SPECIFYING THE STRUCTURAL BOUNDARIES OF THE NUCLEAR LOCALIZATION SEQUENCE IN TORQUE TENO VIRUS VP3

A Major Qualifying Project

Submitted to the Faculty of the
WORCESTER POLYTECHNIC INSTITUTE

In Partial Fulfillment of the Requirements for the
Degree of Bachelor Science

By:

Emma Begbie
Biochemistry and Biology & Biotechnology

Bailey Sarber
Biochemistry and Biology & Biotechnology

Date: April 24, 2013
Approved:

Dr. Destin Heilman, Primary Advisor
Department of Chemistry and Biochemistry

Contents

Table of Figures

Abstract

The Torque Teno Virus (TTV) is a non-enveloped virus. It is single stranded and has a circular genome. TTV contains a third virus protein (VP3) which is known to induce apoptosis. TTV VP3 is known to have two regions of interest, the Nuclear Localization Sequence (NLS) and the Nuclear Export Sequence (NES). The NLS in TTV VP3 is composed of two basic regions separated by a non-specific region. Without the NLS, the peptide is punctate and cytoplasmically localized. Truncation of the NLS was achieved by primer directed mutagenesis separating the two basic regions. This revealed, without the second basic region, the peptide tended to self-aggregate in the cytoplasm. Because this truncated peptide behaved similarly to the protein with full NLS removal, both regions are hypothesized to be necessary for the functionality of the structure and the apoptotic ability of the protein.

Acknowledgements

We would like to thank Professor Heilman for endless help and guidance throughout the year. We would also like to thank Vicki Huntress for assisting us in acquiring the confocal images used in this thesis. Finally, we would like to thank everyone in the Heilman Lab for their collaboration and helpfulness throughout the year.

Introduction

Approximately 90% of the human population is infected with the Torque Teno Virus (TTV), and most people will test positive for at least one of the 27 genotypes of this virus[1, 2]. The virus itself was first discovered in 1997 during a study of patients who contracted hepatitis from blood transfusions[3]. A significant number of these patients were infected with TTV, although no correlation between the hepatitis virus and TTV has yet been determined. From this discovery, studies were then developed to attempt to correlate the presence of TTV with various other diseases in patients. TTV was found to be more prevalent in individuals suffering from cancer, lupus, or idiopathic inflammatory myopathies, but despite this fact, no evidence has been provided to associate the virus and pathologies of these or any other diseases to date[4, 5]. As of now, no study has been able to determine the pathology of the TTV.

The TTV is a non-enveloped virus with a single stranded circular DNA genome[5, 6]. This interesting genomic structure earned it the Latin name Torque Teno, meaning "thin necklace," in 1999. TTV is in the *Circoviridae* family with a tentative genus of *anellovirus* due to its similarities with the Porcine Circovirus (PCV) and the Chicken Anemia Virus (CAV)[5, 7]. Most of the *Circoviridae anellovirus* genomes only contain two genes, one that codes for the capsid protein (cap) that surrounds the virus genome, and another for a replication initiator protein (rep). The genome is ambisense, implying that one gene runs in the 5' to 3' direction and the other runs in the 3' to 5' direction allowing the open reading frames of these essential genes to be inversely overlapped[6]. Once a circovirus has entered a cell, it makes a complementary strand via the rep protein to allow for the translation of the cap protein[6]. Although this is not an

overall characteristic of circoviruses, CAV, PCV, and TTV are a few circoviruses that have a third open reading frame that codes for an apoptosis inducing protein. With TTV, only six of the 27 homologs produce a third protein (VP3) of 105 amino acids. The structures found in this protein are similar to those of the VP3 produced in CAV, known as apoptin[8]. This protein is theorized to induce apoptosis as a way of releasing the virus after replication.

TTV, PCV, and CAV VP3s are interesting because they are known to cause apoptosis almost exclusively in cancerous cells[8]. When tested in healthy mammalian cells, these virus proteins do not induce apoptosis. Apoptin localizes in different regions of a cell depending on the nature of that cell[9]. The protein is known to localize in the nucleus of cancer cells more commonly than in non-cancerous cells, where it is often found in the cytoplasm. This localization in the nucleus is critical to the apoptotic functionality of apoptin. While in the nucleus, apoptin binds to the anaphase promoter complex to induce apoptosis[9]. TTV VP3 is similar to apoptin in structure and has been proven to behave similarly in cells by our lab, though it does not show significant nuclear localization in cancerous cells. TTV VP3 is assumed to follow this pathway to induce apoptosis even though it is not continuously in the nucleus. Despite this assumption, the mechanism for inducing apoptosis has not yet been demonstrated for TTV VP3. Discovering the mechanism of TTV could elucidate more about cancer cells and potentially even assist with early detection of precancerous cells.

While the mechanism of the proteins is currently unknown, apoptin is known to cause apoptosis independent of the p53 pathway[10]. This suggests that, due to their similarities, TTV VP3 will also be mechanistically independent of that pathway. p53 is one of 13 tumor

suppressor genes that are known to be mutated in cancer. This can result in brain tumors, breast, colon, liver, lung, and many other types of cancer[11]. As a tumor suppressor, p53 is responsible for signaling cell cycle arrest or cell death in response to damages in the genome or to the replicatory machinery[12]. It is known as the "guardian of the genome" because it prevents damaged cells from proliferating. Mutations in the expression of this tumor suppressor allow for unchecked proliferation when concurrent with an oncogene mutation[12]. Since p53 mutations occur in approximately 50% of cancer cases, making a treatment that is p53 independent is highly desirable[11].

Both apoptin and TTV VP3 contain two essential regions that are thought to play a major role in the proteins' selective apoptotic ability; they are known as the nuclear export sequence (NES) and the nuclear localization sequence (NLS). The NES is a consensus sequence near the N terminus of the peptide that binds to an export protein allowing it to be transported out of the nucleus. The NLS acts as a target structure for the binding of the β-importin protein that shuttles TTV into the nucleus[13]. An arginine rich sequence, similar to the first basic region in TTV VP3's NLS, is ideal for β-importin trafficking into the nucleus. The non-specific segment can tolerate mutations without affecting the binding to this protein, but both basic regions are required as seen by a loss of function after mutation[14]. The NLS allows TTV VP3 to enter the nucleus. Our lab previously determined that the NLS of TTV VP3 causes limited nuclear localization of the protein in cancerous cells, though the NES structure is more powerful resulting in overall cytoplasmic localization[15]. Apoptin localizes in the cytoplasm in healthy cells and in the nucleus in cancer cells, suggesting the location of the protein in the cell is significant to its selective apoptotic tendencies. With TTV VP3, the peptide always localizes in the

cytoplasm regardless of the cell type, making the mechanism and characteristics of the NLS particularly interesting in TTV VP3.

