Exploring Apoptin's Anticancer Ability: Uncoupling Nuclear Export and Multimerization

A Major Qualifying Project

Submitted to the Faculty
of the
WORCESTER POLYTECHNIC INSTITUTE
In Partial Fulfillment of the Requirements for the
Degree of Bachelor of Science
By

_____
Jessie Brasefield

_____
Doug Tischer

Date: April 28, 2009

Approved:

_____
Professor Destin Heilman, Primary Advisor

Table of Contents

Table of Figures

Abstract

Apoptin, Chicken Anemia Virus VP3, selectively induces apoptosis in transformed cells but not primary cells. In all cases, Apoptin undergoes nucleocytoplasmic shuttling, but partitions into the nucleus of transformed cells and the cytoplasm of primary cells. The N-terminal NES of Apoptin and the domain for multimerization overlap and all previous attempts to uncouple these activities have failed. In this study, we successfully uncoupled nuclear export and multimerization through site-directed mutagenesis of ApI37A and ApI40A. Uncoupling these activities and studying them independently could shed insight into basic biochemical differences between transformed and primary cells that could be exploited for targeted therapeutic purposes.

Introduction

The Chicken Anemia Virus (CAV)

The chicken anemia virus was discovered to be an avian pathogen that causes the eradication of thymocytes and erythroblastoid cells in chicks through the induction of apoptosis (1).It resembles the porcine circovirus (PCV) and is a member of the family *Circoviridae* and the genus *Gyroviridae*. It is a circular, single stranded DNA composed of approximately 2300 nucleotides that has 3 open reading frames (ORFs) which all partially overlap: VP1, VP2, and VP3 (2,3). ORF 3 encodes the 51 kDa protein VP1 which responsible for the assembly of the viral capsid. ORF1 encodes the 28 kDa protein, VP2, a non-structural protein with dual specifity phosphatase (DSP) activity. It has a very unusual signature motif that may play a role in intracellular signaling during viral replication and is the first DSP to be identified in a small viral genome (4). VP2 is expressed at barely detectable levels during infection (5,4). Lastly, ORF2 encodes the 13 kDa protein, VP3, also known as Apoptin (5). In many studies, it is shown that Apoptin selectively induces apoptosis in chicken and human tumor cells and not in most normal cells through a p53 independent pathway (1,6,7). CAV's VP3 alone is necessary and sufficient to induce apoptosis and discriminate between cancerous and normal cells.

CAV VP3: Apoptin

The VP3 of CAV (Apoptin) is composed of 121 amino acids with two proline rich regions and two basic regions. It overlaps with CAV VP2 (7) and has no discernable regular structure (8). Using circular dichroism (CD), Leliveld and coworkers determined that Apoptin was "nearly devoid of an alpha-helical structure" and that the CD spectrum of Apoptin resembled other proteins with an unordered structure (8).

Apoptin has a bipartite basic nuclear localization sequence (NLS) located at the C-terminus of the protein between amino acids 80-121 (1,2,9). To determine whether the bipartite NLS was

functional, Heilman and coworkers induced point mutations K86A, K87A, R88A in the first part of the bipartite NLS, and K116A, R117A, and R118A in the second part of the bipartite NLS (all on Ap-pmNLS). Results show that the mutant Ap-pmNLS mislocated to the cytoplasm of H1299 cells, indicating that the bipartite NLS is indeed functional (2).

Apoptin has a distinct leucine rich nuclear export sequence (LR-NES) located at the N-terminus of the protein between amino acids 33-46 (2). There have been many different studies done to determine the correct NES for Apoptin. The NES was believed to be between amino acids 97 through 105 and that amino acids 33 through 46 contained a leucine rich sequence (LRS) that facilitates nuclear import (10,11). However, this was disproved by Heilman et al., (2006) by inducing point mutations L44A and L46A on GFP-Ap-pmNES which actually mislocated to the nucleus in PFF (primary foreskin fibroblasts) cells showing that the NES (aa 33-46) is functional (2). In addition, they also showed that the putative N-terminus NES was CRM1-dependent which further indicates that the N-terminus NES is functional (2). The NES of Apoptin (IRIGIAGITITLSL) resembles the NES of Rev- or PKIα-. The canonical sequence for NESs is $\Phi(X)_{2-3}\Phi(X)_{2-3}\Phi X\Phi$, where Φ represents hydrophobic amino acids (Lys, Val, Ile, Phe, or Met) and X is any amino acid (12). Apoptin's NES fits the canonical sequence from I37 to L46 (IAGITITLSL), while the rest of the N-terminal domain I33 through G36 (IRIG) are part of the NES, but do not fit the canonical sequence for NES.

Nuclear Export: LR-NESs and CRM1

Leucine rich nuclear export signals (LR-NESs) typically contain alternating hydrophobic residues that fit the canonical NES and are additionally rich in acidic amino acids, which makes them more negatively charged (13). They are also either flexible themselves or near very flexible regions and tend to be in an alpha helix conformation at least at the N-terminal end of the NES (although there are small percentages of NESs in a coil or beta sheet conformation) (12,13). Alpha-helical NESs also have a clear

tendency to have the majority of their hydrophobic residues on one side of the helix, the side which makes contact with their exportin, typically CRM1. La Cour et al. (2004) analyzed the structures of six NESs in an alpha helix conformation and noticed that the structure of the N-terminal hydrophobic residues of NESs tend to stack together on one side of the helix and on the C-terminal end of the NES, the hydrophobic amino acids bend off differently. Their data is not completely representative of all NESs, but are just shown to illustrate a commonality between a large percentage of NESs in alpha helix conformation (13).

The interaction surface of CRM1, the exportin protein that binds to LR-NESs, is a hydrophobic groove formed by 3 helices (H11A, H11B, and H12A) that is structured as follows: Two helices (H11A and H12A) parallel to each other spaced farther apart at the end that bind the N-terminal end of NESs which gradually narrows to a bottle neck towards the C-terminal end of the NES. Past the bottleneck of the hydrophobic pocket that binds the C-terminal end of NESs is a narrow channel. H11B is underneath both H11A and H12A and additionally contributes to the hydrophobicity of the groove. Dong and coworkers comment on the fact that the hydrophobic pocket of CRM1 where the N-terminus of NESs interact with CRM1 is a wider area that besides binding alpha helix NESs, could also accommodate NESs in a coil or beta sheet conformation (which illustrates the degeneracy of NESs) (12).

Nuclear Import and Export: Apoptin

Nuclear import and export is done through nuclear pore complexes (NPCs) with the help of karyopherins (kap) (Figure 1). NPCs allow molecules with a mass of 40kDa or less to pass freely through their aqueous channel. Any molecule larger than 50kDa can actively pass through with the help of a kap. Cleverly, Kaps that mediate import are called importins (IMPs) and kaps that mediate export are called exportins (11).

Apoptin has been shown to be imported into the nucleus via the IMPβ1 protein which uses the RanGTP cycle. IMPβ1 binds to the NLS of Apoptin in the cytoplasm which allows the docking of the VP3

to the NPC. The VP3 is then translocated across the NPC and into the nucleus. In the nucleus, RanGTP binds to IMPβ1 which causes the VP3 to be released (11). Exportins, such as CRM1, mediate nuclear protein export (11,2). CRM1 utilizes the Ran-GTP cycle. RanGTP is bound to CRM1. This complex then binds to the NES of Apoptin, docks the protein cargo to the NPC and is transported out of the nucleus and into the cytoplasm through the NPC. Once in the cytoplasm, the GTP is hydrolyzed to GDP by Ran with the help of Ran GTPase-activating protein, and the Apoptin cargo containing the NES is released from the CRM1 (11).

Apoptin: Nuclear and Cytoplasmic Shuttling (NES and NLS)

Apoptin has a shuttling mechanism that is needed for its apoptotic activity. In transformed cells, Apoptin is partitioned in the nucleus prior to apoptosis, while in primary cells, Apoptin is predominantly partitioned in the cytoplasm (11,2,10). In PFF (Primary Foreskin Fibroblast) cells, Apoptin exhibits a filamentous immune-staining pattern in the cytoplasm which suggests that Apoptin might be tightly associated with cellular filament networks (2)or that it naturally form fibrils. Once Apoptin enters the nucleus, it has been shown to bind or enter promyelocytic leukemia (PML) nuclear bodies (NBs) (11,2,10). PML NBs are structured protein bodies that are associated with DNA replication and repair, transcription regulation, RNA transport and more interestingly, apoptosis (11). Previous studies have shown that in tumor cells, Apoptin is phosphorylated at T^{108} by an unidentified protein kinase while in the PML NBs and in primary cells, T^{108} is not phosphorylated (11). However, conflicting results have also been demonstrated (9). It is also shown that in transformed cells, Apoptin interacts with the APC1 of the APC/C through a domain that overlaps with the NLS. Apoptin recruits APC1 to the PML NBs, and apoptosis is induced at G2/M. PML bodies are not detectable in the absence of Apoptin and during G2/M, PML itself is not heavily present suggesting this is a significant interaction and not artifactual (2). It is known that Apoptin's interaction with the APC1 is extremely important because nuclear localization in the absence of APC1 association does not induce apoptosis (2). Heilman et al. (2006) also

demonstrated that nucleocytoplasmic shuttling of Apoptin is necessary to induce Apoptosis. Inhibition of nuclear export by leptomyosin B in cancer cells doesn't significantly alter Apoptin's nuclear localization, but does eliminate its apoptotic effects. Similarly, artificially stuffing apoptin into the nucleus of primary cells with leptomyosin B does not kill the cells (2). Despite the important of nucleo-cytoplasmic shuttling, it is currently unknown how this process is differentially regulated in transformed and primary cells. One clue is that Apoptin's NES appears to be differentially active, while its NLS appears constitutive (2). One possibility is that shuttling is related to Apoptin's characteristic ability to form large multimers, since its multimerization domain overlaps with the NES.

Apoptin: Multimerization

It has been shown in many different studies that nuclear localization alone is not sufficient for efficient apoptosis of normal cells, which suggests that there is another step involved in apoptosis of transformed cells (6,2,14). Apoptin has a multimerization domain overlapping with the NES (aa 33-46). Using secondary structure prediction, Leliveld et al. (2003) determined that amino acids between Glu^{32} and Leu^{46} (the NES of Apoptin) might fold as an anti-parallel beta sheet with Ala^{38} and Gly^{39} oriented in a beta-turn or -hairpin. If this is the case, since the NES generally has alternating hydrophobic and hydrophilic amino acids, the amino acids would be oriented in such a way that the hydrophobic amino acids would protrude on one side of the hairpin, and the hydrophilic amino acids would do the opposite. They hypothesize that this motif might be responsible for Apoptin's multimerization properties and perhaps the Ile and Leu resides interlock between Apoptin monomers to form the Apoptin multimers (8).

