2020

The Biochemical Characterization of Cellular Retinoic Acid Binding Protein Type 1 (CRABP1)

Emma Magna
Assumption College

The Biochemical Characterization of Cellular Retinoic Acid Binding Protein Type 1 (CRABP1)

Emma Magna

Faculty Supervisor: Laura Marcotte, Ph.D.

Biological and Physical Sciences

A Thesis Submitted to Fulfill the Requirements of the Honors Program at Assumption College

Spring 2020

Abstract

The goal of this research was to express and purify Cellular Retinoic Acid Binding Protein Type 1 (CRABP1). CRABP1 binds to retinoic acid in the cell and shuttles the retinoic acid from one area in the cell to another. The overproduction of CRABP1 appears to be problematic because it can sequester the retinoic acid and prevent it from regulating gene expression. Previous studies have shown that increased levels of CRABP1 can result in tumor-promoting activity and disruption of lipid biology in the cell. Our objective was to express this protein in an inducible bacterial system, so that we could purify the protein and characterize its functions *in vitro*. Initial work to express the protein using a cloning vector and inducible promoter was unsuccessful; however, we were able to clone the *crabp1* gene into other expression vectors and will test transformants for inducible expression in future experiments.

Introduction

One in eight women will contract invasive breast cancer over the course of her lifetime. In the United States alone, it is estimated that in 2020, approximately 276,480 women will be diagnosed with invasive breast cancer, plus another 48,530 women will be diagnosed with non-invasive breast cancer. *Breastcancer.org*, 2019) It is estimated that 30% of cancers diagnosed in women this year will be breast cancer. Besides skin cancer, breast cancer is the most come type of cancer among women ("*Breastcancer.org*, 2020) These startling numbers beg the question, what more can be done? With one of the highest incidences, more research needs to be done on breast cancer to find a cure.

Cancer often arises from mutations in the cellular genome. This mutated genetic information can lead to the production of cell cycle proteins that either lose specific functions or have altered functions, and these mutations may result in the rapid overgrowth and division of cells. Specifically, there are two groups of genes that express cell cycle proteins. The first group is called proto-oncogenes. These are genes that produce proteins that encourage cell proliferation and inhibit cell death (Does, 2004). The other group is tumor suppressors. These genes produce proteins that prevent cell proliferation and induce cell death. These two groups work in opposition to regulate cell growth. One way in which cells become cancerous is with the mutation of a proto-oncogene into an oncogene, causing cell growth to accelerate without control. Mutations of proto-oncogenes are usually found to be dominant. This means that only one out of the two copies of the gene within the cell needs to be mutated in order to get mutated function. If these mutations occur in germ line cells, egg cells, or sperm cells, they will get passed down to offspring (Does, 2004).

Mutations could also occur in tumor suppressor genes. Mutations in these genes cause a loss of function of the resulting proteins. This means that these proteins are no longer able to inhibit cell proliferation, and the proto-oncogenes or oncogenes are able to cause mass cell divisions without the regulation of the tumor suppressors (Does, 2004). These mutations can also be hereditary, but they are recessive so the mutation would have to be present in both copies of the gene to cause an alteration in function. This could happen by the first mutated copy being inherited and the second one mutating itself, or both could mutate without hereditary influence. BReast CAncer 1 and BReast CAncer 2 (*BRCA1* and *BRCA2*) gene mutations are a common example of a hereditary mutation that can lead to breast cancer (Does, 2004). Both of these genes are tumor suppressors. BRCA1 plays a role in cell cycle control, while *BRCA2* is involved in DNA repair processes. These are both crucial functions that when disabled lead to breast cancer (Does, 2004).

Lastly, mutations can occur in DNA repair genes. Mutations in a cell's DNA may arise spontaneously from ultraviolet radiation, ionizing radiation, chemical damage, or mistakes made during DNA replication. Without functional DNA repair proteins, these gene defects could lead to an increased risk of developing cancer (Does, 2004).

Although mutations to proto-oncogenes, tumor suppressors, and the DNA repair machinery can lead to breast cancer, other proteins may also be associated with breast cancer tumorigenesis. One such protein is Cellular Retinoic Acid Binding Protein 1 (CRABP1). CRABP1's normal function is to shuttle retinoic acid, a vitamin A derivative, around the cells (Connolly et al., 2013). Retinoic acid plays an important part in cell differentiation, regulation of growth, and apoptosis (Connolly et. al., 2013). It has been observed that the overexpression of CRABP1 has been associated with breast cancer

incidence (Liu et. al., 2015). In this section I will discuss the function of retinoic acid, the role of CRABP1, how CRABP1 is thought to be related to breast cancer, and how we hope to further investigate the role of this protein in disease progression.

CRABP1 binds to retinoic acid, which is a metabolite of Vitamin A. Vitamin A is a very important vitamin for human survival and is required for development in the embryo through adulthood. In fact, not getting enough vitamin A can lead to vitamin A deficiency (VAD), and this can cause a wide variety of defects including intellectual and physical disabilities. In adults, a lack in vitamin A can impair things like vision, reproduction, and the immune system (Tanoury et al., 2013). Retinoic acid and other retinoid derivatives regulate processes like cell proliferation, cell differentiation, apoptosis, the inflammatory response, embryo growth, and functions in the nervous and immune systems (Napoli et al., 2017). This means that retinoic acid and its derivatives play a key role in stopping the overgrowth of cells and keep the cell cycle in check. For this reason, there has been extensive research done on retinoic acid for cancer treatment. The Food and Drug Administration has approved the use of retinoic acid as a leukemia and lymphoma therapy (Chen et al., 2014). Retinoic acid has the ability to inhibit tumor growth, the formation of new blood vessels, and metastasis (Chen et al., 2014).

Many proteins are involved in the transport of retinoids throughout the cell, including cellular retinol binding proteins 1 and 2 (CRBP1 and CRBP2), CRABP2, and FABP5 (fatty acid binding protein 5). These proteins shuttle retinoids from one cellular compartment to another. CRBP1 carries retinol and retinal in various cell types, while CRBP2 also carries retinol and retinal, but is specific to cells of the small intestine. CRABP2

carries retinoic acid isomers to the skin, uterus, and ovaries. FABP5 shuttles long-chain fatty acids and retinoic acid in many types of cells, but specifically liver cells (Napoli et al., 2017).

