Developing an RNAi based approach to silence the expression of HMGA1 proteins in human mammalian epithelial cells

Patty Martin
UH 450
Spring 2005
Dr. Raymond Reeves

Department of Molecular Biosciences

Honors Thesis
*************************
PASS WITH DISTINCTION
TO THE UNIVERSITY HONORS COLLEGE:

As thesis advisor for Patty Martin,

I have read this paper and find it satisfactory.

Raymond Bauer
Thesis Advisor

April 29, 2005
Date
Table of Contents:

List of Figures and Tables .............................................................................................................. 3
Précis .............................................................................................................................................. 4
Introduction ..................................................................................................................................... 6
Materials and Methods .................................................................................................................. 14
Discussion and Results .................................................................................................................... 30
Conclusions ..................................................................................................................................... 36
Future Directions ............................................................................................................................ 37
References ....................................................................................................................................... 39
List of Figures and Tables:

Figure 1-1: A generic model of DNA.................................................................9
Figure 1-2: The pathway of RNAi.................................................................11
Figure 1-3: The method of DNA insertion into a vector.................................13
Figure 2-1: The difference in sequence between HMGA1a and HMGA1b.........15
Figure 2-2: The nucleotide sequence chosen to down-regulate HMGA1a expression........15
Figure 2-3: The nucleotide sequence chosen to down-regulate HMGA1b expression........15
Figure 2-4: Restriction enzyme recognition sequence and cut site.....................16
Figure 2-5: General structure of an shRNA oligonucleotide.............................17
Figure 2-6: The process of forming the shRNA..............................................18
Figure 2-7: pBS/U6 vector map..................................................................19
Figure 2-8: Agarose gel in gel box connected to electrical source....................22
Figure 2-9: Separation of DNA, RNA, and protein during cell harvest.............25
Figure 2-10: Western Blot Transfer setup....................................................28
Figure 2-11: Transfer gel box....................................................................28
Figure 2-12: Antibody detection complex....................................................29
Figure 3-1: Confirming oligo ligation...........................................................30
Figure 3-2: Confirming vector identity.......................................................31
Figure 3-3: Confirming clone identity I.......................................................32
Figure 3-4: Confirming clone identity II......................................................33
Figure 3-5: Sequencing the possible pBS/U6 clone......................................34
Figure 3-6: Observed and predicted siRNA down-regulation..........................35
Précis:

High Mobility Group (HMG) proteins are located in the nucleus of mammalian cells and have been found to associate with a number of cellular components including DNA and proteins. They have been shown to play a role in cell growth and division as well as cancer formation. In fact, increased levels of HMG are actually associated with more severe tumor progression. Cell lines that over-express HMG proteins are not able to repair themselves as well as cells that do not over-express these proteins. This has led researchers to the idea of targeting these proteins with anti-cancer drugs. Before the drugs can be developed, however, the many effects that HMG proteins play on the cell must be determined.

In the past, the effects of HMG proteins on the cell have been studied in cell lines that have been engineered to over-express the proteins and then compared to cell lines that do not over-express HMG. General differences in expression levels of other proteins can be determined in this way, but it does not allow the researcher to determine the effects of turning off HMG expression in the over-expressing cancer cells. As this would be the consequence of targeting HMG with cancer drugs, this is the goal of current research.

RNA inhibition (RNAi) has been used to successfully silence a number of genes. It involves incorporating a piece of DNA into a vector that can be stably inserted into the cells and be transcribed into a short hairpin RNA (shRNA). Ideally, the shRNA is specific to inhibit production of HMG proteins, and has no effect on other proteins in the cell. To date, research only shows what happens when the proteins are over-expressed. An RNAi vector would allow researchers to determine the effects of turning off HMG proteins in cells over-expressing them, and would be a significant contribution to the field. The goal of this research was to develop two separate vectors that preferentially down-regulate HMGA1a and HMGA1b protein expression
(two isoforms of HMG) separately through the RNAi pathway. Time restrictions have only allowed the development of the HMGA1a vector.

There are two processes through which RNAi can occur. The first involves incorporating a small piece of DNA into a vector and delivering the vector to the cell as previously described. This method is the most favorable because the protein is continuously down-regulated and a greater fold decrease in protein level is observed. The other method involves directly delivering small interfering RNAs (siRNAs) to the cell, which will only down-regulate the protein until the molecules are used up. This method was examined but determined to be less successful than the vector approach due to cost and limited levels of down-regulation. The data suggests that siRNA down-regulation may approach a maximum down-regulation around 3 fold. Although the vector has still not been tested in the cells, it is predicted to demonstrate a much greater degree of down-regulation.

The potential information that can be gained through the use of this vector would have a huge impact on the HMG field of research. It would not only provide a way to differentiate between two very similar proteins, it will also allow researchers to study the effects of turning off protein expression for several cell generations. This information could eventually lead to the development of new anti-cancer drugs.
Developing an RNAi based approach to silence the expression of HMGAI proteins in human mammalian epithelial cells

Introduction:

High Mobility Group (HMG) proteins function primarily within the nucleus of cells and have been found to interact with DNA and many different types of protein. Due to their ability to associate with such a wide variety of partners, many functions have been linked to the HMGAI proteins. Such examples include regulation of cell growth, proliferation (cell division), differentiation (cell type determination), neoplastic transformation (the onset of tumors), and apoptosis (cell death). Since these proteins regulate such a wide range of activities, the levels of the proteins in the cell are very critical. Maximal expression in mammalian cells has been linked to embryonic development, with a decline to low or undetectable levels in differentiated, non-dividing adult cells (8). This is expected since cells are undergoing the most growth during the embryonic period when a baby is formed in a mere nine months. In addition, growth and differentiation proteins would not be needed in adult cells that are simply in a static, maintenance state with respect to growth. Interestingly, while normal adult cells contain little to no amount of HMG proteins, tumor cells contain high levels of these proteins. In fact, increasing concentrations of HMG are actually diagnostic markers for the potential of the tumor cells to be malignant (cancerous) and metastasize (travel to other parts of the body) (8).

As a result of their vital position in the regulation of both normal and neoplastic growth, HMG proteins have been considered as potential drug targets. Several approaches have been considered in targeting these proteins, the most obvious being to lower the cellular concentration of HMG proteins (7). This approach may work for two reasons; first, since the proteins are over-expressed in oncogenic cells and barely present in normal cells, the oncogenic cells could
potentially be targeted while leaving the normal cells untouched (7). In addition, lowering the concentration of HMG proteins in cancer cells may in effect convert the oncogenic cells back to normal cells. Reports have already shown that lowering the concentration of HMG proteins in over-expressing cells begins to decrease tumor severity. Decreasing the HMG concentration has been shown to inhibit neoplastic transformation, suppress cell growth rate, decrease the ability to metastasize, and preferentially induce apoptosis in cancer cells while leaving normal cells untouched (7). These results show promise for the approach of lowering the effective concentration of HMG proteins in the cell as a means of treating cancer.

