

Utah State University

DigitalCommons@USU

Undergraduate Honors Capstone Projects

Honors Program

5-2010

Genetic and Biochemical Studies of Plasmid pIR52-1 in *Lactobacillus helveticus*

Cody Alexander Tramp
Utah State University

Follow this and additional works at: <https://digitalcommons.usu.edu/honors>



Part of the [Cell Biology Commons](#), and the [Molecular Biology Commons](#)

Recommended Citation

Tramp, Cody Alexander, "Genetic and Biochemical Studies of Plasmid pIR52-1 in *Lactobacillus helveticus*" (2010). *Undergraduate Honors Capstone Projects*. 61.

<https://digitalcommons.usu.edu/honors/61>

This Thesis is brought to you for free and open access by the Honors Program at DigitalCommons@USU. It has been accepted for inclusion in Undergraduate Honors Capstone Projects by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



**GENETIC AND BIOCHEMICAL STUDIES OF PLASMID
pIR52-1 In *LACTOBACILLUS HELVETICUS***

by

Cody Alexander Tramp

**Thesis submitted in partial fulfillment
of the requirements for the degree**

of

**HONORS IN UNIVERSITY STUDIES
WITH DEPARTMENTAL HONORS**

in

**Biology (Cellular/Molecular Emphasis)
in the Department of Biology**

Approved:

Thesis/Project Advisor
Dr. Dennis Welker

Thesis/Project Advisor
Dr. Gregory Podgorski

Departmental Honors Advisor
Dr. Kim Sullivan

Director of Honors Program
Dr. Christie Fox

**UTAH STATE UNIVERSITY
Logan, UT**

Spring 2010

1. Introduction

Lactobacillus helveticus is a species of lactic acid bacteria. These bacteria, which produce lactic acid as a major product of carbohydrate metabolism, are used industrially to produce cheese and other fermented dairy products (Ebringer et al., 2008). Many species of lactic acid bacteria also possess probiotic characteristics and when ingested potentially confer increased immune function, regulate gut microbiota, and improve digestion in the host (Reid, 2008). Genetic studies of these probiotic effects and other characteristics of *L. helveticus* and related species have been hindered by the lack of stably replicating plasmid vectors. Plasmid vectors for several *Lactobacillus* species have been constructed, but these are often derived from broad host range plasmids and may not be stably maintained over many generations (Fang et al., 2008; Thompson et al., 2001). In *L. helveticus* specifically, a small number of attempts have been made at adapting plasmid vectors for use in genetic studies, but as the source plasmids were not native to *L. helveticus*, they were either unable to replicate extrachromosomally or be maintained over many generations (Hagen et al., 2010; Thompson et al., 2001).

This study utilizes the cryptic plasmid pIR52-1 from *L. helveticus* strain R0052 as the basis for novel plasmid vectors capable of long-term stable replication in *Lactobacillus* species. Plasmid pIR52-1 has been sequenced and annotated by Hagen et al. (2010). This effort identified several regions that may serve as the plasmid origin of replication and be involved in plasmid maintenance, including a truncated *repA_N* replication gene (*repA'*), a full-length *repA_N* homolog, a *repB* homolog, several conserved homologous nonrepeating regions (HN), and several repeat regions. The full-length *repA_N* gene in pIR52-1 is homologous to genes encoding replication proteins in several other plasmids from lactic acid bacteria. In these plasmids the homologous replication proteins bind to repeat regions within the origin of

replication and then recruit proteins that replicate the plasmid (del Solar et al., 1998; Ito et al., 2009; Fang et al., 2008; Kranenburg et al., 2005; Kwong et al., 2004; Weaver et al., 2009). The pIR52-1 *repA_N* gene contains several repeat regions, suggesting that it may also function as the origin of replication for this plasmid (Hagen et al., 2010; Weaver et al., 2009). In *L. planatarum*, a plasmid vector containing only the *repA_N* gene was capable of replication, but the addition of regions flanking the *repA_N* gene increased the stability of replication (Fang et al., 2008). The RepA', RepA_N, and RepB proteins may also aid in the partitioning of the plasmid to the daughter cells, which is necessary to ensure stable, long-term maintenance (Weaver et al., 2009). Due to the findings of these researchers, this study focuses on the *repA_N* gene and its repeat regions as likely candidates for the minimum origin of replication, but will also test the effects of flanking regions on the stability of long-term maintenance.

In addition to the localization of the origin of replication and the regions contributing to stable maintenance of the plasmid, this study seeks to identify the binding sites of the RepA_N and RepB proteins. We will attempt to verify that the RepA_N protein binds to sequence repeats within the *repA_N* gene on pIR52-1. Additional sites that the RepA_N or RepB proteins could bind to include the HN2 and HN3 conserved homologous nonrepeating regions, and the start codon – ribosome binding site – putative mRNA hairpin loop region found at the 5' end of the *repA_N* gene (Hagen et al., 2010). The identification of the binding regions will be conducted using an electrophoresis mobility shift assay (Garner and Revzin, 1981). This assay utilizes the fact that a DNA fragment's progress through an agarose or polyacrylamide gel will be retarded by the binding of a protein to the fragment. Recently, advances in protein and DNA dyes have allowed the detection of binding interactions through staining for both the DNA and protein. These tools have been utilized in other studies defining interactions of replication proteins with

their DNA binding sites (Jing et al., 2003; Jing et al., 2004) and potentially allow larger DNA fragments to be screened for binding, as the shift in the protein band can also be assessed if the DNA fragment shift is not easily detected.

2. Materials and Methods

2.1 Bacterial strains and growth conditions

L. helveticus R0052 and CNRZ32 (a plasmid-free strain) were propagated at 37 °C in MRS broth or on MRS agar. *Escherichia coli* DH5 α and BL21 (DE3) were propagated at 37 °C in LB broth or on LB agar plates or in Formedium lactose autoinduction media (LAIM) without trace elements. Glycerol stocks for all species were stored at -80 °C.

2.2 Subcloning of pIR52-1 regions

In order to test regions of the pIR52-1 plasmid for the origin of replication, the plasmid vector pAE1 was constructed (Figure 1). The Genbank accession numbers for pIR52-1 and pAE1 are FJ851149 and GQ894741, respectively. The pAE1 vector contains the origin of replication and the ampicillin resistance gene from the *E. coli* pUC19 plasmid and the erythromycin resistance gene from the pG+host9 plasmid. These fragments were generated by PCR using the PFUSION™ High-Fidelity DNA Polymerase from New England Biolabs (Table 1). The ampicillin resistance gene was flanked upstream by *Bam*HI and downstream by *Hind*III.

Fragments of the pIR52-1 plasmid were generated by PCR using the PFUSION™ High-Fidelity DNA Polymerase (Table 1, Table 2, and Figure 2). The upstream primers incorporated *Bam*HI and *Sal*I sites near the 5' primer ends and the downstream primers incorporated a *Hind*III site near the 5' primer ends. The amplified pIR52-1 PCR products were double digested with *Bam*HI and *Hind*III and ligated into the large *Bam*HI and *Hind*III digested pAE1 vector fragment so that the cloned pIR52-1 region replaced the ampicillin resistance gene of pAE1 and maintained both the *Bam*HI and *Sal*I sites at the 5' end of the insert.

The test plasmids were then transformed into *E. coli* DH5a cells using a Bio-Rad Gene Pulser and protocols supplied by the manufacturer (Bio-Rad Laboratories, Richmond, CA). Transformants were screened for erythromycin resistance on LB agar plates containing 100 µg/mL erythromycin. Resistant transformants were further screened for ampicillin sensitivity on LB agar plates containing 50 µg/mL ampicillin. Plasmid DNA extractions were performed on the erythromycin resistant and ampicillin sensitive clones to verify correct plasmid vector size. Representative plasmids of correct size were then sequenced to verify that the pIR52-1 regions did not contain inadvertent mutations; all eight constructs were correct. The plasmids were labeled pTW1 through pTW8. In the pTW1-pTW8 plasmids, the *E. coli* origin was flanked by *Sal*I sites, which allowed its excision as a control for the pUC19 origin function in *L. helveticus*. The excision of the pUC19 origin by *Sal*I digestion and re-ligation of the larger fragment to itself created an additional eight plasmids labeled pTW9, pTW10, pTW13-15, and pTW16-18 (corresponding to pTW4, pTW5, pTW1-3, and pTW6-8, respectively). These vectors replicate in *L. helveticus* but not in *E. coli*.