Because retinoic acid is a largely hydrophobic molecule, it is not very soluble in aqueous compartments of the cell. In order for the retinoic acid to travel through the cell, it binds to CRABP1 and is shuttled around within the cell. CRABP1 is capable of accommodating retinoic acid in a non-polar binding pocket (Napoli et al., 2017). The CRABP1 protein is structured like a barrel (Figure 1). It has hydrophobic amino acids facing the interior of the barrel where the hydrophobic tail of the retinoic acid molecule binds. The outside of the barrel has polar amino acids lining it so that it can interact with its aqueous surroundings. The carboxylate side of the retinoic acid molecule hydrogen bonds to the polar amino acids in the binding pocket.

CRABP1 has two hypothesized roles. First, CRABP1 has a high affinity for cytoplasmic retinoic acid. Researchers hypothesized that CRABP1 decreases the concentration of available retinoic acid for the nucleus while also deactivating it. When CRABP1's expression was altered alterations and mutations were executed, researchers observed that high concentrations of CRABP1 caused cells have a decrease in responsiveness to retinoic acid. That is, cells were less affected by retinoic acid as a gene expression regulator with an increase of cellular levels of CRABP1 (Wei, 2012). Another theory proposed for CRABP1's function is that it aids in the conversion of retinoic acid into other, more polar, derivatives that help regulate gene expression by binding to retinoic acid receptors in the nucleus (Wei, 2012). This pathway is not fully understood, but when cellular levels of CRABP1 were increased in one type of cell it caused an increase in the

conversion of retinoic acid into its other derivatives. This could mean CRABP1 is at least partly responsible for the modification of retinoic acid in cells (Wei, 2012).

To further test the importance of CRABP1, researchers disabled the gene in young, developing mice. These young mice then died. Contradictory to this, when CRABP1 was disabled in normal adult mice, few changes were seen in the health of the mice(Wei, 2012). It was hypothesized, however, that the adult mice may have had and alternative mechanism for the transportation and regulation of retinoic acid and its pathway (Wei, 2012). This could suggest that CRABP1, which we know shuttles retinoic acid into the cells, is needed much more in developing organisms.

Researchers also found that CRABP1 promotes pregnancy-associated breast cancer and have hypothesized that changes in glandular structure during the stages of pregnancy, specifically during lactation, may have an important role in pregnancy-associated breast cancer (McCready et al., 2014). Mammary adipose stromal cells were isolated from mice at different stages of mammary development: post-pubertal, pregnancy, lactation, involution (remodeling of the mammary gland back to its pre pregnant state), and regression. It was found that ASCs that were acquired during lactation (ASC-Ls), but not other stages of mammary development, promoted the growth of carcinoma cells (McCready et al., 2014). Upon comparing gene expression of these cells to the cells in other developmental stages, it was also determined that the CRABP1 gene was more highly expressed in ASC-Ls compared to ASCs at other developmental stages. Inhibition of *crabp1* expression in ASC-Ls restored its ability to accumulate lipids, and incubation of these inhibited cells with carcinoma cells resulted in smaller, slower growing tumor cells. These findings suggest that CRABP1

overexpression in ASC-Ls has some association with the onset or progression of pregnancy-associated breast cancer (McCready et al., 2014).

From the experiments described above, there appears to be a link between the expression of CRABP1 and breast cancer. The relationships between CRABP1, lipid binding and lipid sequestration in breast cancer is unclear. In an effort to understand CRABP1's function in these processes more completely, our research objective was to conduct biochemical assays with purified CRABP1 protein. This thesis discusses our approach to express CRABP1 and to clone the gene into several alternative expression plasmids.

Methods

Expression of CRABP1

A pBlueScript (pBS) plasmid containing the *crabp1* gene was obtained from the Tabin Laboratory at Harvard Medical School (Boston, MA). The *crabp1* gene is under the control of a *lac* operator, and its expression can be induced in the presence of isopropyl β-D-1-thiogalactopyranoside (IPTG). The plasmid was transformed into DH5α cells (see Table 1 for antibiotic resistance), and a mini prep of the cells was done to yield a concentrated amount of the CRABP1 plasmid DNA. Upon completion of the miniprep, the concentration was measured.

The plasmid was then transformed into BL21(DE3) cells, cells that typically yield large amounts of expressed protein. To induce the expression of CRABP1 in these cells, an overnight culture of BL21(DE3) cells in Luria broth (LB) was first made. Overnight cultures were made by using a sterile flask of LB broth (100mL). 100 µL of ampicillin (100 mg/ml) was put into the flask. A sterile inoculating loop was used to pick up multiple colonies from the respective plated culture and inserted into the flask of LB broth. The flask was then put in a shaking incubator at 37°C at 200 rpm for about 12 hours. A sample was taken from this culture and diluted to a ratio of 1:100 in fresh media (100 mL). This new sample grew to an optical density (OD_{600}) of approximately 0.5. Once the sample reached optimum density, 1 mL of 1 mM IPTG was added to the media to induce gene expression, and the cells were induced at 25°C. Four 1-mL samples were taken: (1) uninduced sample (i.e., before IPTG was added), (2) a sample take 3 hours after the induction of expression, (3) a sample taken 6 hours after the induction of expression, and (4) a sample of the culture after the expressed was induced overnight. The samples were centrifuged (discarding all

supernatant), and 100µL of 2x SDS buffer was added to each sample. Because the ODs of all the samples were the same, the same amount of buffer was added to each sample to ensure the same concentration of cell contents material per sample. These samples were heated at 95°C for 15 minutes, vortexing occasionally to lyse the cells. Next, the samples were centrifuged for 10 minutes, and only the supernatant was saved. These samples were then used to run two 15% polyacrylamide gels: one gel for Coomassie staining and another gel for a Western blot. For gels that were transferred for a Western blot, the gel was transferred onto nitrocellulose membrane using a Mini-Vertical PAGE/Blotting System (Bio-Rad). The gels were transferred for 45 minutes at 100 V. The membrane was placed in a cassette box with a blocking solution made with 5 g dry milk, 100mL Tris-Buffer Saline (TBS), and 100 µL of Tween. This was incubated on a rocker at 15°C for 45 minutes. Next, the membrane was rinsed three times with TBS. Then, the primary antibody solution was added, which contained 10 mL TBS, 10 µL of Tween and 5 µL of primary mouse antibody for a 1:2000 dilution ratio This was incubated on a rocker at 15°C for at least 1 hour and sometimes overnight. After incubation with primary antibody, mouse antibody, the membrane was rinsed with TBS-T three times. The secondary antibody solution was made with 10 mL of TBS-T, 0.5 g dry milk and 2 µL of anti-mouse secondary antibody. This secondary antibody solution was added and incubated on a rocker at 15°C for at least one hour. The secondary antibody was then removed. It was then rinsed three times with TBS and washed for one hour in TBS on a rocker at 15°C. The blot was then developed by adding 700 µL of Luminol Enhancer and 700 µL of Peroxide Solution and visualized on the GelDoc.

