DEVELOPMENT OF A NOVEL PLASMID-BASED GENE INTEGRATION SYSTEM FOR
LACTOBACILLUS REUTERI FOR THE PERSISTENT TREATMENT 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 Biological Sciences

By

Jeremy Keith LaBarge

May 2010

©2010

Jeremy Keith LaBarge

ALL RIGHTS RESERVED

COMMITTEE MEMBERSHIP

TITLE: DEVELOPMENT OF A NOVEL PLASMID-BASED
GENE INTEGRATION SYSTEM FOR
LACTOBACILLUS REUTERI FOR THE PERSISTENT
TREATMENT OF CELIAC DISEASE

AUTHOR: Jeremy Keith LaBarge

DATE SUBMITTED: MAY 2010

COMMITTEE CHAIR: Michael Black, Ph.D

COMMITTEE MEMBER: Christopher Kitts, Ph. D

COMMITTEE MEMBER: Kenneth Hillers, Ph.D

This work was supported through the generosity of the Office of Naval Research

Abstract

Development of a Novel Plasmid-based Gene Integration System for *Lactobacillus reuteri* for the Persistent Treatment of Celiac Disease

Jeremy LaBarge

Celiac disease (CD) is an autoimmune disorder that affects approximately 1% of the population [55]. CD is characterized by intestinal villus atrophy after consumption of gluten from wheat, barley, or rye. Patients with CD often experience abdominal pain, diarrhea, malnutrition, fatigue, and a failure to thrive. There is currently no treatment for CD. Patients must live on a strict lifelong exclusion of dietary gluten. Due to the high content of gluten in western diets and poor labeling of gluten content, adherence to a gluten free diet (GFD) is difficult [15]. Nearly all the enzymes that can digest the gluten peptide are sensitive to the stomach's low pH. As a result, dietary supplementation with enzymes to digest gluten has yet to produce a viable alternative treatment to a GFD.

We propose to use a resident microbe of the human intestinal tract to express a peptidase to digest the immunoreactive gluten fragments. The bacteria, *L. reuteri*, will colonize the host's intestines and digest the gluten peptides before causing an autoimmune response. To accomplish this task, this thesis describes a food grade, plasmid based system to integrate genes into the genome of *L. reuteri*. The plasmid system utilizes an origin of replication that requires a protein, RepA, to propagate itself. A helper plasmid provides the RepA protein in *trans* to an integration plasmid that cannot provide RepA to itself. The integration plasmid carries a homologous region to the genome of *L. reuteri* allowing for targeted genomic integration. The integration plasmid will not replicate on its own, and will be integrated into the genome if the helper plasmid is absent. To select for these genomic integrants the integration plasmid expresses an erythromycin resistance marker. Using the Cre/Lox system the antibiotic resistance will be removed from the bacterial genome to re-establish the *L. reuteri*'s food grade status. This thesis describes the construction and verification of the above mentioned plasmid tool kit containing the helper, integration, and Cre expression plasmids to integrate genes into the *L. reuteri* genome.

ACKNOWLEDGEMENTS

I thank Dr. Black, Dr. Kitts, and Dr. Hillers for helping me lay a scientific foundation I am proud to build from.

Contents

List of Tables	ix
List of Figures	x
1 Introduction	1
1.1 Symptoms of Celiac Disease Pathology	1
1.2 Prevalence of Celiac Disease	4
1.2.1 Swedish Epidemic of Celiac Disease	4
1.2.2 Breast Feeding's Effect on Celiac Disease	5
1.2.3 Gluten Toxicity	6
1.2.4 Mechanism of Celiac Disease Pathology	7
1.3 Treatment of Celiac Disease	9
1.4 Probiotics	9
1.4.1 Evidence or Probiotic Benefits	10
1.4.2 Previous Molecular Biology in Lactic Acid Bacteria	11
1.4.3 Delivery of Probiotics	12
1.5 Goals of this Work	13
2 General Materials and Methods	14
2.1 Bacterial Strains and Media	14
2.2 Plasmid Isolation and Propagation	16
2.3 Molecular Work	16
2.4 Restriction Digests	17
2.5 Ligations and Digestion Independent Cloning (DIC) Reactions	17
2.6 Transformation	23
2.7 Plasmid Construction	24
3 Creating ClpP Expression System to Drive Peptidase Expression	26
3.1 Previous Peptidase Promoter Systems	26

3.2	New Potential Promoters for Peptidase Expression	27
3.3	Clp Promoter Expression	28
3.4	Reporter Genes	29
3.5	Methods to Create Expression Cassette	30
3.5.1	Building ClpP and CtsR	30
3.5.2	Building GusA constructs	30
3.5.3	Addition of His-Tag to Clp and CpC cassettes	31
3.5.4	Building pCR2.1 Clp-CtsR	34
3.5.5	GusA Assay	34
3.6	Results and Discussion for Expression Cassette	34
4	Plasmid Based Integration System	36
4.1	Methods for Helper and Integration plasmids	39
4.1.1	Methods for Building Helper Plasmid	39
4.1.2	Methods for Building the Integration Plasmids	40
4.2	Integration Plasmid System Results and Discussion	42
4.2.1	Helper Plasmid Construction	42
4.2.2	Construction of the Integration Plasmid	43
4.2.3	Propagation of Integration Plasmids	45
5	Cre Expression Vector	46
5.1	Methods for Cre Expression Vector	50
5.1.1	Building Cre Expression Plasmid	50
5.1.2	Activity Check of pGK5'Erm-Cre	50
5.2	Results and Discussion of Inducible Cre Shuttle Vector	51
6	Summary of Work	54
7	Future directions	56

List of Tables

1	Bacterial Strains	15
2	Primer Table	18
3	Plasmids used in this study	24

List of Figures

1	Intestinal biopsies of CD patients	2
2	Oat phylogeny	3
3	Construction of CPC cassette	32
4	Construction of pGk-Gus-His plasmid(JL116)	33
5	Activity of GusA in <i>E.coli</i>	35
6	Clp-Pro activity in <i>E. coli</i>	38
7	Dual plasmid integration schematic	40
8	Construction of helper plasmids	42
9	MCS and vector map of pCR4 sequencing plasmid	44
10	Construction of integration plasmid	47
11	Integration plasmid digests	48
12	pInt-Clp-GFP (JL120) activity	48
13	pInt-Clp-GusA (JL119) activity	49
14	pGK5'ErmCre Construction	52
15	pGK-5'ErmCre construction and activity digest	54

1 Introduction

CD is often described as an “allergy to wheat”. This colloquial definition is misleading for a two reasons. First, CD is an autoimmune disorder, similar to type I diabetes, not an allergic reaction. Second, the disorder is triggered by the ingestion of gluten, which can come from barley, rye and wheat. Thus, while wheat is the most common source of gluten, it is not the sole source of gluten. Though convenient in describing CD to the lay person, “an allergy to wheat” is actually a poor description of the true nature of CD.

1.1 Symptoms of Celiac Disease Pathology

Outward symptoms of CD vary among patients and can result in a wide spectrum of severities. The disease most often manifests itself with intestinal problems resulting in diarrhea, weight loss, and/or malnutrition [39]. Diarrhea is the main mode of presentation, found in 85% of individuals with CD. Some patients have very severe intestinal problems, whereas other patients have mild to no symptoms of the disease. Other, but less common symptoms of the disease include anemia and osteoporosis. Patients with CD are also more prone to autoimmune disorders like type I diabetes and other malignancies such as non-Hodgkin’s lymphoma, melanoma, and esophageal carcinomas [27],[14]. Many absorption problems are a result of changes in the villus, the absorptive part of the intestine, in the duodenum (the uppermost part of the intestine). Patients with active CD exhibit highly reduced villi, or villus atrophy (Figure: 1 A and B), in their upper intestinal lining and have reduced surface area to absorb nutrients. Furthermore, patients with celiac disease have increased levels of lymphocytes and plasma cells in the lamina propria. Patients also exhibit increased levels of cellular mitosis in the crypts (Figure: 1 C) of the villi, resulting in reduced crypt depth or crypt hyperplasia [39].

The above mentioned symptoms and pathologies are only present when an individual with CD consumes gluten from wheat, barley, or rye [39]. These three grass varieties have a common ancestral origin and all express similar gluten proteins that are activators of CD (Figure: 2). Conversely, gluten obtained from oats, which is more distantly related to the other grains, seldom causes CD

Figure 1: Intestinal biopsies stained with hematoxylin-eosin. (A) Normal healthy patient with normal villi architecture. Note the villus appendages extending into the intestinal lumen. (B) Patient with CD exhibiting atrophy of the villi. Note the smoothing of the villi across the intestinal lining. In both images the intestinal lumen is the top white space and the intestinal layers are on the bottom. (C) Stylized cross section of healthy intestinal villus to show the location of the lamina propria and villus crypts. Figures adapted from [58] [53].

symptoms [39].

Figure 2: Brief phylogeny of grains common in western diets. Wheat, barley, and rye are all descendants of the Triticeae. Conversely oats are more related to the Aveneae. The ancestral similarity of wheat, rye and barley is thought to explain why their gluten proteins trigger CD and oats do not. Figure adapted from [39].

Currently the treatment of CD consists of a strict adherence to a gluten free diet (GFD) for the life of the patient. If the patient does remove gluten from his/her diet, many of the ailments associated with CD reduce in severity or completely disappear. However, living under a gluten free regime is exceedingly difficult due to the high amounts of gluten in western diets. Additionally, foods that are labeled gluten free or foods that typically do not contain gluten are often contaminated with trace amounts of gluten from the manufacturing facility [15] [32]. North America and Southern European countries have established levels of 20 ppm gluten intake as safe for those with CD. Northern Europe has set the threshold at 200 ppm gluten. Catassi et al. (2007), recommends levels closer to 20ppm gluten to protect individuals with heightened sensitivities to gluten. The safe consumption of oats by patients with CD has been debated in the scientific literature. Some studies have shown the consumption of oats to be safe for people with CD [37], [36],[31]; however, different studies have shown some individuals with CD to be sensitive to oat consumption [46]. Holm et al. (2006), conducted a long term study with children to determine if consuming oats resulted CD symptomologies. Diets of children with CD were monitored to ensure a gluten free diet. The children were then asked to consume approximately 50 g of oats per day along with other

gluten free foods. Patients who strictly adhered to the gluten free diet, but include oats in their diet did not exhibit symptoms (reduced villus height crypt depth ratio of and increase in intraepithelial lymphocytes) of CD after 2 years of oat consumption. During the 2-year trial volunteers were given a purified oat product that had been confirmed to be gluten free. A subset of the children involved reported no symptoms after 5 yrs of consuming commercially available oat products. When the patients consumed gluten symptoms of CD arose, indicating the children still had CD [32].

1.2 Prevalence of Celiac Disease

The exact reason for the triggering of CD is unknown. However, there is strong evidence that both genetic and environmental factors affect the onset of CD. Approximately 1% of western populations are thought to be currently affected by CD [55]. Studies have also shown that over the past two decades, CD has increased in Finnish populations. Between 1978-80 CD was found in 1.05% of the population; however by 2000-01 the number of individuals with CD had risen to 1.99% [45]. This statistic was determined by testing stored serum samples from 1978-80 for CD and comparing the samples to fresh samples from 2000-01. Evidence suggests the nearly doubling of CD is due to a true increase in disease prevalence, and not to increased detection of CD. Lohi et al. (2007) completed blind tests on stored serum samples which resulted in the same CD diagnosis compared to the sample's original diagnosis, ensuring increased detection was not the cause of the observed increase in CD prevalence.

1.2.1 Swedish Epidemic of Celiac Disease

In the late 1970's through the 1980's, Sweden had an increase in the number of children and infants diagnosed with CD. The incidence of CD fluctuated so quickly that a change in the genetic profile of the Swedish population could not explain the rapid fluctuation of CD incidence. The rapid change in CD incidence allowed researchers to examine the environmental factors associated with CD [35]. As infants experienced the greatest change in CD, researchers were able to examine the affects of early diet on the likelihood of developing CD.

In the 1980s the Swedish government recommended that new mothers not feed gluten to their

children before 6 months of age. After 6 months, mothers were asked to introduce gluten into their diets. Traditionally Swedish, mothers cease breastfeeding after 6 months, resulting in infants experiencing high levels of dietary gluten without breast milk. Then in the mid 1990s the Swedish government recommended that mothers breastfed for 4 months and then slowly introduce gluten into the diets of infants. After that shift in policy, the rates of CD rapidly dropped. Introducing the infants to gluten slowly while still breast feeding supports the infants' ability to generate an immunological tolerance to gluten. Data gathered from the Swedish CD epidemic has resulted in the encouragement of mothers to prolong infant breastfeeding in hopes of preventing CD [56], [8], [35].

1.2.2 Breast Feeding's Effect on Celiac Disease

The findings from the Swedish epidemic lead to the hypothesis that the timing of initial gluten intake affects a person's chances of developing CD. Researchers have found if breast milk is given in conjunction with initial gluten uptake the infant is 52% less likely to develop CD [5]. Four potential mechanisms are hypothesized to explain why breastfeeding reduces the chance of developing CD. 1) Breastfed children have less gluten in their diets than children who are not breastfed [5]. Reduced exposure to gluten decreases the risk of developing CD in infants. 2) Breast milk protects the infant's intestines from bacterial infections [5]. Breast milk contains lactoferrin, lysozyme and IgA antibodies to protect against microbial infections [51]. Intestinal infections result in increased intestinal permeability which is an important step in generating CD [21]. 3) IgA antibodies from the breast milk may clump or agglutinate gluten preventing uptake of the antibody complexes [5]. Agglutinated gluten complexes would be excluded from passing through the intestinal barrier because of their size. 4) Breast milk may suppress the activity of T-cells and prevent the overreaction of the infant's immune system [5], [51]. Recent research has shown breast milk to modulate an infant's intestinal physiology [26]. It is therefore possible that breast milk may assist in the generation of gluten tolerance while the infant's immune system and gastrointestinal systems are maturing.

