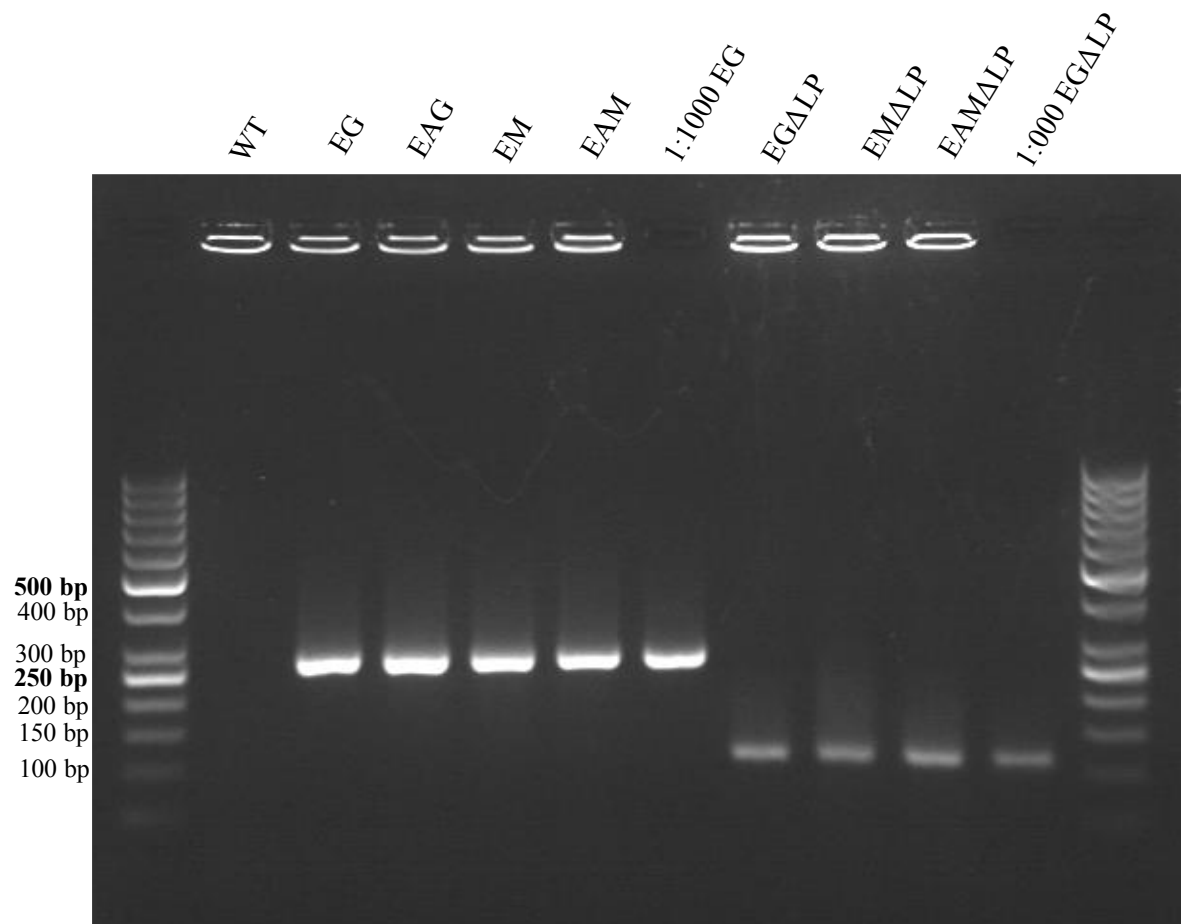


Figure 3-6A. *L. reuteri* Colony PCR to Confirm the Presence or Absence of the *Lac* Promoter in the pGKMCS Expression Vectors. As expected, vectors with the lac promoter yielded bands at 270 bp, and vectors without the lac promoter yielded bands at 112 bp. Dilutions (1:1000) of the EG and EGΔLP vectors were included as positive controls.

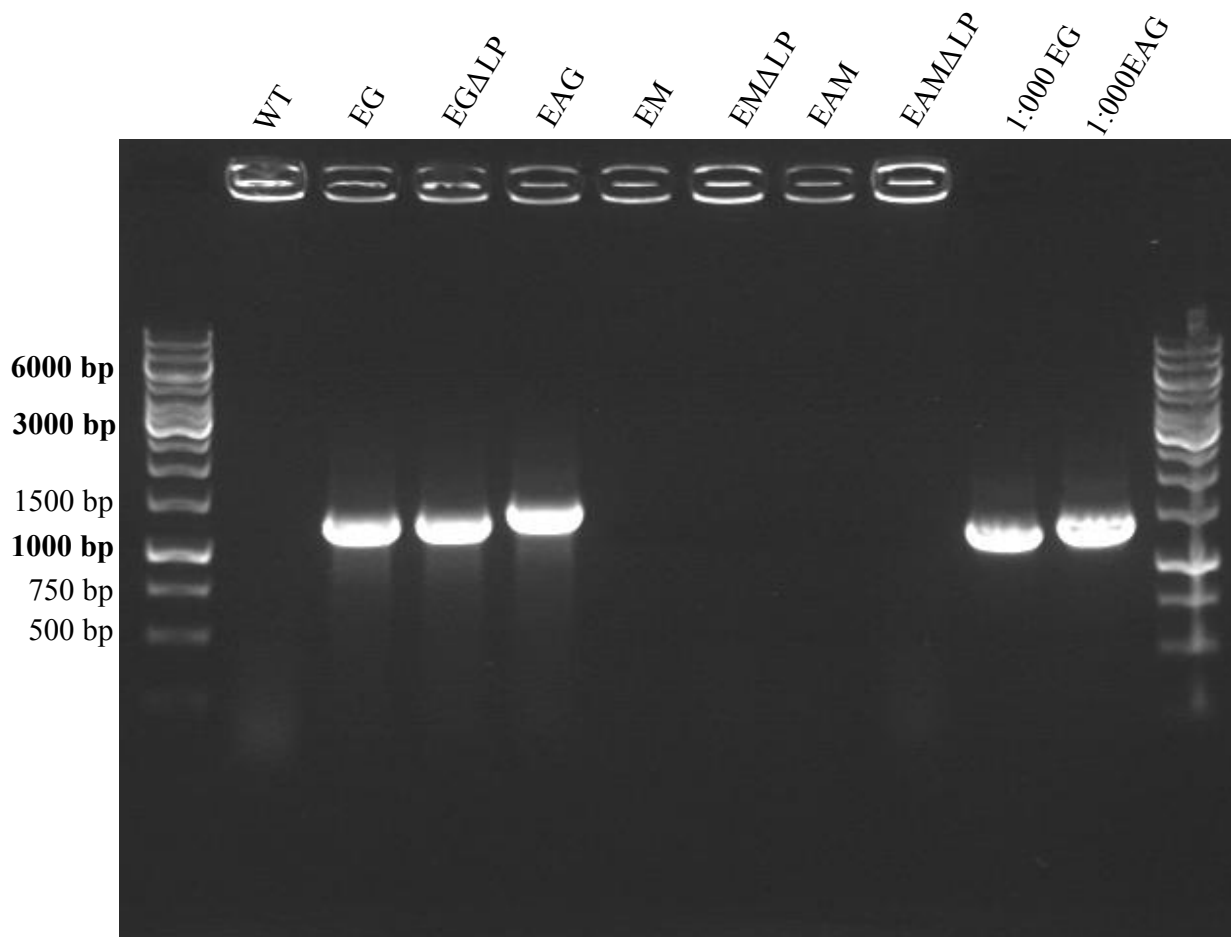


Figure 3-6B. *L. reuteri* Colony PCR to Confirm the Presence of *ErmB*, *Amyl* Signal Sequence, and GFP in pGKMCS Expression Vectors. As expected, vectors containing *ermB* and GFP resulted in a 1,299 bp band, and vectors with the *amyl* signal sequence yielded bands at 1,414 bp. Dilutions (1:1000) of the EG and EGΔLP vectors were included as positive controls.