Optimization of Protein Expression Using Glucose and pLysS

Two different variations of protein expression were also attempted to halt the leaky expression if CRABP1 using the same plasmid. The first method was inducing the expression of CRABP1 in the presence of glucose. For expression in the presence of glucose, we conducted this experiment as described above except in the presence of 0%, 0.5%, 1%, and 2% glucose. From each culture, two samples were taken: (1) an uninduced sample and (2) a sample taken 3 hours after induction. The other method used was to induce expression in the presence of lysozyme. This was done using BL21(DE3) cells that contain a pLysS plasmid that expresses T7 lysozyme that inhibits expression of inducible genes prior to adding IPTG. Samples taken were again used to run two protein gels per test, using one for Coomassie staining and one for a Western blot.

Subcloning *crabp1*

Next, we obtained three different expression vectors with the aim of cloning the *crabp1* gene into these vectors. The expression vectors that we used were pET15b, pRSF, and pRSF Duet. The pET15-b and pRSF Duet include a His_6 affinity tag that can be added to the expressed protein. This His_6 affinity tag can be used later during the protein purification process using affinity chromatography. The cloning process involved multiple steps: polymerase chain reaction (PCR) to amplify the gene of interest using vector-specific primers, gel purification of the PCR products, restriction digest of the amplified *crabp1* gene and the vectors of interest using the appropriate restriction enzymes, gel purification of the digested products, ligation of the digested vector and *crabp1* gene, and

transformation of the ligated vector and gene into DH5α cells. Any positive transformants were tested through screening via colony PCR.

First, PCR was used to amplify *crabp1* present in the plasmid template. In a PCR tube, the following items were added: 2 µL of template DNA, 5 µL 10x Taq Buffer, with 4 µL $MgCl_2$, 1 µL dNTP mix, 2.5 µL forward primer, 2.5 µL reverse primer, 0.2 µL Taq DNA polymerase, and 32.8 µL sterile water. A control was made by replacing the template DNA with sterile water. A standard PCR protocol was used. PCR primers were made to contain restriction enzyme sites that corresponded to each vector (Table 1).

The PCR products were then visualized using agarose gel electrophoresis. The bands from the gel electrophoresis that corresponded to the molecular weight of the *crabp1* PCR product were then gel purified using a Wizard SV Gel and PCR Clean-Up System (Promega). This was followed by a restriction digest using the proper restriction enzymes specific to each vector. Both the *crabp1* gene and vectors were digested. Each was combined with 1 uL of each desired restriction enzyme, 3 µL of CutSmart, and the remaining volume was deionized water (30 µL total volume). This mixture was put in a 37°C bath for at least one hour.

The digests were again run on agarose gels using the same gel electrophoresis protocol from the PCR gel electrophoresis, and the resulting bands from the digested *crabp1* genes and digested vectors were then gel purified. Next, *crabp1* was ligated into each of the vectors in a series of experiments. To insert the *crabp1* gene into the desired vector, different ratios of the gene and desired vector were used and subsequently combined with DNA ligase. The ratios included: a 1:1 plasmid to insert ratio, a 3:5 plasmid

to insert ratio, and a 1:7 plasmid to insert ratio. The tubes were then incubated at room temperature for at least two hours.

To see if the ligations were successful, each ligation reaction was transformed into DH5α cells. The cells were plated on LB plates with proper antibiotic resistance (see table 1). Along with plating the experimental ligations, positive and negative control ligations were also plated. Positive-control ligations were made by using an uncut vector, and negative controls were made using digested vectors without an added *crabp1* gene. Any colonies that grew on the plates were tested via colony PCR to determine if the gene of interest was present. Colony PCR is performed by picking colonies that grew on the resulting plated cultures with a small pipette tip and inserting a part of the colony into a small sample of hot water for 10 minutes. This sample is then used as the "DNA" sample for the previously mentioned PCR protocol.

Results

The purpose of these experiments was to overexpress and purify CRABP1. To start this, IPTG was added to a liquid culture of BL21(DE3) cells transformed with the pBS *crabp1* plasmid in order to induce expression of CRABP1. A Western blot showing the expression of the CRABP1 protein is shown in Figure 2 with a single band observed at approximately 16 kDa, the presumed molecular weight of the CRABP1 protein. As the induction time increases, the CRABP1 bands become lighter and lighter. Though the function of IPTG is to induce expression, the highest expression appears to occur before it is added. It is hypothesized that the CRABP1 protein is produced before induction and is degrading over time.

Because there was no increase or overexpression of CRABP1, the experiment was repeated with the addition of glucose. Different concentrations of a glucose solution (0.5%, 1%, and 2% glucose) were added to liquid cultures with the goal of inhibiting leaky expression of the CRABP1 protein. Two protein gels were run: one for a Coomassie stain (Figure 3) and one for a Western blot (Figure 4). In Figure 3, no discernable differences could be observed in the Coomassie gel when comparing expression before and after induction. In the lanes with 0.5%, 1%, and 2% glucose concentration, the overall protein expression of the cells decreases from pre to post induction. Since anti-CRABP1 from mice was used again and the mass of CRABP1 is 16 kDa, it can be assumed that the protein in the bands of the western blot is CRABP1. Similarly to Figure 2, pre-induction saw more expression than post-induction. Since glucose is supposed to prevent leaky expression, there should be very little to no CRABP1 present in the pre-induction lanes.

BL21(DE3) cells with an added pLysS plasmid, which expresses lysozyme in order to inhibit RNA polymerase before the cells are induced with IPTG, were induced with IPTG. Samples from before and 3 hours after induction were taken and run on an SDS-PAGE gel. Figure 5 shows Coomassie stained protein gel with the uninduced in lane 2 and 3 hours post induction in lane 3. There was little change in protein expression from before to after induction.

When inducing with different forms of leaky expression inhibition failed to show an increase in expression, the *crabp1* gene was subcloned into new plasmids in order to see if those might work better for expression. Our objective was to subclone the *crabp1* gene into three new vectors: pET-15b, pRSF-1 and pRSF-Duet.

