Assessing the Functionality of Localization Sequences Isolated from PCV1 VP3

A Major Qualifying Project

submitted to the faculty of

Worcester Polytechnic Institute

in partial fulfillment of the requirements for

Degree of Bachelor of Science

by

__________________________
Jameson Kokolis

__________________________
Lauren Spada

April 29, 2010

Approved:

__________________________
Dr. Destin Heilman, Advisor
Abstract

Studies have elucidated the cancer-cell killing abilities of proteins expressed by the third open reading frame (VP3) of members of the *Circoviridae* virus family. Porcine Circovirus 1 (PCV1) was discovered to possess a VP3 homologous to those of other circoviruses, suggesting that it too may have apoptotic activity. Previous studies have shown that cellular localization may be responsible for the killing action of the other VP3 proteins. In order to better study the activity of PCV1 VP3, truncation mutants have been developed, isolating localization sequences. Furthermore, the apoptotic abilities of wild type vector constructs were assessed for the first time. The assay illustrates that apoptosis occurs by a caspases-3-7 pathway; leading us to question the impact of localization, and what structural differences of PCV1 may be responsible for functionality.
Acknowledgements

We would like to thank our advisor, Dr. Destin Heilman, for allowing us to conduct research in his laboratory (WPI Department of Chemistry and Biochemistry, Worcester, MA). His knowledge and guidance were invaluable. We would also like to thank Peter Conerly and Molly McShea, whose contributions to the study of PCV1 VP3 provided the basis of our research.
Table of Contents

Abstract ......................................................................................................................................................... 2
Acknowledgements ........................................................................................................................................ 3
1 Background ........................................................................................................................................... 5
   1.1 The Porcine Circoviruses: Taxonomy ............................................................................................ 5
   1.2 Genomic Structure ........................................................................................................................ 5
   1.3 Protein Function: Rep and Rep’ .................................................................................................... 7
     1.3.1 Protein Function: Cap ........................................................................................................... 8
   1.4 Virulence ....................................................................................................................................... 9
   1.5 Virus – Induced Apoptosis ............................................................................................................ 9
   1.6.1 Apoptosis in Circoviruses ............................................................................................................... 11
   1.6.2 The CAV VP3 Mechanism of Apoptosis .................................................................................. 12
   1.7 Apoptosis Induced by PCV2 VP3 ................................................................................................. 13
   1.8 Nucleocytoplasmic shuttling ....................................................................................................... 13
   1.9 The NLS and NES of Apoptin ....................................................................................................... 15
2 Materials and Methods ....................................................................................................................... 16
   2.1 PCV 1 ORF3 Sequence Analysis ................................................................................................... 16
   2.2 Isolation of PCV1 ORF3 Mutants ................................................................................................. 16
   2.3 Maintaining and Transfecting the H1299 cell line ............................................................................ 17
   2.4 Apoptosis Assay .......................................................................................................................... 18
3 Results ................................................................................................................................................. 18
   3.1 Circovirus VP3 Sequence Alignments Illustrate Homology of Functional Sequences and the PCV1
     Elongated C-Terminus ......................................................................................................................... 18
   3.2 PCR Primers Truncated PCV1 VP3 into Mutants Capable of Demonstrating Localization
     Functionality ............................................................................................................................................. 20
   3.3 Molecular Cloning to an Intermediate Vector .............................................................................. 21
   3.4 Transfection of H1299 Cells with Wild Type PCV1 in p3X-FLAG Vector Suggests PCV1 Induces
     apoptosis ............................................................................................................................................... 22
4 Discussion ............................................................................................................................................ 23
5 Figures ................................................................................................................................................. 27
6 References .......................................................................................................................................... 31
1 Background

1.1 The Porcine Circoviruses: Taxonomy

The Porcine Circoviruses (PCV's) belong to family of viruses known as Circoviridae, and are further classified as belonging to the Circovirus genus (8). Members of this viral family, including PCV Type 1, PCV Type 2, the chicken anemia virus (CAV) and psittacine beak and feather disease virus (BFDV) are typically 17-22nm in diameter, icosahedral in shape and lacking envelopes (8). The virions within genus Circovirus contain DNA of a circular shape that are covalently closed, always single-stranded and between 1.7 and 2.3 kilobases in size (8).

PCV1, discovered in 1974 as a contaminant of the porcine kidney cell line PK/15, has only recently been studied, following the discovery of virulence in pigs associated with its homolog, PCV2 (1,39). Since then, research has provided insight into many of the intricacies of the virus, including genomic structure, virulence, strategies of replication and potential apoptotic activity. This study will focus mainly on the exploration of the potential apoptotic activity in PCV1. In 2004, Heilman et. al. elucidated potential mediation of the CAV protein, Apoptin, to induce cell death by apoptosis in transformed cells. This discovery sparked the interest and ambition of researchers to further explore the family of circoviruses, including PCV1, in search of homologous activity, in the hopes that that unique activity might be ultimately harnessed and clinically evoked as a potential treatment for human cancer. The intricacies of apoptosis, and the structure and homology of the viruses are explored in a more detail below.

1.2 Genomic Structure

The PCV1 genome contains 1759 nucleotides which are believed to be organized into seven separate open reading frames (ORFs) (32). Of seven potential reading frames however, only two are known to encode proteins actually produced by PCV1; they are known simply as ORF1 and ORF2 (29).
The two reading frames, aligned head to head and transcribed divergently, are located on either side of the origin of replication. The layout of the genome can be seen in Figure 1.

Figure 1: Genomic Layout of PCV1. ORF1 and ORF2 are located on either side of the origin of replication and are transcribed divergently. ORF3 of PCV1 is 621 nucleotides in length and overlaps ORF1, although it is transcribed in the opposite direction.