Apoptin has been shown to be heavily multimerized and insoluble in normal cells, while in tumor or transformed cells, it is less multimerized and more soluble. Apoptin spontaneously forms multimers composed of approximately 30 to 40 subunits *in vitro* (8). Apoptin's heavily multimerized

state in normal cells along with its proposed association with cellular filament networks might explain the difference in solubility and location of Apoptin in normal cells versus transformed cells (2).

Cellular Concentration of Apoptin and Rate of Cellular Division Affects Programmed Cell Death

A study by Wadia et al. (2004), illustrated that the characteristic ability for Apoptin to localize in the nucleus of tumor cells is not the sole mechanism by which apoptosis is induced. They believe that programmed cell death (PCD) is dependent on a high cellular concentration of Apoptin to be able to form multimers large enough to then localize to the nucleus and induce apoptosis. They transfected ras-transformed mouse 3T3 cells with increasing amounts of Apoptin DNA and based on scoring of their microscopic images, they deduced that increased expression levels of Apoptin in tumor cells lead to better nuclear accumulation and finally apoptosis. However, their hypothesis was disproved by Poon et al. (2005) who quantitatively analyzed expression levels at the single-cell level rather than assumed increased expression levels based on the amount of DNA transfected. Poon and coworkers illustrated that there was no relationship between the level of Apoptin expression and the extent of nuclear localization in tumor cells. In fact, they were able to conclude from their data that increased levels of expression was indirectly proportional to nuclear accumulation in SAOS-2 cells (15).

The rate of cell division has been illustrated to affect Apoptin's ability to induce apoptosis in transformed cell lines. Tumor cells rapidly divide uncontrollably, although different tumor cell lines have different rates of replication. Indeed, faster growing tumor cells such as Hela cells which have population doubling times (PDT) of 24 hours, were killed within 2 days by Apoptin. Slower growing cells such as Saos-2 transformed cells whose PDT is 42 hours, took 5 days for Apoptin to induce apoptosis (16). Evidently, the rate of cellular division plays a role in inducing apoptosis of transformed cells.

Besides trying to discern the exact mechanism by which Apoptin induces apoptosis in cancer cell lines, even more perplexing is that Apoptin can indeed induce apoptosis in certain non-transformed cells (e.g. normal breast epithelial tissue) under certain reaction conditions (16). Clearly there is much to be discovered about the mechanism by which Apoptin induces apoptosis in different tumor cells versus particular non-transformed cell lines.

The Cell Cycle and Cell Cycle Regulators

The eukaryotic cell cycle is divided into four phases: G1, S, G2 and M. The G1 "gap" phase is the longest step in the cell cycle, typically consuming about 18 hours of a typical 24 hour cycle. The G1 phase is when a cell carries out its "normal" metabolic duties, such as energy metabolism and most protein synthesis. If a cell is no longer actively dividing (quiescent), it will enter an indefinite gap phase known as G0. Examples of cells in G0 are terminally differentiated muscles cells and neurons. Following G1 is the S, or synthesis, phase in which the DNA is replicated. S phase is followed by a smaller gap phase, G2, and then the cell proceeds into mitosis in M phase. After the chromosomes have been separated into two nuclei, the cell undergoes cytokinesis, yielding two daughters cells and beginning the cycle over again.

Cells have evolved tight regulation of the cell cycle and have checkpoints that ensure the cell is ready to progress to the next stage in the cell cycle. For instance, it would be undesirable for a cell to progress to G2 and then M phase before all DNA replication had been completed in S phase. In metazoans, cells also retain tight control of the cell cycle to prevent unchecked cell growth, leading to cancer.

At the heart of cell cycle regulation are the cyclin-dependent kinases (Cdk) which allow for the expression of genes needed in the next phase of the cell cycle. Not surprisingly, a Cdk's activity is dependent on the binding of a cyclin, protein's whose concentrations rise and fall in a cyclic nature with

the cell cycle. There are many different cyclins, with each set corresponding to different stages in the cell cycle. Once bound by a cyclin, the Cdk enters the nucleus and phosphorylates serine and theronine residues on histones. The DNA surrounding the histones changes packing and conformation, allowing for the expression of genes needed for the next stage of the cell cycle (Figure 2). Thus Cdks allow for the controlled, selective expression of genes that allow the cell cycle to continue to the next stage. But Cdks are also negatively regulated by cyclin depedent kinase inhibitors (CKI). As the name suggests, these inhibitors block the function of Cdks and are dependent on the presence of other cyclin proteins. CKI can prevent the cell from progressing to the next stage in the cell cycle, and can thus be used by the cell as a way of checking against cancer and uncontrolled cell growth.

Policing the Cell Cycle

Eukaryotic cells have evolved elaborate mechanics to check uncontrolled cell growth. Perhaps the most important cell cycle regulator is the p53/Mdm-2 system. p53 is an antitumor protein that acts as a central hub in the cell, being at the cross roads of many different proliferative pathways. If any one of these pathways signals that the cell is in danger of becoming cancerous, p53 is activated and arrests the cell cycle until the problem is resolved. If the damage is unrepairable, prolonged p53 activation commits the cell to apoptosis, killing the cell for the health of the organism. As evidence of p53's effectiveness, half of all cancers have mutated or deleted p53, sidestepping the watchful eye of this master cell cycle regulator (2,17). In terms of cancer treatment, a mechanism that kills cancer in a p53-independent manner would be very significant.

p53 is a transcriptional activator that activates expression of certain CKIs ($p21^{Cip1}$). Depending on the level and duration of expression, the CKI can either lead to a G2 cell cycle arrest or if the signal is sustained, CKI expression can commit the cell to apoptosis (17). This system gives the cell sometime to try to repair the damage that caused the cell cycle arrest. However if the damage is too great, the cell kills itself to prevent the spread of cancer. The action of p53 is kept in check in healthy cells by its

antagonist, Mdm-2, an E3 ubiquitin ligase that specifically tags p53 for proteolytic destruction (17) (Figure 3). Pathways converging on the p53/Mdm-2 system can induce cell arrest or apoptosis either by inhibiting Mdm-2 or activating p53. The Ras pathway responds to growth signals and can become hyperactive from mutations in receptor tyrosine kinases (RTK), like epidermal growth factor receptor (EGFR). When overactive, the RAS pathway inhibits Mdm-2, the result being an accumulation of p53 which induces expression of CKI leading to cell cycle arrest and apoptosis (17). This p53 check is important since roughly 30% of cancers have a hyperactive Ras pathway (18). Conversely, certain kinases (ATM and Chk2) directly phosphorylate p53, making it immune to Mdm-2 ubiquitylation. As before, p53 activates expression of CKI, causing cell cycle arrest and apoptosis.

The Anaphase Promoting Complex/Cyclosome

The anaphase promoting complex (APC) is a master regulator of several key checkpoints during mitosis and is being explored for its regulatory activity in other aspects of the cell cycle. Here, the focus will be on the APC's role in regulating mitosis. The APC is a large, 13 subunit E3 ubiquitin ligase. It can be found in two different forms, depending on the timing in mitosis. Early in mitosis, the APC^{Cdc20} form (the APC core bound to the protein CDC20) dominates. Later in mitosis, the APC^{Cdh1} form dominates. Each still has E3 ubiquitin ligase activity, differing only in their substrate specificity. The APC^{Cdh1} form prefers substrates with D-boxes and KEN boxes. The consensus sequences for these domains are R-X-X-X-L-X-X-X-X-N and K-E-N-X-X-X-E/D/N respectively (19).

Cyroelectron microscopy and immuno-locating experiments have begun to shed some light on the structure and organization of the APC, in both of its forms (Figure 4). The APC is formed around the central protein Apc1. On one side of Apc1 extends the "TRP arm," composed of subunits Cdc23, Cdc16, Cdc26, Apc9 and Cdc27. The TRP arm associates with the variable Cdh1 or Cdc20 subunits, linking it to the main body of the APC. The exact stoichiometry of the arm in relation to the rest of the APC is unknown. It is possible that the number of TRP arms is variable. Current estimates range from one to

three arms. In principle, each one could be capable of associating with Cdh1 or Cdc20. The variable subunits Cdh1 and Cdc20 are thought to be the subunits that interact with and recruit the substrate, thereby giving the APC is different substrate preferences. It is still open to investigation weather Cdh1 and Cdc20 first capture the substrate and then recruit it to the APC, or if the APC core proteins first capture Cdh1 or Cdc and then recruit the substrates. An attractive explanation as to why Cdh1 and Cdc are at the end of a long arm is so that the APC can accommodate an array of difference substrates sizes (20).

Opposite the Cdh1 and Cdc20 subunits lies the E3 ubiquitination domain of the APC, specifically the subunit Apc11. This subunit is responsible for taking an activated ubiquitin molecule from UBC (an E2 protein) and transferring it to whatever substrate happens to be bound. Apc11 has a zinc RING finger motif which it used to transfer ubiquitin to its substrate without itself ever forming a covalent bond to the ubiquitin. RING finger motifs are seen in other E3 ubiquitin ligases, such as SCF (Skp/Cullin/F-box) (20).

One of the most well known activities of the APC is promoting the transition from metaphase to anaphase by promoting the separation of sister chromatids. Once all of the mitotic spindles have attached to the kinetochores, APC^{Cdc20} ubquitinates securin, which normally inhibits the protease separase. Sepase then degrades the cohesin protein bridge linking sister chromatids, allowing them to separate and anaphase to progress. The system is so sensitive that one unattached mitotic spindle is enough to shut down the pathway, ensuring that each daughter cell will receive one copy of each chromosome (19,20).

It is easy to see how disruption of this master cell cycle regulator could cause significant damage to the cell, ultimately leading to cell cycle arrest and apoptosis. However, not all cell deaths are equal. Apoptosis refers to a very specific and regulated form of cell suicide and differs significantly from

necrosis, in which a cell explodes in an uncontrolled manner. These differences are important when discussing both cancer and viral infections.

