Carlee Luttrell

Honors Thesis

Honors College Degree

Biochemistry and Molecular Biology

Ruhl Laboratory

December 17, 2017

Acknowledgements:

I would like to recognize my research mentor, Dr. Donald Ruhl, for his continued support and constant guidance with experimentations and directions of research throughout my time working in his laboratory. Without his commitment to providing learning opportunities for his students, my work would not be possible. I would like to thank Dr. Jessica Matts and Judy Hall for their day-to-day efforts and advising in the laboratory setting as well as Dustin Steele for his research contributions that directed my work. Finally, I would like to thank the Freshman Research Scholars and CASNR – OAES scholarship programs for assistance in funding my undergraduate research project.

Abstract:

This study is focused on understanding the phosphorylation sites within INI-1 DNA binding domains to understand the INI-1 protein's connections to atypical teraroid/rhadboid tumor (AT/RT), an aggressive pediatric neural cancer. A truncation (1-186 amino acids) of the INI-1 plasmid was discovered to bind DNA with the same functionality as the full-length plasmid. Other studies directed interest within this same truncation toward a phosphorylation site, serine-129, to consider as a specific active site in the mysterious functions of INI-1. This identified truncation, as well as the full-length protein, was manipulated by site directed mutagenesis. The two mutations, serine to alanine and serine to aspartic acid, serve as phosphorylated and non-phosphorylated mimics, and provide quantifiable data regarding its DNA binding activity through EMSA gels.

Project Title:

Phosphorylation Effects on Cancer Protein's DNA Binding

Goal:

Identify portion of INI-1 protein containing active DNA binding site to refine protein function and structure relationship

Introduction:

With the rapid nature of disease development, a simple number cannot be placed on the types of cancer known among the human race. Of the hundreds of cancers characterized a minor portion of them are classified as neurological, and of the neurological forms of cancer very few are defined as pediatric. Within this focus of pediatric neural cancers resides atypical rhabdoid tumors (AT/RT). AT/RT roughly constitutes one to two percent of all pediatric central nervous system tumors [1]. Cancer, in general, is a devastating disease, but AT/RT is especially horrendous with its aggressive hindrance on the cerebellum, brain stem, and spinal cord of children predominantly between six months and three years of age. Unfortunately, atypical rhabdoid tumors are extremely difficult to diagnose due to indistinct symptoms being associated with a whole host of other diseases. In a majority of cases, AT/RT is lethal and will result in debilitating, life-threatening symptoms within weeks of diagnosis. Although extraordinarily rare, atypical rhabdoid tumors desperately need further research invested to provide these unfortunate young with a fighting chance.

Little is known about the direct cause of AT/RT development, but there is evident correlation between the tumors and INI-1 protein. All AT/RT cells suffer from a truncation or complete loss of INI-1. The lack of understanding behind this relationship comes from an inability

to determine INI-1's specific function or structure. Although it is understood that INI-1 participates in DNA repair within the SWI/SNF protein complex, its exact role has yet to be determined. It is clear that INI-1 is responsible for binding directly to the DNA, but overall function and specific mechanisms occurring once bound to DNA is yet a mystery. Pinpointing specific functionality portions of the protein, such as DNA binding sites, will enable a clearer understanding of the relationship between INI-1, AT/RT cancer, and potential pathways to a cure.

Experimentation:

My hypothesis is that by mutating a potential DNA phosphorylation site within a known binding region this will link a specific location to a primary function of INI-1. To test this hypothesis, I will mutate the protein for further testing of phosphorylation sites.

Binding Region Analysis for Mutation Site

Before mutagenesis can even begin, a site for mutation must be determined. Now, INI-1 consists of 376 amino acids, the selection for mutation could not be merely random and yield any success. While the three-dimensional structure is still unknown, there are four general domains defined within INI-1 (Figure 1). Based on previous projects conducted within Dr. Ruhl's laboratory, the plasmid was truncated into two fragments. The N-terminus truncation consisted of amino acids 1-186 and the C-terminus consisted of the remaining amino acids in the domain, 186-376. The two fragments were tested on an electrophoretic mobility shift assay compared to full length INI-1.

