THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

THE EPIGENETIC REGULATION OF PEPTIDYLARGININE DEIMINASE 4 AND ASSOCIATED CELLULAR RESPONSE PATHWAYS GIVE INSIGHT INTO NOVEL CANCER THERAPEUTICS

ALEX RADAOUI
SPRING 2017

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Biology
with honors in Biochemistry and Molecular Biology

Reviewed and approved* by the following:

Yanming Wang
Associate Professor of Biochemistry and Molecular Biology
Thesis Supervisor

David Gilmour
Professor of Molecular and Cell Biology
Honors Advisor

Scott Selleck
Professor Biochemistry and Molecular Biology
Department Head

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

PAD4 is a protein found in the nucleus of many cells that promote post-translational modification of arginine residues on histones and other nuclear transcription factors. By modifying arginine residues to citrulline through deamination reactions, the transcription of specific genes might be turned on or off that are required for proper cellular function. This regulatory mechanism can provide insight into the maintenance and preservation of homeostatic epigenetics inside the cell. In addition, PAD4 has been known to corepress the cell cycle check point regulator p53, affecting its role in conserving proper cellular growth and proliferation. The link between the loss of p53 function in malignant tumors and overexpression of PAD4 has potential in providing significant knowledge into cancer biology. The purpose of my experimentation is to compile research regarding the epigenetic regulation of PAD4 on cellular homeostasis by analyzing a variety of tumor suppressor genes. I would like to know what PAD4's role is in regulating cell cycle signals. Through the research and techniques I have accumulated over the past two and a half years, I hope to show insightful data by implementing a PAD4 knockout plasmid into U2OS cells through CRISPR-Cas9 biotechnology, as well as perform qualitative PCR on treated SMAD3 knockout cells in order to better understand cellular maintenance signaling. Furthering the scientific communities knowledge of epigenetics is paramount in understanding prevalent immunological models like cancer and providing insight into possible drug design targets that can mitigate the effects of these diseases.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like thank my personal investigator, Dr. Yanming Wang. Not only am I appreciative of his wisdom, knowledge, and gentle touch in academia, but I am also grateful for the spark he ignited in me towards research in molecular biology and medicine. As a young student who was not sure where he wanted to go, Dr. Wang accepted me with open arms and helped me develop the skills of a true scientist through multiple hours of technical procedures and critical discussions in our laboratory meetings. Also I would like to thank Dr. David Gilmour for taking me on as my honors advisor in such short notice and being critical of my work. This has made me a better researcher and I could not be more appreciative. With this milestone comes the patience and time not many individuals have to spare, and for this, I am grateful. Our interactive and collaborative learning environment in the Center for Eukaryotic Gene Regulation taught me how to see the overall picture without getting distracted in heavy detail.

I would also like to thank the wonderful graduate students and undergraduates I have worked with over the past two and a half years. Dr. Xiangyun Amy Chen was one of my first mentors and taught me all the lab techniques one needs to know to be successful in molecular biology, Dr. Jinquan Sun who was always willing to give a helping hand or advice while also working on the final touches of her graduate thesis dissertation, and Lai Shi who I have had the privilege to learn from this past year and maximize my skills in molecular biology. In addition, Nicholas Schultheis, Jiayuan Zhou, Lu Yang, and Haonan Xu were all important parts of my experience in the Wang lab and have always provided a pedestal of motivation for me.

Lastly I would like to thank my family. Although distant, I felt your love and strength every day I have been at Penn State. Without you, none of this would be possible.

Chapter 1 – Introduction

I. Epigenetic Regulation and Histone Modification in Disease

One of the major regulators of gene expression involves the alteration or modification of histones in the nucleus. Histones work as a packaging and regulatory mechanism for coded DNA in the genome. This allows for certain genes to be activated or shut off through these specific modification mechanisms on different histones. Different modification mechanisms of histones can include acetylation, deacetylation, methylation, citrullination, phosphorylation, and ubiquitination, to name a few. Depending on the regulatory mechanism of modification, certain genes might be turned on or upregulated, or even shut off or downregulated in terms of transcriptional expression. Studying histone modification provides important insight on many diseases and the immune response in mammalian organisms including cancer and inflammation. Specifically, the family of protein arginine deiminases (PADs) regulates the citrullination of arginine and methylarginine residues on certain histones or downstream regulatory factors in order to promote or inhibit specific gene expression. Once a modification is made, an alteration in transcription factors occurs. An example of this regulation is PAD4's tendency to citrullinate histones H3 and H4 near promoter regions of genes. This post-translational modification results in the decondensation of chromatin, increasing the accessibility RNA polymerase machinery has with DNA and the likelihood of transcription. This holds potential in cancer epigenetics regarding possible tumor suppressor genes or oncogenes that may be affected by such modifications.

II. Genetic Engineering and Transcriptional Regulation

There are several methods in bioengineering that work to edit or artificially induce a change in the genetic code, which is an important aspect of biochemical research when attempting to understand the molecular expression of specific pathways in the cell. Three common methods are the use of transcription activator-like effector nucleases (TALENs) and zinc finger nucleases (ZFNs), as well as the more recent, efficient, and favored CRISPR (clustered regularly interspaced short palindromic repeats) method. All of these approaches involve sequence specific nucleases that facilitate genetic manipulation at the experimenter's desired locus of interest or study. These gene editing methods are favored for experiments involving clone isolation, gene knock-downs or knock-ins, as well as protein expression. The novel characteristic of these bioengineering methods is the induction of double stranded breaks at the loci of interest and then the utilization of the host's molecular repair mechanisms that include non-homologous end joining (NHEJ) or homology directed repair (HDR) to implement and incorporate the change in the genome.

There are two major outcomes desired in genome editing: adequate flexibility of defined target sequences by the experimenter as well as endonuclease recognition of the target sequence being researched. Although TALENs and ZFNs are still highly used depending on the experiment of interest, they have their setbacks. ZFNs are expensive and complicated to design as well as lack exact sequence specificity, whereas TALENs leave sticky ends after double stranded breaks by the nuclease and results in minimal off-target effects.[6] The CRISPR-Cas9 system addresses these setbacks by creating a more efficient and cost effective mechanism to introduce or delete certain genetic sequences in the genome.

Derived from bacteria's innate immune response mechanism to foreign or invading nucleic acids, the CRISPR-Cas9 system was manipulated for genetic engineering to work in eukaryotic organisms. The complex contains the Cas9 endonuclease as well as two crRNAs (crispr RNAs) and a trcRNA (trans-acting antisense RNA). The crRNAs work together as guide RNA (gRNA) to localize the Cas9 nuclease to the specific target DNA sequence being researched, and thus can help to mediate a break in the genome for insertions or deletions.[6] The double stranded breaks that occur are induced from Cas9's cleavage domains, RuvC and HNH, which have site specificity for the engineered sequence. In addition, a downstream protospacer adjacent motif, or PAM sequence, on the target gene is necessary for Cas9 recognition. In addition to its convenience when protein engineering is not possible, the CRISPR-Cas9 system is useful in studying the modification of human methylated DNA, which ZFNs and TALENs cannot do. The importance rises from the necessity to maintain proper transcriptional regulation within the nucleus of a cell. Should there be a specific histone regulatory protein that prevents the methylation of DNA through a separate post-translational modification, this can awaken genes that were not meant to be transcribed and disturb cellular homeostasis. Many diseases arise from the loss of proper epigenetic control such as cancer. Therefore, the innovation of the CRISPR-Cas9 system can help researchers understand gene silencing or activation by implementing knock-out or knock-in specific sequences that can prevent the post-translational modification of histones.

III. Protein Arginine Deiminase (PAD) Family

PADs are a family of enzymes that regulate an immune response in eukaryotic organisms and have several isoforms that work at different functionalities within the cell. Their roles vary from cell differentiation and proliferation in epidermal tissue to control of transcriptional regulation at the nuclear level. The family consists of PADs 1-4 and PAD6. Downstream signals of this enzyme can include proinflammatory responses, stimulation of angiogenesis and thrombin, as well as apoptotic regulation and gene expression.[1] Genetic research has highlighted this important class of enzymes due to their varied expression and distribution throughout the cell as well as their potential as targets for genetic manipulation and therapeutics in certain diseases such as cancer. The classification of PAD family isozymes and their specific cell and tissue expression patterns, as well as physiological roles can be found below:

PAD Isozyme	Substrates	Subcellular Localization, Cell and Tissue-specific Expression Patterns	Known and Putative Physiological Roles	Refs.
PAD1	keratin K1	• cytoplasm • epidermis, uterus, keratinocytes	• skin differentiation • terminal differentiation of keratinocytes	[3]
PAD2	vimentin, myelin basic protein (MBP), glial fibrillary acidic protein (GFAP)	• cytoplasm • skeletal muscle, brain, pancreas, glial cells, macrophages, bone marrow, muscle, breast, colon, embryo, eye, kidney, epidermal, uterus, thymus	• may play a role in brain development	[3], [29], [30], [31]
PAD3	trichohyalin	• cytoplasm • hair follicles, keratinocytes	• skin differentiation • hair follicle formation • terminal differentiation of keratinocytes	[1], [3]
PAD4	histones H2A, H3, H4, vimentin, p300	• cytoplasmic granules and nucleus • normal ○ eosinophils ○ neutrophils ○ granulocytes ○ macrophages • cancerous ○ cancer cell lines (MCF7, A549,SKOV3, U937) ○ cancerous tissues (breast carcinomas, lung adenocarcinomas, hepatocellular carcinomas, esophageal squamous cancer cells, colorectal adenocarcinomas, renal cancer cells, ovarian adenocarcinomas, endometrial carcinomas, uterine adenocarcinomas, bladder carcinomas, chondromas)	• cellular differentiation • transcriptional corepressor for the estrogen receptor and p53 • neutrophil extracellular traps (NETs) formation	[3], [4], [5], [6], [7]

Table 1: PAD Family Isozyme Classifications and Respective Expression Patterns [8]

Overall, PAD family members carry out the citrullination of arginine or methylarginine residues on corresponding enzymes, or on histones for PAD4, specifically. At high calcium levels, a requirement for PAD-mediated catalysis, the enzyme deiminates the arginine-like residue through hydrolysis resulting in a newly citrullinated residue and the release of ammonia.

In some assays, the activity of PADs can be monitored through the measured accumulation of ammonia in the sample. The catalytic reaction can be simplified as such:

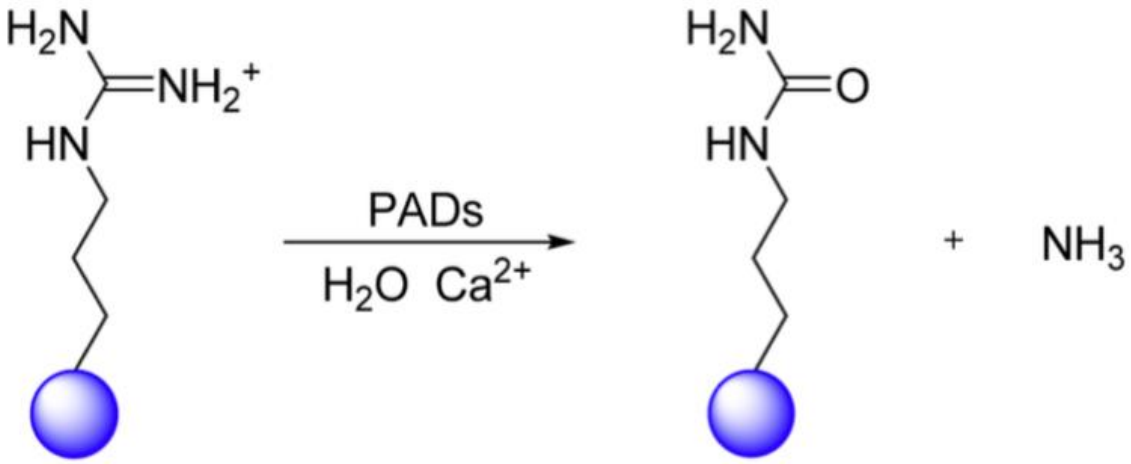

Figure 1: Reaction of PAD Mediated Catalysis [8]

Although the domain structures of all PADs are roughly 50% conserved throughout the family, PAD4 is one of the most highly studied isozymes due to its unique expression patterns in the cell.[12] Specifically, PAD4 is the only isozyme known to have a nuclear localization signal (NLS) and thus is the only member that is found in large amounts in nucleus. This is of interest in genetic engineering due to its potential to regulate or modify gene expression.

