2013

The cloning and expression of histone genes from the archaeal organisms Pyrococcus furiosus and Methanobrevibacter smithii in the polycistronic vector pST44.

Doha Abdullah Alqurashi
Pittsburg State University

THE CLONING AND EXPRESSION OF HISTONE GENES FROM THE ARCHAEAL ORGANISMS *PYROCOCCUS FURIOSUS* AND *METHANOBREVIBACTER SMITHII* IN THE POLYCISTRONIC VECTOR pST44

A Thesis Submitted to the Graduate School
in Partial Fulfillment of the Requirements
for the Degree of
Master of Science

Doha Abdullah Alqurashi

Pittsburg State University

Pittsburg, Kansas

December 2013

THE CLONING AND EXPRESSION OF HISTONE GENES FROM THE ARCHAEAL ORGANISMS *PYROCOCCUS FURIOSUS* AND *METHANOBREVIBACTER SMITHII* IN THE POLYCISTRONIC VECTOR pST44

Doha Abdullah Alqurashi

APPROVED:

Thesis Advisor
_____
Dr. Irene Zegar, Department of Chemistry

Committee Member
_____
Dr. James McAfee, Department of Chemistry

Committee Member
_____
Dr. Khamis Siam, Department of Chemistry

Committee Member
_____
Dr. Xiaolu Wu, Department of Biology

ACKNOWLEDGMENTS

First of all, I would start by thanking God for giving me the strength to accomplish this project in order to graduate. God graces me with great countless blessings that made me a person of ability and strength to achieve what I set my mind on it. Secondly, I am thankful to my great family for their continued support, love, encourage and being beside me all the way to the end of my degree path. To my mother Kawthar Hazazi for her emotional support and unstoppable love that gave me the hope to complete my thesis and make her proud. Moreover, I am grateful to my dear father Abdullah Alqurashi for his encourage and inspiration that had a great influence on completing my thesis. Furthermore, I would like to appreciate and thank my brothers Naher, Bader and Taher Alqurashi for being beside me all the time and aim me to make it through, without my dear brothers I would not have gotten the self-confident to come to an end to my thesis. This project might not come alive without my dear husband Yaseen Alaidarous who stood beside me through the heavy storms and show me his great love and support that gave a hope to get steps forward toward my ultimate degree.

In addition, I would like to express my gratitude to my wonderful friends who were more of a family to me while I am on the other side of earth from where my family stays, there were the most influential people in my life in my academic success. Also, I would like to acknowledge my thesis advice Dr. James McAfee for his aim this thesis with his limitless knowledge that contributes greatly to this project. A great thank you to Dr. Irene Zegar for lighting my path through her great help and for her appreciated trust to

experimenting and searching to complete this thesis successfully. The thesis would not be written without Dr. Khamis Siam whom I dedicate my grateful to, for being for provide his knowledge and support that was needed so my thesis could come to an end nicely. Very special thanks go to Dr. Xialo Wu for accepting my request to be one of the committee members.

THE CLONING AND EXPRESSION OF HISTONE GENES FROM THE ARCHAEAL ORGANISMS *PYROCOCCUS FURIOSUS* AND *METHANOBREVIBACTER SMITHII* IN THE POLYCISTRONIC VECTOR pST44

An Abstract of the Thesis by

Doha Alqurashi

The Euryarcheota branch of the Archaea contains histone proteins that are highly homologous to eukaryotic histones. Virtually all species within this group have two histone proteins (designated H1A and H2B). These proteins contain a canonical histone fold motif that stabilizes homo- and perhaps hetero-dimeric interactions. Molecular modeling calculations previously carried out in our laboratory revealed that both homo- and hetero-dimers of histones from two euryarchaeal organisms are equally plausible. We have also shown through molecular modeling that inter-species hetero-dimers are structurally identical to existing solution structures archaeal histones. An interesting question that needs to be addressed is whether homo- and hetero-dimeric species exist in the living cell and if they do, whether they are both physiologically functional. To address these questions we cloned the H2A and the H1A genes from *Methanobrevibacter smithii* (MS-H2A) and *Pyrococcous furiosus* (PF-H1A) in the polycistronic vector, pST44. When cloned individually in the pST44 vector, *Escherichia coli* BL21 (DE3) PLysS transformants were viable while *E. coli* BL21 (DE3) RIPL were not. However, when both genes were placed in pST44 and expressed simultaneously neither strain survived. Since expression of both genes is under the control of the *LacI* repressor

increased protein expression should not be significant enough to account for cell death. It is more reasonable to suggest that hetero-dimer formation has led to the formation of a functionally distinct protein that is lethal in insignificant amounts (i.e. in the presence of LacR repression as well as PLysS repression). Furthermore, we successfully expressed both genes in the *E. coli* BL21 (DE3) PLysS. However, when attempting to purify the resulting histones after gene expression, the histones which contained affinity tags failed to bind to the corresponding affinity resin. We propose that this might be due to the possibility that the tags which are present on the N-terminus of the proteins are buried in a hydrophobic pocket that is formed by the proline-proline tetrad motif which exists in all archaeal histones.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

TABLE PAGE

LIST OF FIGUERS

FIGURE PAGE

CHAPTER I

INTRODUCTION

Historical Perspective

Understanding the kinetics and thermodynamics of protein folding has been a subject of intense investigation since Christian Anfinsen first demonstrated that the information for protein folding is present in the primary amino acid sequence of proteins (*1*). The paradigm of folding through random sampling of conformational space could be rapidly excluded based upon the predicted slow rate of this process. Cyrus Levinthal pointed out this paradox in 1968 through simple calculations where an amino acid in a peptide of a given length was restricted to three different conformations (*2*). Levinthal calculated that sampling of random conformational space would take longer than the age of the universe (*3*). Based upon this simple analogy it was obvious that proteins did not sample conformation space but quickly collapsed into intermediates that resembled the final folded state Without going into an exhaustive discussion of the field of protein folding suffice it to say that we have made tremendous strides towards understanding the problem primarily at the single polypeptide level that consist of a single structural domain (*4*). However, many proteins in the cell contain multiple domains and many exist

as multi-subunit structures (*5*). The formation of quaternary interaction adds a second level of complexity to the "protein folding" paradigm (*6*).

Molecular Pathology of Protein Folding and the Recombinant DNA Era

The significance of protein folding took on a new meaning with the discovery that a single amino acid substitution of valine for glutamate in the amino acid sequence of the β-globin protein led to sickle cell anemia a disease resulting from alteration in the quaternary structure of the hemoglobin molecule (*7*). Since this initial discovery the molecular basis of numerous diseases has been linked to aberrant protein folding (*8*). For example many cancers, diabetes, cystic fibrosis, just to mention a few results from "protein folding gone wrong" (*9*).

The development of gene cloning in the 1970s offered tremendous hope with regard to producing proteins with pharmaceutical, industrial, and scientific research applications (*10*). However, numerous recombinant proteins fail to fold into the correct native structure and as a result they either are proteolytically degraded or exist as partially unfolded intermediates that are found as aggregates in the cell (*11*). The recombinant DNA era also offered the hope of engineering proteins that would have physical and chemical properties superior to the native species. Specifically proteins that are more catalytically efficient have been engineered, along with those that have enhanced thermo stability (*12*). However, the rate limiting step for the implementation of all of these initiatives is the ability of the proteins to fold into their native state inside of a host cell or

in an in vitro expression system. Thus, the applications outlined above are limited by the ability of the protein to achieve the correct three dimensional shapes (*13*).

A Model System for Exploring Quaternary Interactions

Our specific interest lies in identifying the structural determinants that stabilize homo- versus hetero-dimer formation between orthologous or paralogous proteins where the promoter units are structurally similar. To investigate this aspect of protein folding we have turned to the euryarcheota, one of the two groups within the domain Archaea. All members of this group have histone-like proteins that are structurally similar to the eukaryotic H3 and H4 histones. Specifically, the euryarcheota have at least two genes that encode two very small proteins that have a mean size of 66 amino acids and are typically designated as an A or α and B or β types. The crystal structures of both the α and β types from both mesophilic and thermophilic organisms have been solved and all of the proteins consist of a helix- loop- helix- loop- helix structure that matches the typical histone fold found in eukaryotic histones as well as a few other proteins. The length of the central helix is as long as one would find outside of fibrous proteins (29 amino acids), and one can easily surmise that the α or β monomer might be unstable in water due to competition for the hydrogen bonds. Thus, oligomerization may be required to stabilize the monomers. (*14-17*)

All of these proteins have been characterized as homo-dimers. However, we believe that homo-dimers have only been observed because laboratories only clone either α or β in a cell, and due to instability of the monomer, homo-dimerization is

thermodynamically forced. In other words, if both proteins were in the cell at the same time we believe that homo- and hetero-dimers would form, though there might be differences in their stability. Evidence for artifactual quaternary structure in the absence of all promoters being available has numerous precedence in the literature. For example HU α and Hu β form hetero-dimers predominantly in *E. coli* (*18*). However, *E. coli* mutants null for the expression of one of the orthologs results in homo-dimer formation (*19*). In vertebrates the hnRNP C1 and C2 proteins, that result from alternative splicing of the same gene form hetero-tetramers with a promoter composition of 3 hnRNP C1 polypeptides to 1 hnRNP C2 polypeptide in vivo. However, when hnRNP C1 or hnRNP C2 are expressed in *E. coli,* only homo-tetramers result (*20*). More dramatic examples that are associated with pathological conditions include α−thalassemias that result from the lack of expression of one of the α subunits of hemoglobin (*21*) leading to an abundance percent of β subunits which lead to the formation of homo-dimeric structures to replace the physiologically functional hetero-dimeric structures.

We have demonstrated the efficacy of hetero-dimer formation through the use of protein-protein molecular docking calculations. The results of these calculations indicate that hetero- and homo-dimers are structurally indistinguishable. And even more interesting analysis was obtained when we used molecular docking calculations to look at homo- vs. hetero-dimer formation between paralogs which are histones from two different organisms, namely the thermophilic archeon *M. fervidus,* and the mesophilic archeon *Methanobrevibacter smithii*. The results of those calculations also showed that chimeric homo- and hetero-dimers are structurally identical to one another as well as to

the actual crystal structure (unpublished data of research conducted previously carried out in our laboratory).

Significance

The molecular modeling results discussed above seem to suggest that homo and heterodimer formation between these proteins form randomly. However, with regard to physiological relevance nothing is random. If one structure was preferred over the other it is obvious that either the α or β genes would be lost through a lack of evolutionary selection pressure. Thus, it is clear that within the cell there exists either a preference for one or the other species or both exist and their concentrations are regulated.

The essential question that must be answered is what occurs in the cell with regard to the state of oligomerization of these proteins. The significance of this question has been amplified in light of recent data that showed that members of the basic leucine zipper family of transcription factors, of which many of these are proto-oncogenes, have the ability to form homo-dimeric and hetero-dimeric complexes with structural homologues, and each species has a different function (*22*). For example, the protein Max has the ability to form homo-dimers or hetero-dimers with members of the c-Myc family (*23*). In both cases the DNA binding specificity of both proteins are the same. However, the hetero-dimer is a transcriptional activator and the homo-dimer is a transcriptional repressor. Recently, Max was shown to hetero-dimerize with two other members of the basic leucine zipper class of transcription factors and the functional significance of each has not been assessed.

As indicated earlier despite the structural and primary sequence homology between both α and β orthologs and paralogs all the proteins that have been cloned and expressed in *E. coli*, have been shown to be homo-dimers. Interestingly, the eukaryotic paralogs of these proteins all form hetero-dimeric species. Since both the A and B proteins are highly similar with regard to primary, tertiary, and quaternary structure it would be interesting to explore the structural determinants in both proteins that stabilize homo-dimeric structure or inhibit hetero-dimeric formation.