ORF1, the larger of the reading frames, provides the genetic coding for production of the replication initiator proteins Rep and Rep’ and is transcribed in the clockwise direction. ORF2, smaller and oriented for counterclockwise transcription, encodes the structural protein Cap (29). The promoter regions for both coding sequences are strategically placed in the genome: the Cap promoter lies within the Rep coding region and the Rep promoter lies within the origin of replication (29). A third reading frame has also been identified. In 2009, Worcester Polytechnic Institute students Molly McShea and Peter Connerly, with the help of classmate Doug Tischer, were the first to identify and clone the hypothetical reading frame into the green fluorescent protein vector, and Sigma’s 3xFLAG vector. Referred to as ORF3 (or VP3), this region of the genome, as well as the previously developed vectors mentioned above,
will be primary focus of research. It is hypothesized that based on virus homology of PCV1 VP3 and CAV VP3, that this region may in fact also encode for the mechanism of apoptosis (4).

The 111 nucleotide origin of replication region too has been the target of relatively recent research ventures, and has been used as evidence to imply a rolling circle replication mechanism for PCV1 (31). Researchers identified a hairpin, or stem loop, structure in PCV1 that has several highly conserved sequences, including a nonanucleotide, 5'-TAGTATTAC-3' closely resembling the 5'-TAATATTAC-3' loop peak sequence found in DNA viruses belonging to a family of rolling-circle replicators called Geminiviridae (30, 31). This conservation, along with the putative replication proteins mentioned above, likely suggests extremely similar if not identical strategies of replication. Crucial to all plasmids replicating by rolling circle replication is the action of the Rep proteins, Rep and Rep’. The proteins bind and nick the origin of replication. The mechanisms by which these proteins initiate replication are explored in detail in the following section.

1.3 Protein Function: Rep and Rep’

As previously mentioned the rep gene for PCV1 is located in ORF1 and encodes two proteins, Rep and Rep’. The Rep protein, 312 amino acids in size, is considered ‘full-length,’ while Rep’ at 168 amino acids is its truncated isoform (10, 30, 36). The conservation of three amino-acid motifs generally conserved by all rolling-circle replication viruses within the Rep proteins, along with a highly conserved loop peak binding site, strongly suggests such a mechanism of replication for the porcine circoviruses (36). Unlike other rolling replication viruses however, PCV1 is unable to initiate replication with just the Rep protein, and instead requires the joint action of both Rep and Rep’ (30). Mankertz and Hillenbrand showed by real-time PCR and sequence analysis that both replication proteins are in fact simultaneously transcribed, and hypothesized by extrapolation and comparison to similar RC replication viruses that it is strategic splicing of messenger RNA that eventually determines the production of Rep or Rep’ (30). Point mutations or truncations affecting any of four conserved sequence motifs of the rep gene in are
sufficient to knock out the function of the entire protein complex before any replication can commence (30).

When both Rep and Rep’ have been properly translated however, the mechanism of initiation of replication is generally understood. The stem-loop structure at the origin of replication serves as the binding site for both proteins and contains the conserved 5’-TAGTATTAC-3’ apex sequence, as well as the hexamer 5’-CGGCAG-3’, in quadruple repeat, referred to individually as H1, H2, H3 and H4 (36,37). The proteins show no affinity for binding motifs on resting state single stranded genome of PCV1, and therefore, researchers have concluded, bind to a double stranded intermediate formed following viral infection (36, 37). After adhesion of the proteins to the minimal binding site (MBS) composed of the 3’ end of the conserved sequences and two ‘H-’ groups, the hypothesized mechanism purports a Rep-induced nick between nucleotides seven and eight of the conserved nonamer, releasing a free 3’ hydroxyl group, allowing replication to ensue (21,36). Lastly, the Rep protein (actuator of all appreciable action in this mechanism) is proposed to re-join cleaved products at the completion of the rolling-circle replication (21).

1.3.1 Protein Function: Cap

The smaller of two known reading frames, ORF2 holds the genetic material responsible for the expression of the 234 amino acid Cap protein, the capsid forming protein and the sole structural component of a porcine circovirus (37). Several studies have illustrated localization of the Cap protein in the nucleus and the presence of an arginine rich N-terminus in both strains of PCV. Also, while these studies have illustrated divergence between strain 1 and 2 cap proteins by varied protein-protein interaction, they have also shown that divergence to be insufficient in determining the virulence of the individual strain (9, 11).
1.4 Virulence

All members of the *Circoviridae* family are known to cause diseases affecting the host’s immune system except for PCV1 (28). PCV2 is widely known to inflict Postweaning Multisystemic Wasting Syndrome (PMWS) upon pigs ranging from 5-18 weeks of age, and brings about symptoms including steady weight loss, fever and lymphadenopathy (28, 35). Despite a 70% sequence homology with PCV2, PCV1 has not been linked to PMWS nor any other known affliction, though recent studies suggest it may play a sort of secondary role when both PCV-1 and -2 are present (1, 28, 35). Not until the definitive report of PCV-2 as the exclusive inducing factor of PMWS in 2001 however, did both PCV-1-and-2 fall under the intensive research scrutiny under which they currently exist (19, 35). These continued investigations have given scientists the opportunity to further investigate not only the virulence of PCV’s but also their apoptotic activities and potential usefulness as therapies outside of veterinary sciences.

1.5 Virus – Induced Apoptosis

Apoptosis is a physiological process by which multicellular organisms regulate cell death and tissue development and repair. It also regulates the elimination of cells that have genetic damage. This process is known as “programmed cell death (28).” The initiation of apoptosis is marked by the condensation of genomic DNA, the formation of bulges in the nuclear and plasma membranes, the loss of microvilli (in those cells that have them) and the detachment of affected cell from the surrounding cells (28). The progression of apoptosis leads to the disruption of the cytoskeleton, condensation of the cytoplasm and compaction of the cellular organelles and cleavage of the DNA. In the final stages of apoptosis, the cell shrinks and forms fragmented, membrane bound vesicle-like structures called “apoptic bodies.” These bodies are cleared from the tissue by macrophages (28).