1.2.3 Gluten Toxicity

Gluten is storage protein composed of two components, gliadin and glutenin, in wheat, barley and rye. Gluten is most associated with baked goods such as bread. In baked goods gluten imparts a soft or elastic texture to the product. Gluten may also be used as a stabilizer in foods where gluten is unexpected such as ice cream or ketchup. The gluten molecules of wheat, rye, barley, oats are all rich in glutamine and proline residues. When the gluten is partially digested in the stomach the resulting protein fragments are enzymatically modified by tissue transglutaminase (tTG). Human tTG is a 76 kDA protein made of 686 amino acids. The primary functions of tTG are protein cross-linking, cellular signaling, and the addition or removal of amines from amino acids [24]. Levels of tTG rise after tissue damage because it cross-links proteins during tissue repair [75]. The glutamine residues of gluten are converted to glutamic acid residues deamidation by tTG [52]. The change in glutamine residues results in a more negatively charged oligopeptide which has shown greater toxicity to individuals with CD [70]. The modified gluten fragments are more immunogenic because their negative charge allows for a tighter interaction adhesion to the MHC II complexes on CD4+ cells. The modified gluten oligopeptide is able to strongly activate immune cells in the intestine. Vader et al. (2002), examined the activity of tTG to determine why certain grains become toxic when acted on by tTG. By analyzing synthetic polypeptides with different amino acid residue combinations around the glutamine residue, a set of rules were established that could predict the locations of the deaminating activity of tTG. Proline's location was found to greatly affect the targeting of tTG on the oligopeptide. If a proline residue was found immediately after a glutamine or 3 residues following the glutamine, tTG did not deaminate the glutamine to a glutamic acid. However, if a proline was 2 amino acid residues after the glutamine, tTG converted the glutamine to a glutamic acid. Vader et al. (2002) used the specificity rules for to identify potential MHC binding peptides after being acted on by tTG. The partially digested gluten molecules of wheat, rye, and barley were all predicted as targets of tTG. Furthermore, their algorithms showed the gluten molecules of oats did not serve as tTG targets. Vader et al. (2002) also proposed the activity of pepsin, cleavage after phenylalanine, tyrosine, leucine, and isoleucine, promotes the production of

tTG targets in gluten residues [73].

1.2.4 Mechanism of Celiac Disease Pathology

CD is a result of many different environmental and genetic factors. CD has a 10% prevalence among first degree relatives and a 70% concordance with monozygotic twins indicating CD has a strong genetic component [70]. The primary genetic factor associated with CD is related to the HLA II locus. These genes are a subset of the major histocompatibility complex (MHC) class II genes, and are responsible for the presentation of foreign antigens to immune cells. Nearly all individuals correctly diagnosed with CD carry the HLA alleles HLA-DQ2 or HLA-DQ8. The HLA-DQ2 allele confers CD susceptibility in most cases (approximately 90-95%), and the HLA -DQ8 allele is found in the remaining individuals affected by CD [39].

When gluten is consumed it is digested in the stomach. However, due to high levels of glutamine and proline (>30% and >15% respectively) its 3D structure is resistant to much of the proteolytic processes normally associated with gastric fluids. These partially digested peptides then travel into the intestinal duodenum. For CD to manifest, these gluten peptide fragments must pass through the intestinal barrier and enter the lamina propria. There are a number of proposed mechanisms to explain how the peptide fragments pass through the gut barrier, two will be presented below. One model proposes the creation of localized lesions in the gut lining by infections. These lesions increase gut permeability and allow gluten fragments to enter the tissue. Gut lesions would also increase immune cell activity near the lesion. These immune cells may then inadvertently generate the cytokine environment thought to be necessary to generate an autoimmune disease. Furthermore, tTG levels and activity would also be increased to repair the lesion increasing the chances of creating an autoantigen [39]. The second model is based on the activity of the protein zonulin. Zonulin is a 47 kDa protein that has been found to reversibly loosen the tight junctions of the intestine. Zonulin levels are higher in individuals with active CD compared to individuals without CD suggesting the immune response may influence zonulin expression [21]. Immunogenic gluten peptides may readily pass the gut barrier in patients with active CD because of loosened tight junctions in the intestine. Though the precise mechanism used for gluten peptides to pass the intestinal boundary

is unknown, it is widely accepted that these peptides do enter the lamina propria. Once through the intestine, the gluten peptides encounter tTG and are enzymatically modified and become more immunogenic [52].

Autoantibodies are thought to be generated when a tTG/gliadin peptide complex binds to a rare B-cell receptor specific to tTG (cell bound antibodies) and taken up. Upon cellular processing, the tTG and gliadin peptides would be loaded onto MHCII complexes. These gliadin/tTG fragments are then presented on the B-cell's HLA-DQ2 complex. The modified tTG fragments of gliadin have an extremely high affinity to the HLA-DQ2 MHCII protein complex. even capable of displacing previously loaded proteins while in the endosome [77]. A T-cell with a T- cell receptor specific to gliadin (which is thought to be common) must then interact with the presented gliadin fragment on the HLA-DQ2 molecule and activate the B-cell [71]. The activated B-cell then produces an autoantibody against tTG. The consumption of dietary gluten ensures the presence of tTG/gliadin complexes in the intestine. These protein complexes will continue to activate B-cells as mentioned above as long as gluten is a component of the person's diet [71].

The role of auto-tTG antibodies in CD pathology is not clear. One hypothesis is the localized increase in anti-tTG antibodies may result in the inhibition of tTG in the surrounding area. The inhibited tTG would not be able to activate transforming growth factor- β (TGF- β) [6]. TGF- β is involved in the differentiation of epithelial cells. Therefore, if TGF- β is not activated, because of the loss of the tTG by the anti-tTG antibody, cell differentiation might be impeded resulting in mucosal transformations [6].

The tissue damage observed due to CD is a result of the inflammatory response from surrounding CD4+ T-cells. T-cell activation occurs when antigen presenting cells (APC) take up digested gluten peptides (via phagocytosis) and presents the processed fragments on HLA-DQ2 or -DQ8 molecules. A T-cell specific to the gliadin molecule, is then activated by interacting with the APC. The activated CD4+ T-cell then secretes interferon γ which directs the immune system toward a cell mediated response which mediates tissue damage [39] [6], [70]. Structural rearrangements of the intestine occur through a number of poorly characterized steps that are known to be associated with matrix metalloproteinases. The matrix metalloproteinases are able to destroy connective

tissues in the gut. The loss of the connective tissue *via* metalloproteinases results in changes of cell proliferation and cellular migration. These two features result in the villus atrophy of the intestine observed when a patient with CD has consumed gluten [70].

1.3 Treatment of Celiac Disease

There is no cure for CD. Knowledge of the molecular pathways involved with CD has opened new avenues to treat CD. Blocking the activity of tTG, destroy gluten specific T-cells, and altering the cytokine environment in the gut are all possible methods to treat CD [16]. However, the side effects of targeting the previously mentioned elements of CD are thought to be too severe to be considered for current treatments for CD. A strict, life long gluten free diet is the only currently accepted treatment for CD. Investigators have suggested supplementation of prolyl endopeptidases, to digest the gluten peptides before entering the duodenum, into the diets of individuals with CD [66]. Many prolyl endopeptidases are inactivated in the low pH environments of stomach and are sensitive to digestion with pepsin [72]. A prolyl endopeptidase found in *Aspergillus niger* was shown to be resistant to gastric conditions, and may be suitable for dietary supplementation. However, the peptidases are inactive in basic environment found in the intestines. If an enzyme could be identified that survives the low pH of the stomach but is active in the intestines, it would need to be taken before or with a meal much like lactose intolerant patients take lactase with dairy products. Additionally if the peptidase was encased in a protective capsule the enzyme could survive passage through the stomach [17]. Another alternative is to use bacteria in the gut to produce and secrete enzymes to digest immunogenic gluten peptides as they leave the stomach. As a result the bacteria would deliver a virtually unlimited source of the enzyme at the site of celiac pathology.

1.4 Probiotics

Though often overlooked, the microbial communities that inhabit the body serve important functions. In fact, the intestine is the most densely populated microbial community in the human body, and is thought to be where microbial communities exert their greatest effect on the host. For example, using mice with intestines colonized with only *Bacteroides thetaiotaomicron*, researchers

found the bacteria helps the host with digestion, absorption, metabolism of xenobiotic compounds, angiogenesis, and the fortification of the intestinal barrier. No single definition for probiotics has been accepted throughout the literature. However, the common link between proposed definitions is: when probiotic microbes are introduced to the host they will confer some additional benefit to the host or enhance a currently existing process in the host. It has also been found that probiotic microbes help their host by crowding out potentially harmful bacteria. Mice that lacked microbes in their intestines required a higher caloric intake than mice with microbial communities in their intestines [34]. Due to the health benefits of the probiotic organisms, industry is introducing probiotic bacterial strains into foods e.g. yogurt, cheese and ice-cream [54]. Common probiotic genera of bacteria include *Lactobacillus*, *Bifidobacterium*, and *Streptococcus*.

1.4.1 Evidence or Probiotic Benefits

It has been estimated that roughly 300-400 cultivatable species from approximately 190 genera of bacteria live inside the human intestinal tract. Furthermore, the predicted number of bacteria in the intestines is 10^{14} , which equates to 10 times as many bacterial cells as eukaryotic cells in the entire human body [33]. Host organisms significantly benefit from the dietary supplementation of probiotic bacterial strains. Probiotics are often administered to enhance or restore the normal gut flora. When patients receive antibiotics, the antibiotic will often disrupt the microbial balance in the intestines. As a result, some patients suffer from diarrhea after the cessation of antibiotic treatment. The addition of probiotic strains to the diets of those who take antibiotics reduces the chances of antibiotic associated diarrhea. Hickson (2007) found the probiotic strains *Lactobacillus casei*, *L. bulgaricus*, and *Streptococcus thermophilus* reduced the chances of diarrhea when introduced into the diet. Patients, previously on an antibiotic medication, who took a probiotic supplement were ~20% less likely to show signs of diarrhea when compared to individuals given a similar sterilized supplement [30].

Diarrhea in infants is a common cause of infant mortality. Infants are more prone to intestinal diseases because their immune system and microbial gut flora have not fully developed [61]. Researchers have proposed using probiotics to support the infants intestinal tract to defend against

common pathogens like the rotavirus. Saavedra and Bauman (1994) found that adding *Bifidobacterium bifidum* and *Streptococcus thermophilus* to infant formula dropped the rates of rotavirus diarrhea from 19% to 7%. Additionally, they found that the amount of rotavirus shed by the infants taking the probiotic formula was reduced compared to infants who took the same formula without the probiotic strains. After two days 74% of children who took a probiotic supplementation with *Lactobacillus reuteri* showed no signs of rotavirus associated diarrhea. Conversely, only 16% of children showed no signs of diarrhea without the probiotic drink [64].

1.4.2 Previous Molecular Biology in Lactic Acid Bacteria

Probiotics are often common intestinal bacteria that show positive effects on human health. Some researchers have proposed that the benefits of probiotics may be enhanced by modifying these bacteria to perform specific targeted functions. One of the proposed functions of modified probiotic lactic acid bacteria (LAB) is for the delivery of oral vaccines. Oral vaccines are important in large scale vaccination programs because they can be administered to large populations without trained personnel [50]. Different species of LAB (*Streptococcus gordonii* and *Lactococcus lactis*) have been used to present antigens on their surface because of the inherent colonizing behavior of each strain. *S. gordonii* is known to colonize the oral cavities of humans and only transiently pass through the intestines. Conversely, *L. lactis* does not colonize the human intestinal tract and will be excreted after consumption. Researchers have yet to determine if a persistent (due to colonization) or an acute (due to transient exposure) dose of antigen generates a stronger intestinal mucosa immune response. Mercenier et al. (2000) suggest that using different species of *Lactobacillus* may prove the most fruitful in developing engineered probiotics. *Lactobacillus* is a diverse group of bacteria that interacts with the human gut in many different ways. Some species do not colonize the gut and could be used as transient delivery systems, whereas others colonize specific locations of the gut and allow for directed applications of the probiotic.

1.4.3 Delivery of Probiotics

For many applications a probiotic strain is only useful if it survives and multiplies in the intestinal tract of the host. The bacterial strain must pass through the mouth, esophagus, stomach and pass the bile ducts before it reaches the intestines. In the stomach, the bacteria are exposed to a low pH environment and a myriad of peptidases. The bacteria must then resist the bile salts secreted by the liver in the small intestine. Without food or another carrier molecule, bacteria often experience a rapid die off after consumption [25]. To protect the bacteria from harmful digestive enzymes and chemicals, probiotic strains are often given in conjunction with dairy products such as, cheese [25], yogurt [63], and ice cream [29]. The dairy products buffer stomach acids and raise the pH of the stomach so the bacteria are able to survive. Furthermore, it is believed the and lipids polysaccharides in the dairy products help the bacteria survive [25]. There are also a number of technologically advanced methods of encapsulating the bacteria into microcapsules to protect the bacteria. However, many of these methods reduce the viability of the bacteria by several-fold. For review of some of the existing methods see [7]. Beyond using dairy products, a method showing promise is the use of cellulose acetate phthalate (CAP) as a microencapsulating agent. The solubility of CAP is dependent on the external pH of the environment. If the pH is below 6, CAP is not soluble. However, if the pH is greater or equal to 6, CAP becomes soluble. CAP could be used to create protective capsules around the probiotic bacteria and shield the bacteria from the acidic environment of the stomach. Then, when the bacteria pass into the intestine, and the pH approaches 7, the microcapsules would dissolve and release the probiotic directly into the intestinal lumen. Using CAP to protect *Lactobacillus acidophilus* has already shown promise [22]. When encased in a CAP microcapsule *L. acidophilus* was able to survive the acid environment of the stomach with a single log unit loss of viability compared to a complete loss of viability when not protected. *L. acidophilus* is normally quite tolerant to bile salts, so the study was not able to assess if microencapsulation helped with bile tolerance. However, other methods of encapsulating *L. acidophilus* have been known to sensitize the bacteria to bile, while CAP microencapsulation did not sensitize *L. acidophilus* to bile.

1.5 Goals of this Work

The goals of this project are to develop an engineered probiotic that will be an enduring treatment for CD. We propose to engineer a common resident microbe of the human intestinal flora to digest the gluten peptides before they can trigger damage the gut lining. The bacteria will be modified to express a peptidase that digests the immunogenic fragments of gluten on its surface. We chose to anchor the peptidase to the surface of the probiotic instead of secreting the peptidase because the localization of the peptidase would maximize the effectiveness of the probiotic. If the peptidase is secreted from the bacteria it will constantly be lost from the site of pathology. However, if we anchor the peptidase to the surface of the bacteria, it can colonize the site of pathology and localize the peptidase on the intestinal walls where CD pathology is observed. The peptidase gene will be integrated and expressed from the genome of the bacteria to ensure persistent gene expression. If the peptidase were expressed from a plasmid, the plasmid would be lost to the environment without constant antibiotic selection.

The primary site of celiac pathology is the duodenum; therefore we sought a probiotic that would colonize that area of the small intestine. Studies that track the gut microbial flora often use fecal samples to detect bacteria. However, using fecal samples does not identify where the bacteria originated in the intestine. The intestinal colonization site of *L. reuteri* was verified using fluorescence *in situ* hybridization [74]. Valeur et al. (2004) analyzed intestinal biopsies from different intestinal locations from volunteers before and after a dietary regimen of *L. reuteri*. Samples were then histologically prepared and probed for *L. reuteri*. *L. reuteri* was found localized to the human stomach, duodenum and ileum [74]. Therefore, we expect *L. reuteri* will colonize to the site of celiac pathology upon oral administration.