Figure 1

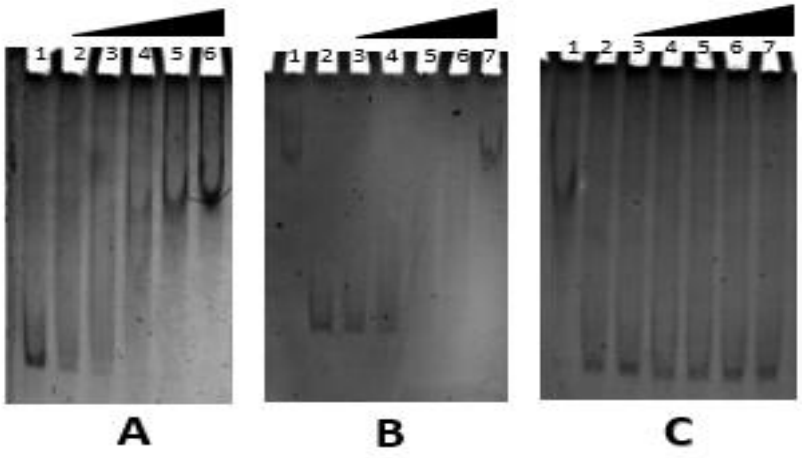

Figure 1: Three EMSA gels that indicate different protein banding patterns across gel from lane 1-6 or 7. Banding pattern for gel A and gel B are very similar where gel C lacks differentiation from lane 1 through lane 7.

The truncations were ran on an electrophoretic mobility shift assay (EMSA), which allows the visualization of protein and nucleic acid interaction. This is possible because EMSA gel shifts separate the protein/nucleic acid complexes by charge and shape, giving qualitative representation toward the complex's ability to move through the gel matrix. If the protein were to bind DNA it would naturally make the complex larger and hinder its movement down the gel. Whereas proteins with inhibited DNA binding would allow for the unbound nucleic acids to flow relatively quickly through the matrix.

The two protein truncations, 1-186 and 187-376, were then combined at varying concentrations with DNA ranging from 2:1 to 1:2 ratios of DNA to protein. These increasing ratios of protein to DNA allowed control of the potential capacity of DNA binding activity per sample. Concentrations are presented in the gels with increasing protein going from 1 to 7 (or left to right). Figure 1A is full length INI-1 with evident binding visible in lanes 4 through 6 as the protein to DNA ratio began to favor greater concentrations of protein. Figure 1B is truncation 1-186 with a

full-length control in lane 1 and similar binding occurring in lanes 5 through 7. Figure 1C is truncation 187-376 presenting no change in banding patterns across the gel, thus representing no DNA binding protein is present. This EMSA gel narrowed the active binding site location by more than 50% with evidence of the C-terminus binding domain exhibiting no visible binding activity where the N-terminus portion bound DNA in similar proportions to the full-length protein. The target site for mutation, the active binding site of INI-1, narrowed to a mere 186 amino acids.

The work of Matsuoka et. al. (2016) took what little is known about proteins involved in DNA damage response and identified hundreds of phosphorylation sites from hundreds of different proteins [2]. This study established a database from identifying phosphorylation sites that were induced in response to DNA damage. After studying trends of 900 proteins and 700 specified sites this DNA damage phosphorylation database developed into a rather useful tool for further research on the damage response proteins exhibit in repairing DNA. The database included information on INI-1 and directed this mutation study to a specific serine within the binding domain previously determined. The point mutation site was selected for serine-129 after reviewing the data from Matsuoka's expansive project. This specific site occurs within the same truncation shown to conserve is binding functionality when truncated from the full-length protein.

Figure 2

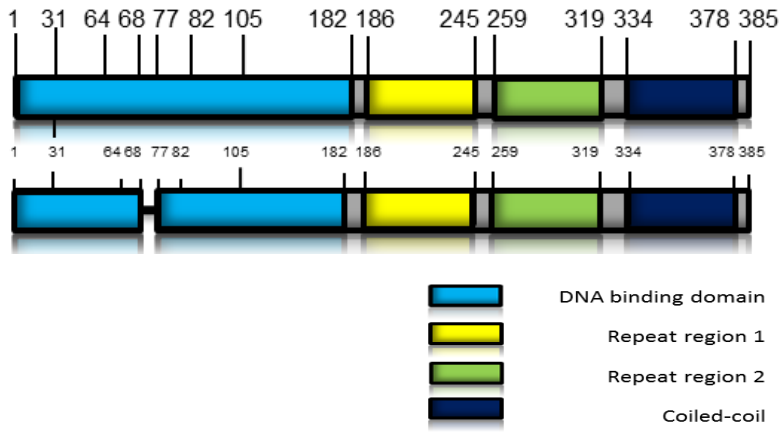

Figure 2: Representation of full length INI-1 and four different domains including amino acid markers for size reference. Serine 129 (S129) identified as site of mutagenesis.