Research Plan

All the structures and data on Archaeal histones are based upon homo-dimers. It is clear from our molecular modeling experiments that the hetero-dimer is as thermodynamically feasible, and homo-dimers form because of the inherent instability of the monomer. In other words, oligomerization of these proteins is an obligate interaction. To determine the actual distribution of homo–versus hetero-dimer formation requires the co-expression of the proteins from a cis-recombinant construct in a polycistronic vector, and compare that to distribution occurring from trans expression of each monomer on a different plasmid. The vector that we will use to express the MS-H2A and the PF-H1A proteins from a polycistronic message was developed by Dr. Song Tan's laboratory at Pennsylvania State University (*24*). The plasmid, pST44, contains unique restriction sites that facilitate the cloning of up to four different genes each containing a separate ribosome binding site (Shine_Dalgarno Sequence) (*25*) upstream of a start codon. Each of the four "cassettes" contains termination codons. Transcription of the polycistronic

message is driven from a T7 RNA polymerase promoter (*26*) and transcription is terminated by a highly efficient bacteriophage terminator. The T7 bacteriophage promoter is one of the strongest promoters for gene expression, and the polymerase is one the most processive enzymes in *E. coli*. The use of a T7 promoter ensures that the recombinant genes will not be expressed unless there is a source of T7 RNA polymerase. This is essential since many DNA binding proteins are lethal when expressed in *E. coli.* As a result, it is undesirable to express genes from promoters that are recognized by *E. coli's* RNA polymerase. Therefore, we typically archive our genes initially in *E. coli* strain DH5α and then transfer the plasmids to an expression strain that has the gene for T7 RNA polymerase inserted into the *E. coli* chromosome under the control of the *LacI* repressible promoter. Since leaky expression from the lac promoter is often lethal, we will initially transfer the plasmid into Bl21 DE3 PLysS a strain that in addition to having the T7 RNA polymerase gene has an additional plasmid that encodes lysozyme an inhibitor of T7RNA polymerase. In this system extremely lethal genes remain in an off state until induced with the allolactose analog IPTG. Our laboratory has a great deal of experience in using T7 based vectors and we already have a number of *E. coli* BL21 expression strains that facilitate expression from T7 promoters.

Initially, to use the pST44 system the genes will be placed in different transfer vectors which allow one to choose from a wide array of affinity tags that also have epitopes that are recognized by commercially available antibodies. The transfer vector also facilitates directional cloning using a pair of restriction enzymes to facilitate insertion of the recombinant sequence in one of the four expression cassettes of pST44. Once the different histone genes are placed in their prospective transfer vectors, they will

be excised along with the affinity tag sequences and the Shine_Dalgarno Sequence which is necessary for gene expression.

CHAPTER II

MATERIAL & METHODES

Growth and Maintenance of Bacterial Strains

All *E. coli* strains were grown in LB media (10 g of tryptone, 5 g of NaCl, and 5 g of yeast extract per Liter of media) at 37°C. Cultures were aerated by vigorous shaking of the culture at 100 to 200 rpm. Where needed the media was supplemented with 25 μg/mL of ampicillin. For the growth of Bl21 (DE3) strains chloramphenicol was added to a final concentration of 24μg/ml. Bacterial strains were stored by adding sterile glycerol to a final concentration of 20% (vol/vol) followed by storage at -80°C.

Plasmid Purification

For plasmid isolation, recombinant *E. coli* cells were grown overnight (*ON*) as previously described. Plasmid was purified on affinity spin columns purchased from Qiagen, and following the manufacturer's instructions. Briefly, approximately 5 mL of cells were pelleted by centrifugation at 17,000 xg in a blank centrifuge (AccuSpin Micro 17, Fisher Scientific, 2012). The media was then discarded and cells were resuspended in buffer P1. After the mixing, the cell suspension was mixed with buffer P2 to lyse the

cells and denature the DNA. The addition of buffer N3 facilitated renaturation of the plasmid DNA. The cell membrane and precipitated chromosomal DNA was removed by centrifugation for 10 min at 17,000 xg. Plasmid DNA was then applied to a Qiagen DNA affinity spin column followed by centrifugation for 1 min at 17,000 xg. The column was then washed with 750 μL of ethanol (Buffer PE) and was again centrifuged as before. After discarding, the excess ethanol the column was spun again to remove residual ethanol. The DNA was detached from the column by adding either 50 μL of water or elution buffer (10 mMTris, pH 8.3). Furthermore, the QlAprep® Spin Miniprep Kit (250) by Qiagen was used for DNA plasmid purification following.

PCR Reactions

PCR reactions were carried out on a Gene Amp PCR Systems 2400 thermal cycler made by Applied Biosystems. PCR reactions contained 50μl of 2XTAQ mixture (New England BioLabs), 44μl distilled water, template DNA (< 1 μg) and 2μl each of a gene specific primers tailed with a BamHI site on the 5' end (sense primer) or NgoMIV site on the 5' end (anti-sense primer). Primer sequences were based upon the DNA sequence of the gene obtained from the National Center of Bioinformatics (NCBI). The thermal cycler was programmed for 30 cycles using the following parameters: 95 °C for 30 seconds, followed by DNA annealing at 50 °C for 30 seconds, and extension at 72 °C for 30 seconds. PCR reactions on intact cells contained the same reagents described above but lacked template. In whole cell PCR,

E. coli colonies were stabbed with a sterile toothpick which was then placed in the PCR reaction.

Agarose Gel Electrophoresis

DNA was electrophoretically separated using agarose gels. Molecular biology grade agarose was purchased from Fisher Scientific. All agarose gels were prepared by mixing the appropriate amount of agarose with Tris-Acetate-EDTA (1X TAE buffer) (See Appendix). Gels were cast in an 8 cm x 8 cm gel casting apparatus containing an 8-well 2 mm x 4 mm comb. Upon polymerization, the gel was transferred to an electrophoresis chamber containing 400mL of 1XTAE buffer with 20µl of ethidium bromide (10 mg/mL). Electrophoreses were carried out at a constant voltage (100V) and DNA was visualized using a hand held UV lamp (260 nm) or a UV trans-illuminator (304 nm).

Isolation and Purification of DNA from Agarose Gels

All plasmids that were digested with restriction endonucleases were subject to gel purification to remove any un-digested plasmid. Undigested plasmid, a DNA ladder, and 2 samples of digested plasmid were separated on a 1% wt/vol agarose gel (See Appendix). We have previously shown that DNA exposed to ethidium bromide is refractory to further enzymatic activities like ligation; that is why two samples of digested DNA are separated. One sample will be stained and used as a guide to cut out

the other sample that will not be subjected to ethidium bromide staining. After electrophoretic separation, the gel was cut in half between the two digested lanes. The half of the gel containing undigested DNA, DNA ladder, and digested plasmid was placed in an aqueous solution containing 0.5 µg/mL of ethidium bromide. After confirming that the DNA had been cut, the digested plasmid was removed with a razor blade. The gel was then reconstructed and the un-stained digested DNA was cut out of the gel by extrapolation. The mass of the gel slice was then determined on an analytical balance and three times this mass of a proprietary chaotropic salt solution in 0.5 µg/mL was then added (QG buffer) and the gel was incubated at 50 °C for 10 min or until the gel dissolved (i. e. 300 mL for every 100 µg of gel). Isopropanol was then added (One times the mass of the gel volume in 0.5 µL). This solution was then applied to a Qiagen affinity spin column and the DNA was purified as previously described for other Qiagen affinity columns. The gel purification kit from Qiagen was used for this purification.

PCR Purification

After PCR amplification the DNA was purified using a PCR purification kit from Qiagen and following the manufacturer's instructions. First of all, the PCR reaction was mixed with 5 volumes of PB buffer (the chaotropic salt solution allows the DNA to bind). The column was then spun for 60 s at 17,000 xg. As previously described the DNA was then washed with PE buffer followed by elution with 35µL of water or EB buffer.

Restriction Endonuclease Digestion of DNA

All restriction enzymes were purchased from New England BioLabs (NEB) and used according to the manufacturer's instructions. Typically, digest contained 1-10 U of enzyme (1 U is the amount of enzyme required to digest 1 μg of plasmid in 1 hour), in a volume that did not exceed 40 μL. The volume of enzyme used did not exceed 1% of the final volume. This was done to prevent the glycerol concentration in the reaction from exceeding 5%. Each restriction enzyme from NEB is supplied with a 10X buffer that is diluted in the reaction to a final concentration of 1X. Restriction digestion was carried out in a 37 °C water bath for at least 1 hour.

Ligation

DNA ligase was purchased from Fisher Scientific. Ligation reactions contained 2μL of 10X ligase buffer (10X buffer for T4 DNA ligase), 2μl of the digested vector, 7μl of digested PCR product, and 2 U of T4DNA ligase (One unit is defined as the amount of enzyme required to give 50% ligation of HindIII fragments of λ DNA [5′ DNA termini concentration of 0.12 μM, 300- μg/ml] in a total reaction volume of 20 μl in 1 hour at room temperature in 1X T4 DNA Ligase Reaction Buffer). A negative ligation control containing only 2μl of digested vector, 15μl of dH2O, 2μL of 10X ligase buffer, and 2U of T4DNA ligase was also performed.

Preparation of Competent *E .coli* and Transformation

An *ON* culture of *E. coli* was prepared in SOB media (See Appendix). 1 mL of the *ON* culture was then used to inoculate 50 mL of SOB and the culture was grown to an optical density measured at 550 nm of 0.45-0.55. The culture was then placed on ice for 15 minutes. Prior to centrifugation the volume of culture was recorded. Each 2.5 mL of cell culture corresponds to a single transformation reaction. Cells were then pelleted by centrifugation at 4000 rpm in a Sorvall RT7 for 10-15 minutes. After discarding the supernatant, cells were resuspended in 1 mL of TFB (See Appendix) for each transformation reaction followed by incubation on ice for 10 minutes. Cells were then recovered by centrifugation for 15 minutes as described earlier. The cell pellet was then resuspended in 200 μL of TFB for each transformation reaction. DMSO was then added to the cell suspension (7μL/transformation) and the cells were placed back on ice. After 5 minutes. β-mercaptoethanol (in 10 mM K-Mes) was added (7μL/transformation). After 10 minutes the identical volume of DMSO was then added and cells were kept on ice for up to 48 hours. For immediate use 210 μL of competent cells were added to pre-chilled sterile 15 mL conical tubes. DNA was then added to the cells and the mixture was incubated on ice for at least 30 minutes. The cell suspension was then placed in a 42 °C water bath for 90 seconds followed by placing on ice for 2 minutes. Cells were then transferred to room temperature and 800 μL of SOC (See Appendix) was added to each tube. Cells were incubated at 37° C for 1 hour to allow expression of the antibiotic resistance gene. After 1 hour cells were pelleted by centrifugation for 5 minutes at 4000 rpm. The media was then discarded and cells were resuspended in 200 μL of SOC. Cells were then plated on media containing the appropriate antibiotics (See Appendix).

Whole Cell PCR Reaction

A single colony from PF/PSt44-DH5α and MSM-H2A/D2-DH5α plate was placed in eppendorf tube that contain 1µl T7 reverse primer, 1µl STO720 forward primer, 2µl 2X master mix TAQ, and 16µl distilled water (total of 20µl). PCR was conducted using the protocol described previously. The PCR product was analyzed on a 2% agarose gel (See Appendix).

Gene Expression

For protein expression recombinant plasmids were transferred to *E. coli* BL21-CodonPlus® (DE3)-RIPL cells or *E. coli* BL21 (DE3)-PLysS. *ON* cultures of these strains were then used to inoculate larger cultures. These cultures were grown to an optical density measured at 600 nm of 0.8 by using a Bausch & Lomb Spectronic 20. Before inducing gene expression with isopropyl-β-D-thio-galactoside (IPTG), a 1.5 mL sample of the culture was removed and the cells were pelleted by centrifugation at 17000 xg for 1 minute. The culture was then induced by adding IPTG to a final concentration of 0.1 mM To optimize for gene expression, 1.5 mL aliquates of the post-induced cultures were collected every half hour interval for a total of 2 hour incubation period. The cells were pelleted by centrifugation and subsequently suspended in 1X SDS-PAGE loading dye by vortexing vigorously. The solutions were then placed in a boiling water bath for 5 minutes. The samples were applied on a 15% SDS-polyacrylamide gel to test for gene expression. The remaining cultures were harvested by centrifugation at 4000 rpm until

the supernatant was clear. The media was then removed and cells were stored at -80 °C until protein extraction.