Before a drug is designed to down-regulate HMG proteins, the entire spectrum of proteins regulated and affected by HMG proteins must be explored. The effect of HMGA1 expression levels has been studied by comparing the differences between two differently expressing cell lines. The first is derived from human breast epithelial tissue and has low levels of HMGA proteins. The cell line is referred to as the parental line and is named MCF-7. The second cell line is derived from the same breast cancer tumor tissue. The line is a clone of MCF-7 but is transgenic for HMGA1a therefore over-expressing this isoform. This transgenic HMGA over-expressing cell line is named HA-7C-Cs. While this approach gives a general idea of the differences in protein expression due to varying concentrations of HMGA protein, it does not reveal what happens to oncogenic cells when the levels of HMGA protein are decreased. In addition, a protein that appears to be regulated by HMG due to a change in level between the two cell lines may actually be caused by the artificial over-expression when the transgenic lines were established. For example, the HA-7C-Cs line was exposed to a transfecting agent, new DNA, forced to integrate this DNA and then to express this new gene while the MCF-7 line did not face any of these situations. For this reason, a method in which the concentration of HMG proteins
could be diminished in an over-expressing cell line would be of great value to the laboratory. While several methods are plausible including using a dominant negative protein, antibodies or antisense down-regulation, it was determined that using RNAi would provide a quick, effective means of down-regulating the HMG proteins and possibly distinguishing between the two main subtypes.

There are many proteins in the HMG protein category, and the exact functions of these proteins are still being determined. The entire group of HMG proteins was recently renamed to avoid confusion with non-nuclear HMG proteins (2). Of relevance to this project are the HMGA1 proteins, specifically the HMGA1a and HMGA1b subtypes. These proteins result from alternative splicing of 33 nucleotides (11 amino acids) from the mRNA transcript of the *HMGA1* gene. If the segment is spliced out, HMGA1b protein is translated. If no splicing occurs, HMGA1a protein is translated (5). To better understand what this is implying, a brief discussion of DNA, RNA, and their relationship to protein expression is in order.

Deoxyribonucleic acid (DNA) is found inside the nucleus (the largest intercellular compartment) and mitochondria (another inner compartment) of every mammalian cell. It is composed of two complimentary chains that run in opposite directions (5'→3' pairs with 3'→5') and twist around each other to form a double helix (Figure 1-1). Each chain is composed of nucleotide subunits, which are respectively composed of a ribose sugar group, a phosphate group, and a base. The 5' end contains a free phosphate group while the 3' end contains a hydroxyl group.
Figure 1-1: A generic molecule of DNA. Two strands composed of nucleotide subunits run in opposite directions 5' → 3' and are paired through hydrogen bonds between the base pairs.

The nucleotides on one chain are paired through hydrogen bonds with the nucleotides on the opposing chain, which can be seen in the above diagram as the thin lines connecting the two chains. The nucleotides pair according to the Watson-Crick base pairing rules: nucleotides with adenine and thymine bases pair with each other leaving nucleotides with guanine and cytosine bases to pair together. All of the genetic information of an individual is contained in the sequence of the DNA, meaning that no two people have the exact same sequence. When it is time for a protein to be expressed, DNA must first be converted to ribonucleic acid (RNA).

There are two differences between the two compounds: the sugar group in RNA contains oxygen and the thymine base of DNA is replaced by uracil. The resulting product of the conversion from DNA to RNA is referred to as messenger RNA (mRNA), and the process is known as transcription. It is also important to note that mRNA is single stranded while DNA is double stranded. After transcription into mRNA, the protein can then be formed via ribosomes. Transfer RNA (tRNA) brings amino acids to the growing protein in accordance with the triplet code. The triplet code simply states that three consecutive RNA nucleotides form a codon, which codes for a specific amino acid. An amino acid is the basic building block of protein. Some amino acids are coded for by more than one codon, but two amino acids are never coded for by the same codon. This avoids confusion. The amino acids are then assembled together in a single chain to form a protein. This process is known as translation. When a protein is
"expressed," it has been successfully transcribed from DNA to RNA and then translated into a protein. It is therefore possible to down-regulate protein expression by either preventing transcription in the first place (DNA→mRNA) or by simply degrading the mRNA before it can be translated into protein.

RNA inhibition (RNAi) is a process in which protein is not expressed due to the degradation of mRNA. It is a highly conserved, naturally occurring pathway thought to have evolved to protect cells against viruses and other foreign invaders. RNAi is present in a wide range of eukaryotes ranging from invertebrates and fungi to plants and mammals (3). The process was first discovered in 1998 in Caenorhabditis elegans (nematode worms) (4). At first it appeared to simply be a biological anomaly, but it was soon recognized as a useful method of down-regulating specific proteins. Using C. elegans as a model, the RNAi pathway was elucidated, but to date, some details are still being worked out.

The generally accepted pathway (Figure 1-2) is composed of two main steps: the initiation step and the effector step. The process begins with the introduction of short hairpin RNA (shRNA) molecules complimentary to the mRNA of interest into the cell. The complementary shRNA is then incorporated into the Dicer protein complex, which processes the shRNA (500+ nucleotides long) into short dsRNA subunits (6). These shorter subunits 21-22 nucleotides long have come to be known as small interfering RNA (siRNA) because of their role in the pathway. Each siRNA is composed of a sense and an antisense strand. The sense strand has the exact same sequence as the mRNA to be degraded. It is also known as the coding strand, as this is the strand that is read by the tRNA when making protein. The antisense strand has the complementary sequence, meaning that it can bind with the target mRNA as described by the base pairing rules (adenine with uracil, thymine with guanine). The siRNA is next unwound, and
the antisense strand is incorporated into the RNAi silencing complex (RISC). RISC then uses the antisense strand to direct it to the complementary sequence on the target mRNA. Once found, RISC cleaves the mRNA into pieces thereby preventing protein translation (11). If synthetic siRNAs are introduced into the cell, they simply incorporate into the RISC complex, bypassing Dicer. While this pathway yields some protein down-regulation, for reasons still unknown it is not as effective as shRNA.

Figure 1-2: The pathway of RNAi. ShRNA is introduced into the cell and incorporated into the Dicer complex. Dicer uses ATP to cleave the strands into an siRNA duplex, which is then incorporated into the RISC-siRNA protein complex. The complex uses ATP to unwind the strands and uses the anti-sense strand to scan target mRNA for a match. Once a complementary segment of DNA is found, RISC cleaves the mRNA yielding it inactive thereby successfully stopping protein translation.