Caspase activity has been linked to many of the biochemical and morphological features of apoptosis described above. Caspases are proteases containing cystine residues in their active sites that
cleave substrate after specific tetrapeptide motifs (34). They cleave sites following aspartic acid residues. The caspases are expressed as zymogens. These zymogens are activated by the initiation of apoptosis when cytochrome c is released into the cytoplasm (15). The effector capases, capases -3 and -7, are responsible for cell demolition during apoptosis. The effector capases initiate a pathway that leads to the destruction and disposal of key cell structures. Once activated, capase-3 is also responsible for the degradation of cellular DNA into fragments thus making it a biomarker of apoptosis (34).

The apoptosis inducing protein known as Apoptin is typically regulated by caspases which function as apoptosis initiators. Caspases are activated via two signaling routes: the extrinsic or the intrinsic (27). In the extrinsic pathway, death receptors are ligated. This triggers the recruitment of FADD and initiator caspase-8 into a death-inducing signal complex. In the intrinsic mitochondrial pathway, cytochrome c is released into the cytosol. This process is normally regulated by Bcl-2. Cytochrome c binds to Apaf-1 and results in the activation of initiator caspases-9 in the complex apoptosome. The two pathways converge when effector capases are activated and result in cell death due to cleavage of cellular substrate (27).

Capases are not the only pathway through which apoptosis can be initiated. It is known that the p53 transcription factor enhances the rate of transcription for a set of genes. Some of the genes are those that promote apoptosis. Under normal cell conditions, p53 levels are kept low by protease-mediated degradation. Low levels of p53 have also been linked to the sequestration of the protein (23). The transcription factor is negatively regulated by MDM2, which is encoded for by one of the genes which p53 enhances. MDM2 partially binds to the DNA binding domain of p53; this causes the degradation of the p53 to accelerate (15).

When cell damage is detected, expression of p53 is upregulated. Phosphorylation of p53 in the cytosol results in an extended half-life for the transcription factor and it enhances p53 activity and inhibits the binding of MDM2 (15). Stabilizing p53 results in the upregulation of the protein p21. The
role of p21 is to bind and inhibit cyclin-dependant kinases, such as cyclinD and CDK4. CyclinD and CDK4
normally complex and phosphorylate retinoblastoma protein (Rb). The phosphorylation of Rb allows the
cell to progress from G₁ to S phase (15). G₁ arrest allows the cell machinery time to attempt to repair the
damage that lead to the upregulation of p53. If the damage is irreparable, p53 guides the cell down an
apoptotic pathway (23).

The apoptotic pathway is regulated by the family of proteins known a Bcl-2/Bax. These proteins
are essential to the regulation of the pathway. Bax, found in the cytoplasm, promotes the pathway
while Bcl-2, bound to the mitochondrial membrane, inhibits it. Bcl-2 regulates ion transport in and out of
the mitochondria and protects the membrane from damage. It is believed that p53 regulates the level
of expression of these two proteins. If the production of Bax is favored, it is thought to bind to Bcl-2 on
the mitochondrial membrane and results in a loss of selective ion permeability and the release of
cytochrome c and apoptosis inducing factor (AIF). Once in the cytoplasm, cytochrome c and AIF induce
apoptosis by activating caspases and chromatin condensation (15).

1.6.1 Apoptosis in Circoviruses

Two of the Circoviridae viruses are known to induce apoptosis. PCV2 was shown to induce
apoptosis in PK-15 porcine kidney cells post-infection. The protein encoded in the novel ORF3 of PCV2,
called the VP3 protein, is thought to be responsible (19). Similarly, the Chicken Anemia Virus (CAV) also
induces apoptosis via production of a VP3 product, aptly named ‘Apoptin’ (22). CAV Apoptin, as well as
PCV2 VP3 protein, have been shown to affect only transformed (cancer) cells, while failing to destroy
normal cells, drawing attention to itself as a potential cancer-fighting drug (20, 22). Because a third
reading frame has been identified in PCV1, homology suggests that a similar apoptotic mechanism may
lie somewhere in that region. The potential of an apoptosis-inducing VP3 protein in PCV1 provides a
starting point for experimentation.
1.6.2 The CAV VP3 Mechanism of Apoptosis

The ORF3 of CAV encodes the 13.5 kDa VP3 protein known as Apoptin. Apoptin’s 121 amino acids include two proline stretches and two positively charged regions (8). The protein is responsible for apoptosis induction (8). The protein inhibits the anaphase-promoting complex/cyclosome before entering into the nucleus. This leads to cell cycle arrest and the induction of apoptosis (8).

APC/C is a large multiprotein complex. It acts as an E3 ligase that selectively ubiquinates cyclins by targeting them for proteasome degradation. This results in the cell progressing through mitosis (30). APC/C also targets securin, a protein that inhibits the enzyme separase. Separase causes the premature unbinding of sister chromatids (36). Destruction of this targeted activity leads to the activation of seperase and therefore the separation of the sister chromatids (36).

APC/C therefore is an important target for viral protein because interrupting its activity leads to G2/M arrest and the induction of apoptosis (10). Apoptin may induce G2/M arrest and apoptosis by associating with APC 1, APC/C’s largest subunit. The association results in the disruption of the complex and the accumulation of the substrates usually targeted by APC/C (31). The association between Apoptin and the anaphase promoting complex has not been observed in primary cells (31).

The complete mechanism by which Apoptin induces apoptosis is still unknown. However, it is known that CAV Apoptin is unique in two ways. First, Apoptin is p53-independent. The p53 pathway is the mechanism by which most cancer treatments induce apoptosis in tumor cells. However, approximately half of tumors develop mutations in p53 during growth. This results in a poor response to traditional therapies (5).