The sixteen plasmids, along with pAE1 lacking a pIR52-1 region were electroporated into CNRZ32 cells. Transformants were screened on MRS agar plates containing 1 µg/mL

erythromycin. PCR amplification with primers specific to pIR52-1 regions was used to confirm presence of plasmid and correct insert size. Extrachromosomal replication of the plasmids was confirmed by the presence of plasmid bands of correct size in agarose gels of DNA extracts from the CNRZ32 transformant colonies. Additional verification was obtained from successful re-transformation of *E. coli* DH5a cells with DNA extracts from the pTW1-8 CNRZ32 colonies.

Plasmid maintenance was studied in the CNRZ32 colonies. Three independent transformants for each plasmid were grown overnight in MRS broth containing 1 µg/mL erythromycin. Cultures were passaged for 20 consecutive subculturings using 10 µL of the previous subculture to inoculate 1 mL MRS broth lacking any antibiotic. Cultures were incubated at 37 °C without shaking in 1.5 mL microcentrifuge tubes. Cells from the twentieth subculturing were recovered by centrifugation and were re-suspended in 1 mL sterile distilled water. Cell densities were standardized using a 1:10 dilution of cells to an O.D.₆₀₀ of 0.50 ± 0.02 . Further dilutions were made, and 100 µL samples of the $1:10^{-4}$, $1:10^{-5}$, and $1:10^{-6}$ dilutions were plated on four replicate MRS agar plates without antibiotic and four replicate MRS agar plates containing 1 µg/mL erythromycin. After three days of incubation, the number of colonies per plate was counted for the dilution plate containing between 30 and 300 colonies, and colony counts values were converted to CFU/mL. Plasmid retention was calculated by averaging the number of colonies on the four replicate plates with erythromycin, dividing by the average number of colonies on the four replicate plates without antibiotic, and multiplying by 100%. The plasmid retentions for the three independent transformants of each plasmid were then averaged together and the standard deviations calculated.

2.3 Electrophoresis mobility shift assays

The *repA_N* and *repB* regions from pIR52-1 were amplified by PCR (Table 1). The primers upstream of the genes incorporated an *NdeI* site at the 5' end. The primers downstream of the genes incorporated a 6-histidine tag sequence in front of the stop codons, and a *HindIII* site for the *repA_N* primer and a *BamHI* site for the *repB* primer. Double restriction digests with *NdeI* and *HindIII* for *repA_N* and *NdeI* and *BamHI* for *repB* were conducted, followed by ligation into the pET22b *E. coli* expression vector downstream of the T7 promoter. The vectors were labeled pTW19 and pTW20, for the vectors containing *repA_N* and *repB*, respectively. As a control, a pET22b derivative vector, pBL1, containing a gene from *Brevibacterium linens* fused to a 6-His tag was used. Two additional plasmids were constructed to verify that the 6-histidine tags attached to the RepA_N and RepB proteins did not interfere with their functions *in vivo*. Plasmid pTW11 contained only the 6-histidine tagged *RepA_N* region. Plasmid pTW12 contained the 6-histidine tagged *RepA_N*, 6-histidine tagged *RepB*, and intergenic region. Both regions were ligated into a pAE1 vector backbone after restriction digests with *BamHI* and *HindIII*. Five independent clones of each were passaged for 20 consecutive subculturing. Plasmid retention rates and standard deviations were calculated as described in the above section.

The pTW19 and pTW20 vectors were transformed into *E. coli* BL21 (DE3) cells by electroporation. Transformants were screened for ampicillin resistance on LB agar plates containing 50 µg/mL ampicillin. Plasmid presence was verified by correct plasmid band size in DNA extracts run on agarose gels. Additional PCR verification of inserted regions and ampicillin resistance genes was performed (Table 1). Sequencing of inserted genes verified sequences and that 6-His tag was in frame.

Protein expression was induced using either IPTG induction or lactose autoinduction using procedures provided by the pET22b System Manual 9th edition (Novagen) and from Formedium, respectively. For IPTG induction, cells were grown overnight in LB broth containing 50 µg/mL ampicillin from glycerol stocks in a shaker at 37 °C and 180 rpm. Dilutions of 1:10, 1:20, 1:30, and 1:40 were made and O.D.₆₀₀ measured to determine cell density. Based on the cell density measurements 50 mL of LB broth was inoculated at an O.D.₆₀₀ of 0.10 ± 0.01. Cultures were grown for 2 hours at 37 °C and 180 rpm, reaching O.D.₆₀₀ of 0.5 to 0.6. Induction of protein was initiated by addition of 100 mM IPTG stock solution to final concentrations of 0.4 mM, 2 mM, or 4 mM, depending on the experiment. Cultures were incubated for 1, 2, or 3 hours, depending on the experiment. Uninduced control cultures were incubated similarly except for the addition of IPTG. Cultures were divided into 12 mL aliquots and cells collected by centrifugation. For lactose autoinduction, cultures were grown in lactose autoinduction media (LAIM) with 50 µg/mL ampicillin for 12 or 24 hours. Uninduced controls were grown in LB media with 50 µg/mL ampicillin for the same periods. Cultures were divided into 12 mL aliquots and cells collected by centrifugation.

Proteins were extracted using a procedure modified from the pET System Manual 9th Edition (Novagen). Cell pellets were subjected to osmotic shock by resuspension in 0.6 M sucrose, 30 mM Tris-HCl pH 8.0, and 1 mM EDTA solution. Suspensions were incubated on ice in 1.5 mL microcentrifuge tubes for ten minutes, mixing every two minutes by gentle inversion. Cells were pelleted by centrifugation for 30 seconds in a microcentrifuge at 10,000 rpm, and resuspended in 30 mM Tris-HCl pH 8.0 and 1 mM EDTA solution. Cells were incubated on ice for ten minutes, mixing every two minutes by gentle inversion. After the ten minute incubation the cells were again pelleted by centrifugation for 5 minutes at 10,000 rpm. The cell pellets were

washed once by resuspension in 1 mL LEW buffer (50 mM sodium phosphate monobasic, 300 mM sodium chloride, pH 7.0) followed by centrifugation for 5 minutes at 10,000 rpm. After pelleting, the cells were resuspended in 1200 μ L of fresh LEW with 30 mM Tris-HCl, pH 7.0 and then lysed. Cells were lysed by addition of 20 μ L lysozyme stock solution (50 mg/mL) and 10 μ L DNaseI stock solution (500 μ g/mL). The lysate was incubated for 30 minutes at 30 $^{\circ}$ C, mixing by gentle inversion every 10 minutes. Soluble phase proteins were prepared by centrifugation for 10 minutes at 10,000 rpm and collection of the supernatant. The insoluble phase proteins were collected by washing the pellet twice with 750 μ L LEW buffer and resuspending the washed pellet in 1 mL 1% SDS solution.

Total cell protein samples were generated from the cells recovered from 1 mL of induced and uninduced cultures after pelleting for 5 minutes at 10,000 rpm. The cell pellets were resuspended in 100 μ L 50 mM sodium phosphate, pH 7.0 and the cells lysed by addition of 100 μ L 2x TCP Sample Buffer, 2 μ L lysozyme stock solution and 1.5 μ L DNaseI stock solution. The 2x TCP Sample Buffer contains 100 mM DTT, 2% SDS, 80 mM Tris-HCl, 15% glycerol, and 0.006% Bromophenol Blue. Samples were incubated for 30 minutes at 30 $^{\circ}$ C, mixing by gentle inversion every 10 minutes.

Soluble phase protein samples were further purified using either a PrepEase $\text{\textcircled{C}}$ Histidine-Tagged Protein Purification Kit – High Specificity (USB) or a column containing BDTM TALON Metal Affinity Resin (BD Biosciences Clontech), using the protocols provided by the manufacturers. Samples were eluted from the columns with varying concentrations of imidazole. For some experiments samples purified through columns were concentrated using ammonium sulfate precipitation (SAFC Biosciences). Protein samples were denatured by heating to 70 $^{\circ}$ C for 3 minutes in a thermocycler before running on 6% Novex $\text{\textcircled{R}}$ Tris-Glycine gels, using the SDS

PAGE settings provided by the manufacturer (Invitrogen). Protein samples were stained with SYPRO® Ruby EMSA stain using the protocols provided by the manufacturer (Invitrogen). Gels were photographed on a (Find out what system the Broadbent lab has) using a SYPRO ® photographic filter.

DNA fragments for use in the EMSA were amplified by PCR using the primers listed in Table 1. Features of the amplified regions are detailed in Table 3 and Figure 3. Fragments were purified by gel electrophoresis, dialysis of gel slices, followed by chloroform extraction and ethanol precipitation of the eluted DNA. DNA concentrations were determined using a Nanodrop spectrophotometer (Thermo Scientific).