Necrosis Versus Apoptosis

Necrosis and apoptosis are two general ways in which eukaryotic cells are killed. Necrosis typically takes place under toxic conditions, such as during a pathological reaction, and is considered to be a passive degeneration of cells. This type of cell death is characterized by loss of integrity of the cell membrane and the loss of control of cellular functions of internal cellular organelles. By contrast, apoptosis can be induced either by the cells themselves (intrinsic) or by cell of the immune system (extrinsic), as a type of "cell suicide." Apoptosis is characterized by cell shrinkage, perinuclear condensation of chromatin, fragmentation of nuclei, and fragmentation of chromosomal DNA (21). After these events take place, the cell fragments form into membrane-enclosed bodies and are phagocytosed by neighboring cells (1). The ability of neighboring cells to phagosize the apoptotic cells allows apoptosis to be more efficient and less detrimental to cell tissue than necrosis (7).

two routes through which apoptosis can occur, via the extrinsic or intrinsic pathway (Figure 5). In the extrinsic pathway, an outside cell directs the target cell to commit to apoptosis through the interaction of membrane proteins. The inducing cell presents the membrane protein FasL (Fas Ligand), which docks with the target cell's Fas transmembrane proteins, which forms a trimer. The cytoplasmic protein FADD (Fas Associating Death Doman) binds to the cytoplasmic side of the Fas transmembrane proteins, and serves as an adapter protein for recruiting procaspase-8 and procaspase-10. These inactive initiator caspases require protolytic processing before they themselves become active proteases. Due to their physical proximity, procaspase-8 and procaspase-10 are able to cleave and activate each other, producing Caspase-8 and Caspase-10. Caspase-8 and Caspase-10 then proteolytically convert

procaspase-3 to caspase-3, the so called executioner caspase because it degrades many of the cellular targets of apoptosis (22).

The route of intrinsic apoptosis is less well known. What is certain is that it is triggered by the release of cytochromeC from the mitochondrial intermembrane space. Together, cytochromeC and Apaf-1 (Apoptin protease activating factor) form the apoptosome, which cryoelectron microscopy has shown to resemble a wheel with six bent spokes. The apoptosome activates procaspase-9, which in turn activates procaspase-3 using similar proteolytic processing strategies as for the extrinsic pathway. Caspase-3 then degrades it cellular targets. The events that link apoptotic stimuli, such as DNA damage, to the downstream events of the intrinsic pathways are not yet established (22). Thus the molecular events linking CKI expression by p53 and cell cycle arrest to apoptosis are not well understood.

Virus-induced Apoptosis

Cell use apoptosis as a means of limiting viral egress. Results from a study done with mutant insect viruses (21) show that virus-induced apoptosis can be a host defense mechanism as a result of premature lysis of infected cells that have not finished replicating their genome. This leads to a loss or suppression of virus multiplication in the host organism (21). This hypothesis, however, seems to be increasingly challanged by the fact that animal viruses have developed two strategies for overcoming premature apoptosis or apoptosis itself: rapid multiplication and an antiapoptosis gene. Through rapid multiplication, viruses are able to completely replicate their genome before apoptosis takes place (21). RNA viruses are particularly good at rapid multiplication, completing their life cycle before apoptosis-mediated suppression is effective. From an evolutionary standpoint, it makes sense that RNA viruses have developed the means for rapid multiplication before apoptosis because they have a relatively small number of genes compared to DNA viruses so it would be more difficult to acquire another gene (21). DNA viruses on the other hand, have an antiapoptosis gene that prevents apoptosis in infected cells. Examples of such DNA viruses with an antiapoptosis gene would be poxviruses, herpesviruses and

adenoviruses. HIV has an antiapoptosis gene even though it is an RNA virus. HIV is a retrovirus that has a "long incubation period (for reverse transcription, integration into host chromosome and expression of virus genome with fine regulations), [so] it seems easy and reasonable to have an antiapoptosis gene instead of rapid multiplication of virus" (21).

However, apoptosis is not always used a defense mechanism of the cell. Viruses have the ability directly induce apoptosis for their advantage. The first virus found to induce apoptosis directly was an adenovirus mutant (E1B-19K). It was found that the adenovirus mutant induces extensive apoptosis in infected cells and that the adenovirus contains an antiapoptosis gene. Over the years, many RNA and DNA viruses have been discovered to induce apoptosis in eukaryotic cells (21). DNA viruses induce apoptosis differently than RNA viruses. Most large DNA viruses only induce apoptosis when their genome acquires a mutation and they lose a particular function of a gene while most RNA viruses induce apoptosis under conditions that allow sufficient virus multiplication (21).

There are viruses that are able to selectively induce apoptosis in cancer cells while leaving most normal cells untouched. Tumor necrosis factor-related apoptosis-inducing ligand (TRAIL), or Apo 2 ligant (Apo2L), is able to do so through an extrinsic pathway (see Figure 5). It is a member of a tumor necrosis factor (TNF) family that is able to induce apoptosis through the binding of its ligand to a death receptor. It also induces apoptosis independently of the p53 tumor-suppressor gene, which is important since this pathway is mutated in approximately half of all cancer (23). Thus transformed cell can be told to "shut down" even if they can no longer regulate the cell cycle. The chicken anemia virus (CAV) is different in that it does itself infect cancer cells. However, intriguing evidence has shown that its Apoptin protein is able to selectively kill cancer cells via an intrinsic pathway independent of the p53 tumor-tumor suppressor gene and cannot be blocked by Bcl-2 (1,24).

In order to determine how exactly Apoptin induces apoptosis of transformed cells, further study needs to be done on the roles of the rate of cell division, shuttling, and multimerization. It has been

shown that nuclear localization in tumor cells is required for apoptosis. It has also been shown that in order for apoptosis to take place, Apoptin needs to interact with the APC1 and does so most efficiently when multimerized (2). The goal of this project was to discover how nucleocytoplasmic shuttling and multimerization are related, and what roles they play in selectively inducing apoptosis in cancer cells. This was done through a more expansive probing of the NES with point mutations (I40A, T43A, L44A, L46A) to try to uncouple nuclear export and multimerization. A senior thesis project done by DeConti and Medeiros (2005) illustrated that they had retained multimerization by inducing a point mutation in the NES (I37A) (25). This project also conducted additional testing on ApI37A to discover if nuclear export was retained or lost.

Materials and Methods

Site Directed Mutagenesis

Site directed mutagenesis was accomplished using the "megaprimer" approach. Forward (F) and reverse (R) primers of ~40nt in length were designed to be complementary to Apoptin, except for the desired base changes, which were located in the center of the primer. N and C terminal fragments of Apoptin were generated by PCR using forward and reverse primers for Apoptin. Primers ApF and SDMR generated the N-terminal fragment, while ApR and SDMF generated the C-terminal fragment. N and C-terminal fragments (the "megaprimers") were gel extracted and combined in an approximate stoichiometric ratio of 1:1. A second PCR was done with ApF and ApR primers, and 1:10 serial dilutions of the megaprimers. The resulting mutated full length Apoptin from the PCR with the most dilute megaprimers was used in subsequent cloning. All other cloning was accomplished with standard techniques.

SDM Primers:

Ap F	5' atgaacgctctccaagaagatactccaccc 3'
Ap R	5' ttacagtcttatacgccttttgcggttcggg 3'
I37A F	5' cagagagatccggattggtgccgctggaattacaatcactc 3'
I37A R	5' gagtgattgtaattccagcggcaccaatccggatctctctg 3'
I40A F	5' gatccggattggtatcgctggagccacaatcactctatcgctgtg 3'
I40A R	5' cacagcgatagagtgattgtggctccagcgataccaatccggatc 3'
T43A F	5' gtatcgctggaattacaatcgccctatcgctgtgtggctg 3'
T43A R	5' cagccacacagcgatagggcgattgtaattccagcgatac 3'
L44A F	5' gctggaattacaatcactgcctcgctgtgtggctgcgcgaatgc 3'

L44A R	5’ gcattcgcgcagccacacagcgaggcagtgattgtaattccagc 3’
L46A F	5’ ggaattacaatcactctatcggcctgtggctgcgcgaatgctcg 3’
L46A R	5’ cgagcattcgcgcagccacaggccgatagagtgattgtaattcc 3’

Yeast 2-Hybrid Assay

The yeast 2-hybrid assay was carried out in Y190 yeast cotransformed with the appropriate pACT2 and pGBK plasmids and selected for on SD –tryptophan –leucine plates, as described elsewhere (26). Colonies were lifted on VWR Blotting paper 703, freeze fractured in liquid nitrogen and β-gal was detected using X-gal in Z-buffer as described elsewhere (26). The assay was incubated overnight at 30C before imaging.

Fluorescence Microscopy

H1299 cells were grown on cover slips for 24 hrs and transfected using Effectene from Qiagen per manufacturer’s instructions. The plasmid pEGFP-C1 contained truncations of the Apoptin mutants so as to lack the NLS. (Truncation primer: Ap244R) 24 hrs later, cells were washed with PBS and fixed with 2% paraformaldehyde in PBS on an orbital shaker for 15 minutes. The cells were again washed with PBS and mounted onto slides with a drop of mounting media (50% glycerol, 100 mM Tris pH 7.5, 2% DABCO and DAPI). Cells were imaged by fluorescence microscopy.

Results

Site-Directed Mutagenesis of Apoptin NES

In order to uncouple multimerization from nuclear export, point mutations were induced at the most conserved residues of the canonical NES of Apoptin. Inducing mutations of a canonical multimerization sequence was not possible because there is no known canonical multimerization sequence. Sequence alignments of other various canonical viral and mammalian NESs with Apoptin NES were done to illustrate the most conserved residues that likely play a key role in nuclear export (Figure 6, section (A)). ApL44 and ApL46 are the most conserved residues among all of the selected NESs of comparison. ApT43 of is also conserved among HTLV-1 Rex NES, MDM2 NES, p63 NES, and p73 NES except that their residues are all serine. Since threonine and serine are both very similar amino acids in that their side chains are uncharged and polar they are essentially interchangeable amino acids. ApI40 is also highly conserved among all of the NESs. As is the case with ApT43, ApI40 is interchangeable with the leucine in the same position of all the selected NESs (except for MDM2 NES) because both amino acids have large non-polar side chains and behave similarily. ApI37 appears to be less conserved when compared to the conservation of the other canonical amino acids; however, there is still conservation between ApI37and HIV-1 Rev NES, p63 NES, p73 NES and p53 NES.

In this study four canonical residues of Apoptin's NES, isoleucine 40, threonine 43, leucine 44 and leucine 46, were all individually mutated to alanine through PCR site-directed mutagenesis. The NES of Apoptin has been shown to interact with the major exportin Crm1. Previous studies in which Crm1 dependent canonical NES residues were mutated to alanine, a small neutral amino acid, and showed no interaction with Crm1 therefore alanine was the chosen amino acid to mutate the canonical residues to. Also, in addition to L44 and L46 being highly conserved residues of NESs, previous studies have shown that an Apoptin double mutant, ApL44A-ApL46A, results in the loss of nuclear export and

multimerization. This reveals that L44 and L46 are likely key players in Crm1 mediated nuclear export and multimerization. Isoleucine 37 was previously mutated to alanine and was additionally used in this study. I40, T43, L44 and L46, along with I37 were the five amino acid residues most conserved that would most likely have an effect on nuclear export and multimerization when mutations were induced at these residues (Figure 6, section (B)). Once ApI40A, ApT43A, ApL44A and ApL46A were created, DNA sequencing results verify successful point mutations and homology throughout the rest of the Apoptin gene for all four Apoptin mutants (Figure 6, section (C)).