Second, Apoptin only induces apoptosis in transformed cells. These discoveries lead to a study of the localization of the protein. The localization of Apoptin differs between normal and transformed cells (5). In normal cells, the Apoptin localizes in the cytoplasm. In transformed cells, it localizes in the
nucleus. This localization is an important aspect of the apoptotic activity of the protein. This is due to a change in the nuclear transport system in transformed cells (5). It has been shown that Apoptin associates with the subunit 1 of the anaphase-promoting complex in transformed cells.

1.7 Apoptosis Induced by PCV2 VP3

PCV2 ORF 1 encodes for Rep protein while PCV2 ORF2 encodes for Cap protein (24). PCV2 VP3 is known to be a direct cause of viral pathogenesis in vivo (24). It has also been shown to induce apoptosis in H1299 cells but not primary cells. Therefore it is capable of cell-type specific apoptosis (3). The difference in localization of PCV2 VP3 relative to Apoptin suggests that viral protein induced apoptosis in transformed cells is propagated by different mechanisms.

A mechanism has been proposed. It is believed that PCV 2 VP3 binds to porcine E3 ubiquitin ligase, also known as pPirh2, in the cytosol of transformed cells. The interaction results in the destabilization of pPirh2 and a decrease in the cellular proteins associated with a rise of p53. This triggers an onset of apoptosis (24).

1.8 Nucleocytoplasmic shuttling

It has been shown that the localization of Apoptin in the nucleus of the transformed cells is required for the association of the protein with APC/C and for the induction of G2/M arrest and apoptosis (31). Translocation of Apoptin to the nucleus occurs during interphase and requires transiting the nuclear envelope. Movement of proteins across the nuclear envelope is highly regulated (37). Therefore, the cellular machinery and the mechanisms of nucleocytoplasmic shuffling are essential to Apoptin-induced apoptosis in transformed cells.

Nucleocytoplasmic transport occurs through a unit known as the nuclear pore complex (NCP). The NPC channel spans the inner and outer membranes of the nuclear envelope and serves as a portal between the cytoplasm and the nucleoplasm (8). The NPC is made of up proteins known as nucleoporins
or nups. There are about 30 different nucleoporins found in humans which display a small array of functional diversity. A single NPC may be composed of up to 1000 nups (8). Certain nups consist of FG repeats that bind directly to transport receptors carrying substrates through the NPC (19).

Generally, ions and molecules smaller than 40 kDa diffuse passively through the NPC. Larger molecules are translocated by an interaction between the molecule’s transporter and the FG domain bearing nups (37). The exact mechanism of this process of translocation is unknown but it has been proposed that the transiting the NCP requires the transporter-molecule complex to overcome physical or energetic barriers (37). Most nups are members of the kayropherin-β/importin-β family. The proteins are classified as either importins or exportins. The classification is based on the direction of transport they are involved in relative to the nucleus (24). The movement of molecules via these transporters is regulated by Ran, a guaninsine triphosphatase (37).

A Ran-GTP gradient across the nuclear envelope seems to be responsible for the direction of transport. The gradient is established by the activities of RCC1, a Ran guanine nucleotide exchange factor, and the RanGAP, which promotes Ran-mediated GTP hydrolysis (24). RCC1 is nucleoplasmic while RanGAP is cytoplasmic (24).

Nuclear import begins in the cytoplasm and is mediated by karopherin. The importin and the molecule enter the nucleus via the NPC. Once inside the nucleus, the N-terminus of importin binds to Ran-GTP. This releases the bound molecule. The importin-Ran-GTP complex transports back to the cytoplasm where RanGAP causes Ran to hydrolyze the GTP and the complex is disassembled. NTF2, a Ran importer, binds to the Ran-GDP complex and transports it to the nucleus. In the nucleus, RCC1 replaces GDP with GTP and the cycle of import begins again (24).

However, the regenerated Ran-GTP may follow another pathway. The complex could associate with the N-terminus of an exportin-molecule complex and begin a cycle of export from the nucleus.
Once in the cytoplasm, Ran is stimulated by RanGAP to hydrolyze GTP and disassemble the complex. Exportin and Ran are then transported back to the nucleus (24).

Exportin and importin require stimulation in order to bind to a molecule. Importins recognizes and binds to a protein bearing a nuclear localization signal (NLS) while exportins recognize and bind to proteins bearing a nuclear export signal (NES) (37). Generally, an NLS is comprised of a region of protein enriched in basic amino acids. They can be either monopartite or bipartite (33). For some proteins, importin-β binds directly to the NLS. For others, importin-α recognizes and binds to the NLS and importin-β binds to importin-α to enable transport (33). An NES generally is composed of a single of a protein that is rich in hydrophobic amino acids (3). Most of the NES-bearing proteins are transported by Crm-1, an exportin (3).

1.9 The NLS and NES of Apoptin

The study of Apoptin has identified both an NLS and an NES sequence motif. It is believed that the NES lies in the N-terminus of the protein. It spans residues 37-46. The NLS is believed to lie within the C-terminus, spanning residues 70 to 121 (8). In both transformed and primary cells, the NLS and the NES of Apoptin are functional. It has also been found that the export of Apoptin is Crm-1 dependent. Therefore, nucleocytoplasmic shuttling must be responsible for the cell-type specific localization of Apoptin (8).