Furthermore, for the engineered probiotic to have commercial success it must be deemed safe by the scientific and food communities. The Lactobacilli have maintained a generally recognized as safe (GRAS) rating since 1997 [2]. The safety of *L. reuteri* has been further investigated among patients who are immune compromised [76]. Wolf et al. (1998) found HIV+ individuals could tolerate 1×10^{10} CFU per day of *L. reuteri* without harm.

Initial administration of this study’s engineered probiotic will likely take place in dairy products, due to the lower cost and the availability of skilled personnel at Cal Poly’s Dairy Products Technology Center. However, using microencapsulation technology might provide a more efficient delivery system as the product moves toward industrial scales. Regardless of the delivery method, it is important that enough probiotic bacteria are given to the host to allow proper colonization of the probiotic strain. The standard amount of bacteria thought to be required for proper colonization is between 10^7 - 10^8 viable cells/g or ml of product [25]. The delivery method used to deliver the engineered probiotic should be judged based on its ability to deliver enough bacteria to successfully colonize the host.

2 General Materials and Methods

2.1 Bacterial Strains and Media

The strains of *Escherichia coli*, *Lactobacillus reuteri* and *Bacillus subtilis* are described in Table 1. *E. coli* and *Bacillus subtilis* was grown in, LB Miller media(FisherBrand). *L. reuteri* was grown in MRS (Oxid). The media was autoclaved at 116°C for 15 min at 121 psi to prevent caramelization. Cultures of *L. reuteri* were grown at 37°C in a candle jar to provide a low oxygen environment. When transforming *L. reuteri* LCM [3] was used. When selecting for plasmids, the appropriate antibiotics were used. For *E. coli*, ampicillin was used at 100 $\mu\text{g/ml}$, chloramphenicol at 20 $\mu\text{g/ml}$, erythromycin at 250 $\mu\text{g/ml}$, and kanamycin 30 $\mu\text{g/ml}$. For *L. reuteri* chloramphenicol was used at 8 $\mu\text{g/ml}$ and erythromycin at 5 $\mu\text{g/ml}$. When using pGK12 or its derivatives, resistance to chloramphenicol and/or erythromycin were induced in *E. coli* and *L. reuteri* using 500 ng/ml and 100 ng/ml respectively. Antibiotic resistance was induced for one hour in *E. coli*, and three hours in *L. reuteri*.

Table 1: Strains of bacteria used

Species	Strain	Genotype
<i>E. coli</i>	dH5 α T1R	F'proA+B LacIq Δ (lacZY-argF)U169 recA1 endA1 HsDR17 (rk-,mk+)poA supE44 thi-1 GyrA96 tonA (confers resistance to phage T1)
	Turbo	F' proA+B+LacI1 Δ lacZ M15/fhuA2 Δ (lac-proAB) glnV gal R(zgb-210 :: Tn10) TetS EndA1 thi-1 Δ (hsdS-lue)7697galU GalK rpsL(Str ^R) endA1 λ^-
	Top10F'	F'[lacI ^q Tn109(Tet ^R)] mcrA Δ (mrr-hsdRMS-mcrBC φ 80LacZ Δ M15 Δ lacX74 deoR nupG RecA1 araD Δ (ara-leu)7697 GalU GalK rpsL (Str ^R endA1 λ^-
	BL21 (DE3)	F ⁻ ompT gal dcm lon hsdS _B (r _B ⁻ m _B (r _B ⁻)) λ (DE3 [LacI LacUV5-T7 gene 1 ind1 sam7 nin5])
	EC101	JM101 <i>E. coli</i> Expressing RepA <i>via</i> plasmid integration. Select with Kan _{40μg/ml} to keep the RepA integrant. [42].
	K12	Wild type strain
<i>L. reuteri</i>	100-23C ATTCC 23272	
<i>B. subtilis</i>	168	

2.2 Plasmid Isolation and Propagation

To isolate plasmids from *E. coli*, cells were grown overnight in LB supplemented with the necessary antibiotics. Plasmids were isolated using the manufacturer's protocol and supplies from the Zippy[®] Plasmid Miniprep Kit from Zymo Research[™]. To isolate plasmids with the ColE1 origin, 600 μ l of cells was harvested before using the Zippy[®] Plasmid Miniprep Kit. However, pGK12 and its derivatives required 1.2 ml of cells. To isolate plasmids from *L. reuteri*, cells were grown under antibiotic selection overnight in MRS. Approximately 2-3 ml of cells were harvested and resuspended in 600 μ l of sterile water. The cells were incubated for 30 min at 37°C with 10 μ l of 10 μ g/ml lysozyme and 5 μ l of 10 KU/ml mutanolysin. The *L. reuteri* cells were then processed using the Zippy[®] Miniprep Kit from Zymo Research[™] using the manufacture's protocol and supplies.

All of the built plasmids were generated and propagated in *E. coli* and then transformed into *L. reuteri*.

2.3 Molecular Work

PCR was accomplished using one of two enzyme systems, Taq polymerase or Phusion[™] polymerase (Finnzymes). With Taq PCR, the GoTaq[®] buffers (Promega) and enzymes were used. A typical reaction contained: 1 μ l template DNA, 0.2 μ M each dNTP, 2 μ M MgCl₂ 1X green GoTaq buffer, 0.5 μ M forward primer 0.5 μ M reverse primer 1 U GoTaq enzyme and nanopure water up to 25 μ l final reaction volume. The following parameters were used to amplify the DNA: 95°C - 2 min followed by 25 cycles of 95°C - 30 sec, 55°C - 30 sec, 72°C - 1.5 min with a final extension of 72°C - 5 min. The extension time would change depending on the length of the amplicon(1 min/kb of amplicon), and the annealing temperature was optimized for each primer set based upon T_m (~2°C below) (Table 2). A typical reaction for the phusion polymerase would consist of: 1 μ l template DNA, 1X HF buffer, 0.5 μ M each dNTP, 0.5 μ M forward primer, 0.5 μ M reverse primer, 0.2 U Phusion[™](Finnzyme) polymerase, and brought to 20 μ l with nanopure water. The reaction would then be run with the following parameters: 98°C - 30 sec followed by 25 cycles of 98°C - 10sec, 55°C - 30 sec, 72°C - 1 min, then final extension of 72°C - 5 min. The block would be preheated to 98°C

to provide a hot start. Annealing temperatures were optimized for each primer set based upon T_m (55°C or 3°C above) (Table 2). Extension times would vary and were adjusted to ~15 sec/kb of amplicon. For both Taq and Phusion reactions, the cycles could be increased depending on starting amount of DNA and final amount of amplicon desired.

2.4 Restriction Digests

Restriction digests were performed using NEB restriction enzymes and buffers. A typical reaction would be run at 37°C for 1-2 hrs. Digests were always done with less than 5% total volume contributed from the enzyme to ensure low star activity. Heat inactivation of the enzymes was accomplished using the recommendations from NEB, usually 65°C to 80°C for 20 min. Single enzyme digests were done in 10 or 20 μ l reactions, while double digests were always done in ≥ 20 μ l reaction volumes.

2.5 Ligations and Digestion Independent Cloning (DIC) Reactions

Ligations of plasmids was accomplished using NEB or Promega T4 ligase enzymes. The reaction was done in NEB digestion buffers (preferably buffer 4) supplemented with 1mM ATP. Incubation times and temperatures varied depending on the supplier of the enzyme and the nature of the cut (sticky or blunt ends). Using an agarose gel image to estimate molar ratios of insert and vector, approximately 3:1 insert to vector molecules were added to the ligation mixture. When ligating inserts with the same cut site on either end, the vector backbone was treated with NEB Antarctic Phosphatase following manufactures recommendations and heat inactivated at 65°C before ligations. Restriction enzymes were heat inactivated before treating the vector with a phosphatase. If heat inactivation was not possible, the reaction was either column cleaned with the Zippy[®] Column Clean Kit or the Zippy[®] Gel Purification Kit. Orientation of inserts were determined using colony PCR with Taq polymerase enzymes.

Table 2: Primers used in the project.

Project	Primer Number	Primer Name	Tm°C	Complementary binding (Tm°C)	Original Template	Notes	Sequence 5'→3'
General Primers	1	M13R	53.8°C	53.8°C		N/A	CAGGAACAGCTATGACC
	2	M13F	62.9°C	62.9°C		N/A	GTTTTC ^{ccc} CAGTCACGAC-GTTGTA
	3	T3	54.3°C	54.3°C		N/A	ATTAACCCTCACTAAAGG
	4	T7	54.3°C	54.3°C		N/A	GA TAATACGACTCACTATAG GG
Helper Plasmid	5	RepA AmpF	73.8°C	64.1°C	pGK12	<i>XhoI</i>	cctcgag ^{ccg} ATTTAAGCAC ACCTTTATTCC
	6	RepA AmpR	69.6°C	56.5°C	pGK12	<i>HindIII</i>	gggaagcttATACGAGTTTTC GCTACTTGTT
	7	RepA Del-F	75.2°C	65.2°C	pGK12	<i>HindIII</i>	ccaagcttCCTTCTAAAAAC GCGAGCTTTCGC
	8	RepA Del-R	69.9°C	58.4°C	pGK12	<i>XhoI</i>	gctcgag ^{cc} ATTTTGCGTAA TAAGAAAAAGG
	9	RepA amp Dic-F	53.8°C	53.8°C	pGK12	N/A	ATAAGCGAAAAAGCTCGGG

10	RepA amp Dic-R	58.4°C	58.4°C	pGK12	N/A	T GTCATAACCTGAAGGAA GATCA
11	RepA del Dic-F	55.7°C	55.7°C	pGK12	N/A	CAAAAGAAAAAGAGTTAT TTGCTG
12	RepA del dic-R	59.4°C	59.4°C	pGK12	N/A	CATAGATACGCCAAAC TCTC
13	Erm del-F	71.2°C	64.1°C	pGK12	dimer	<u>cgaattc</u> GTTAAGGGATGC ATAAACTGCATCC
14	Erm del-F2	59.2°C	44.7°C	pGK212	<i>Eco</i> RI	<u>cgaattc</u> TTATCTCTTTTC AATAGC
15	Erm del-R	63.4°C	52.8°C	pGK12	<i>Eco</i> RI	<u>cgaattc</u> TATAACCCTCTT TAATTGGTT
16	Erm del-R2	73.2°C	65.9°C	pGK12	<i>Eco</i> RI	<u>cgaattc</u> AAAGCCCTTGAG CGTAGCGTGG
17	pLacIPZ-F2	66.3°C	51.6°C	<i>E. coli</i>	<i>Eco</i> RI	<u>gcgaattc</u> AATGGCGCAAAA CCTTTC
18	PLacIPZ-R	82.8°C	56.1°C	<i>E. coli</i>	<i>Eco</i> RI	TAACCAACCACCCGCC <u>Ggaattc</u> CTACAGCCAGTT TGAGGG
pGK-Cre						

Plasmid	19	Cre-F	65.2°C	54.3°C	pMB145	<i>EcoRI</i>	<u>cgaattc</u> ATGTCCAATTTA CTGACCGT
	20	Cre-R	71.8°C	64.0°C	pMB145	<i>EcoRI</i>	cgaattcCTAATCGGCATC TTCCAGCAGG
	21	Erm5'Dic CreF	70.0°C	54.3°C	pMB145	N/A	cagaaataaacttggtATGTCC
	22	pGK5'CreR	58.4°C	58.4°C	pGK12	N/A	AATTTACTGACCGT AACAAAGTTTATTTTCTG TAGTTTTCG
Integration							
Plasmid	23	pGK-CpC Gus Replace-R	70.9°C	53.0°C	CpC	N/A	catatgtatatctccttCCTGATT TGCACCTTATACG
	24	pGK-CpCCtsR Replace-R	68.4°C	51.6°C	pGK12	N/A	catatgtatatctccttCACCTTT TAGGT
	25	Lox66Ori-F2	70.5°C	52.4°C	pGK12	PstI	cccctgcagctaccgttcgtataatgt atgctatacgaagttatCGCGAG CTTTGCGCTTAT
	26	Lox71Erm-R	70.4°C	53.8°C	pGK12	<i>PstI</i>	cccctgcagcataaacttcgtataatgt atgctatacgaacggtaGCTCCA GATCGATTACCA
	27	LacA-Int-F	80.2°C	60.5°C	<i>L. reuteri</i>	<i>NotI</i>	tcgtgcggccgcACGCCGCCA TGTATTTGGTG

	28	lacA-Int-R	81.2°C	57.3°C	<i>L. reuteri</i>	<i>SpeI/NotI</i>	ccactagtg <u>cgccgc</u> CCTGGG TAATACCATCGAC
	29	LacSpeIF	73.2°C	60.5°C	<i>L. reuteri</i>	<i>SpeI</i>	ccactagtgACGCCGCCATG TATTTGGT
pET30 Plasmid	30	pET-HisTag-R	56.1°C	56.1°C	pET30	N/A	GTTAGCAGCCGGATCTCA
	31	pET-Nde-F	47.0°C	47.0°C	pET30	N/A	GAAGGAGATATACATATG
Expression Cassette	32	CtsR-Dic-F	78.9°C	50.7°C	<i>B. subtilis</i>	N/A	acctaaagggaagga ggagcattccATGGGACAT AATATTTCTGACAT tggcgcgccctectect
	33	CtsR-R	78.9°C	54.8°C	<i>B. subtilis</i>	N/A	GATTTGCACTTATACCA gcaaatcaggaggagggg cgcgccATGTTACGTCCT GTAGAAACC
	34	Gus-Dic-F	84.4°C	57.8°C	<i>E. coli</i>	N/A	agtcgtacgacttattag gcgcgcATTGTTTGCC TCCCTGCTGC
	35	Gus-Dic-R	83.5°C	54.9°C	<i>E. coli</i>	N/A	

36	PClp-Dic-F	81.8°C	59.4°C	<i>B. subtilis</i>	N/A	<i>ccgccagtgctgctggaattcg</i> AGGC
37	PClp-R	67.9°C	67.9°C	<i>B. subtilis</i>	N/A	CAGCTTTTGTGTTGACC GGAATGCTCCTCCTTCA
38	pGK5pMCS-R	65.2°C	65.2°C	pGKMCS	N/A	CCTTTTAGG CGAATTCCAGCACACTG
39	TermF	59.3°C	59.3°C	pGKMCS	N/A	GCGG TAATAAGTCGTAGCACTA ACGC

Underlined text identifies restriction enzyme cut sites. *Italicized* text is complementary region used in DIC reactions.