Secondary Structure Predictions

As virtually no structural data is available for Apoptin, two secondary structure prediction programs (JPRED and SSPRO) were used to get a crude ideas of any secondary structures present in Apoptin. Both programs agreed that Apoptin is largely devoid of secondary structure, except for an extended peptide near the N-terminus and two α-helices at the C-terminus (Figure 7) . Perhaps the general lack of secondary structure could explain the difficult in crystallizing Apoptin, as well as the inconclusive results of the circular dichroism. Significantly, the extended peptide entirely covers the NES and extends several upstream residues. To check that the prediction programs could correctly determine the secondary structure of known NES, the programs were asked to predict the secondary structures of SNUPN and p53, both of which have crystal structures of their NES. Both programs correctly predicted the SNUPN and p53 NESs to be in an α-helix, in aggrement with the crystal structures (Figure 7). While these controls do not prove that JPRED and SSPRO are 100% accurate in their secondary structure predictions of an NES, it does significantly boost the confidence of their predictions.

There is less correlation between secondary structure and the NLS. Apoptin's NLS is known to be bipartite, comprised of two sites of containing three adjacent basic residues. The most C-terminal basic site overlaps partially at the end of the most C-terminal predicted α-helix, while the other basic site and other predicted α-helix do no align at all (data not shown). While there is poor correlation

between Apoptin's NLS and the predicted α-helices, it is significant the predicted extended peptide in a largly structureless protein correlates extremely well with Apoptin's NES. While these results are not experimental, it is tempting to speculate that there might be a reason why the extended peptide overlaps with the NES, which is also found in Apoptin's multimerization domain. A more complete consideration of this point, and the relation between the NES and multimerization, can be found in the discussion section.

Y190 Yeast 2-hybrid Assay

The yeast 2-hybrid assay assessed the retention of multimerization through a double transformation of Y190 yeast with pGBKT7 and pACT2 plasmid vectors containing Apoptin and/or the Apoptin mutants. All five mutants, ApI37A, ApI40A, ApT43A, ApL44A, and ApL46A were ligated into pGBKT7, a plasmid vector that contains the TRP1 gene and a DNA binding domain (DBD), half of the necessary transcription factor to turn on the Gal4 promoter. ApWT was previously ligated into pACT2, the second plasmid vector used in the yeast 2-hybrid system. pACT2 contains the second half of the Gal4 transcription factor, the transactivating domain (TAD) that binds to RNA polymerase II for transcription initiation, and also the LEU2 gene. The Gal4 promoter is also under control of *LacZ* which when transcribed, produces β-galactosidase whose artificial substrate when cleaved yields galactose and 5-bromo-4-chloro-3-hydroxyindole, a product that stains colonies blue. The DBD of the pGBKT7 vector and the TAD of the pACT2 vector when brought into close proximity to one another by a protein-protein interaction, will recruit RNAPII to the Gal4 promoter and transcribe the *LacZ* (Figure 8, section (B)). Y190 yeast were transformed with both pGBKT7 and pACT2 vectors and plated onto SD –Leu/-Trp plates to confirm that the yeast indeed were transformed with both vectors. The transformed yeast that took up both vectors would express the LEU2 gene from the pACT2 vector and the TRP1 gene from pGBKT7 and be able to grow on the plates without leucine or tryptophan. Untransformed Y190 did not grow on the

SD–Leu/-Trp plates verifying that Y190 yeast require both plasmids to express leucine and tryptophan(Figure 8, section (A), panel 2B).

For the second half of the yeast 2-hybrid assay, we lifted the colonies off each SD-Leu/-Trp plate with 3mm whatman paper, flash fractured them in liquid nitrogen, and soaked the fracture colonies in X-gal to induce B-galactosidase activity if present. The positive control, pACT2-ApWT /pGBKT7-ApWT, showed β-galactosidase acitivy (Figure 8, section (A), panel 1A). Wild-type Apoptin was used as the positive controls for this assay because it is well known that Apoptin extensively multimerizes with itself. The negative control –pACT2/-pGBKT7, showed no β-galactosidase activity (Figure 8, section (A), panel 1B). This negative control illustrated no β-galactosidase activity due to the missing inserts in both of the vectors so the protein-protein interaction was unable to occur to allow the DBD and TAD to turn on the Gal4 promoter for transcription. All five of the mutants illustrate β-galactosidase activity and as a result they all have retained the ability to multimerize (Figure 8, section (A), panels 2A, 3A/B, and 4A/B)

Subcellular Localization of Apoptin mutants.

Fluorescence imaging of mutants Apoptins fused to GFP revealed that point mutation of conserved residues in the NES change subcellular localization. As compared to GFP alone, GFP-ApT43A, GFP-ApL44A and GFP-ApL46A showed a clear cytoplasmic localization consistent with a functioning NES (Figure 9). The cytoplasmic localization is consistent with the localization of GFP-Apwt as reported previously (2). This data indicates that T43, L44 and L46 alone are not essential for nuclear export. It is interesting to note, however, that the double mutant L44AL46A has been shown to abrogate nuclear export. Either L44 alone or L46 alone is required for nuclear export, but the presence of both is not required.

Mutant ApI37A showed a strong reduction of nuclear export, adopting a diffuse localization (Figure 9). The partitioning between the nucleus and the cytoplasm was more equal than even the GFP control. The GFP control shows a slight bias towards a nuclear localization. This difference could indicate

very weak NES in the I37A mutant. However, the significant result is that mutation of a single conserved hydrophobic residue in the NES can almost completely knock out nuclear export of Apoptin. Because Ap-I37A was shown to still multimerize in the yeast 2-hybrid assay, these results show that it is possible to uncouple nuclear export from multimerization in Apoptin despite the fact that they have overlapping dowmains.

Mutant ApI40A shows partial loss of nuclear export. In some cells, GFP-ApI40A adopted a diffuse localization, identical to the localization of GFP-ApI37A. In other cells, nuclear export was only partially disrupted, with significant amounts of GFP-ApI40A observed in the nucleus (Figure 9). The levels of GFP-ApI40A in the nucleus were noticeably higher than for mutants GFP-ApT43A, GFP-ApL44A or GFP-ApL46A (Figure 9). It is possible that the observed reduction in NES activity and the cell to cell variability in localization are controlled by two different mechanisms. That is, one mechanism controls *if* GFP-ApI40A is excluded from the nucleus and another mechanism controls *how much* is excluded from the nucleus. However, it is more tempting to speculate that these two observations may be functionally related, with the activity of nuclear export existing along a continuous spectrum. However, the trivial explanation that CRM1 is differentially expressed at different stages of the cell cycle is unlikely, as it has been shown that CRM1 protein levels are constant throughout the cell cycle, although CRM1 mRNA levels do rise and fall with the cell cycle (27). Perhaps an extra cellular factor modifies Apoptin to regulate its export in a cell-cycle dependent manner. This would be in line with the majority of CRM1 substrates, which are differentially exported because of their role in cell cycle control (27).

Discussion

Apoptin is a protein from the Chicken Anemia Virus that has the unique ability to induce apoptosis in transformed and cancerous cells, but not primary cells. Apoptin triggers apoptosis through a p53-independent pathway, which is significant since roughly half of all cancers are insensitive to p53-mediated apoptosis, owing to a mutated pathway. A dazzling array of mutations in any one of a sea of regulatory proteins can cause cancer. The study of Apoptin is important because the protein is somehow able to distinguish between a wide variety of transformed and primary cells, recognizing some common biochemical difference. Not only can it specifically recognize cancerous cells, but it can kill them in a manner that it not dependent on the specific mutation that caused the cancer. Thus Apoptin's killing abilities are specific enough to leave primary cells unharmed, but broad enough to recognize cancers caused by different genetic mutations.

Two significant characteristics of Apoptin have provided clues as to Apoptin's selection mechanism. First, Apoptin shuttles between the nucleus and cytoplasm in both primary and transformed cells using an N-terminal NES and C-terminal NLS. However, in primary cells, Apoptin partitions into the nucleus and in transformed and cancerous cells, Apoptin partitions into the nucleus. Previous studies suggest differential activity of the NES is responsible for the characteristic partitioning (2). Second, Apoptin forms large multimers *in vivo* and is able to shuttle between the nucleus and cytoplasm at least as a dimer. Significantly, the multimerization domain overlaps with the NES, opening the possibility that multimerization and nuclear export are functionally related. In line with this idea, prior to this study all mutations that disrupted nuclear export also disrupted multimerization, suggesting nuclear export might depend on multimerization. However, this study mutated a greater number of conserved hydrophobic residues in the NES and found that two of the mutants, I37A and I40A, lost nuclear export while retaining multimerization.

Generating Mutants and Assessing Multimerization

Apoptin contains a canonical NES that overlaps with the domain for multimerization and this study illustrates that these activities can definitively be uncoupled. Through sequence alignments of viral and mammalian NESs with the NES of Apoptin, we were able to identify the most conserved residues that were likely to affect recognition by Crm1 in an effort to uncouple nuclear export from multimerization. The most conserved residues, mostly hydrophobic, were I37, I40, T43, L44, and L46. It was hypothesized that a mutation of one these key residues would lead to the loss of nuclear export and/or multimerization. A previous study showed that the replacement of the NES of Apoptin with another NES lead to the loss of multimerization. The same study also illustrated that a double mutation of L44 and L46 to alanines lead to the loss of both nuclear export and multimerization. This result left the interesting question of if single mutations at L44 and L46 could lead to the retention or loss of nuclear export and/or multimerization.

A yeast 2-hybrid assay was first done in this study to assess the protein-protein interaction of the five Apoptin mutant NESs with ApWT. Each mutant was fused to the DNA biding domain (DBD) of GAL4 and ApWT was fused to the trans-activiating domain (TAD) of GAL4; when the DBD and TAD are brought into close proximity through protein-protein interactions, the Gal4 promoter is turned on, expressing the reporter protein β-galactosidase. After freeze fracturing the yeast in liquid nitrogen, the β-galactosidase can be easily detected, as it the breakdown of X-gal into galactose and the blue dye 5-bromo-4-chloro-3-hydroxyindole. Results of this assay showed that all five mutants had β-galactosidase activity, indicating the mutant Apoptins retained multimerization. This indicates that single point mutations at these key hydrophobic residues does not disrupt the domain for multimerization. This could be because either A) point mutations do not disrupt the secondary structure of the domain for multimerization so the interaction surface between Apoptin molecules is still intact or B) the mutated residues do not participate directly in the Apoptin-Apoptin interactions.