The initiation of cell-type specific apoptosis by Apoptin is dependent of the process of nucleocytoplasmic shuttling. If the functionality of the NLS is lost, Apoptin is rendered incapable of entering the nuclei of transformed cells and therefore it is unable to induce apoptosis by interacting with APC1. Additionally, the loss of a functional NES also inhibits Apoptin’s ability to induce apoptosis in transformed cells. It has been proposed that the loss of NES eliminates multimerization which may be critical to the process of apoptosis induction (8).
2 Materials and Methods

2.1 PCV 1 ORF3 Sequence Analysis

The clustal alignment feature of CLC Free Workbench 4.0.2, developed by CLC bio A/s, was used to compare the ORF3s of PCV1, CAV, TTV and PCV2, type A and B. The comparison was used to construct two truncation mutants of PCV1 ORF3. The first mutant contained the NES and the second contained the NLS.

2.2 Isolation of PCV1 ORF3 Mutants

Each mutant was isolated and amplified by PCR from PCV1 ORF3 template DNA using the following primers: Post NES F and Post NES R in conjunction with PCV1 VP3 R and PCV1 VP3 F, respectively. The Post NES F and R primers were designed based on the putative location of the NES sequences as determined from by the clustal alignments. The Post NES Forward primer contained the EcoR1 restriction site and was engineered as follows: 5’-GCG AAT TCA CGT TAC AGG GAA CTG CTC GGC-3’. The Post NES Reverse primer contained the BamHI restriction site and was engineered as follows: 5’-GCG GGA TCC TCA GCC GAG GAC TTC CCT GTA ACG-3’. The PCV1 VP3 primers were designed by Conerly and McShea. The PCR reaction mixture contained 2 µL of 10X PCR buffer (50 mM KCl; 10 mM Tris-HCL (pH 8.3); 15 mM MgCl2), 1 µL of each primer, 1 µL of Taq DNA polymerase, 1 µL dNTPs and 13 µL of ddH2O. The reactions mixtures were run through the following PCR cycle: initial denaturing at 95° for 4 minutes followed by 30 amplification cycles (95° for 30 seconds, 55° for 30 seconds, 72° for 1 minute). The reaction mixtures were held at 72° until retrieved from the PCR machine.

The reaction components were then separated by gel electrophoresis on a 0.9% agarose gel containing 0.5 ug/mL ethidium bromide. Bands corresponding to a 0.300 kb DNA sequence was identified and excised from the gel. The gel slice was purified using the Promega Wizard Plus Midiprep PCR Purification System. The purified DNA was inserted into a T-vector, by ligation, using the Promega
pGEM-T Vector kit. DH5α competent E. coli were transformed with the T-vector ligation products and grown over night at 37° on agar plates containing 0.100 mg/mL ampicillin, 0.1 mM IPTG and 40 µg/mL Xgal. The T-vector contains the β-galactosidase gene, which spans the point of insertion. From the agar plates, single white colonies were selected and transferred to 15 mL tube containing 3 mL of LB media contain 1X ampicillin. The tubes were incubated overnight at 37° with vigorous shaking (220 rpm).

E. coli cells containing the T-vector were isolated from the overnight culture by centrifugation and the plasmid DNA was harvested by minipreparation. To confirm the presence of PCV1, a restriction digest was performed with BamH1 and EcoR1. The digested fragments were isolated by gel electrophoresis on a 0.9% agarose gel contain 0.5 µg/µL ethidium bromide.

2.3 Maintaining and Transfecting the H1299 cell line

H1299 non-small cell lung carcinoma cells were purchased from the American Type Culture Collection (ATCC), and grown initially in a large scale culture until they were of a great enough confluence and stability to be aliquotted into several smaller populations. The cells were grown in DMEM/High Glucose with 10% Fetal Bovine Serum and PSF (100 units/mL Pen G sodium; 100 mg/mL streptomycin sulfate; 0.25 mg/mL amphotericin B). Cells were passed ten times prior to transfection. A single passage included first equilibrating the new media into which growing cells would be placed, by placing flasks containing that media into a 37° C and 5% carbon dioxide incubation for about a half an hour. Old media in which the H1299 cells had been growing was then aspirated from the cell culture flask by vacuum. Cells were washed with 1X PBS and PBS was immediately aspirated. Cells were then dislodged from the floor of the flask by an immediate wash with Trypsin, which was also quickly aspirated. Cells lifted from the flask following several forceful knocks against the lab counter, and immediately bathed with the 5mL new, equilibrated media. 1mL of the washed cells were transferred to a new flask, and again allowed to incubate until they were ready for another passage.
Following passage 10 of the H1299 cells, 1mL of newly bathed cells was transferred to a tube containing 10mL of warmed, equilibrated media. From this 11mL, 200uL were added to wells of a 96 well reaction plate. The cells were transfected with p3XFLAG PCV1ORF3 vector (produced by Peter Conerly and Molly McShea) following the protocols of the proprietary Qiagen Effectene Transfection Reagent Kit. A sample containing no vector DNA was used as the negative control for the transfection, and another sample containing wild-type Apoptin served as a positive.

2.4 Apoptosis Assay

The 96-well plate containing the transfected H1299 cells was incubated at 37°C with 5% carbon dioxide for 18 hours. Following incubation, the Promega Apo-ONE Homogeneous Caspase-3/7 Assay kit was used to assess the apoptotic activity of the samples. The Caspase Substrate and Apo-One© Caspase-3/7 Buffer were combined. The reagent was added to the samples and the plate was incubated for 18 hours at room temperature. The fluorescence of each well was measured with a fluorescence spectrometer.

3 Results

3.1 Circovirus VP3 Sequence Alignments Illustrate Homology of Functional Sequences and the PCV1 Elongated C-Terminus

Prior to the physical manipulation of PCV1 ORF3 to assess functionality, it was first necessary to visualize and better understand homologous features among the virus and its apoptotic family members, to facilitate identification of potential localization sequences and sequence abnormalities that might be investigated. The CLC Free Workbench Sequence Viewer software was used to create a clustal alignment of the VP3 sequences of five circoviruses including PCV1, PCV2A, PCV2B, Chicken Anemia Virus and the Human Torque Teno Virus. The alignment can be seen in Figure 2 which has aligned the amino acid residue sequences by the highest calculated homology, where residues are color coded by
charge and polarity. A consensus sequence is also provided at the bottom of each five-sequence line up.