IV. Crystallographic Data and Catalytic Function of PAD4

In addition to being known as the only PAD isozyme to be in the nucleus, PAD4 is also the only family member with a known crystal structure. Data suggests the enzyme's quaternary structure to be a homodimer with head-to-tail interaction, with the monomer having less catalytic activity (25%-50% of the dimer). This would be indicative of dimerization for maximum

catalysis. Furthermore, the enzyme tends to show cooperativity when binding calcium to one of its five potential pockets, increasing catalytic efficiency as calcium concentrations increase.[11] The crystallographic layout of the homodimer structure with calcium binding pockets and catalytic residues can be observed below:

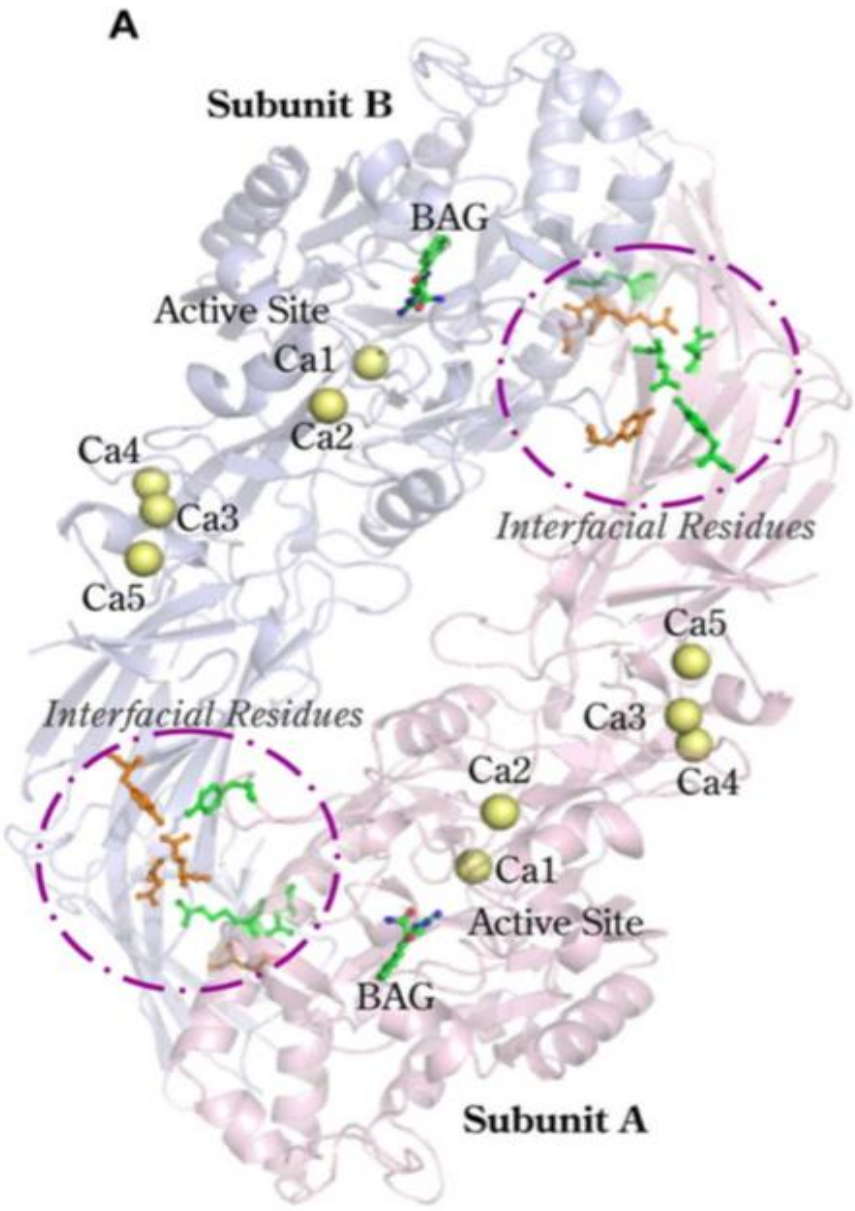

Figure 2: Crystal Structure of PAD4 [11]

There are four key residues involved in catalysis that include Cys-645, His-471, Asp-350, and Asp-473. The hypothesized catalytic mechanism in which arginine residues are converted to citrulline can be seen below:

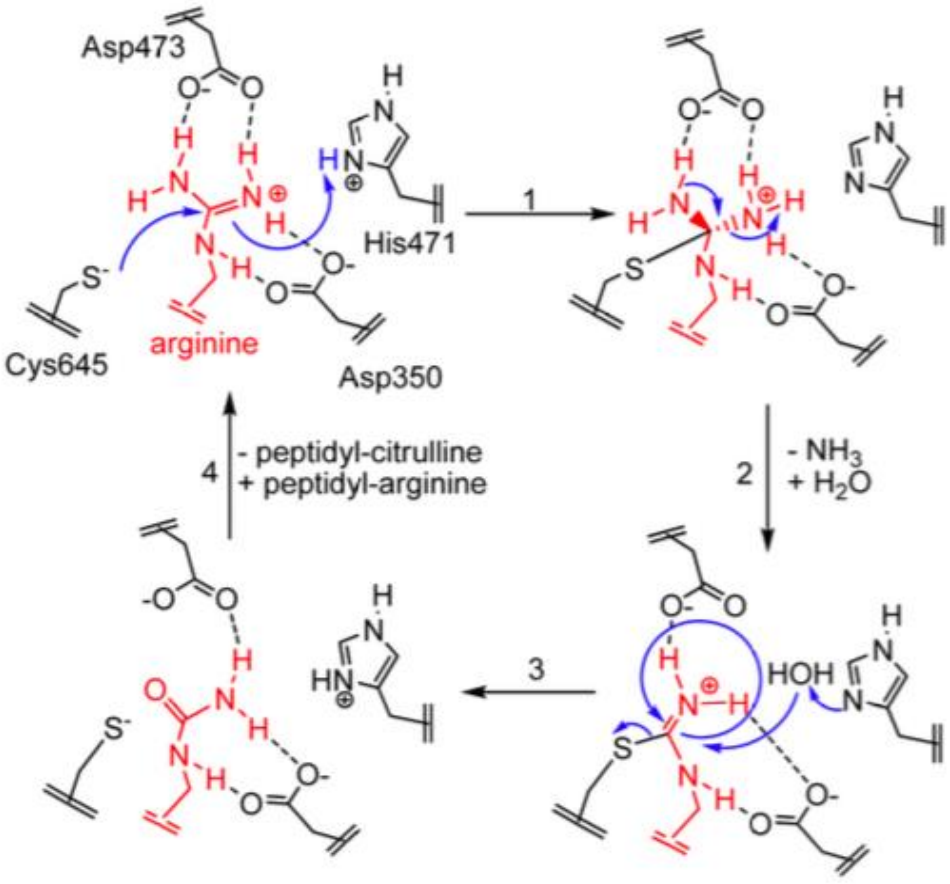

Figure 3: Catalytic Mechanism of PAD4 [8]

In the above mechanism, deamination most likely occurs through the cysteine reside acting as a nucleophile on the guanidium carbon to form a tetrahedral intermediate, almost similar to a serine protease mechanism. After the tetrahedral intermediate collapses from assisting aspartic acid residues, hydrolysis finishes the mechanism off creating a newly formed citrulline residue.[8] The histidine residue acts as a general acid or base in the mechanism, known

as a reverse protonation, which is common in thiolate reactive enzymes. Understanding the mechanism and binding efficiencies of PAD4 and its respective substrates is necessary in understanding the catalytic role of the enzyme in gene expression. From this, research can help to identify potential drug substrates or unknown targets of gene expression for PAD4 catalysis.

Some of PAD4's downstream targets include the posttranslational modifications of histones H2A, H3, and H4 through deamination of their arginine residues to citrulline, thus affecting gene expression. These modifications allow for regulatory effects over other intracellular pathways, DNA repair, apoptosis, and cellular growth and differentiation. Some of these induced or inhibited pathways include the ability of PAD4 to act as a transcriptional corepressor of p53, which can result in the likelihood of tumorigenesis.[8] The application behind inhibiting PAD4 in order to maintain homeostasis in the immune response is currently under intensive research due to its potential in having therapeutic qualities in diseases such as cancer.

V. PAD4 Expression in Cancer Biology and Associated Stress Response Pathways

PAD4 was first discovered as an increased signal in HL-60 cells in differentiating granulocytes and monocytes. It was soon observed that PAD4 is specifically expressed in a majority of immune cells as well as several cancer cell lines and in tumors. Research has even shown that PAD4-deficient mice did not show any formation of neutrophil extracellular traps (NETs) and were more susceptible to infection from pathogens.[9] The high nuclear expression levels of PAD4 work to promote an inflammatory immune response in order to activate a downstream series of multi-signal transduction cascades such as the ATF4 endoplasmic

reticulum (ER) stress response pathway or the mammalian target of rapamycin complex 1 (mTORC1) pathway.[25]

The ATF4 cellular stress response pathway is activated when there is a cellular imbalance in metabolism and can be induced by events such as a high ratio of improper protein folding that would stress cytoplasmic organelles like the ER. Furthermore, activation of the ATF4 pathway upregulates gene transcripts like sestrin-2 (SESN2) and DDIT4, which work in conjunction to inhibit the mTORC1 signaling pathway, which normally inhibits autophagy and stimulates cell growth, protein synthesis, and proliferation. When cellular stress is prevalent or compartmentalization machinery is not working properly, it is innate for the cell to prevent proliferation or growth in order to fix the malfunction or promote apoptosis. Sestrin-2 indirectly promotes the inhibition of the mTORC1 pathway by activating a signal cascade of AMPK and TSC1 and TSC2.[25] A western blot analysis correlating the ATF4 ER stress response with the mTORC1 signaling pathway under treatment with PAD4 inhibitor YW3-56 is shown below:

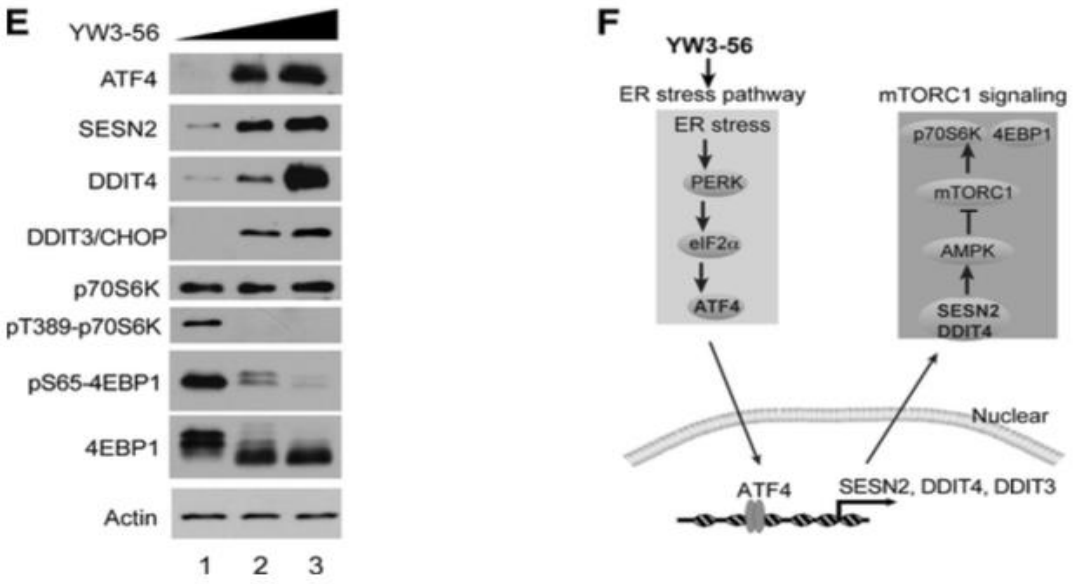

Figure 4: Western Blot of PAD4 Drug Inhibition and Activation of the ATF4 Gene Network [24]

After performing a microarray analysis of genes in the ER stress response network after YW3-56 treatment, Shu Wang and researchers compared the potential overlapping of ATF4 and CEBPB candidate genes through a ChIP-exo assay. A sequence tag density further confirmed that ATF4 binds much stronger at 4 highly inducible genes after YW3-56 treatment: SESN2, DDIT3, DDIT4, and CHAC1.[24] It is clear from the data presented in Figure 4 that by increasing the dosage of the PAD4 inhibitor, YW3-56, there is a higher presence of ATF4 downstream target signals like SESN2 and DDITs 3 and 4. This increase in expression further decreases mTORC1 signaling, which is evident from the decrease in phosphorylation signals of the pT389-p70S6K and pS65-4EBP1 on the western blot. Inhibiting mTORC1 kinase signaling increases autophagy, indicating the therapeutic potential in the PAD4 inhibitor, YW3-56. These conserved self-check mechanisms are highly important for the cell in order to prevent the growth of mutated or failing cells. Other relevant homeostatic cell maintenance pathways include the induction of protein 21 (p21) for cell cycle arrest, PUMA for activating proapoptotic genes like Bcl-2 or NOXA, and even OKL38 for the DNA damage response, which works to release mitochondrial cytochrome C and eventual apoptotic signals.[28]

In tumorigenesis, the well-known protein, p53, is over 50% mutated in cancer cells indicating a strong regulatory contribution to cell growth and maintenance. There are 6 major hallmarks regarding cancer cell proliferation: the evasion of programmed cellular death (apoptosis), insensitivity to anti-growth signals, self-sufficiency or autostimulation of growth, limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis. It is clear from these major characteristics that numerous homeostatic pathways are involved in hijacking normal cellular function. Many transcription factors and nuclear trans-activating regulators can work as tumor suppressors to ensure cellular homeostasis and prevent

tumorigenesis, or as proto-oncogenes that are innately silent, but have potential to be activated and promote cancer. More specifically, signals such as the TGF-β pathway that is involved with developmental cell growth and differentiation, or the PI3K-AKT pathway, which is involved with cell cycle and growth regulation, proliferation, and metabolic survival, are both important for regular and healthy cell function. What all these pathways have in common is p53, which seems to be a very prevalent transcriptional regulator in monitoring cell homeostasis. Numerous downstream signals of these pathways interact with p53 in order to ensure normal cellular function, and thus makes p53 a versatile and highly influential target for genetic regulation.

When the cell senses an imbalance in metabolic substrates, improper protein folding, or disruption of proper genetic regulation, p53 is recruited by the aforementioned pathways to upregulate gene transcripts that inhibit the growth of the cell by halting the cell cycle and activating apoptotic factors. This is a conserved feature of most cells in order to prevent the spread of diseases that have mutated cells like in cancer. In the TGF-β pathway, TGF-β cytokine activates a series of transcription factors known as SMADs in order to activate cellular proliferation and growth. SMADs normally interact with p53 and p63 to activate tumor suppressor factors and inhibit growth, however, mutant p53 has been shown to reinstate downstream TGF-β transcripts and reactivate the pathway.[4] Furthermore, in many cancers the PI3K-AKT pathway is over active, stimulating both the mTORC1 pathway and the TGF- β pathway reducing the likelihood of apoptosis and increasing cell proliferation.

Because PAD4 works as a corepressor with p53 as well as works with histone deactylases (HDACs), it is a novel target for genetic engineering. Furthermore it is also known that PAD4 is highly expressed in malignant but not benign tumors.[8] Possible gene therapy by knocking down or knocking out PAD4 has shown potential in decreasing tumorigenesis by rescuing normally

inhibited tumor suppressor genes that p53 normally interacts with such as p21 or even non-histone substrates like ING4 (inhibitor of growth 4). The western blot analysis below demonstrates the increased signaling of the p53 transcriptional regulator and its associated protein p21 when PAD4 is knocked down with small interfering RNA (siRNA):

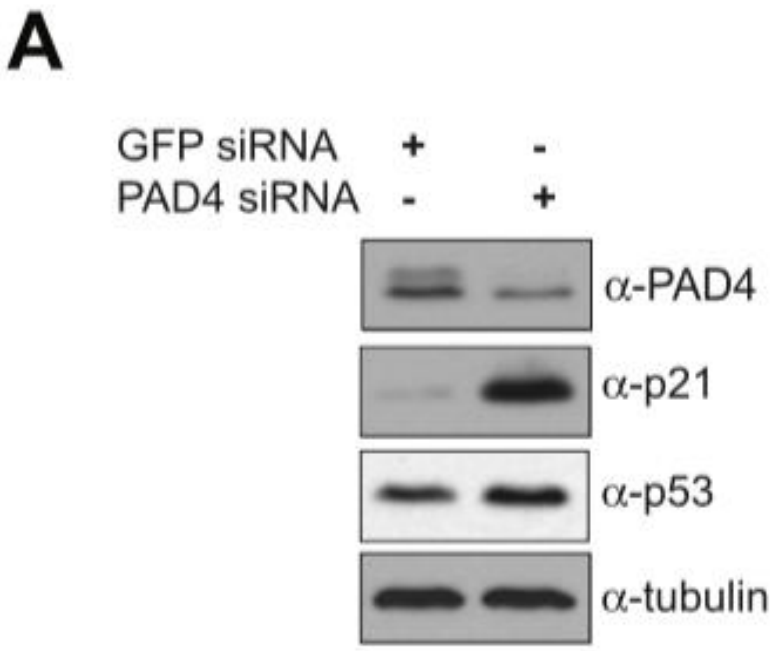

Figure 5: Western Blot of PAD4 Knockdown and Associated Tumor Suppressor Factors [11]

Other research by Qin Guo and Walter Fast demonstrated the binding of ING4 to p53 via its nuclear localization signal and thus works as a tumor suppressor gene to inhibit cellular growth. In addition, they had found that PAD4 citrullinates ING4 at the same nuclear localization signal site thus disrupting ING4-p53 interaction and decreasing p53 acetylation and p21 expression.[5] Signal disruptions such as these, which inhibit tumor suppressor regulators, are ones that could allow the cell to divide uncontrollably and potentially cause cancer.

p53 has also been shown to interact and localize with the promoter region of OKL38, another tumor suppressor gene. OKL38 is an important immune response regulator of apoptosis when DNA damage occurs inside the nucleus. Proapoptotic genes like BAX, NOXA, and the p53-AIP1 complex are turned on and OKL38-p53 localizes to the mitochondria in order to promote the release of cytochrome C. Through a chromatin immunoprecipitation (ChIP) assay, it was shown that there were heavy histone arginine modifications at the OKL38 promoter region and that PAD4 interacted with p53 in order to decrease the expression of OKL38.[28] Jones and researchers reported that expression of OKL38 is normally low in MCF7 breast cancer cells through the deamination of promoter associated nucleosomes, but post PAD4 siRNA knock-down or Cl-amidine (a nonspecific PAD inhibitor) treatment, OKL38 expression increased and induced apoptosis.[8] The extensive research done on the genetic regulation of PAD4 and its associated substrates has demonstrated its importance in regulating an immune response and cellular homeostasis. Due to its strong correlation with tumor suppressor regulatory genes such as p53 and its prevalence in tumorigenesis, PAD4 remains a very intriguing area of research for drug therapies that inhibit its signaling.

VI. PAD Inhibitors and Potential Therapeutics in Disease

Novel drug targeting in molecular medicine research has taken off with more proficiency due to the recent knowledge and understanding of genetic regulation. In cancer, tumor suppressor genes are turned off by enzymes such as PAD4 that modify histones and thus have become prime targets for inhibitory treatment. Other histone regulators such as HDACs and DNA methyltransferases have been studied in order to develop small inhibitor reagents that can

reactivate these tumor suppressor genes in cancer chemotherapy.[25] Many types of PAD4 inhibitor drugs have been developed from the base analogue molecule benzoylarginine amide due to its substrate specificity for the enzyme. Extensive research eventually confirmed that the terminal acetamidine functional group was responsible for the inhibition of PAD4 and that the incorporation of a haloaceteamidine warhead in place of the guanidium group, increased inhibitory potential.[8] Inactivation of PAD4 was discovered to work by the warhead's alkylation of the cysteine residue 645.[14] Inhibitors also preferentially bind to the active, calcium-bound form of PAD and that selectivity as well as improved inactivator potency could be achieved by modifying the non-warhead positions in order to stabilize enzyme-inhibitor interactions.[8] Many novel PAD inhibitors have been extensively studied and chemically designed for specificity, however there are yet drugs to be discovered that promote total PAD4 specificity. For example, Cl-amidine shows no selectivity in the cell and also targets other PAD family members. An image of some PAD inhibitors being currently studied is shown below:

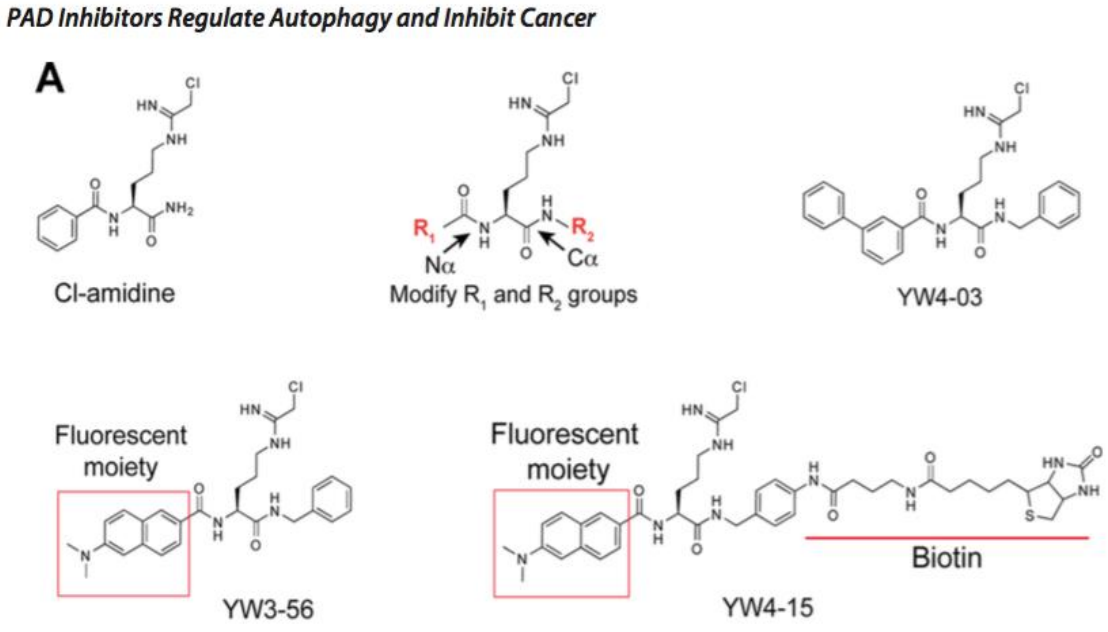

Figure 6: Novel PAD Inhibitors and Their Organic Structures [25]