Subcellular Localization of Apoptin Mutants.

Next we assessed the effect each mutation had on nuclear export H1299 cells, a non-small-cell-lung-carcinoma line . We truncated Apoptin to remove the NLS, fused it to the C-terminus of GFP and transformed each of the tagged truncation mutants into H1299 cells. Fluorescence imaging of each mutant Apoptin fused to GFP revealed that point mutations of conserved residues in the NES change subcellular localization. As compared to GFP alone, GFP-ApT43A, GFP-ApL44A and GFP-ApL46A showed a clear cytoplasmic localization consistent with a functioning NES. The cytoplasmic localization is consistent with the localization of GFP-Apwt as reported previously (2). This data indicates that T43, L44 and L46 alone are not essential for nuclear export. It is interesting to note, however, that the double mutant L44AL46A has been shown to abrogate nuclear export. Either L44 alone or L46 alone is required for nuclear export, but the presence of both is not required.

Mutant ApI37A showed a strong reduction of nuclear export, adopting a diffuse localization. This significant result shows that mutation of a single conserved hydrophobic residue in the NES can completely knock out nuclear export of Apoptin. Because Ap-I37A was shown to still multimerize in the yeast 2-hybrid assay, these results show that it is possible to uncouple nuclear export from multimerization in Apoptin, despite the fact that they have overlapping domains.

Mutant ApI40A shows partial loss of nuclear export. In some cells, GFP-ApI40A adopted a diffuse localization, identical to the localization of GFP-ApI37A. In other cells, nuclear export was only partially disrupted, with significant amounts of GFP-ApI40A observed in the nucleus. Significantly, the levels of GFP-ApI40A in the nucleus were noticeably higher than for mutants GFP-ApT43A, GFP-ApL44A or GFP-ApL46A. One possibility is that Ap-I40A partitions differently depending on the stage in the cell cycle. This would be in line with the majority of CRM1 substrates, which are differentially exported because of their role in cell cycle control (27). If this is true, the trivial explanation that CRM1 is differentially expressed at different stages of the cell cycle is unlikely, as it has been shown that CRM1 protein levels

are constant throughout the cell cycle, although CRM1 mRNA levels do rise and fall with the cell cycle (27). Perhaps an extra cellular factor modifies/interacts with Apoptin to regulate its export in a cell-cycle dependent manner. However, post translation modifications (PTM) seem difficult to rationalize, given that no PTM targets isoleucine and that no other mutants show variable subcellular partitioning.

Apoptin's NES: N-terminus versus C-terminus

The consensus sequence for NESs has been proposed by numerous sources to be $\Phi(X)_{2-3}\Phi(X)_{2-3}\Phi X\Phi$, where Φ represents hydrophobic amino acids (Lys, Val, Ile, Phe, or Met) and X is any amino acid (12,13). Apoptin's NES is IRIGIAGITITLSL and residues I37 through L46 (IAGITITLSL) fit the canonical sequence, while N-terminal residues I33 through G36 (IRIG) do not fit the canonical sequence, even though those residues have been experimentally validated as part of the NES. Sequence alignments of various known NESs with Apoptin's NES illustrates that the most conserved residues are the C-terminal hydrophobic residues L44 and L46 of Apoptin (ΦXΦ with respect to the canonical NES) (Figure 6, section (A)). One would expect that these C-terminal residues, L44 and L46, are the most important residues of NESs since they are most conserved and when double mutants are produced at these residues, nuclear export is lost. That would make the C-terminal end of NESs more important for NES recognition by CRM1. Contrary to this analysis, our data suggests that the less conserved N-terminus of Apoptin's NES (ApI37A and ApI40A) is more important for CRM1 binding, while the more conserved C-terminus (ApL44A and ApL46A) is less important for CRM1 binding. Since it took a double mutant to knock out nuclear export at the C-terminus (ApL44A-L46A) and single mutants more towards the N-terminus (I37A and I40A), it appears that the C-terminus is more tolerant to mutations while the N-terminus is more sensitive to mutations. Nonetheless, our data shows the degree of conservation of a residue between Apoptin's NES and the canonical NES does not correlate well a residue's importance in nuclear export. This results could hint at the possibility that Apoptin's NES binds CRM1 in a manner different than the

canonical leucine rich NES (LR-NES). In line with this idea are the results of secondary structure predition programs, which predict an α-helix in the canonical SNUPN NES, but predict an exended peptide in Apoptin's NES.

Single mutations at residues ApL44 and ApL46 each showed the retention of export, illustrating that CRM1 still has the ability to recognize Apoptin's NES. As previously stated, double mutants at these residues completely abolishes nuclear export. One explanation for the difference in this data between the double and single mutants is that perhaps CRM1 recognition of an NES requires a general hydrophobic residue in the area of extreme C-terminal end of an NES. By having either ApL44 or ApL46 present CRM1 is still able to make contact with one of the hydrophobic residues at the C-terminus of Apoptin and recognize it as an NES. Another explanation for the difference in nuclear export between the double mutants of ApL44A-ApL46A and single mutants of the same residues is that perhaps knocking out both residues in a double mutant changes the way the NES folds and therefore inhibits the interaction between CRM1 and Apoptin's NES. The double mutant NES could fold in such a manner as to completely abolish any interaction with CRM1 at the C-terminal end. Such a change in secondary structure might also be responsible for ablating multimerization in L44AL46A double mutant. On the other hand, the single mutants ApL44 and ApL46A retained nuclear export, which could be because the single mutants may not have an effect on the secondary structure of the NES; the secondary structure could still be in its normal form in which CRM1 most efficiently binds the NES. It is frustrating to leave this at speculation, but all previous attempts to determine Apoptin's crystal structure have failed and circular dichroism revealed no secondary structure.

Apoptin's canonical NES and CRM1: Sequence and Structure

Using the nuclear export data obtained in this study, the canonical NES of Apoptin, and information about the canonical NES of SNUPN (who's NES is also recognized by CRM1) we have deduced the key residues in Apoptin's NES needed for CRM1 mediated nuclear export. We also

postulate a possible secondary structure for Apoptin's NES that could explain Apoptin's ability to shuttle as a dimer.

The exportin CRM1 binds LR-NES. The following summary of CRM1 binding is based on the recent cocrystalization of CRM1 with SNUPN (a canonical LR-NES containing protein)(12). It is the only cocrystalization of CRM1 with an NES to be produced thus far. The interaction surface of CRM1 that binds to the NES is a hydrophobic groove formed by 3 helices. The hydrophobic groove starts out wide and then gets progressively narrower until it is just a narrow channel. The widest part of the hydrophobic groove binds the N-terminal region of the NES which is in an α-helix, while the narrowest part of the hydrophobic channel binds the C-termial region of the NES, which has been squeezed into an extended peptide conformation. Analysis of crystallographic structures LR-NESs show that LR-NESs typically contain hydrophobic residues spaced 2-3 amino acids apart that fit the canonical NES, are rich in acidic amino acids which make contact with flanking basic residues on CRM1, and are very flexible themselves or are near very flexible regions of the protein. A large percentage of NESs tend to be in an α- helix conformation, although there are a small percentage of NESs in an extended peptide conformation such as a B-sheet or coil. The hydrophobic amino acids align all on the same side of the helix that contacts CRM1. LR-NES are also enriched so serines, which could help solvate the back of the NES.

Snurportin 1 (SNUPN) is a protein who's LR NES is also recognized by Crm1. SNUPNs NES is located at the extreme N-terminus of the protein and consists of residues 1-16 (MEELSQALASSFSVSQ). The N-terminal region of the NES is in an α- helix conformation (MEELSQALASS) while the extreme C-terminal end (FSVSQ) is part of a very long extended peptide/loop structure. The loop structure at the C-terminus of the NES makes the NES itself a very flexible structure. Like Apoptin's NES, only a portion of the NES fits the canonical NES. Residues LSQALASSFSV are part of the canonical NES, while the extreme N-terminal NES residues, MEE, and extreme C-terminal residues, SQ, are not part of the canonical NES,

but are still part of the entire NES. Residues M1, L4, L8, F14 and V16 of SNUPN's NES interact directly with a hydrophobic groove of CRM1. Acidic side chains (E2 and E3) of SNUPN's NES help to interact with basic residues (L560 and L522) that flank the hydrophobic groove of CRM1. SNUPN additionally contains a second NES epitope consisting of mainly basic residues that help to stabilize the main NES with an acidic patch on CRM1. The N-terminal α- helix of SNUPN's NES binds to the wider area of CRM1's hydrophobic pocket and the extended loop of the C-terminal end of SNUPN's NES fits in the narrow channel after the bottleneck of CRM1s NES binding pocket.

As illustrated by mutants ApI37A and ApI40A who both lost nuclear export, both hydrophobic residues ApI37 and ApI40 play a role in CRM1 recognition of Apoptin's NES. Double mutant ApL44A-ApL46A lost nuclear export in another study (2), while in this study single mutants of L44 and L46 retained nuclear export, which indicates that the C-terminal end of Apoptin's NES does play a role in CRM1 recognition of Apoptin's NES. Apoptin's canonical NES is IAGITITLSL, and corresponds to residues 37 through 46. However there are more NES residues at the extreme N-terminal end of the entire NES and within that region, there are 2 hydrophobic residues, I33 and I35 that could potentially interact with CRM1. SNUPN's residues M1, L4, L8, F14, and V16 interact directly with CRM1; from our experimental data and from data previously done in another study, one could deduce that Apoptin's residues I37, I40, L44 and/or L46 directly interact with CRM1. One could also speculate that either I33 or I35, residues of the N-terminus of Apoptin's NES that do not match the canonical NES, could interact with CRM1 since SNUPN's M1, which is not a part of the canonical NES, also interacts with CRM1 (Figure 10).