Mutagenesis

The mutagenesis process began with determining ideal two mutations for experimentation: fixed phosphorylated and fixed non-phosphorylated mimics. Where serine is a small amino acid with a neutral charge and phosphorylatable hydroxyl group, an amino acid of similar size and charge is needed to serve as the fixed unphosphorylated mimic (Figure 3). Two commonly used mimics for serine are methionine and alanine. After further examination, alanine was selected as the unphosphorylated mimic because it has been consistently beneficial in previous phosphorylation studies whereas methionine was more commonly used as a mimic in acetylation and other activation method studies [3]. Where a phosphorylated serine is larger and negatively charged due to the new addition of the phosphate group, a charged amino acid is needed to serve as the phosphorylated mimic (Figure 3). The ideal amino acid to use for this mimicry is aspartic acid.

Figure 3

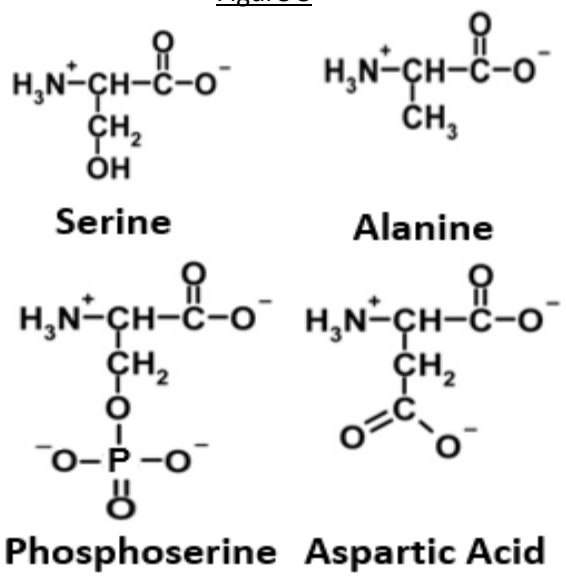

Figure 3: Amino acid structure of serine in natural and phosphorylated states with selected mimic structures next to each.

After mimics were carefully selected, plasmids of full length and 1-186 truncations were prepared. The full-length plasmid was prepared in advance for the project in a TEV 1155 vector. The truncation was prepared using a digest with restriction enzymes Xhol and Xbal. These restriction enzymes were selected for their specificity to cut the truncation needed. After the digestion, the digested sample was ran on an electrophoresis gel with a Kb+ ladder serving to quantify results and uncut plasmid as a control. When the digested plasmid presented two bands measuring around 300 and 5,000 base pairs a successful truncation was confirmed. The band at approximately 300 base pairs was the truncation needed.

Both truncated and full-length INI-1 plasmids will receive mutations on the Ser129 site. Mutants were prepared using a Q5 site-directed mutagenesis kit with primers designed to enable the mutations for fixed phosphorylation mimics in the INI-1 truncation as well as the full-length protein. Four mutations will be created in total: Full-length 129 alanine, full-length 129 aspartic

acid, 1-186 truncation 129 alanine, and 1-186 truncation 129 aspartic acid. Each mutation was prepared one at a time in the following steps from PCR to final DNA sequencing.

The first step in this process is polymerase chain reaction (PCR) amplification. The PCR reaction consists of thermocycling the DNA with primers, DNA template (TEV 1155), and DNA polymerase mixture to amplify the double stranded DNA plasmid. The primers are designed with nucleotide regions surrounding the mutation site to encourage mutation. For the specific primers used, the annealing temperature was at 64 °F. This is necessary to amplify amount of viable DNA for further testing.