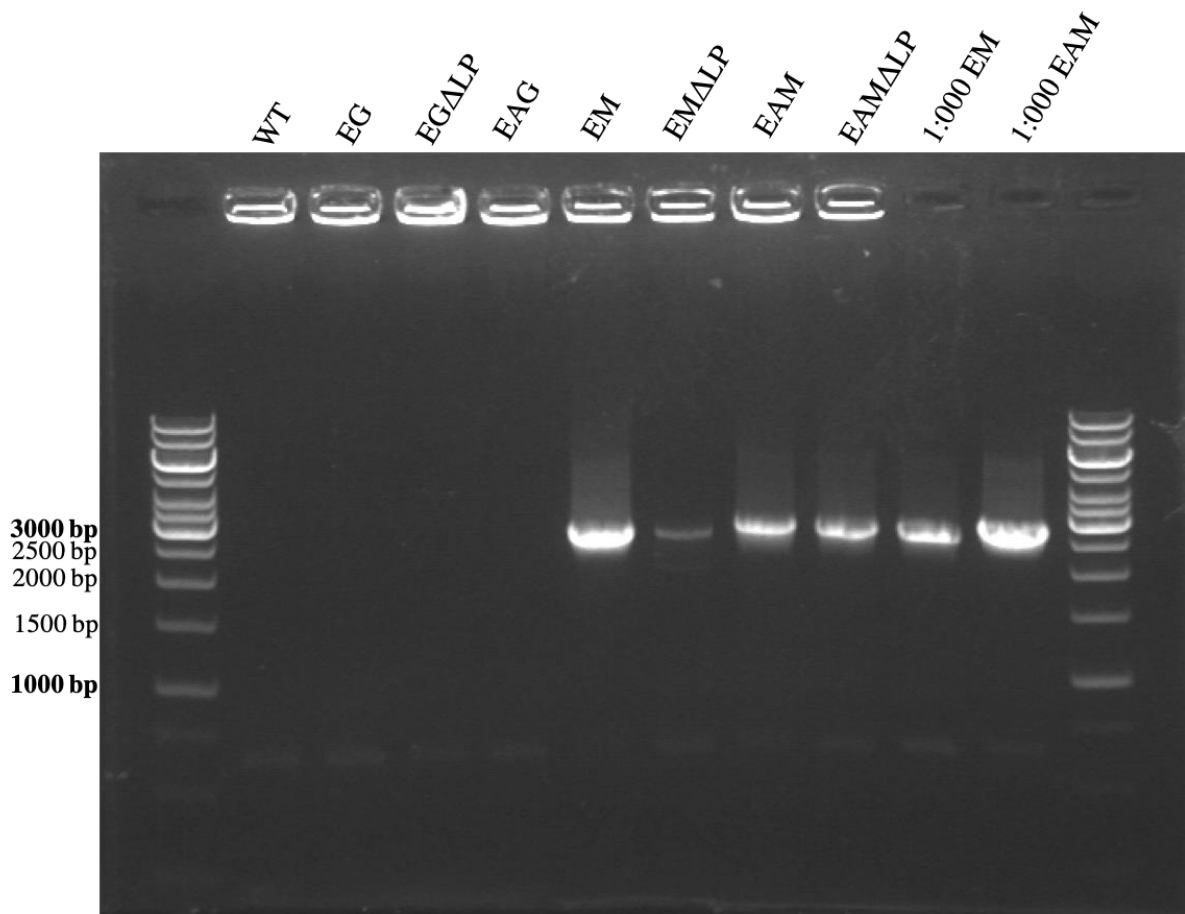


Figure 3-6C. *L. reuteri* Colony PCR to Confirm the Presence of *ErmB* and MxPEP in pGKMCS Expression Vectors. As expected, vectors containing *ermB* and MxPEP resulted in a band ~2,700 bp. The vectors containing the *amyl* signal sequence (2,768 bp) cannot be distinguished from those without (2,653 bp) in this colony PCR. Dilutions (1:1000) of the EM and EAM vectors were included as positive controls.

After transformation into *L. reuteri*, expression of GFP and MxPEP was assessed by SDS-PAGE. Whole cell lysates from *L. reuteri* with pGKMCS EG, EG Δ LP, EM, or EM Δ LP were incubated with Ni-NTA resin. Ni-NTA purified samples from *E. coli* BL21(DE3) pET30 GFP and pET30 MxPEP were included as positive controls. Both positive controls were detected at the appropriate sizes; however, GFP and MxPEP expression from *L. reuteri* was not detected (data not shown). Because cytosolic expression of GFP and MxPEP was not observed, *L. reuteri* cultures with pGKMCS EAM and EAM Δ LP were not analyzed for secreted MxPEP.

3.3.4 Sequencing of pGKMCS EG, pGKMCS EAG, pGKMCS EAM, and pGKMCS EG Δ LP

To determine if the lack of protein expression in *L. reuteri* was due to any mutations within the promoter region or coding region, pGKMCS EG, pGKMCS EAG, pGKMCS EAM, and pGKMCS EG Δ LP were sequenced by GENEWIZ. All vectors sequenced with the MCS-R primer revealed that the *ermB* promoter was identical to the sequence from the Lizier et al. 2014. Moreover, pGKMCS EAG sequenced with GFP-AscI-New-R revealed that GFP did not contain any mutations. Samples sequenced with the ErmB-AmyL-SS-R and the Mx-opt-R primers did not generate viable sequences because these primers did not anneal to the vector of interest. Only 16 of the 45 nt in ErmB-AmyL-SS-R annealed to pGKMCS EG and pGKMCS EG Δ LP. This region of the primer had a $T_m=46.1^\circ\text{C}$, thus the annealing temperature may have been too low allow for sufficient binding to the template. Because pGKMCS EM, EM Δ LP, EAM, and EAM Δ LP were constructed from the same MxPEP template, only pGKMCS EAM was used for sequencing. Three reverse

MxPEP primers were available for pGKMCS EAM (Mx-opt-R, Mx-opt-AI-R, and Mx-opt-AI-R2). Of these three primers, Mx-opt-R does not bind to the 3' end of MxPEP, therefore, the wrong reverse primer was used for sequencing.

3.4 Discussion

In this study an expression cassette containing the *ermB* promoter derived from *E. faecalis* was employed to drive expression of GFP and MxPEP, and the signal sequence from the *B. licheniformis amyl* gene was incorporated to generate secreted versions of these recombinant proteins in *L. reuteri* (Wu and Chung 2006; Lizier et al. 2010). Successful construction and transformation of these vectors into *E. coli* was confirmed through diagnostic digests and the presence of cytosolic GFP and MxPEP. The same vectors were transformed into *L. reuteri*; however, cytosolic GFP and MxPEP were not detected. Because the *L. reuteri* cell lysates did not contain the heterologous proteins of interest, an analysis of secreted GFP and MxPEP was not performed. Although cytosolic GFP and MxPEP expression was confirmed in *E. coli*, stronger GFP fluorescence was observed when the reverse complement of GFP was inserted relative to the *ermB* promoter. This observation along with the absence of heterologous protein expression in *L. reuteri* led to a deeper investigation into the backbone of the pGKMCS vector.