Apoptin's amino acid sequence was run through a secondary structure prediction program and was predicted to have an extended peptide conformation in the NES (CITE). As secondary structure prediction programs are not near 100% accurate, we put SNUPN and p53 through the same prediction program as a control. SNUPN was predicted to have an α-helix at the N-terminal, but not C-terminal end of its NES, which matches with the crystal structure. Also p53 was correctly predicted to have an α-helix

throughout its NES, which also matches solved crystal structures (28) (Figure 7). As found by Leliveld et al. (2003) through circular dichroism, Apoptin is largely devoid of helical structures (8), so it is possible that Apoptin's NES is in a different conformation than most of its LR-NES counterparters that are in an α-helix conformation. Although the NESs of SNUPN and Apoptin may have different secondary structures, it seems safe to assume that all proteins that are exported by CRM1 must interact with CRM1 in a very similar manner. Key hydrophobic residues in the NES need to interact with CRM1's hydrophobic residues for export. As of now, there is no reason to believe that the correct positioning of NES hydrophobic residues could only be accomplished by an α-helix. It's possible that Apoptin does adopt an extended peptide conformation since it is believed that NESs in a non α-helix conformation could still be recognized by CRM1 due to the wideness of CRM1's N-terminal NES interaction surface. Apoptin could adopt an extended peptide conformation such as a B-sheet for the N-terminal region of the NES (residues I33 through T43), and an extended peptide (either in a coil or a B-sheet conformation) at the extreme C-terminus of the NES (L44 through L46) to fit within CRM1's narrow hydrophobic channel that binds the C-terminal ends of NESs. In addition, it seems almost mandatory that the hydrophobic residues of an NES that interact with CRM1 be on one side of its secondary structure in order to interact with CRM1's hydrophobic interface. Indeed, the key hydrophobic residues of SNUPN and p53's NES are on one side of the α-helix and are then buried within CRM1's hydrophobic groove at the CRM1-SNUPN interface. In sum, although the secondary structure was predicted to be different than SNUPN, it is speculated that Apoptin could potentially adopt an extended peptide conformation with key hydrophobic residues on one side of the extended peptide available for interaction with CRM1.

Noncanonical Aspects of Apoptin's NES

Apoptin's NES also differs from the canonical NES with regards to its non-hydrophobic residues. Alignments of experimentally validated NESs have revealed an enrichment of aspartic acid, glutamic acids and serine residues in the NES. Dong et al. (2009) have speculated that the acidic residues enhance

NES binding by being complementary to basic residues flanking CRM1's NES binding site. The serine residues, they speculated, would cover the back of the NES α-helix to help solvate it. Analysis of Apoptin's NES shows no aspartic or glutamic acid residues, either directly in the sequence or in the surrounding amino acids, suggesting that it does not use CRM1's basic residues to aid its binding. This would make sense if Apoptin's NES adapted an extended peptide conformation, as predicted by different secondary structure prediction programs. Assuming that the hydrophobic residues are all facing downward to contact the CRM1 NES binding site, an extended peptide conformation would orient the intervening residues facing up. Being more exposed to solvent than to the flanking side chains of CRM1, the orientation would prohibit the interaction of any acidic residues with the flanking basic residues. Conversely, when an NES adopts the more usual α-helix conformation, residues can be positioned such that they jut out to the side and make easier contact with nearby CRM1 residues. Thus the lack of aspartic acid and glutamic acid residues could be explained by the predicted extended peptide conformation being unable to position the acid residues to productively contact CRM1's basic residues.

Another amino acid that is enriched in NES's is serine. Based on the crystal structure of SNUPN binding to CRM1, Dong et al. (2009) suggested that polar residues along the solvent exposed side of the NES α-helix would help to solvate it (12). In agreement with this idea, an enrichment in serine residues has been identified at positions within NESs (13). Taken together, these observations suggest that at least some NESs function as a flexible, independent tag that is not significantly bound up in the tertiary structure of the protein. If the α-helix of the NESs were a face of a globular protein, there would be no need to coat the back side with hydrophilic residues. While Apoptin's NES has only one serine residue, it does have two nearby threonine residues. Again assuming Apoptin's NES adopts the predicted extended peptide conformation, the three polar residues would be aligned to the same side, helping to solvate

the NES. The presence of the polar residues could be an indication that Apoptin's NES functions as in independent motif, spatially separated from the main Apoptin fold.

Thus, while Apoptin's NES likely does not have an α-helix, a putative extended peptide conformation could allow for similar residue-positioning strategies. It would allow the hydrophobic residues to align downward, fitting into the hydrophobic pockets of CRM1, while allowing the hydrophilic residues to align upward, solvating the back of the NES. A limitation of the extended peptide conformation as compared to an α-helix conformation is that it would not allow for acidic residues to be positioned sideways to interact with CRM1 basic residues. Consistent with this idea, no acidic residues are found in or near Apoptin's NES.

In addition to the LR NES of SNUPN there is also an NES region rich in basic residues in SNUPN that binds with an acidic patch on CRM1 near the hydrophobic NES binding site and is used to help stabilize and strengthen the weak SNUPN-CRM1 interface. Interestingly, there are an abundance of basic residues which make up the bipartile NLS of Apoptin that is located near the extreme C-terminus of the protein. Dong et al. (2009) speculate that since the proximity of the LR NES binding site and the acidic patch are so close on CRM1, that many CRM1 substrates may contain, in addition to their LR-NES, basic NES regions (12). It is complete speculation, but perhaps when CRM1 binds the NES of Apoptin, it induces a conformational change in Apoptin to essentially swing over the NLS to bind with the acidic patch on CRM1 to help stabilize the CRM1-ApNES interface. This could help explain why Apoptin is believed to have little regular secondary structure since it has yet to be crystallized and circular dichroism has shown no secondary structure. NESs and NLSs are required to be flexible portions of proteins for easy access to exportins and importins. Since Apoptin is such a short protein (121 amino acids), and the NES is located at the N-terminal end and the NLS is located at the C-terminal end, it begs the question of how much regular structure can the protein have if the ends needs to be flexible and mobile.

Intermediary Export of I40A

Site directed mutagenesis of Apoptin's NES resulted in a spectrum of NES activity. Mutations ApT43A, ApL44A and ApL46A showed complete retention of export, while mutation ApI37A showed a complete loss of transport. Intermediary between these two extremes is mutation ApI40A, which in some H1299 cells shows complete loss of export, while in other H1299 cells, it shows varying amounts of nuclear export. These results show that nuclear export of Apoptin isn't strictly on or off, and that in H1299 cells the same mutant can adopt different sub-cellular localizations. This observation may hint at more complex regulatory mechanism that mediate Apoptin export beyond simply binding to CRM1. Although more data is certainly needed to tease out the regulatory details, we will speculate on some possibilities here.

One attractive possibility is that Apoptin's export is somehow linked to the cell cycle. Cells in culture are at different stages in the cell cycle, and so could exhibit different Apoptin localization, if its export were coupled to the stage in the cell cycle. Indeed, a major class of proteins with NES are involved in cell cycle control. These proteins obviously need to be exported only at certain times in the cell cycle, and so it seems possible that Apoptin could be regulated by some of the same cell-cycle dependent mechanisms. But if the variability in localization of ApI40A is due to the stage of the cell cycle, why then is no variability observed in wild type Apoptin or any of the other mutants? This could be explained if a cell cycle dependent helper protein could assist Apoptin in binding to CRM1. The necessity of this helper protein would depend on the strength with which Apoptin's NES binds to CRM1. In the wild type Apoptin or in mutations ApT43A, ApL44A or ApL46A, the NES binding to CRM1 would be strong, and would be stable regardless of the presence of a helper protein. Thus wild type, T43A, L44A and L46A Apoptin would always be exported. Conversely, the I37A mutation could be so destabilizing that it could not bind CRM1 even in the presence of the helper protein. In the case of I40A, the mutation could have an intermediary effect on CRM1 binding. The mutations could be disruptive enough such

that Apoptin doesn't bind Apoptin on its own, but still has enough affinity for CRM1 that it can bind with the assistance of the helper protein. Thus Apoptin would only be exported at the stages in the cell cycle during which the helper protein is expressed.

Alternatively, the mutations in Apoptin affect how well it competes for CRM1 binding. Apoptin mutants T43A, L44A and L46A could have equal CRM1 affinity as does wild type Apoptin. These mutants would always strongly bind CRM1 and be exported. In contrast, ApI40A could have intermediary affinity for CRM1, and could be outcompeted for CRM1 binding by other cell-cycle dependent proteins. Thus a diffuse ApI40A could correlate to times of high CRM1 dependent export in which CRM1 is saturated and so preferentially exports strong binders. Similarly, nuclear exclusion of ApI40A could correlate to times of low nuclear export during which CRM1 can export low affinity binders. Clearly, further experiments are needed to provide data to distinguish among these possibilities. One particularly useful experiment would be to transfect a synchronous population of H1299. If ApI40A localization is still variable, then the difference could not be correlated to the stage of the cell cycle and is likely due to other cell to cell variablity. However if ApI40A localizes similarly in the synchronous cells, then it would point to a dependence on the stage of the cell cycle. It would be particularly convincing if the cells could be caught at the stage in the cell cycle where ApI40A is diffusely localized, since this is normally a relatively minor localization of ApI40A.

Post Translational Modification of Apoptin's NES

It is unlikely that mutated residues that caused a change in Apoptin localization are the targets of differential post-translational modifications (PTM), which could influence localization. Both of the mutated residues that cause a change in nuclear export were isoleucines, which are not subject to post-translational modification. It is possible that mutations I37A and I40A could distort the secondary structure, which could inhibit binding to the enzyme that adds the PTM. However, this is entirely

speculative and it is simpler to think that alteration in the secondary structure would directly inhibit CRM1 binding, rather than inhibiting binding of an enzyme that adds a PTM that prevents CRM1 binding.

Relationship Between the NES and Multimerization

The most significant result of this study is that multimerization and nuclear export of Apoptin are separable functions. Our findings are in contrast to previous studies in which all mutations that eliminated nuclear export also eliminated multimerization. Taken together, the data shows that nuclear export and multimerization are functionally related, and that only certain mutations are able to separate them. The simplest explanation for nuclear export and multimerization being related is that they both rely on the same amino acid residues to function. For example, the hydrophobic residues of p53's NES align along one side of an α-helix and form a hydrophobic strip that participates in the tetramerization domain interactions. Thus mutating a conserved hydrophobic residue would not only prevent CRM1 binding, but could also reduce or eliminate the affinities of the tetramerization domain. However, secondary structure prediction programs predict Apoptin's NES to be in an extended peptide conformation, not an α-helix. Significantly, the extended peptide of the NES is the only secondary structure predicted in the multimerization domain. The lack of any predicted secondary structure elsewhere in the multimerization domain makes it even more tempting to speculate that the multimerization activity maps to the extended peptide of the NES, as it is difficult to see how a proteins could multimerize at an interface without a regular secondary structure.