Before loading EMSA samples, samples were prepared by mixing 200 ng DNA and 10 µg of each appropriate protein preparation, after which the proteins were allowed to bind to the DNA fragment during a 20 minute incubation at 37 °C. Samples were run on 6% Novex ® Tris-Glycine gels, using the settings provided by the manufacturer (Invitrogen). DNA was stained using SYBR ® Green EMSA stain using protocols provided by the manufacturer (Invitrogen).

3. Results

3.1 Localization of the pIR52-1 origin of replication to the *repA_N* gene and its promoter region.

The pAE1 control vector lacking a pIR52-1 insert was not capable of replicating in CNRZ32 cells, as pAE1 plasmid DNA was not recovered in DNA extracted from numerous putative transformant clones. The pTW1-8 vectors that carried a pIR52-1 region were all capable of extrachromosomal replication in CNRZ32 transformants as shown by the recovery of circular

plasmid DNA corresponding to each vector from transformant colonies. Therefore the pIR52-1 origin of replication must lie within the portion of pIR52-1 shared by these vectors, which corresponds to the *repA_N* gene and a portion of the intergenic region between *repA_N* and *repB* that includes the *repA_N* promoter.

3.2 Retention of Plasmids in *L. helveticus*

Retention rates of the pTW1-8 vectors after 20 consecutive subculturings are shown in Figure 4 and in Table 4. Clones lacking the *E. coli* origin of replication showed similar rates of retention to their corresponding pTW1-8 vectors, with the exceptions of pTW14 and pTW16 which had higher rates of retention than pTW2 or pTW6 respectively. This indicates that the *E. coli* origin was not functioning as an origin of replication in *L. helveticus*, although the additional size of the plasmid due to its inclusion in pTW2 and pTW6 may have contributed to lower retention of these vectors. The *repA'* gene and the conserved homologous non-repeating region HN1 do not seem to contribute significantly to stability, as retention rates of pTW1-4 (including HN1) and pTW5-8 (lacking HN1) did not differ significantly. Similarly, the *repB* region similarly did not significantly affect retention of the plasmid in *L. helveticus*, but additional studies in *L. casei* indicate that it is required for high frequency transformation in that species (Hagen et al., 2010). The inclusion of the intergenic region between *repA_N* and *repB*, which includes the HN3 region, increased the retention rate of the plasmid by up to 10 fold.

3.3 Retention of 6-histidine tagged *repA_N* and *repB* containing plasmids in *L. helveticus*

Retention rates of the pTW11 (containing tagged *repA_N*) and pTW12 (containing tagged *repA_N* and tagged *repB*) are shown in Table 4. These plasmids contained the *E. coli* origin of replication, and had retention rates within one standard deviation of the retention rates for the pTW5 and pTW8 vectors, to which they were identical except for the addition of the 6-histidine tags on the Rep A_N and RepB proteins. The tags therefore do not appear to affect the proteins' contributions to replication and stable maintenance of the plasmid.

3.4 Protein Induction and Purification

Both IPTG and lactose autoinduction yielded similar total cell protein (TCP) profiles (data not shown). The pBL1 control gene clearly showed increased expression of an approximately 50 kDa band (Figure 5), consistent with the predicted molecular weight of 50.2 kDa for this protein. Bands of approximately 40 kDa were observed in some induced pET22b only controls and the pGL1 controls (as well as in induced RepA_N and RepB preparations). The origin of these bands is unclear. They do not correspond to the expected 99 kDa size of T7 RNA polymerase which should also be induced in these experiments using BL21 DE3 cells. Intriguingly, the lactose repressor protein, which is expressed from a pET22b encoded gene in these cells, has a size of 38.6 kDa; perhaps the approximately 40 kDa bands correspond to the lactose repressor protein. RepA_N should be 43.3 kDa and RepB should be 21.7 kDa. It is unclear whether the ~40 kDa band in the induced RepA_N TCP preparations is in fact the RepA_N protein, as the band also appears in the control lanes (Figure 6). RepB was also not readily apparent in the induced TCP preparations (Figure 6) perhaps due to low expression levels and the induced protein bands overlapping in size with other cellular *E. coli* protein bands. The

pBL1 control protein and RepA_N were mostly in the insoluble phase, although modification of the extraction procedure shifted some of the protein bands to the soluble phase (Figure 7). The distribution of RepB was not tested. Protein samples purified on columns showed similar band patterns to the soluble phase samples, if any sample was eluted at all (Figure 8). Column purification procedures need to be optimized further to yield specific binding and elution of the desired protein.

3.5 EMSA

A preliminary EMSA was conducted with the soluble phase lysates from induced RepA_N and RepB expressing cells. Since this test used the soluble phase lysates of these cells instead of purified protein, soluble phase lysate from induced cells containing an empty pET22b vector was used as a control for binding reactions between the other soluble proteins and the DNA (Figure 9). Band shifts (white arrows) were observed with the RepA_N and RepB preparations with region B (containing the R2 and R3 repeat regions and much of the center of *repA_N*). These bands were not present in the DNA only control lanes. The results with region C (containing the HN3 region, RBS, and 5' end of *repA_N*) are less clear. The DNA only control for the region C fragment appeared to be composed of two initial DNA bands (orange dashed arrow). However, the lanes with the RepA_N and RepB lysates contained this additional initial band as well as the shifted bands. The empty pET22b vector control lane appeared to have a continuous smear above the main DNA band suggesting that other proteins in the lysate or the DNaseI used during lysis were binding to the DNA. Further purification of the PCR amplified

regions and the RepA_N and RepB lysates will need to be performed in order to obtain more clear EMSA results.

4. Discussion

A plasmid vector that can be used in *L. helveticus* requires an origin of replication that is able to ensure the stable replication and maintenance of the vector over multiple generations. The analysis pIR52-1 plasmid regions for the location of its replication origin showed that the *repA_N* gene and its upstream promoter region were sufficient to allow extrachromosomal replication of vectors in *L. helveticus* CNRZ32 cells. However, this region alone does not allow stable, long-term replication, as only 7% of the cells retained the plasmid after 20 subculturing. The regions necessary for stable maintenance in *L. helveticus* include *repA_N* and the entire region between the *repA_N* and *repB* genes. Together, these regions increase plasmid retention more than 10-fold to approximately 100% after 20 rounds of subculturing. For high frequency transformation of *L. casei*, the *repB* gene was also found to be necessary, and its inclusion does not negatively affect the stability of plasmid replication in *L. helveticus*. Thus the shuttle vectors pTW7 and pTW8 constructed in this work should be useful in further genetic studies of *L. helveticus* and related Lactobacillus species.

Plasmids with a pAE1 backbone containing 6-His tagged *repA_N* (pTW11) or both 6-His tagged *repA_N* and 6-His tagged *repB* genes (pTW12) were constructed. These plasmids were shown to have similar retention rates to the pTW5 and pTW8 plasmids, to which they were identical except for the addition of the 6-histidine tags. This demonstrated that the addition of the tags did not interfere with the ability of the proteins to replicate and maintain the plasmid *in vivo*. The *repA_N* and *repB* genes were then successfully cloned into pET22b expression vectors and

sequenced to verify that they were inserted in-frame, had correct 6-His tags, and had no mutations. The total cell protein (TCP) profiles of the induced cells showed increased expression of a ~40 kDa band, indicating induction did occur (Figures 5 and 6). However, the ~40 kDa induced band appeared in the induced pET22b and pBL1 control TCP profiles, as well as the RepB induced TCP profile, suggesting that it is another protein responding to the induction. Since the samples were denatured and treated with SDS, it is unlikely that the band in RepB TCP lanes is a dimer of the 21.7 kDa RepB protein. The induced band also does not correspond to the expected 99 kDa size of T7 RNA polymerase which should be induced in these experiments using BL21 DE3 cells. The ~40 kDa band could potentially be the lactose repressor protein, which is expressed from a pET22b encoded gene in these cells, and which has a size of 38.6 kDa. Purified protein samples eluted from the affinity columns would help to identify which bands in the TCP profiles correspond to the RepA_N and RepB proteins. However, these purified protein samples have not yet been obtained.

One of the major factors complicating purification is the presence of much of the induced proteins in the insoluble phase. Modification of the lysis procedure to include an osmotic shock treatment increased the protein in the soluble phase from almost nothing to a detectable amount, but large amounts of the protein still remain in the insoluble phase (Figure 7). A different approach to the purification protocol might be needed, and different cellular expression systems may need to be explored in order to obtain highly soluble protein. Denaturing the insoluble protein and re-folding may be another approach to yield purified protein for EMSA analysis. Using a different protein expression cell line may also increase the yield, as the protein bands that were induced were approximately the same intensity as other cellular proteins, whereas in most expression studies the induced bands are significantly more intense. This reduced

expression may be due to the structure of the *repA_N* gene, which we hypothesize to contain the binding site for the RepA_N protein. As intracellular levels of RepA_N rise, the proteins might bind to the site within *repA_N* more frequently, inhibiting mRNA transcription, and preventing the accumulation of higher levels of RepA_N in the cell.