The *crapb1* gene was amplified by PCR using primers that had specific restriction enzyme sites that corresponded to the different plasmid restriction enzyme sites into which we hoped to insert the gene. Figure 6 shows the PCR products of *crabp1* with pET-15b and pRSF-1 specific primers, on an agarose gel. For lanes 2-7, there were bright bands below 500 base pairs, assumed to be *crabp1.* Figure 7 shows the PCR product of *crabp1* using primers specific for the pET-Duet vector. Lanes 2-4 contain the PCR product, which all show *crabp1* sized bands.

After the presumed *crabp1* DNA bands were gel purified, the crabp1 and vectors were digested with their corresponding restriction enzymes. Figure 8 displays the results of the restriction digest done on the pET-Duet and pRSF-1 plasmids, as well as the *crabp1* PCR product. Figure 9 shows the restriction of pET-15b and the *crabp1* PCR product. The pET-Duet digestion did not yield a significant amount of digested vector, even when concentrations of vector were increased in the digest reactions (Figure 8). For this reason,

we focused our subcloning efforts on the pRSF-1 and pET-15b vectors from this point forward.

After the pRSF-1 and pET-15b vectors and *crabp1* were successfully digested with their respective restriction enzymes, the samples were gel purified, ligated, and transformed into DH5α cells. The ligation of *crabp1* into pET-15b was not successful, as no plated cultures grew. The plated cultures of the ligation of *crabp1* did accumulate colonies on all of the plates. Figure 12 shows the results of the colony PCR of *crabp1* ligated into pRSF-1 from the transformed DH5α cells. Lanes 2-7 (gel on the left) and lanes 2-7 (gel on the right) show faint bands right below the 500 base-pair mark, thought to be *crabp1*. The positive and negative control lanes did not show any bands corresponding to *crabp1*.

Discussion

The goal of this research was to express and purify a recombinant form of the CRABP1 protein with the objective of gaining a better understanding of the biochemistry of the protein.

The first objective to be completed was to express CRABP1 in BL21(DE3) cells using IPTG to induce its expression. When IPTG was added to the BL21(DE3) cells containing the pBS plasmid encoding the *crabp1* gene, the expression of CRABP1 decreased, as shown in the Western blot (Figure 2). The bands were observed at a molecular weight between 15 and 16 kDa, which is consistent with the molecular weight of the CRABP1 protein, 15.556 kDa. However, it was expected that protein expression would increase upon induction and that we would observe darker bands post-expression, indicative of more protein being produced in the cells. Instead, the addition of the IPTG had no effect, as protein expression either decreased or the protein began to degrade over time. It is not entirely clear why

IPTG induction was not successful with this plasmid. In *E. coli*, the lac operon contains the gene *lacI*, which expresses the lac repressor. The lac repressor binds to the lac operator and inhibits transcription from moving to the promoter. When IPTG is added to *E. coli*, expression is induced at the promoter (Hansen et al., 1998).

In an attempt to increase the expression of CRABP1 protein in these cells transformed with the pBS *crabp1* plasmid, glucose was added to the culture media. Glucose is known to prevent "leaky" protein expression before induction (Novy and Morris, 2001). Varying percentages of glucose were added (0%, 0.5%, 1% and 2%) to the cell cultures prior to induction with IPTG. The addition of glucose did not appear to prevent expression prior to induction. In fact, we observed a decrease in expression over the course of time, regardless of the percentage of glucose added to the media (Figures 3 and 4). These results were similar to that of the protein expression trials in the absence of glucose (Figure 2), in that the expression seemed to decrease post-induction or there was protein degradation over time.

Next, BL21(DE3) pLysS cells were used, again in an attempt to prevent pre-induction expression of the protein. This was executed by transforming the pBS plasmid with *crabp1* into cells containing the pLysS plasmid. This plasmid expresses T7 lysozyme, which inhibits the expression of inducible genes before IPTG induction. The pLysS experiments were conducted by collecting samples pre-induction and three hours post-induction with IPTG. As observed in the previous trials, there was still expression before induction and no visible increase in expression after induction. In fact, protein expression seemed to decrease over time (Figure 5).

Because we were unable to achieve inducible expression of the CRABP1 protein in the pBS plasmid, our next step was to clone the gene into a different plasmid, which would serve two aims: (1) achieve inducible protein expression and (2) add an affinity tag to the protein, which would ultimately lead to more efficient purification. We had three expression vectors available to us: pET15b, pRSF-1, and pET-Duet.

The *crabp1* gene was first amplified using pET15b, pET-Duet and pRSF-1 specific primers. Primer design took time and care, and this process had to be repeated after our first attempt, as the *crabp1* gene was not in-frame with the encoded purification tags on our first attempt at primer design (data not shown). However, we were eventually able to amplify *crabp1* using appropriately designed primers for the plasmids we wanted to use.

Our attempts at digesting the different plasmids with our restriction enzymes of interest were, in part, successful. The pRSF-1 and pET-15b plasmids had the most successful digests, whereas the pET-Duet restriction digest did not work as well as seen by the unconcentrated double band seen on the gel(Figure 8), possibly due to the use of an older aliquot of the restriction enzyme. Once both the plasmids and *crabp1* were digested and gel-purified, we next attempted ligation.

The digested *crabp1* insert was ligated with the digested plasmids. We attempted these ligations many times using a variety of ratios of insert-to-plasmid. Additionally, we tried varying lengths of time for these ligations and different aliquots of the ligase to see if this would aid in successful ligation reactions. These ligations were transformed into BL21(DE3) cells. The ligation only seemed to be successful using the pRSF-1 plasmid, as no colonies grew on the plates with the pET-15b plasmid and insert. In addition, we also noted

that the positive control plates had many colonies grow, while the negative control plates had very few colonies.

Cells from the resulting pRSF-1 colonies were collected for colony PCR. The results of the colony PCR for the pRSF-*crabp1* colonies showed bands at the correct molecular weight for *crabp1* from all tested colonies except for the negative and positive controls. Though the bands were faint, we grew up two of these colonies and performed minipreps to get concentrated plasmid DNA. Our last experiments included testing the expression of one of the pRSF-*crabp1* plasmids, though we did not see overexpression of the CRABP1 protein in Coomassie gels (data not shown). Further tests with respect to these isolated plasmids will include sequencing the plasmids and also performing a Western blot to see if the CRABP1 protein is expressed.

Once the gene is successfully incorporated into an expression plasmid, the CRABP1 protein will then be purified. The protein will be overexpressed, the cells lysed using a sonicator, and the lysate applied to a nickel resin that specifically binds to the His-tagged protein. Additional purification steps may be necessary depending on the purity of the protein, which can be assessed by running a Coomassie protein gel. Should we not have efficient lysis of the cells, an alternative method for acquiring the CRABP1 protein is to use a signal sequence to secrete CRABP1 from the cells. A signal sequence is an extension on the N-terminus of a protein, which conveys to the cellular machinery where to be "sent" (Martoglio,1998). Such a sequence could be used to secrete the protein into the periplasmic space, which would allow for easier extraction from the bacterial cell.