The most notable results of the alignment are the general demonstration of homology among the sequences, the indication of similarly aligned putative localization sequences and a clear visualization of the extremely elongated tail of Porcine Circovirus 1.

The first significant result of the alignment is the demonstration of appreciable sequence conservation among Porcine Circroviruses Type 1,2A and 2B, and to a lesser extent, among PCV1, TTV and CAV. Among the first ~105 residues of PCV1, PCV2A and PCV2B, very few discrepancies exist, and many of the charge-polarity motifs are conserved despite a residue change; for example, residues 40-70 among the three sequences are nearly identical. This conservation immediately suggests that Porcine Circovirus 1, though known to be non-virulent in pigs, may likely still encode functional sequences homologous to those in virulent strains and further, be capable of expressing a homologous VP3 protein capable of inducing apoptosis in transformed cells.

The second interesting aspect of the sequence alignments is the identification of putative localization sequences, both NES and NLS, in PCV1; conclusions drawn from assessing conservation in sequence regions known to encode the functional regions in PCV2A, PCV2B and CAV. First, the putative NES of PCV1 was identified by comparing the Proline, Leucine and Isoleucine-rich non polar regions of the PCV2A and B, CAV and TTV at residues 40-46, with the similarly encoded region in PCV1. It has been suggested that this region in PCV1 is functionally identical to the canonical export sequences of the other circoviruses. Previous research of PCV1 corroborates this hypothesis. The location of a putative NLS is less clear because of the abnormal C-Terminal region of PCV1. Interestingly however, several highly polar, Arginine and Lysine rich motifs in the PCV1 sequence (residues 145-147, 167-169) demonstrate a striking resemblance to the accepted bipartite NLS region of the apoptotically active Chicken Anemia Virus (Arginine and Lysine rich residues 87-89, 116-188). This result is two-fold in that it
provides evidence of sequence conservation and perhaps localization functionality as well as evidence to support the hypothesis that PCV1 possess a mechanism for apoptosis.

The alignment also gave insight to the extraordinarily long C-terminal tail of PCV1. PCV1 is 102 residues longer than PCV2A and B, 100 longer than TTV, and 85 longer than CAV. With such an expansive section of unstudied sequence, one must question whether or not the virus has maintained any functionality might have had, or even further, if it may have gained any nominal abilities

Following visualization of the sequence, and determination by homology of potential functional regions, it was necessary to design Polymerase Chain Reaction primers that could truncate the VP3 sequence into amplicons that when expressed would provide answers to the questions the above result presented.

3.2 PCR Primers Truncated PCV1 VP3 into Mutants Capable of Demonstrating Localization Functionality

The simplest way to assess the presence and or functionality of localization sequences is to uncouple the sequences via site directed mutagenesis; in essence cutting the NES and NLS apart, so that their simultaneous expression is knocked-out in vivo. Uncoupling the sequences was necessary in order to better understand the functionality of the VP3 protein. Thus, reaction primers were designed to initiate polymerase activity in the forward and reverse directions at Guanine 197 of PCV1 VP3, directly following the putative nuclear export sequence (See Figure 3A). Cutting and amplifying sequence at this location (with both designed primers from Figure 3A, and existing primers amplifying forwards from the VP3 start and backwards from the end of VP3) resulted in two truncation mutants of VP3; one containing the first 197 bases of the reading frame and hypothetically containing PCV1 NES, the second containing remaining portion of the VP3, including the abnormally elongated C-Terminal region, and what is supposed by sequence homology to be PCV1 NLS. The restriction sites EcoRI and BamHI, were
engineered into the primers so that the PCR products could be ligated into a T-vector containing those specific restriction sites.

Figure 3, panel B represents a UV-captured image of a gel electrophoresis of the PCR product, confirming that sequences of the desired truncation size have been cut and amplified at the intended site. In the figure, the far left column represents a 2-Log DNA Hyper Ladder, where the sizes of each band of the ladder are listed in kilobases along the left of the image. Column 1 represents the NLS amplicon (the Post NES region encompassing the lengthy C-terminus), and Column 2 is indicative of the NES amplicon (from the VP3 beginning to Guanine 197, including putative NES). Both amplicons ran to a reasonably expected size between 200 and 300 bases, indicating that the PCR was successful, and putative NES and NLS sequences had in fact been uncoupled.

3.3 Molecular Cloning to an Intermediate Vector

After uncoupling localization sequences, it was necessary to transfer the amplicons into an intermediate vector capable of housing the PCV1 fragment during transformation into, and replication as a plasmid in *Escheria Coli*. As such, the amplified mutants were inserted into Promega p-GEM-T Easy® Vector System, transformed and cloned in DH5α competent *Escheria Coli* (See Figure 4). Successfully transformed cells were visually selected from a LB/ Ampicillin/ IPTG /X-Gal growing media by a blue white screen, as positively cloned cells containing the PCV1 VP3 truncation plasmid would have inserted into the T-Vector in such a manner that the bacteria were no longer capable of cleaving X-Gal (remained white in color), but still maintained their inherent Ampicillin resistance. Positive colonies were selected and used to inoculate LB media containing also Ampicillin. After a period of 24 hour growth and incubation, cells were harvested from the culture and the plasmid DNA was lysed from the cells. The plasmid DNA was restricted using EcoRI and BamHI restriction enzymes, in an attempt to release truncated VP3 from T-vector at engineered cut sites, and ultimately confirm positive ligation. Though
restriction has yet to be confirmed by gel electrophoresis, linearized T-Vector had been observed, at least suggesting some partial ligation/restriction activity.