DIC reactions were accomplished by first amplifying the insert and vector with Phusion™ polymerase. The amplified products were purified away from primers in the previous reaction with either Zippy spin columns or gel purification. The amplicons were run out on an agarose gel to estimate relative molar ratios. An approximate 3:1 insert to vector molar ratio was added to each DIC reaction. The DIC reaction contained as follows: X μ l vector DNA, Y μ l insert DNA, 1X HF buffer, 0.2 μ M each dNTP, 0.2U Phusion™ polymerase, and fill to 20 μ l with nanopure water. The reaction was then run at: 98°C-30 sec. followed by 10 cycles of 98°C - 10 sec, 55°C - 30 sec, 72°C - 2 min, then a final extension of 72°C - 5 min. The annealing temperature was determined by the amount of homology between the insert and vector (Table 2). For details on the DIC procedures see [65]

2.6 Transformation

Transformation of *E. coli* was accomplished using chemically competent cells. Cells were prepared using the Z-competent® *E. coli* transformation kit from Zymo Research™. Following the manufacturers protocol and aliquoted into 100 μ l aliquots. Cells were used immediately or frozen at -80°C (snap freezing with liquid nitrogen was found unnecessary). If frozen, the cells were thawed on ice for 5-10 min or until completely thawed. Sample plasmid or ligation DNA (no greater than 5% total volume of cells) was added to 50 μ l of chemically competent cells. The cells were incubated on ice for approximately 10 min and transferred to a 42°C heat block for 30 seconds, followed by an immediate quenching on ice for ~5 min. Cells were removed from ice and 250 μ l of SOC was added and immediately placed at 37°C with shaking (~220 rpm) for 15 min. If the cells were selected with ampicillin, they were plated after the 15 min outgrowth period. If selected with chloramphenicol, erythromycin, or kanamycin, the cells were given 1 hr before plating. Furthermore, if pGK12 or its derivatives were used, resistance was induced using lower concentrations of antibiotic (see bacterial strains and media) for 1 hr. Once the liquid had soaked into the media, the beads were removed and the plates incubated at 37°C overnight or until growth was observed. Often cultures selected with chloramphenicol or erythromycin took ~48 hrs to grow. To transform *L. reuteri*, electroporation was used with a modified protocol from Ahrn et al. (1992). An overnight culture of *L. reuteri* was

grown in LCM [3] at 37°C with low oxygen. Fresh LCM (10 ml) was then inoculated 1/100 from the overnight and allowed to grow an OD₆₀₀= 0.6-0.8. Cells were harvested by centrifugation (4,000 Xg for 5 min at 4°C) and washed twice with equal volumes of sterile ice cold water. After the second wash, the cell pellet was resuspended in 1/10 the original volume with EHR buffer (1mM HEPES 0.5M Raffinose, pH 7.0) [3]. Cells were then centrifuged as before and resuspended in 1/100 original volume EHR buffer. *L. reuteri* cells were then aliquoted (100 µl) into ice cold 1.5 ml tubes. About 1 µg (5 µl) of circular plasmid DNA was then added to the cells and allowed to incubate for 10 min on ice. To each electrocuvette (0.2 cm gap), 80 µl of cells was added and single pulse of 2.5 kV at 246 Ω was applied. The pulse lengths were ~10 ms. Immediately following electrophoration cells were then in 1 ml of pre-warmed (37°C) LCM broth supplemented with induction levels of erythromycin or chloramphenicol (see bacterial strains and media and Table 3) and incubated at 37°C for 2 hrs. 50 µl of cells was spread on a half plate of LCM with full antibiotic selection. The remaining cells were added to 4 ml LCM with full antibiotic selection (Table 3) and allowed to grow at 37°C in a candle jar for 24-48 hrs or until colonies were visible. Electrocompetent *L. reuteri* cells can also be frozen for later use by resusping them in a 1/100 volume of water after the second wash and freezing at -80°C. Aliquots were then thawed and washed in 1 ml water followed by a wash in 1 ml EHR. Cells were then centrifuged as above and resuspended in 100 µl EHR and electroporated as above.

2.7 Plasmid Construction

The construction of each plasmid will be discussed separately in future chapters of the thesis.

Table 3: Plasmids used in this study

Name	Plasmid Description	Resistance Markers
JL101	pGK12 ErmC ORF removed LacIPZ ligated with <i>EcoRI</i> in <i>same</i> orientation as ErmC ORF	Cm [†]
JL102	pGK12 ErmC ORF removed lacIPZ ligated with <i>EcoRI</i> in the <i>opposite</i> orientation as the ErmC ORF	Cm [†]

JL103	pGK12 ErmC ORF removed and self ligated with <i>EcoRI</i>	Cm [†]
JL106	PLST from SlpA blunt cloned into pCR4 sequencing vector in the sense orientation	Amp, Kan
JL108	JL106 with LacA(from <i>L. reuteri</i> clone in using <i>SpeI</i>	Amp, Kan
JL110	JL108 with the pGK12 origin and ErmC inserted with <i>PstI</i>	Amp, Erm [‡] ,Kan
JL111	JL110 with CpC Gus Ligated in with <i>EcoRI</i>	Amp, Erm [‡] ,Kan
JL114	GFP from pET30 GFP is replaced with GusA (via) <i>AscI</i>	Kan
JL115	ClpP-GFP-His ligated into <i>EcoRI</i> sites of pGK-PLT	Erm [‡]
JL116	ClpP-Gus-His ligated into <i>EcoRI</i> sites of pGK-PLT	Erm [‡]
JL117	pCR2.1 with Clp-CtsR TA cloned in sense orientation	Amp, Kan
JL119*	Gus-His integration plasmid containing LacA (from <i>L. reuteri</i>) and pGK12 origin and ErmC gene	Erm [‡]
JL120*	GFP-His integration plasmid containing LacA (from <i>L. reuteri</i>) and pGK12 origin and ErmC gene	Erm [‡]
JL121	pGK derivative with Cre ORF (from PMB145) fused to first 78 aas of ErmC gene	Cm [†]

†: Chloramphenicol must be induced with 250 ng/ml chloramphenicol for before adding 20 μ g/ml chloramphenicol.

‡: Erythromycin must be induced with 100 ng/ml erythromycin before adding 250 μ g/ml erythromycin.

*: Plasmids must be propagated in EC101.

3 Creating ClpP Expression System to Drive Peptidase Expression

3.1 Previous Peptidase Promoter Systems

Initial work to develop an expression system for our peptidase was done by Matthew Shurtleff. The criteria for the expression system is constitutive high expression of the peptidase. Using the criteria, the promoter elements of the gene SlpA from *Lactobacillus acidophilus* was chosen, as SlpA is constitutively expressed at high levels (about 10 percent of the overall cellular protein) [10]. Furthermore, the S-layer protein is actively excreted outside of the bacterial cell and anchored so that it creates a matrix around the bacterial cell [68]. Due to the large numbers of S-proteins found on cellular surfaces, researchers have estimated that an average of approximately 500 S-layer proteins must be synthesized, translocated and anchored every second [68]. With the previous criteria satisfied, Shurtleff replaced the Slp-A gene, leaving the leader and anchor sequences, with different peptidases or GFP. In doing this, he retained the use of the Slp-A promoter, Shine-Dalgarno sequence, leader sequence, anchor sequence, and transcriptional terminator. The cloning of these constructs into *E. coli* was exceedingly difficult. Restriction digests of isolated plasmid constructs with the Slp-A expression system displayed unexpected banding patterns, indicating the plasmid possessed deletions in the open reading frame (ORF) after the SlpA promoter. Shurtleff learned the anchor sequence of the Slp-A gene had been shown to have murein hydrolase activity [1]. This discovery led him to believe the anchor's enzymatic activity was digesting the cell wall of *E. coli*. The problem was further exacerbated because the cell wall anchor sequence was targeted to the periplasmic space of *E. coli*. To alleviate the problem, Shurtleff truncated the C terminus of the anchor sequence. Previous studies have shown that C terminal end of the Slp-A protein was sufficient to anchor proteins into the cell wall of lactic acid bacteria [69]. With the truncated anchor, he was able to build plasmid constructs with GFP and different peptidase cassettes using the Slp-A genetic elements. However, these plasmids were unstable, leading to the belief the plasmids may still be toxic to *E. coli*. To determine the source of the toxicity, Shurtleff cloned different combinations of the Slp-A genetic elements into the pET30 vector. The pET30 vector is a commonly used expression system that has tightly controlled regulation of gene expression. The pET30 plasmid uses IPTG to

induce expression of proteins cloned into the plasmid. In these expression experiments, he discovered that when the Slp-A leader and anchor are expressed alone they are highly toxic to *E. coli*. If a peptidase was inserted between the leader and anchor sequences of Slp-A, toxicity to the cells was not observed, however the recombinant protein was found in inclusion bodies. Two hypotheses were proposed to explain the loss of toxicity: 1) the inclusion bodies were no longer capable of exerting the toxic effect, or 2) the addition of the peptidase may cause a conformational change in the Slp-A leader sequence or the truncated anchor sequence which negates the toxicity. For a more complete explanation of the Slp-A gene structure and the difficulties in protein expression using Slp-A genetic elements, see the Masters Thesis of Matthew Shurtleff [65].

3.2 New Potential Promoters for Peptidase Expression

To alleviate the toxicity problems when using the SlpA promoter elements, a different expression system was chosen. The elements of Slp-A that were suspected to be toxic were the promoter, the anchor, or the leader. Therefore, we decided that for our proof of principle we would use an inducible promoter. Furthermore, the new expression system would not target the expression cassette to the cell surface. Without protein targeting and anchoring sequences, we could mitigate the toxicity potentially imposed upon *E. coli* during the construction of the integration cassettes. Furthermore, it was possible the Slp-A promoter strength was toxic to *E. coli*. The Slp-A promoter may have resulted in the overproduction of its cassette which could overwhelm the protein chaperons or deplete cellular resources. For our new expression system, we sought high expression in *L. reuteri* with minimal expression in *E. coli* to eliminate the toxicity during plasmid construction.

Although gene transcription is well understood in Gram-negative bacteria such as *E. coli*, there has been limited research on gene transcription in Gram positive bacteria [49]. Many inducible systems require multiple genetic elements that are optimized for use in *E. coli*, but may not work in *L. reuteri*. Therefore if we were to use an inducible promoter we would need a system that has been shown to work in Gram positive bacteria. An approach used to obtain constitutive *Lactobacillus* promoters is creating synthetic promoter libraries [59]. To generate the new promoters, Rud et al. (2006) modified known ribosomal RNA promoters. By aligning known rRNA promoters

consensus with known elements such as the -10 and -35 sequences, different strength promoters were developed. Rud et al. (2006) used PCR primers containing a random mixture of nucleotides to change the spacing between known sequences. These new promoters were then used to drive a reporter gene (GusA). By randomizing the nucleotides between the known DNA sequences, promoters were generated with strengths which varied ~ 3 -4 log units [59]. While the generation of a customized promoters may be a valuable tool in optimizing the expression of our peptidase cassette in *L. reuteri*, we chose to seek a known promoter system. Screening through different synthetic promoters would divert us from generating the gene integration system in a timely manner.

3.3 Clp Promoter Expression

Instead of generating new promoters to drive peptidase expression, we chose to use a previously described promoter clpC. In McCracken et.al. (2000) P-clpC was compared to the P-slpA along with a number of different promoters in a variety of species of lactic acid bacteria. In the study the group *Lactobacillus fermentum*, a close relative to *L. reuteri* [23]. P-clpC drove strong expression of the reporter gene, GusA, but at lower levels than reported for P-slpA. With this information in hand, we looked into the Clp family of promoters for a potential promoter we could use to drive expression of our reporter protein.

The Clp family of proteins is involved in mitigating environmental stress, particularly heat stress. The regulation of heat stress in Gram(+) and Gram (-) bacteria is quite different. In the Gram(-) bacteria, heat stress is tolerated by the switching to a set of heat shock proteins. In *E. coli*, a temperature shift from 30°C to 42°C results in a dramatic increase in the levels of σ^{32} to activate over 20 genes related to heat stress [12]. These genes include a number of chaperones and proteases that help refold or destroy misfolded proteins. The shift to higher levels of σ^{32} production is controlled at the translational level. Through a change in the σ^{32} mRNA secondary structure resulting in higher levels of translation [12]. With Gram(+) cells the heat shock response is more complex. There are ~ 200 genes that are induced that fall into six general classes of proteins. Though the specifics of these classes are not important for this thesis, the combined effect is to increase the amounts of chaperones and proteases during stressful episodes. The P-clpP natively

drives the expression of proteins classified as class III stress proteins. These class III stress proteins have play a roll

3.4 Reporter Genes

Before expressing the peptidase with ClpP we expressed benign reporter genes to ensure the integration system worked. One of the reporter genes we have chose to use is the protein β -glucuronidase (GusA). A common reporter in *E. coli* is β -galactosidase. We felt β -galactosidase would not work in our system because *L. reuteri* has β -galactosidase activity. Therefore, it would be difficult to determine if the β -galactosidase activity was from our reporter or if it was background activity from the native β -galactosidase expressed in the genome. GusA activity has not been shown in *L. reuteri*. Therefore, detected GusA activity would be attributed to our expression system.

β -glucuronidase (GusA or uidA) is a 603 aas acid hydrolase, found in many organisms including in *E. coli* and is known to cleave a variety of β -glucuronides. It has been widely used in plants [4] and microorganisms [59], [49], [57]. The GusA reporter systems had many different substrates that allow for enzymatic activity detection of using chromogenic substrates. This study used X-Gluc and PNP-Gluc which are analogues to X-Gal and ONPG to the β -galactosidase substrates that are commonly used when attempting α -complementation. When X-gluc is cleaved by GusA, it forms a blue precipitate on agar plates. This allows for rapid screening of the bacterial clones which contain GusA activity. When PNP-Gluc is cleaved by GusA, it forms a yellow soluble compound that can be detected using spectrophotometric techniques. Therefore, the enzymatic units of the cells can be determined with simple chromogenic assay procedures.

We also wanted an immediate assay that could be checked *in vivo*. We chose to use green fluorescent protein (GFP) as a secondary reporter protein. GFP offers a number of benefits that complement the GusA reporter system. Cellular clones with GFP can be quickly screened by detecting visible fluorescence when exposed to long or short wave ultraviolet light (UV). If necessary, there are also antibodies against GFP that can quantify the amount of GFP in the cell. Using these two systems (GusA and GFP) we can generate both a readily detectable qualitative measure with GFP, and a simple procedure for quantitative measurements using the chromogenic assay for GusA.

3.5 Methods to Create Expression Cassette

3.5.1 Building ClpP and CtsR

Genomic DNA was harvested using the ZR Fungal/Bacterial DNA Kit™(Zymo Research). To amplify Clp-P, primers PClp-DicF and PClp-R (Table 2) were used to amplify a 106 bp fragment from the *B. subtilis* genome. These primers were based off primers TM-197 and TM-187 from Derr et al. (1999). The 5' end of the P-Clp-P amplicon contains a homologous region to the pGK-MCS plasmid and an *EcoRI* cut site. CtsR was amplified using the primers CtsR-DIC-F and CtsR-R to generate a 659 bp fragment. The CtsR primers were based off the primers ID-1 and ID-2, also from Derr et al. (1999). The non conventional start codon (GTG) of CtsR was replaced with an ATG [19]. Furthermore, the distance from the Shine-Dalgarno sequence to the start codon was increased to 8 nt from 6 nt in the hopes of increasing the translational efficiency of the CtsR transcript. The CtsR transcript would be made as a polycistronic mRNA, therefore a 167 bp was added between the GusA gene and the CtsR protein to act as a spacer for the two open reading frames.