SDS-Polyacrylamide Gel Electrophoresis

Protein samples were fractionated on a 15% discontinuous SDS-PAGE gel using the Laemmli method (*27*). Then a little amount of 500 ml culture cell pellet was resuspended with 1 mL of B-PER (Thermo scientific) and adds 2µl of lysozyme solution to lyse the cells. In addition, the cell was incubated on ice for ½ hour or until mixture become egg white like consistency. 1µl of DNAse was added and then incubated it for ½ hour or until it's become water like consistency. Centrifuging it at 15,000×g for 5 minutes, then split the pellet and the supernatant in two eppendorf tubes. 80µl of the supernatant was taken and added to it 20µl of 4X loading buffer, and 500µl of 1X loading buffer to the pellet followed by boiling for 5-10 minutes prior to being analyzed on the 15%SDS-PAGE gels to insure complete protein denaturation. The SDS-PAGE gels were prepared according to the protocol outlined in (See Appendix). The samples were electrophoresed at voltage of 100 V. Protein bands were visualized using coomassie blue stain (See Appendix).

Protein Purification of PF-H1A (His-bind Resin)

A cell pellet containing His-tag-PF-H1A obtained from 25 mL-culture was suspended in 1 mL of His-bind 1x-binding buffer (See Appendix). 1µl of DNAse

(indicate concentration) was then added to the mixture and allowed to incubate while agitating on ice for ½ hour or until the mixture gains water like consistency. Clear lysate was collected by centrifuging the crude extract at 3000rpm for ½ hour or until the supernatant became clear. The supernatant was then mixed with ½ ml of His-bind resin (50 % slurry) and incubated while agitating on ice for 1 hour. The supernatant / His-bind resin slurry was then packed by gravity into the column provided in the His-Bind Kit (Novagen). The packed resin was carefully washed with 3mL of 1X binding buffer containing 1 µL/mL of triton (See Appendix). The resin was then washed with 3ml of 1X washing buffer (See Appendix). Finally the protein was eluted from the column with 6×0.5ml volumes of 1X elution buffer (See Appendix). 750 µl aliquots of the flow through, the washing, and the elution samples were mixed with 250 µl of 4X SDS-PAGE loading dye (See Appendix) for analysis on SDS-PAGE.

Protein Purification of CBP tag-MS-H2A

A cell pellet containing CBP tagged-MS-H2A histone obtained from 25 mL-culture was suspended in 1 mL of CBP-binding buffer (See Appendix). The supernatant was obtained in identical manner to that described in the previous section. The supernatant was mixed with 1 mL of 50% suspension of CBP-resin and incubated with shaking on ice for 1 hour. The supernatant/resin suspension was then poured into a small column and allowed to pack by gravity. The flow through was kept for further analysis. The resin was then washed with 3ml aliquots of two CBP-washing buffers containing different concentrations of Ca^{2+} (See Appendix). The protein was then eluted from the column with 3 volumes of CBP-elution buffer (See Appendix). 750 µl aliquots of the

flow through, the washing 1, washing 2 and the elution samples were mixed with 250 µl of 4X SDS-PAGE loading dye for analysis on SDS-PAGE.

CHAPTER III

RESULTS

Cloning of the Histone Genes into pST50Trc3-CBPDHFR (D2)

The MS-H2A gene was first cloned into transfer vector 3 (pST50Trc3-CBPDHFR) of the polycistronic vector suite between the BamHI and NgoMIV restriction sites. In addition to providing the Shine-Dalgarno sequence for translational initiation signals and translational enhancers, this vector introduces a calmodulin binding protein tag to the MS-H2A gene which will be used in the purification of the recombinant protein. This tag consists of a 28 amino acid peptide that is added to the N-terminus of the recombinant protein increasing its molecular weight by 3.2 kDa.

Primer Design

The forward and reverse primers needed for the sub-cloning of the MS-H2A gene in the pST50Trc3-CBPDHFR vector were designed based on the MS-H2A gene sequence shown in figure 1. The forward primer contains the first 18 nucleotides of the 5' end of the sense strand. Also included in the sequence is the recognition site of BamHI restriction enzyme (GGATCC). In addition, the sequence, AAATAAA is added to the 5'

end of the forward primer to prevent mispriming of the primer to unintended sequences in the gene. The reverse primer contains 18 nucleotides that are complementary to the 3'-end of the sense strand. In addition, the sequence for the recognition site of NgoMIV is included (GCCGGC), along with the sequence, AAATAAA to prevent mispriming. These primer sequences are shown in Table I.

5′-
ATGTCTGAAATACCAAAAGCTCCTATCGCAAGGATTATTAAAGATACTGGTGCTGAAAGAGTTAGTGAAG
ATGCTAAAGCTGAATTAGCTGAATATCTTGAAGAAGTAGCTCGTGACGTTGCAATTGAAGCAAACAATGT
TGCTAAAATCGCTAAACGTAAAACTATCAAACCAGAAGATATTAAATTAGCTATTAAAAATTTAGAATAG-
3′

Figure 1: The sequence of the MS-H2A gene (the sense strand with the start and stop codons shown in green and red, respectively)

Table I: The Forward and reverse primers for the cloning of MS-H2A in pST50Trc3-CBPDHFR

Primer	Sequence
Forward Primer (MS-H1A-F-BamHI)	5'-AAATAAAGGATCCATGTCTGAAATACCAAAA-3'
Reverse Primer (MS-H1A-R-NgoMIV)	5'-AAATAAAGCCGGCCTATTCTAAATTTTTAATA-3'

PCR Amplification

The MS-H2A gene contains 207 nucleotides. After purification of the PCR product using the protocol outlined in materials and methods, we analyzed the product of the PCR amplification reaction on a 2% agarose gel which is shown in lane 3 of Figure 2. Also shown in lane 1 of figure 2 are bands that represent standards of DNA molecules of known sizes with the number of kilobases for each band shown on the left hand side of the figure. The agarose gel functions as a molecular sieve and is able to separate DNA molecules based on size with smaller molecules moving faster and farther than larger ones. However, this technique can be effectively used to determine the molecular size only for linear, double-stranded DNA which migrates through the gel matrix at a rate that is inversely proportional to the $\log_{10}$ of the number of base pairs. Based on this, we can estimate the size of the PCR product to be between 100 and 200 bp, which agrees with the MS-H2A gene sequence that we amplified.

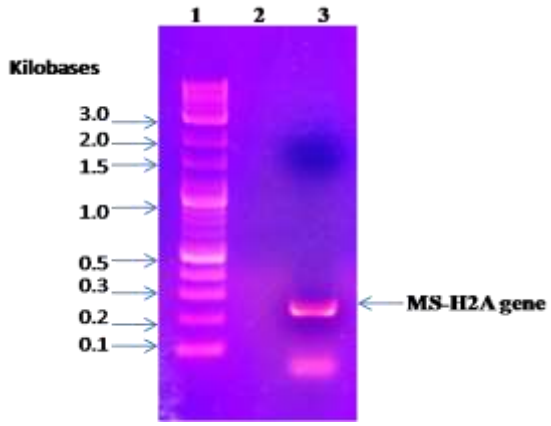

Figure 2: A 2% agarose gelto assess the success of the PCR amplification procedure. Lane 1 shows bands of the DNA molecular weight marker (TriDye 2-log DNA ladder 0.1-10.0 kb). In addition, lane 3 shows the product of amplification of the MS-H2A gene.

Plasmid Purification

To propagate a DNA sequence or gene in *E. coli* requires that we link the gene to a replicon. A replicon is a DNA molecule that contains an origin of replication (ori), a DNA sequence that is necessary to initiate DNA replication. The replicons used extensively for cloning in *E. coli* are plasmids, circular DNA molecules that contain origins of replication as well as selectable markers that allow one to identify recombinant *E. coli* from non-recombinant organisms. Another useful feature to look for when choosing a cloning vector is the presence of an affinity tag that can be used to select for the recombinant gene. The cloning vector used here is pST50Trc3-CBPDHFR (D2), a transfer vector that contains a promoter, translational initiation signals as well as transcriptional termination sites which will ultimately be transferred to the third coding

region in the polycistronic vector, pST44. Figure 3a shows a general map for the third transfer vector. As indicated in this map this transfer vector contains an ampicillin resistant gene (Amp^R) and an ori site (not shown). The region highlighted in yellow represents the cloning region. The region labeled N represents the N-terminus affinity tag which in the case of D2 corresponds to the CBP fragment. The MS-H2A gene is cloned directly after this fragment between the BamHI and NgoMIV restriction sites. A detailed sequence of this region showing these restriction sites is shown in Figure 3b.

The D2 vector was purified according to the procedure described in the materials and methods. To assess the purity and presence of this vector, we applied a small aliquot of the purified solution on a 1% agarose gel. The results obtained are shown in Figure 4. Lane 1 in this figure represents the TriDye 2-log DNA ladder described previously. And lane 2 represents the band that corresponds to the purified D2 transfer vector. As can be seen in this figure the D2 band lies between the 3 kilobase and the 2kilobase marker bands, with the D2 band appears closer to the 2 kilobase band. Although, the actual size of the D2 transfer vector is 2.8 kilobases, its band appears to run closer to the 2 kilobase marker band than to the 3 kilobase band. This is due to the fact that circular DNA migrates faster through the agarose gel matrix than linear DNA with the same size.

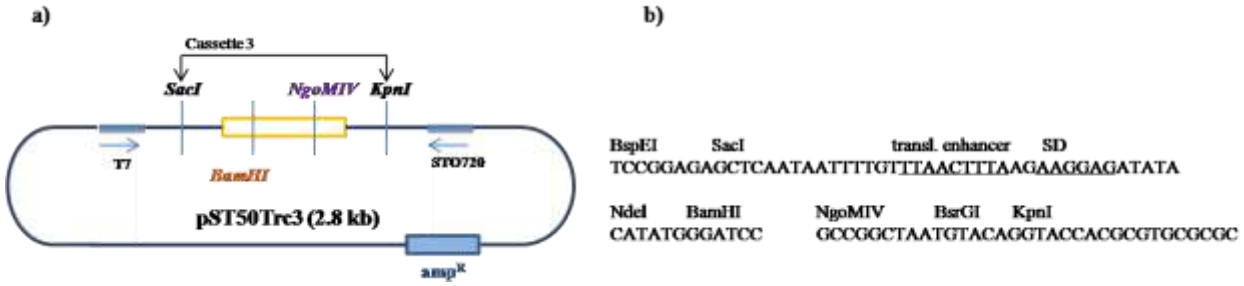

Figure 3: The map of transfer vector, D2. Panel (a) is the complete map showing the positions of the cloning region (yellow), and the amp^R site. Panel (b) is the detailed sequence of the cloning region showing the BamHI and NgoMIV restriction sites.

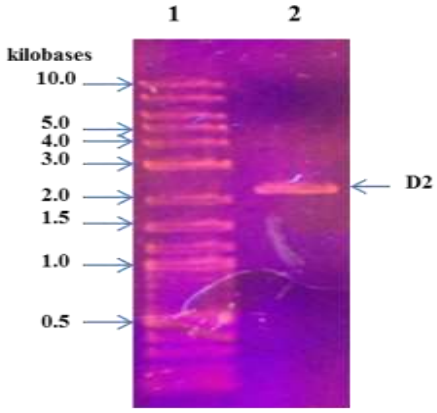

Figure 4: A 1% agarose gel showing the migration of the D2 vector through the agarose gel matrix (lane 2), where lane 1 represents the migrations of the DNA fragments of the TriDye 2-log DNA ladder.