Though we spent most of the time in the lab working on the subcloning of the *crabp1* gene into different plasmids, there are many experiments that we hope to pursue

upon successful cloning and purification of the CRABP1 protein. The experiments that we envision would first be to assess the protein's folding using techniques that might include circular dichroism or differential scanning fluorimetry. Non-bacterial proteins that are expressed in these microorganisms are sometimes misfolded, and before testing CRABP1's ability to bind to ligands or inhibitors we would first want to make sure that it was folded. The purpose of finding a small molecule to bind to CRABP1 would be to free up retinoic acid. If, using a small molecule library, we could find a molecule that has a higher affinity to CRABP1 than retinoic acid, retinoic acid would be freed up to accumulate in the adipose stromal cells. The affinity of the small molecules to CRABP1 could be tested by using a thermal shift assay. Lastly, we would make mutations to CRABP1 in order to find a mutation that causes CRABP1 to have the inability to bind to retinoic acid. This would be done by systematically mutating single amino acids within the binding pocket for retinoic acid. Once a promising mutation is found, the mutated CRABP1 would be tested in the adipose stromal cells to assess lipid accumulation.

Next, the purified and properly folded CRABP1 protein could be screened with a small-molecule library to find small molecules that have a high affinity for CRABP1. It was found in a previous study that CRABP1 prevents the accumulation of retinoic acid in adipose stromal cells by binding to and sequestering it. The goal in this part of the experiment would be to find small molecules that bind to CRABP1 and have a higher affinity for CRABP1 than retinoic acid and CRABP1. The affinity of the small molecules to CRABP1 could be tested via thermal shift assay or isothermal titration calorimetry (ITC). Then, tests could be done with these molecules *in vitro* using techniques such as a lipid accumulation assay to see whether the cells can regain the ability to accumulate lipids. It is

hypothesized that CRABP1 would bind with a molecule that has a higher affinity, leaving retinoic acid to be free to accumulate in the adipose stromal cells.

Lastly, we could attempt to mutate specific amino acids to see if any mutations cause CRABP1 to have the inability to bind to retinoic acid. In a normal CRABP1 and retinoic acid interaction, the hydrophobic β-sheets in the CRABP1 protein encase the hydrophobic tail while the carboxylate head group of the retinoic acid hydrogen bonds to the polar amino acids in the CRABP1. We could systematically make single mutations of amino acids, specifically within the retinoic acid binding pocket of CRABP1, to determine if these mutations cause a decrease in affinity between CRABP1 and retinoic acid. To do this, site-directed mutagenesis could be used on the plasmid sequence followed by expression and purification of the mutated CRABP1. We could then determine if the mutated protein is folded correctly and test the affinity of retinoic acid to the mutated CRABP1 protein. Finally, this protein could then be tested in ASC-Ls or other supporting cells to observe the effectiveness of mutating the CRABP1 on lipid accumulation in adipose stromal cells. Although this research leaves many open-ended questions about the relationship between CRABP1 and pregnancy associated breast cancer and the specific role that it plays, these questions have potential to encourage research to improve our understanding of this complex disease.

References

Chen, M., Lin, H., Yang, T. Retinoic acid and cancer treatment. *BioMedicine* vol. 4,4 (2014): 22.

Connolly, R.M., Nguyen, N.K., and Sukumar, Set al. Molecular pathways: current role and future directions of the retinoic acid pathway in cancer prevention and treatment. *Clin Cancer Res.* **2013**.

Does, A., Johnson, N., Thiel, T. "Rediscovering Biology: Molecular to Global Perspectives." *Annenberg Learner*, **2003**, learner.org/series/rediscovering-biology-molecular-to-global-perspectives/.

Fogh, K., Vorhees, J., Astrōm, A. Expression, Purification, and Binding Properties of Human Cellular Retinoic Acid-Binding Protein Type I and Type II. *Arch of Biochem Biophys.***1993**.

Goh, A., Coffill, C., Lane, D. The Role of Mutant p53 in Human Cancer. *J Pathol*. **2011**.

Gudas L., Wagner J. Retinoids regulate stem cell differentiation. *J cell*. **2011**.

Hansen, L., Knudsen, S., Sørensen, S. The Effect of the LacY Gene on the Induction of IPTG Inducible Promoters, Studied in Escherichia Coli and Pseudomonas Fluorescens. *Curr Microbiol*. **1998**.

Kainov, Y., Favorskaya, I., Delektorskaya, V., Chemeris, G., Komelkov, A., Zhuravskaya, A., Trukhanova, L., Tavitian, B., Dyakova, N., Zborozskaya, I., Tchevkina, E. CRABP1 Provides High Malignancy of Transformed Mesenchymal Cells and Contributes to the Pathogenesis of Mesenchymal and Neuroendocrine Tumors. *Cell Cycle*. **2014.**

Liu, R., Garcia, E., Glubrecht, D., Poon, H., Mackey, J., Godbout, R. CRABP1 Is Associated with a Poor Prognosis in Breast Cancer: Adding to the Complexity of Breast Cancer Cell Response to Retinoic Acid. *Mol Canc*. **2015**.

Martoglio, B., Dobberstein, B. Signal Sequences: More than Just Greasy Peptides. *Trends Cell Biol*. **1998**.

McCready, J., Arendt, L.M., Glover, E., Iyer, V., Briendel, J.L., Lyle, S.R., Naber, S.P., Jay, D.G., Kuperwasser, C. Pregnancy-associated breast cancers are driven by differences in adipose stromal cells present during lactation," *Breast Cancer Res.* (**2014**).

Napoli, J.L. "Cellular retinoid binding-proteins, CRBP, CRABP, FABP5: effects on retinoid metabolism, function and related diseases," *Pharmacol Ther.* **2017**.

Novy R., Morris B. Use of glucose to control basal expression in the pET system innovations. *BioTechniques*. **2001.**

Nbcf. Triple Negative Breast Cancer : The National Breast Cancer Foundation. *Www.nationalbreastcancer.org*, www.nationalbreastcancer.org/triple-negative-breast-cancer.

Tanoury, Z., Piskunov, A., Rochette-Egly, C. "Vitamin A and Retinoid Signaling: Genomic and Nongenomic Effects." *J Lipid Res*. **2013**.