The RNAi process was successful in down-regulating protein expression in several experimental systems including plants, \textit{C. elegans}, \textit{Drosophila melanogaster}, and even human embryo cells. Problems arose, however, when this technique, while effective in other organisms,
was used in somatic (typical bodily) mammalian cells because gene silencing was not observed. In fact, the ultimate outcome of such experiments was apoptosis. This is due to dsRNA being an inducer of nonspecific, antiviral cell defense responses in mammalian somatic cells. Eventually it was found that dsRNA molecules shorter than 30 nucleotides and mimicking the siRNA molecules produced by the Dicer protein were short enough to escape the defense response, yet long enough to induce gene-specific silencing (3). A successful method to achieve RNAi in mammalian cells had thus been developed, but several downfalls still existed. The cost of the synthesized, specific siRNA molecules prevented many laboratories from utilizing this method and testing many different genes was still out of reach. In addition, the transiently transfected siRNAs only down-regulated protein expression for a limited time; reactivation normally occurred after a few days. This did not allow the effects to be studied for extended periods of time. Finally, the amount of down-regulation observed using the siRNA molecules was much less than the down-regulation observed with shRNA.

Researchers then developed a method to express the siRNAs in vivo (inside the cell) and overcome the problems of cost and reactivation of expression. The solution was a vector-based approach. Using this method, short, synthetically engineered nucleotides of DNA (oligonucleotides) complimentary to the target mRNA are ligated (inserted) into a vector (10). A vector, or plasmid, is a small, circular piece of DNA from bacteria that can be inserted into cells and expressed as if it were the cells’ own DNA.

Vector DNA is the method by which antibiotic resistance is passed between bacteria (Figure 1-3). Vectors contain a promoter, or region of DNA at which RNA transcription can begin. The RNA polymerase III (Pol III) promoter was chosen as the vector promoter for the pBS/U6 vector used in this research because termination occurs after a repeated stretch of 4-5
thymidines, which can easily be inserted into the shRNA. This easy, effective initiation and termination make it possible to use the DNA oligonucleotides to synthesize small RNAs similar in structure to the siRNA produced by Dicer protein (10). Vectors also contain a multiple cloning site (MCS) following the promoter, which contains several cut sites for restriction enzymes. This allows the vector to be cut and allows the oligonucleotide a place to insert. After the vector and oligonucleotide are ligated together, and proper orientation is confirmed by DNA sequencing, the plasmid can then be transfected into the cell.

![Diagram of DNA insertion into a vector](image.png)

**Figure 1-3: The method of DNA insertion into a vector.** A circular vector containing a promoter to start transcription followed by an MCS is cut with a restriction enzyme to linearize the DNA. The MCS is thus cut into two. The DNA to be inserted contains ends complementary to the ends of the vector that were just cut. This allows the new DNA to be ligated to the vector and close it up again resulting in a larger, circular vector that can then be transfected into bacteria.

Transfection is the process by which the plasmid, free in solution, enters the nucleus of the cells of interest. Here, the inserted sequence is transcribed into a short hairpin RNA (shRNA) that can then enter the Dicer complex, allowing siRNA to be made, thereby silencing expression of the target proteins. Since the vector enters into the nucleus it is not degraded like the siRNAs. This allows the possibility of tracking the effects of protein down-regulation into the progeny (offspring). The plasmid can also be transfected into bacterial cells and then harvested, enabling amplification of the vector thus keeping costs down. Another advantage of this method is that it typically yields a greater down-fold regulation than simply introducing siRNA molecules into the cell.
HMGAla and HMGAlb differ in 33 nucleotides, as previously stated. Since RNAi is specific in its down-regulation, hypothetically two plasmids could be made that would specifically target either HMGAla or HMGAlb. This would allow the differences in the proteins associated with HMGAla and HMGAlb to be determined. While siRNA specific to HMGAl has been made by a laboratory in Santa Cruz and used to down-regulate HMG protein expression in overexpressing cell lines, there currently is no method to selectively down-regulate only one of the two subscripts. One of the transgenic cell lines (HA7C) overexpresses HMGAla, the tumor cell line (HS578T) overexpresses both HMGAla and HMGAlb, and the parental cell lines (MCF-7 and HS578bst) hardly express either of the HMGAl proteins. This vector would also allow the effects of down-regulating the tumor line for an extended period of time to be determined. The goal of this thesis was to design two vectors that would specifically target the mRNA sequence that is different between HMGAla and HMGAlb to allow down-regulation of one protein at a time. While the HMGAla vector is almost complete, the HGMA1b vector has not been started due to time constraints.

**Materials and Methods:**

**Construction of oligonucleotide sequence**

The oligonucleotide sequence for the HMGAla vector was constructed by exploiting the region of mRNA that is spliced out in the translation of HMGAlb and incorporating key restriction enzyme sequences onto each side of the target sequence (Figure 2-1). Since HMGAla is not spliced during transcription, a region of nucleotides located just before the splice and continuing into the nucleotides that are spliced out was chosen as the siRNA sequence (Figure 2-2). This allows HMGAla to be preferentially inhibited, as this sequence is not present in HMGAlb. Likewise, to down-regulate HMGAlb a sequence beginning before the deletion
and continuing through the deletion was chosen (Figure 2-3). This sequence will not be present in HGMA1a, and should therefore selectively down-regulate HMGA1b.

**Figure 2-1: The difference in sequence between HMGA1a and HMGA1b.** The HMGA1a protein is 33 nucleotides longer. This corresponds to 11 aminio acids since 3 nucleotides are needed to code for 1 amino acid. The sequence that is spliced out in the transcription of HMGA1b is bolded in the HMGA1a sequence above and represented by an underline in the HMGA1b sequence. The P before the line in HMGA1b would be next to the K at the end of the line in the actual mRNA.

**Figure 2-2: The nucleotide sequence chosen to down-regulate HMGA1a expression.** The underlined nucleotides show the 21 nucleotide sequence chosen to be incorporated into the shRNA oligo. The bolded letters represent the nucleotides that would be cleaved out in the formation of HMGA1b. The sequence starts just before and continues into the portion that would be cleaved, ensuring that only HMGA1a should be affected.

**Figure 2-3: The nucleotide sequence chosen to down-regulate HMGA1b expression.** The underlined sequence will be ligated into the vector. The bolded letters are the area from which the deletion occurs. The deletion begins after the 'g' and ends before the 'a.' The sequence starts before the deletion and continues through to end with nucleotides from the other side of the deletion. As this sequence is not contained in HMGA1a, this should only down-regulate expression in HMGA1b.