Restriction Enzyme Digestion

The next step in the cloning process is to digest the MS-H2A gene and the D2 vector using the restriction enzymes BamHI and NgoMIV. Both of these enzymes produce overhangs in the DNA sequences of the gene and the vector (sticky ends) which makes the ligation of the digested MS-H2A gene to the digested D2 vector more efficient than blunt end ligation. Figure 5 is a 1% agarose gel which shows the product of the D2 digestion. Lanes 1 and 2 represent the TriDye DNA ladder and the undigested D2 vector described in figure 4. After the digestion, D2 is linearized and it should migrate through the agarose matrix at the predicted rate for its molecular size as seen in lane 3 of figure 5. A faint band close to 0.5 kilobases is also seen in lane 3. This band corresponds to the fragment that results from the digestion of D2 with BamHI and NgoMIV restriction enzymes (486 bp).

Careful examination of lane 3 revealed the presence of a faint band just below the band that represents the digested D2 plasmid. This faint band has the same migration pattern as the undigested D2 plasmid (see lane 2). In order to isolate the digested from the undigested D2, we excised the digested band taking care not to contaminate it with the undigested species. DNA was eluted from the excised gel segment according to the procedure outlined in the Materials and Methods section, "Isolation and Purification of DNA from Agarose Gels". The purification of the digested PCR product was simply performed by following the procedure outlined in PCR Purification section in Materials and Methods.

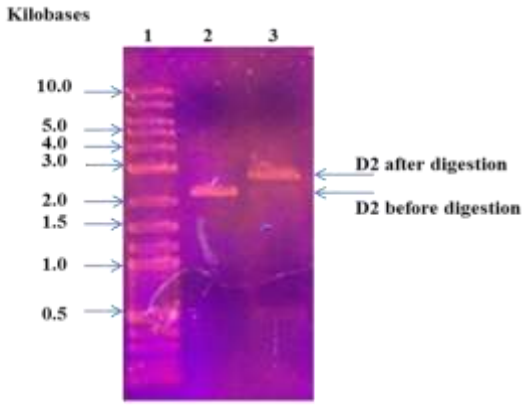

Figure 5: A 1 % agarose gel assessing the degree of digestion of the D2 plasmid. Lane 1 is the TriDye 2-log DNA ladder 0.1-10.0 kb. Lane 2 is a sample of the undigested D2 plasmid. Lane 3 is a sample of the restriction enzymes digested D2 plasmid.

Ligation and Transformation

The digested and PCR purified MS-H2A gene was ligated to the digested and gel purified D2 plasmid using T4-DNA ligase according to the protocol outlined in Materials and Methods. To test for the efficiency of restriction enzyme digestion, we also carried out a ligation reaction that contains the digested plasmid without the digested MS-H2A gene. If one or both restriction enzymes were not 100% efficient, the ligation reaction would then lead to the formation of a circular plasmid which is more efficiently taken up by *E. coli* than linear plasmid. The latter ligation mixture was used as a negative control. We also tested for transformation efficiency of the *E. coli* cells by carrying out a separate transformation reaction using undigested D2 (positive control). The products of the ligation reaction of the MS-H2A gene to the digested plasmid and the controls were then

used to transform the DH5α *E. coli* strain. The transformation mixtures were then applied to Agar plates that select for ampicillin resistance. Figure 6 shows the results of these transformation reactions. As can be seen in panel A, high transformation efficiency was obtained indicating that the *E. coli* cells used in this reaction are highly competent. Furthermore, the absence of colonies on the negative control plate shown in panel B indicates that D2 was efficiently digested leading to no transformants. Therefore the results shown in panel C indicates that it is highly likely that the colonies present on that plate represent *E. coli* cells that contain a recombinant MS-H2A gene.

Tentatively, clones were first screened by whole cell PCR using the protocol outlined in Materials and Methods. The primers used in the whole cell PCR reactions are the T7 and STO720 primers (see table II). The sequences of these primers are present on the D2 plasmid whether or not the MS-H2A gene has been incorporated (see figure 3). However, if whole cell PCR is performed on a colony from the positive control plate which contains only undigested D2 a 761 bp fragment will be produced. Whereas a whole cell PCR reaction on an *E. coli* colony containing MS-H2A/ D2 will result in a 475 bp fragment. To more concretely confirm the presence of a clone, we need to excite the gene by digesting with two restriction enzymes that border the gene sequence. Digesting with SacI and KpnI will result in a 344 bp fragment in the presence of the MS-H2A gene or a 630 bp fragment in the absence of the MS-H2A gene. Figure 7 is a 2% agarose gel that shows the results of the SacI and KpnI digestion of a plasmid purified from the colony that gave the tentatively positive whole cell PCR result (lane 2). Based on the standard molecular weight markers (lane 1) the band found in lane 2 (between 300 and 400 bp) supports the presence of the MS-H2A gene in D2.

Table II: The T7 and STO720 primer sequences used in the whole cell PCR reactions

Primers	Sequences
T7 primer	TAATACGACTCACTATAGGG
STO720	GCTAGTTATTGCTCAGCGG

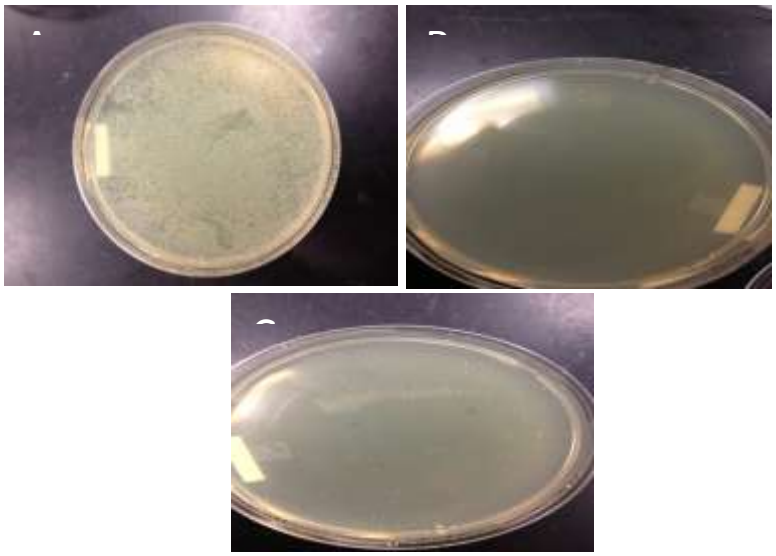

Figure 6: Ampicillin selective agar plates to test for the success of the cloning of the MS-H2A gene in D2. Panels A and B represent the transformation results of the positive and negative controls, respectively. Panel C represents the transformation result of the MS-H2A gene/D2 ligation.

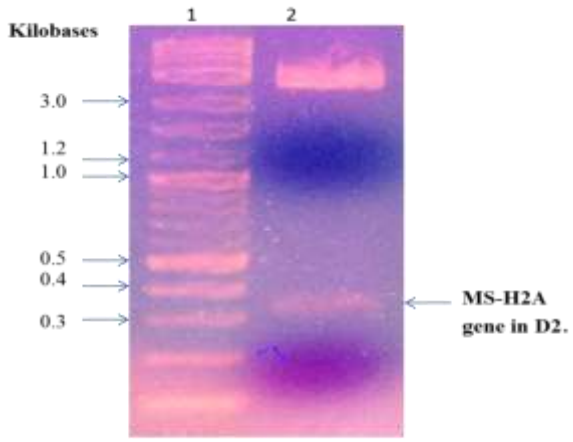

Figure 7: A 2% agarose gel that outlines the results of the SacI and KpnI digestion of the D2 plasmid containing the MS-H2A gene. This plasmid was purified from a colony that gave a positive whole cell PCR result (lane 2). Lane 1 is the TriDye 2-log DNA ladder 0.1-10.0 kb.

Sub-cloning of the MS-H2A Gene in pST44 Containing a Histone Gene from *Pyrococcous furiosus* and in pST44.

The ultimate objective of this project is to introduce multiple histone genes from the same archaeal organism as well as from different organisms in the pST44 polycistronic vector system in order to investigate whether these histones form hetero- or homo-dimers. A schematic diagram of the pST44 plasmid is shown in figure 8. Panel A shows the pST44 in the absence of a cloned gene whereas panel B shows pST44 containing the *pyrococcous furiosus* histone H1A (PF-H1A) in cassette 1 (PF-H1A/pST44). This clone was previously generated in our laboratory (*28*). The next step would then be to introduce the MS-H2A gene to cassette 3 of PF-H1A/pST44.

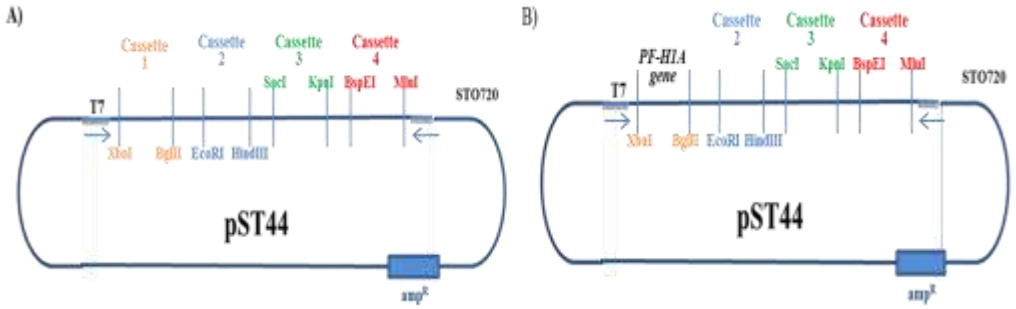

Figure 8: A schematic diagram of the polycistronic vector, pST44 adapted from (*24*). Panel A shows the pST44 in the absence of a cloned gene whereas panel B shows pST44 containing the *pyrococcous furiosus* histone H1A (PF-H1A) in cassette 1 (PF-H1A/pST44).

To carry out this sub-cloning step, we purified the MS-H2A/D2 plasmid from a DH5α colony that gave a positive clone (see previous section). The MS-H2A gene was excised out of this plasmid using the restriction enzymes specific for cassette 3 on the pST44 vector which are SacI and KpnI (see figure 8). The excised MS-H2A fragment (344 bp shown in lane 2 of figure 7) was purified from the digested mixture on a 2% agarose gel as previously described in the Materials and Methods. The pST44 plasmid that contains the PF-H1A gene (PF-H1A/pST44) was also digested with SacI and KpnI restriction enzymes. Figure 9 shows the results of this digestion. Lane 2 shows the undigested PF-H1A/pST44 (the band between 2 and 3 kilo bases). The migration pattern observed here agrees with what was discussed earlier regarding circular plasmid. A less intense band at a molecular size between 3 and 4 kilobases is also observed in this lane. This band appears to have the migration pattern expected for a linearized PF-H1H/pST44. Although the sample applied in this lane is assumed

to be undigested, we sometimes observe bands reflecting linear plasmid migration patterns. We believe that when performing plasmid purification we obtain a certain population of linearized vectors. Lane 3 shows the band resulting from the digestion with SacI and KpnI. Here we see a predicted migration consistent with a DNA fragment slightly above 3000 bp which is the size of pST44 containing the PF-H1A gene. The digested PF-H1A/pST44 was gel eluted on a 1% agarose gel to purify it from possible undigested plasmid. The gel eluted MS-H2A was then ligated to the gel eluted PF-H1A/pST44 plasmid as previously described. The ligation mixture was then transformed into DH5α. The transformation reactions were applied to ampicillin selective Agar plates. The results are shown in figure10. The plate in panel (a) of figure10 shows the transformation reaction of undigested PF-H1A/pST44. This reaction is used as a positive control which should result in a large number of colonies indicating high level of transformation efficiency of the DH5α. On the other hand, the plate in panel (b) shows the transformation reaction of a sample of digested PF-H1A/pST44 (negative control), where no colonies were observed indicating that the plasmid was completely digested. Panel (c) represents the transformation reaction of the plasmid that resulted from ligating the digested PF-H1A/pST44 with the digested MS-H2A (clone plate). The large number of colonies observed in comparison to the negative control implies the likelihood that we have successfully cloned the MS-H2A gene in cassette 3 of PF-H1A/pST44 (PF-MS/pST44).