Ullah, I., Subbarao, R., Rho, g."Human Mesenchymal Stem Cells - Current Trends and Future Prospective." *Bioscience Reports*. **2015**.

U.S. Breast Cancer Statistics. *Breastcancer.org*, **2019**.

Wei L. Chromatin remodeling and epigenetic regulation of the CrabpI gene in adipocyte differentiation. *BBA-MOL CELL BIOL L.* **2012.**

"1CBR." CRYSTAL STRUCTURE OF CELLULAR RETINOIC-ACID-BINDING PROTEINS I AND II IN COMPLEX WITH ALL-TRANS-RETINOIC ACID AND A SYNTHETIC RETINOID *RSCB Protein Data Bank*. **1995**.

Figures and Captions

Table 1.

Plasmid	Antibiotic Resistance	Restriction Enzymes
pBS	Ampicillin	N/A
pRSF^{-1} Duet	Kanamycin	PstI and HindIII
pET^{-1} Duet	Ampicillin	EcoRI and HindIII
pET15b	Ampicillin	NDEI and BamHI

Table 1. This table shows the specific antibiotic resistances and restriction digest enzymes needed for all plasmids used.

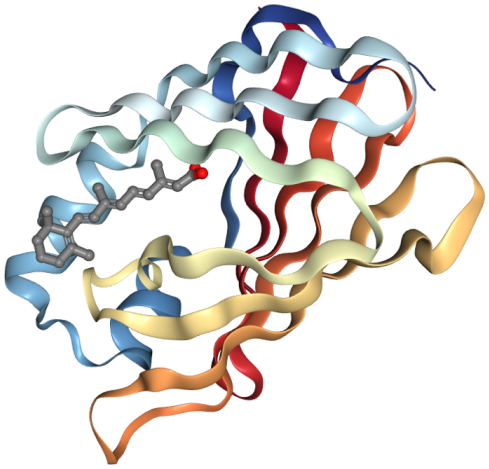


Figure 1. Structure of CRABP1 bound to retinoic acid. Image generated by WebGL Protein Viewer using the atomic coordinates from PDB ID 1CBR.

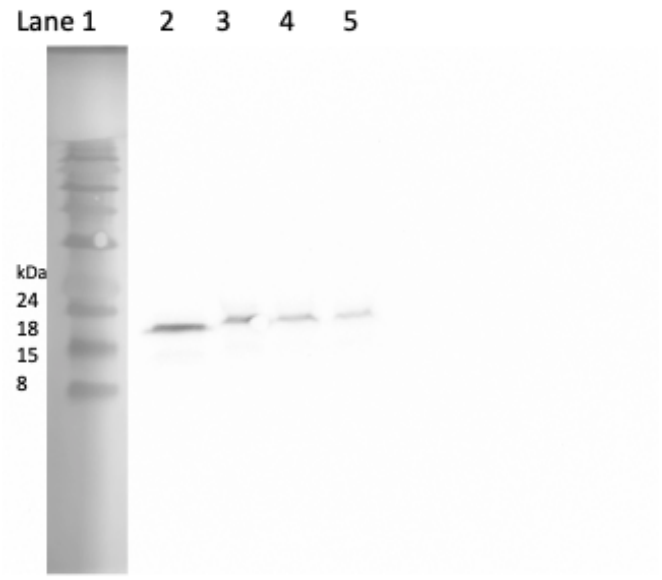
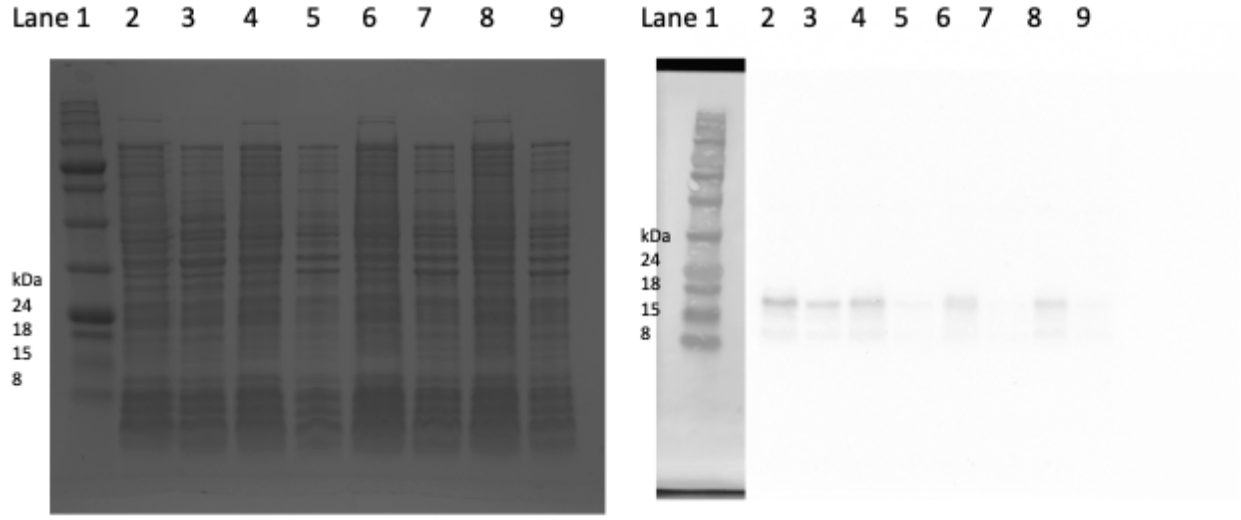


Figure 2. The expression of CRABP1 in BL21(DE3) cells. Lane 2 are uninduced cells, Lane 3 are cells lysed 3 hours post-induction, Lane 4 are cells lysed 6 hours post-induction, and Lane 5 are cells lysed after being left overnight post-induction.



Figures 3 and 4. The expression of CRABP1 in BL21(DE3) cells with the addition of glucose. The lanes for both are as follows: (1) molecular weight ladder , (2) 0% glucose uninduced, (3) 0% glucose 3 hours post-induction, (4) 0.5% glucose uninduced, (5) 0.5% glucose 3 hours post-induction, (6) 1% glucose uninduced, (7) 1% glucose 3 hours post-induction, (8) 2% glucose uninduced, and (9) 2% glucose 3 hours post-induction. Figure 2 (left) shows the Coomassie stain of all protein expression while Figure 3 (right) shows the expression of just the CRABP1 protein. For both figures, all lanes show a noticeable decreased in expression from before to after induction.