3.5.2 Building GusA constructs

The GusA gene was amplified using the primers Gus-DIC-F and Gus-DIC-R to amplify the protein DNA sequence. *AscI* sites were engineered before the start codon and before the stop codon of GusA. Therefore, GusA could be interchanged with restriction enzymatic cleavage with other peptidase ORFs. To generate the functional expression system, the P-Clp-P, CtsR and GusA amplicons were column purified using the DNA Clean and Concentrator Kit™(Zymo Research) and diluted 100 fold. These amplicons were then added to a single PCR reaction containing the pClp-DicF and Gus-DIC-R primers. The resulting amplicon was 2,573 bp long. To add the Slp-A terminator to the P-Clp-P expression system, an existing plasmid pGK (MCS)PGT [65] was utilized. (The 'T' of PGT is the SlpA terminator.) The primers Term-F and pGK5pMCS-R were used to amplify the plasmid backbone (4617 bp in length) where the expression cassette would be inserted. P-ClpP-CstR-GusA (amplified with PClp-DicF and Gus-DicR) was then switched into the pGK plasmid backbone *via* DIC. This plasmid was then transformed into chemically competent *E. coli* (Turbo (NEB) or dh5 α -

T1R (Invitrogen) cells), and cells were then screened for positive transformants using colony PCR with M13R and CtsR-R primers to generate an amplicon of 802 nt. The completed expression cassette containing the Clp promoter, CtsR repressor protein and SlpA transcriptional terminator will be called *CpC* from this point forward. The cassette between the *AscI* sites will be after the *CpC* designation, the aforementioned sequence would be called *CpC-GusA*. The *CpC-GusA* fragment was then cloned into the commercial vector pCR4 (Invitrogen) containing LacA gene fragment and ErmC-pGKori sequence (discussed in more detail in the integration vector section 4.1.2).

3.5.3 Addition of His-Tag to Clp and CpC cassettes

The histidine(His) tag for each cassette was derived from the pET30 vector (Invitrogen). The pET30 plasmid has both C- and N-terminal His tags engineered into the plasmid vector. The lab has a pET30-GFP plasmid built with *AscI* sites flanking the ORF of GFP which is in frame with the C terminal His-Tag of the plasmid. For the construction of pET30-GFP see [65]. GusA was inserted into pET30 by replacing the GFP coding region with GusA at the *AscI* sites. Transformants were screened for clones with GusA in frame with the His-tag using the primers Gus-DIC-F and the pET30 T7-terminator (Invitrogen) for an amplicon of 1952 nt. To insert Gus-His or GFP-His into a pGK-CpC Gus primers were developed to amplify the vector so we could include or remove the CtsR repressor (Figure 4). Primers were also developed to amplify the cassette out of pET30 and retain the His-tag. To create pGK-CpC-GusA-His, pGK-CpC-GusA was amplified using the primers pGK-CpC-R and pGK-CpC-F to yield an amplicon of 5342 bp. Gus-His or GFP-His were amplified from their respective pET30 plasmids using pET Nde-F and pET-His-top-R, which generated a band at 1923 bp for GusA and 828 bp for GFP. DIC was used to stitch either GusA or GFP into the pGK CpC backbone. To remove CtsR, the primer set pGK-CpC-F and pGK-Clp-NoCtsR-R were used to generate an amplicon at 4710 bp. This plasmid backbone would then be combined with GusA or GFP using DIC as with the CPC inserts. When the *CtsR* gene was removed from the expression system, the designation changed from *CpC* to *Clp*.

Figure 3: Construction of the CPC expression cassette. Primer numbers refer to primers from Table 2. For more details see materials and methods for expression cassette construction.

Figure 4: Construction of pGk-Gus-His plasmid (JL116). To keep the CtsR repressor protein, use primer 6 in place of primer 5. Primer numbers refer to primers from Table 2. For more details see materials and methods in the expression cassette construction section. To generate create JL115, use the same primers as shown above but use pET30 GFP instead of JL114.

3.5.4 Building pCR2.1 Clp-CtsR

Clp-CtsR was amplified with Taq polymerase from pGK-CpC-Gus using CtsR and PClp-Dic-F for an amplicon 738 bp long. The amplicon was then ligated into pCR2.1 using the original T/A Cloning Kit (Invitrogen) following manufactures instructions and cloned into TURBO *E. coli* (NEB). Colonies were screened using colony PCR using M13R and CtsR-R for an amplicon of 870 bp. The resulting plasmid was names JL117 or pCR 2.1 Clp-CtsR.

3.5.5 GusA Assay

The Gus assay is a modified protocol from [57]. Cultures were grown to an OD₆₀₀ between 0.4 and 0.8 and put on ice for 5 min. Cells were pelleted by centrifugation (16,000 Xg, 2 min.), resuspended in Gus buffer (50 mM NaHPO₄(pH 7), 10 mM β -mercaptoethanol, 1 mM EDTA, 0.1 % Triton X-100), and lysed by 3 rounds of sonication for 20 seconds (intensity level 6 on the Fisher Scientific ultra-sonicator). Lysates were centrifuged for 5 min at 16000 Xg max speed to pellet cell debris. 50 μ l of cell lysate was then added to 0.5 ml of Gus assay buffer supplimented with 1.5 μ M *p*-Nitrophenyl- β -D-galactopyranoside (pNPG)(Acros Organics). The sample was then incubated at 37°C until yellowing was observed or 30 min had passed. To stop the reaction, 100 μ l of the sample was removed and added to 0.8 ml 2 mM carbonate buffer(pH2). Read the sample at 405 nm. The extinction coefficient used to calculate the PnP⁻ was 18,700M⁻¹ at 405 nm [67].

3.6 Results and Discussion for Expression Cassette

Our goal was to create an expression system that would drive the expression of our reporter genes (GusA and GFP). We were able to properly stitch together the bacterial components as shown in Figure 3. We had initially hoped that our new reporter system would not be active in *E. coli*, yet be active in *L. reuteri*. However, we saw strong activity of the promoter in *E. coli* harboring pGK-clp-Gus or pGK-clp-GFP (Figures 5 and 6). Reduced Gus activity observed in the presence of the CtsR protein (middle column compared to right column in Figure 5) lead us to believe the CtsR protein inhibits much of the Clp-P activity in *E. coli*. However, it does not completely inhibit

Figure 5: Activity of GusA in *E. coli*. Assays were done in a TURBO *E. coli* strain background. Enzymatic units of GusA were defined as $\mu\text{M pNP}^-/\text{min}$. No statistical significance could be applied to these samples because too few replicates.

the Clp-P promoter's activity. Therefore, if future peptidase cassettes are toxic in *E. coli*, the Clp-P promoter system may not be suitable. When building the CpC cassette (Clp-P-CtsR-GusA/GFP) there were concerns that the ~ 150 nt spacer between the CtsR protein and the expression cassette might contain a transcriptional terminator sequence. In Figure 5, the cells that contain pGK-CpC-GUS show higher GusA activity ($\sim 6\text{X}$) than cells without the pGK plasmid. This piece of evidence supports increased levels of GusA as a result of the plasmid's gene expression. When the CtsR protein is removed from the CpC cassette (pGK-clp-Gus), there is a dramatic increase in the levels of Gus activity. The pGK Clp-Gus construct shows $\sim 141\text{X}$ and $\sim 812\text{X}$ more activity when compared to pGK CpC-Gus and WT strains of *E. coli* (Figure 5). The reduced Gus activity observed in the presence of the CtsR protein (middle column compared to right column in Figure 5) lead us to believe the CtsR protein effectively inhibits much of the Clp-P activity. However, it does not completely inhibit the promoters activity.

GFP expression under Clp-P was also monitored (Figure 6). Strong fluorescence was seen in cells containing pGK-clp-GFP. However, there was background fluorescence observed in *E. coli* with pGK-clp-Gus. The fluorescence is so much brighter with pGK-Clp-GFP that we are confident the fluorescence seen is due to GFP expressed by our plasmid, and not background fluorescence.

Approximately a third of the cells show GFP fluorescence. Previous experiments have shown plasmids with the pGK12 origin of replication to be highly unstable in *E. coli* (data not shown). Shortly after the removal of antibiotic selection pGK plasmids are lost from the *E. coli* population. Therefore, the plasmid instability may generate a population of cells with or without the pGK-clp-GFP plasmid, and explain the low proportion of fluorescing cells.

We also co-transformed the pCR2.1 containing the clp-CtsR cassette into cells containing pGK-clp-GFP (right most panel). We hypothesized we could express CtsR on a secondary plasmid and have the CtsR protein inhibit the Clp-promoter on our expression plasmid (in this case pGK-clp-GFP). The far right panel in figure 6 shows highly reduced fluorescence when compared to pGK-clp-GFP (middle right panel). The reduced levels of fluorescence indicate CtsR may be working in *trans* on the expression plasmid and may reduce its promoter activity. It is possible we could express cassettes that are toxic in *E. coli* if the expression is truncated by providing CtsR in *trans*.

4 Plasmid Based Integration System

If the peptidase used to digest gluten was expressed from a plasmid the plasmid would be lost shortly after antibiotic selection was removed from the bacteria. As a result, patients using the probiotic would need to be on an antibiotic regime for the probiotic to persistently express the peptidase. Additionally, FDA regulations do not grant food grade status to products which can introduce antibiotic resistance to the environment. Therefore our probiotic must stably express the peptidase without antibiotic selection. To accomplish this goal the peptidase must be expressed from the genome of *L. reuteri* and not a plasmid. Integration into the genome occurs when a bacterial cell takes up a foreign DNA and incorporates the DNA into its genome. For these two events to happen in *L. reuteri* DNA has to pass through the thick cell wall, and then recombine with the bacterial gDNA. Integration of genes into bacteria has been accomplished by [38] one-step genomic integrations [44] followed by screening. In general screens require a readily observable product and many man hours to produce positive genomic integrants. *L. reuteri* has a low transformation frequency which further decreases the likelihood of obtaining positive genomic

integrants. Therefore, the one step integration system does not provide a feasible method for genomic integration. Another method used to integrate genes into the genome of Gram positive bacteria uses two plasmids that are dependent on each other for proper replication[60].

Figure 6: Activity of the Clp-promoter in *E. coli*. All plasmids are in dH5α cells to account for possible native fluorescence. Exposure times are in the lower right corner of each image and are reported in seconds or fractions of a second. Magnification is 1000X.

Our system will be based off the plasmid pGK12, [40] which harbors two inducible antibiotic markers: chloramphenicol acetyl transferase (CAT) and erythromycin resistance marker C (ErmC) [20]. The pGK12 plasmid is able to replicate in a wide host range including both Gram positive and negative bacteria such as, *L. reuteri* and *E. coli*. The origin of replication in pGK12 is from the plasmid pWV01 [62], [43]. For pGK12 to replicate, the RepA, protein must be expressed to bind to the origin of replication on the plasmid. If RepA is not present, plasmids with the pWV01 origin will not replicate inside a cell. Our system takes advantage of the RepA dependence by providing RepA in *trans*: a helper plasmid will express RepA while a second plasmid, the integration plasmid, will have the RepA dependent origin of replication, but will not express the RepA protein. Therefore, the integration plasmid will not propagate itself unless the helper plasmid supplies the RepA protein (Figure 7A). Built within the integration plasmid will be a genomic integration targeting region (LacA), the expression cassette (peptidase), and a selectable marker (ErmC). Using this system, we increase our chances of a genomic integration by selecting for the presence of both plasmids to generate a population of cells that contain both the helper and integration plasmids. Genomic integration will then be promoted by removing antibiotic selection for the helper plasmid (chloramphenicol) and destabilizing RepA with thermal stress [60] (Figure 7B). If RepA is absent, the integration vector will no longer be able to replicate. Therefore the integration plasmid faces one of two outcomes 1) integrate itself into the host's genome *via* homologous recombination or 2) be lost by the cells during cell division. The first option can be favored by continuing selection for the antibiotic marker on the integration vector (erythromycin). Therefore cells that do not integrate the plasmid will lose their antibiotic resistance and will be selected against.

4.1 Methods for Helper and Integration plasmids

4.1.1 Methods for Building Helper Plasmid

Erm-Del-R and Erm-Del-F2 primers were used to amplify the vector backbone from pGK12, the amplicon was 3732 bp. LacIPZ (1-92 aas of the LacZ /*alpha* fragment) was then amplified from the gDNA of *E. coli* K12 using LacIPZ-F2 and PLacIPZ-R, yielding an amplicon of 1584 bp.

Figure 7: Dual plasmid integration schematic. (A) The helper plasmid provides RepA to both the integration plasmid in *trans*. By selecting with different antibiotics, both plasmids are maintained in the cells to generate a population cells with both plasmids. (B) Using thermal stress and removing the antibiotic used to select the helper plasmid genomic, integrants are selected for.

These two products were then digested with *EcoRI* and the pGK12 vector amplicon was treated with Antarctic phosphatase (NEB). The two fragments were ligated together and transformed into chemically competent *E. coli* DH5 α or TURBO. Blue transformants were screened for the orientation of the LacIPZ insert using colony PCR. Colonies that were white were screened and verified to be pGK12 without the erythromycin marker.

4.1.2 Methods for Building the Integration Plasmids

To act as a place holder for the blunt cloning site of the pCR4 vector. the SlpA expression vector developed by Matthew Shurtluff was used. Primer sequences are defined in [65]. Briefly, Pro-Dic-F and ProL-R were used to amplify up the Slp-A promoter and leader(PL) with an expected size of 375 bp. The SacI anchor sequence was amplified using A-TermF and Sac-1-term-R for a 165 bp amplicon. The SlpA terminator region was amplified using Term-Dic-F and A-Term-Dic-R with an expected size of 135 bp. The SacI anchor(S) and the terminator(T) regions were then gel purified and stitched together using DIC to create 'ST'. To combine 'PL' and 'ST', they were both digested using *AscI* and ligated together. The ligation product was amplified by PCR using the primers Pro-Dic-F and A-Term-Dic-R. This amplicon was TOPO cloned into pCR4 blunt sequence vector using the manufactures protocol and transformed into Turbo *E. coli* cells (Invitrogen). The colonies

were screened using colony PCR to determine the orientation of the PLST insert. Pro-Dic-F and M13R were used for an amplicon of ~ 800 bp in the sense orientation to create JL106 (Figure 10).