Previous work done in this lab cloned the multiple cloning site of pCR2.1 into a pGK12-derived vector to construct pGKMCS. It was discovered that the pCR 2.1 multiple cloning site contained the *lac* promoter, and it was hypothesized that any expression from the *ermB* promoter was dulled due to the strength of the *lac* promoter. The *lac* promoter was removed

from all pGKMCS *ermB* expression vectors, but the absence this promoter had negative ramifications on the stability of the vectors. Initial positive clones for all pGKMCS *ermB* ΔLP vectors produced expected digest patterns with *XhoI* and *NcoI*; however, the patterns were altered after being subcultured for a total of 5 days. These results suggest that there may have been toxicity associated with the vectors without the *lac* promoter. *E. coli* MC1061 is *RecA*⁺, responsible for DNA repair and maintenance, which supports the notion that selective pressure provoked the accumulation of random mutations within these pGKMCS expression vectors. Cells containing the unaltered and potentially lethal variant of the vector do not survive. Contrastingly, cells with the mutated and benign vector may maintain erythromycin resistance, thereby becoming the dominant culture in the media.

In addition to the overall toxicity associated with the removal of the *lac* promoter, pGKMCS *ermB amyI MxPEP* ΔLP appeared to be more unstable than pGKMCS *ermB MxPEP* ΔLP . The frequency of transformants that contained unaltered pGKMCS *ermB amyI MxPEP* ΔLP was much lower than that of pGKMCS *ermB MxPEP* ΔLP . Furthermore, mutations developed earlier over the course of 5 days within pGKMCS *ermB amyI MxPEP* ΔLP than pGKMCS *ermB MxPEP* ΔLP . These mutations resulted in a more drastic deviation from the original digestion pattern in pGKMCS *ermB amyI MxPEP* ΔLP than pGKMCS *ermB MxPEP* ΔLP . Altogether, these data further suggest that the presence of both the *amyI* signal sequence and MxPEP may have inflicted a higher degree of toxicity upon *E. coli*. However, it cannot be confirmed whether or not the *amyI* signal sequence made a significant contribution to the instability of the vectors. To assess the impact of the *amyI* signal sequence on the stability of the expression vectors, pGKMCS *ermB amyI GFP*

ΔLP must be constructed and the diagnostic digest experiment should be repeated. If the resulting data reveal more drastic mutations within pGKMCS *ermB amyl GFP ΔLP* than pGKMCS *ermB GFP ΔLP* , then it may indicate the *amyl* signal sequence amplifies the toxicity associated with these expression vectors.

These data demonstrate the need for a tightly regulated promoter that would help alleviate the toxicity inflicted upon *E. coli* harboring these vectors. Thus, it is necessary to develop a controlled expression system to allow for stabilization of the expression vectors in both *E. coli* and *L. reuteri* and detection of MxPEP in *L. reuteri*. An inducible expression system of interest for *L. reuteri* is the nisin-controlled-gene expression (NICE) system derived from *L. lactis*. The NICE system, a two-component signal transduction system that relies on a histidine-protein kinase (*nisK*) and response regulator (*nisR*) to induce expression from the *nisA* promoter, is one of the most widely adapted and characterized for regulated gene expression in Gram positive bacteria. Expression of recombinant proteins from NICE is induced by subinhibitory concentrations of nisin, a broad-spectrum antimicrobial peptide that is widely used as a preservative in the food industry. This system was adapted in *L. reuteri* by integrating the *nisA* promoter, *nisK*, and *nisR* into the *E. coli*-*L. reuteri* shuttle vector pSTE32 to generate pNICE (Wu et al. 2006). It was reported that *L. reuteri* harboring pNICE successfully produced the heterologous α -amylase when induced with nisin, and induction resulted in a 6.9-fold increase of α -amylase compared to non-induced cultures. Since the NICE system is dependent on an antimicrobial compound for gene expression, pursuit of an alternative system may be of interest for inducible gene expression within the GI tract. Recent investigations have looked into the use of a bile-

induced promoter for the expression of recombinant proteins in LAB. The bile-responsive *lactate dehydrogenase (ldh)* promoter from *Lactobacillus johnsonii* resulted in a 1.8-fold increase of recombinant β -galactosidase in *Lactobacillus plantarum* (Chae et al. 2019). In another study, the *L. casei* promoter P16090 was successfully induced by 0.05% bile salts and 0.025% cholic acid in *L. reuteri* to express an anaerobic fluorescent reporter gene (Martínez-Fernández et al. 2019). As an inhabitant of the duodenum of the small intestine, a bile-induced system may be of interest for recombinant gene expression in *L. reuteri*.

Recombinant protein expression was not observed from *L. reuteri*, therefore, vectors that were transformed into *L. reuteri* were sequenced to determine if the *ermB* promoter, GFP, and MxPEP sequences were conserved. Sequence data revealed that the *ermB* promoter and GFP sequences did not contain any mutations. These data suggest that *L. reuteri* was transformed with the appropriate sequences for GFP expression. It could not be determined if the MxPEP sequence was preserved because the wrong MxPEP primer was used for sequencing. Because the resulting *ermB* promoter sequence was identical to that of Lizier et al. 2014, expression of heterologous proteins from this promoter was possible. However, it could not be established if these vectors remained unchanged in *L. reuteri* because the resources required to isolate plasmids from Gram positive bacteria were not available. Moreover, *L. reuteri* 100-23 is also a RecA⁺ strain, therefore, it is possible that clones with the unaltered vector were selected against, and clones confirmed by colony PCR harbored the mutated version of the vector.

The absence of recombinant protein expression in *L. reuteri* may be further explained by the length of the spacer, the distance between the Shine-Dalgarno (S-D) sequence and start codon. It has been established that translation initiation is the most critical step in protein production (Chen et al. 1994; Berwal, Sreejith, and Pal 2010). When translation is initiated two parts of the ribosome, the 16s rRNA and P-site, interact with two sites on the mRNA, S-D sequence and start codon. Because these mRNA sites are related to the distance between the 3' end of the 16s rRNA (complementary to the S-D sequence) and the P-site of the ribosome (holds the fMet-tRNA), the spacer has an impact on the efficiency of translation initiation. For example, protein expression in *Bifidobacterium longum* was evaluated with varying spacer lengths (4 nt to 9 nt), and it was reported that the loss of one nucleotide in the spacer (from 5 nt to 4 nt) resulted in 85% loss of protein expression (He, Sakaguchi, and Suzuki 2012). Additionally, previous studies have reported an average spacer length of 7-8 nt for endogenous genes, and a reduction in recombinant protein expression when the spacer ranged from 10-12 nt (Chen et al. 1994; Cao et al. 2015). The length of the spacer for all pGKMCS *ermB* expression vectors was 11 nt, therefore, it is plausible that this spacer was sufficient enough for expression in *E. coli* but not in *L. reuteri*. These 11 nt may have caused the distance to be too great for the 16s rRNA and fMet-tRNA to efficiently interact with the mRNA. Ultimately, expression may from the *ermB* promoter in both *E. coli* and *L. reuteri* may be enhanced if the length of the spacer was reduced to 7-8 nt.