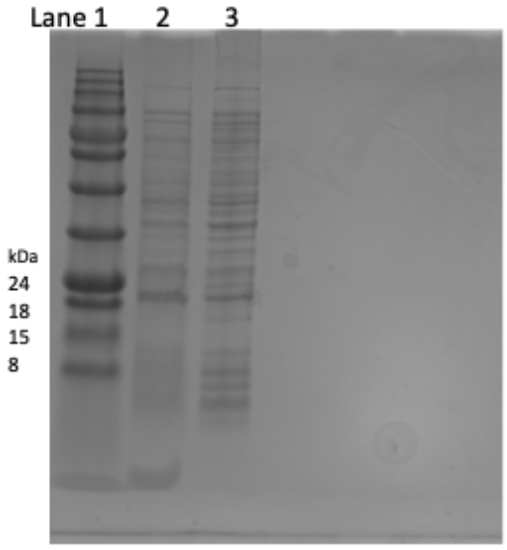


Figure 5. The Expression of CRABP1 in pBS with BL21 cells with the addition of pLysS plasmid. Lane 1 contains the molecular weight marker, lane 2 contains the uninduced sample, and lane 3 contains the 3-hour post-induction sample. There are discernable bands before and after induction with little change in concentration of the bands.

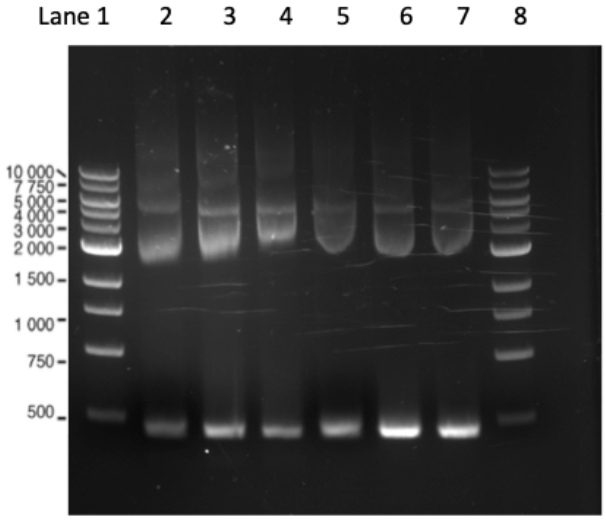


Figure 6. PCR amplification of *crabp1* using pET-15b- and pRSF-1-specific primers. Lanes 1 and 8 contain a 1kb ladder. Lanes 2-4 contain the amplified *crabp1* using primers designed for the pET15b plasmid. Lanes 5-7 contain the amplified *crabp1* using primers specific for the pRSF-1 plasmid. Lanes 2-7 all show bright bands right below 500 base pairs corresponding to the *crabp1* gene.

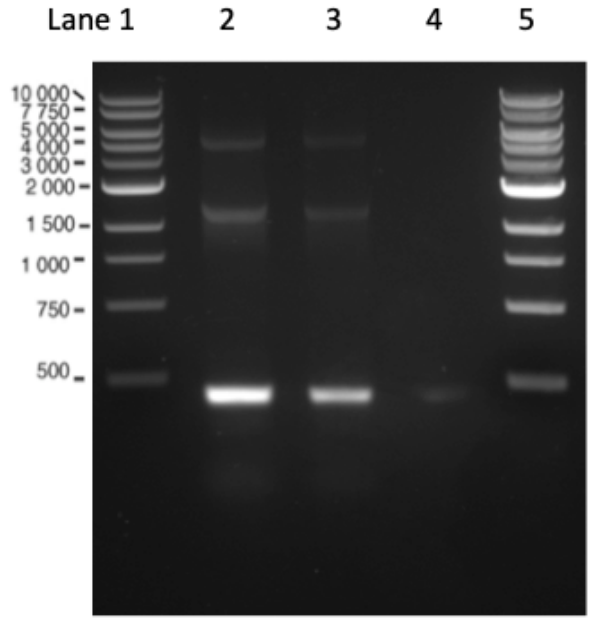


Figure 7. PCR amplification of *crabp1* using pET-Duet specific primers. Lanes 1 and 5 contain a 1kb ladder. Lanes 2-4 contain the amplified *crabp1* using primers specific for the pET-Duet plasmid. The most intense bands in lanes 2-4 are located below 500 base pairs, again corresponding to the *crabp1* gene.

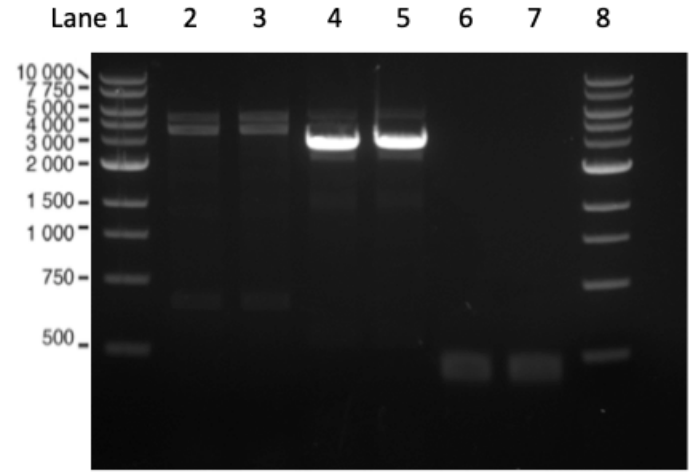


Figure 8. Restriction digest of pET-Duet, pRSF-1 and *crabp1* with EcoRI and HindIII. Lanes 1 and 8 contain a 1kb ladder. Lanes 2 and 3 contain the digested pET-Duet plasmid, lanes 4 and 5 contain the digested pRSF-1 plasmid and lanes 6 and 7 contain the digested *crabp1*.

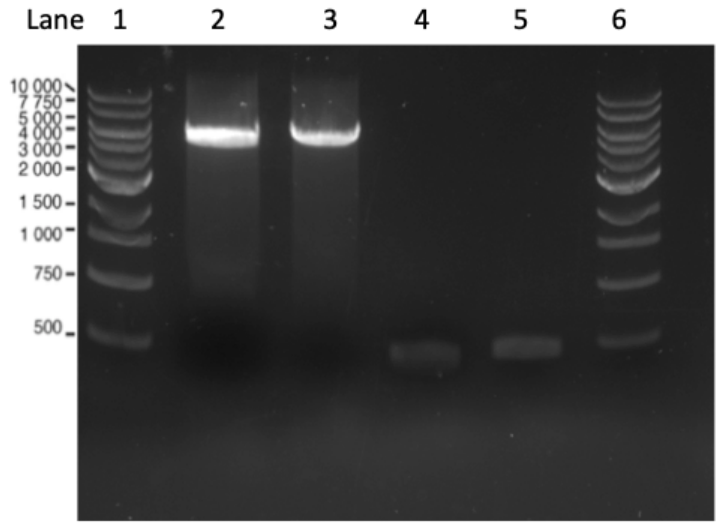


Figure 9. Restriction digest of pET-15b plasmid and *crabp1*. Lanes 1 and 6 contain 1kb ladders. Lanes 2 and 3 contain the digested pET15b plasmid and lanes 4 and 5 contain the digested *crabp1*.