One speculative idea is that the extended peptide of the NES contributes to multimerization in a manner similar to amyloid plaque formation (Figure 11). Nelson et al. showed short extended peptides can stack in the standard Pauling-Corey parallel arrangement and form fibrils that show the essential features of amyloid plaques (29). The fibril is composed of two sheets that are held together by tightly interlocking amino acid residues, with each sheet being composed of a parallel stacking of extended peptides held together vertically by hydrogen bonding. The two sheets are arranged around a 2_1 screw

axis, so that they are rotated 180 degrees around the central axis, and staggered one half the length of the vertical repeating unit. A similar quaternary structure could be formed by the extended peptide of the NES. Most significantly, it would explain the Apoptin fibrils observed in the cytoplasm but not the nucleus under immunofluoresense. GFP fused to the N-terminus of Apoptin are known to instead form punctate aggregates in the cytosol, possibly a consequence of GFP interfering with the stability of large stacks of the extended peptides. The amyloid-like formation would also explain the fact that the appearance of Apoptin foci only occurs after a certain period of time and are only observed in the cytoplasm. Nelson et al. proposed that the rate limiting step in amyloid formation is the nucleation step, which requires high protein levels to occur (29). Only after a certain period of time would Apoptin's concentration be high enough for productive nucleation. Additionally, active export from the nucleus would reduce Apoptin's concentration below the threshold needed for nucleation. This mechanism is in contrast to the possibility that Apoptin associates with existing fibrils in the cytosol, like actin filiments, which are present in the cytoplasm but not the nucleus. If Apoptin did associate directly with existing filiments, it would not explain the time it takes for Apoptin foci to form. An immunofluorscence time-course assay would be useful to distinguish the two possibilities. If Apoptin filaments are found from the beginning of expression, when Apoptin's concentration is low, then Apoptin most likely associates with existing cytosolic filaments. However if the filaments only appear after a certain amount of time, it would point to an amyloid-like Apoptin filament. Most significantly though, the possibility that an extended peptide could pull double duty in both the NES and multimerization domains could explain why mutations in the NES also disrupt multimerization. A mutation in one of the hydrophobic residues needed for CRM1 binding could also destabilize the interactions between the extended peptides forming the amyloid-like fiber. However, given the preliminary data at this point, it is difficult to rationalize a structural basis as to why some mutations (ApI37A and ApI40A) eliminate nuclear export and not mutlimerization, while other mutations (ApL44AL46A) eliminate both nuclear export and

multimerization. While it is certain that each hydrophobic residues contributes differently to nuclear export or multimerization, at this point we are unable to say *why* each residues contributes differently. Perhaps the mutation of certain hydrophic resides disrupts the secondary strucuture, leading to loss of both export and multimerization, while other mutations only remove the residues needed for CRM1 binding, while leaving the secondary structure unchanged.

Having the multimerizaton domain overlap with the NES domain could potentially result in an antagonistic relationship between multimerization and nuclear export if the NES is buried in the aggregated state. This is similar to the functioning on p53's NES, which is buried in the protein-protein interface in the tetramer state. However a series of experiments have provided convincing evidence that Apoptin can shuttle between the nucleus and cytoplasm in at least a dimerized state, and potentially as larger aggregates. In the study, two differentially labeled mutant Apoptins were created which either lacked a functional NLS (GFP-Ap-pmNLS) or lacked a functional NES (dsRED-Ap-pmNES) (Heilman et al., 2006). As expected, GFP-Ap-pmNLS always localized to the nucleus and dsRED-Ap-pmNES always localized to the cytoplasm, regardless if they were expressed in cancerous (H1299) or primary (PFF) cells. However coexpression of the mutants resulted in wild type localization patterns, partitioning into the nucleus of cancer cells and partitioning into the cytoplasm of primary cells. Restoration of wild type partitioning is presumably due to a functional NES and NLS being brought together by Apoptin dimerizing or multimerizing. As stated above, mutlimerization of p53's would sandwhich the NES, masking its activity. Apoptin could not multimerize this way if its mutlimerization and NES map to the same residues. However if Apoptin multimerized similarly to amyloid fibers, a functional NES would be left exposed at the end of the stack, allowing for nuclear export of Apoptin aggregates.

Conclusions and Future Recommendations

Apoptin, a viral protein from the chicken anemia virus, has the unique ability to selectively induce apoptosis in transformed cells but not primary cells. Apoptin contains both an NLS and NES, which promote the nucleocytoplasmic shuttling which is necessary to induce apoptosis. Overlapping the NES is the multimerization domain, which facilitates Apoptin's aggregation into large, insoluble multimers. Because protein multimerization has the obvious potential of masking an NES, and because previous studies have shown that Apoptin can shuttle at least as a dimer, the question arose as to weather multimerization and nuclear export were functionally linked. All previous attempts to uncouple multimerization from nuclear export using site directed mutagenesis have failed. However this study has shown that mutlimerization and nuclear export can be uncoupled. Mutations I37A and I40A retain multimerization while loosing nuclear export.an obvious way of regulating NES activity via mas

It will be necessary in the future to test each of the mutants for their ability to induce apoptosis in transformed cell lines such as H1299 cells, especially with the Apoptin mutants I37A and I40A. Previous studies have shown that chemical inhibition of nuclear export nullifies Apoptin mediated apoptosis. Therefore, it would be expected that ApI37A will show no apoptotic activity and that ApI40A will show cell to cell variability in apoptosis, mirroring its cell to cell variability in nuclear export. It will also be necessary to assess the ability for each of the Apoptin mutants that retained nuclear export (T43A, L44A, L46A) to induce apoptosis to verify that these mutations do not affect Apoptin's ability to induce apoptosis even though nuclear export is functional.

Although this study mutated many of the hydrophobic residues in Apoptin's NES, several remain to be mutated. To assess the ability of additional hydrophobic residues of Apoptin's NES to interact with CRM1, mutants ApI33A, ApI35A, ApI42A should be generated. Additionally double mutants ApI40AI42A and ApI33AI35A could be generated to see if those amino acids interact with CRM1 like L44 and L46 appear to do. If double mutants ApI40AI42A and ApI33AI35A act like L44 and L46, then they should both

loose NES and multimerization activity, while their single mutants (ApI33A, ApI35A, ApI40A, ApI42A) should retain NES and multimerization (except ApI40A which has partial NES activity as verified by this study). In these regions of the NES, perhaps both hydrophobic residues are positioned similarly and so only one is needed to interact with CRM1. These same single and double mutants would also be useful to try to distinguish the exact domain for multimerization and to try to determine which amino acids are required for the multimerization interface. Is it possible to knock out multimerization with single mutants to disrupt the multimerization interface, or are double mutants required to so (perhaps by disrupting the secondary structure and disrupting the multimerization interface then)?

Apoptin mutants that eliminate nuclear export while retaining multimerization will be indispensible in future studies of Apoptin's mechanism of action. However, equally valuable would be mutants that retain nuclear export but lack multimerization. Potentially, mutations I33A and I35A could accomplish this. Residues I33 and I35 lie just upstream of the canonical NES but still within the predicted extended peptide. If Apoptin does form multimers similar to amalyoid fibrils, mutations I33 or I35 could disrupt the interactions between the extended peptides, without necessarily changing the position of the conserved hydrophobic residues in the downstream NES.

Apoptin's cancer killing ability, once revealed, could be utilized as a guide to future cancer therapy. Apoptin's mechanism of action is particularly valuable since it is able to kill a broad spectrum of cancer cells in a novel p53-independent manner, while leaving most normal cells unharmed. It is significant that Apoptin kills via a p53 independent pathway, as roughly half of all cancers have a mutated p53 and are insensitive to therapies targeting this cellular policing pathway. If recapitulated in a cancer therapeutic, Apoptin's broad yet specificity against cancer cells would be the holy grail of cancer treatment. The mutants generated in this study will aid in the future dissection of Apoptin's mechanism of action and selectivity. If understood well enough, Apoptin research could provide a roadmap to designing new powerful, broadly acting anticancer therapeutics.

Figures

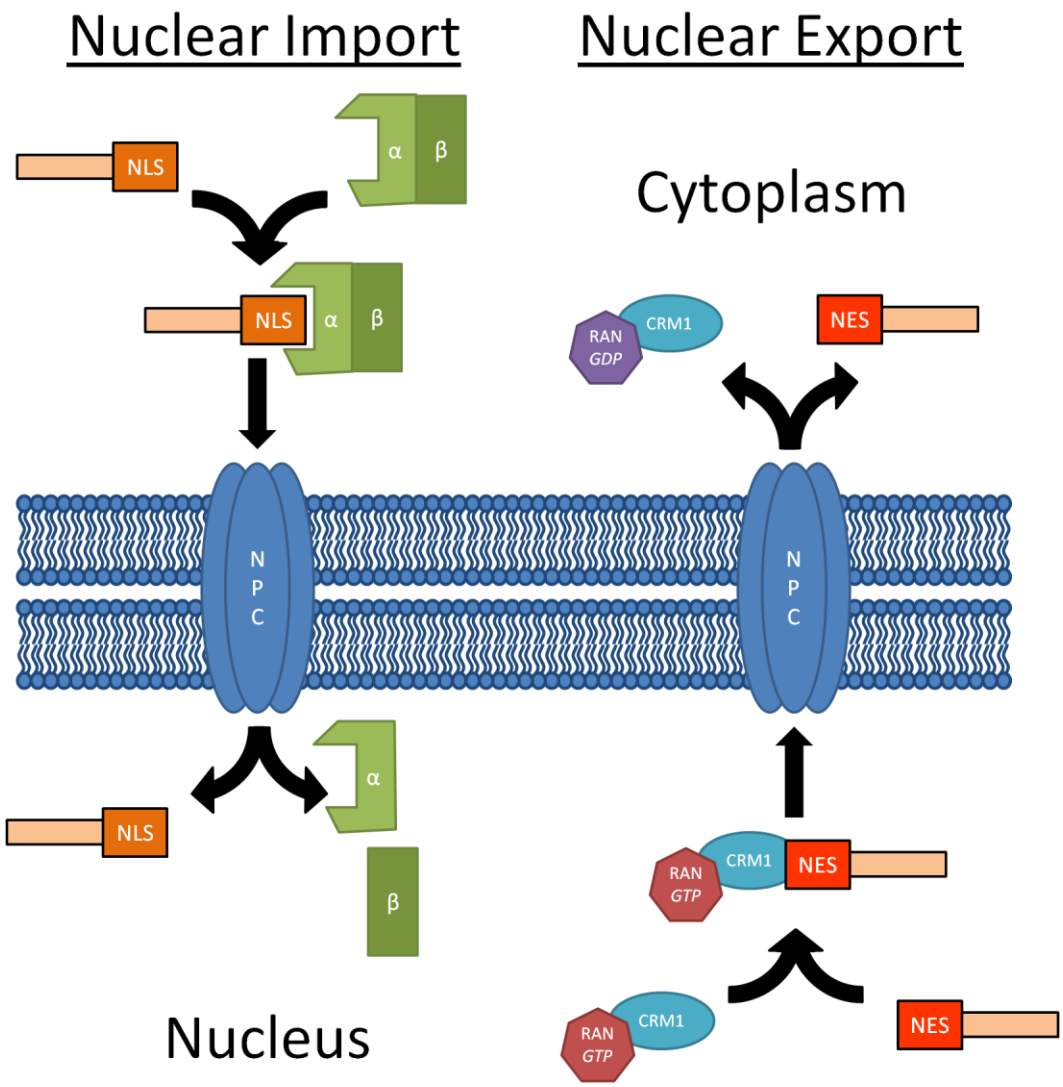

Figure 1 – Nuclear import and export of the RAN-GTP cycle through the nuclear pore complex (NPC). (Key: NLS – Nuclear Localization Sequence. NES – Nuclear Export Sequence. Green proteins: Importin α and β, respectively. NPC – Nuclear Pore Complex.) (11)