Another major factor complicating purification is the low yield of protein from the affinity columns. The column purification procedures will need to be optimized to extract only specifically-binding 6-His tagged proteins. So far, the elutions from the BD ® TALON Resin and PrepEase ® columns appear to contain all of the same bands present in the soluble phase lysate solution, which was the solution added to the columns for purification, just at a reduced intensity (Figure 8). This implies that the column is still non-specifically binding other cellular proteins. This is still occurring despite washes at many different imidazole concentrations intended to remove these non-specifically binding proteins, while leaving the 6-His tagged protein bound to the column. Additionally, procedures to concentrate the proteins eluted off the column need to be optimized, allowing the eluted proteins to be clearly identified on a gel, unlike the blurred concentrated lane in Figure 8.

The preliminary test of the EMSA using whole cell protein samples instead of purified protein samples also did not provide conclusive results (Figure 9). The shifted DNA bands present in the RepA_N and RepB sample lanes also appear in the smear present in the pET22b vector only control lanes. This suggests that the binding interactions detected by the gel are due to other proteins present in the cell or in the DNaseI enzyme preparation used to digest genomic DNA during protein extraction. The gel verifies that the EMSA assay can detect shifts in DNA fragments of approximately 500 bp, which are much larger than the fragments used in other

studies (Jing et al., 2003), suggesting that useful data could be recorded once the purified protein sample is obtained.

Overall, this study has yielded some successes, some promising results, and some clear failures. The origin of replication and regions contributing to the stability of replication of the pIR52-1 plasmid were identified. The pTW7 and pTW8 shuttle vectors that are capable of replication in *E. coli*, *L. helveticus*, and other related *Lactobacillus* species were created. The RepA_N and RepB coding regions were successfully cloned into an expression vector system, and the vector was successfully transformed into the *E. coli* BL21 (DE3) expression system. The 6-histidine tagged RepA_N protein retained its normal function, allowing stable replication of its plasmid in *L. helveticus*. Induction of gene expression was observed in this system, although the induced proteins might not be the proteins of interest. The expression and purification methodologies still need to be optimized for increased soluble protein retrieval. An EMSA gel using 500 bp DNA fragments verified that the assay may work with large regions; however, the crude protein lysate used does not indicate if the binding interactions observed were due to the proteins of interest or other proteins present in the samples. The successful construction of the plasmid shuttle vectors pTW7 and pTW8 and the progress toward analysis of the replication protein binding sites mark the major accomplishments of this study.

5. Acknowledgements

This project was funded in by an Undergraduate Research and Creative Opportunities (URCO) Grant, a Utah State University College of Science Research Minigrant, the Department of Biology at Utah State University, a grant from Institut Rosell, and a Utah State University Center for Integrated Biosystems uCIBR Grant. We would also like to thank Dr. Joanne E.

Hughes, Dr. Gregory Podgorski, Dr. Sean Johnson, Dr. Joseph K.-K. Li, and Craig Brighton for their advice and assistance during this project.

6. References

- del Solar G, Giraldo R, Ruiz-Echevarría MJ, Espinosa M, and Díaz-Orejas R. 1998. Replication and control of circular bacterial plasmids. *Microbiology and Molecular Biology Reviews* : MMBR 62: 434-464.
- Ebringer L, Ferencík M, and Krajcovic J. 2008. Beneficial health effects of milk and fermented dairy products--review. *Folia Microbiologica* 53: 378-394.
- Fang F, Flynn S, Li Y, Claesson MJ, van Pijkeren J, Collins JK, van Sinderen D, and O'Toole PW. 2008. Characterization of endogenous plasmids from *Lactobacillus salivarius* UCC118. *Applied and Environmental Microbiology* 74: 3216-3228.
- Garner MM, and Revzin A. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Research* 9: 3047-3060.
- Hagen KE, Tramp CA, Altermann E, Welker DL, and Tompkins TA. 2010. Sequence analysis of plasmid pIR52-1 from *Lactobacillus helveticus* R0052 and investigation of its origin of replication. *Plasmid* 63: 108-117.
- Ito Y, Kawai Y, Arakawa K, Honme Y, Sasaki T, and Saito T. 2009. A Conjugative Plasmid from *Lactobacillus gasseri* LA39 That Carries Genes for the Circular Bacteriocin Gassericin A-Production and Immunity. *Applied and Environmental Microbiology* 75: 6340-6351.

- Jing D, Agnew J, Patton WF, Hendrickson J, and Beechem JM. 2003. A sensitive two-color electrophoretic mobility shift assay for detecting both nucleic acids and protein in gels. *Proteomics* 3: 1172-1180.
- Jing D, Beechem JM, and Patton WF. 2004. The utility of a two-color fluorescence electrophoretic mobility shift assay procedure for the analysis of DNA replication complexes. *Electrophoresis* 25: 2439-2446.
- Kranenburg RV, Bongers R, Leer RJ, Vos WM, Siezen RJ, and Kleerebezem M. 2005. Functional Analysis of Three Plasmids from *Lactobacillus plantarum*. *Applied and Environmental Microbiology* 71: 1223-1230.
- Kwong SM, Skurray RA, and Firth N. 2004. *Staphylococcus aureus* multiresistance plasmid pSK41: analysis of the replication region, initiator protein binding and antisense RNA regulation. *Molecular Microbiology* 51: 497-509.
- Reid G. 2008. Probiotics and prebiotics – Progress and challenges. *International Dairy Journal* 18: 969-975.
- Thompson JK, McConville KJ, Nicholson C, and Collins MA. 2001. DNA cloning in *Lactobacillus helveticus* by the exconjugation of recombinant mob-containing plasmid constructs from strains of transformable lactic acid bacteria. *Plasmid* 46: 188-201.
- Weaver KE, Kwong SM, Firth N, and Francia MV. 2009. The RepA N replicons of Gram-positive bacteria: A family of broadly distributed but narrow host range plasmids. *Plasmid* 61: 94-109.

7. Tables

Table 1. Primers used for investigation of pIR52-1 replication, pAE1 construction, *repA_N* and *repB* cloning for expression, and investigation of RepA_N and RepB binding sites.