The work of my previous graduate teacher, Dr. Xiangyun A. Chen, demonstrated the potential of metallorganic inducing drug inhibitors. Her western blot data and quantified cell viability shows the success of implementing a gold nano-particle derived inhibitor, called Nano-6E, into U2OS cancerous cell samples:

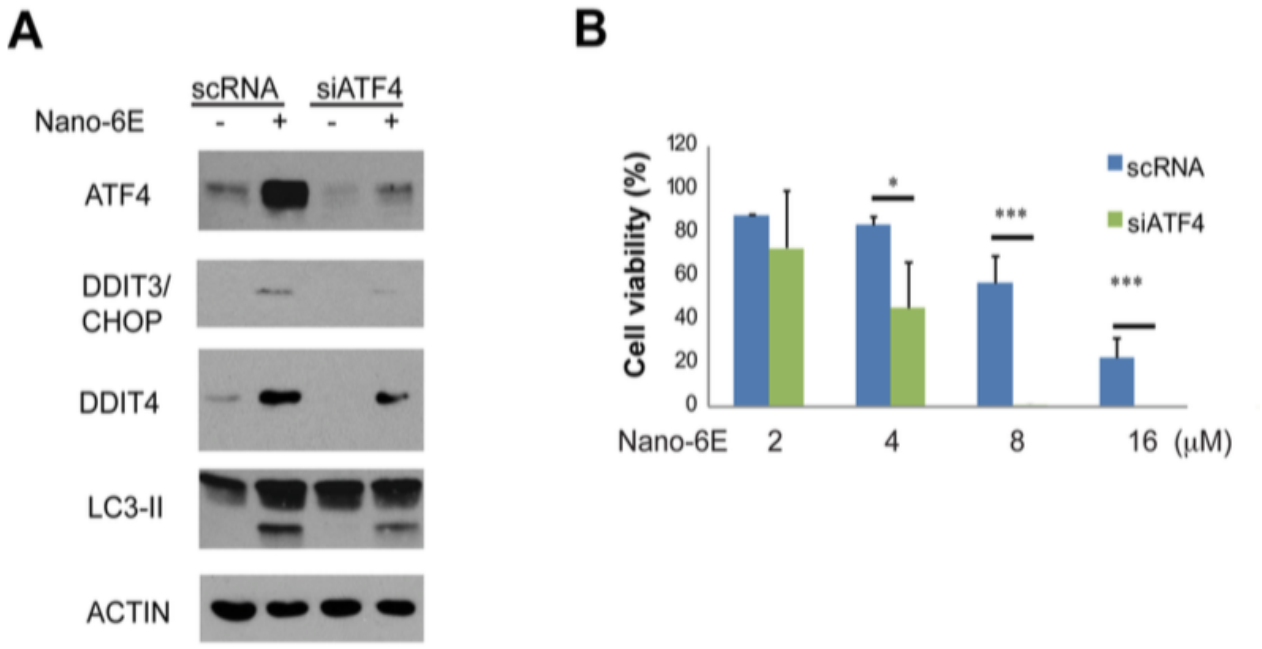

Figure 7: Western Blot and Cancerous Cell Viability After Treatment with Nano-6E [3]

In Figure 7, Chen demonstrates in a western blot that after Nano-6E treatment in U2OS cells, ATF4 knockdowns from either siATF4 or scRNA were attenuated in their DDIT3 and DDIT4 downstream signals. Cells were then transfected in a 96-well plate with siATF4 or scRNA and treated with Nano-6E and DMSO as a control in order to demonstrate cell growth by performing an ATP titer. This data concludes that ATF4 plays a critical role in the survival of U2OS cells after Nano-6E treatment.[3]

More recently efforts in research have been made in order to design drug inhibitors that target the PAD family but with sole specificity for the nuclear localized PAD4 isozyme. Specifically companies have recently designed a PAD inhibitor known as GSK-484 that supposedly has the potential to specifically target PAD4. Its molecular structure and IC_{50} values from Lewis and researcher can be found below:

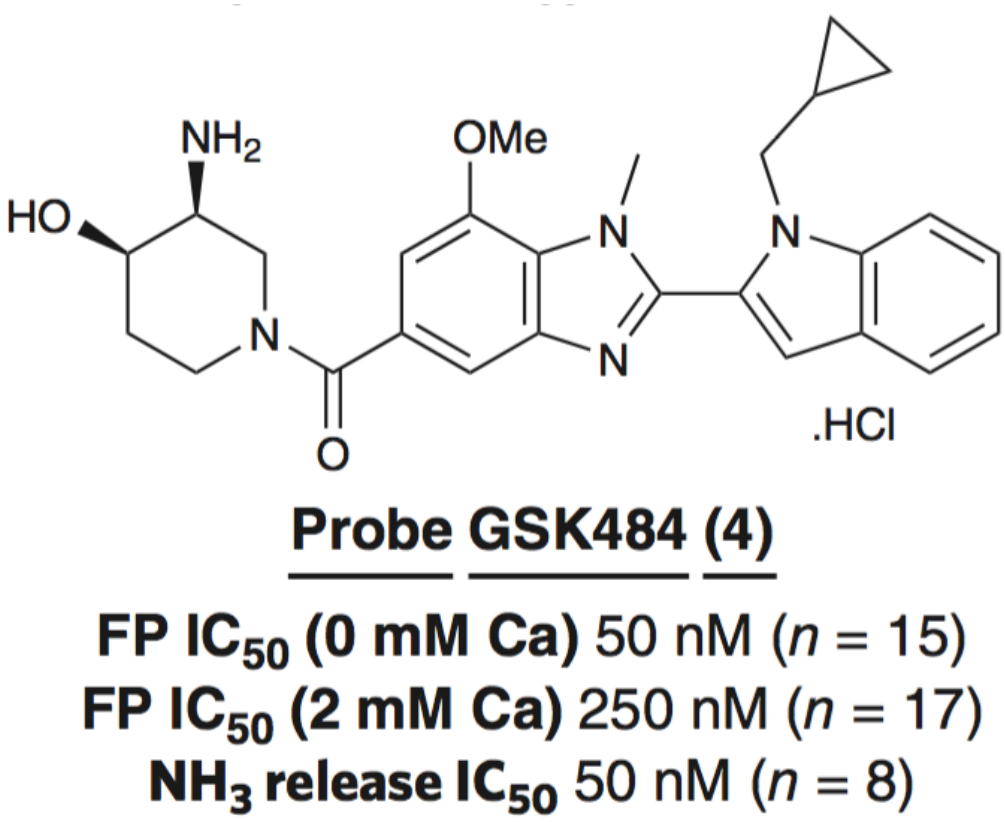

Figure 8: A Potential PAD4 Specific Inhibitor and IC_{50} Dosages [9]

Researchers have shown GSK-484 to mimic the phenotypic properties of PAD4 deficient mice in which there is a decrease in expression of neutrophil extracellular traps (NETs) of primary immune cells.[9] Molecules such as GSK-484 that have gone under extensive structural modification could have the potential to revolutionize the way drug inhibitors work under

conditions of cellular stress. Moreover, their relationship with PAD4 deactivation and mechanisms of inhibiting the hallmarks that compromise the characteristics of cancer, show promise in future chemotherapy and therapeutics regarding tumorigenesis.

VII. Experimentation and Methodical Approach

The goal of this background research is to give a better understanding of epigenetic regulation as well as the versatile mechanisms that compromise the homeostasis and balance of proper cellular function. The potential in studying these mechanisms can help scientists further understand the prevalence of diseases associated with gene awakening or silencing through the use of bioengineered methods such as the CRISPR-Cas9 system. Not only is the CRISPR system cost-effective, but it helps ameliorate the imperfect characteristics of utilizing small interfering RNA (siRNA), a widely applied technique. siRNA is helpful in knocking down a gene of interest, but never fully eradicates it from the genome due to partial instability. The CRISPR system maximizes upon these setbacks with a more efficient mechanism of delivery into the cell's genome, which enables the full knockout of genotypes (or knock-ins). Overall, methods such as siRNA knock-downs are transient and only manipulate levels of mRNA expression, whereas CRISPR directly targets and influences the genome resulting in longer lasting effects on the cell.

From this, it is convenient to use a CRISPR system when attempting to quantify the phenotypic differences between knock-in and knockout genotypes for a specific cell line and can provide insight into the cell's developmental regulation. Over the past the years I have compiled data that has included constructing a CRISPR knockout plasmid vector and performing a western

blot analysis on tumor suppressor gene factors by transfection into U2OS cells. Furthermore, I was also fortunate enough to learn and report on real time PCR (qPCR) technique by analyzing the effects of wild type and SMAD3 knockout cDNAs treated with SMAD pathway and DNA damage inducing reagents. Studying cell signaling patterns in regards to tumor suppressor function in conjunction with CRISPR biotechnology could help to establish a stable cell line that can be further used to study significant genes. Should a cell line not take for transfection, CRISPR technology can help to understand the genetic makeup of a mutated cell line and the reasons behind its instability. From working in the Yanming Wang laboratory at the Pennsylvania State University over the past two and a half years, I finally present and report on the experimental knowledge I have gained in regards to epigenetic regulation and its prevalence in cancer biology. It is of question to me, how the CRISPR-Cas9 system can be used to establish a knockout plasmid of PAD4, and if it will implement the recovery of normal cellular maintenance responses, or tumor suppressors, involved in the proliferation and growth of cancerous cell lines.

Chapter 2 – Results

I. The CRISPR-Cas9 Plasmid Vector Construction System

The regulatory properties of PAD4 as a p53 corepressor are of ultimate importance in the study of cancer biology. Because p53 is inherently a major checkpoint gene for cell cycle maintenance, researching PAD4 and its effects on tumor suppressor factors like p53 offer insight into possible cancer therapeutics that can attenuate the effects of uncontrollable growth. The application of a CRISPR-Cas9 system that implements the knockout of the PAD4-specific sequence can not only help in understanding the effects on cell cycle maintenance and tumor suppressing signals, but also test the application and efficiency of cutting edge bioengineering. From this, it is possible to better understand the mechanisms of cancer and potential treatments regarding-PAD4 deficient genotypes.