The NLS of both proteins are bipartite sequences of several basic amino acids split up by short non-specific regions. Since these sequences in apoptin and TTV VP3 are similar, it suggests similar mechanisms for the proteins. It is hypothesized that without the full sequence together, there will be even less nuclear localization in the cancer cells than with the wild type protein. Previous studies by our lab have indicated that the complete removal of the NES from TTV VP3 resulted in nuclear localization of the protein in cancerous cells, proving the functionality of the NLS[13]. When paired with the lack of a consensus sequence for the NLS and the bipartite characteristic of the NLS, there is a need to study this sequence further. Our lab also observed that without the NLS structure, TTV VP3 tended to multimerize, or self-aggregate, at a higher rate. This is also a characteristic of apoptin after the removal or knockout of its NLS structure. It was posited that removal of part of the NLS would result in self-aggregation of the protein in cancer cells.

The functionality of the bipartite NLS is believed to be essential to the apoptotic ability of the protein, due to its similarities with apoptin. The lack of a consensus sequence makes the NLS an interesting sequence to study although it is believed that the sequence contains two basic regions separated by a non-specific region. Separating these two basic segments would allow for the study of the functionality of the two basic NLS regions on their own. This study focuses on the putative NLS and analyzes the function of the two basic regions by separating the two segments.

Materials and Methods

TTV VP3 Construct Design

Primers were designed to truncate TTV VP3 at a selected location within the non-specific region of the bipartite NLS, while preserving the NES. The gene was truncated after base pair 234 to create a 78 amino aci3d protein and a 27 amino acid protein. Restriction sites for Bam HI and Eco RI were engineered into the gene sequence via these primers. The primer design was based on the gene sequence found by our lab in 2009. The first construct was designed with two different primers in order to create a protein that contained 78 amino acids.

The first primer was an N-terminal extension of TTV VP3, which was a 53 base pair oligonucleotide sequence of 5' gc tct aat gcg cta gct agc tat agG AAT TCa atg atc aac act acc tta act 3' containing an Eco RI restriction site (indicated by the capital letters) followed by a start codon. The second primer for this sequence was a reverse complement of a short region of TTV VP3 with an artificial stop at nucleotide 234. This was a 53 base pair oligonucleotide sequence of 5' gct aat gta cgc tag cta gct aaG GAT CCt tac acc agc cat agg cca tgg tg 3' containing a Bam HI restriction site (indicated by the capital letters) followed by a stop codon.

The second construct was a 27 amino acid protein created using two different primers. The first primer created an artificial start at nucleotide 234 of TTV VP3, which was a 53 base pair oligonucleotide sequence of 5' gc tct aat gcg cta gct agc tat agG AAT TCa gcg ccg aag gag aag acg gtg 3' containing an Eco RI Restriction site (indicated by the capital letters). The second primer was a reverse compliment of the C-terminal extension of TTV VP3 which was a 53 base pair oligonucleotide sequence of 5' gct aat gta cgc tag cta gct aaG GAT CCc tag cag gtc tgc gtc ttc

gg 3’ containing a Bam HI restriction site (indicated by the capital letters) followed by a stop codon.

With the use of the primers, the following PCR protocol was followed: |95°C, 4min, 1x| [95°C, 30sec] [55°C, 30sec] [72°C, 1min], 30x| 72°C, 4min, 1x| 10°C, ∞|. The PCR products were then run through a 1% agarose gel with TAE buffer solution at 90 mV for an hour. The resulting bands were purified with the Wizard SV Gel and PCR Clean-Up kit (#TB308) to obtain pure samples of the PCR product. The protocol is available from Promega. The samples were restricted using Bam HI and Eco RI restriction enzymes for two to three hours at 37°C. The restriction was confirmed by running the samples through a 1% agarose gel with TAE buffer solution at 90 mV for an hour. The resulting bands were purified with the Wizard SV Gel and PCR Clean-Up kit (#TB308). Each sample was then ligated into a pEGFP vector (Clontech) with kanamycin resistance using T4 DNA ligase and 1X ligase buffer for three hours. The plasmid vector was intended to create a fusion protein of the GFP fluorophore N terminally fused to the construct proteins.

Molecular Cloning

The plasmids were transformed into competent E. coli. Colonies were grown on LB agar plates with 50 mg/mL kanamycin for 18-24 hours in a 37°C incubator. Individual colonies were selected and incubated in liquid LB with 50 mg/ml kanamycin. The flasks remained in a shaking incubator at 37°C overnight. The flasks were then processed to retrieve and purify the DNA using the PureYield Plasmid Midiprep System (#A2492). The DNA was then diluted and quantified through UV spectroscopy and sequenced by Macrogen USA.

Tissue Culture

H1299 non-small cell lung carcinomas were maintained in sterile DMEM with 10% FBS in a 37°C incubator with 5% CO_2. The purified and sequenced plasmids were transfected into H1299 non-small cell lung carcinomas using a standard procedure for the FuGENE® HD Transfection Reagent (#TM328). The cells were then fixed on glass cover slips using 4% paraformaldehyde in 1X PBS and 70% ethanol. The cover slips were stained with DAPI, mounted on glass slides, and sealed in place. The slides were viewed and photographed using a Leica SP5 confocal microscope with an inverted Leica DMI 6000 CS microscope base. This was visualized using Leica Application Suite Advance Fluorescence Lite from Leica Microsystems.

Determination of Cytoplasmic/Nuclear Localization and Punctate Analysis

The determination of the cytoplasmic to nuclear localization was analyzed through a program called Image J. Running average histograms were created using the rectangular tool in Image J. The box was placed across a portion of the cytoplasm and a portion of the nucleus to analyze the pixel density across the two regions using the plot profile function. The measurements were set to display area, standard deviation, integrated density, area fraction, and mean gray value for all experiments. Punctate analysis was conducted through Image J by drawing a line across the cytoplasm of a cell using the straight tool. Each point on this line was analyzed for pixel density to create a histogram using the plot profile function. The localization of each construct was quantified by using the freehand selection tool to trace the nucleus and the cytoplasm separately, and by using the plot profile function. These expression readings were normalized by the GFP expression and formatted into ratios and fractions using Microsoft Excel.