LacA was amplified from *L. reuteri* (23272) gDNA, obtained using the ZR Fungal/Bacterial DNA kitTM (Zymo Research), using the primers LacSpe1-F and LacA-Int-R for a band that is 981 bp in size. The LacA amplicon was inserted into pCR4 by digesting the plasmid and LacA amplicon with *SpeI*, and ligating the two pieces together. To screen for the correct orientation of the LacA insert, colony PCR was used with the primers LacSpe1F and M13R for a band of 1036 bp in size to create JL108. The origin of replication and ErmC (Erm-Ori) marker was amplified from pGK12 using the primers Lox66-Ori-F2 and Lox71-Erm-R to yield a band 2,053 bp in size. This amplicon along with PCR4-LacA-PLST(JL108) was digested with *PstI*. The vector was then treated with Antarctic Phosphatase(NEB), and the two fragments were ligated together. The ligation was transformed into *E. coli* (TURBO). The orientation of the ErmC-Ori insert was determined using colony PCR and the primers M13R and Lox71-Erm-R to yield a band 2,109 bp in size to create JL110. Once the pCR4 plasmid contained both LacA, the ErmC marker, and the Origin of replication, it was designated pCR4 LAEO or JL110. To insert CpC GusA into JL110, pGK-CpC-GusA and JL110 were digested with *EcoRI*, gel purified using the Zymo Research Gel Purification kitTM and ligated together. The ligation was transformed into *E. coli* (TURBO), and orientation was determined using Gus-Dic-F and T7 for an amplicon 2022 bp. The resulting plasmid was designated pCR4-LAEO-CpC-Gus or JL111. To build the integration plasmid with His-tagged cassettes, JL111 was digested with *NotI* and self ligated to remove the pCR4 scaffold. The ligation was then used as a template with the primers pGK-cpc-F and CpNoCtsR-R or pGK-cpc-R. If CpNoCtsR-R was used, the CtsR repressor was removed along with GusA from the integration vector and an amplicon of 7,192 bp was expected. Conversely if pGK-CpC-R was used, CtsR remained intact and GusA was removed and a band 7,824 bp was amplified. GusA-His and GFP-His were amplified using pET30 Gus or GFP plasmids with Pet-NdeF and pET-His-tag-R. The GusA-His and GFP-His were then separately stitched into the integration vector using DIC. The reaction was then transformed into *E. coli* (EC101). Transformants were screened using colony fluorescence and diagnostic digests with GFP-His. pInt-Gus was verified using colony PCR, using the primers pET-Nde-F and Aterm-

Figure 8: Construction of the three helper plasmids. Numbers above the primer refer to the primer table (Table 2).

DicR with positive colonies having an amplicon of 2013 bp. Positive colonies were then verified by restriction digests of the plasmid preps (Figure 11).

4.2 Integration Plasmid System Results and Discussion

4.2.1 Helper Plasmid Construction

As stated previously, pGK12 contains two antibiotic resistance markers to erythromycin and chloramphenicol. The helper plasmid was designed to include the chloramphenicol marker with the erythromycin resistance marker deleted. During our preliminary work it was discovered that *E. coli*

loses pGK derivative plasmids quickly when there is no antibiotic selection. To determine plasmid stability we cloned LacZ into the ErmC locus under the control of the *E. coli* LacP. We found that at 37°C, approximately 50% of the colonies showed signs of colony sectoring when plated on a LB-X-gal/IPTG plate without antibiotic selection. Furthermore, when the temperature was increased to 42°C nearly all the colonies were white, indicating the plasmid was lost, after being plated on a LB-X-gal/IPTG plate (data not shown). Three plasmids that could serve the function of the helper plasmid were built: a plasmid with LacIPZ the same orientation as the erythromycin promoter (pGK- Δ Erm-LacZ_{sense} or JL101), a second with LacIPZ in the opposite orientation as the erythromycin promoter (pGK- Δ Erm-LacZ_{antisense} or JL102), and a third where the erythromycin ORF was removed and the ends of the plasmid were ligated together (pGK- Δ Erm or JL103) (see Figure 8).

We were able to construct all three versions of the helper plasmid (Figure 8). We chose to use the pGK Δ Erm lacZ_{sense} plasmid as our helper plasmid because it offers more methods of detection while in *E. coli* (via blue/white screening). Furthermore, the LacZ cassette provides more primer binding sites to be used during future verification processes.

4.2.2 Construction of the Integration Plasmid

Construction of the integration plasmid provided us with some interesting challenges: 1) propagating a plasmid without a functional origin of replication, 2) the final integrated product had to be food grade, and 3) the plasmid must contain the site for integration into the *L. reuteri* genome. To deal with the propagation issue we decided to build the integration plasmid using the pCR4 vector (Invitrogen) as a scaffold. The blunt end sequencing version of the plasmid was chosen to host the components of our integration plasmid. This version of pCR4 was chosen because of the paucity in restriction sites within the plasmid (Figure 9). Fewer restriction sites allowed the system to maintain high versatility. If the plasmid contained too many restriction sites, future integration sequences might be limited because they contain restriction sites found within the multiple cloning site (MCS) of the pCR4 plasmid. When construction of the integration plasmid started, we planned on using the SlpA expression system to drive peptidase expression. Therefore, constructs were built

Figure 9: MCS and vector map of pCR4 sequencing plasmid. Note the locations of the native *SpeI*, *PstI* and *NotI* restriction sites. Figure adapted from Invitrogen website, <http://www.invitrogen.com>

using PLST (from Slp-A) as inserts to blunt end clone into pCR4 [65]. PLST was used as a placeholder in the TOPO Cloning site while the remaining parts of the integration vector were inserted into pCR4.

The original plan was to use the *SpeI* and *PmeI* sites to hold our screenable marker, the origin of replication, and the site of integration (see diagram of pCR4 vector). The fragment ligated into the *SpeI* site would have a *NotI* site located at its 5' end downstream the *SpeI* site. The pCR4 plasmid contains a native *NotI* site downstream of the DNA sequences we will insert. Using these two *NotI* sites (native and engineered) we can cut away the pCR4 backbone and self ligate the integration plasmid together. This system allows us to build an integration vector in the context of a high copy number plasmid and then isolate the integration vector once it is constructed.

As stated in the expression vector section above, the CpC-GusA cassette was inserted into the pCR4 scaffold vector using *EcoRI* sites. Furthermore, the GFP or GusA cassettes possess a His-tag

acquired from pET30 MCS. Built into the pET30 MCS is an additional *NotI* cut site before the His-tag coding sequence. This additional *NotI* site is incompatible with our plan to excise our suicide vector from PCR4, as it would result in the loss of the terminator region of our expression cassette. To solve the problem, we delayed inserting our His-tagged marker until the final stage and used a digestion independent cloning step to complete the construction (see details in methods section and/or Figure 10).

4.2.3 Propagation of Integration Plasmids

The integration plasmids must have an external source of the RepA protein to replicate. We found the helper plasmid was unable to provide adequate levels of RepA in *E. coli* to propagate both plasmids. To solve this issue, we obtained a RepA expressing strain of *E. coli* (EC101) which has been used in the construction of a similar integration vector [42]. Using this strain we were able to successfully construct and propagate the integration plasmid. To confirm the construction of the integration cassettes (GFP or GusA), diagnostic digests were used. Both plasmids were digested using *HpaI* and *XhoI* (Figure 11). *HpaI* cuts near the 3'-end of the ErmC ORF, while *XhoI* cuts within at the 3'-end expression cassette and is introduced via the pET30 vector. The expected digest products for pInt-Clp-Gfp with *HpaI* and *XhoI* are 2,669 bp and 1,398 bp. The digest of pInt-Clp-GFP shows 2 bands at approximately 1.4 kb and 2.6 kb (Figure 11). The expected digest products for pInt-Clp-Gus with *HpaI* and *XhoI* are 3,765 bp and 1,398 bp, which were also observed. Therefore, these digests confirm both pInt-Clp-Gus and pInt-Clp-GFP are built correctly and contain the expression cassette from pET30 and the ErmC ORF.

To prepare *L. reuteri* for gene integration the pGK Δ Erm lacZ_{sense} helper plasmid was transformed into *L. reuteri* 10-23C. After inserting the helper plasmid, we successfully transformed pInt-clp-GFP into *L. reuteri*, with Δ Erm lacZ_{sense} providing RepA. A slight increase in GFP fluorescence was observed in *L. reuteri* containing pInt-clp-GFP (data not shown). The CtsR repressor protein has been found in many Gram positive bacteria including *Lactobacillus* [19]. *L. reuteri* may be expressing CtsR from its genome reducing the activity of the clp promoter in the integration plasmid. At this time, we have not verified genomic integration of the plasmid into the LacA gene.

However, we do have a strain of *L. reuteri* which exhibits a continued resistance to erythromycin and sensitivity to chloramphenicol (data not shown). We are continuing our investigations to verify and increase the integration events.

A greater proportion of the EC101 with JL120 (int-Clp-GFP) cells fluoresce than do the dH5 α cells with JL115(pGK-Clp-GFP) (compare glowing cells between Figure 12 and Figure 6). As stated previously, the pGK plasmid is not stable within *E. coli*; therefore, it is expected that the population of cells may contain a subset of cells that have lost the pGK plasmid derivative. EC101 expresses RepA from its genome which may increase the stability of plasmids with the pGK origin based upon expression levels. An increased stability of the pGK derivatives in EC101 would explain the greater proportion of GFP positive cells, than pGK derivatives in other *E. coli* strains. Individually cells harboring pInt-Clp-GFP(JL120)(*E. coli* EC101) fluoresced less brightly than cells harboring to pGK-clp-GFP(JL115)(dH5 α). The different levels of expression could be due to the background strain of *E. coli* (dH5 α vs EC101).

The pInt-clp-gus plasmid(JL119) shows increased GusA activity compared to bacteria without the plasmid (Figure 13). In fact, the EC101 harboring JL119 has $\sim 590X$ more GusA activity as EC101 alone. However, pGK-clp-Gus (JL116) has $\sim 4X$ the activity as seen with JL119 in EC101 (compare Figure 13 and Figure 5). The differences in enzymatic units produced may be due to differences in the plasmid copy number allowed by each strain. Another possibility is that dH5 α may have higher levels of transcription and/or translation than EC101.

5 Cre Expression Vector

For our antibiotic marker in the integration vector we chose to use the ErmC gene from pGK12. However if we wanted our probiotic to be food grade, we recognized the fact that our end product cannot harbor resistance to erythromycin [18]. Therefore we developed a system that allowed us to use the antibiotic resistance as a selectable maker during the construction of our bacteria and remove the antibiotic resistance after the plasmid integrated into the genome. To accomplish this, the Cre/Lox system was used. Cre is an integrase protein that is found in the coliphage P1 [13].

Figure 10: Construction of the integration plasmid. The numbers above the primers refer to Table 2. To build the integration plasmid with GFP, use the same primers and methods, however use pET30 GFP to combine with the integration vector backbone instead of JL114.

Figure 11: Verification of integration plasmid construction *via* enzymatic digestion. Enzymes used are in parentheses. Expected band sizes for pInt-clp-GFP are 2669 bp and 1398 bp, pInt-clp-Gus are 3765 bp and 1398 bp.

Figure 12: Verification of pInt-Clp-GFP (JL120) activity. The plasmid is harbored within EC101. The exposure time is in seconds and is reported in the lower right corner.

Figure 13: Verification of pInt-Clp-GusA (JL119) activity. The plasmid is harbored within EC101. Enzymatic units are defined as $\mu\text{M pNP}^-/\text{min}$.

Cre recognizes two 34 bp site (LoxP) and catalyzes the recombination and removal of the contents between the LoxP sites. This system provides an opportunity to utilize the convenience of antibiotic selection, as well as ultimately producing a food grade product. Lox sites were added using the primers created to amplify the ErmC antibiotic marker and origin of replication from pGK12. With standard LoxP sites, an active LoxP site still exists within the genome after the recombination event occurs. In order to eliminate this factor, modified LoxP sites were used [41]. By modifying 5 nt on the 5' end of one LoxP (creating Lox66) site and the 3' end of the other LoxP site (creating Lox71), the LoxP site is inactivated (designated Lox72) after the recombination by the Cre enzyme. The modification of the Lox site is considered to be an important step as it eliminates a possible genomic recombination hot spot in the genome of *L. reuteri*. After the Lox sites we also engineered *PstI* sites on the ends of the ErmC-Origin amplicon. These cut sites allowed for the insertion of the antibiotic marker and RepA dependent origin into the of the pCR4 plasmid.

For this study we needed an expression system for Cre recombinase that would propagate itself in *L. reuteri*. Using pGK12 as a backbone ensured the system would propagate itself in *L. reuteri*. It was also important that the expression of the Cre protein was inducible. Other studies have used a nisin inducible system to induce control the expression of Cre [13]. The nisin inducible system

was not used for this study because it requires the nisin operon to be included into the plasmid system. Instead, the inducible *ErmC* regulatory region was chosen to control the expression of the Cre ORF. With our system cells that have undergone a genomic integration will be resistant to erythromycin. Therefore the CAT gene was required to select for the cre plasmid. Therefore, the *ErmC* promoter was chosen to control Cre expression to maintain an unaltered chloramphenicol resistance marker. The helper plasmid also contains the CAT gene, however it will have been cured from cells that contain genomic integrants before transforming in the Cre plasmid. Therefore, these two plasmids resistance markers will not interfere with one another because they will not be in the same cell.

5.1 Methods for Cre Expression Vector

5.1.1 Building Cre Expression Plasmid

The Cre gene was obtained from the plasmid pMB145 (lab stock). The gene was amplified using *Erm5'Dic* Cre-F and Cre-R with an expected amplicon of 1056 bp. The plasmid backbone was amplified from pGK12 with pGK5'Erm-R and *ErmDel-F2* yielding a band of 3908 bp. The two amplicons were then gel purified with the Zymo Gel Pure Kit™ (Zymo Research) and digested with *EcoRI*. The digests were ligated together to attach the 3' end of the Cre cassette to the plasmid backbone. The ligation was column cleaned using the Zymo Cleaner Concentrator Kit™. DIC was used to stitch the 5' end of the Cre gene to the plasmid backbone. The annealing step of the Cre DIC reaction was 40°C due to a small stretch of homology. The plasmid was then transformed into *E. coli* TURBO. Colony PCR was used to verify the construction of pGK5'Erm-Cre (JL1212) with Cre-R and *Lox-Ori-F2* to generate a 2.4 kb amplicon for positive colonies. (See Figure 14)

5.1.2 Activity Check of pGK5'Erm-Cre

The pGK 5'Erm-Cre plasmid was transformed into *E. coli* (dH5α T1R). The pUG6 plasmid [28] was transformed into the dH5α containing pGK5'Erm-Cre(JLE121) cells and selected for using both erythromycin and ampicillin(100 µg/ml). Colonies were picked and grown in liquid broths.

Plasmids were obtained using the ZymoResearch Plasmid Prep Kit and screened by diagnostic digests for positive transformants. To isolate pUG6 from pGK5'Erm-Cre, cells were plated on LB-agar plates with ampicillin and grown at 42°C for four successive generations. Plasmids were digested with *Nde*I, *Sac*I, *Xho*I, and/or *Hind*III and DNA banding patterns were analyzed using agarose gel electrophoresis to confirm Cre activity (Figure 15).