Furthermore, finding the right SP may prove to be a substantial challenge in the overall scope of this study. A previous study reported that it was essential for the signal sequence

to contain non-optimized codons to slow the rate of translation to allow the protein to fold correctly (Zalucki and Jennings 2007). Wu and Chung 2006 and Malik et al. 2013 both amplified the *amyl* signal sequence directly from the *B. licheniformis* ATCC 27811 genome, but in this present study the sequence was optimized for expression *L. reuteri*. In the sequence sequenced used in this study, 37 of the 38 codons generated a relative adaptiveness value of 100% and the remaining codon had a value of 96%. Although it could not be determined in this study, it is possible that this optimized *amyl* signal sequence would not have successfully secreted recombinant proteins. If the α -amylase SP is of interest for the future of this study, then the original sequence from *B. licheniformis* ATCC 27811 should be obtained. Another obstacle to SP selection may be a variation in translocation efficacy for different recombinant proteins. One study determined that the junction between the C-terminus of the SP and N-terminus of the heterologous protein influences translocation efficiency of the SP (Brockmeier et al. 2006). Another study further reported that SPs that worked well in *Lactobacillus plantarum* to secrete a staphylococcus nuclease did not have the same effect on a *Lactobacillus amylovorus* α -amylase (Mathiesen et al. 2009). Thus, a SP that secretes GFP with great efficacy may not have the same effect on MxPEP. Ultimately, a SP-GFP fusion may not be of great interest for this study, and future researchers may only want to consider evaluating the efficiency of SPs with a N-terminal portion of MxPEP. Because a wide range of predicted SPs should be screened and evaluated with MxPEP, it is important that a robust *L. reuteri* expression cassette is developed to allow signal sequences to be easily swapped in and out of the vector. Additionally, an assay that is a cheap and time-saving alternative to purification of secreted MxPEP would be of great value to this study. For example, SPs that successfully

secrete α -amylase are identified by observing a zone of clearing on starch agar with the addition of Gram's iodine. Transformants can be eliminated with ease if a zone of clearing is not observed. Once promising candidates are selected, a more thorough assessment of the SP secretion efficiency can be performed through purification of the recombinant protein from the culture medium.

BIBLIOGRAPHY

- Abadie, Valérie, and Bana Jabri. 2014. "IL-15: A Central Regulator of Celiac Disease Immunopathology." *Immunological Reviews* 260 (1): 221–34.
- Al-Bawardy, Badr, D. Chamil Codipilly, Alberto Rubio-Tapia, David H. Bruining, Stephanie L. Hansel, and Joseph A. Murray. 2017. "Celiac Disease: A Clinical Review." *Abdominal Radiology* 42 (2): 351–60.
- Arentz-Hansen, Helene, Roman Körner, Øyvind Molberg, Hanne Quarsten, Willemijn Vader, Yvonne M. C. Kooy, Knut E. A. Lundin, et al. 2000. "The Intestinal T Cell Response to α -Gliadin in Adult Celiac Disease Is Focused on a Single Deamidated Glutamine Targeted by Tissue Transglutaminase." *Journal of Experimental Medicine* 191 (4): 603–12.
- Asmar, Rahzi El, Pinaki Panigrahi, Penelope Bamford, Irene Berti, Tarcisio Not, Giovanni V. Coppa, Carlo Catassi, and Alessio Fasano. 2002. "Host-Dependent Zonulin Secretion Causes the Impairment of the Small Intestine Barrier Function after Bacterial Exposure." *Gastroenterology* 123 (5): 1607–15.
- Baranwal, Ajay Kumar, and Narinder K. Mehra. 2017. "Major Histocompatibility Complex Class I Chain-Related A (MICA) Molecules: Relevance in Solid Organ Transplantation." *Frontiers in Immunology* 8 (February).
- Bauer, Groh, Wu, Steinle, Phillips, Lanier, Spies. 1999. "Activation of NK Cells and T Cells by NKG2D, a Receptor for Stress-Inducible MICA" *Science* 285 (5428): 727–729.
- Berwal, Sunil K., R.K. Sreejith, and Jayanta K. Pal. 2010. "Distance between RBS and AUG Plays an Important Role in Overexpression of Recombinant Proteins." *Analytical Biochemistry* 405 (2): 275–77.
- Bodis, Gergely, Victoria Toth, and Andreas Schwarting. 2018. "Role of Human Leukocyte Antigens (HLA) in Autoimmune Diseases." *Rheumatology and Therapy* 5 (1): 5–20.
- Brockmeier, Ulf, Michael Caspers, Roland Freudl, Alexander Jockwer, Thomas Noll, and Thorsten Eggert. 2006. "Systematic Screening of All Signal Peptides from *Bacillus Subtilis*: A Powerful Strategy in Optimizing Heterologous Protein Secretion in Gram-Positive Bacteria." *Journal of Molecular Biology* 362 (3): 393–402.
- Cao, Jicong, Manish Arha, Chaitanya Sudrik, Abhirup Mukherjee, Xia Wu, and Ravi S. Kane. 2015. "A Universal Strategy for Regulating MRNA Translation in Prokaryotic and Eukaryotic Cells." *Nucleic Acids Research* 43 (8): 4353–62.

- Chae, Jong Pyo, Edward Alain Pajarillo, In-Chan Hwang, and Dae-Kyung Kang. 2019. "Construction of a Bile-Responsive Expression System in *Lactobacillus Plantarum*." *Food Science of Animal Resources* 39 (1): 13–22.
- Chen, Hongyun, Matthew Bjerknes, Ravindra Kumar, and Ernest Jay. 1994. "Determination of the Optimal Aligned Spacing between the Shine – Dalgarno Sequence and the Translation Initiation Codon of *Escherichia Coli* m RNAs." *Nucleic Acids Research* 22 (23): 4953–57.
- Cortazzo, Patricia, Carlos Cerveñansky, Mónica Marín, Claude Reiss, Ricardo Ehrlich, and Atilio Deana. 2002. "Silent Mutations Affect in Vivo Protein Folding in *Escherichia Coli*." *Biochemical and Biophysical Research Communications* 293 (1): 537–41.
- Di Sabatino, Antonio, Alessandro Vanoli, Paolo Giuffrida, Ombretta Luinetti, Enrico Solcia, and Gino Roberto Corazza. 2012. "The Function of Tissue Transglutaminase in Celiac Disease." *Autoimmunity Reviews* 11 (10): 746–53.
- Elsing, Christoph, Jörg Placke, and Wilhelm Gross-Weege. 2005. "Ulcerative Jejunoileitis and Enteropathy-Associated T-Cell Lymphoma." *European Journal of Gastroenterology & Hepatology* 17 (12): 1401–1405.
- Fasano, Alessio. 2011. "Zonulin and Its Regulation of Intestinal Barrier Function: The Biological Door to Inflammation, Autoimmunity, and Cancer." *Physiological Reviews* 91 (1): 151–75.
- Ferranti, Pasquale, Gianfranco Mamone, Gianluca Picariello, and Francesco Addeo. 2007. "Mass Spectrometry Analysis of Gliadins in Celiac Disease." *Journal of Mass Spectrometry* 42 (12): 1531-1548.
- Greco, L. 2002. "The First Large Population Based Twin Study of Coeliac Disease." *Gut* 50 (5): 624–28.
- Groh, Veronika, Rebecca Rhinehart, Julie Randolph-Habecker, Max S. Topp, Stanley R. Riddell, and Thomas Spies. 2001. "Costimulation of CD8 $\alpha\beta$ T Cells by NKG2D via Engagement by MIC Induced on Virus-Infected Cells." *Nature Immunology* 2 (3): 255–60.
- He, Jianlong, Kouta Sakaguchi, and Tohru Suzuki. 2012. "Determination of the Ribosome-Binding Sequence and Spacer Length between Binding Site and Initiation Codon for Efficient Protein Expression in *Bifidobacterium Longum* 105-A." *Journal of Bioscience and Bioengineering* 113 (4): 442–44.
- Hedstrom, Lizbeth. 2002. "Serine Protease Mechanism and Specificity." *Chemical Reviews* 102 (12): 4501–24.