Results

The Torque Teno Virus (TTV) is a member of the *Circoviridae* family along with viruses such as the Chicken Anemia Virus (CAV). In a few homologs of these viruses, a third open reading frame codes for a protein that selectively induces apoptosis in cancerous cells. An interesting characteristic of these viruses is that they contain a nuclear export sequence (NES) and a nuclear localization sequence (NLS). The NES is a consensus sequence that binds to an export protein in the nucleus that moves the virus protein three (VP3) into the cytoplasm. The NLS is believed to play a key role in inducing apoptosis by causing nuclear localization of the VP3. This is shown in apoptin with a loss of apoptotic ability after removal of the NLS function. TTV VP3 is thought to contain an NLS even though there is no consensus sequence that clearly defines the basic bipartite regions. TTV VP3 has a short series of arginine residues believed to make up the first portion of the bipartite sequence. After a short non-specific sequence of amino acids, two lysine residues are believed to make up the second basic region (Figure 1). This uncertainty of the NLS structure and location in TTV VP3 made this an ideal sequence to study. The purpose of these experiments was to create two recombinant protein constructs from TTV VP3 to specify the structural boundaries of the NLS and the protein's localization and aggregation tendencies.

Creation of Constructs

The tendencies of TTV VP3 with an NLS mutation were best studied through truncation for these experiments. Primers were designed with the CLC Workbench sequence analysis software and were used to truncate TTV VP3 into two smaller genes. One construct contained the open reading frame for TTV VP3, coding for the entire protein through half of the non-

specific region of the NLS, ending at amino acid 78, including the first set of basic amino acids. The second construct contained the open reading frame for the second half of the non-specific region including the second set of basic amino acids in the NLS (Figure 2). Primer directed truncation was achieved via PCR to cut the gene at base pair 234 between the two basic regions of the bipartite NLS. When analyzing the electrophoretic gel of the PCR products, the resultant bands matched with the expected sizes for the truncated genes, with the first construct being 234 base pairs and the second construct being 81 base pairs (Figure 3).

After the verification by electrophoresis, the constructs were ligating into an expression plasmid pEGFP that attaches the gene for a fluorescent marker to the construct sequence. The resulting ligated plasmid was transformed into *E. coli* cells (Figure 4). After confirmation of the correct construct sequences by Macrogen USA, the restricted samples were electrophoresed to further confirm the product (Figure 5). The first restricted construct appeared on the gel at approximately 200-250 base pairs indicated the presence of the 234 base pair gene. A band appeared at approximately 100 base pairs when compared to the DNA ladder, suggesting that the second 81 base pair construct was achieved. For both constructs, the vector was visible near the top of the gel between the 3500 and 4000 base pair markers.

Determination of Localization

The sequenced plasmids were consistent with the expected product of these mutations, allowing the localization of the construct proteins to be analyzed. The plasmids were transfected into H1299 non-small cell lung carcinomas, which are p53 deficient and are immortalized cell lines. These were viewed under a confocal microscope to determine the

localization tendencies of both constructs. It was observed that, with the GFP control slide, there was equal distribution between the nucleus and the cytoplasm of the cell, a typical dispersion pattern of GFP (Figure 6A). The TTV VP3 wild type was visualized in the cytoplasm with minimal nuclear localization, which is also typical of the wild type protein (Figure 6B). The first mutated construct consisting of amino acids 1 to 78 appeared to have localized mainly in the cytoplasm with an insignificant nuclear signal (Figure 6C). This seemed to indicate that the second half of the NLS was needed for the limited nuclear expression of the protein. The second construct, consisting of amino acids 79 to 105, displayed no specific localization tendencies and was distributed equally throughout the cell. This suggested that the second basic region was not a functional NLS on its own (Figure 6D).

Although localization tendencies could be visually estimated, a running average analysis was needed to show the localization tendencies of each construct across part of the cytoplasm and nucleus (Figure 7). These histograms were overall consistent with the visual data suggested by the confocal images. This signal analysis suggests that the TTV VP3 wild type had more cytoplasmic signal than TTV VP3 amino acid 1-78. The data from both GFP and the second construct TTV VP3 amino acid 79-105 showed that the proteins were distributed more throughout the cell instead of localized in just one area, though they both showed slight favoritism to the nucleus. This also indicated that the second basic region of the NLS might not be functional on its own.

A quantitative analysis that is more sensitive than the human eye was gathered using Image J to further confirm the observed trends. The results of the confocal microscopy were

quantified to establish cytoplasmic to nuclear ratios for the expression of the constructs to obtain a more precise analysis. Image J measured the pixel density of the GFP fluorophore to determine the amount of protein that localized in each area of the cell with a higher sensitivity than the human eye. This was done for three cells from each construct, averaging the ratios together to obtain a representative ratio. GFP was assumed to have no specific localization so the ratio for GFP was normalized to 1:1 for cytoplasmic to nuclear localization. TTV VP3 wild type showed more cytoplasmic localization with a ratio of 31.31:1 ± 7.15 supporting the visual observation. The first construct from amino acid 1 to 78 had a ratio of 10.12:1 ± 0.85. The second construct from amino acid 79 to 105 had a ratio of 0.70:1 ± 0.09. In addition, the ratios were graphed in order to visualize the differences between the fractions for each protein (Figure 8). The ratios indicated that TTV VP3 amino acid 1-78 had a smaller cytoplasmic to nuclear fraction than the wild type VP3. It also showed that TTV VP3 amino acid 79-105 slightly favored the nucleus.

In addition to the localization studies, the multimerization of the constructs was determined by analyzing the aggregation tendencies of the protein in the representative cells for each construct. Using image J to produce histograms, GFP, TTV VP3 wild type, and TTV VP3 amino acid 79-105 displayed low aggregation tendencies as seen by the brief sharp peaks (Figure 9). A previous study in our lab determined that TTV VP3, with only the NES structure, had a high aggregation tendency, which was similar to that of TTV VP3 amino acid 1-78. Both constructs showed wide regions displaying high signal, suggesting that there were isolated areas of multimerized protein concentrated throughout the cytoplasm.