The system to incorporate the CRISPR-Cas9 vector was set up by using a 9,175 base pair sequence called pSpCas9 (BB) – 2A – Puro (PX459) V2.0. After PAD4-specific cDNA was made through PCR, a gel was run to confirm appropriate band size. The cDNA was then used to perform reverse transcription PCR in order to get the desired template required for the construct. Then the restriction enzymes BamH1 and EcoR1 were then implemented in a restriction digest that would create the two fragments necessary for ligating the desired plasmid construct. After the annealing reaction, the ligated plasmid construct was treated with, Bbs1 and Not1 restriction enzymes to determine whether an ineffective digestion took place with BamH1 and EcoR1. An

image from Addgene can be found below in order to illustrate the complexity and layout of the plasmid used for this experiment:

Figure 9: Graphic Design of the pSpCas9(BB)-2A-Puro (PX459) Plasmid for a PAD4-KO

Because the restriction site for Not1 is not associated with the location of insertion, it would not be found in the newly ligated construct. However, Bbs1 is in the portion of the plasmid that was removed by digestion and therefore should be present in the desired plasmid construct. Therefore, should an agarose gel be run and two different bands appear, this would be indicative

of an ineffective or incomplete digestion due to the plasmid being cut at both Bbs1 and Not1 sites. An agarose gel was run in order to confirm the successful incorporation of the PAD4 knockout into the plasmid after purification. Figure 10 below shows one band roughly in-between 9 and 10 kilobases, which would indicate the successful incorporation of the desired PAD4 sequence into the construct:

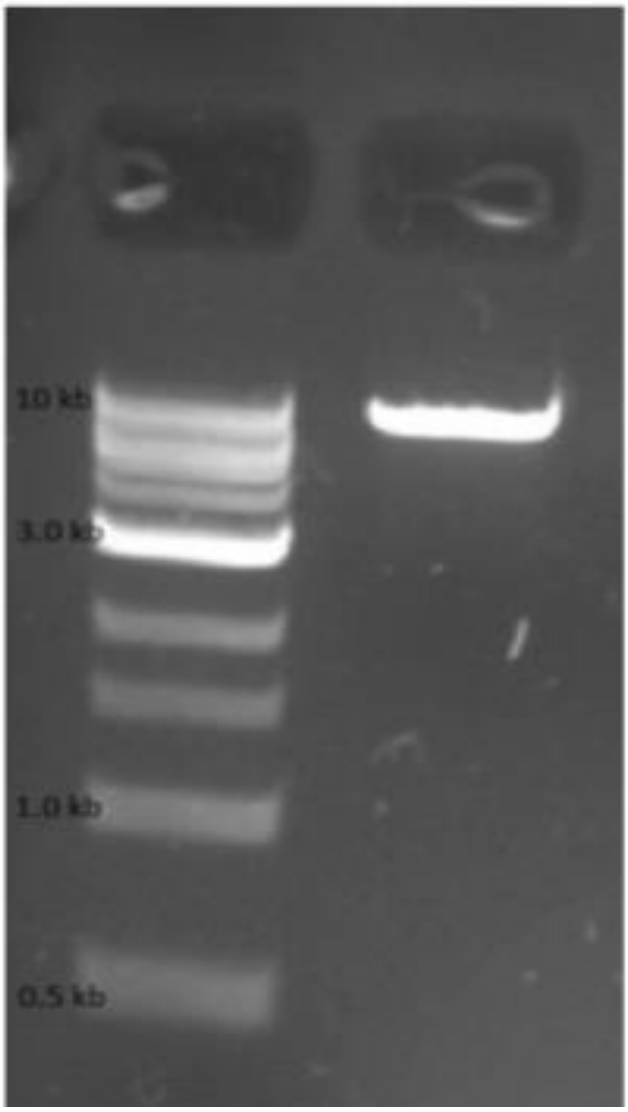

Figure 10: Gel Electrophoresis of CRISPR-Cas9 and PAD4 Knockout Plasmid Vector

Before the plasmid was transfected into U2OS cells, the purified plasmid was transformed in to bacterial *E. coli* cells for cloning purposes. In addition to the desired PAD4 target sequence, the plasmids also maintained an ampicillin resistant gene. *E. coli* cells were then grown in LB medium with added ampicillin. The surviving colonies were then selected and grown in order to perform a DNA isolation and extraction for sequencing. The plasmid DNA was extracted using the E.Z.N.A Plasmid DNA Mini Kit spin protocol and sent to the Genomics Core Facility at the Huck Institute of the Life Sciences at Penn State, which further confirmed the presence of the desire plasmid through sequencing.

II. Quantification of mRNA Signals By Spectrophotometry

Four different colonies that conferred ampicillin resistance in regards to the CRISPR construct were selected and inoculated in growth medium. After the growth period the four cultures were selected and the plasmid DNA was purified and extracted using the OMEGA Bio-tek Plasmid DNA Mini Kit. After the PAD4 knockout Crispr-Cas9 plasmid construct was purified and extracted from the bacterial cells, each of the four samples were quantified for their DNA using spectrophotometry (NanoDrop-1000 Spectrophotometer). The following data can be found below:

Sample	Concentration (ng/ µL)	Absorbance: 260/280	Absorbance: 260/230
PAD4-KO 1	318.3	1.88	2.32
PAD4-KO 2	199.8	1.89	2.09
PAD4-KO 3	239.8	1.88	2.01
PAD4-KO 4	312.3	1.89	2.36

Table 2: NanoDrop-1000 Spectrophotometry of PAD4 Knockout Plasmid Concentration

It was important to implement the purification and extraction of more than one plasmid for further study. By quantifying more than one colony, a stock of purified knockout plasmids can be saved in order to further investigate the correlation between PAD4 and other cellular signals involved in growth and proliferation. For example, it would be of interest to test the influence a PAD4 knockout cell line has on the expression of other p53-mediated regulatory signals like PUMA or GADD45. From this, scientists can slowly unveil downstream pathways and their influence over other targets inside the cell.

III. Analysis of Protein Signals in a Western Blot

After the PAD4 Knockout CRISPR-Cas9 construct was made and transfected into U2OS cells, they were set aside to grow. PAD4-KO 1 from Table 2 was the plasmid sample used for transfection due to it having the highest concentration. After the growth period was finished, a protein isolation and extraction was performed on the transfected U2OS sample and a western

blot was run against normal, wild-type U2OS cells in order to have a control to compare to. The aim was to quantify and demonstrate the effects that the loss of PAD4 has on an extensively proliferative cell line. Primary antibodies were used for β-actin, p21, and p53. It is important to note that unfortunately there was not enough of the PAD4 primary antibody to include in the western blot. Secondary antibodies included were mouse –HRP. The blotting signals can be found in the western blot below with U2OS wild-type cells and transfected PAD4 knockout cells:

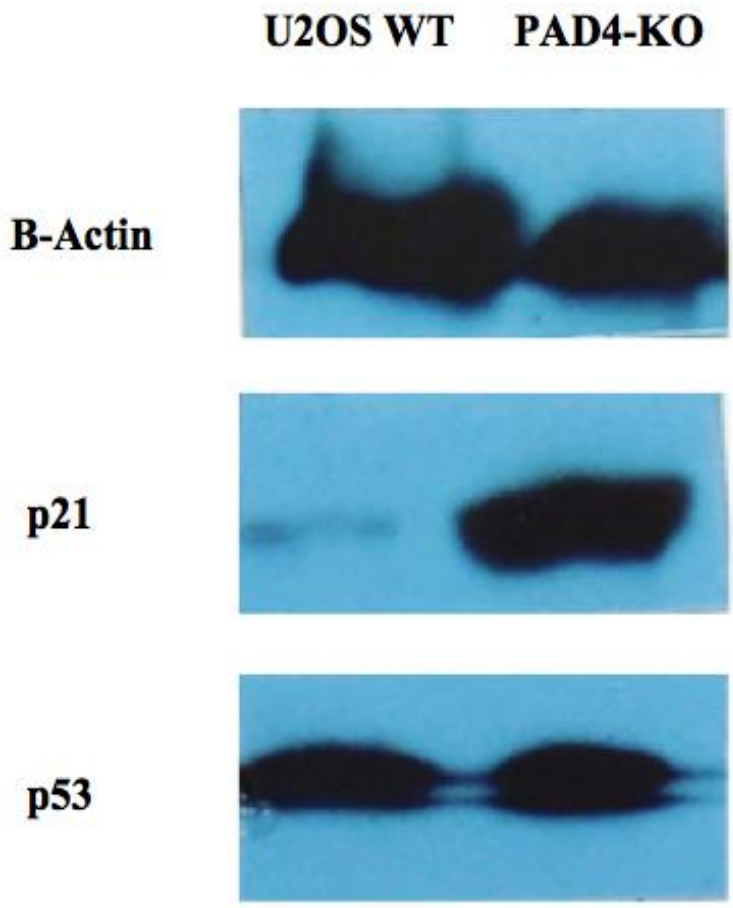

Figure 11: Western Blot of PAD4 Knockout CRISPR-Cas9 Plasmid Construct

IV. Analysis of mRNA Signals from RT-qPCR Heat Curve Analysis

It was of interest to my graduate student (Xiangyun Chen) and I to quantify the effects of mRNA expression of cell cycle regulatory transcripts in two different cell lines: a wild type mammary tumor cell line known as MMTV-PyMT (mammary tumor virus-driven polyoma middle T), and its respective SMAD3 knockout cell line. In addition, each cell line sample was treated with two reagents: 5FU and TGFβ-1. 5-Fluoricil works to induce DNA damage in cells by negatively affecting pyrimidine synthesis through metabolite inhibition in the cell and thus acts as an antitumorigenic reagent in halting the cell cycle. TGFβ-1 works in cell growth and proliferation and is an upstream regulator of the cell cycle. These reagents have potential to invoke or upregulate cellular stress responses and induce a p53-dependent pathway signal for apoptosis or repair. DMSO was used for untreated samples as a control.

Three sets of qPCR experiments were run for 6 cDNA's that were constructed from the reverse transcription of mammary tumor cells in MMTV-PyMT (mammary tumor virus-driven polyoma middle T) transgenic mice after PCR using the qScript cDNA Synthesis Kit (Quanta Biosciences). Lastly, a heat curve analysis was carried out using a real time - quantitative polymerase chain reaction (RT-qPCR) system. Each of the three experiments that were run have three graphical representations of the qPCR analysis: one for a wild type and two that analyzed the effects of the SMAD3 Knockout. Each experiment set had a respective set of primers that were being analyzed. The wild-type qPCR analyzed the effects of actin (control), ATF3, and ATF4 primers, whereas the SMAD3 knockout RT-qPCR analyzed the effects of the aforementioned three primers as well as GAPDH (control), p21, p53, CXCR4, CDK1, and PAD4. The data can be found below:

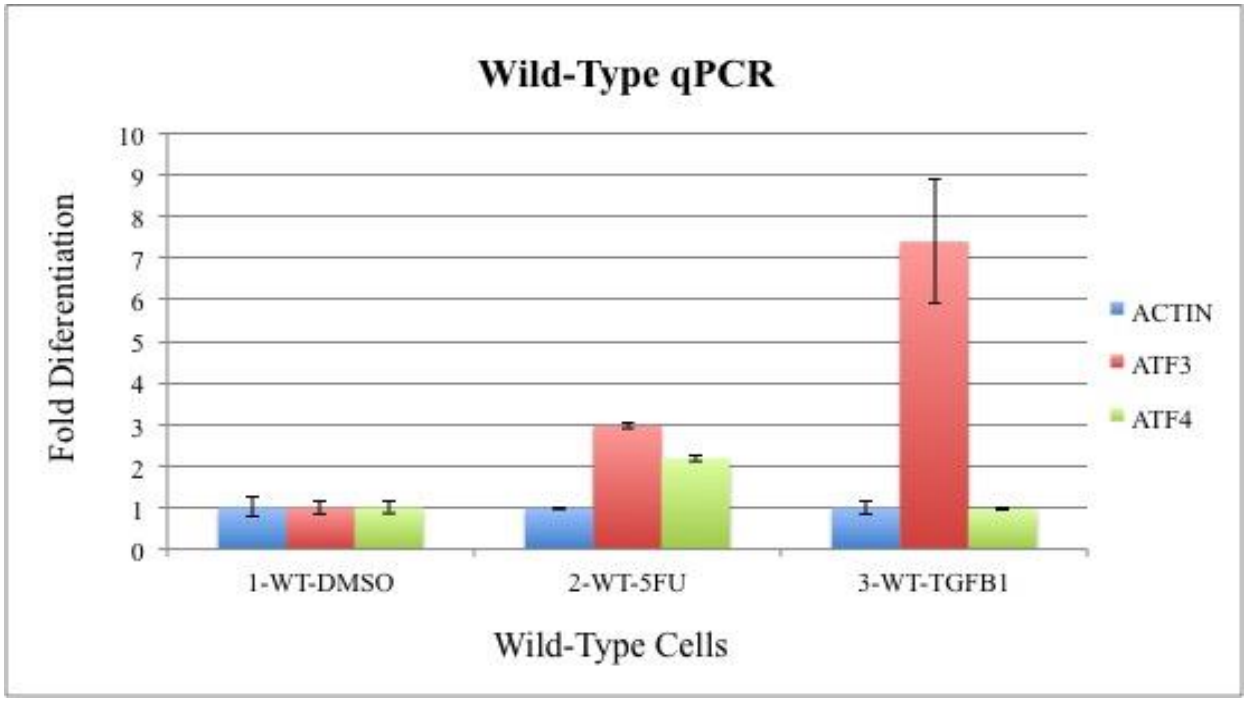

Figure 12: Real Time Quantitative PCR in Wild Type Cells and 3 Primers

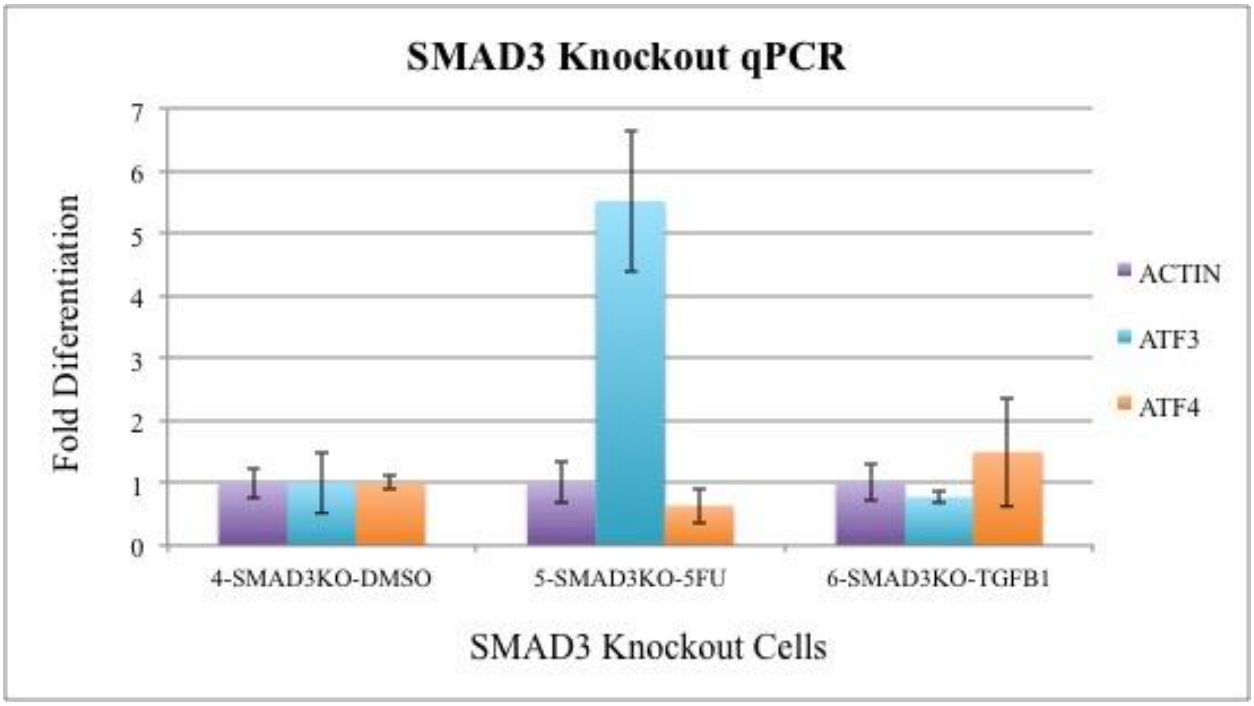

Figure 13: Real Time Quantitative PCR in SMAD3 Knockout Cells for 3 Primers

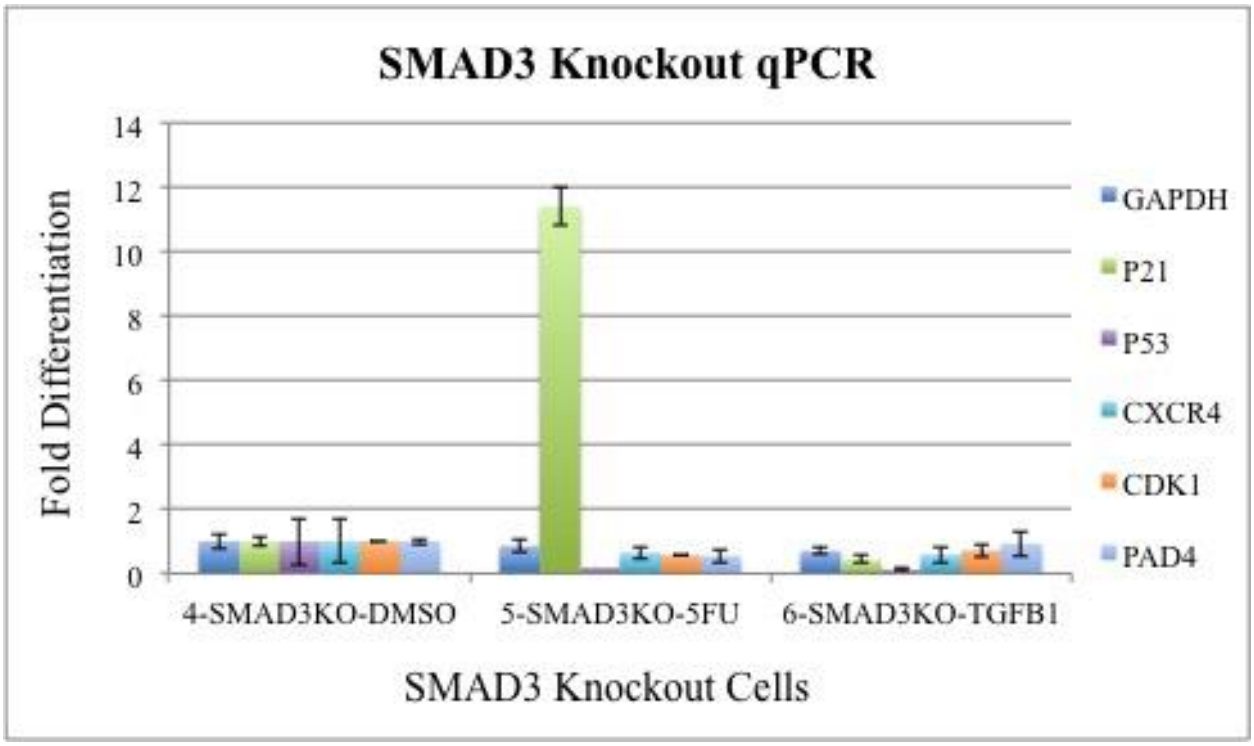

Figure 14: Real Time Quantitative PCR in SMAD3 Knockout Cells for 6 Primers

Chapter 3 – Discussion

Once the plasmid starting vector of pSpCas9(BB)-2A-Puro (PX459) was digested with EcoR1 and BamH1, the PAD4 knockout reverse transcription PCR product was combined with the fragments in order to perform a ligation reaction to form the desired new plasmid construct. In order to check if the correct construct was formed, an agarose gel was run using electrophoresis. In Figure 10, the gel shows a single band around 9 kilobases, which is indicative that the ligation reaction worked. If the construct digestion and ligation reactions did not work, two separate, smaller bands would appear on the gel indicating an ineffective digestion from BamH1 and EcoR1restriction enzyme cutting at the two different locations. Because there was a single band, this shows the effective cut with Not1 restriction enzyme and removal of the Bbs1 restriction enzyme site to form the desired plasmid vector of a PAD4 knockout.

When quantifying the real time quantitative PCR data through the NanoDrop Spectrophotometer, the four PAD4 knockout plasmid DNA sample concentrations were well within a usable and efficient concentration, which is visible in Figure 2. The absorbances for the 260/280 ratio for the experiment showed values just under 2. This is generally accepted as relatively pure DNA and therefore the extraction and purification of the samples worked well. Furthermore the 260/230 ratio was around 2 for the plasmid DNA quantification with the exceptions of samples PAD4-KO 1 and PAD4-KO 4, which gave absorbances around 2.32 and 2.36, respectively. This might be indicative of slight protein contamination or organic compounds that stayed in solution.

The importance of establishing a CRISPR plasmid construct with a desired mutation such as knocking out PAD4 is paramount in deciphering the limitations of plasmid transfection into mammalian cells. There are several genes in genetic studies that are deemed essential and

without them the organism cannot survive; an example would be the ORI site in bacterial plasmids, which without, the bacterium would not be able to proliferate due to the lack of a start site for DNA replication. Therefore, depending on the gene of focus, certain sequences of other genes could be involved in concert with the one of interest and may not be applicable for the organism in terms of their survival. This could be a reason as to why a certain cell line is unstable or not viable. A lot of considerations must be accounted for bioengineering a way to silence or overexpress a gene in certain cells.

In the real time quantitative PCR experiments, the reagents 5FU and TGFβ-1 were used to induce differential effects on the initial cellular line as previously mentioned. In the first set of experiments of real time quantitative PCR for wild type cells, treatments showed increases in mRNA transcripts for ATF3 by 3 fold for 5FU and roughly 7 fold for TGFβ-1 in Figure 11. Because TGFβ-1 works in concert with ATF3 in the ATF gene network pathway during the cellular response of controlling growth, it would seem like that there would be a fold increase in ATF3. Many types of cancers are known to have a lost sensitivity to the growth inhibitory actions of TGFβ-1 and thought to be an important step in the process of oncogenic transformation.[21] In addition, it is important to understand the high versatility of ATF3 due to the fact that it is not only activated a big number of cellular responses under stress and growth inhibition, but also influences many separate and different downstream signals to invoke cellular homeostasis.[7]

Furthermore, ATF3 transcripts surpassed a 5 fold increase after treatment with 5FU in the SMAD3 knockout cells of Figure 12. Research has previously shown that 5FU induces DNA damage and therefore works to halt cellular metabolism and the cell cycle in order to induce repair mechanisms or initiate apoptosis. ATF3 is also known to work in conjunction with p53 in

order to promote tumor suppressor functions, however in Figure 13 it is evident that p53 is actually downregulated in SMAD3 knockout cells with the addition that p21 (a cyclin-dependent kinase inhibitor) is over expressed. How could one explain this anomaly?