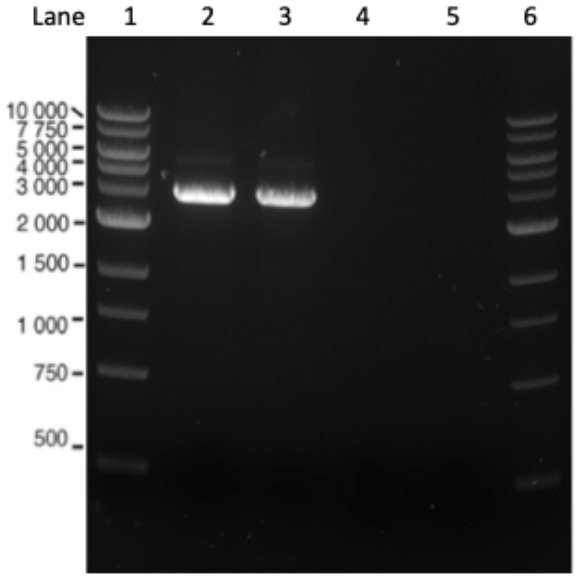


Figure 10. Restriction digest of pRSF-1 plasmid and *crabp1*.
Lanes 1 and 6 contain 1kb ladder, lanes 2 and 3 contain the digested pRSF-1 plasmid with bright bands at about 4000 base pairs and faint bands between 4000 and 5000 base pairs. Lanes 4 and 5 contain the digested *crabp1*, with no discernable bands. Though no digested *crabp1* is observed on this gel, we were able to digest the gene is subsequent experiments to be utilized in ligation reactions (data not shown).

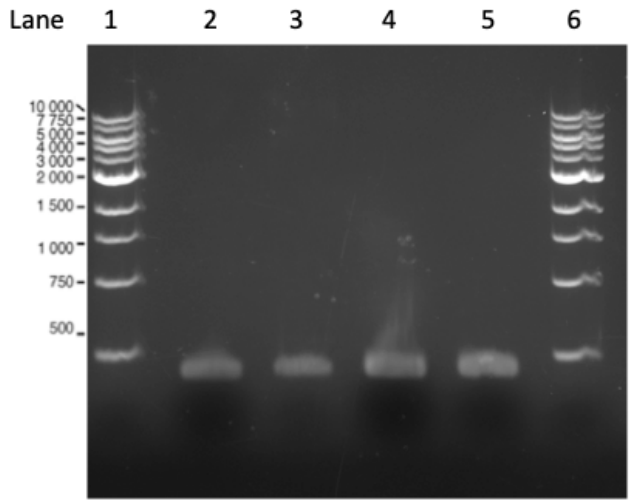


Figure 11. Restriction digest of *crabp1* using pRSF-1 and pET-15b enzymes. Lanes 1 and 6 contain 1kb ladder, lanes 2 and 3 contain the digested *crabp1* using the BamHI and NDEI restriction enzymes and lanes 4 and 5 contain the digested *crabp1* using HindIII and PstI.

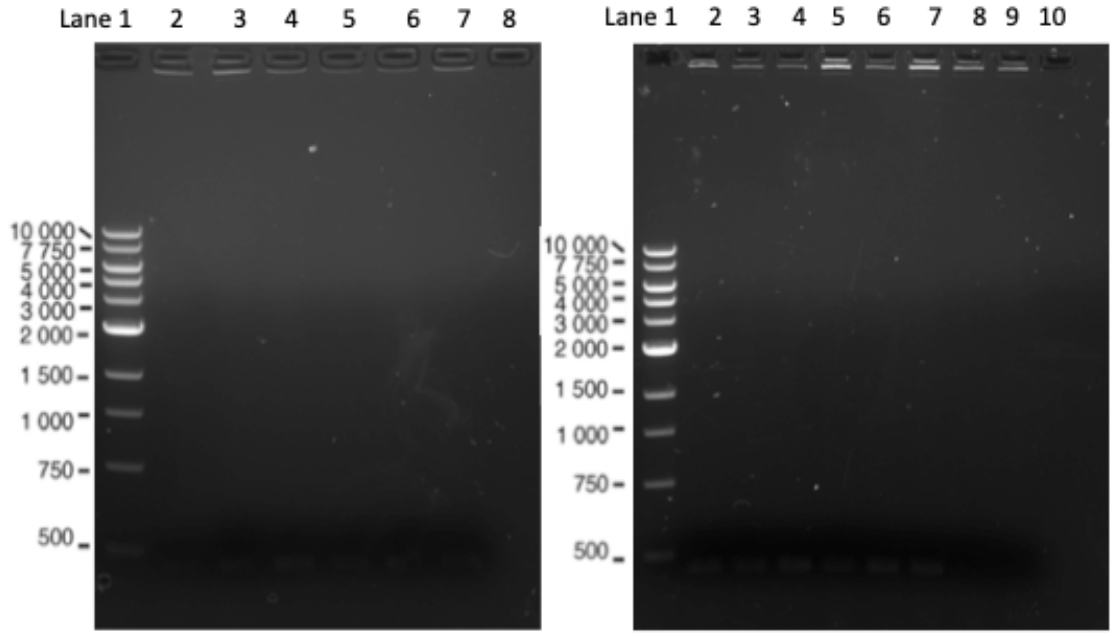


Figure 12. Colony PCR of *crabp1* ligation into pRSF-1. In the 8-well gel (left) lane 1 contains a 1kb ladder, lanes 2-4 contain 1:1 plasmid to insert ligation samples and lanes 5-7 contain 3:5 plasmid to insert ligation samples. In the 10-well gel (right), lane 1 contains a 1kb ladder, lanes 2-4 contain 1:7 ligation samples plasmid to insert, lane 8 contains a positive control and lane 9 contains a negative control. There are faint bands below the 500 base-pair mark in lanes 2-7 of both gels. Lanes 8 and 9 of the 10-well (right) gel do not show any bands.