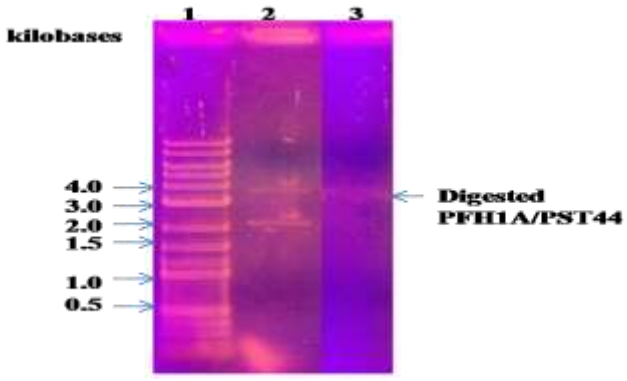

Figure 9: A 1% agarose gel showing the digestion of PF-H1A/pST44. Lane 1 represents the TriDye 2-log DNA molecular size ladder. Lane 2 is the undigested PF-H1A/pST44. Lane 3 is the PF-H1A/pST44 digested with SacI and KpnI restriction enzymes.

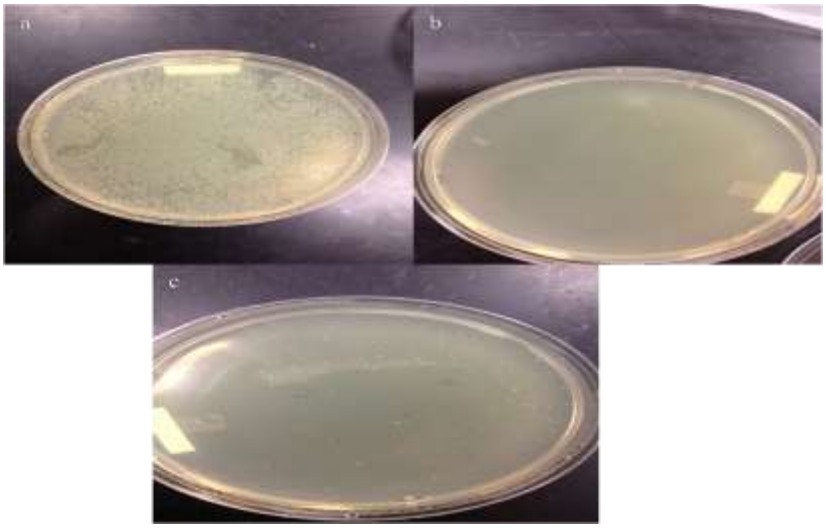

Figure 10: Ampicillin Selective agar plates to test for the success of the cloning of the MS-H2A gene in PF-H1A/pST44. Panels a and b represent the transformation results of the positive and negative controls, respectively. Panel c represents the transformation of the MS-H2A gene in PF-H1A/pST44.

To confirm the success of the cloning process we carried out whole-cell PCR on a number of colonies from the clone plate described above. Whole-cell PCR was conducted using the T7 and STO720 primers as discussed before. If the *E. coli* colony used to carry out the whole-cell PCR procedure contains both the PF-H1A and the MS-H2A genes we should expect the PCR reaction to produce a 930 bp DNA fragment. However, if the plasmid only contains the PF-H1A gene then PCR will produce a 586 bp DNA fragment. Figure 11 is a 2% agarose gel that shows the results of this experiment where lanes3 and 6 show the results of the whole cell PCR conducted on two colonies from the clone plate. Upon comparison of the band obtained from the PCR reaction using the standard molecular size DNA marker (lane 1) we conclude that the fragment size is between 900-1000 bp which is consistent with the presence of the MS-H2A gene along with the PF-H1A gene in pST44. To check the efficacy of whole cell PCR as a screening method, we also conducted the procedure by picking several regions on the cloning plate which are devoid of colonies. Due to the sensitivity of PCR, if those regions contain the ligation solution that was used for transformation then there is a high probability that the bands we discussed previously could be resulting from pST44 in the ligation solution. The results of those PCR reactions are shown in lanes 2, 4, and 5. The absence of any DNA bands in those lanes confirms that the PCR products found in lanes 3 and 6 are due to cloned pST44. Further validation of the whole cell PCR was also obtained by carrying out the procedure on a colony from the positive control that contains pST44 void of any cloned genes (Panel a of figure 10). There we expected a PCR fragment of 244 bp, which is exactly what we see from the product found in lane 7.

Further confirmation of the success of the cloning of the MS-H2A gene in the PF-H1A/pST44 plasmid is to enzymatically digest plasmid purified from several colonies (5 colonies) that grew on the clone plate (Panel C) Figure 10. The restriction enzymes used here are SacI and KpnI (cassette 3 restriction enzymes). We also digested the plasmid from the same colonies using XbaI and BglII (cassette 1 restriction enzymes) to check for the presence of the PF-H1A gene. The digestions were carried out as previously outlined in the Materials and Methods. As discussed in the Ligation and Transformation section, in both cases we should obtain a 344 bp fragment. The digestion products were analyzed on a 2% agarose gel. Figure 12 shows the results of these digestion reactions. Panels A and B correspond to the digestion with SacI and KpnI, and with XbaI and BglII, respectively. The SacI and KpnI digestion reactions were loaded in lanes 2-6 of Panel A. The presence of a small band at 300-400 bp in lanes 2-4 and 6 indicate that the corresponding plasmids contain the MS-H2A gene. Whereas, the absence of this fragment from lane 5 indicates that the PF-H1A/pST44 plasmid does not contain the MS-H2A gene. The larger fragment obtained in each of those lanes represents the expected linear pST44. The results of XbaI and BglII digestion are shown in Panel B of figure 12. As we expected, a band that corresponds to the PF-H1A gene is found at 300-400 bp is shown lanes 2 and 4-6. The absence of this band in lane 3 maybe due to the fact that the plasmid used in that digestion is very dilute as you can see from the faint band at about 3000 bp which corresponds to linear pST44.

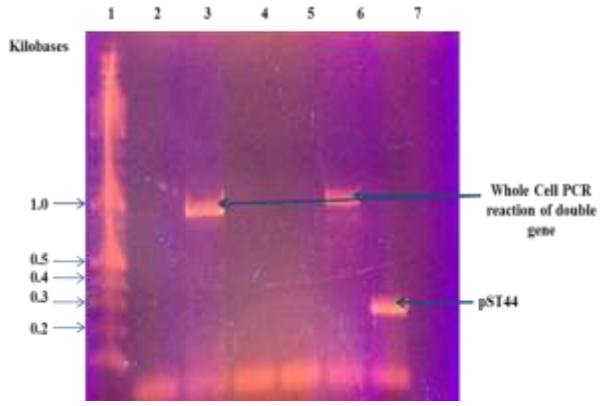

Figure 11: A 2% agarose gel demonstrating the results of whole cell PCR reactions to test the success of the cloning of MS-H2A in pST44 containing the PF-H1A gene. Lane 1 represents the TriDye 2-log DNA molecular size ladder. Lanes 2, 4 and 5 represent mock whole cell PCR reactions carried out on plate positions devoid of colonies (Figure 10, panel c). Lanes 3 and 6 represent whole cell PCR reactions carried out on two colonies from the cloning plate (Figure 10, panel c.), whereas Lane 7 represents whole cell PCR on a colony taken from a plate containing pST44 devoid of clones (Figure 10, panel a).

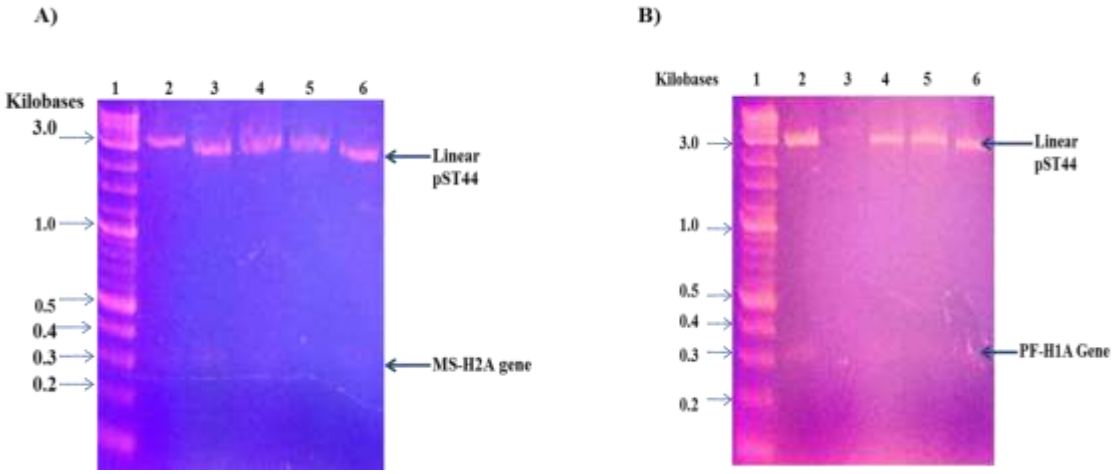

Figure 12: A 2% agarose gels demonstrating the results of enzymatic digestion of pST44 extracted from colonies presumed to contain the MS-H2A gene in cassette 3 and the PF-H1A gene in cassette 1. Panels A and B show the results of the digestions of the plasmid using cassette 1(SacI and KpnI) and cassette 3 (XbaI and BglII) restriction enzymes, respectively. Lane 1 in each panel represents the TriDye 2-log DNA molecular size ladder. Lanes 2-6 represent the results of the restriction digestions of plasmids extracted from 5 positive colonies. In both panels the band at ~ 3.0 kilobases is that of the linear pST44 resulting from the digestion, whereas the faint bands at ~0.3 kilobases represent the MS-H2A gene (panel A) and the PF-H1A gene (panel B).

Sub-cloning the MS-H2A Gene in Cassette 3 of pST44.

We also cloned the MS-H2A gene in cassette 3 of pST44 plasmid devoid of the PF-H1A gene in order to compare the effect of the presence of multiple genes in pST44 on levels of gene expression. Using the exact method discussed in the previous section, we sub-cloned the MS-H2A gene in cassette 3 of pST44 and transformed the

recombinant plasmid (MS-H2A/pST44) in DH5α. The results of the transformation are shown in Figure 13 where panels A and B are transformations of the undigested pST44 (positive control) and the digested pSt44 (negative control) in DH5α, respectively. Panel C shows the transformation result of the MS-H2A/pST44 plasmid in DH5α (clone plate). Again, high transformation efficiency of the cells is apparent from the high density of colonies observed in the positive control plate. The lack of colonies on the negative control along with the large number of colonies on the clone plate implies that there is a high probability that those colonies contain the recombinant MS-H2A/pST44 plasmid. We screened some of the colonies on the clone plate for the presence of the MS-H2A gene using whole cell PCR as described in the previous section. If the colony tested contains the recombinant MS-H2A/pST44 then we should see PCR fragment of ~344 bp Figure 14 is a 2% agarose gel illustrating the results obtained from screening two colonies from the clone plate. As seen in lanes 2-6, a band between 500-600 bp is observed, which confirms the presence of the MS-H2A gene. Lane 7 in this gel represents the product obtained for whole cell PCR carried out on cells containing pST44 (as discussed in figure 11).

To more conclusively confirm the presence of the MS-H2A gene, restriction digestions using the same method outlined in the previous section using SacI and KpnI were carried out. The digestion products were examined on a 2% agarose gel which showed a band between 300-400 bp confirming the presence of the MS-H2A gene (data not shown).