Discussion

The Torque Teno Virus (TTV) is a circovirus containing a third open reading frame that produces a protein similar to apoptin. It is postulated that the nuclear localization sequence (NLS) of TTV VP3 contributes to the apoptotic ability of the protein in cancerous cells, due to the protein's similarities with apoptin. In TTV VP3, the NLS is a bipartite sequence composed of two basic regions separated by a non-specific region. The purpose of this experiment was to determine whether both basic regions of the NLS were necessary for the protein to localize in the nucleus and to prevent self-aggregation.

The confocal images acquired, as well as the graphs produced by Image J, suggested that TTV VP3 amino acid 1-78 had diminished nuclear localization when compared to TTV VP3 wild type, most likely due to the loss of the second basic region. Loss of function of the NLS in apoptin is known to result in cytoplasmic localization in cancer cells. From this, it can be speculated that the second basic region of the NLS of TTV VP3 is essential for the functionality of the sequence. Apoptin is also known to lose the ability to induce apoptosis without a functional NLS. The NLS region of apoptin interacts and binds to the anaphase promoter complex to induce apoptosis, indicating that nuclear localization is essential for apoptosis to occur. For TTV, the mechanism has not yet been determined though it is speculated to follow the same pathway. If this is the case, the lack of nuclear localization seen in this construct would also indicate a loss of apoptotic ability. A pixel analysis comparing the cytoplasmic to nuclear localization suggested that this construct had a smaller ratio than the TTV VP3 wild type. The high level of expression may have forced some protein into the nucleus due to diffusion. The Image J software works by measuring pixel density in the confocal images. The

confocal images are a single horizontal slice of the cells that transects the nucleus of the cell as well as the surrounding cytoplasm. The TTV VP3 wild type fusion protein was dispersed evenly throughout the cytoplasm, allowing an accurate representative slice to be imaged. The expression of the first construct protein was high and concentrated in multiple dispersed groupings throughout the cytoplasm. If the clusters of signal were above or below the image taken, the photographed slice would not be representative of the true cytoplasmic expression. This may have prevented Image J from correctly determining the amount of protein in the cytoplasm, therefore skewing the ratio. An analysis of the entire cytoplasmic signal of the cell compared to the entire nuclear signal would provide a more definitive and accurate ratio. Image J is currently unable to process three-dimensional images preventing this complete analysis.

The cytoplasmic signal was surprisingly punctate compared to the TTV VP3 wild type protein as seen by the concentrated clusters of expression throughout the cytoplasm. The first construct was expected to display self-aggregation tendencies to a certain degree, due to its similarities with apoptin, but not in significant quantities. Apoptin is known to multimerize in its wild type version and even more so with the loss of its NLS. The multimerization region is predicted to contain the NES in apoptin though the NLS limits this interaction. Concurrent with self-aggregation, the protein without the NLS displays a distinct punctate-ness to the effect of large clusters of apoptin grouping throughout the cytoplasm. This process is hindered by the presence of the NLS in its functional form. Our lab previously determined that, in the absence of the NLS, the peptide had a significant self-aggregation tendency similar to apoptin. The

moderate punctate-ness of TTV VP3 amino acid 1-78 suggests that the second basic region limits self-aggregation in the full-length protein.

The second construct consisted of amino acids 79 to 105. The confocal images indicated a high level of expression, with a consistent dispersion of the protein throughout the cell. This was expected because the NLS in TTV VP3 is weak compared to apoptin's NLS, which is a stronger sequence that induces nuclear localization. The second half of the TTV VP3 NLS, present in this construct, contains only two basic amino acids compared to the second half of apoptin's NLS, which contains many more basic amino acids. The analysis of pixel density indicated that the protein actually favored the nucleus, although only slightly. The protein may have also entered the nucleus to satisfy equilibrium and due to the lack of an NES, the protein was not shuttled from the nucleus into the cytoplasm. This localization may also be due to the fluorophore attached to the protein. GFP alone displayed a slight localization in the nucleus that may have caused TTV VP3 amino acid 79-105 to favor nuclear localization. The data for the constructs were normalized to this GFP control ratio of 1:1, which suggests that the localization is caused by the protein and not the fluorophore.

No multimerization was expected for the second construct, TTV VP3 amino acid 79-105, due to the structural similarities with apoptin. In apoptin, the multimerization sequence includes the NES structure. This construct should not have this multimerization sequence as it is assumed to be in the same region of TTV VP3 as in apoptin. The punctate analysis of this construct did not suggest any self-aggregation in either the nucleus or the cytoplasm. The

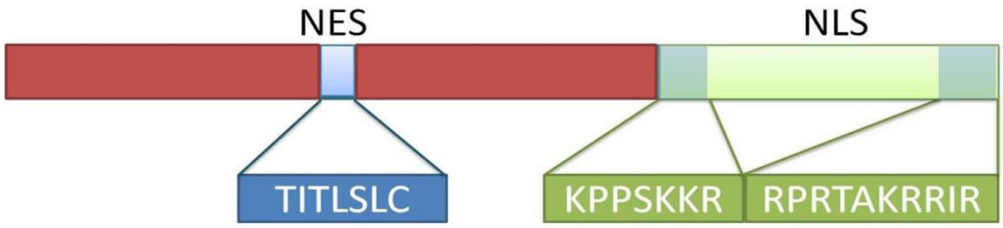
protein contained only a few basic amino acids of the NLS and was too small to participate in any multimerization interactions.

Future Investigations

For future studies it could be useful to perform an apoptosis assay with the created constructs from this study. As of now it is believed that the localization in the nucleus is what signals apoptosis in CAV VP3, though it has not yet been proven necessary in TTV. Learning more about the NLS in TTV VP3 could lead to a better understanding of the apoptotic capabilities of the protein. Although this study suggests that both halves of the NLS are needed for the sequence to function, it would be interesting to see if an apoptosis study showed otherwise. Another study could be to shorten the NLS on either side of the sequence to isolate exactly where the NLS of TTV VP3 begins and ends, since this sequence has no consensus as of now. The first construct of TTV VP3 amino acid 1-78 could be further mutated to remove the NES and determine more about both basic regions of the NLS on their own.