Once the target nucleotide sequence was chosen, the entire sequence of the oligo was constructed. As the oligo sequence needed to form a hairpin inside the nucleus, it needed a 6
nucleotide spacer inserted at the end of the sequence of interest to serve as a means for the hairpin to fold back upon itself. A HindIII restriction enzyme cut site was incorporated into this spacer sequence as a marker for the hairpin structure. This cut site also served as a means to later select for colonies that contain the hairpin sequence insert by a simple HindIII restriction enzyme digestion of DNA isolated from various colonies. The reverse complement of the target sequence was then constructed. This involved beginning at the last base in the sequence and constructing a new sequence based on the original sequence compliments (according to Watson-Crick base pairing rules). At the end of the inverted repeat, a polyT (many thymine) tail consisting of 5 thymine residues was added to the 3’ end to terminate the PolIII reaction as previously explained. The ends of the hairpin were then made to look as though they had already been cut by restriction enzymes chosen from the plasmid multiple cloning site. Restriction enzymes cut at specific locations in the DNA. The enzyme recognizes a specific sequence of typically 6 nucleotides and always cuts between the same two nucleotides in the sequence (Figure 2-4).

Figure 2-4: Restriction enzyme recognition sequence and cut site. Restriction enzymes typically recognize a 6 nucleotide sequence in the DNA and cut somewhere in the middle. Both XhoI and EcoRI leave ‘sticky’ ends that can easily religate through complementarity either to a new piece of DNA or back to itself. The vertical lines indicate between which two nucleotides the enzymes cut and the horizontal line indicates the overhang that will be left on each end.

This allowed the synthetic oligo hairpin cut with restriction enzymes to ligate into the
digested plasmid vector. For example, The C to the left of the cut site on the top line of XhoI would be left in the vector while the TCGAG would be incorporated into the oligo. A second oligo was made through complementarity to the first, apart from the restriction enzyme sequences. This provided a double stranded compliment to anneal to the first, which was necessary for the oligo to ligate into the double stranded plasmid. The restriction enzyme cut sites in the second strand were also made to look as though the enzymes had already digested them. For XhoI, the GAGCT would be left in the vector while the C would be incorporated onto the end of the oligo. The general structure of the oligonucleotide had now been designed (Figure 2-5).

![Diagram of shRNA oligonucleotide](image)

**Figure 2-5: General structure of an shRNA oligonucleotide.** The ends of the shRNA are complementary to the overhang ends left on the vector so that the shRNA can ligate with the linear vector DNA and reform the circular vector. The shRNA then contains a 21 nucleotide siRNA sequence and its inverted complementary DNA. Finally, the 6 nucleotide spacer sequence containing a restriction enzyme cut site is incorporated to allow the shRNA to fold into structure and allow a means of distinguishing the clones. The vertical lines indicate where the shRNA will be cut by Dicer to form the siRNA duplex.

It is important to note that while the diagram appears only single stranded, the actual hairpin is double stranded DNA that will encode RNA. This double-strandedness again allows it to incorporate into the double stranded plasmid vector. This hairpin structure ended up being 59 base pairs long. The company from which the oligos were ordered could not synthesize a 59
base pair oligo, so each hairpin sequence was separated in the HinDIII cut site into 27 and 32 base pair oligos (Figure 2-6). Upon arrival, these oligos were first annealed together to form two short double stranded molecules. Since the oligos were synthetically produced, no phosphates were present on the ends of the molecules. Phosphates are necessary to form phosphodiester bonds, which are the types of bonds between nucleotides. The short double stranded molecules were then treated with T4 kinase enzyme, which uses ATP (a form of energy that contains phosphates) to put phosphates onto the ends of the molecules. The oligos were then ligated together. Unlike annealing, which occurs spontaneously, forming phosphodiester bonds through a ligation reaction requires T4 DNA ligase.

Figure 2-6: The process of forming the shRNA. shRNA molecules 1 and 3 contain the end complementary to XhoI and were annealed together while molecules 2 and 4 containing the end complementary to EcoRI were annealed together in a separate reaction. The 27 and 32 base pair double stranded molecules were then ligated together. The area ligated together on the molecules contains 6 nucleotides that correspond to the sequence that HindIII recognizes. This formed a 59 base pair double stranded molecule that was ready to be inserted into the prepared vector. The actual nucleotide sequence is also shown.
The double stranded oligo with phosphates on the ends was now ready to be inserted into the pBS/U6 vector once the vector prepared.

**Construction of pBS/U6+oligo vector**

The pBS/U6 and pBS/U6-GFP plasmids were obtained from the laboratory of Dr. Guangchao Sui at Harvard Medical School (Figure 2-7). There were many restriction enzyme sites following the promoter in this plasmid, but EcoRI and XhoI were chosen since they were both available in the lab. These were therefore the restriction enzyme cut sites that were incorporated onto the ends of the oligo design previously described (Figures 2.5 and 2.6).

![pBS/U6 vector](image)

Figure 2-7: **pBS/U6 vector map.** The vector map of pBS/U6 shows the promoter region in dark blue. A few of the restriction enzyme cut sites in the multiple cloning site can then be seen directly following the arrow. Many more cut sites are located in the MCS than are stated on the map above. For a complete list, refer to the Ambion website where the vector is now sold as pSilencer 1.0 (1).

Since only a small amount of the vector was sent, the first step was to amplify the vector. Amplification consists of transforming (incorporating) the vector into bacteria, growing the bacteria on plates with antibiotics, and finally harvesting the vectors from the bacteria. The
bacterial cells used in this process were calcium chloride treated DH5α cells. Treating the bacteria with calcium chloride in the early log phase of growth makes the cell membrane permeable to plasmids of DNA in a process that is still being worked out. Somehow the extreme changes in temperature between hot and cold cause competent cells to uptake the plasmids. The plasmid being transfected typically contained an antibiotic resistance gene. This allowed the cells that had taken in the plasmid to be selected for by growing the cells in media containing that antibiotic. pBS/U6 contains ampicillin resistance. As the bacterial cells replicated and multiplied, the plasmid was replicated and multiplied as well. The plates were incubated overnight and the cells were harvested the next day. Harvesting consisted of lysing (breaking) the bacterial cells open and then separating the plasmid DNA from the chromosomal DNA, RNA, protein and cell debris.