Once PCR is complete, the PCR product is cleaned up using a KLD reaction. This reaction contains an enzyme mixture of kinase, ligase, and Dnp1. During incubation for this reaction the plasmid is made circular again and the template is broken down so at the end of the KLD reaction all that should remain is INI-1 plasmid. Now that we have the plasmid isolated and back to its circular form, it is then transformed into *E. coli* bacteria competent cells. This allows the plasmid to grow rapidly with preference toward the mutated plasmids forming colonies. The bacteria colonies selected are transferred to broth followed with a final plasmid prep. Post transformation and plasmid prep the samples were run through an electrophoresis gel to ensure plasmid was in fact in the sample then the samples with the most plasmid present in the gel image were sent for DNA sequencing. Sequencing ultimately revealed a successful or unsuccessful attempt at the point mutations. On occasion, a sample would make it all the way through this process to sequencing and still be the original plasmid with no mutation. In this circumstance, mutagenesis was performed again until the point mutation was successful.

Purification

After transformation into *E. coli* cells, purification of INI-1 is necessary to remove all the bacterial molecules and proteins from the sample. A common purification technique known as column chromatography was used to separate the protein of interest from the rest of the cell debris. Before the sample can be purified the cell membrane must be broken to release the protein. This was achieved through sonication. Sonication uses high frequency soundwaves to lyse the cell and disrupt the membrane, thus freeing the protein needed for purification.

Now that the cell is essentially destroyed the sample is ready for chromatography. The chromatography process produces three main products: a flowthrough, a wash, and several elutions. The flowthrough is unpurified sample ran through the column to allow protein binding to the resin column. This flowthrough also serves as a starting frame for analyzing purification progress later on the SDS-PAGE, sodium dodecyl sulfate polyacrylamide, gel (Figure 4, lane 2). The wash product is the remainder of sample pulled through the column with a wash buffer at a relatively neutral pH. The debris is washed from the resin while the protein remains bound (Figure 4, lane 3 and 4). The elutions are a series of aliquots coming off the column as an elution buffer works its way through the column. The elution creates a significant pH change, creating a conformational change in the substrate. The lowering of pH induces the protein of interest to be released from the column. Thus, the purified protein should appear in one or multiple elutions when quantified on the SDS-PAGE gel. Six elutions were collected to ensure the plasmid was eluted out of the column (Figure 4, lanes 5-10).

Figure 4

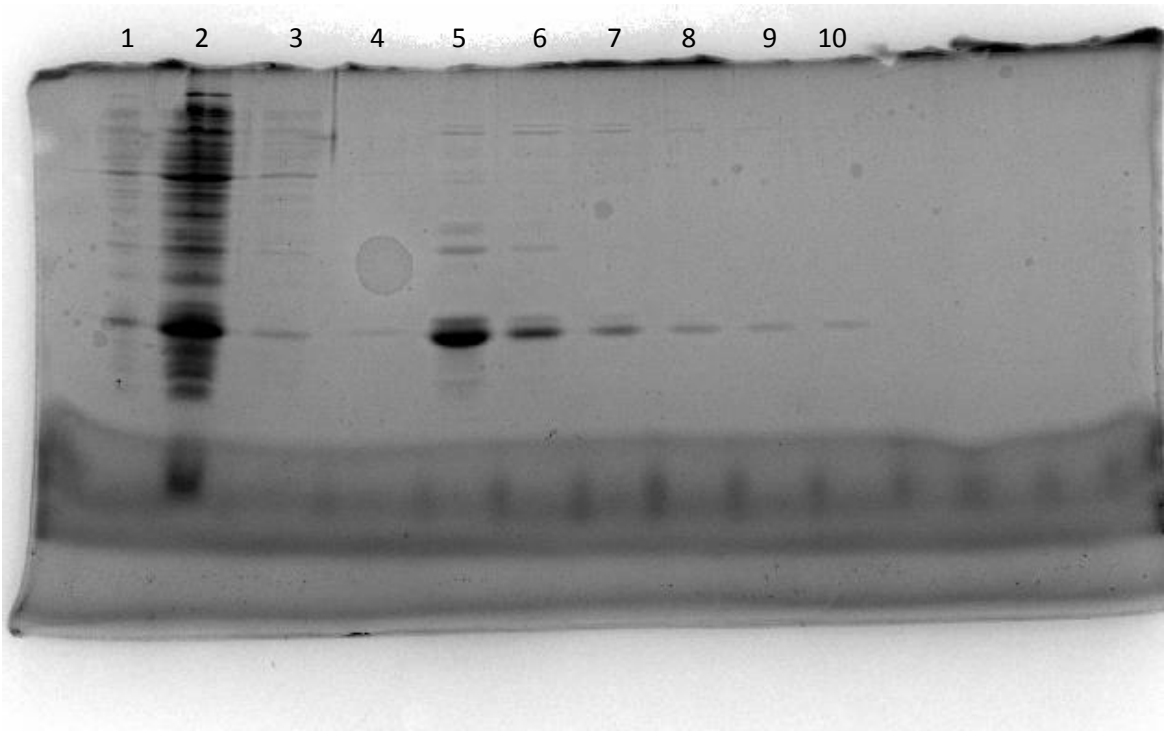

Figure 4: Image of SDS-PAGE gel for alanine mutation in 1-186 truncation.