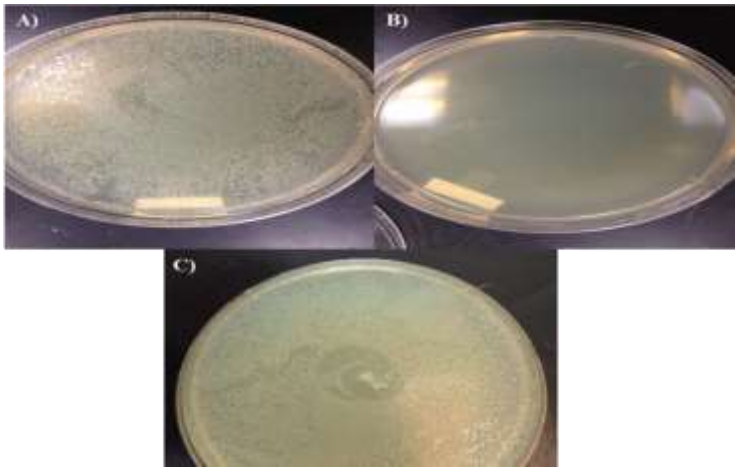

Figure 13: Ampicillin Selective agar plates showing transformation results of the MS-H2A/pST44 clone in DH5*α*. Panels A and B represent the transformation results of the positive and negative controls, respectively. Panel C represents the transformation of the MS-H2A gene in pST44.

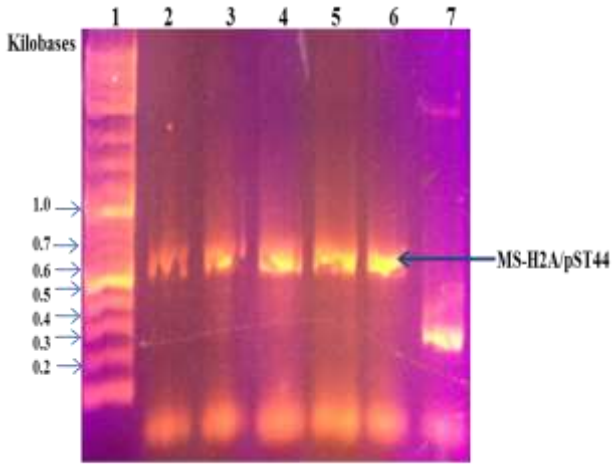

Figure 14: A 2% agarose gel demonstrating the results of whole cell PCR reactions to test the success of the cloning of MS-H2A in pST44. Lane 1 represents the TriDye 2-log DNA molecular size ladder. Lanes 2- 6 represent whole cell PCR reactions carried out on colonies from the cloning plate (Figure 13, panel c.), whereas Lane 7 represents whole cell PCR on a colony taken from a plate containing pST44 devoid of clones (Figure 13, panel A).

Transformation of PF-MS/pST44 in *E. Coli* BL21-CodonPlus® (DE3)-RIPL Cells and *E. coli* BL21 (DE3)-PLysS

So far we have cloned the MS-H1A gene along with the PF-H1A gene in the polycistronic vector pST44 in DH5α, an *E. coli* strain that lacks the ability to express recombinant genes. To carry out gene expression we then must transform the recombinant pST44 which contains both the MS-H2A and the PF-H1A genes in an expression *E. coli* strain. Two strains that we commonly use in our laboratory are *E. coli* BL21-CodonPlus® (DE3)-RIPL cells which is used for high level of protein

expression and *E. coli* BL21 (DE3)-PlysS that is specifically designed for the expression of lethal genes. Specifically, *E. coli*BL21 (DE3)-PLysS contains the PLysS plasmid that encodes for lysozyme which inhibits T7 polymerase therefore, eliminating leaky expression. PlysS plasmid also contains a chloramphenicol resistant gene a feature that can use for screening for *E. coli* BL21 (DE3)-PLysS. The *E. coli* BL21-CodonPlus® (DE3)-RIPL cells contain apACYC-like plasmid which encodes for arginine, isoleucine, proline and leucine codons that are rarely found in *E. coli* thus enhancing the expression of genes that contain such codons. This plasmid also encodes for the chloramphenicol gene.

We extracted the PF-MS/pST44 plasmid from frozen DH5α culture containing the recombinant plasmid. The plasmid was then transformed into both of the cells type mentioned above as previously discussed in the Materials and Methods. The transformation mixtures were plated on Agar plates that contain both ampicillin and Chloramphenicol. Figure 15 depicts agar plates that show the results of this transformation experiment. In this figure, the transformation results of *E. coli* BL21 (DE3)-PLysS and *E. coli* BL21-CodonPlus® (DE3)-RIPL with pST44 plasmid lacking either gene are shown in panels A and B, respectively. The large number of transformants seen in these plates indicates that these cells contain high levels of transformation efficiency. Panels C and D show the results of the transformation of the *E. coli* BL21 (DE3)-PLysS and *E. coli* BL21-CodonPlus® (DE3)-RIPL cells with the PF-MS/pST44 plasmid, respectively. No colonies were observed on either of those two plates indicating that these two genes (MS-H2A and PF-H1A) are lethal to both *E. coli* BL21 (DE3)-PLysS and *E. coli* BL21-CodonPlus® (DE3)-RIPL cells.

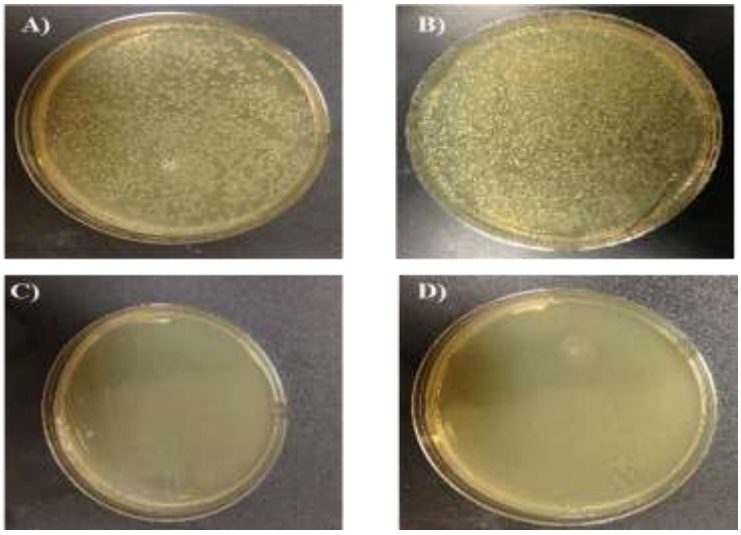

Figure 15: Ampicillin/chloramphenicol selective agar plates showing the results of transforming the PF-MS/pST44 in the *E. coli*strainsBL21 (DE3)-PLysS and *E. coli* BL21-CodonPlus® (DE3)-RIPL. Panels A and B represent the transformation results of the positive controls in BL21 (DE3)-PLysS and *E. coli* BL21-CodonPlus® (DE3)-RIPL, respectively. Panels C and D represent the transformation results of the PF-MS/pST44 clone in BL21 (DE3)-PLysS and *E. coli* BL21-CodonPlus® (DE3)-RIPL, respectively.

Transformation of MS-H2A/pST44 in *E. Coli* BL21-CodonPlus® (DE3)-RIPL cells and *E. coli* BL21 (DE3)-PLysS

To address the question of whether the lethality observed above is due to both genes present in the same plasmid or are the individual genes equally lethal for the *E. coli* cells used, we transformed the MS-H2A/pST44 plasmid (extracted from frozen cells that tested positive for the recombinant plasmid) into *E. coli* BL21 (DE3)-PLysS

and *E. coli* BL21-CodonPlus® (DE3)-RIPL cells. The transformation reactions we screened on agar plates containing ampicillin and chloramphenicol. We also tested for the transformation efficiency of the *E. coli* cells as described in the previous section. High transformation efficiency was observed (positive control) (results not shown). Similarly we tested for the efficiency of the restriction digestion (negative control) and no colonies were observed (results not shown). The results of the transformation of the recombinant MS-H2A/pST44 plasmid in both *E. coli* cells are shown in figure 16, where panels A and B show the transformation of MS-H2A/pST44 in *E. coli* BL21 (DE3)-PLysS and *E. coli* BL21-CodonPlus® (DE3)-RIPL cells, respectively. Unlike the results observed for the double recombinant PF-MS/pST44 plasmid (see figure 15 above), the single recombinant MS-H1A/pST44 produced several *E. coli* BL21 (DE3)-PlysS transformants (figure 16, panel A). However, no transformants were obtained when using *E. coli* BL21-CodonPlus® (DE3)-RIPL cells. These results are expected for moderately lethal genes being transformed into *E. coli* BL21 (DE3)-PLysS which has protection against gene lethality, whereas *E. coli* BL21-CodonPlus® (DE3)-RIPL cells, which do not have any such protection, will not survive in the presence of lethal genes. However, introducing both genes in the same plasmid generated extreme lethality that even *E. coli* BL21 (DE3)-PLysS could not combat. This extreme lethality could stem from the fact that cloning two moderately lethal genes in the same plasmid simply increases their concentrations to a threshold level that is not tolerated by *E. coli* BL21 (DE3)-PLysS cells.

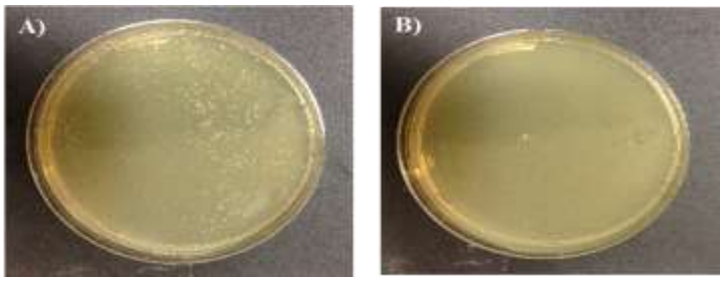

Figure 16: Ampicillin/chloramphenicol selective agar plates showing the results of the transformation of the recombinant MS-H2A/pST44 plasmid in *E. coli* BL21 (DE3)-PLysS (panel A) and *E. coli* BL21-CodonPlus® (DE3)-RIPL cells (panel B).

Gene Expression in *E. coli* BL21 (DE3)-PLysS Cells

The ultimate objective of this research project in our laboratory is to study protein-protein interactions between various archaeal histones in order to determine factors that affect quaternary structure formation. That was the primary reason for using the polycistronic vector, pST44, to clone the PF-H1A and MS-H2A genes in the same vector system. This would allow the co-expression of both genes simultaneously. However, because we were unable to transform the double recombinant plasmid, PF-MS/pST44 in either of the BL21 strains, we decided to carryout gene expression of each histone gene separately. This was carried out as described in the Materials and Methods. The expression of both genes was examined on a 15% SDS-PAGE. Shown in figure 17 are the results of this experiment where lanes 2-6 are the samples collected for the MS-H2A gene expression and lanes 7-10 are the samples collected for the PF-H1A expression. Lanes 2 and 7 are samples collected prior to IPTG addition (pre-induction) from cells containing the MS-H2A/pST44 and

PF-H1A/pST44, respectively. Lanes 3-6 are samples collected after 0.5, 1, 1.5 and 2 hours post-addition of IPTG (post-induction) of the MS-H2A/pST44. Similarly, lanes 8-10 are samples of the PF-H1A/pST44 collected after 0.5, 1, and 2 hours of post-induction. Also included in this gel is a sample of a broad molecular weight range protein marker (10-225kDa) which is used to confirm the presence and to estimate the molecular weight of the PF-H1A and the MS-H2A recombinant proteins. To confirm whether or not either or both genes have been expressed we compared the protein bands observed in lanes 2 and 7 to those found in the post-induction lanes of the corresponding cells. Since samples used in those two lanes are pre-induction samples, all protein bands present are native to the BL21 (DE3)-PLysS cells *E. coli* cells. Any additional protein bands that appear in the post-induction samples will most likely represent the recombinant proteins. Comparison of lanes 2-6 reveals a protein band slightly below 10 kDa that is only present in the post-induction sample and appears to increase in intensity as a function of induction period. It is most likely that this band represents the MS-H2A recombinant protein.