There are many protocols to harvest plasmids from the bacterial cells. The first method employed in this project was to use a Mini-Prep® kit from QIAGEN®. The kits reduced the amount of time needed as the solutions are pre-made, however, sometimes the kits do not give the best yield. After several attempts, the best yield obtained was ~1µg/µL which was unacceptable for the uses of this project. Another method was then used in which the solutions were all made by hand. The first solution contained 25mM Tris buffer, pH 8, 10mM EDTA, 1% glucose, and lysozyme to break the cells open. The next solution contained 0.2N sodium hydroxide and 1% SDS. This solution caused the double stranded linear chromosomal DNA to denature by breaking the hydrogen bonds between the two strands while leaving the circular double stranded plasmid DNA untouched. The next solution contained 3M sodium acetate pH 4.8. This caused the chromosomal DNA to randomly renature, which forms a clump with the cell debris. The clump was then centrifuged out of the solution. Centrifugation involves
spinning the samples at thousands of times the force of gravity in a machine called a centrifuge. This caused molecules to migrate to the bottom of the tube due to size and density. The larger, heavier molecules centrifuged out of solution before the smaller ones, and the longer the solution is centrifuged the more molecules fall out of solution. The DNA-cell debris clumps were therefore centrifuged out while the vector DNA was left in solution. A phenol extraction was then performed to precipitate the protein out of solution. Finally, 70% ethanol was added to precipitate the plasmid DNA and concentrate it. The plasmid DNA was then resuspended in TE buffer with RNAse to break up any RNA left in solution. This method yielded 11.35\mu g/\mu L of pBS/U6.

To ensure that pBS/U6 DNA had been obtained, a series of restriction enzyme digests were carried out and then run on a 1% agarose gel. The agarose gel provided a matrix that the DNA were forced to travel through. The gel was placed in the gel box filled with TE buffer. The DNA was loaded into wells at the top of the gel and exposed to an electric field with 70 to 100V of electricity running through (Figure 2-8). The electric current forced the negatively charged DNA (due to the phosphates in the backbone) to travel toward the positively charged cathode. Since the samples were forced to travel through the meshwork of agarose, the gel separated the DNA based on size. The shorter DNA traveled further down the gel than the longer DNA. This is logical; it is easier for a small child to fit through a small maze than a 7-foot tall basketball player.
After the identity of the plasmid DNA was verified, the plasmid DNA was prepared for cloning. The plasmid was digested with both EcoRI and XhoI and then treated with shrimp alkaline phosphatase to take the phosphates off the ends. This prevented the plasmid from simply closing back upon itself and increased the chance of ligation success. This resulted in a plasmid with overhanging ends that correlated to the overhanging ends of the doublestranded ligated oligo already prepared. Since the oligo had been treated with a kinase, it had phosphates on the ends and was therefore able to form a phosphodiester bond with the corresponding vector.

For the ligation reaction, twice as much oligo DNA as vector DNA was incubated in ligation buffer, T4 DNA ligase, and water. The reason for incorporating more oligo DNA than vector DNA was to increase the change of having the oligo insert into the vector. After the ligation reaction, the newly formed plasmid was digested with EcoRI, XhoI, and EcoRI+XhoI. The EcoRI + XhoI digest should have been the reaction to confirm whether or not the oligo had inserted. This particular digest should cut the 59 base pair oligo out of the vector resulting in two bands, a larger vector (~3150bp) and a smaller oligo band (~59bp). DNA is visualized on an agarose gel by intercalating ethidium bromide between the nucleotide bases. In addition, pBS/U6 uncut, pBS/U6 digested with AlwNI, and pBS/U6 digested with AlwNI + HindIII were also ran as controls. These final tests show that the enzymes are cutting, that AlwNI is a cut site
in the plasmid, and that cutting with HindIII and AlwNI result in two separate larger bands. It is important to note that the HindIII site in pBS/U6 is not the same as the one in the ligated vector. The HindIII site was cut out in preparing the vector and then reincorporated into the vector through the oligo. HindIII should not cut unless the oligo had incorporated.

After a sample passed this test, it was then sequenced. Sequencing involved using a PCR thermocycler to create different length tagged copies of the vector. In the reaction, normal adenine, guanine, cytosine, and thymine were added in order to extend the DNA molecules. In a much smaller amount, fluorescently tagged adenine, guanine, cytosine, and thymine were also included. In addition to the nucleotides, synthase enzyme to extend the DNA, as well as vector DNA and template DNA were added. The template DNA served as a starting point for the elongation reaction. Nucleotides cannot begin to synthesize DNA without a free 3' nucleotide already annealed to the template DNA. The primer was designed to be complementary to a portion of the vector and would therefore anneal to provide a free 3' end for elongation. The reaction was then placed in the thermocycler. The time in the thermocycler was composed of 96°C for 10 seconds, followed by 50 °C for 15 seconds, and finally 60 °C for 4 minutes. This repeated for 25 cycles. 96 °C caused the double stranded DNA to denature into single stranded copies. 50 °C caused reannealling, which is when the complementary primers annealed to the single stranded DNA pieces. Finally, 60 °C caused elongation and was when free nucleotides (adenine, guanine, cytosine, and thymine) were incorporated onto the free 3' ends of the primer. In this manner, many copies of the DNA were made. Every so often a fluorescent tagged nucleotide was incorporated into the elongating DNA. This tagged nucleotide stopped elongation, but also provided a way to scan and see which nucleotide was present (each nucleotide fluoresces at a different wavelength). Therefore, after the thermocycler had
completed all 25 cycles, the samples were sent down to the LBB1 sequencing laboratory where they were run through a laser beam able to detect which nucleotide had incorporated in each position for the entire vector DNA sequence. The sequence from the laboratory was then compared to the theoretical sequence that should be obtained if the oligo had correctly inserted.

**Cell culture**

Before vectors or siRNA could be tested, mammalian tissue culture cells needed to be grown. The adherent cells were grown to a confluency of between 60 and 70 percent. This means that 60-70% of the plate was covered in cells when looked at under a microscope. The cell lines being compared were MCF-7 (regularly expressing cells) and HA7C (overexpressing cells). The cells were grown in 10% fetal bovine serum in Dulbecco’s Modification of Eagle Medium (DMEM) as this is the optimal media for these cell lines. In addition, penicillin, streptomycin, and L-glutamine were added to control contamination and supplement the media. Twenty-four hours before the transfection, cells were washed with phosphate buffered saline (PBS) to ensure that all antibiotic was washed away, as no antibiotic should be present during transfections. The media was then replaced and the cells incubated without antibiotic or serum.

**Transfection**

The first transfection carried out was 40nmol siRNA from Santa Cruz Laboratory©. The siRNA transfection was carried out in order to compare the down-regulation efficiency of a high concentration of siRNA to the efficiency of a lower concentration that was previously carried out by PhD candidate Jennifer Adair. To transfec the mammalian cells, 40nmol of siRNA was diluted into serum free DMEM media. Lipofectamine was also diluted into serum free DMEM media and allowed to sit at room temperature 5 minutes. Lipofectamine is a reagent that facilitates in the delivery of the siRNA molecules into the mammalian cells with minimal levels
of cell death. The two solutions were then combined and allowed to incubate 20 minutes at room
temperature. Serum free DMEM was then added to the solution and applied to the cells. The
cells were incubated 5 hours at which time 20% serum media without antibiotic was added. The
cells were then allowed to incubate 48-72 hours. The plates were then washed again with PBS,
and were harvested.