Figures

Apoptin Protein Structure



TTV VP3 Protein Structure

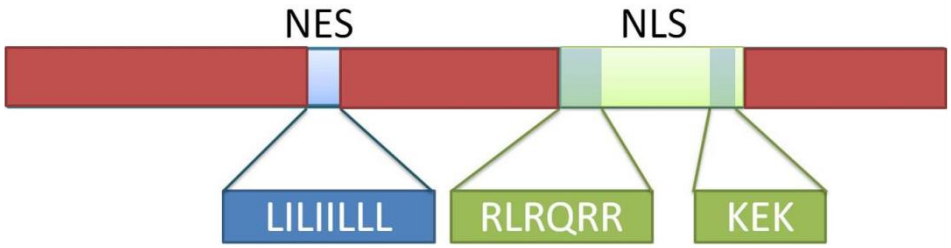


Figure 1: Location of the NES and NLS structures in apoptin and TTV VP3. The blue boxes indicate the amino acid sequence of the NES. The green boxes indicate the amino acid sequences of the basic regions of the NLS.

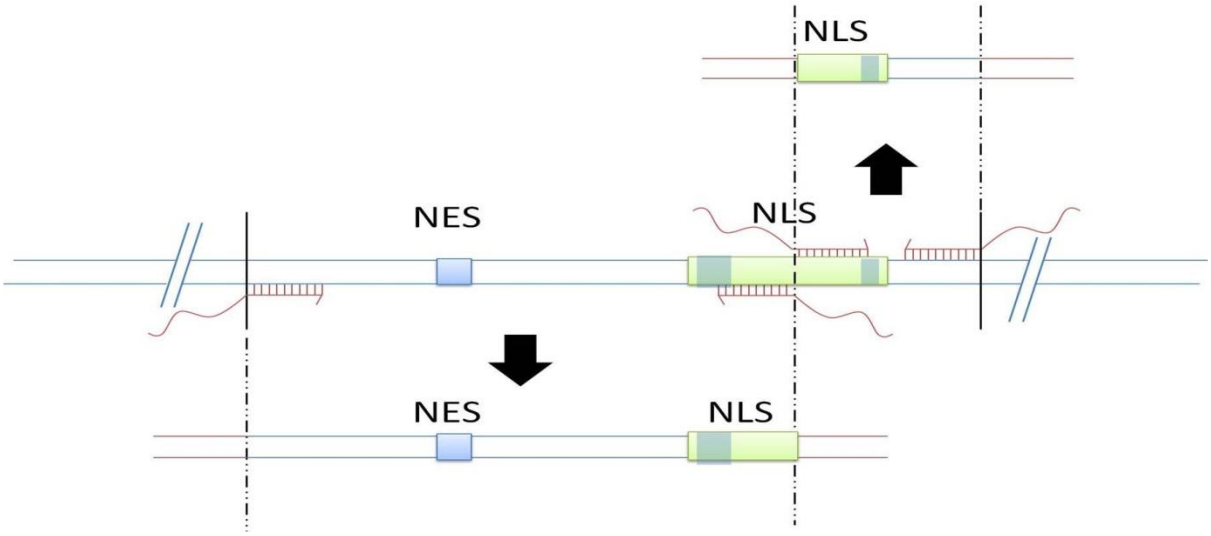


Figure 2: Primer directed mutation of TTV VP3 to design two novel constructs. TTV VP3 wild type (center) was truncated into two DNA constructs. The coding region for TTV VP3 amino acid 1-78 contained the NES structure and the first basic region of the NLS ending amino acid 78 (lower). The coding region for TTV VP3 amino acid 79-105 contained the coding region for the latter half of the NLS, including the second basic region, along with the C-terminus of the protein (upper).

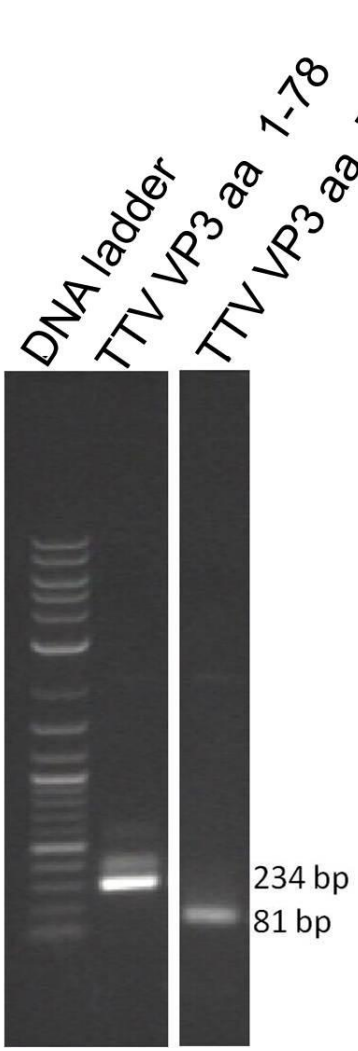


Figure 3: Gel electrophoresis image of the two mutated PCR inserts. Gel electrophoresis confirms that the result of the PCRs were the desired mutants of TTV VP3. Lane 1 is a DNA ladder. Lane 2 is TTV VP3 amino acid 1-78. Lane 3 is TTV VP3 amino acid 79-105. Both constructs are the correct length in base pairs confirming that the truncation was successful. (Note: The third lane was from the same gel but not originally next to the second lane).