It seems that there could be some affinity for SMAD3-p53 complexation in the cell to regulate and maintain cellular homeostasis due to the fact that 5FU and TGFβ-1 treatments to stimulate cellular repair and maintenance have decreased p53 levels compared to the DMSO control. Buenemann reports that TGFβ-1 induces SMAD signaling to promote cell cycle arrest and apoptosis.[2] It is my belief that due to the absence of SMAD3, nuclear localization of SMAD4 cannot be activated in order to activate tumor suppressor factors and inhibit growth and therefore a p21-induced signal must be overcompensated to take on the lost cellular maintenance that the SMAD pathway normally carries out. Therefore p21 helps growth arrest that precedes cellular terminal differentiation in an attempt to stop the cell cycle. There is no upregulation of p21 in TGFβ-1 treated SMAD3 knockout cells due to the fact that its main role was to activate the SMAD3 pathway but there is only a decreased cellular response from transcript measurements in Figure 13. Because no DNA damage was induced (like 5FU treatments), there was no increase in transcript levels of p21. ChIP assays showed that TGFβ-1 normally stimulates p53, SMAD4, and SMAD2/3 binding in an effort to initiate apoptosis.[26] Protein 53 is known to work with the SMAD4 promoter in order to inhibit cell growth or proliferation. However, because SMAD3 was absent from the treatment, there would be no need to upregulate p53 transcripts because they cannot work in concert with absent SMADs to invoke a response. Therefore without SMAD3, SMAD4 cannot be activated to interact with p53 and using TGFβ-1 treatment does not seem to help this.

The western blot in Figure 14 shows β-actin working as a control for both samples: the wild type U2OS cells and the transfected CRISPR-Cas9 PAD4 knockout plasmid that was constructed. It is evident that in the wild type U2OS cells, p21 had very low expression levels whereas in thePAD4 knockout sample, there is an intense increase in the blot indicating higher expression levels in the manipulated sample. It is curious, however, to note that p53 expression levels were unaffected by the PAD4 knockout. As mentioned in the background material, PAD4 acts as a corepressor for p53 and thus prevents it from localizing to promoters of specific tumorigenic suppressing transcription factor genes such as p21. It is known that PAD4 is widely expressed in the nucleus of many cancerous cells or proliferative cell lines that lose control of maintaining their cell cycles. Protein 21 is a cyclin-dependent inhibitor of the cell cycle and helps maintain proper growth homeostasis so uncontrollable proliferation does not occur. This specific feature of lose in control of growth is one of the aforementioned cancer hallmarks. It seems that dysregulating PAD4 allowed for p53 to recover to localizing to the promoter region that it normally controls for p21's expression and thus reinstates p21 maintenance of the cell cycle. In the blot the recovery is evident. Because PAD4 only complexes (corepresses) with p53 it does not affect the transcriptional levels that occur inside the cell, which could be a potential explanation for the unchanged expression levels of p53. (SMAD3 was also tested in this blot, but no signals were detected, which most likely occurred due to improper antibody dilutions)

The application of genomic engineering with methods like CRISPR holds potential in furthering our understanding of cancer biology. Pulmonary metastasis is observed in 80% - 94% of tumor bearing female mice in addition to them developing palpable and mammary tumors as early as five weeks of age. Establishing knock-in and knockout cell lines that are stable enough to implement into certain cancerous cell lines can give detail on the growth rate of the separate

genotypes. Three established genotypes of MMTV-PyMT tumor-bearing mice have given statistically significant insight on the length of survival over multiple days in which growth has slowed. $PAD4^{+/+}$ transgenic mice have been observed to show minimal survival at 93 days regarding cancerous growth when compared to PAD4 heterozygous mice at 122 days, and $PAD4^{-/-}$ observed to survive until day 138.[3] This poses the question as to why the tumors are growing slower. Could the prolonged survival work as a mechanism to promote a longer time span for potential tumor growth or metastasis? This becomes difficult to discern due to the fact that PAD4 is a versatile gene with many downstream targets. Although PAD4 overexpression in cancerous tissue is an established trend in research, other observations have resulted in questioning the roles of PAD4's regulatory properties in the immune response. Researchers have observed that the dysregulation of PAD4 in breast cancer cells causes the loss of GSK3β activity in suppressing tumor invasiveness and results in epithelial-to-mesenchymal transition.[22] With a loss in GSK3β activity, the TGFβ pathway is then reinstated resulting in growth. This raises question to PAD4's role in cellular differentiation and in this case seems to cause metastasis. The role of PAD4 activity is still not fully understood, but is rapidly becoming a prevalent area of research.

Further studies of cancer biology and PAD4 gene editing could involve establishing a PAD4-KO PyMT cell line to use for a xenograft into wild type and PAD4 knockout transgenic sibling mice. Another area of potential focus is quantifying the role of PAD2 in embryonic growth and differentiation of cells in conjunction with PAD4. Experimenting with CRISPR to establish a viable cell line of both PAD2 and PAD4 knockout genotypes in MCF7 or 4T1 breast cancer cells could help to further understand this intriguing class of enzymes. By further performing a metastasis assay of the breast cancer cells one can quantify the effects of the

bioengineered knockouts on their growth and proliferation while simultaneously studying cell maintenance signals through western blot analysis.

Overall molecular genetics and bioengineering provide valuable insight on prevalent diseases like cancer. The genetic makeup of cancer is one that is confusing More specifically, the study of the loss of p53 as well as PAD4's counteracting characteristics at the molecular level will help biologists and pharmacologists better understand the mechanisms at play when cellular homeostasis is disturbed. From this, novel cancer therapeutics can be implemented through drug specific inhibition or engineering ways to implement efficient gene editing.

Chapter 4 – Materials and Methods Supplement

CRISPR-Cas9 Plasmid Vector Construct for PAD4 Knock-out

First a 100 μL restriction digest was set up using 1 μg of pCas9 vector with the restriction enzymes BamH1 and EcoR1. Primers specific to the desired PAD4 construct were ordered from Integrated DNA Technologies. After the digestion, the annealing of a top and bottom oligo-specific sequence was carried out by using their respective single-guide RNAs (sgRNA). The total PCR reaction system was diluted to 200 μL and utilized T4 PNK ligase and its buffer. After using the PTC-200 Peltier Thermo Cycler in PCR to make the necessary PAD4-specific cDNA, reverse transcription PCR was implemented in order to make the required DNA template for the ligation (Quanta Biosciences). A ligation reaction was done by combining the annealed oligo-sequence and the digested Cas9 plasmid vector with the cDNA target sequence in a 25 μL system. The reaction sat for 2.5 hours before ligation was complete. After the ligation was complete the construct was added to competent bacterial cells. The sample was iced for 25 minutes and then heat shocked in a 42°C water bath for 45 seconds. Immediately following the heat shock the sample was put back on ice for 2 minutes. Next 200 μL of ampicillin resistant bacterial growth medium was added to tubes and placed in a 37°C shaker for 20 minutes. Post incubation, the sample was then spread on prewarmed ampicillin resistant LB plates and placed in the incubator for roughly half a day. After this transformation, colonies were picked form each plate and inoculated with liquid bacterial culture and placed in the 37°C incubator over night. Left over plasmid samples were stored in the -80°C freezer for future use. A similar procedure was done for the transfection of the CRISPR plasmid construct into U2OS cells for further western blot analysis.

CRISPR-Cas9 Plasmid Purification and Isolation

Inoculated bacterial colonies were collected from over night incubation at 37°C and prepped for purification. The E.Z.N.A Plasmid DNA Mini Kit spin protocol was used for this step in order to extract the construct of interest (OMEGA Bio-Tek). After this procedure, the samples were taken to the NanoDrop -1000 Spectrophotometer to quantify the concentration of the recovered plasmid. Lastly, a double enzyme restriction digest was performed on the plasmid construct and a 2% agarose gel ran to confirm the successful ligation and incorporation of the knocked out sequence. A single intense band was confirmed at roughly 9 kb indicating a successful transformation. The sample of the plasmid was sent to the Genomics Core Facility at the Huck Institute of the Life Sciences at Penn State for further sequencing and confirmation of the desire plasmid.

U2OS Cellular Growth and Maintenance

Cells were retrieved from liquid nitrogen cryopreservation and cultured in DMEM (Gibco) that was supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin antibiotics in a 5% CO_2 incubator at 37°C. Cells were monitored and cleaned with DPBS (Gibco) accordingly. After the cells reached roughly 85% - 90% confluency, they were split for further use. Cells were then isolated for protein and RNA extraction after respective transfections.

Immunoprecipitation and Purification of Protein and mRNA from Mammalian Cells

For protein extractions, reagents that include IP buffer, and four protease inhibitors (aprotinin, leupeptin, pepstatin, and PMSF) were gathered and place on ice in preparation for immunoprecipitation. Meanwhile, adherent MDA-MB-231 breast cancer cells were washed with PBS and trypsinized. The cell density was calculated for the appropriation of reagent volumes in the protein extraction. The samples were split in two for an RNA extraction and a protein extraction and centrifuged. The volume of PMSF added was 1:100 in relation to the amount of IP buffer added and all other protease inhibitors were added in a 1:1000 fashion. After the reagents were combined with the cell pellet, the sample was sonicated for 15 minutes and then put on ice. During protein sonication, RNA was extracted from the second sample using PureLink RNA Mini Kit (Ambion – Life Technologies). Then the RNA samples were quantified using the NanoDrop -1000 Spectrophotometer in preparation for western blotting and RT-PCR. The concentration levels measured for RNA were found to be within an optimal range for further experimentation.

Quantitative Reverse Transcriptase – Polymerase Chain Reaction (Q-RT-PCR)

RNA was isolated from three different samples: wild-type, SMAD3 knockout, and p53 knockout cells. These three samples were initially treated each with 5FU and TGFβ-1 reagents. The PureLink RNA Mini Kit was used for the extraction (Ambion - Life Technologies). 0.5 μg of the RNA was converted to cDNA using the qScript cDNA Synthesis Kit (Quanta Biosciences) and PTC-200 Peltier Thermo Cycler. The cDNA was then diluted in a 1:10 fashion and used for quantitative PCR using SYBR Green and the Applied Biosciences RT-PCR machine. A

combination of 9 forward and reverse primer sets was implemented (actin, ATF3, ATF for one experiment and GAPDH, p21, p53, CXCR4, CDK1, and PAD4 for another). Results show the fold change in expression levels of the different RNAs of interest.

Western Blotting

The protein concentration of the extracted sample was analyzed by running a gel and fixing it in solution, followed by coomassie blue staining. Destaining followed. It was decided to use a 13.2% SDS-PAGE gel due to the proteins of interest being under 80 kDa. The extracted protein samples were loaded into the wells of the gel and then ran at 220 volts for roughly 40 minutes or until the bromophenol blue dye has run out of the gel. The gel was transferred onto nitrocellulose with transfer buffer (20% methanol in 1 x transfer buffer) at a constant voltage at 220 Amps for 1 hour using a TE-77-ECL Semi-Dry Transfer Unit (Amersham Biosciences). Books were put on top to facilitate the transfer. Staining the nitrocellulose was then carried out using Ponceau S dye for roughly 3 minutes followed by destaining with deionized H_2O until the bands appeared clear. The nitrocellulose membrane was then cut according to the antibodies used for the assay (β-actin, p21, p53, and SMAD3) and where the associated protein size markers appeared. The membrane was then washed with TBST buffer. After, a 5% blotto solution was made with 5 grams of dry milk solution in 100 mL of TBST buffer. The membrane was then combined with 4 mL of the blotto solution in sealable pouch pockets and the appropriated amount of antibody. The milk pouches were attached to a rotator in a 4°C fridge for overnight incubation. The next day the membrane was washed in three intervals of 10 minutes in TBST buffer solution. Post wash, a secondary antibody was added to new milk pouches with the

membrane and incubated on the 4°C rotator for an additional two hours. The membranes were then washed three times again with TBST buffer in 10 minute intervals. Lumifluorescent reagents were then applied to the membranes and taken to the dark room to develop the blots on X-Ray films.

BIBLIOGRAPHY

[1] Bicker, K.L. and Thompson, P.R. (2013). The protein arginine deiminases (PADs): Structure, Function, Inhibition, and Disease. *Biopolymers* 99 (2): 155-63.