While the idea behind the harvesting of mammalian cells is the same as bacterial cells,
the protocol used was slightly different since protein and not plasmid DNA was needed.
TRIzol® reagent from Invitrogen was used to isolate the protein in the cells. TRIzol® provides
a way to separate protein from DNA and RNA. TRIzol® was added to the cells and the cells
were then scraped off the plate and forced through a syringe to break up clumps. The cells were
frozen in the −70°C freezer and then allowed to thaw. Chloroform was added, and after mixing
and centrifugation the solution separated into a lower organic phase, an interface, and an upper
aqueous phase (Figure 2-9). This is the same separation that occurs when oil and water mix; the
aqueous water layer sits on top of the oil organic layer and the two will not mix.

The RNA remained in the aqueous phase while the DNA and proteins moved to the
organic layer as well as to the interface between the aqueous and organic layers. The DNA was
precipitated out using 100% ethanol and centrifugation. The supernatent that remained after the
DNA had precipitated contained the protein. Isopropyl alcohol was used to precipitate the protein from the supernatent. After centrifugation, the pellet was washed three times with 0.3M guanidine hydrochloride in 99% ethanol. After the final wash, the pellet was vortexed (shaken) in ethanol and centrifuged. The pellet was finally dissolved in SDS through a series of pipettings and incubations. The protein was then ready for the Western blot procedure.

**Western Blot**

To perform a western blot, a Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis (SDS-PAGE) gel is run and transferred to a nitrocellulose paper. The protein of interest is labeled with a specific antibody and then detected (typically through fluorescence). Through this method, HMG A protein can be detected and the amount of down-regulation can be determined by comparing transfected and non-transfected plates.

To begin, an SDS-PAGE gel was poured. This gel was run in a unit similar to the agarose gel described earlier, but consisted of a lower 12% acrylamide gel with a 15% acrylamide gel on top. The sodium dodecyl sulfate (SDS) in the gel caused the proteins to denature. This made the proteins linear and ensured that the folded native structures of the proteins would not give a false size reading when run on the gel. In addition, the SDS put negative charges onto the proteins. This was essential, since the gel was run in an electrical field toward a positive cathode. If the proteins were not coated in negative charges, a protein with extra positive charges would appear larger than a regular protein of similar size because the cathode would repel the positive charges on the protein. This step is not necessary with DNA, as DNA already has equally distributed phosphates with negative charges. The upper 15% gel is referred to as the stacking gel while the lower 12% gel is referred to as the separating gel. The stacking gel served to compact the protein loaded into the well. This ensured that all proteins,
whether occupying the top or the bottom of the well upon being loaded, were exposed to the electrical current for an equal amount of time. After the proteins were stacked, they entered the 12% separating gel where they were separated based on size. Since the proteins were linear and coated with negative charges they migrated to the positive charge based on size alone. The smaller proteins were able to move easier through the acrylamide gel matrix than larger proteins and were therefore found at the bottom of the gel. A ladder containing many proteins of known sizes was run in a well of the gel as well. Since the size of the protein of interest (HMGA) was known, it was easy to determine where on the gel HMGA was located. The gel was run at 200V for 60 minutes.

Next, the protein transfer was set up. A sponge soaked in buffer was placed down onto the transfer apparatus. Filter paper cut to the exact dimensions of the gel were then placed on top of the sponge followed by the gel (face down), the membrane (to which the protein will be transferred), more filter paper, and a second sponge (Figure 2-10). It was necessary to remove all bubbles between the gel and the membrane to ensure complete transfer of protein. The apparatus was then closed and placed in the gel box casing surrounded by ice. During the transfer, the proteins again migrated toward the positive charge (Figure 2-11). This time, however, the positive charge pulled the proteins from the gel into the membrane instead of simply pulling them down the gel (the gel was vertical instead of horizontal as in the previous gels). The transfer was run at 100V for 60 minutes.
Figure 2-10: Western Blot transfer setup. The setup consists of a sponge onto which filter paper is placed. The gel is then placed face down onto the filter paper and the nitrocellulose membrane placed directly on the gel. Filter paper and another sponge are placed on the membrane, completing the setup.

Figure 2-11: Transfer gel box. The transfer gel box differs from the agarose gel box because the box sits upright and the current is passed in a different plane. The gel is placed closest to the negative charge and the membrane closer to the positive charge allowing the proteins to transfer from the agarose gel to the membrane (9).

Next, the apparatus was disassembled and the gel was stained in a solution consisting of Coomassie brilliant blue dye, acetic acid, isopropanol, and water. The stained gel was used to compare to the blotted gel once the antibody had been recognized. The membrane was soaked in a blocking buffer consisting of milk and TBS/tween. The nitrocellulose membrane attracted proteins, which was why it accepted the proteins during the transfer. When labeling with an antibody, however, if the empty areas of the gel were not blocked with a generic protein (milk), the antibody would have adhered everywhere and the labeled protein of interest would not have been detected. After the membrane was blocked, it was incubated overnight with the primary antibody (MR19; rabbit anti-human HMGA1), which specifically recognized and bound to the protein of interest. The membrane was then washed with TBS/tween to wash away any primary
antibody that was not specifically bound to HMGA. Next, the membrane was washed with the secondary antibody (goat anti-rabbit conjugated to horseradish peroxidase (HRP)), which recognized and bound to the primary antibody. In addition, it contained HRP, which luminesced upon binding. This allowed the protein, primary antibody, secondary antibody complex to be identified on films. Films were then exposed in a dark room at various times and developed. The amount of protein corresponded to the amount of primary antibody and secondary antibody, which corresponded to the amount of light given off and recorded as a dark spot on the developed film (Figure 2-12).

![Diagram of antibody detection complex](image)

Figure 2-12: Antibody detection complex. After the protein of interest is transferred to the membrane, the membrane is blocked with a non-specific protein. The primary antibody that recognizes the protein of interest is then applied, and the excess washed away. A secondary antibody that recognizes the primary antibody is then applied and contains a means of detection, in this case luminescing. The luminescence can then be detected on film paper.

The films were then scanned into a computer and evaluated with a densitometer. The densitometer turned the dark spots on the film into numbers that could then be quantified. This made it possible to determine the exact amount of protein down-regulation due to transfection by comparing the different samples.
Discussion and Results:

The 27 base pair and 32 base pair oligos were annealed together and then ligated together to form a 59 base pair molecule. To ensure that a 59 base pair molecule had actually been produced, the sample was run on an agarose gel (Figure 3-1). If the sample had not ligated, there would have been a 27 base pair molecule and a 32 base pair molecule. These would have shown up below the 50 base pair band in the ladder. Since the molecule ran just above the 50 base pair band, it was concluded that the molecules ligated and the oligo had correctly assembled. The pBS/U6 vector then needed to be prepared so that the vector and oligo could be ligated together.