- Hongying, Fan, Wu Xianbo, Yu Fang, Bai Yang, and Long Beiguo. 2014. "Oral Immunization with Recombinant *Lactobacillus Acidophilus* Expressing the Adhesin Hp0410 of *Helicobacter Pylori* Induces Mucosal and Systemic Immune Responses." Edited by D. L. Burns. *Clinical and Vaccine Immunology* 21 (2): 126–32.
- Hou, Chengli, Xiangfang Zeng, Fengjuan Yang, Hong Liu, and Shiyan Qiao. 2015. "Study and Use of the Probiotic *Lactobacillus Reuteri* in Pigs: A Review." *Journal of Animal Science and Biotechnology* 6 (1).
- Hüe, Sophie, Jean-Jacques Mention, Renato C. Monteiro, ShaoLing Zhang, Christophe Cellier, Jacques Schmitz, Virginie Verkarre, et al. 2004. "A Direct Role for NKG2D/MICA Interaction in Villous Atrophy during Celiac Disease." *Immunity* 21 (3): 367–377.
- Huibregtse, Inge L., Veerle Snoeck, An de Creus, Henri Braat, Ester C. de Jong, Sander J.H. van Deventer, and Pieter Rottiers. 2007. "Induction of Ovalbumin-Specific Tolerance by Oral Administration of *Lactococcus Lactis* Secreting Ovalbumin." *Gastroenterology* 133 (2): 517–28.
- Jabri, Bana, Jeanette M Selby, Horia Negulescu, Leanne Lee, Arthur I Roberts, Andrew Beavis, Miguel Lopez-Botet, Ellen C Ebert, and Robert J Winchester. 2002. "TCR Specificity Dictates CD94/NKG2A Expression by Human CTL." *Immunity* 17 (4): 487–99.
- Kang, Chao, Xiao-Wei Yu, and Yan Xu. 2013. "Gene Cloning and Enzymatic Characterization of an Endoprotease Endo-Pro-*Aspergillus Niger*." *Journal of Industrial Microbiology & Biotechnology* 40 (8): 855–64.
- Kim, Sungmin F., SEUNG J. Baek, and M. Y. Pack. 1991. "Cloning and Nucleotide Sequence of the *Lactobacillus Casei* Lactate Dehydrogenase Gene." *Applied and Environmental Microbiology* 57 (8): 2413–2417.
- Klerk, Nele de, Lisa Maudsdotter, Hanna Gebreegziabher, Sunil D. Saroj, Beatrice Eriksson, Olaspers Sara Eriksson, Stefan Roos, Sara Lindén, Hong Sjölander, and Ann-Beth Jonsson. 2016. "Lactobacilli Reduce *Helicobacter Pylori* Attachment to Host Gastric Epithelial Cells by Inhibiting Adhesion Gene Expression." Edited by S. R. Blanke. *Infection and Immunity* 84 (5): 1526–35.
- Kocadag Kocazorbaz, Ebru, and Figen Zihnioglu. 2017. "Purification, Characterization and the Use of Recombinant Prolyl Oligopeptidase from *Myxococcus Xanthus* for Gluten Hydrolysis." *Protein Expression and Purification* 129 (January): 101–7.
- Kubota, Keiko, Masaru Tanokura, and Kenji Takahashi. 2005. "Purification and Characterization of a Novel Prolyl Endopeptidase from *Aspergillus Niger*." *Proceedings of the Japan Academy, Series B* 81 (10): 447–53.

- Lai, Thung-S., Cheng-Jui Lin, and Charles S. Greenberg. 2017. "Role of Tissue Transglutaminase-2 (TG2)-Mediated Aminylation in Biological Processes." *Amino Acids* 49 (3): 501–15.
- Lammers, Karen M., Ruliang Lu, Julie Brownley, Bao Lu, Craig Gerard, Karen Thomas, Prasad Rallabhandi, et al. 2008. "Gliadin Induces an Increase in Intestinal Permeability and Zonulin Release by Binding to the Chemokine Receptor CXCR3." *Gastroenterology* 135 (1): 194-204.e3.
- Lebwohl, Benjamin, Alberto Rubio-Tapia, Asaad Assiri, Catherine Newland, and Stefano Guandalini. 2012. "Diagnosis of Celiac Disease." *Gastrointestinal Endoscopy Clinics of North America* 22 (4): 661–77.
- Leffler, Daniel, Detlef Schuppan, Kumar Pallav, Robert Najarian, Jeffery D Goldsmith, Joshua Hansen, Toufic Kabbani, Melinda Dennis, and Ciarán P Kelly. 2013. "Kinetics of the Histological, Serological and Symptomatic Responses to Gluten Challenge in Adults with Coeliac Disease." *Gut* 62 (7): 996–1004.
- Li, Jingjie, Wen Zhang, Chuan Wang, Qian Yu, Ruirui Dai, and Xiaofang Pei. 2012. "Lactococcus Lactis Expressing Food-Grade β -Galactosidase Alleviates Lactose Intolerance Symptoms in Post-Weaning Balb/c Mice." *Applied Microbiology and Biotechnology* 96 (6): 1499–1506.
- Liu, R. B., B. Engels, K. Schreiber, C. Ciszewski, A. Schietinger, H. Schreiber, and B. Jabri. 2013. "IL-15 in Tumor Microenvironment Causes Rejection of Large Established Tumors by T Cells in a Noncognate T Cell Receptor-Dependent Manner." *Proceedings of the National Academy of Sciences* 110 (20): 8158–63.
- Lizier, Michela, Pier G. Sarra, Roberto Cauda, and Franco Lucchini. 2010. "Comparison of Expression Vectors in Lactobacillus Reuteri Strains: Constitutive Promoters in L. Reuteri." *FEMS Microbiology Letters* 308 (1): 8–15.
- Ljungh, Asa, and Torkel Wadström. 2006. "Lactic Acid Bacteria as Probiotics." *Current Issues in Intestinal Microbiology* 7 (2): 19.
- Malik, B., N. Rashid, N. Ahmad, and M. Akhtar. 2013. "Escherichia Coli Signal Peptidase Recognizes and Cleaves the Signal Sequence of α -Amylase Originating from Bacillus Licheniformis." *Biochemistry (Moscow)* 78 (8): 958–62.
- Marti, Thomas, Oyvind Molberg, Qing Li, Gary M. Gray, Chaitan Khosla, and Ludvig M. Sollid. 2004. "Prolyl Endopeptidase-Mediated Destruction of T Cell Epitopes in Whole Gluten: Chemical and Immunological Characterization." *Journal of Pharmacology and Experimental Therapeutics* 312 (1): 19–26.