Primer Name	Purpose	Sequence (5'-3')	Target
866	Construction	AAA AGT CGA CCA AAA TCC CTT AAC GTG AGT T	Site 5' of pUC19 ori
867	and sequence	AAA AGG ATC CAA AGA ACA TGT GAG CAA AAG GCC A	Site 3' of pUC19 ori
868	verification of	AAA AGG ATC CGG AAA TGT GCG CGG AAC CCC TAT T	Site 5' of pUC19 ampicillin resistance gene promoter
869	pAE1 vector	AAA AAA GCT TAT GAG ATT ATC AAA AAG GAT CTT C	Site 3' of pUC19 ampicillin resistance gene stop codon
870		AAA AAA GCT TAG ATC TCT GCA GGG TAC CGG AAA CGT AAA AGA AGT TAT GGA	Site 5' of pG+host9 erythromycin resistance gene promoter
871		AAA AGT CGA CGG GAC CTC TTT AGC TCC TTG GAA G	Site 3' of pG+host9 erythromycin resistance gene stop codon
872	Cloning,	AAA AAA GCT TCC CTG GAG ACT GCT AAG CGT ATT TTT AAG	Site 3' of <i>repA'</i> stop codon
873	sequencing,	AAA AAA GCT TTT TAA GAT AAT ACC AAT TCA GCC ACT CAA	Site 3' of <i>repA'</i> stop codon
875	and verifying	CCC GGG ATC CGT CGA CTT TGT AAT TCA GTT AGA GCA TAT C	Site 5' of <i>repA_N</i> promoter
876	pIR52-1 inserts	CCC GGG ATC CGT CGA CAG ATG GAC TAA GTA GCC AAC GGC TAT T	Site in region between <i>repA_N</i> promoter and start codon of <i>repB</i>
877		CCC GGG ATC CGT CGA CTG TTG CAC TGA AAG CGG TGT GCT AAA C	Site in region between <i>repA_N</i> promoter and start codon of <i>repB</i>
878		CCC GGG ATC CGT CGA CAC TAG TAA ATG CTT GTT TCC AAA CTC TCC	Site 3' of <i>repB</i> stop codon
879		TGG TTG AGT ACT TTT TCA CTC	Site 3' of pIR52-1 region insertion site in pAE1
880		GTT ACT TAT GAG CAA GTA TTG	Site 3' of erythromycin resistance gene in pAE1
881		TCA GGG GGG CGG AGC CTA TGG	Site 5' of pIR52-1 region insertion site in pAE1
890	Cloning and	AAA AAG CTT TTA GTG GTG GTG GTG GTG ACT CCA AGT ATG CTC	Site 3' of <i>repA_N</i> stop codon with added 6-histidine tag
891	verifying <i>repA_N</i>	CCA GGA TCC TTA GTG GTG GTG GTG GTG CCA AAA ATG CCA CCA	Site 3' of <i>repB</i> stop codon with added 6-histidine tag
892	and <i>repB</i>	GAT ATA CAT ATG GAT CAA ACA AAT TTC AAG	Site 5' of <i>repA_N</i> promoter
893	sequences in	GAT ATA CAT ATG ACA TTG GCA AAA ACA ATT AAA C	Site 5' of <i>repB</i> promoter
001	Cloning regions	GAT AAT ACC AAT TCA GCC AC	Site 3' of <i>repA_N</i> stop codon past HN2
002	of stably	CTG ATG AAG ACA GTG GCT TAC	Site 5' of R1
003	replicating origin	GTA CGG CAC CAA AAG CTA AG	Site 5' of R1
004	for EMSA	GCT TTA TCA AAA AGC CAT G	Site 5' of R2 and R3 repeat region
005		CCT GTT TTG GGA TTA AAA TG	Site 5' of R2 and R3 repeat region
006		GTA GCC AAC GGC TAT TTA GTC	Site immediately 5' of HN3 and <i>repA_N</i> start codon
007		CAA ATC TGA TAT GCT CTA ACT G	Site in middle of HN3
008		CAA TGT CAA TCA CCA CCT AG	Site 5' of <i>repB</i> promoter
009		CTT TAA GTT CTT TAT CTA GAT CGT G	Site immediately 3' of R2 and R3 repeat region
010		GTG AAT ATG GCC AAA AAG AG	Site immediately 5' of R2 and R3 repeat region
011		CCA TAA ACG TTG TCA GCT TG	Site immediately 3' of HN3 and <i>repA_N</i> start codon

Table 2. Description of amplified regions containing pIR52-1 minimum origin of replication.

Clones pTW1-8 and pTW11-12 contain the pUC19 *E. coli* origin of replication, the remaining clones do not.

Clone Name	5' Primer	3' Primer	Insert Size (kbp)	Description
pTW1, pTW13	872	875	1.75	<i>repA'</i> , HN2, and <i>repA_N</i>
pTW2, pTW14		876	1.87	<i>repA'</i> , HN2, <i>repA_N</i> , and Hn3
pTW3, pTW15		877	2.13	<i>repA'</i> , HN2, <i>repA_N</i> , Hn3, and the remaining <i>repA_N-repB</i> intergenic region
pTW4, pTW9		878	2.88	<i>repA'</i> , HN2, <i>repA_N</i> , Hn3, the remaining <i>repA_N-repB</i> intergenic region, and <i>repB</i>
pTW5, pTW10	873	875	1.29	<i>repA_N</i>
pTW6, pTW16		876	1.41	<i>repA_N</i> and HN3
pTW7, pTW17		877	1.67	<i>repA_N</i> , Hn3, and the remaining <i>repA_N-repB</i> intergenic region
pTW8, pTW18		878	2.42	<i>repA_N</i> , Hn3, the remaining <i>repA_N-repB</i> intergenic region, and <i>repB</i>
pTW11	875	890	1.18	6-histidine tagged <i>repA_N</i>
pTW12	891	890	2.23	6-histidine tagged <i>repA_N</i> , <i>repA_N-repB</i> intergenic region, 6-histidine tagged <i>repB</i>

Table 3. Description of amplified regions for RepA_N and RepB EMSA.

Region	5' Primer	3' Primer	Region Size	Description
A	001	002	499	HN2, R1, and 3' end of <i>repA_N</i>
B	003	004	506	R2 (x2), R3 (x3), and center of <i>repA_N</i>
C	005	006	555	5' end of <i>repA_N</i> , RBS, start codon, HN3
D	007	008	517	half of HN3 and remaining <i>repA_N-repB</i> intergenic region
E	009	010	202	R2 (x2), R3 (x3)
F	011	006	273	RBS, start codon of <i>repA_N</i> , HN3

Table 4. Retention rates of pIR52-1 origin of replication containing plasmids. Retention rates are the average of three independent clones containing each plasmid. Retention rates are shown \pm one standard deviation.

Vector containing <i>E. coli</i> origin	Plasmid retention after 20 consecutive subculturings (%)	Vector without <i>E. coli</i> origin	Plasmid retention after 20 consecutive subculturings (%)
pTW1	6.9 \pm 3.5	pTW13	4.4 \pm 3.6
pTW2	56.9 \pm 17.4	pTW14	89.5 \pm 12.9
pTW3	105.4 \pm 14.3	pTW15	108.7 \pm 5.2
pTW4	101.7 \pm 13.3	pTW9	94.4 \pm 9.9
pTW5	10.9 \pm 7.5	pTW10	23.0 \pm 12.7
pTW6	55.0 \pm 15.2	pTW16	96.5 \pm 9.1
pTW7	81.2 \pm 3.6	pTW17	96.6 \pm 6.4
pTW8	90.3 \pm 8.8	pTW18	88.0 \pm 1.6
pTW11	21.1 \pm 16.8		
pTW12	101.3 \pm 9.9		

8. Figures

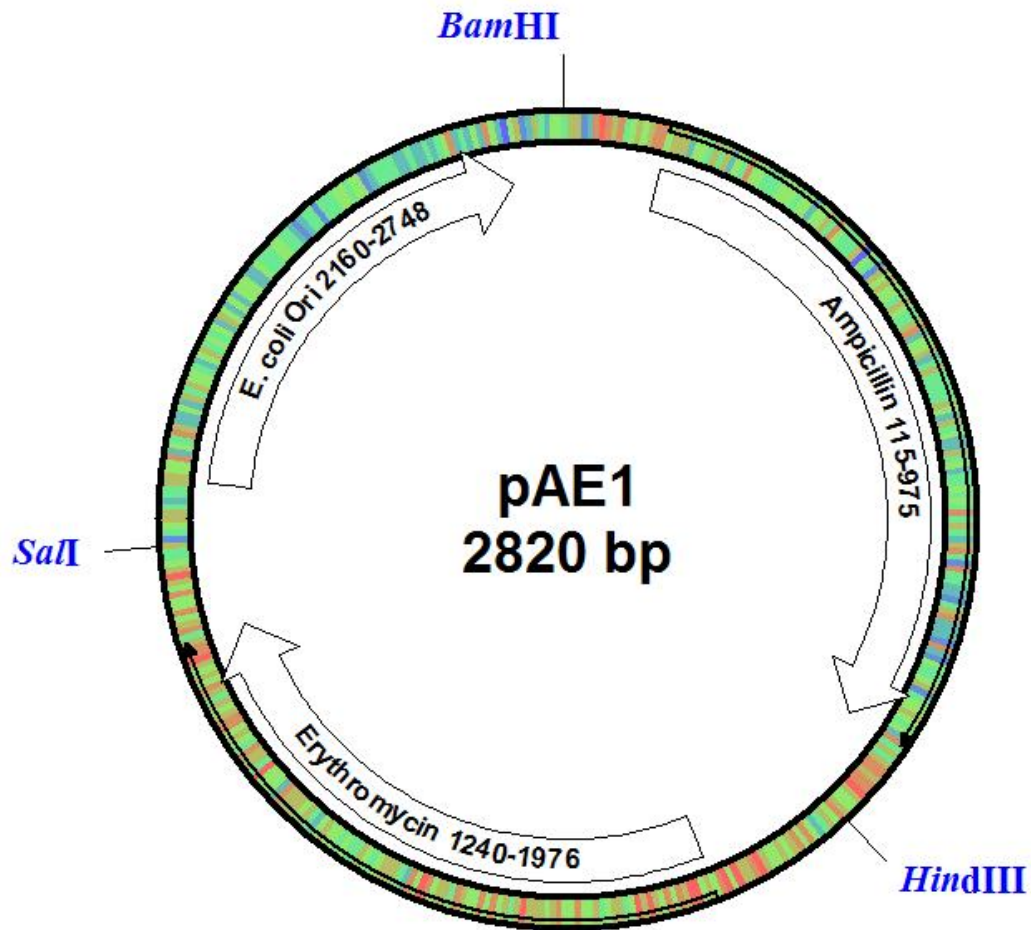


Figure 1. pAE1. Only restriction sites used in cloning are shown. *Bam*HI, *Sal*I, and *Hind*III sites are unique.