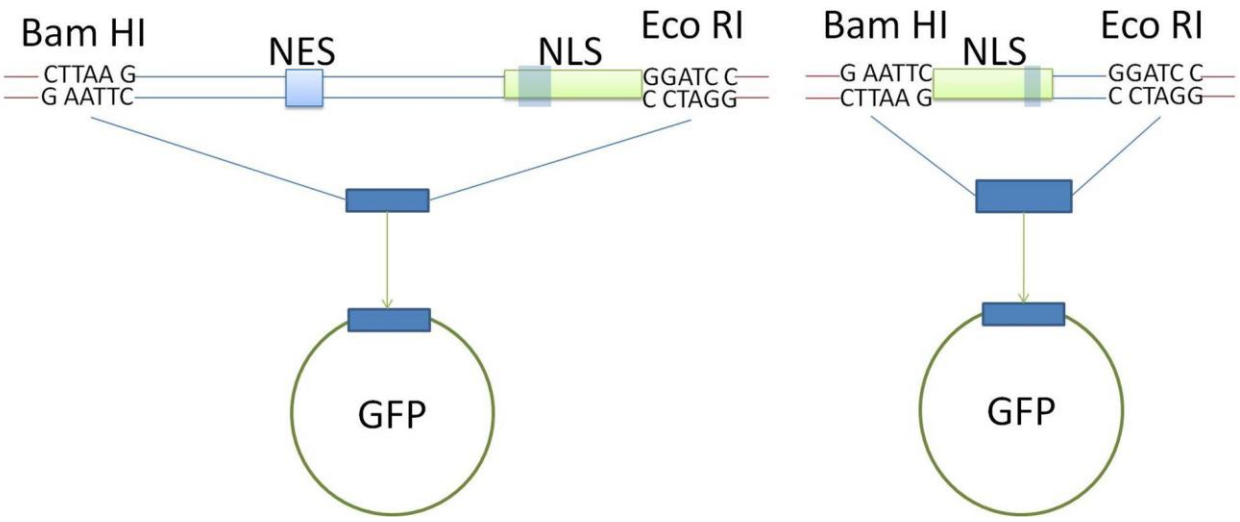


Figure 4: Restriction and insertion of the mutated constructs into a green fluorescent protein (GFP) vector. The gene is shown divided into the two mutated constructs created through PCR. The primers were designed to contain sequences specific to the enzymes Bam HI and Eco RI so the constructs could be successfully cut and ligated into the GFP vector.

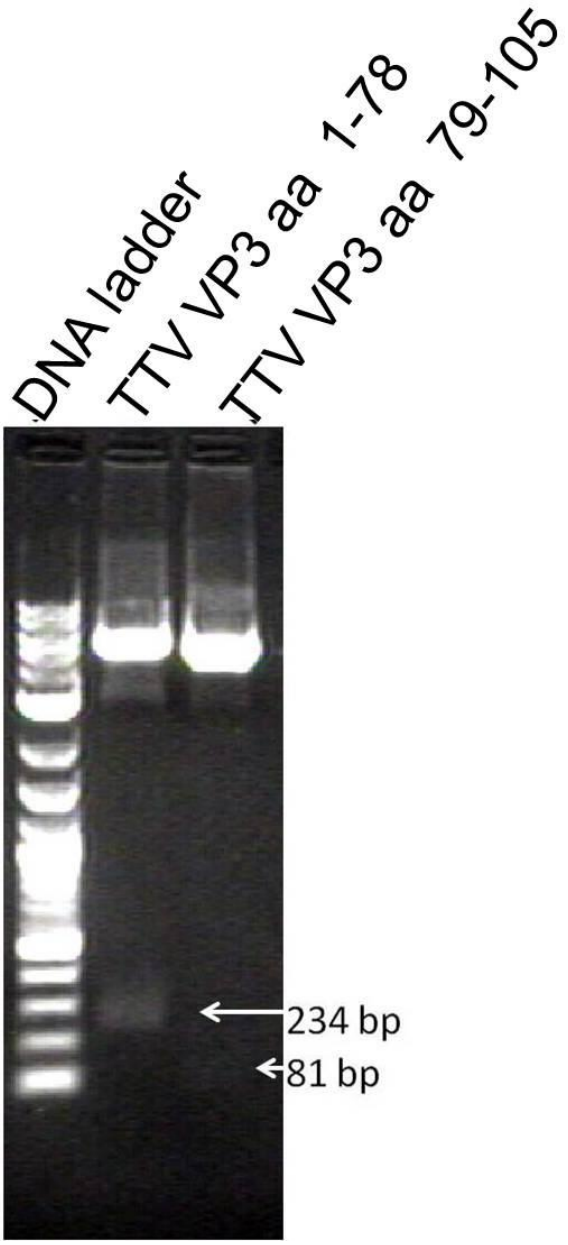


Figure 5: Gel electrophoresis image confirming the positive transformation of two mutated constructs. Gel electrophoresis confirms that the two mutated constructs were successfully created and inserted into a GFP vector. This is shown by enzymatic restriction to remove the insert from the vector. Lane 1 is a DNA ladder. Lane 2 is TTV VP3 amino acid 1-78. Lane 3 is TTV VP3 amino acid 79-105.