- Martínez-Fernández, José Alberto, Daniel Bravo, Ángela Peirotén, Juan Luis Arqués, and José María Landete. 2019. "Bile-Induced Promoters for Gene Expression in Lactobacillus Strains." *Applied Microbiology and Biotechnology* 103 (9): 3819–27.
- Mathiesen, Geir, Anita Sveen, May Brurberg, Lasse Fredriksen, Lars Axelsson, and Vincent GH Eijsink. 2009. "Genome-Wide Analysis of Signal Peptide Functionality in Lactobacillus Plantarum WCFS1." *BMC Genomics* 10 (1): 425.
- Meijer, J. W. R., C. J. J. Mulder, M. G. Goerres, H. Boot, and J. J. Schweizer. 2004. "Coeliac Disease and (Extra)Intestinal T-Cell Lymphomas: Definition, Diagnosis and Treatment." *Scandinavian Journal of Gastroenterology* 39 (241): 78–84.
- Mohan Kumar, B.V., M. Vijaykrishnaraj, Nawneet K. Kurrey, Vijay S. Shinde, and P. Prabhasankar. 2019. "Prolyl Endopeptidase-Degraded Low Immunoreactive Wheat Flour Attenuates Immune Responses in Caco-2 Intestinal Cells and Gluten-Sensitized BALB/c Mice." *Food and Chemical Toxicology* 129 (July): 466–75.
- Montserrat, Veronica, Maaïke J. Bruins, Lippo Edens, and Frits Koning. 2015. "Influence of Dietary Components on Aspergillus Niger Prolyl Endoprotease Mediated Gluten Degradation." *Food Chemistry* 174 (May): 440–45.
- Morita, T., M. Ueda, K. Kubo, and H. Aiba. 2015. "Insights into Transcription Termination of Hfq-Binding SRNAs of Escherichia Coli and Characterization of Readthrough Products." *RNA* 21 (8): 1490–1501.
- Mu, Qinghui, Husen Zhang, Xiaofeng Liao, Kaisen Lin, Hualan Liu, Michael R. Edwards, S. Ansar Ahmed, et al. 2017. "Control of Lupus Nephritis by Changes of Gut Microbiota." *Microbiome* 5 (1): 73.
- Nistico, L. 2006. "Concordance, Disease Progression, and Heritability of Coeliac Disease in Italian Twins." *Gut* 55 (6): 803–8.
- Ozuna, Carmen V., Julio C. M. Iehisa, María J. Giménez, Juan B. Alvarez, Carolina Sousa, and Francisco Barro. 2015. "Diversification of the Celiac Disease α -Gliadin Complex in Wheat: A 33-Mer Peptide with Six Overlapping Epitopes, Evolved Following Polyploidization." *The Plant Journal* 82 (5): 794–805.
- Park, K. S., J. H. Park, and Y. W. Song. 2008. "Inhibitory NKG2A and Activating NKG2D and NKG2C Natural Killer Cell Receptor Genes: Susceptibility for Rheumatoid Arthritis." *Tissue Antigens* 72 (4): 342–46.
- Reuter, Gerhard. 2001. "The Lactobacillus and Bifidobacterium Microflora Of-the Human Intestine:," *Current Issues in Intestinal Microbiology* 2 (2): 43-53.
- Roberts, A. I., L. Lee, E. Schwarz, V. Groh, T. Spies, E. C. Ebert, and B. Jabri. 2001. "Cutting Edge: NKG2D Receptors Induced by IL-15 Costimulate CD28-Negative

- Effector CTL in the Tissue Microenvironment.” *The Journal of Immunology* 167 (10): 5527–30.
- Rosano, Germán L, and Eduardo A Ceccarelli. 2009. “Rare Codon Content Affects the Solubility of Recombinant Proteins in a Codon Bias-Adjusted Escherichia Coli Strain.” *Microbial Cell Factories* 8 (1): 41.
- Salden, B. N., V. Monserrat, F. J. Troost, M. J. Bruins, L. Edens, R. Bartholomé, G. R. Haenen, B. Winkens, F. Koning, and A. A. Masclee. 2015. “Randomised Clinical Study: *Aspergillus Niger* -Derived Enzyme Digests Gluten in the Stomach of Healthy Volunteers.” *Alimentary Pharmacology & Therapeutics* 42 (3): 273–85.
- Sayers, Jon R., Debbie Evans, and James B. Thomson. 1996. “Identification and Eradication of a Denatured DNA Isolated during Alkaline Lysis-Based Plasmid Purification Procedures.” *Analytical Biochemistry* 241 (2): 186–89.
- Shan, Lu, Thomas Marti, Ludvig M. Sollid, Gary M. Gray, and Chaitan Khosla. 2004a. “Comparative Biochemical Analysis of Three Bacterial Prolyl Endopeptidases: Implications for Coeliac Sprue.” *The Biochemical Journal* 383 (Pt 2): 311–18.
- Shan, Lu, Irimpan I Mathews, and Chaitan Khosla. 2005. “Structural and Mechanistic Analysis of Two Prolyl Endopeptidases: Role of Interdomain Dynamics in Catalysis and Specificity,” January, 6.
- Shan, Lu, Øyvind Molberg, Isabelle Parrot, Felix Hausch, Ferda Filiz, Gary M. Gray, Ludvig M. Sollid, and Chaitan Khosla. 2002. “Structural Basis for Gluten Intolerance in Celiac Sprue.” *Science* 297 (5590): 2275–79.
- Singh, Prashant, Shubhangi Arora, Suman Lal, Tor A Strand, and Govind K Makharia. 2015. “Risk of Celiac Disease in the First- and Second-Degree Relatives of Patients With Celiac Disease: A Systematic Review and Meta-Analysis.” *The American Journal of Gastroenterology* 110 (11): 1539–48.
- Sollid, LM., Thorsby E. 1993. “HLA susceptibility genes in coeliac disease: genetic mapping and role in pathogenesis.” *Gastroenterology* 105: 910-22.
- Spinler, Jennifer K., Malai Taweechotipatr, Cheryl L. Rognerud, Ching N. Ou, Somying Tumwasorn, and James Versalovic. 2008. “Human-Derived Probiotic Lactobacillus Reuteri Demonstrate Antimicrobial Activities Targeting Diverse Enteric Bacterial Pathogens.” *Anaerobe* 14 (3): 166–71.
- Stepniak, D. 2006. “Highly Efficient Gluten Degradation with a Newly Identified Prolyl Endoprotease: Implications for Celiac Disease.” *AJP: Gastrointestinal and Liver Physiology* 291 (4): G621–29.