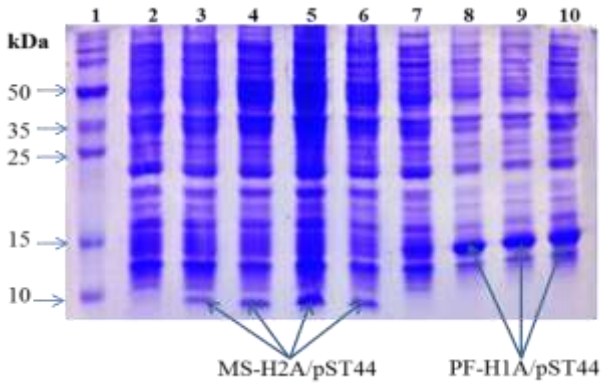

Figure 17: A 15% SDS-PAGE demonstrating the gene expression ofMS-H2A/pST44 and PF-H1A/PST44 in BL21 (DE3)-PlysS *E. coli* cells. Lane 1 shows the broad molecular weight range protein marker (10-225kDa). Lane 2 is crude cell lysate taken from cells prior to IPTG induction of BL21 (DE3)-PLysS cells containing MS-H2A/pST44. Lanes 3-6 are post-IPTG crude cell lysates of the cells containing MS-H2A/pST44 taken after 0.5, 1, 1.5 and 2 hours, respectively. Lane 7 is crude cell lysate taken from cells prior to IPTG induction of BL21 (DE3)-PLysS cells containing PF-H1A/pST44, whereas lanes 8-10 are post-IPTG crude cell lysates of the cells containing PF-H1A/pST44 taken after 0.5, 1, and 2 hours, respectively.

Protein Purification:

The recombinant MS-H2A and PF-H1A histones were purified from BL21 (DE3)-PLysS cells containing the MS-H2A/pST44 and the PF-H1A/pST44, respectively. The MS-H2A histone was purified on a CBP-resin, whereas the PF-H1A histone was purified on a HIS-bind resin according to the procedure described in the Materials and Methods. Samples from the various elution fractions of both

preparations were analyzed on 15% SDS-PAGE. The results of these gels are shown in figure 18. Panel A of this figure shows the results of the purification of the MS-H2A histone, whereas panel B shows the results obtained for the purification of the PF-H1A histone. Lane 7 in panel A represents a sample of the IPTG-induced of the BL21 (DE3)-PLysS cells containing the MS-H2A/pST44 recombinant vector. The band labeled with an arrow represents the MS-H2A protein which appears well below the 17 kDa molecular weight marker band found in lane 1. Lane 2 in this gel represents the flow through solution obtained from packing a column with the CBP-resin containing the recombinant MS-H2A protein. Lanes 3 and 4 represent fractions obtained from washing the column with buffers containing reduced concentrations of $CaCl_2$ in order to remove species with weak binding to the resin. Lane 5 represents the fraction obtained from elution of the column with a buffer containing 2 mM EGTA that is used to dissociate the calmodulin binding peptide from the resin. Lanes 2-4 show a low molecular weight band that has the same size as MS-H2A found in lane 7. These bands most likely represent the MS-H2A protein. The protein appears to elute in the absence of EGTA indicating that it did not bind to the CBP-resin. This is corroborated by the fact that the fraction that corresponds to EGTA elution (lane 5) is devoid of this molecular weight band.

The gel shown in Panel B of Figure 18 shows the results of the purification of the PF-H1A protein from His-bind resin. Lane 1 of this gel represents the broad range protein molecular weight marker. Lane 2 of this gel represents the fraction collected while packing the His-bind resin (flow through). The low molecular weight band (between 10 and 15 kDa) represents the PF-H1A protein. Lane 3 represents the

fraction collected while washing the His-bind resin with a buffer containing 60 mM immidazole, the concentration of imidazole used to remove proteins that bind to Ni^{2+} very loosely. While lane 4 represents the elution fraction collected by passing a buffer containing 0.5 mM immidazole, the concentration needed to elute His-tag proteins from Ni^{2+}. The presence of the PF-H1A band from lane 2 and its absence from lane 4 clearly indicates that this protein did not bind Ni^{2+} even though it contains a His-tag.

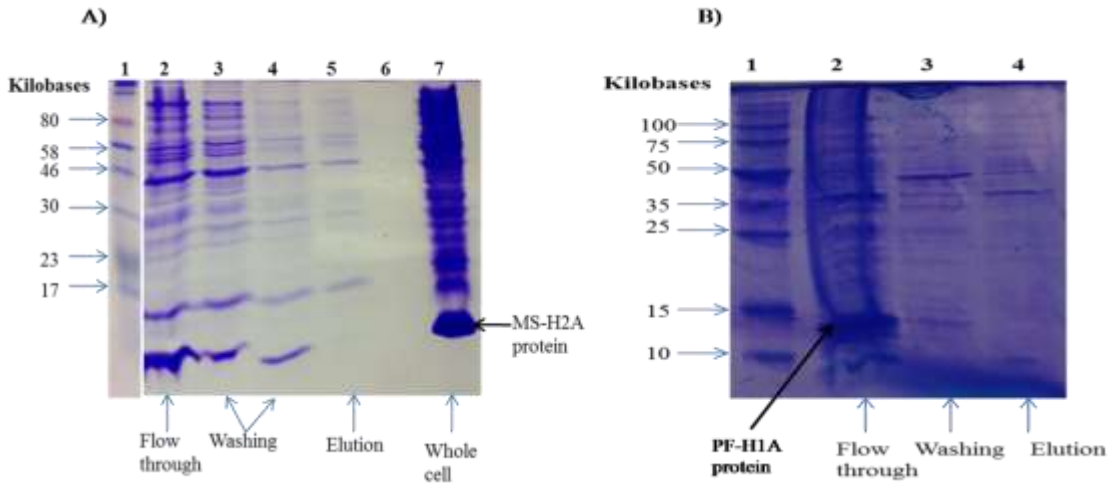

Figure 18: 15% SDS-PAGE of affinity column purifications of the MS-H2A and the PF-H1A proteins. Panel A represents fractions collected throughout the purification of the MS-H2A protein using CBP-resin. Lane 1 of this panel represents the Color-Plus prestained protein marker broad range (7-175 kDa). Lane 2 of this panel represents the fraction collected while packing the resin in the column. Lanes 3 and 4 are the fractions collected after applying the washing buffer. Lane 5 is the fraction collected after applying the elution buffer. Lane 7 represents the post-IPTG crude cell lysate of the BL21 (DE3)-PLysS cells containing the MS-H2A/pST44 gene with the bold black arrow to the right of this lane pointing to the band that represents the MS-H2A histone. Panel B represents fractions collected throughout the purification of the PF-H1A protein using His-bind resin. Lane 1 of this panel represents the broad range protein molecular weight marker (10-225 kDa). Lane 2 of this panel represents the fraction collected while packing the resin in the column. Lane 3 represents the fraction collected after applying the washing buffer. And lane 5 is the fractions

collected after applying the elution buffer. The bold black arrow to the right of this lane is pointing to the band that represents the PF-H1A histone.

CHAPTER IV

DISCUSSION

Our primary objective was to test whether PF-H1A could form a heterodimer with MS-H2A if both proteins were expressed from a single cistron. More specifically, since quaternary interactions might be coupled with translation, co-expression of both proteins from the same RNA would be the best system to evaluate this question. However, PF-H1A cloned into pST44 could not be propagated in *E. coli* strain BL21-CodonPlus® (DE3)-RIPL. This also proved to be the case for the propagation of the MS-H2A gene in this strain. Genes cloned into the *E. coli* BL21-CodonPlus® (DE3)-RIPL strain are under the control of the *LacI* repressor, which blocks synthesis of the T7 RNA polymerase necessary for transcription of the T7 promoters in our plasmids. Since expression from the Lac promoter is somewhat leaky in the presence of the *LacI* repressor, it is not surprising that these two DNA binding histone proteins would be lethal.

One might surmise that the expression of deleterious proteins in a strain that contains the rare codons for the amino acids arginine (R), isoleucine (I), proline (P), and leucine (L), which are typically lacking in *E. coli*, would have a synergistic effect with the presence of leaky *LacI* repressor. To address this problem we transferred both

genes to *E. coli* BL21 (DE3)-PLysS, a strain that lacks the rare RIPL codons and non-specifically inhibits leaky expression of the T7 RNA polymerase protein. Under these conditions both genes could be established in the strain. When each gene was placed in a separate pST44, we were able to transform the constructs in the *E. coli* BL21 (DE3)-PLysS strain, but not in the *E. coli* BL21-CodonPlus® (DE3)-RIPL strain. However, when both genes were placed in the same polycistronic vector, pST44, cells could not be transformed with the construct even under conditions of *LacI* repression and inhibition of T7 RNA polymerase. There are two hypotheses that can explain why we could not transform *E. coli* cells with constructs containing both genes. The first hypothesis, which we term the additive effect, is that enhanced expression resulting from having two copies of a histone gene instead of one led to cell death. The second hypothesis is that both proteins are forming a functionally distinct heterodimer that is much more lethal than the homo-dimers that form in cells when only one gene is present. The latter hypothesis seems more logical since the expression of both genes under conditions of *LacI* repression and inhibition of T7 RNA polymerase is not expected to be significant. This hypothesis is currently being tested in the following manner. pST44 constructs containing either two copies of MS-H2A or two copies of PF-H1A will be generated and used to transform *E. coli* BL21 (DE3)-PLysS. If these constructs are lethal then death probably results from the additive effect of multiple gene copies. If not, then our hypothesis of the formation of a functionally distinct heterodimer gains credence and will be more deeply explored.

Not only were we able to transform single gene constructs in *E. coli* BL21 (DE3)-PLysS, but we were also able to express both MS-H2A and PF-H1A genes to

significant levels that increased with time (maximum amount of MS-H2A and PF-H1A histones were reached after 1.5 hrs, and 2 hrs, respectively). However, when attempting to purify the proteins using their prospective affinity tags, neither histone was found to bind to the corresponding affinity resin. The lack of binding can be attributed to the N-terminus amino acid sequence and the structures of archaeal histones. We carried out sequence alignments of several euryarchaeal histones using Specialized BLAST search engine (*29*). This search revealed highly conserved proline residues at amino acid positions 5 and 8. NMR and crystal structures for *Methanothermus fervidus* histones MFH1 and MFH2 show that these prolines form intermolecular forces that result in what is termed the proline tetrad motif. This motif has been found in many other protein structures and is crucial for the stabilization of protein-protein interactions. An examination of available structures containing proline-proline tetrad motifs reveal that these structures form highly hydrophobic pockets that are being exploited in drug-discovery (*30*). We believe that the affinity tags placed at the N-terminus of MS-H2A and PF-H1A are buried within this hydrophobic pocket making them inaccessible to the affinity resin used for protein purification. This hypothesis can be corroborated by conducting the purification of protein under denaturing conditions where the proteins are completely unfolded which lead to exposing the N-terminus of these proteins and the affinity tags attached to them to the solvent and hence the affinity binding resin. We are currently in the process of carrying out this procedure. We are also in the process of re-cloning the genes with the affinity tags added to the C-terminus which appears to be more solvent exposed.