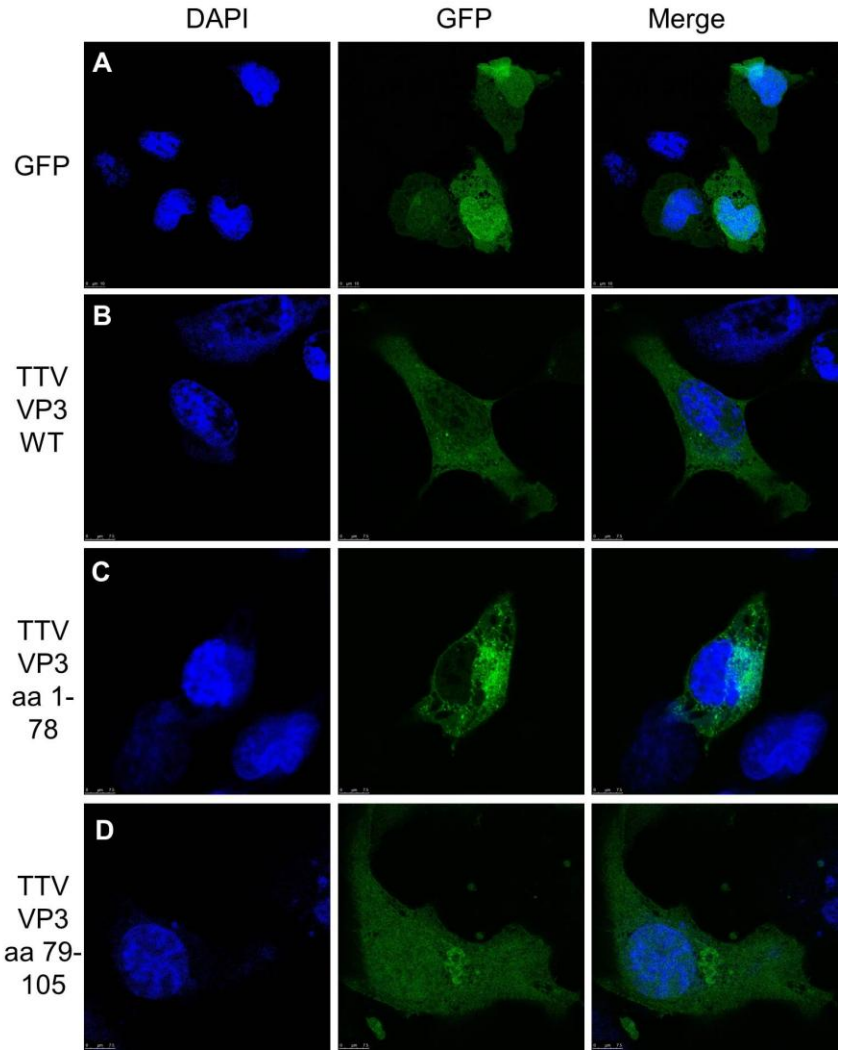


Figure 6: Confocal microscopy images to demonstrate localization tendencies in H1299 cells. The first column is a DAPI staining of the nucleus. The second column is a fluorescence microscopy image of the GFP-construct fusion protein. The third column is the merger of the DAPI image and the fluorescence image. A) A GFP control had no specific localization and showed approximately equal distribution between the nucleus and the cytoplasm. B) The TTV VP3 wild type protein was distributed throughout the cytoplasm with low levels of expression in the nucleus. C) The first construct, TTV VP3 amino acid 1-78, localized mainly in the cytoplasm with a tendency to cluster in the crease of the nucleus. D) TTV VP3 amino acid 79-105, showed no localization tendencies with consistent distribution throughout both the nucleus and the cytoplasm.

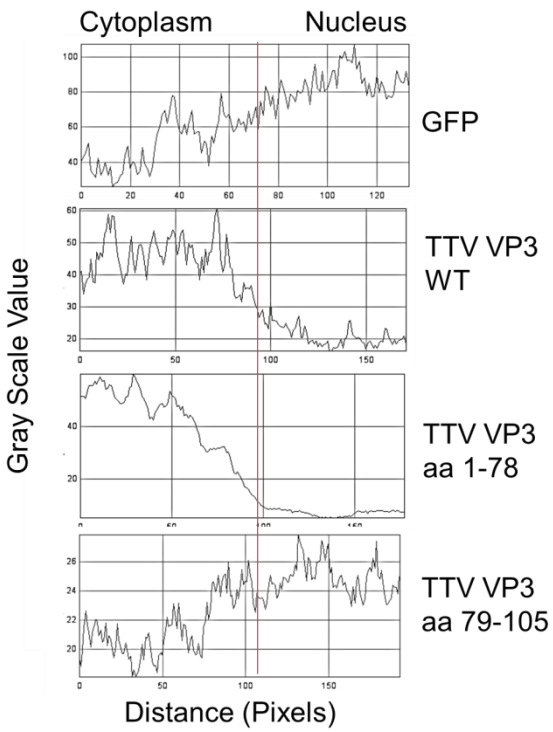


Figure 7: A graphical representation of the localization tendencies of the different constructs. A graphical representation of the amount of protein localizing in the different parts of the cell for each construct are compared. Both TTV VP3 wild type and TTV VP3 amino acid 1-78 localized cytoplasmically, while GFP and TTV VP3 amino acid 79-105 were more equally distributed throughout the cell.

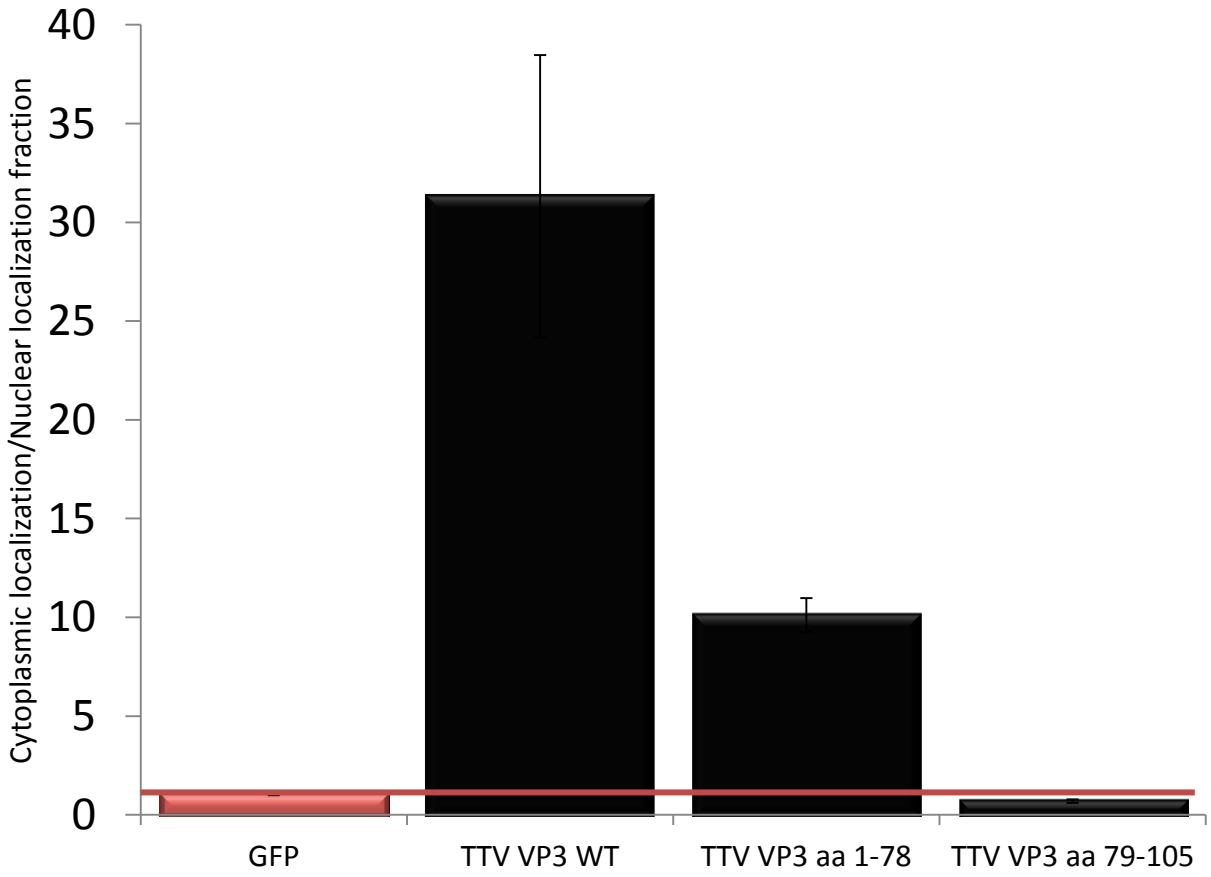


Figure 8: A quantification of cellular localization of the different constructs. To determine the amount of protein localization in the cytoplasm and the nucleus, the microscopy images were quantified using Image J software. GFP was used as a control to set a standard ratio of protein in the cytoplasm to protein in the nucleus. All ratios were normalized using this control. TTV VP3 wild type localizes mostly in the cytoplasm as seen by the large fraction ratio. The TTV VP3 amino acid 1-78 had a smaller cytoplasmic to nuclear localization fraction than the wild type. The TTV VP3 amino acid 78-105 construct showed no specific localization.