- Thomas, K. E., A. Sapone, A. Fasano, and S. N. Vogel. 2006. "Gliadin Stimulation of Murine Macrophage Inflammatory Gene Expression and Intestinal Permeability Are MyD88-Dependent: Role of the Innate Immune Response in Celiac Disease." *The Journal of Immunology* 176 (4): 2512–21.
- Tsiatsiani, Liana, Michiel Akeroyd, Maurien Olsthoorn, and Albert J. R. Heck. 2017. "*Aspergillus Niger* Prolyl Endoprotease for Hydrogen–Deuterium Exchange Mass Spectrometry and Protein Structural Studies." *Analytical Chemistry* 89 (15): 7966–73.
- Yu, Q., C. Tang, S. Xun, T. Yajima, K. Takeda, and Y. Yoshikai. 2006. "MyD88-Dependent Signaling for IL-15 Production Plays an Important Role in Maintenance of CD8 TCR and TCR Intestinal Intraepithelial Lymphocytes." *The Journal of Immunology* 176 (10): 6180–85.
- Van Raemdonck, Katrien, Philippe E. Van den Steen, Sandra Liekens, Jo Van Damme, and Sofie Struyf. 2015. "CXCR3 Ligands in Disease and Therapy." *Cytokine & Growth Factor Reviews* 26 (3): 311–27.
- Walter, J., R. A. Britton, and S. Roos. 2011. "Host-Microbial Symbiosis in the Vertebrate Gastrointestinal Tract and the *Lactobacillus Reuteri* Paradigm." *Proceedings of the National Academy of Sciences* 108 (Supplement_1): 4645–52.
- Wang, W., Uzzau, S., Goldblum, SE., Fasano, A. 2000. "Zonulin as Modulator of Intestinal Tight Junctions." *Journal of Cell Science* 113: 4435-4440.
- Wilkins, T. and Sequoia, J. 2017. "Probiotics for Gastrointestinal Conditions: A Summary of the Evidence." *American Family Physician* 96 (3): 170-179.
- Wilson, Charlotte M., Diane Loach, Blair Lawley, Tracey Bell, Ian M. Sims, Paul W. O'Toole, Aldert Zomer, and Gerald W. Tannock. 2014. "Lactobacillus Reuteri 100-23 Modulates Urea Hydrolysis in the Murine Stomach." Edited by G. T. Macfarlane. *Applied and Environmental Microbiology* 80 (19): 6104–13.
- Wu, Chi-Ming, and Tung-Ching Chung. 2006. "Green Fluorescent Protein Is a Reliable Reporter for Screening Signal Peptides Functional in *Lactobacillus Reuteri*." *Journal of Microbiological Methods* 67 (1): 181–86.
- Wu, Chi-Ming, Chuen-Fu Lin, Yi-Chih Chang, and Tung-Ching Chung. 2006. "Construction and Characterization of Nisin-Controlled Expression Vectors for Use in *Lactobacillus Reuteri*." *Bioscience, Biotechnology, and Biochemistry* 70 (4): 757–67.
- Xia, Jiang, Ludvig M. Sollid, and Chaitan Khosla. 2005. "Equilibrium and Kinetic Analysis of the Unusual Binding Behavior of a Highly Immunogenic Gluten Peptide to HLA-DQ2 α ." *Biochemistry* 44 (11): 4442–49

Zalucki, Yaramah M., and Michael P. Jennings. 2007. "Experimental Confirmation of a Key Role for Non-Optimal Codons in Protein Export." *Biochemical and Biophysical Research Communications* 355 (1): 143–48.

APPENDICES

A. Optimized *Aspergillus niger* Patent PAP (1581 bp)

ATGCGTGCTTTTTCAGCTGTTGCTGCTGCTGCTTTAGCTTTATCATGGGCTTCA
TTAGCTCAAGCTGCTCGTCCACGTTTAGTTCCAAAACCAGTTTCACGTCCAGC
TTCATCAAAATCAGCTGCTACTACTGGTGAAGCTTATTTTGAACAATTATTAG
ATCATCATAATCCAGAAAAAGGTACTTTTTTACAACGTTATTGGTGGTCAACT
GAATATTGGGGTGGTCCAGGTTACCAGTTGTTTTATTTACTCCAGGTGAAGT
TTCAGCTGATGGTTATGAAGGTTATTTAACTAATGGTACTTTAACTGGTGTTT
ATGCTCAAGAAATTCAAGGTGCTGTTATTTTAATTGAACATCGTTATTGGGGT
GATTCATCACCATATGAAGTTTTAAATGCTGAAACTTTACAATATTTAACTTT
AGATCAAGCTATTTTAGATATGACTTATTTTGCTGAAACTGTTAAATTACAAT
TTGATAATTCAACTCGTTCAAATGCTCAAAATGCTCCATGGGTTATGGTTGGT
GGTTCATATTCAAGGTGCTTTAACTGCTTGGACTGAATCAGTTGCTCCAGGTAC
TTTTTGGGCTTATCATGCTACTTCAGCTCCAGTTGAAGCTATTTATGATTATTG
GCAATATTTTTATCCAATTCAACAAGGTATGGCTCAAAATTGTTCAAAAGATG
TTTCATTAGTTGCTGAATATGTTGATAAAATTGGTAAAAATGGTACTGCTAAA
GAACAACAAGCTTTAAAAGAATTATTTGGTTTAGGTGCTGTTGAACATTTTGA
TGATTTTGCTGCTGTTTTACCAAATGGTCCATATTTATGGCAAGATAATGATT
TTGCTACTGGTTATTCATCATTTTTTCAATTTTGTGATGCTGTTGAAGGTGTTG
AAGCTGGTGCTGCTGTTACTCCAGGTCCAGAAGGTGTTGGTTTAGAAAAAGC
TTAGCTAATTATGCTAATTGGTTTAATTCAACTATTTTACCAGATTATTGTGC
TTCATATGGTTATTGGACTGATGAATGGTCAGTTGCTTGTTTTGATTCATATA
ATGCTTCATCACCAATTTATACTGATACTTCAGTTGGTAATGCTGTTGATCGT
CAATGGGAATGGTTTTTATGTAATGAACCATTTTTTTATTGGCAAGATGGTGC
TCCAGAAGGTACTTCAACTATTGTTCCACGTTTAGTTTCAGCTTCATATTGGC
AACGTCAATGTCCATTATATTTTCCAGAACTAATGGTTATACTTATGGTTCA
GCTAAAGGTAAAAATGCTGCTACTGTTAATTCATGGACTGGTGGTTGGGATA
TGACTCGTAATACTACTCGTTTAATTTGGACTAATGGTCAATATGATCCATGG
CGTGATTCAGGTGTTTCATCAACTTTTCGTCCAGGTGGTCCATTAGCTTCAAC
TGCTAATGAACCAGTTCAAATTATTCAGGTGGTTTTTCATTGTTTCAGATTTAT
ATATGGCTGATTATTATGCTAATGAAGGTGTTAAAAAAGTTGTTGATAATGA
AGTTAAACAAATTAAAGAATGGGTTGAAGAATATTATGCTTAA

B. *Lactobacillus acidophilus* PAP (918 bp)

ATGAAAACCTGGTACTAAAATTATCACTTTAGACAACGGTTACCACTTATGGA
CTAATACTCAAGGTGAAGGCGACATTCACCTTATTAGCTCTTCACGGTGGTCCT
GGCGGCAACCACGAATATTGGGAAGACACTGCAGAACAACTAAAAAACAA
GGCTTAGACGTCCAAGTTACCATGTACGATCAACTTGGCTCACTCTACTCAGA
TCAACCTGACTATTCTAATCCTGAAATTGCTAAAAAGTATTTAACTTATGAAT
ACTTCTTAGATGAAGTTGATGAAGTTCGTGAAAAGCTCGGTTTAGACAATATT
TACTTAATCGGTCAAAGTTGGGGTGGGTTATTAGTTCAAGAATACGCCGTTAA
ATATGGTCAGCACTTAAAGGGTGCGATCATTTTCATCAATGGTTGATGAAATC
GACGAATATGTTGCATCAGTTAATCGTAGACGTCAAGAAGTTCTACCACAGA
CTGAAATTGATTTTATGCATGAATGTGAAAAGAACAATGATTACGACAACAA
ACGTTACCAAGATGACGTTCAAATCTTGAACATTAACCTTTGTTGATCGTAAGC
AACCTTCAAAGCTTTACCATCTAAAGGACCTTGGTGGTTCTGCTGTTTACAAC
GCCTTCCAAGGTGATAATGAGTTTGTATCACCAGTAAGTTAAAGGACTGGC
ACTTCAGAGATCAATTACACAAGATCAATGTTCCAACCTTTGCTTACTTTTGGT
GAAAACGAACTATGCCTATTTCAACTGCTAAGATTATGCAAAAGGAAATTC
CTAACTCACGTTTAGTTACTACTCCAGATGGTGGACACCACCACATGGTTGAT
AATCCTACAGTTTACTATAAACACTTGGCTGACTTCATTCGTGAAGTAGAAAA
CGGCACCTTTAAAGGCCAAAATTAA