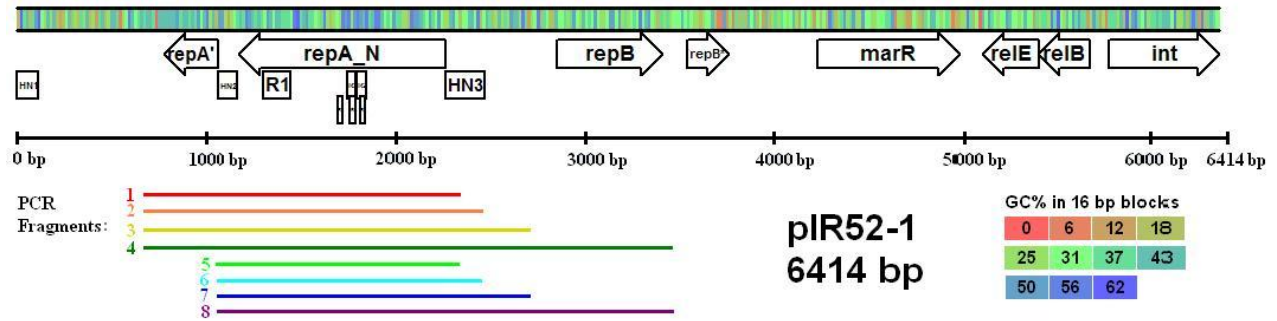


Figure 2. Map of pIR52-1 plasmid with PCR fragments. The circular plasmid is shown linearized for ease of reading. PCR Fragments are shown as colored lines below the map. Lines are color coded to match the bars for plasmid retention rate in Figure 4.

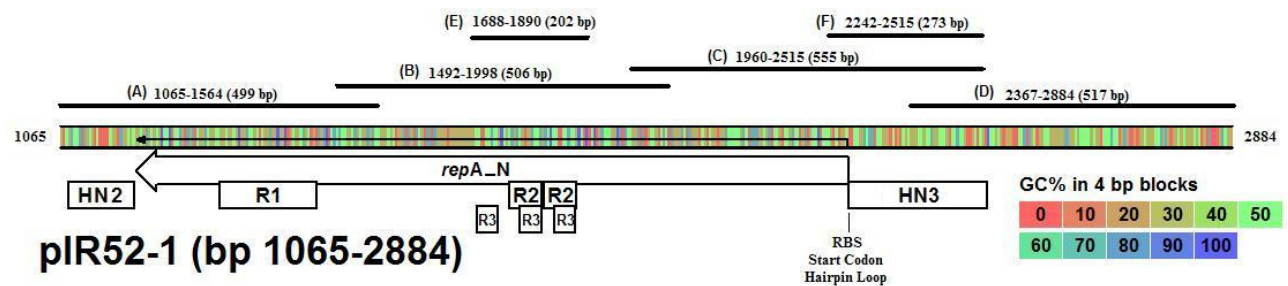


Figure 3. PCR amplified fragments of the stably replicating region from pIR52-1 that were used for EMSA testing of RepA_N and RepB. HN = homologous nonrepeating region, R = repeat region.

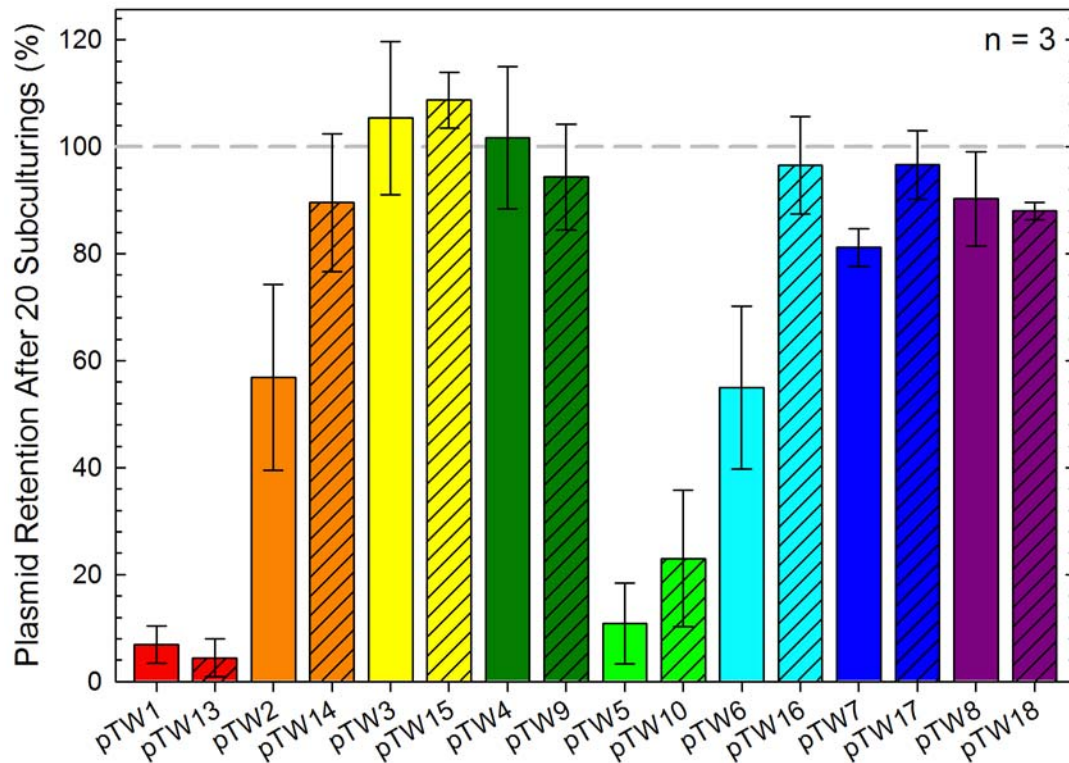


Figure 4. Plasmid retention rates of pIR52-1 origin of replication-containing plasmids. Bars are color coded to match the PCR amplified regions shown in Figure 2. Shaded bars represent plasmids lacking the *E. coli* origin of replication. Retention rates are the average of three independent clones containing each plasmid. Error bars represent \pm one standard deviation.

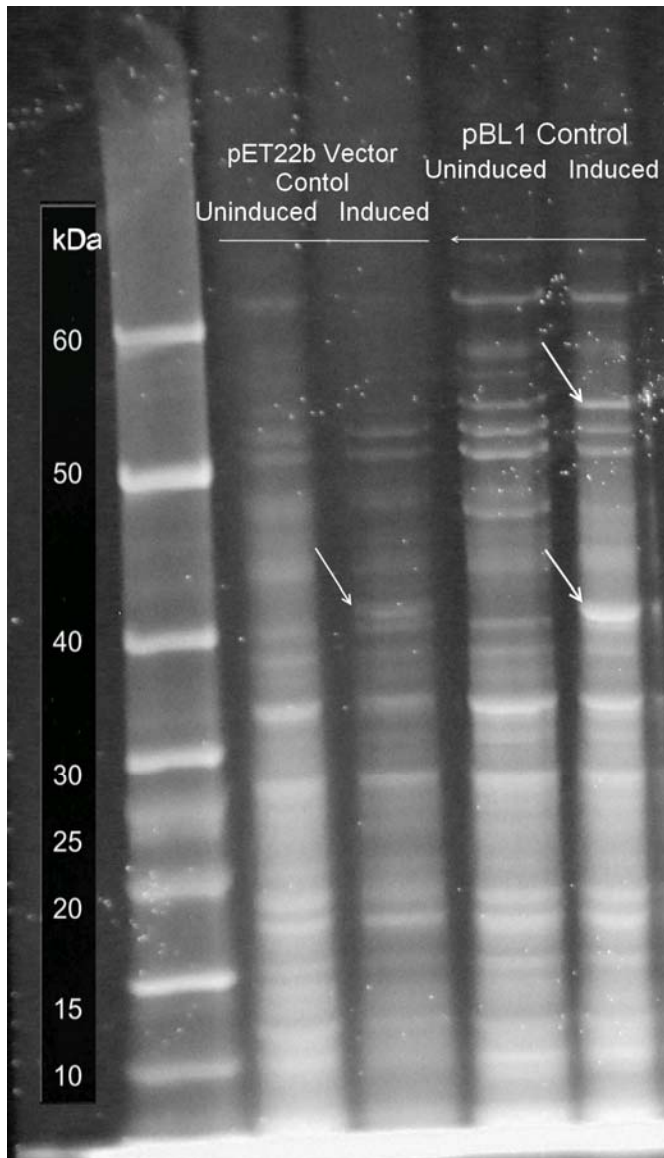


Figure 5. pBL1 control total cell protein (TCP) gel. Cell line transformed with empty pET22b vector used as control. Cells were induced with 0.4 mM IPTG for 3 hours. White arrows indicate protein bands with increased expression after IPTG induction. pBL1 control protein's molecular weight is 50.2 kDa.