This SDS-PAGE gel containing the truncation INI-1 with the alanine mutation clearly has purified protein in the elution 2 lane (lane 5) with a pronounced, single band. That band can be quantified using the ladder in lane 1, providing a check standard to ensure the protein of interest, INI-1 1-186, was protein purified in the chromatography performed. The full-length mutations did not cooperate in the purification process as well as the truncations, so other extensive methods were used to get these proteins purified for further testing.

Binding Study Preparation

With all mutations completed and purified the INI-1 plasmids are near ready for binding activity testing. Using the same type of gel from Figure 1, an electrophoretic mobility shift assay, the binding activity of each mutation will be developed into quantifiable data. Thus providing evidence for or against a specified active binding site in the elusive INI-1 protein. A few final steps are still needed before a gel can be ran: concentrations of DNA and protein need to be calculated to determine ratios, and preparing controls for the gel shift assays. Using quantified gels prepared by Dr. Jessica Matts, calculations were done to determine both protein concentration and DNA concentration in current samples for varying ratios in the binding assays to come.

Future Work:

To continue this project electrophoretic mobility shift assays will be prepared. A total of four assays will be performed: one for each alanine and aspartic acid mutations in both full length and truncated plasmids. All mutations will be tested on an EMSA gel shift and imaged to determine their ability to bind DNA with fixed variations in the phosphoryl site. Native INI-1 plasmids will be prepared to serve as a control on the EMSA gels. Analyses will be performed to determine any variations or binding patterns, if patterns appear, as well as correlations and conclusions will be drawn depending on results per mutation and plasmid size. These kind of results could potentially solidify a known location and specific function of INI-1, giving way to further studies to hone in this region and experiment with functionality components particularly nearby serine-129. The culmination of these final EMSA gels in this study could provide vital information toward fully understanding INI-1's precise function piece by piece.

Results and Conclusion:

The results from this experiment were that I successfully mutated a fixed-positive and fixed-negative phosphoryl form of INI-1 in both truncated 1-186 and full-length plasmid. While I ran out of time within my undergraduate studies to see the whole project to completion, I trust another undergraduate student will pick up where I left off and finish this project for the furthering of specific functionality of INI-1. This activated phosphorylation site might seem mundane in the grand scheme of things, but this functional point could act as a spotlight for discovering novel and specific activity surrounding this serine. Discovering new activity could shed light on the troubling mystery that is currently DNA repair of INI-1 and the SWI/SNF protein complex as a whole. It is not just the unknown functionality of these proteins, but of most proteins involved with damaged DNA and the corrective pathways naturally activated in the cell to fix the error. If DNA damage response can be studied in depth with understanding, how soon after could the drug development industry produce artificial damage response molecules? What seemingly incurable cancers and diseases could potentially be treated or cured by a corrective approach rather than the most common regulative approaches used today?

Bibliography:

1. Biswas, A., Lakhan, K., Aanchal, K., Sarkar, C., Julka, P.K. Atypical teratoid/rhabdoid tumors: challenge and search for solutions. Retrieved November 27, 2017.
2. Matsuoka, S., Balliff, B.A., Smogorzeswka, A., et. al. ATM and ATR Substrate Analysis Reveals Extensive Protein Networks Responsive to DNA Damage. Retrieved September 29, 2016.
3. Choi, J.H., Murray, J.W., Wolkoff, A.W. PDZK1 binding and serine phosphorylation regulate subcellular trafficking of organic anion transport protein 1a1. Retrieved September 07, 2015.
4. Mancinia, J.A., O'Neill, G.P., Baylyb, C., Vickersa, P.J. Mutation of serine-516 in methionine or aspirin r human prostaglandin G/H synthase-2 to acetylation of this residue stimulates I J-R-HETE synthesis. Retrieved September 07, 2015.