C. Optimized *Lactobacillus reuteri* PAP (928 bp)

aggcgcgccATGAAACAAGGCACTAAAATTATTACCCTTGATAATGGCTATCATC
TATGGACGAATACCCAAGGTGAAGGTGATATTCATTTATTGGCTTTGCATGGG
GGTCCTGGTGGCAATCATGAATACTGGGAAGACGCTGCTGAACAATTAAAGA
AGCAAGGTCTGAACGTTTCAGGTAACAATGTATGATCAATTAGGTTCACTCTAT
TCTGATCAACCAGATTTTTCTGACCCTGAGATTGCGAAGAAGTACCTTACTTA
CGAATATTTCTTTGATGAAGTAGATGAAGTACGAGAAAAGCTTGGCTTAGAC
AATTTCTATCTTATCGGTCAAAGTTGGGGTGGCCTTTTAGTTCAAGAATACGC
TGTTAAGTATGGGCAACATCTTAAGGGCGCAATTATTTCTTCAATGGTTGACG
AAATTGATGA_gTAcGTCGAC_{cgt}GTTAATGAATTA_{cgt}GAAAAGACTCTTTCTCCA
GAAGCGGTTGCCTTTATGAAAGAATGCGAAGCCAAGAATGATTACAGTAATC
CTAAGTATCAAGAATGCGTTCAAGTAATGAATGAACAATACGT_cGAC_{cgt}AAG
CAGCCATCCAAGCTTTATCATCTTAAAGACCTTGGTGGCACGGCGGTTTACAA
CGTATTCCAAGGTGATAACGAATTTGTGATTACCGGTAAGCTTAAAGACTGG
CATTTCCGTGACCAACTGAAGAACATTAAGGTGCCAACTTTAATTACATTTGG
TGAACACGAAACGATGCCAATCGAAACTGCTAAGACAATGAATAGTCTCATT
CCAAATTCACAGCTAGTTACTACTCCCGATGGTGGTCACCACCACATGGTAG
ATAACCCCGATGTTTATTACAAGCACCTCGCTGACTTTATT_{cgt}AATGTTGAAA
ATAATACGTTTAATAAT_{cggcgcgcc}t

D. Optimized *Myxococcus xanthus* PEP (2070 bp)

ATGTCATATCCAGCTACTCGTGCTGAACAAGTTGTTGATACTTTACATGGTGT
TCAAGTTGCTGATCCATATCGTTGGTTAGAAAGATGAAAAAGCTCCAGAAGTT
CAAAGTTGGATGACTGCTCAAAATGCTCATGCTCGTGAAGCTTTAGCTAAATT
TCCAGGTCGTGAAGCTTTAGCTGCTCGTTTTTAAAGAATTATTTTATACTGATT
CAGTTTCAACTCCATCACGTCGTAATGGTCGTTTTTTTTATGTTTCGTACTCATA
AAGATAAAGAAAAAGCTATTTTATATTGGCGTCAAGGTGAATCAGGTCAAGA
AAAAGTTTTATTAGATCCAAATGGTTGGTCAAAAGATGGTACTGTTTCATTAG
GTACTTGGGCTGTTTCATGGGATGGTAAAAAAGTTGCTTTTGCTCAAAAACCA
AATGCTGCTGATGAAGCTGTTTTACATGTTATTGATGTTGATTGAGGTGAATG
GTCAAAAGTTGATGTTATTGAAGGTGGTAAATATGCTACTCCAAAATGGACT
CCAGATTCAAAAGGTTTTTATTATGAATGGTTACCAACTGATCCATCAATTAA
AGTTGATGAACGTCCAGGTTATACTACTATTCGTTATCATACTTTAGGTACTG
AACCATCAAAAGATACTGTTGTTTCATGAACGTACTGGTGATCCAACTACTTTT
TTACAATCAGATTTATCACGTGATGGTAAATATTTATTTGTTTATATTTTACGT
GGTTGGTCAGAAAATGATGTTTATTGGAAACGTCCAGGTGAAAAAGATTTTC
GTTTATTAGTTAAAGGTGTTGGTGCTAAATATGAAGTTCATGCTTGGAAGAT
CGTTTTTATGTTTTAACTGATGAAGGTGCTCCACGTCAACGTGTTTTTGAAGTT
GATCCAGCTAAACCAGCTCGTGCTTCATGGAAAGAAATTGTTCCAGAAGATT
CATCAGCTTCATTATTATCAGTTTCAATTGTTGGTGGTCATTTATCATTAGAAT
ATTTAAAAGATGCTACTTCAGAAGTTCGTGTTGCTACTTTAAAAGGTAAACCA
GTTCGTACTGTTCAATTACCAGGTGTTGGTGCTGCTTCAAATTTAATGGGTTT
AGAAGATTTAGATGATGCTTATTATGTTTTTACTTCATTTACTACTCCACGTCA
AATTTATAAACTTCAGTTTCAACTGGTAAATCAGAATTATGGGCTAAAGTTG
ATGTTCCAATGAATCCAGAACAAATATCAAGTTGAACAAGTTTTTTATGCTTCA
AAAGATGGTACTAAAGTTCCAATGTTTGTGTTTCATCGTAAAGATTTAAAACG
TGATGGTAATGCTCCAACCTTTATTATATGGTTATGGTGGTTTTAATGTTAATAT
GGAAGCTAATTTTCGTTTCATCAATTTTACCATGGTTAGATGCTGGTGGTGT
ATGCTGTTGCTAATTTACGTGGTGGTGGTGAATATGGTAAAGCTTGGCATGAT
GCTGGTCGTTTAGATAAAAAACAAAATGTTTTTGATGATTTTCATGCTGCTGC
TGAATATTTAGTTCAACAAAAATATACTCAACCAAAACGTTTAGCTATTTATG
GTGGTTCAAATGGTGGTTTATTAGTTGGTGCTGCTATGACTCAACGTCCAGAA
TTATATGGTGCTGTTGTTTGTGCTGTTCCATTATTAGATATGGTTCGTTATCAT
TTATTTGGTTCAGGTCGTACTTGGATTCCAGAATATGGTACTGCTGAAAAACC
AGAAGATTTTAAAACCTTTACATGCTTATTCACCATATCATCATGTTTCGTCCAG
ATGTTTCGTTATCCAGCTTTATTAATGATGGCTGCTGATCATGATGATCGTGTT
GATCCAATGCATGCTCGTAAATTTGTTGCTGCTGTTCAAAATTCACCAGGTAA
TCCAGCTACTGCTTTATTACGTATTGAAGCTAATGCTGGTCATGGTGGTGCTG
ATCAAGTTGCTAAAGCTATTGAATCATCAGTTGATTTATATTCATTTTTATTTC
AAGTTTTAGATGTTCAAGGTGCTCAAGGTGGTGTGCTGCTCAAGGTCGTAA