5.2 Results and Discussion of Inducible Cre Shuttle Vector

Early attempts to build the Cre expression plasmid involved ligating the Cre ORF into pGK12 under the control of the ErmC promoter. However, no successfully transformed cells were obtained after multiple attempts. It was suspected that the Cre cassette was being expressed from the ErmC locus at levels high enough to be toxic to the cells, as the expression of Cre has been reported as toxic to *E. coli* [47]. Mähönen et al. (2004) engineered an intron into the Cre ORF thus making the Cre enzyme inactive in bacteria. When expressed in eukaryotic cells, the intron was removed and Cre was active. This option was not available for this study as our product was not intended to be expressed in a eukaryotic background.

In an attempt to further reduce expression from the ErmC promoter, the mechanism of control was further investigated. The regulation of gene expression in ErmC is at the translational level. The mRNA of the ErmC contains two Shine-Dalgarno (SD) sequences that are separated by a 19 amino acid leader sequence [48]. When the ErmC gene is transcribed, the mRNA secondary structure forms a hairpin that leaves the first SD sequence open; however, the secondary SD (which starts translation of the methylase) sequence is inactivated by the mRNA secondary structure. Upon exposure to low levels of erythromycin, erythromycin bound ribosomes stalls translation on the leader peptide. The ribosomal stalling results in a change in mRNA secondary structure and exposes the second SD sequence. Ribosomes not bound to erythromycin then bind to the second SD sequence and translate the methyltransferase [9]. The mRNA secondary structure that creates the inducible properties of the ErmC gene requires the 5' region of the ErmC gene. It is possible our original construct disrupted the secondary structure of the mRNA so the inducible properties of ErmC were disrupted. Other groups have constructed fusion proteins with the ErmC gene and

Figure 14: Construction of pGK-5'ErmCre. Numbers above primers refer to Table 2. pMB145 is a lab stock provided by M. Black

maintained the native regulation [9],[48], [11], [20]. To ensure the Cre cassette did not interfere with the regulation of the *ErmC* gene, we inserted the Cre ORF after the 73rd amino acid encoded by *ErmC* [20]. The fusion protein enabled the successful creation of a plasmid bearing the Cre coding cassette (Figure 14). The plasmid construction was confirmed by digesting it with *NdeI* and *SacI*. The digest should yield a banding pattern of 3560 bp and 1379 bp. The digestion of pGK-5'ErmCre with *NdeI* and *SacI* yielded three bands approximately 4.8, 3.6, and 1.3 kb in size (Figure 15). The largest DNA fragment is assumed to be undigested pGK-5'ErmCre plasmid based on size and intensity. The smaller two fragments are close to the expected band sizes of pGK-5'ErmCre and indicate the plasmid is properly built.

To assess if functional Cre recombinase was produced from this plasmid, the plasmid was co-transformed with the pUG6 plasmid [28]. The pUG6 plasmid (Figure 15 middle lane) contains two antibiotic resistance markers, ampicillin and kanamycin. The kanamycin marker is flanked by *LoxP* sites which are the target of Cre recombinase. By co-transforming pUG6 with our pGK-5'ErmCre plasmid, we were able to determine if Cre was active *via* screens for the loss of kanamycin resistance followed by diagnostic digest verification. If active Cre recombinase was expressed from pGK 5'Erm-Cre the Cre enzyme would interact with pUG6 and recombine pUG6 and remove the kanamycin resistance marker. It was found that the kanamycin resistance marker was rapidly removed from pUG6 even before induction of Cre expression with erythromycin (Figure 15 far left lane). The pUG6 plasmid was isolated from Kan sensitive strains and digested with *HindIII* and *XhoI*. The digest of the parental pUG6 with these enzymes theoretically yields three bands 2484 bp, 1009 bp, and 516 bp. If the kan marker is removed from pUG6, the plasmid is linearized (2502 bp) when digested with *HindIII* and *XhoI*. The digest of pUG6 before cotransformation with pGK-5'Erm-Cre revealed three bands of approximately 2.5 kb, 1 kb, and 0.5 kb (Figure 15, second to last lane). After cotransforming pUG6 with pGK-5'Erm-Cre digests with *HindIII* and *XhoI* reveal a single band at ~2.5 kb (Figure 15 last lane). The loss of the kanamycin marker (1 and 0.5 kb bands) from the digest suggest the Cre recombinase expressed from pGK-5'Erm-Cre is functional in *E. coli*.

These data confirm suspicions that the *ErmC* expresses its transcripts at a low level regardless of erythromycin's presence. Another possible explanation is that the chloramphenicol (another

Figure 15: Construction and activity of pGK-5'ErmCre. Enzymes used to digest the plasmids are in parentheses. The pGK-ErmCre digest produced two bands with expected sizes of 1379 bp. and 3560 bp. The pUG6 linearized gives an expected band of 4009 bp. The pUG6 w/ kan digest has expected band sizes of 516 bp, 1009 bp, and 2484 bp. The pUG6 w/o kan should yield a single band of 2502 bp.

translational inhibitor) has some cross-induction of the ErmC locus. The cross induction is unlikely to express high levels of ErmC-controlled genes at high levels based upon previous studies that have used chloramphenicol as their negative control when examining ErmC induction [9].

6 Summary of Work

The end goal of the project is to create a probiotic strain of *L. reuteri* that will act as a persistent treatment for CD. The current treatment for CD is a life long gluten free diet, which is difficult to accomplish with western diets. This project aimed to create a system to efficiently integrate genes into the *L. reuteri* genome. The work contained in this thesis has generated a body of necessary tools to integrate useful genes, such as a prolyl endopeptidase, into the *L. reuteri* genome and maintain food grade status.

Previous researchers have used a similar dual plasmid system to drive genomic integration into LAB [60]. However, the papers did not mention the need for a strain of *E. coli* that expresses RepA from its genome. As a result, early investigations attempted to use the helper plasmids generated to propagate the integration plasmids in *E. coli*. To alleviate the possible plasmid incompatibility

encountered, from using pGK12 plasmid derivatives, RepA was expressed from plasmids with the ColE1 origin of replication (pCR2.1 and pCR4). Multiple attempts to transform the integration plasmids into strains of *E. coli* containing the non-pGK helper plasmids failed. It is possible ColE1 plasmids were not expressing RepA at high enough levels to properly support replication of the integration plasmids. It is also possible cells containing a plasmid with the ColE1 origin do not efficiently take up pGK derivative plasmids. To successfully propagate the integration plasmids we requested a strain of *E. coli* (EC101) from a research group in the Netherlands [42]. The Kok lab quickly responded and generously agreed to send us the strain of *E. coli*. After two months nothing had arrived and I contacted the grad student who sent me the strain. After some investigation we learned the grad student had promptly sent the strain, but upon arrival to the Los Angeles U.S. Customs Office the strain was destroyed. The Kok lab generously sent the strain again through another carrier, and we received the *E. coli* strain and we were able to propagate the integration plasmids.

Another problem that plagued this project was the transformation of *L. reuteri*. Researchers have been able to transform *L. reuteri* with high efficiency [3], however in our hands we have had extremely low transformation efficiencies. A potential problem is the autoclaves in at our facilities seemed to caramelize the media we used to grow the bacteria. The caramelization of the media is thought to greatly decrease the transformation efficiency in *L. reuteri* (personal communications with Matt Shurtluff). The problem was eventually solved by reducing the temperature of the autoclave below 121°C typically used to sterilize liquids and changing the media from MRS(Oxid) to LCM [3].

Initial plans to build the pGK-Cre expression vector involved using a simple replacement of the ErmC ORF with the Cre ORF. However, multiple attempts to ligate Cre into the pGK- Δ -Erm backbone failed. When the ErmC ORF was removed from the plasmid the inducible properties or the ErmC promoters were disrupted. Therefore, the ErmC promoter constitutively expressed the Cre enzyme. The high levels of Cre may have been toxic to *E. coli* [47] and prevented successful transformation. Only by creating a Cre fusion protein with the 5' end of the ErmC ORF could the pGK-Cre shuttle vector be created.

Though the project was full of unforeseen complications an integration system was developed which will be instrumental in creating a probiotic strain of bacteria to treat CD. This thesis has developed a plasmid based system to integrate genes into *L. reuteri*. Furthermore, the system has the ability to utilize the convenience of antibiotic selection during the gene integration process, then remove the resistance marker via Cre recombinase. With the antibiotic resistance removed, the probiotic bacteria is more suitable as a food grade product. The integration system also maintains a high degree of plasticity because the integration site can be changed *via* by removing the LacA site and ligating in a new target sequence. Also, the gene to be integrated to can be changed by cutting out the current gene(GusA or GFP) with *AscI* and ligating in a new gene. Though originally developed for uses in *L. reuteri* this plasmid system can be utilized to integrate genes into any bacterial species able to host plasmids with the pGK origin of replication. As a result, this system may be use to create other probiotic bacteria which could host enzymes such as lactase or α -galactosidase. The proteins responsible for generating immunity could be integrated into the bacterial genome and the bacteria could be used as a delivery mechanism for vaccines development.

7 Future directions

After confirming successful integration of our reporter genes, the integration plasmid should be re-examined to find a better integration site within the *L. reuteri* genome. It is possible interrupting LacA will result in a competitive disadvantage for the bacteria when colonizing the duodenum. If a competitive disadvantage is observed a different integration site should be used to ensure the bacteria can successfully colonize the intestine. Furthermore, a prolyl endopeptidase must be chosen to replace our reporter genes (GFP and GusA). Within our lab, we are currently working with a number of peptidases to determine which enzyme most effectively digests the gluten fragments known to trigger CD. This peptidase must also contain a leader sequence to direct it to the bacterial cell surface and an anchor sequence to keep the peptidase from being excreted into the intestinal lumen. Once these hurdles are overcome, the dairy department at Cal Poly should be consulted on how to best deliver the probiotic bacteria to consumers. They have extensive experience in working

with food grade microbes and likely will provide key insights for our probiotic delivery techniques. Once a delivery system for our probiotic is developed the probiotic should begin small scale tests to ensure safety and effectiveness of the probiotic to alleviate CD symptomologies.

References

- [1] Mariano Prado Acosta, María Mercedes Palomino, Mariana C. Allievi, Carmen Sanchez Rivas, and Sandra M. Ruzal. Murein hydrolase activity in the surface layer of *Lactobacillus acidophilus* ATCC 4356. *Appl Environ Microb*, 74(24):7824–7827, 2008.
- [2] M. R. Adams and P. Marteau. On the safety of lactic acid bacteria from food. *Int. J. Food Microbiol.*, 27:263–264, 1995.
- [3] Siv Ahrné, Gřan Molin, and Lars Axelsson. Transformation of *Lactobacillus reuteri* with electroporation: Studies on the erythromycin resistance plasmid pLUL631. *Curr Microbiol*, 24:199–205, 1992.
- [4] Richard A. Jefferson, Tony A. Kavanagh, and Michael W. Bevan. Gus fusions: β -glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO*, 6(13):3901–3907, 1987.
- [5] A. K. Akobeng, A. V. Ramanan, I. Buchan, and R. F. Heller. Effect of breast feeding on risk of coeliac disease: a systematic review and meta-analysis of observational studies. *Arch Dis Child*, 91:39–43, 2006.
- [6] Armin Alaedini and Peter H.R. Green. Narrative review: Celiac disease: Understanding a complex autoimmune disorder. *Ann Intern Med*, 142:289–298, 2005.
- [7] Anil Kumar Anal and Harjinder Singh. Recent advances in microencapsulation of probiotics for industrial applications and targeted delivery. *Trends Food Sci Tech*, 18:240–251, 2007.
- [8] Salvatore Auricchio, Daniela Follo, Giorgio de Ritis, Annamaria Guinta, Donata Marzorati, Luigia Prampolini, Nicoletta Ansaldi, Patrizia Levi, Domenico Dall’Olio, Anna Bossi, Ivan Cortinovis, and Ettore Marubini. Does breast feeding protect against the development of clinical symptoms of celiac disease in children? *J Pediatr Gasrt Nutr*, 2:428–433, 1983.
- [9] Marne Bailey, Tobin Chettiath, and Alexander S. Mankin. Induction of *erm*(C) expression by noninducing antibiotics. *Antimicrob Agents Ch*, 52(3):866–874, 2008.

- [10] Hein J. Boot, Carin P. A. M. Kolen, Johannes M. Van Noort, and Peter H. Pouwels. S-layer protein of *Lactobacillus acidophilus* atcc 4356: Purification, expression in *Escherichia coli*, and nucleotide sequence of the corresponding gene. *J Bacteriol*, 175(19):6089–6096, 1993.
- [11] F. Breidt and D. Dubnau. Identification of *cis*-acting sequences required for translational autoregulation of the *ermC* methylase. *J Bacteriol*, 172(7):3661–3668, 1990.
- [12] Bernd Bukanu. Regulation of the *Escherichia coli* heat-shock response. *Mol. Microbiol.*, 9(4):671–680, 1993.
- [13] Nathalie Campo, Marie-Line Daveran-Mingot, Kees Leenhouts, Paul Ritzenthaler, and Pascal Le Bourgeois. Cre-*LoxP* recombination system for large genome rearrangements in *Lactococcus lactis*. *Appl Environ Microb*, 68(5):2359–2367, 2002.
- [14] Carlo Catassi, Italo Bearzi, and Geoffrey K. T. Holmes. Association of celiac disease and intestinal lymphomas and other cancers. *Gastroenterology*, 128(4):S79–S86, 2005.
- [15] Carlo Catassi, Elisabetta Fabiani, Giuseppe Iacono, Cinzia D’Agate, Ruggiero Francavilla, Federico Biagi, Umberto Volta, Salvatore Accomando, Antonio Picarelli, Itali De Vitis, Giovanna Pianelli, Rosaria Gesuita, Flavia Carle, Alessandra Mandolesi, Italo Bearzi, and Alessio Fasano. A prospective, double-blind, placebo-controlled trial to establish a safe gluten threshold for patients with celiac disease. *An J Clin Nutr*, 85:160–166, 2007.
- [16] Nadine Cerf-Bensussan, Tamara Matysiak-Budnik, Christophe Cellier, and Martine Heyman. Oral proteases: a new approach to managing coeliac disease. *Gut*, 56:157–160, 2007.
- [17] Jun Chen, Seongbong Jo, and Kinam Park. Polysaccharide hydrogels for protein drug delivery. *Carbohydr Polym*, 28(1):69–76, 1995.
- [18] Willem M. de Vos. Safe and sustainable systems for food-grade fermentations by genetically modified lactic acid bacteria. *Int. Dairy J.*, 9:3–10, 1999.