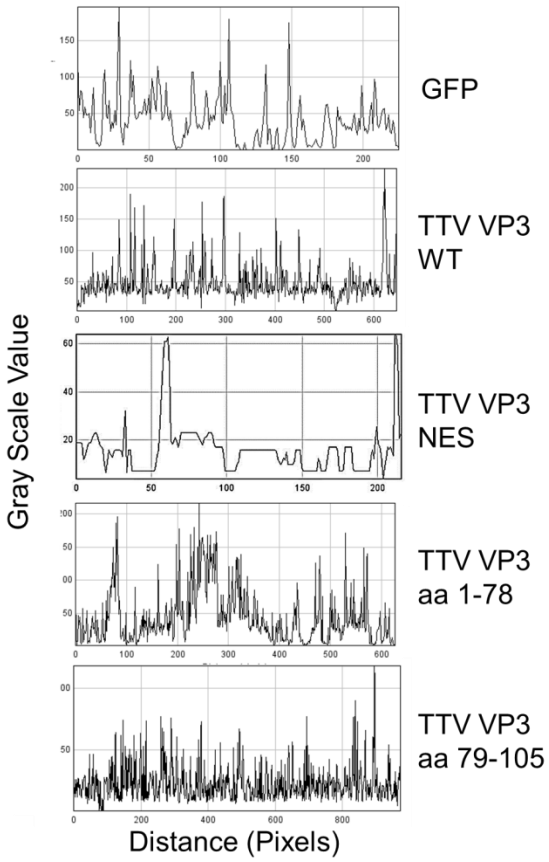


Figure 9: A graphical representation of the punctate-ness of the different constructs. The aggregation tendencies of the constructs were quantified in the cytoplasm of each cell. The GFP, TTV VP3 wild type, and TTV VP3 amino acid 79-105 had low aggregation tendencies while TTV VP3 NES and TTV VP3 amino acid 1-78 showed high aggregation tendencies.

References

1) Hino, S. TTV, a new human virus with single stranded circular DNA genome. Rev. Med. Virol. 2002; 12: 151–158.

2) Handa, A., Dickstein, B., Young, N.S., Brown, K.E. Prevalence of the newly described human circovirus, TTV, in United States blood donors. Transfusion 2000; 40, 2: 245-251.

3) Tsutomu, N., Hiroaki, O., Keiko, K., Hiroshi, Y., Yuzo, M., Makoto, M. A Novel DNA Virus (TTV) Associated with Elevated Transaminase Levels in Posttransfusion Hepatitis of Unknown Etiology. Biochemical and Biophysical Research Communications 1997; 241, 1: 92-97.

4) Zhong , S., Yeo, W., Tang, M. W., Lin, X. R., Mo, F., Ho, W. M., Hui, P., Johnson, P. J. Gross Elevation of TT Virus Genome Load in the Peripheral Blood Mononuclear Cells of Cancer Patients. Annals of the New York Academy of Sciences 2001; 945: 84–92.

5) Hino, S., Miyata, H. Torque teno virus (TTV): current status. Rev. Med. Virol. 2007; 17: 45–57.

6) Delwart, E., Li, L. Rapidly expanding genetic diversity and host range of the Circoviridae viral family and other Rep encoding small circular ssDNA genomes. Virus Res 2012; 164: 114–121.

7) Brassard, J., Gagné, M. J., Lamoureux, L., Inglis, G. D., Leblanc, D. , Houde A. Molecular detection of bovine and porcine Torque teno virus in plasma and feces. Veterinary Microbiology 2008; 126, 1: 271-276.

8) Noteborn, M. H., Todd, D., Verschueren, C. A., de Gauw, H. W., Curran, W. L., Veldkamp, S., Douglas, A. J., McNulty, M. S., van der EB, A. J.,Koch, G. A single chicken anemia virus protein induces apoptosis. Journal of Virology 1994; 68, 1: 346.

9) Heilman, D. W., Green, M. R., Teodoro, J. G. The Anaphase Promoting Complex: A Critical Target for Viral Proteins and Anti-Cancer Drugs. Cell Cycle 2005; 4, 4: 560-e82.

10) Hollstein, M., Sidransky,D., Vogelstein,B., Harris, C. C. p53 Mutations in Human Cancers. Science 1991; 253, 5015: 49-53.

11) Cooper, G. M. The Cell: A Molecular Approach. 2nd edition. Tumor Suppressor Genes. Sunderland (MA): Sinauer Associates; 2000.

12) Oren, M. Regulation of the p53 tumor suppressor protein. J. Biol. Chem.1999; 274: 36031–36034.

13) Truant, R., Cullen, B. R. The Arginine-Rich Domains Present in Human Immunodeficiency Virus Type 1 Tat and Rev Function as Direct Importin β-Dependent Nuclear Localization Signals. American Society for Microbiology: Molecular and Cellular Biology 1999; 19, 2: 1210-1217.

14) Robbins, J., Dilworth, S. M., Laskey, R. A., Dingwall, C. 2 Interdependent Basic Domains in Nucleoplasmin Nuclear Targeting Sequence – Identification of a Class of Bipartite Nuclear Targeting Sequence. Cell 1991; 64: 615-623.

15) DeCicco, M. A., Schiemer, J. S. Exploring Multimerization and the Cancer-Killing Potential of Human Torque Teno Virus VP3O: An Apoptin Analog, A Major Qualifying Project, Worcester Polytechnic Institute 2010.