![Figure 3-1: Confirming oligo ligation.](image)

Lane 1 contains a Hi-Lo molecular weight ladder. Lanes 2 and 3 both contain samples of annealed, ligated oligo that should be 59bp in length. As the bands are above the 50bp band in the marker, the oligo is assumed to have correctly ligated.

The identity of the pBS/U6 vector needed to be determined before it could be prepared for ligation. This was accomplished through a series of restriction enzyme digests of EcoRI, XhoI, EcoRI+XhoI, BamHI, XbaI, and HindIII (Figure 3-2). EcoRI, XhoI, and HindIII all had one restriction site in the plasmid and therefore resulted in one band of ~3200 base pairs. EcoRI and XhoI together cut two times in the vector, but are only 34 base pairs apart. This small difference was not seen on the gel. Finally, BamHI and XbaI cut the vector twice resulting in bands of ~400 base pairs and ~ 2800 base pairs. In addition, uncut pBS/U6 plasmid was run to ensure that the enzymes were cutting. Vector DNA is naturally supercoiled meaning that it is
twisted upon itself many times over resulting in a tiny ball. This ball is able to pass through the matrix easier than the linearized DNA. The ladder contained linearized DNA, so the uncut pBS/U6 vector DNA ran to a size that appeared smaller according to the ladder than it actually was. From the gel, it was concluded that pBS/U6 vector had been obtained because all of the expected bands showed up on the gel.

Figure 3-2: Confirming vector identity. A Hi-Lo molecular weight ladder was run to confirm the size of the vector digestion bands. Uncut pBS/U6 was run to ensure that the restriction enzymes were indeed cutting. This supercoiled DNA appears to be 2500 bp according to the ladder, but upon linearization it is obvious that the vector is ~3200 bp. EcoRI, XhoI and HindIII each cut once and therefore yield a linear DNA molecule ~3200 bp long. EcoRI and XhoI cut twice, but since the sites are so close, a significant difference is not apparent on the gel. Finally, BamHI and XbaI each cut the gel twice yielding bands of ~400bp (which is hard to visualize but is contained in the box) and ~2800 bp which are shifted down from the EcoRI and XhoI bands.

After the identity was confirmed, the vector was digested with XhoI and EcoRI, treated with shrimp alkaline phosphatase, and ligated to the prepared oligo. The ligated vector was then treated with XhoI, EcoRI, and EcoRI + XhoI. The last digestion should have cut out the 59 base pair insert resulting in two bands. Unfortunately, the XhoI+EcoRI digest did not result in two
bands (Figure 3-3). This could possibly have been due to the small DNA band not intercalating enough ethidium bromide to be visualized on the gel. To ensure that the absence of the band was not due to an inability to visualize the band, another set of restriction enzyme digestes were set up.

![Figure 3-3]

Figure 3-3: Confirming clone identity I. Three colonies that grew on ampicillin plates were tested through a series of restriction enzyme digestions. Lane 1 contains the Hi-Lo molecular weight ladder, Lanes 2, 6, and 10 contain EcoRI digested vector DNA, lanes 3, 7, and 11 contain XhoI digested DNA, lanes 4, 8, and 12 contain XhoI and EcoRI digested DNA, and lanes 5, 9, and 13 contain uncut vector DNA. Lanes 2-5 contain DNA from the same colony, lanes 6-9 contain DNA from second colony, and lanes 10-13 contain DNA from a third colony. A band should have shown up in lanes 4, 8, or 12 around 50bp if the insert had been cut out but it was too small to visualize. Some of the lanes contain 2 smaller bands that correspond to knicked DNA (the highest band) and supercoiled DNA (the lowest band) due to incomplete digestion.

In the next set of digestions, AlwNI and HindIII were used (Figure 3-4). HindIII should not be present unless the oligo had ligated into the vector. AlwNI cuts away from the multiple cloning site, which should have resulted in two bands that both would be easy to visualize. Six different colonies that grew under ampicillin selection were tested. In addition, uncut pBS/U6, pBS/U6 digested with AlwNI, and pBS/U6 digested with AlwNI and HindIII were run on the gel as controls. HindIII should not cut the plasmid if the oligo had not incorporated. As can be seen
below, it appears that all but one of the colonies tested contained cloned plasmid with oligo DNA.

![Figure 3-4: Confirming clone identity II](image)

Figure 3-4: **Confirming clone identity II.** Uncut vector DNA was run in lane 3 to ensure the enzymes were cutting. The DNA of six different colonies was digested with AlwNI and HindIII and run in lanes 4-6. According to the vector map, this would result in two bands of ~2300bp and ~900bp. This was the case for all but one of the colonies. Since a HindIII cut site should not be present unless the oligo ligated in, it was concluded that the oligo had ligated into 5 of the 6 colonies. DNA from lanes 8 and 9 was then sequenced. Lane 10 contains pBS/U6 DNA digested with AlwNI, lane 11 contains pBS/U6 DNA digested with HindIII, and lane 12 contains pBS/U6 DNA digested with both AlwNI and HindIII to see what the digestion would look like on pBS/U6 DNA not containing the 59bp insert.

Since the restriction enzyme digests suggested that the vector was likely cloned and isolated, the DNA from two of the colonies was then sequenced. A T7 primer was used, since this is the primer advised by the manufacturer. It took several attempts for the sequencing reaction to finally work. From the sequencing reaction, a printout was obtained that contained colored peaks corresponding to each nucleotide base. In the area where the insert should have been, all of the restriction enzyme sequences were found (EcoRI, HindIII, and XhoI) but the oligo sequence did not match up. In addition, an incorrect number of bases were found. Upon further examination, it was found that the sequenced vector matched up exactly with the pBS/U6 sequence (Figure 3-5). During the transfection of the vector that was hopefully ligated into the DH5α bacterial cells, not all of the vectors contained insert. The two vectors that were sequenced were plain pBS/U6 vectors without the insert.
Figure 3-5: Sequencing the possible pBS/U6 clone. The clone was sequenced and the colorful bottom portion of the figure was obtained. The arches on the bottom correspond to light given off by the tagged nucleotides and determine which nucleotide is where. When compared to the predicted sequence above in black, some similarities as well as differences were found. Sequences with a '*' signify similarities in sequence. First, The XhoI sequence (CTCGAG) can be seen corresponding at the beginning of both sequences, and the EcoRI cut site (GAATTC) can be seen toward the end of both sequences. In addition, the HindIII cut site can be seen in the middle (AAGCTT). Unfortunately, the oligo sequences and sizes do not match up, which is designated by the boxes.