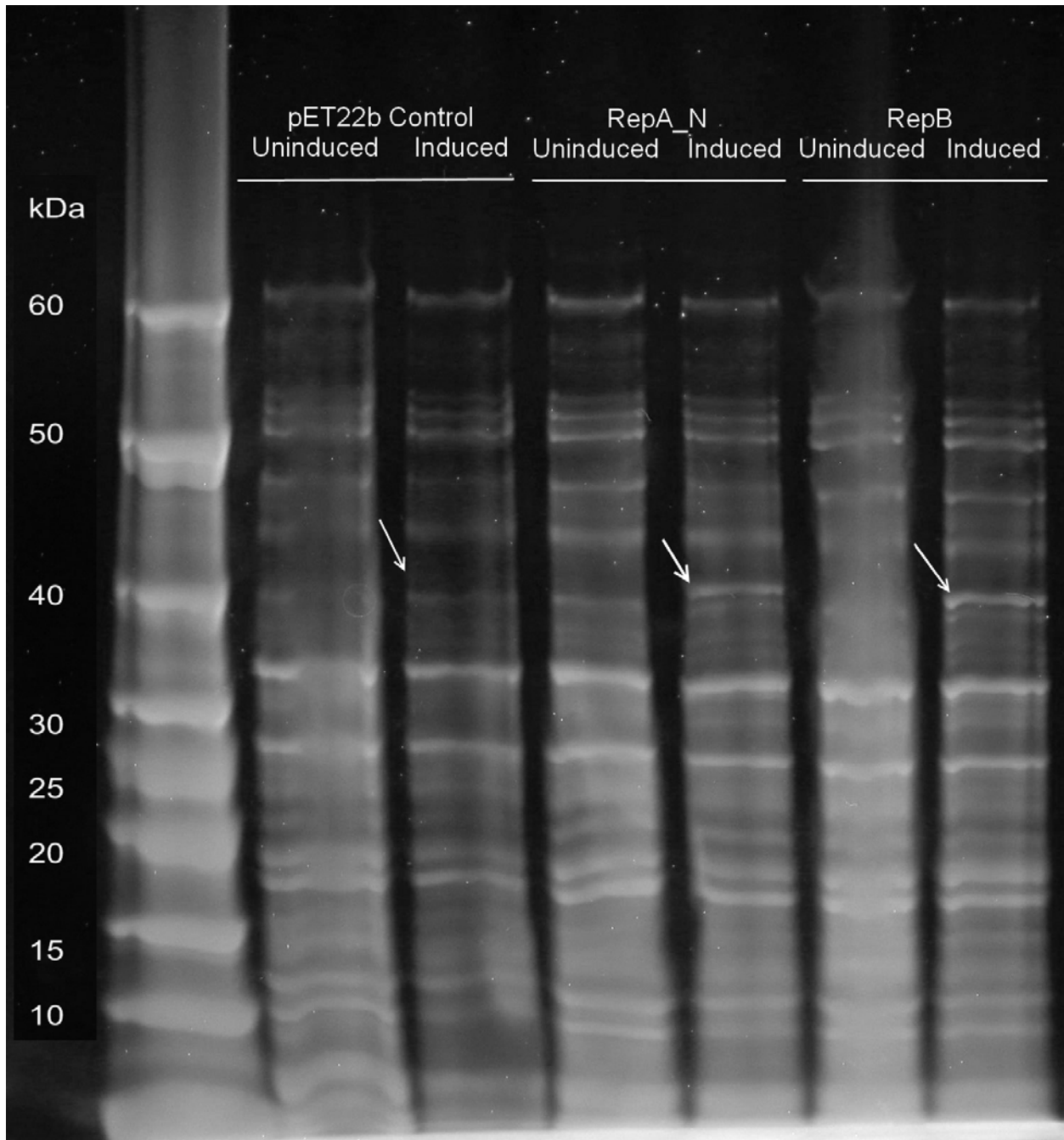


Figure 6. RepA_N and RepB total cell protein gel. The first two lanes after the marker are controls of exacts from cells transformed with empty pET22b vector. Cells were induced with 0.4 mM IPTG for 3 hours. White arrows indicate bands with increased expression after induction. RepA_N molecular weight is 43.3 kDa and RepB molecular weight is 21.7 kDa.

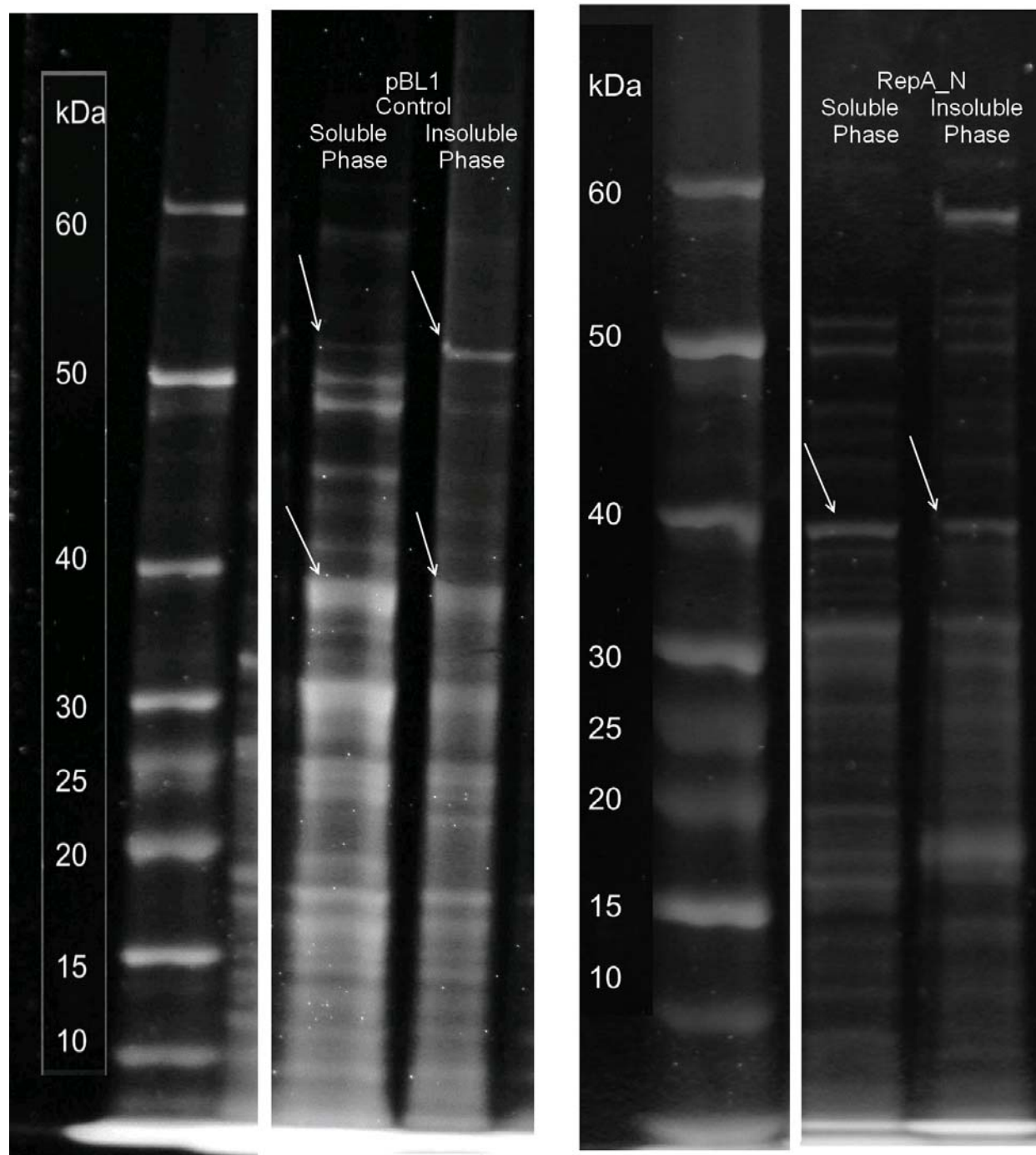


Figure 7. Soluble and insoluble proteins fractions of cells containing pBL1 control and RepA_N expression plasmids. White arrows indicate bands that showed increased expression with IPTG induction.