Despite lack of ligation confirmation, extensive rounds of T-Vector screening consistently produced positive colonies. This suggests that the progress of the experiments towards the final green fluorescent protein vector (pEGFP) is not far from completion, and only a positive restriction and simple amplicon insertion stands between future researchers and a confident confirmation and negation of localization activity in PCV VP3.

Inducing Apoptosis in H1299 Cells

3.4 Transfection of H1299 Cells with Wild Type PCV1 in p3X-FLAG Vector Suggests PCV1 Induces apoptosis

In parallel to the experiments truncating PCV1 VP3 to assess localization functionality, separate experiments were conducted with wild-type, uncut PCV1 VP3 to assess its apoptotic activity in transformed cells. It was previously established that the Chicken Anemia Virus VP3 (Apoptin) and Porcine Circovirus 2 Type B VP3 induce apoptosis in transformed, however it has not to this point been proven or disproven whether or not PCV1 does the same. Non-small cell lung carcinoma (H1299 line) cells were transfected with p3XFLAG-myc-CMV-26-PCV1 vector, and assessed for cell death after an 18 hour incubation period using the Promega Apo-ONE Homogeneous Capases -3/7 assay. For the very first time, PCV1 was demonstrated in wild-type form as capable of inducing apoptosis in transformed cells by a caspase-3-7 pathway. The data presented in Figure 5 is representative of both runs that were indicative of apoptosis. Note that the top six curves on the graph are representative of apoptotic activity, with expected peaks at 521nm; of the six curves, the first, fifth and sixth (from the top) are PCV1 VP3, while the second, third and fourth are positive controls (wild-type apoptin). The lowest three curves on the graph are clustered closely together, contain no peaks, and are indicative of no apoptotic activity; they are negative controls.
This information is exceedingly interesting, in that for the very first time anywhere in the scientific community, Porcine Circovirus 1 has been demonstrated capable of inducing apoptosis in transformed cells, and it now too may be considered in future research as a potential cancer therapeutic.

4 Discussion

Over 1.3 million new cases of cancer are diagnosed in the United States every year, and more than 550,000 Americans die every year from the disease (1). The effects of cancer are far reaching and there is a clear and obvious need for a clinical solution to the problem; a solution which has not yet been found. That is why ever expanding research that slowly uncovers each and every potential treatment is so essential to bettering the health of the human population. Such thorough research involves not only developing new, cutting edge synthetic drugs, but also studying what nature has already offered us; the biochemical pathways of naturally-occuring viruses, like Porcine Circovirus Type 1, that might hold the key to the eventual eradication of cancer.

PCV1 VP3 was compared, by alignment of amino acid residues, to that of other members of the Circoviridae family of viruses: CAV, PCV2 A, PCV2 B and TTV. Although PCV1 is nonvirulant, its striking homology to the other members of the Circoviridae family, as illustrated in Figure 2, suggested that it, like CAV, may be capable of inducing apoptosis. In fact, the results of the apoptosis experiments, seen in Figure 5, performed as part of this study showed that it does indeed kill transformed cells in a p53-independent manner with approximately the same efficiency as the positive control, wild type Apoptin. This suggests that PCV1 VP3 protein may be useful as cancer therapy.

Although PCV1 shows 65% homology to CAV, there is one notable difference in the sequence of the ORF3 of PCV1 as compared to the other Circoviridae viruses. As seen in Figure 2, PCV1 has an abnormally long C-terminus; most likely the result of genetic drift. Sequence analysis reveals that the
region is rich in hydrophobic amino acid residues. These hydrophobic residues suggest a clump-like formation of the C-terminus tail. Because the functionality of wild type PCV1 VP3 was not lost, it suggests that the tail either has no effect on apoptosis or it is promoting the activity.

Further studies should investigate the functionality of the tail. We believe that if the tail is removed the protein will not lose its functionality because both the NLS and NES regions have been maintained but we suspect that some efficiency will be lost. This claim is based on the fact that some of the samples of PCV1 VP3 induced apoptosis at a rate that was higher than the wild type Apoptin (see Figure 5).

Studies of CAV suggest that the localization sequences are essential to the virus’s ability to induce apoptosis (7). If further research finds that removing the long C-terminus of PCV1 VP3 inhibits apoptosis, it would suggest that the region is indeed promoting selective killing. However, that does not mean that the NLS and NES regions are nonessential to the functionality of the VP3 protein.

CAV is known to localize in the nucleus of transformed cells (7). Further studies of PCV1 VP3 should investigate the localization of the protein in cancer cells. The VP3 region of PCV1 was analyzed and the location of the NLS and NES were determined based on homology and the location of the same regions in the other Circoviridae viruses. Using the primers designed during this research study, the VP3 should be broken down into two fragments; one containing the NLS and the other containing the NES. The fragments should be cloned into the terminal GFP vector so that the localization may be determined. We believe that the vector containing only the NES will not exhibit apoptotic activity as the protein will be located in the cytoplasm. We believe that the vector containing the NLS fragment will exhibit apoptotic activity as the protein will localize in the nucleus. If the results of such an experiment are not as expected, it would suggest a difference in the mechanism by which PCV1 VP3 induces apoptosis as compared to its counterpart, CAV.
In addition to dividing the genome in half, the functionality of the NES should be studied further by uncoupling the proposed region from the rest of the VP3 region to assess the existence of a nucleocytoplasmic shuttling mechanism and determine its potential role in apoptosis induced by VP3. To do this, researchers should study the activity of the NES alone and the VP3 without the NES.