For the last part of the project, siRNA was transfected into cells at a concentration of 40nmol. Jennifer Adair, a PhD candidate in the Reeves laboratory, had already used the siRNA in a transfection of 20nmol and had seen indiscriminant down-regulation of both HMGA1a and HMGA1b of 2-fold. The goal of this experiment was to determine if an increased concentration gave greater down-regulation. In addition, this experiment gave another set of data with which to compare the vector down-regulation data once the cloned vector was obtained. It was predicted that doubling the concentration would double the down-regulation in a linear relationship. This was not observed. siRNA transfected at 40nmol yielded 2.5 to 3 fold down-regulation. This suggests that the 40nmol concentration is approaching the maximum down-regulation that can be achieved using siRNA in this cell line. This would mean that a hyperbolic relationship exists between concentration and down-regulation. This is emphasized when the data is extrapolated past the 40nmol concentration and compared to a linear relationship (Figure 3-6). Since the HMGA1 proteins are overexpressed up to 20 times the normal value in some of
the cell lines, this is not acceptable. It is anticipated that the vector, once obtained, will yield an even greater down-regulation. This shows the necessity of the vector.

Before the fold down-regulation was calculated, the data was normalized against H1, a house-keeping gene that is expressed and is not affected by HMG protein concentrations. This normalization accounts for the possible difference in down-regulation being due to changes in transcription levels between the four different populations of cells tested (transfected and non-transfected for over-expressing and parental lines). For example, if the non-transfected over-expressing cell population had more transcription and translation than the transfected population, it would appear as though the transfection was more successful than it actually was. By normalizing the data, this discrepancy is eliminated.

![Observed and Predicted down-regulation from siRNA transfection](image)

**Figure 3-6: Observed and Predicted siRNA down-regulation.** If a linear relationship existed between concentration and down-regulation, the graph would be similar to the blue line (predicted down-regulation). The data already obtained is shown on the pink line up to 40nmol concentration. The 60nmol and 80nmol data on the pink line was extrapolated to emphasize that the siRNA down-regulation is reaching a maximum.

To give another reason to obtain this vector, Jennifer Adair performed survival assays on the siRNA transfected cells. HMGA1a over-expressing cells and parental cells were both transfected with siRNA and then irradiated with UV light. As expected, the parental cell
survival rate did not change with transfection of siRNA. The survival of the over-expressing cells was less than that of the parental cells for non-transfected cells, but was better than that of the parental cells for the transfected cells. While this phenomena cannot be explained, it suggests that there is a future to the down-regulation of HMG proteins leading to the return of normal cell function.

Conclusions:

While it is frustrating that the vector was not obtained, all was not lost. A protocol for cloning has been set up for the next student who continues with this project. Ample amounts of sample material have been stored in the freezer at all points along the way. This will save the next student a lot of time in preparing materials. In addition, it was found that while there are many methods to select for the cloned vector, these do not ensure that the clone has been obtained. The growth on ampicillin, restriction enzyme digest with HindIII and AlwNI, and restriction digest with XhoI and EcoRI will all result in bands of similar size whether the insert is present or not. The ampicillin plates will grow colonies if any type of pBS/U6 vector finds its way into the bacterial cells. The EcoRI, XhoI restriction enzyme digest (while difficult to visualize the small bands) will result in a 40 base pair band if the oligo does not insert and a 59 base pair band if the oligo is present which will be difficult to distinguish on a gel. Finally, as the HindIII site is present in the normal pBS/U6 vector, if the vector escapes being cut during preparation with XhoI and EcoRI, the HindIII site will still be present. This conclusion can possibly lead to an explanation of Lane 5 in Figure 3-4. The vector may have had the HindIII site cut out, but then ligated back together without an insert. While this digest suggests that the insert was not contained in that sample, it does not say for sure whether the other samples
contain an insert either. The only way to be sure the vector contains the insert is to sequence colonies that pass the three tests described above until a clone is found.

Although the vector was not obtained, the project was still a success. During the year of research that went into obtaining these results, a procedure has been set up and mistakes along the way have been worked out. The next student to take over this project should have a much easier time and should be able to obtain these much-needed vectors. Once obtained, they will be very useful in the field of HMGAI protein research, as far as studying the effects on cancer, and the usefulness of these proteins as potential anti-cancer drug targets.

**Future Directions:**

The procedures described need to be repeated and colonies need to continue to be tested. Once the correct vector identity is obtained by sequencing through future work of those following, it will be transfected into the mammalian HMGAIa and HMGAIb over-expressing cells. If the vector down-regulates HMGAI protein expression, it is anticipated that the expression will be even more dramatic than simply introducing the siRNAs. Since it is already known that the siRNAs work to a certain extent, this is very exciting. Unlike the siRNAs, the suppressed protein expression will continue and should be passed on to the progeny as well. This will allow the effects of the suppression of HMGAI expression in the cell to be studied for several generations. It will also determine if the suppression of HMGAI in cancer cells can lead to a change in oncogenic cells back to regularly dividing cells. This research could lead to potential targets for anti-cancer drugs. Finally, the vector should have not affect the HMGAIb protein in the cells, as the vector was designed to specifically target HGMA1a. At this point, the second vector will be developed following the same procedure and using the sequence in Figure 2-3. This will give a way to differentiate between HMGAIa and HMGAIb, and will allow the
role of the two closely related proteins to be deciphered. This vector has the potential to be very useful once obtained, which demonstrates the significance of this research.
References:


REQUEST FOR PERMISSION TO INCLUDE YOUR HONORS STUDENT'S THESIS IN THE WSU RESEARCH EXCHANGE
Washington State University Libraries

Permission is requested for a non-exclusive license to post the Honors thesis described below in digital form in the Honors College community within the WSU Research Exchange <https://research.wsulibs.wsu.edu:8443/dspace/community>. Posting in the Research Exchange will make the material publicly available as part of the Washington State University Research Exchange digital repository of research-related documents. Additional information about Research Exchange can be viewed at <http://research.wsulibs.wsu.edu>.

Thesis Author                  Martin, Patty

Thesis Title                  Developing an RNAi Based Approach to Silence the Expression of HMGA1 Proteins in Human Mammalian Epithelial Cells

Date of Thesis               Spring 2005

X I grant a non-exclusive right to include this item in the Research Exchange. All other rights under copyright law are retained.

Permission granted by:

Raymond Reeves / reevesr@wsu.edu
Name (please print)                      Telephone / Email

Signature (not computer generated)       Date of Signature

Please return the form with your written signature via fax, email, or postal service to:

Kay Vyhnanek, Scholarly Communication Librarian
120F Terrell Library / PO Box 645610
Washington State University
Pullman, WA 99164-5610 USA
Fax: +1-509-335-672