- [19] Isabelle Derré, Georges Rapoport, and Tarek Msadek. CtsR, a novel regulator of stress and heat shock response, controls *clp* and molecular chaperone gene expression in Gram-positive bacteria. *Mol. Microbiol.*, 31(1):117–131, 1999.
- [20] David Dubnau. Induction of *ermC* requires translation of the leader peptide. *EMBO*, 4(2):533–537, 1985.
- [21] Alessio Fasano, Tarcisio Not, Wenle Wang, Sergio Uzzau, Irene Berti, Alberto Tommasini, and Simeon E Goldblum. Zonulin, a newly discovered modulator of intestinal permeability, and its expression in coeliac disease. *Lancet*, 355:1518–1519, 2000.
- [22] C. S. Fávaro-Trindade and C. R. F. Grosso. Microencapsulation of *L. acidophilus* (la-05) and *B. lactis* (bb-12) and evaluation of their survival at the pH values of the stomach and in bile. *J Microencapsul*, 19(4):485–494, 2002.
- [23] Giovanna E. Felis and Franco Dellaglio. Taxonomy of lactobacilli and bifidobacteria. *Curr. Issues Intestinal Microbiol*, 8:44–61, 2007.
- [24] Laszlo Fesus and Mauro Piacentini. Transglutaminase 2: an enigmatic enzyme with diverse functions. *Trends Biochem Sci*, 27(10):534–539, 2002.
- [25] G. Gardiner, C. Stanton, P. B. Lynch, J.K. Collins, G. Fitzgerald, and R.P. Ross. Evaluation of cheddar cheese as a food carrier for delivery of a probiotic strain to the gastrointestinal tract. *J. Dairy Sci*, 82:1379–1387, 1999.
- [26] Armond S. Goldman. Modulation of the gastrointestinal tract of infants by human milk. interfaces and interactions. An evolutionary perspective. *J. Nutr.*, 130:426S–431S, 2000.
- [27] Peter H. R. Green. The many faces of celiac disease: Clinical presentation of celiac disease in the adult population. *Gastroenterology*, 128(4):S74–S78, 2005.
- [28] U. Güldener, S. Heck, T. Fiedler, J. Beinhauer, and Hegemann J.H. A new efficient gene disruption cassette for repeated use in budding yeast. *Research*, 24:2519–2524, 1996.

- [29] Sharareh Hekmat and Donald J. McMahon. Survival of *Lactobacillus acidophilus* and *Bifidobacterium bifidum* in ice cream for use as a probiotic food. *J. Dairy Sci*, 75(6):1415–1422, 1992.
- [30] Mary Hickson, Aloysius L DSouza, Nirmala Muthu, Thomas R Rogers, Susan Want, Chakravarthi Rajkumar, and Christopher J Bulpitt. Use of probiotic lactobacillus preparation to prevent diarrhoea associated with antibiotics: randomised double blind placebo controlled trial. *BMJ*, 335(80), 2007.
- [31] L. Högberg, P. Laurin, K. Fälth-Magnusson, C. Grant, E. Grodzinsky, G. Jansson, H. Ascher L. Browaldh, J.A. Hammersjö, E. Lindberg, U. Myrdal, and L. Stenhammar. Oats to children with newly diagnosed coeliac disease: a randomized double blind study. *Gut*, 53(5):327–331, 2004.
- [32] K. Holm, M. Mäki, N Vuolteenaho, K. Mustalahti, M. Ashorn, T. Ruuska, and K. Kaukine. Oats in the treatment of childhood coeliac disease: a 2-year controlled trial and a long-term clinical follow-up study. *Aliment Pharmacol Ther*, 23:1463–1472, 2006.
- [33] Wilhelm H. Holzapfel, Petra Haberer, Johannes Snel, Ulrich Schillinger, and Jos H.J. Huis in’t Veld. Overview of gut flora and probiotics. *Int J Food Microbiol*, 41:85–101, 1998.
- [34] Lora V. Hooper, Melissa H. Wong, Anders Thelin, Lennart Hansson, Per G. Falk, and Jeffrey I. Gordon. Molecular analysis of commensal host-microbial relationships in the intestine. *Science*, 291:881–884, 2001.
- [35] Anneli Ivarsson. The swedish epidemic of coeliac disease explored using an epidemiological approach-some lessons to be learnt. *Best Pract Res Cl Ga*, 19(3):425–440, 2005.
- [36] E. K. Janatuinen, T. A. Kempainen, P. H. Pikkarainen, K. H. Holm, V-M Kosma, M. I. J. Uusitupa, M. Mäki, and R. J. K. Julkunen. Lack of cellular and humoral immunological responses to oats in adults with coeliac disease. *Gut*, 46:327–331, 2000.

- [37] Esko K. Janatuinen, Pekka. H. Pikkarainen, Tarja A. Kemppainen, Veli-Matti Kosma, Ritva M. K. Järvinen, Matti I.J. Uusitupa, and Risto J.K. Julkunen. A comparison of diets with and without oats in adults with celiac disease. *N Engl J Med*, 333(16):1033–1037, 1995.
- [38] Soo Jin Jang, Min Seok Ham, Jung Min Lee, Sung Kyun Chung, Hyong Joo Lee, Jeong Hwan Kim, Hae Choon Chang, Jong-Hoon Lee, and Dae Kyun Chung. New integration vector using a cellulase gene as a screening marker for *Lactobacillus*. *FEMS Microbiol Lett*, 224:191–195, 2003.
- [39] Martin F. Kangoff. Overview and pathogenesis of celiac disease. *Gastroenterology*, 128(4):S10–S18, 2005.
- [40] Jan Kok, Jos M. B. M. Van Der Vossen, and Gerard Venema. Construction of plasmid cloning vectors for lactic streptococci which also replicate in *Bacillus subtilis* and *Escherichia coli*. *Appl Environ Microb*, 48(4):726–731, 1984.
- [41] Jolanda M. Lambert, Roger S. Bongers, and Michiel Kleerebezem. Cre-*lox*-based system for multiple gene deletions and selectable-marker removal in *Lactobacillus plantarum*. *Appl Environ Microb*, 73(4):1126–1135, 2006.
- [42] Jean Law, Girbe Buist, Alfred Haandrikman, Jan Kok, Gerard Venema, and Kees Leenhouts. A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. *J Bacteriol*, 177(25):7011–7018, 1995.
- [43] Kees J. Leenhouts, Jan Kok, and Gerard Venema. Lactococcal plasmid pWV01 as an integration vector for lactococci. *Appl Environ Microb*, 57(9):2562–2567, 1991.
- [44] Laurence Leloup, S. Dusko Ehrlich, Monique Zagorec, and Françoise Morel-Deville. Single-crossover integration in the *Lactobacillus sake* chromosome and insertional inactivation of the *pstI* and *lacL* genes. *Appl Environ Microb*, 63(6):2117–2123, 1997.

- [45] S. Lohi, K. Mustalahti, K. Kaukinen, K. Laurila, P. Collin, H. Rissanen, O. Lohi, M. Gasparin E. Bravi, A. Reunanen, and M. Mäki. Increasing prevalence of coeliac disease over time. *Aliment Pharmacol Ther*, 26:1217–1225, 2007.
- [46] K. E. Lundin, E. M. Nilsen, H. G. Scott, E. M. Løberg, A. Gjøen, J. Bratlie, V. Skar, E. Mendez, A. Løvik, and K. Kett. Oats induced villous atrophy in coeliac disease. *Gut*, 52:1649–1652, 2003.
- [47] Anssi J. Mähönen, Kari J. Airenne, Miia M. Lind, Hanna P. Lesch, and Seppo Ylä-Herttuala. Optimized self-excising cre-expression cassette for mammalian cells. *Biochem Biophys Res Commun*, 320:336–371, 2004.
- [48] Mark Mayford and Bernard Weisblum. Conformational alteration in the ermC transcript *in vivo* during induction. *EMBO J*, 8(13):4307–4313, 1989.
- [49] Andrea McCracken, Mark S. Turner, Phil Giffard, Louise M. Hafner, and Peter Timms. Analysis of promoter sequences from *Lactobacillus* and *Lactococcus* and their activity in several *Lactobacillus* species. *Arch Microbiol*, 173:383–389, 2000.
- [50] Annick Mercenier, H. Müller-Alouf, and Corinne Grangette. Lactic acid bacteria as live vaccines. *Curr. Issues Mol. Biol.*, 2(1):17–25, 2000.
- [51] L. Mincheva-Nilsson, M. L. Hammarström, P. Juto, and S. Hammarström. Human milk contains proteins that stimulate and suppress T lymphocyte proliferation. *Clin. Exp. Immunol*, 79:463–469, 1990.
- [52] Øyvind Molberg, Stephen N. Mcadam, Roman Körner, Hanne Quarsten, Christel Kristiansen, Lars Madsen, Lars Fugger, Helge Scott, Ove Norén, Peter Roepstorff, Knut E.A. Lundin, Hans Sjöström, and Ludvig M. Sollid. Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived t cells in celiac disease. *Nat. Med.*, 4(6):713–717, 1998.
- [53] Allan M. Mowat. Anatomical basis of tolerance and immunity to intestinal antigens. *Nat rev Immunol*, 3:331–341, 2003.

- [54] Arthur C. Ouwehand and Seppo J. Salminen. The health effects of cultured milk products with viable and non-viable bacteria. *Int. Dairy Journal*, 8:749–758, 1998.
- [55] Natalia Periolo and Alejandra C. Cherañvsky. Coeliac disease. *Autoimmun Rev*, 5:202–208, 2005.
- [56] Ulrike Peters, Sebastian Schneeweiss, Elke A. Trautwein, and Helmut F. Erbersdobler. A case-control study of the effect of infant feeding on celiac disease. *Ann Nutr Metab*, 45:135–142, 2001.
- [57] Christ Platteeuw, Guus Simons, and Willem M. De Vos. Use of the *Escherichia coli* β -glucuronidase (gus)A gene as a reporter gene for analyzing promoters in lactic acid bacteria. *Appl Environ Microb*, 60(2):587–593, 1994.
- [58] Presutti RJ, Cangemi JR, Cassidy HD, and Hill DA. Celiac disease. *Am Fam Physician*, 76(15):1795–1802, 2007.
- [59] Ida Rud, Peter Ruhdal Jensen, Kristine Naterstad, and Lars Axelsson. A synthetic promoter library for constitutive gene expression in *Lactobacillus plantarum*. *Microbiol*, 152:1011–1019, 2006.
- [60] W.M. Russell and T.R. Klaenhammer. Efficient system for directed integration into the *Lactobacillus acidophilus* and *Lactobacillus gasseri* chromosomes via homologous recombination. *Appl Environ Microb*, 67(9):4361–4364, 2001.
- [61] Jose M. Saavedra, Nancy A. Bauman and I Oung, J A Perman, and Yolken R H. Feeding of *Bifidobacterium bifidum* and *Streptococcus thermophilus* to infants in hospital for prevention of diarrhoea and shedding of rotovirus. *Lancet*, 344(8929):1046, 1994.
- [62] Jos F. M.L. Seegers, Adam C. Zhao, Wilfried J. J. Meijer, Saleem A. Khan, Gerard Venema, and Sierd Bron. Structural and functional analysis of the single-strand origin of the replication from the lactococcal plasmid pWV01. *Mol Gen Genet*, 249:43–50, 1995.

- [63] Nagendra P. Shah, Lankaputhra E. V. Warnakulasuriya, Margaret L. Britz, and Kyle S. A. William. Survival of *Lactobacillus acidophilus* and *Bifidobacterium bifidum* in commercial yoghurt during refrigerated storage. *J. Dairy Sci*, 5:512–521, 1995.
- [64] Aino-Vieno Shornikova, Ivan A. Casas, Erika Isolauri, Hannu Mykkänen, and Timo Vesikari. *Lactobacillus reuteri* as a therapeutic agent in acute diarrhea in young children. *J Pediatr Gastr Nutr*, 24(4):399–404, 1997.
- [65] Matthew J. Shurtleff. Designing a Microbial Prolyl Peptidase Delivery System for the Treatment of Celiac Disease. Master’s thesis, California Polytechnic State University, San Luis Obispo, May 2009.
- [66] Matthew Siegel, Michael T. Bethune, Jonathan Gass, Jennifer Ehren, Jiang Xia, Alexandre Johannsen, Tor B. Stuge, Gary M. Gray, Peter P. Lee, and Chaitan Khosla. Rational design of combination enzyme therapy for celiac sprue. *Chemistry and Biology*, 13:649–658, 2006.
- [67] Christopher W. Simmons. Transient co-expression of post-transcriptional gene silencing and β -glucuronidase in harvested lettuce leaf tissue does not improve recombinant protein accumulation *in planta*. *Biotechnol Lett*, 29:641–645, 2007.
- [68] Uwe B. Sleytr. I. basic and applied s-layer research: An overview. *FEMS Microbiol Rev*, 20:5–12, 1997.
- [69] Egbert Smit and Peter H. Pouwels. One repeat of the cell wall binding domain is sufficient for anchoring the *Lactobacillus acidophilus* surface layer protein. *J Bacteriol*, 184(16):4617–4619, 2002.
- [70] Ludvig M. Sollid. Molecular basis of celiac disease. *Annu. Rev. Immunol.*, 18:53–81, 2000.
- [71] Ludvig M. Sollid, Øyvind Molberg, Stephen N. McAdam, and Knut E.A. Lundin. Autoantibodies in coeliac disease:tissue transglutaminase-guilt by association? *Gut*, 119:851–852, 2000.

- [72] Dariusz Stepniak, Liesbeth Spaenij-Dekking, Cristina Mitea, Martine Moester, Arnoud de Ru, Renee Baak-Pablo, Peter van Veelen, Luppó Edens, and Frits Koning. Highly efficient gluten degradation with a newly identified prolyl endoprotease: implications for celiac disease. *Am J Physiol Gastrointest Liver Physiol*, 291:G621G629, 2006.
- [73] Willemijn L. Vader, Arnoud de Ru, Yvonne van der Wal, Yvonne M. C. Kooy, Willemien Benckhuijsen, M. Luisa Mearin, Jan Wouter Drijfhout, Peter van Veelen, and Frits Koning. Specificity of tissue transglutaminase explains cereal toxicity in celiac disease. *J. Exp. Med*, 195(5):643–649, 2002.
- [74] Nana Valeur, Peter Engel, Noris Carbajal, Eamonn Connolly, and Karin Ladefoged. Colonization and immunomodulation by *Lactobacillus reuteri* ATCC 55730 in the human gastrointestinal tract. *Appl Environ Microb*, 70(2):1176–1181, 2004.
- [75] E. A. M. Verderio, T. Johnson, and M. Griffin. Tissue transglutaminase in normal and abnormal wound healing. *Amino Acids*, 26:387–404, 2004.
- [76] B.W. Wolf, K.B. Wheeler, D.G. Ataya, and K. A. Garleb. Safety and tolerance of *Lactobacillus reuteri* supplementation to a population infected with the human immunodeficiency virus. *Food Chem Toxicol*, 36:1085–1094, 1998.
- [77] Jiang Xia, Ludvig M. Sollid, and Chaitan Khosla. Equilibrium and kinetic analysis of the unusual binding behavior of a highly immunogenic gluten peptide to hla-dq2. *Biochemistry-US*, 4:4442–4449, 2005.