[2] Buenemann, C.L., Willy, C., Buchmann, A., Schmiechen, A., and Schwarz, M. (2001). Transforming growth factor-β1-induced Smad signaling, cell-cycle arrest and apoptosis in hepatoma cells. *Carcinogenesis* 22 (3): 447-52.

[3] Chen, X.A. (2015). Novel self-assemble nanoparticles to inhibit cancer growth - Effects on ER stress, autophagy, and multidrug resistance by a small molecule. *A Thesis in The Molecular Cellular and Integrative Bioscience Graduate Program*: 1-180.

[4] Elston, R and Inman, G.J. (2012). Crosstalk between p53 and TGF-β Signaling. *Journal of Signal Transduction* 2012: 1-10.

[5] Guo, Q. and Fast, W. (2011). Citrullination of Inhibitor of Growth 4 (ING4) by Peptidylarginine Deiminase 4 (PAD4) Disrupts the Interaction between ING4 and p53. *The Journal of Biological Chemistry* 286 (19): 17069-7078.

[6] Gupta, S.K. and Shukla, P. (2016). Gene editing for cell engineering: trends and applications. *Critical Reviews in Biotechnology*: 1-14.

[7] Jiang, H., Wek, S.A., McGrath, B.C., Lu, D., Hai, T., Harding, H.P., Wang, X., Ron, D., Cavener, D.R., and Wek, R.C. (2004). Activating Transcription Factor 3 Is Integral to Eukaryotic Initiation Factor 2 Kinase Stress Response. *Molecular and Cellular Biology* 24 (3): 1365-377.

[8] Jones, J., Causey, C., Knuckley, B., Slack-Noyes, J.L., and Thompson, P.R. (2009). Protein

arginine deiminase 4 (PAD4): current understanding and future therapeutic potential. *Current Opinion in Drug Discovery and Development* 12 (5): 616-27.

[9] Lewis, H.D., Liddle, J., Coote, J.E., Atkinson, S.J., Barker, M.D., Bax, B.D., Bicker, K.L., Bingham, R.P., Campbell, M., Chen, Y.H., Chung, C., Craggs, P.D., Davis, R.P., Eberhard, D., Joberty, G., Lind, K.E., Locke, K., Maller, C., Martinod, K., Patten, C., Polyakova, O., Rise, C.E., Rüdiger, M., Sheppard, R.J., Slade, D.J., Thomas, P., Thorpe, J., Yao, G., Drewes, G., Wagner, D.D., Thompson, P.R., Prinjha, R.K., and Wilson, D.M. (2015). Inhibition of PAD4 activity is sufficient to disrupt mouse and human NET formation. *Nature Chemical Biology* 11: 189-94.

[10] Li, P. (2010). Analysis of Peptidylarginine Deiminase 4 and Histone Citrullination in Transcriptional Regulation and Innate Immunity. *A Dissertation in Genetics* 1-236.

[11] Li, P., Yao, H., Zhang, Z., Li, M., Luo, Y., Thompson, P.R., Gilmour, D.S., and Wang, Y. (2008). Regulation of p53 Target Gene Expression by Peptidylarginine Deiminase 4. *Molecular and Cellular Biology* 28 (15): 4745-758.

[12] Liu, Y., Chiang, Y., Liu, G., and Hung, H. (2011). Functional Role of Dimerization of Human Peptidylarginine Deiminase 4 (PAD4). *PLoS ONE* 6 (6): 1-11.

[13] Lu, C., Jain, S.U., Hoelper, D., Bechet, D., Molden, R.C., Ran, L., Murphy, D., Venneti, S., Hameed, M., Pawel, B.R., Wunder, J.S., Dickson, B.C., Lundgren, S.M., Jani, K.S., Jay, N.D., Papillon-Cavanagh, S., Andrulis, I.L., Sawyer, S.L., Grynspan, D., Turcotte, R.E., Nadaf, J., Fahiminiyah, S., Muir, T.W., Majewski, J., Thompson, C.B., Chi, P., Garcia, B.A., Allis, C.D., Jabado, N., and Lewis, P.W. (2016). Histone H3K36 promote sarcomagenesis through altered histone methylation landscape. *Science* 352 (6287): 844-49.

[14] Luo, Y., Arita, K., Bhatia, M., Knuckley, B., Lee, Y.H., Stallcup, M.R., Sato, M., and Thompson, P.R. (2006). Inhibitors and inactivators of protein arginine deiminase 4: functional and structural characterization. *Biochemistry* 45 (39): 11727-1736.

[15] Macleod, K.F., Sherry, N., Hannon, G., Beach, D., Tokino, T., Kinzler, K., Vogelstein, B., and Jacks, T. (1995). p53-Dependent and independent expression of p21 during cell growth, differentiation, and DNA damage. *Genes & Development* 9: 935-44.

[16] Park, J., Wysocki, R.W., Amoozgar, Z., Maiorino, L., Fein, M.R., Jorns, J., Schott, A.F., Kinugasa-Katayama, Y., Lee, Y., Won, N.H., Nakasone, E.S., Hearn, S.A., Küttner, V., Qiu, J., Almeida, A.S., Perurena, N., Kessenbrock, K., Goldberg, M.S., and Egeblad, M. (2016). Cancer cells induce metastasis-supporting neutrophil extracellular DNA traps. *Science Translational Medicine* 8 (361): 1-12.

[17] Pfister, S.X., Markkanen, E., Jiang, Y., Sarker, S., Woodcock, M., Orlando, G., Mavrommati, I., Pai, C., Zalmas, L., Drobnitzky, N., Dianov, G.L., Verrill, C., Macaulay, V.M., Ying, S., La Thangue, N.B., D'Angiolella, V., Ryan, A.J., and Humphrey, T.C. (2015). Inhibiting WEE1 Selectively Kills Histone H3K36me3-Deficient Cancers by dNTP Starvation. *Cancer Cell* 28: 557-68.

[18] Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* 8 (11): 2281-308.

[19] Rochette, L., Guenancia, C., Gudjoncik, A., Hachet, O., Zeller, M., Cottin, Y., and Vergely, C. (2015). Anthracyclines/trastuzumab: new aspects of cardiotoxicity and molecular mechanisms. *Trends in Pharmalogical Sciences*: 1-23.

[20] Sabulski, M.J., Wang, Y., and Pires, M.M. (2015). PAD2 Activity Monitored via a Fluorescent Substarte Analog. *Chemical Biology and Drug Design* 86: 599-605.

[21] Satterwhite, D.J. and Moses, H.L. (1995). Mechanisms of transforming growth factor-beta 1-induced cell cycle arrest. *Invasion Metastasis* 14 (1-6): 309-18.

[22] Stadler, S.C., Vincent, C.T., Fedorov, V.D., Patsialou, A., Cherrington, B.D., Wakshlag, J.J., Mohanan, S., Zee, B.M., Zhang, X., Garcia, B.A., Condeelis, J.S., Brown, A.M.C., Coonrod, S.A., and Allis, C.D. (2013). Dysregulation of PAD4-mediated citrullination of nuclear GSK3β activates TGF- β signaling and induces epithelial-to-mesenchymal transition in breast cancer cells. *Proceedings of the National Academy of Sciences* 110 (29): 11851-1856.

[23] Tohme, S., Yazdani, H.O., Al-Khafaji, A.B., Chidi, A.P., Loughran, P., Mowen, K., Wang, Y., Simmons, R.L., Huang, H., and Tsung, A. (2015). Neutrophil Extracellular Traps Promote the Development and Progression of Liver Metastases after Surgical Stress. *Cancer Research*: 1-40.

[24] Wang, S., Chen, X. A., Hu, J., Jiang, J., Li, Y., Chan-Salis, K. Y., Gu, Y., Chen, G., Thomas, C., Pugh, B.F., and Wang, Y. (2015). ATF4 Gene Network Mediates Cellular Response to the Anticancer PAD Inhibitor YW3-56 in Triple-Negative Breast Cancer Cells. *Molecular Cancer Therapeutics* 14 (4): 877-88.

[25] Wang, Y., Li, P., Wang, S., Hu, J., Chen, X.A., Wu, J., Fisher, M., Oshaben, K., Zhao, N., Gu, Y., Wang, D., Chen, G., and Wang, Y. (2012). Anticancer Peptidylarginine Deiminase (PAD) Inhibitors Regulate the Autophagy Flux and the Mammalian Target of Rapamycin Complex 1 Activity. *The Journal of Biological Chemistry* 287 (31): 25941-5953.

[26] Wilkinson, D.S., Ogden, S.K., Stratton, S.A., Piechan, J.L., Nguyen, T.T., Smulian, G.A., and Barton, M.C. (2005). A direct intersection between p53 and transforming growth factor beta pathways targets chromatin modification and transcription repression of the alpha-

fetoprotein gene. *Molecular Cell Biology* 25 (3): 1200-212.

[27] Yan, C. and Boyd, D.D. (2006). ATF3 Regulates the Stability of p53. *Cell Cycle* 5 (9): 926-29.

[28] Yao, H., Li, P., Venters, B.J., Zheng, S., Thompson, P.R., Pugh, B.F., and Wang, Y. (2008). Histone Arg Modifications and p53 Regulate the Expression of OKL38, a Mediator of Apoptosis. *The Journal of Biological Chemistry* 283 (29): 20060-0068.

[29] Yin, X., Wolford, C.C., Chang, Y., McConoughey, S.J., Ramsey, S.A., Aderem, A., and Hai, Tsonwin, H. (2010). ATF3, an adaptive-response gene, enhances TGFβ signaling and cancer-initiating cell features in breast cancer cells. *Journal of Cell Science* 123: 3558-565.

[30] Yoo, C.B. and Jones, P.A. (2006). Epigenetic therapy of cancer: past, present and future. *Nature Reviews Drug Discovery* 5 (1): 37-50.

ACADEMIC VITAE

Alexander Braun Radaoui
abr5163@psu.edu

Education

- Schreyer Honors College at The Pennsylvania State University
    i. Biology (Neuroscience) Major, Biochemistry and Molecular Biology Minor, and Chemistry Minor (Fall 2013 – Spring 2017)

Research

- Yanming Wang Lab
    i. Center for Eukaryotic Gene Regulation (2014 – 2017)
    ii. Plasmid vector CrisprCas9 construct, PAD protein knockouts, breast cancer metastasis, cancer cell drug (PAD-inhibitor) treatment and response (thesis work)

Honors

- The Dean's List
    i. Eberly College of Science (Fall 2013/2015/2017, Summer 2014/2015, Spring 2015)

Activities

- Cardiology/Plastic Surgery Shadowing Rotations (67 hours: June – August 2016)
- Disease Dynamics Research in Tanzania, Africa for 3 weeks (May/June 2016)
- Tropical Field Ecology Research in Costa Rica for 2 weeks (January 2016)
- Cancer Biology Research/Academic Abroad in Shanghai, China (May/June 2015)
- Public Relations Director for Science Lion Pride (2015 – 2016)
- Historian Chair for Science LionPride (2016 – 2017)
- Springfield (2016 – 2017)
- THON volunteer (Rules & Regulation, 2015 – 2016)
- Tennis Club

Volunteer Experience

- Student Tutor (2014 – 2017)
    i. Biochemistry & Molecular Biology and Organic Chemistry/General Chemistry
- Haiti Donations
    i. Sociology 119 at Penn State; raising awareness for those affected

Work Experience

- Family Medicine Associates of Wyomissing
    ii. Medical filing and shadowing of family doctors
- Lifeguarding at the Wyomissing Pool (Summers, 2012 & 2013)
- Penn State Chemistry Department TA (Fall 2016)