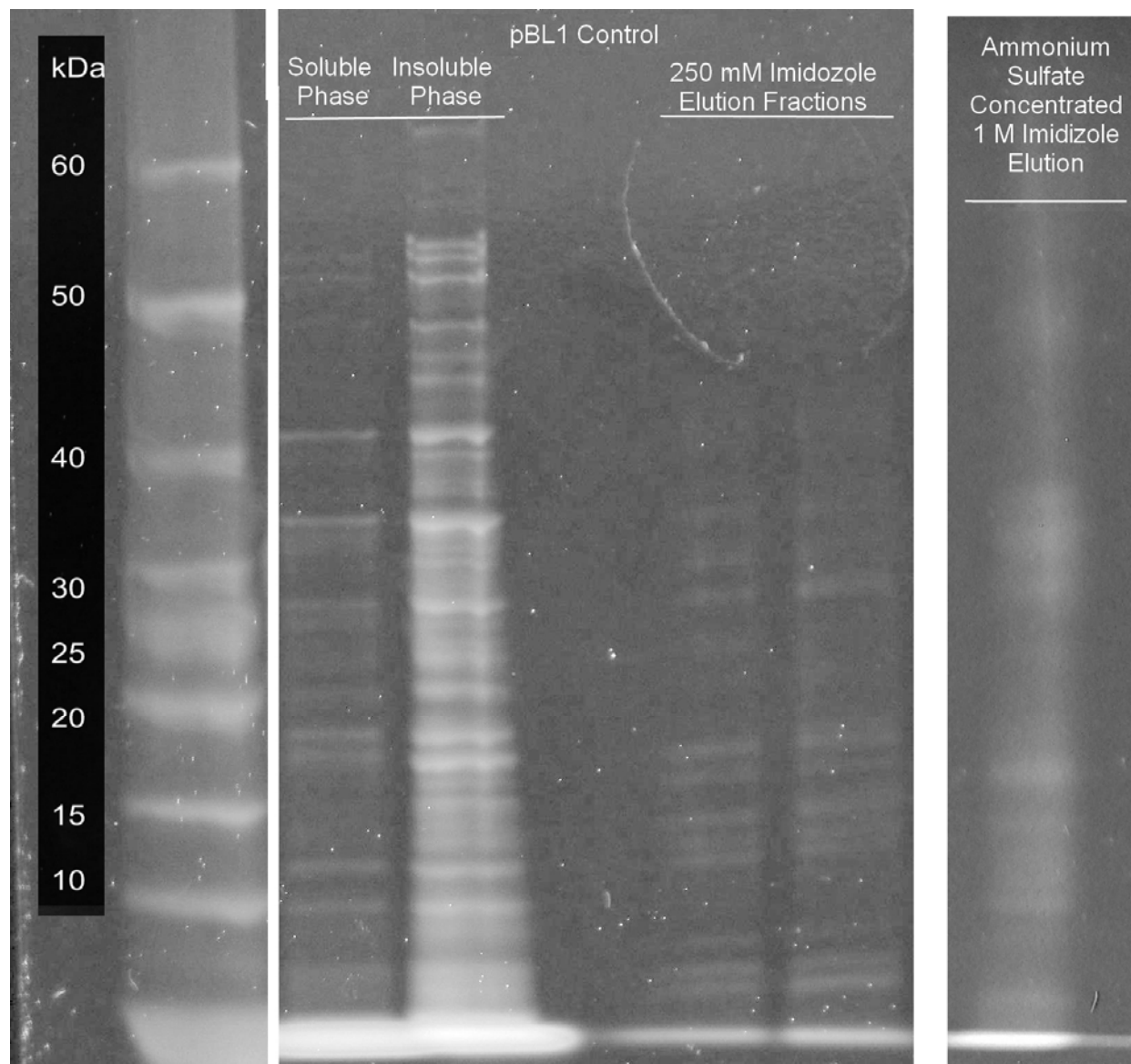


Figure 8. BD FALCON column purification of soluble phases. Two separate experiments are shown, one using elution with 250 mM imidazole and the other using elution with 1 M imidazole. Both elutions followed a 5 mM imidazole wash to elute non-specifically bound proteins. Ammonium sulfate precipitation was used in 1 M imidazole elution to increase concentration of protein in final sample.

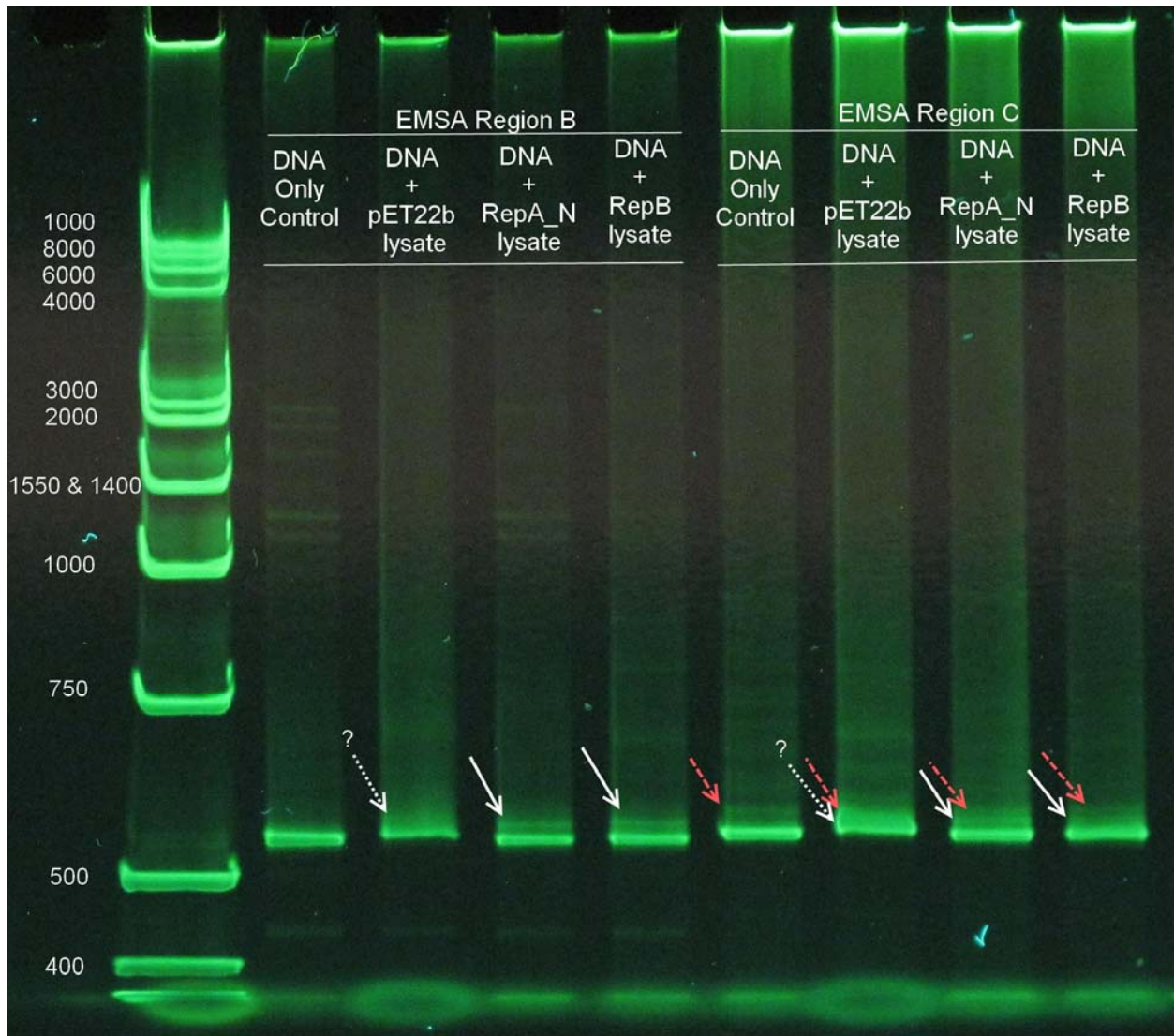


Figure 9. Preliminary EMSA of regions B and C with RepA_N and RepB crude lysates. Gel stained with SYBR® Green EMSA stain. Protein lysate from cells containing empty pET22b vector was used as a control. Solid white arrows indicate shifted DNA band. Dashed orange arrows indicate secondary bands present in DNA only sample. Dotted white arrows indicate potential shifted bands in smears in pET22b control lanes.