REFERENCES

1. Anfinsen, C. B.; Haber, E. Bringing an Enzyme Back To Life. *Journal of Biological Chemistry*, 1961, 236-1362.
2. Szilagyi, A.; Kardos, J.; Osvath, S.; Barna, L.; Zavodszky, P. Protein Folding. In *Handbook of neurochemistry and molecular neurobiology;* 3rd ED. Lajtha, A.; Banik, N. Springer : New York, c2007; volume editors, 303-344.
3. Debe, D. A.; Carlson, M. J.; Goddard, W. A. 3rd. The Topomer-Sampling Model of Protein Folding. *Proceedings of the National Academy of Sciences of the United States of America*, Mar 16, 1999, 96(6), 2596-601.
4. Uversky, V. N. Amyloidogenesis of Natively Unfolded Proteins. *Current Alzheimer Research*, June 2008, 5(3), 260–287.
5. Alberts, B.; Johnson, A.; Lewis, J. et al. Protein Function. In *Molecular Biology of the Cell*; 4th ED. Garland Science: New York, 2002 http://www.ncbi.nlm.nih.gov/books/NBK26911/ (accessed July 15, 2013).
6. Quaternary Structure http://www.acsu.buffalo.edu/~sjpark6/pednotes/Quaternary%20Structure.pdf (accessed July 15, 2013).
7. Alain, J.; Marengo-Rowe, M.D. Structure-Function Relations of Human Hemoglobins. *Proceedings (Baylor University Medical Center*) [Online] July 2006, 19(3), 239–245 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1484532/ (accessed July 15, 2013).
8. Hartl, U. Protein Folding: Mechanisms and Role in Disease http://www.mpg.de/36350/bm10_Proteinfolding-basetext.pdf (accessed August 5, 2013).
9. Stefani M. Protein Misfolding and Aggregation: New Examples in Medicine and Biology of the Dark Side of the Protein World. *Biochimica et Biophysica Acta*, December 2004, 24, 1739(1), 5-25.
10. Institute of Medicine (US) Forum on Microbial Threats. The Science and Applications of Synthetic and Systems Biology: Workshop Summary. National Academies Press (US): Washington (DC), 2011, ISBN-13: 978-0-309-21939-6 http://www.ncbi.nlm.nih.gov/books/NBK84445/ (accessed January 29, 2013).
11. Baneyx, F.; Mujacic, M. Recombinant Protein Folding and Misfolding in Escherichia Coli. *Nature Biotechnology*, November 2004, 22(11), 1399-408.
12. Song, J. K.; Rhee, J. S. Simultaneous Enhancement of Thermostability and Catalytic Activity of Phospholipase A_1 by Evolutionary Molecular Engineering. Applied and Environmental Microbiology, March 2000, 66(3), 890–894.
13. National Research Council (US) Committee on Intellectual Property Rights in Genomic and Protein Research and Innovation. Genomics, Proteomics, and the Changing Research Environment. In *Reaping the Benefits of Genomic and Proteomic Research: Intellectual Property Rights, Innovation, and Public Health*;

Merrill, S. A.; Mazza, A.M., Eds; National_Academies Press (US): Washington (DC), 2006
http://www.ncbi.nlm.nih.gov/books/NBK19861/ (accessed July 15, 2013).
14. Sandman, K.; Reeve, J. N. Archaeal Histones and the Origin of the Histone Fold. Current Opinion in Microbiology [Online] October 2006, 9(5), 520-525. Epub 2006 Aug 22.
http://www.ncbi.nlm.nih.gov/pubmed/16920388 (accessed August 5, 2013).
15. White, M. F.; Bell, S. D. Holding it Together: Chromatin in the Archaea. *Trends Genetics* [Online] December 2002, 18(12), 621-6
http://www.ncbi.nlm.nih.gov/pubmed/12446147 (accessed August 5, 2013).
16. Ding, Y.; Cai, Y.; Han, Y.; Zhao, B. Comparison of the Structural Basis for thermal Stability Between Archaeal and Bacterial Proteins. *Extremophiles* [Online] January 2012, 16(1), 67-78. DOI:10.1007/s00792-011-0406-z. Epub 2011 Oct 21
http://www.ncbi.nlm.nih.gov/pubmed/22015540 (accessed August 5, 2013).
17. Bailey, K. A.; Chow, C. S.; Reeve, J. N. Histone Stoichiometry and DNA Circularization in Archaeal Nucleosomes. *Nucleic Acids Research*, January 15, 1999, 27(2), 532-6.
18. Bonnefoy, E.; Rouvière-Yaniv, J. HU and IHF, Two Homologous Histone-Like Proteins of Escherichia coli, form Different Protein-DNA Complexes with Short DNA Fragments. *EMBO Journal*, March 1991, 10(3), 687–696.
19. Vimr, E. R.; Kalivoda, K. A.; Deszo, E. L.; Steenbergen, S. M. Diversity of Microbial Sialic Acid Metabolism. *Microbiology and Molecular Biology Reviews,* March 2004, 68(1), 132–153.
DOI:10.1128/MMBR.68.1.132-153.2004.
20. McAfee, J. G.; Soltaninassab, S. R.; Lindsay, M. E.; LeStourgeon, W. M. Proteins C1 and C2 of Heterogeneous Nuclear Ribonucleoprotein Complexes Bind RNA in a Highly Cooperative Fashion: Support for their Contiguous Deposition on pre-mRNA During Transcription. *Biochemistry*, January 30, 1996, 35(4), 1212-22.
21. Pilon, A. M. et al. Alterations in Expression and Chromatin Configuration of the Alpha Hemoglobin-Stabilizing Protein Gene in Erythroid Kruppel-Like Factor-Deficient Mice. *Molecular and Cellular Biology,* June 2006, 26(11), 4368-77.
22. Ecevit, O.; Khan, M. A.; Goss, D. J. Kinetic Analysis of the Interaction of b/HLH/Z Transcription Factors Myc, Max, and Mad with Cognate DNA. *Biochemistry*, March 30, 2010, 49(12), 2627-35. DOI:10.1021/bi901913a.
23. Baudino, T. A.; Cleveland, J. L. The Max Network Gone Mad. *Molecular and Cellular Biology*, February 2001, 21(3), 691-702.
24. Tan, S.; Kern, R. C.; Selleck, W. The pST44 Polycistronic Expression System for Producing Protein Complexes in Escherichia coli. *Protein Expression and Purification*, April 2005, 40(2), 385-95.
25. Shine, J.; Dalgarno, L. The 3′-Terminal Sequence of Escherichia coli 16S Ribosomal RNA: Complementarity to Nonsense Triplets and Ribosome Binding Sites. *Proceedings of the National Academy of Sciences of the United States of America,* April 1974, 71(4), 1342–1346.
26. Tabor, S.; Richardson, C. C. A Bacteriophage T7 RNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes. *Proceedings of the*

National Academy of Sciences of the United States of America, February 1985, 82(4), 1074–1078.
27. SDS-PAGE (Laemmli). http://camelot.bioc.cam.ac.uk/~marko/methods/sdspage.pdf (accessed June 1, 2013).
28. Soni, S. Molecular Basis of Lethality of the *Purococcus Furiosus* H1A Gene Product in Escherichia Coli. Master thesis, Pittsburg State University, 2013.
29. Decanniere, K.; Babu, A. M.; Sandman, K.; Reeve, J. N.; Heinemann, U. Crystal structures of Recombinant Histones HMfA and HMfB from the Hyperthermophilic Archaeon *Methanothermus fervidus*. Journal of Molecular Biology, October 13, 2000, 303(1), 35-47.
30. Kay, B. K.; Williamson, M. P.; Sudol, M. The Importance of Being Proline: the Interaction of Proline-Rich Motifs in Signaling Proteins with their Cognate Domains. *FASEB journal*, February 2000, 14(2), 231-41.

APPENDIX

LB (Luria-Bertani) (per liter): 10 g Bacto-Tryptone, 5 g Bacto-Yeast Extract, 5 g Sodium Chloride.

50X Tri-accetate-EDTA (TAE)(per liter): 242 g Tris Base (MW=121.14 g/mol), 57.1 mL 17.4 N glacial acetic acid, 100 ml 0.5 M EDTA (pH 8.0).

1% agarose gel: 0.3 g of agarose dissolved in 30 mL of TAE buffer by microwaving solution in 30 second intervals. The gel is cast immediately.

10X ligase buffer: 50 mMTris-HCl pH 7.5, 10 mM $MgCl_2$, 10 mMdithiotheitol, 1 mM ATP, and 25 µg/mL bovine serum albumin.

1X T4 DNA ligase buffer: 50 mMTris-HCl, 10 mM$MgCl_2$, 1mM ATP, 10 mM DTT, pH 7.5 at 25°C.

SOB medium (per liter): 20g Bacto-Tryptone, 5g Bacto-Yeast Extract, 0.5g NaCl, 0.18 g KCl. Addition of 5 mL of sterilized 2M $MgCl_2$ added before use.

TFB buffer: 30mMKOAc,50mMMnCl2,100mMRbCl,10mM CaCl2,15% Glycerol.

SOC medium (per liter): 20 g Bacto-trytone, 5 g Bacto-Yeast Extract, 0.18 g KCl, 0.5 g NaCl autoclaved and the addition of 5 mL of sterilized 1 M $MgCl_2$ is added prior to use.

2% agarose gel: 0.6 g of agarose dissolved in 30 mL of TAE buffer by microwaving solution in 30 second intervals. The gel is cast immediately.

4X Loading Buffer per 10 mL: 2.0 mL 1 M Tris-HCl (pH 6.8), 0.8 g SDS, 4 mL glycerol, 0.4 mL β-mercaptoethanol, 1.0 mL 0.5 M EDTA, 8.0 mg bromophenol blue, 2.6 mL H_2O.

Coomassie Blue Stain (per 2L): 0.1 g Coomassie blue, 200 mL Methanol, and 200 mL Acetic acid.

SDS-PAGE gel: The 15% Acrylamide SDS-PAGE gels used for electrophoretic studies were prepared using both the EZ-Run Continuous Gel Kit (Fisher) and a discontinuous gel. Using a gel apparatus, the resolving layer of the discontinuous gel is made with 9 mL distilled water, 5 mL of 4X Lower Tris, 14 mL of 30% acrylamide solution, 130 µL 10 vol% ammonium persulfate (APS), and 28 µL TEMED. This is quickly poured into the apparatus, covered with water saturated butanol and allowed to fully polymerize. After the bottom layer is polymerized, the butanol is discarded and chamber is rinsed with distilled water. The upper stacking layer is made with 6 mL distilled water, 3 mL of 4X Upper Tris, 2 mL of 30 % acrylamide solution, 40 µL 10 vol% ammonium persulfate (APS), and 20 µL TEMED. This solution is used to fill the remaining volume in the chamber making sure that bubbles are not formed. This solution is then allowed to polymerize.

Destain: 10 vol% Methanol and 10 vol% Acetic acid.

Lysozyme: 10 mg/mL in 10 mMTris-HCl (pH 8.0)

4X SDS-PAGE loading dye: 2.0 ml 1M Tris-HCl pH 6.8, 0.8 gSDS, 4.0 ml 100% glycerol, 0.4 ml 14.7 M β-mercaptoethanol, 1.0 ml 0.5 M EDTA, 8 mg bromophenol Blue

For protein purification of PFU-H1A (His-bind Resin):

8X Binding Buffer: 4 M NaCl, 160 mMTris, 40 mMImmidazole.

1X Binding + Triton Buffer per 80 mL: 80 µL Triton and 10 mL 8X Binding Buffer.

8X Wash Buffer: 4 M NaCl, 160 mMTris, 480 mMImmidazole.
8X Elute Buffer: 2 M NaCl, 160 mMTris, 4mMImmidazole.
For protein purification of CBP tag-MS-H2A:
CBP Binding buffer (per liter)**:** 10 mMTris-HCl, 150 mMNaCl, 2 mM $CaCl_2$.
CBP Washing buffer 1: 50 mM Tri-HCl (pH 8.0), 150 mMNaCl, 2 mM $CaCl_2$.
CBP Washing buffer 2: 50 mM Tri-HCl (pH 8.0), 150 mMNaCl, 0.1 mM $CaCl_2$.
CBP Elution buffer: 50 mM Tri-HCl (pH 8.0), 150 mMNaCl, 2 mM EGTA.