There is currently no evidence to suggest that wild type PCV1 causes any known disease. For the purposes of this study, the ORF3, which overlaps ORF1, as seen in Figure 1, was excised from the genome and forcibly expressed in transformed cells. Because PCV1 VP3 was shown to induce apoptosis but wild type PCV1 is nonvirulant, we believe that the VP3, encoded in the ORF3, is not expressed in nature. Our claim is further supported by the fact that ORF3 is not necessary for viral replication. The proteins necessary for replication of PCV1 are encoded in ORF1 and 2. Interestingly, CAV is known to require ORF3 for replication. This, coupled with the fact that ORF3 has experienced genetic drift and thus developed an abnormally long C-terminus, suggest that the VP3 is functionally dead in nature. If this is the case, the virus may, in the future, evolve in such a way that the ORF3 is lost. Further studies should investigate whether or not the protein is expressed by the virus in both primary and cancer cells.

If the results of such an experiment show that the virus does not express the VP3 protein, it would confirm our suspicions. However, if the virus does express the protein, but does not kill primary cells, it would speak to the selectivity of the protein which has not been previously studied. We believe that, like the other members of the Circoviridae family, PCV1 VP3 kills transformed cells selectively. Further studies should investigate this claim by transfecting primary cells with the both the wild type PCV1 virus and the excised VP3 region.

The Circoviridae viruses analyzed through the course of this research study are not the only ones to show great homology with PCV1. In fact, in 2008 researchers discovered a striking resemblance between PCV1 and the Pigeon Circovirus (PiCV). A sequence alignment analysis of twelve variations of PiCV revealed PCV1 to be the most homologous. A phlyogenetic tree tracing
the origins of the PiCVs to direct divergence from PCV1 was constructed (40). Interestingly, all PiCVs encode the hypothetical third reading frame (called C2 instead of ORF3 in the literature) of 126 amino acid residues. The Dove Circovirus encodes a hypothetical C2 of 199 residues that is remarkably close to the 202 residues encoded by the elongated PCV1 (40). This groundbreaking discovery could lead to a better understanding of PCV1, a virus whose genomic makeup was once believed to be unique. The team of scientist behind the discovery does not appear to be aware of the connection between PCV1 and PiCV and are most certainly are not aware that the elongated ORF3 of PCV1 is capable of encoding a cancer destroying VP3 protein. Study of the PiCV in parallel to PCV1 is perhaps the most exciting of the opportunities presented by the results of these experiments. If PiCV exhibits apoptotic abilities, it would support our belief that the elongated tail is providing the virus with some evolutionary advantage. If the elongate tail is not providing a benefit, it is unlikely that the mutation would be conserved and it even more unlikely that two viruses would have such similar mutations and without conservation of functionality. Further research should investigate the functionality, localization, and apoptotic abilities of PiCV and its variants.

We have shown, for the first time, that PCV1 VP3, a homolog of Apoptin, has the ability to induce apoptosis when expressed in transformed cells. This activity makes the protein of great interest as a therapeutic agent and a possible solution in the ongoing fight against cancer. The activity was shown, through a capases 3-7 pathway, to be independent of p53, the tumor suppressor gene. As many as 50% of cancerous tumors are known to have a mutated p53; meaning they are unresponsive to traditional cancer therapies (S). Thus, the implications of this study are far reaching. A continued effort to understand PCV1 VP3 and its apoptotic activity is an important step in the right direction; in the direction of a world without cancer.
5 Figures

Figure 2: Clustal alignments illustrate homology and suggest localization sequence locations. Using the Clustal Alignment feature of CLC Workbench software, VP3 sequences of CAV and PCV1, as well as those of Circoviridae family members Porcine Circovirus 2 (PCV2) types A and B, and Torque-Teno Virus (TTV), were first translated and consequently aligned to establish sequence homologies, as well as to help identify characteristic localization sequences. The alignment also illustrates the extraordinary length of the C-terminal region of PCV1 in comparison to its counterparts.
Figure 3: Polymerase chain reaction (PCR) primers strategically truncate PCV 1 VP3. (a) Reaction primers were designed to initiate polymerase activity in the forward and reverse directions at Guanine 197 of PCV1 VP3, directly following the putative Nuclear Export Sequence (NES) resulting in two mutants: One containing the first 197 bases of the reading frame and most likely containing PCV1 NES, the second containing remaining portion of the VP3, including the abnormally elongated C-Terminal region, and potentially the Nuclear Localization Sequence (NLS). (b) Gel Electrophoresis experiments following PCR confirms that the third open-reading frame has in fact been cut and amplified at the intended site. Lane 1 (NLS) contains the 424 bp segment from Guanine 197, while Lane 2 contains the 197 bp fragment from ORF3 start to the end of the NES Reverse primer. The fragments both ran at about 0.3kb; minor differences in size are indiscernible.
Figure 4: Schematic illustrates ligation of PCV1 VP3 truncation mutants into p-GEM-T Easy Vector System; Into pEGFP Vector. Following PCR, amplified mutants were inserted into Promega p-GEM-T Easy® Vector System, transformed and cloned in DH5α competent Escheria Coli. Successfully transformed cells were visually selected from a LB/ Ampicillin/ IPTG /X-Gal growing media, and re-cultured. After a period of 24 hour growth and incubation, cells were restricted using EcoRI and BamHI restriction enzymes. Restriction is confirmed by gel electrophoresis. Positive restrictions are ligated into pEGFP vector for examination of localization functionality and potential apoptotic activity via killing assays.
Figure 5: Apoptotic activity of wild type p3XFLAG-myc-CMV-26-PCV1 is assessed by killing assay of H1299 non-small cell lung carcinoma cells. H1299 Cells transfected with non-truncated wild type PCV1 ORF3 in the 3X-FLAG vector underwent apoptosis. The brown, violet and purple curves between 340 and 400nm are positive controls (CAV Apoptin). The teal, green and blue curves on either side of positive controls are those for PCV1 VP3. Negative controls are those closely grouped at about 200nm. The data suggests that PCV1 does induce apoptosis by a caspase 3-7 pathway.
6 References