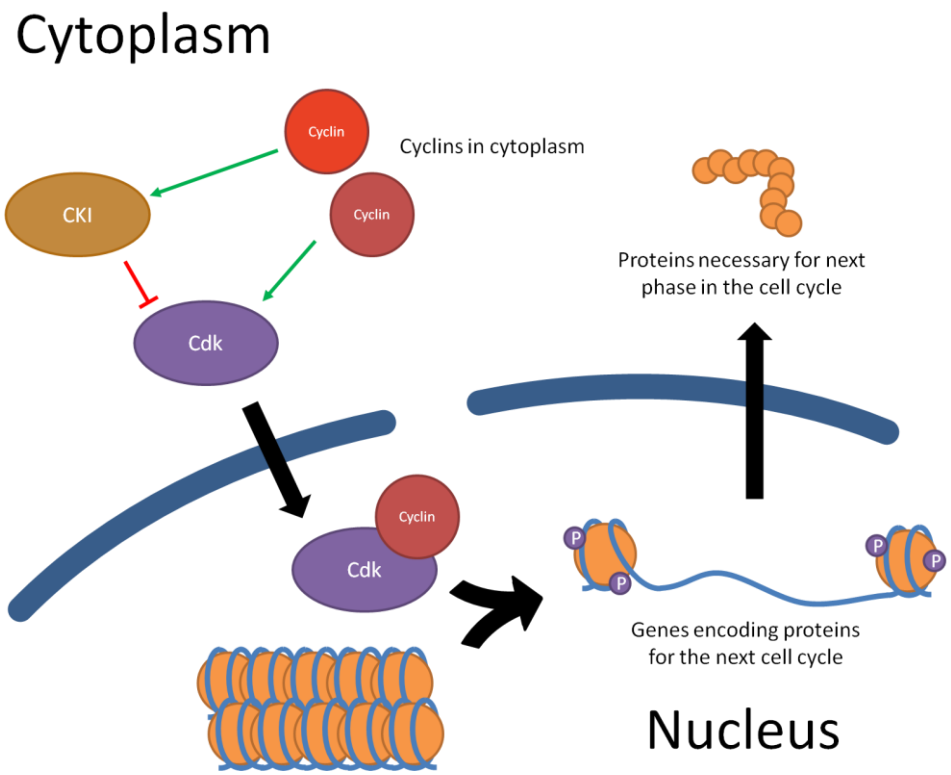

Figure 2 – Regulation and effects of Cdk. Upon binding to the proper cyclin, the Cdk migrates to the nucleus where it phosphorylates the tails of histones surrounding particular genes. The DNA unwinds, allowing transcriptional machinery access to the gene. Once expressed, the gene products promote the next phase of the cell cycle. Green arrows represent activation, red blunt arrows represent inhibition, and black arrows represent physical associations or movement.

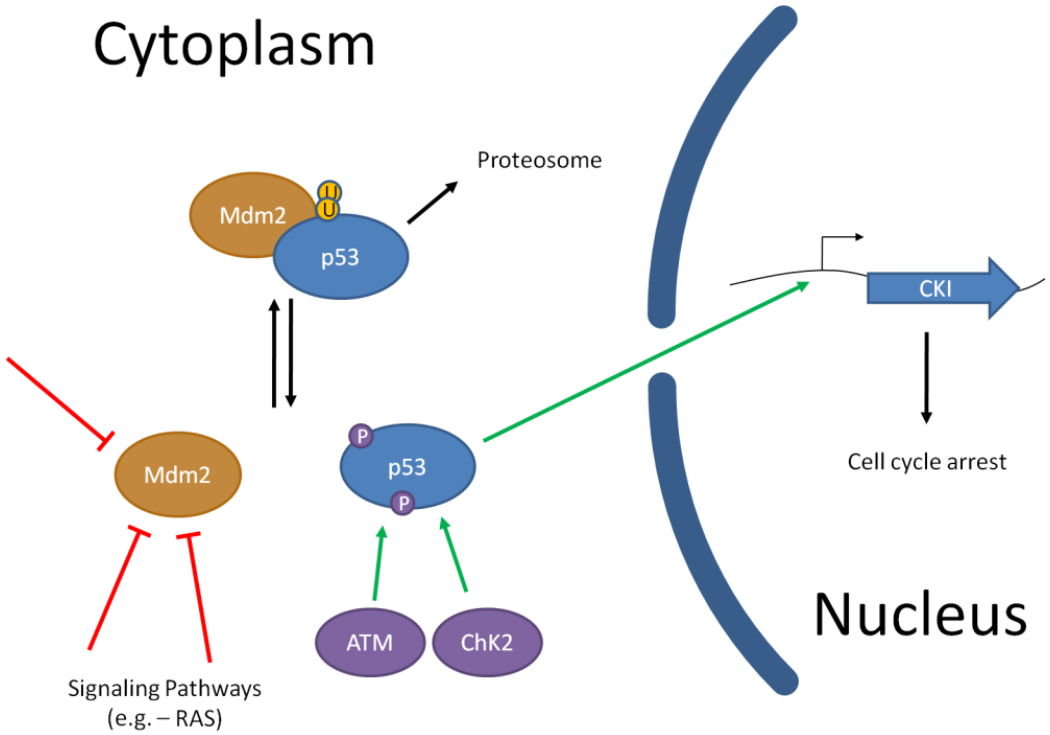

Figure 3 – The p53/Mdm2 anticancer pathway. Mdm2 normally antagonizes p53 through ubiquitylation. This process is blocked if kinases (e.g. ATM or Chk2) phosphorylate p53 or if Mdm2 is inhibited by antiproliferitive pathways (e.g. RAS). p53 promotes expression of CKIs, which arrest the cell cycle and can eventually induce apoptosis. Green arrows represent activation, red blunt arrows represent inhibition, and black arrows represent physical associations or movement (17).

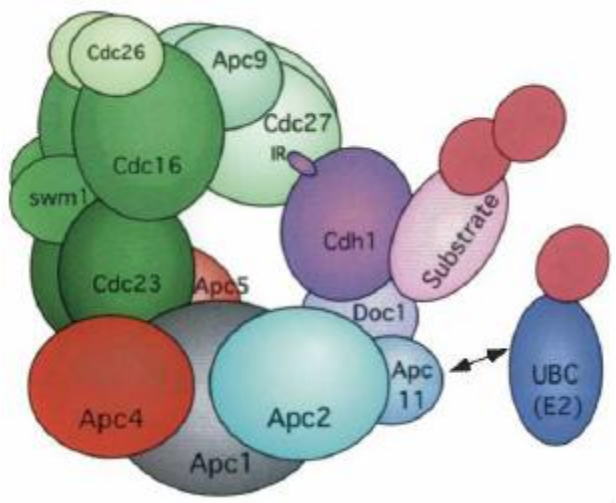

Figure 4 – Structural arrangement of the subunits of the APC/C as proposed by Thornton et al. (20).

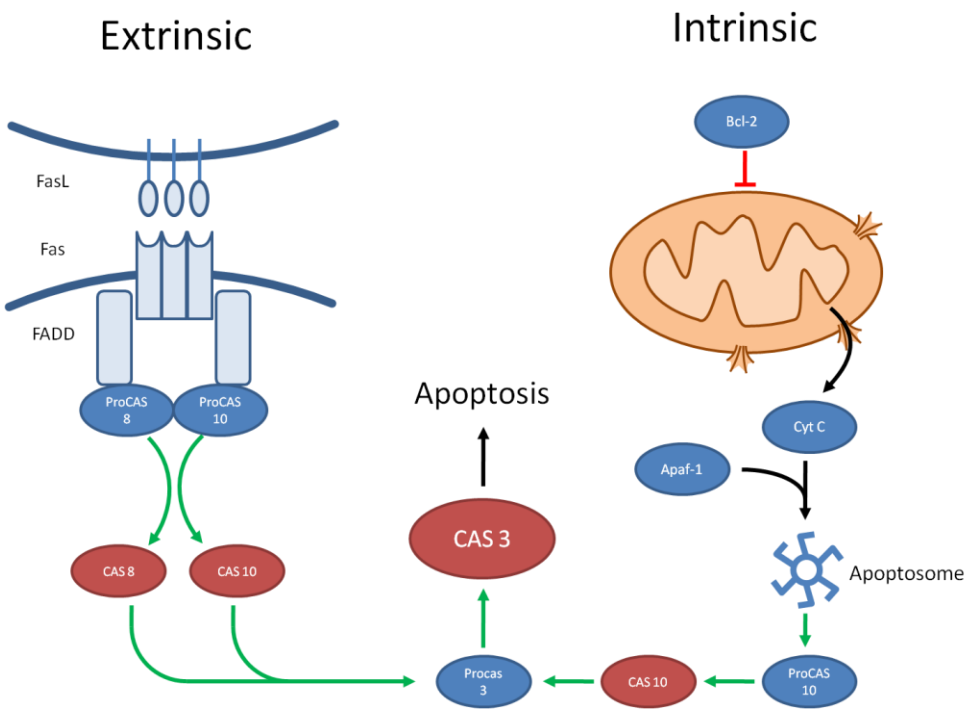

Figure 5 – Extrinsic and intrinsic apoptotic pathways. Green arrows denote activation, while black arrows denote physical associations or translocations (22).

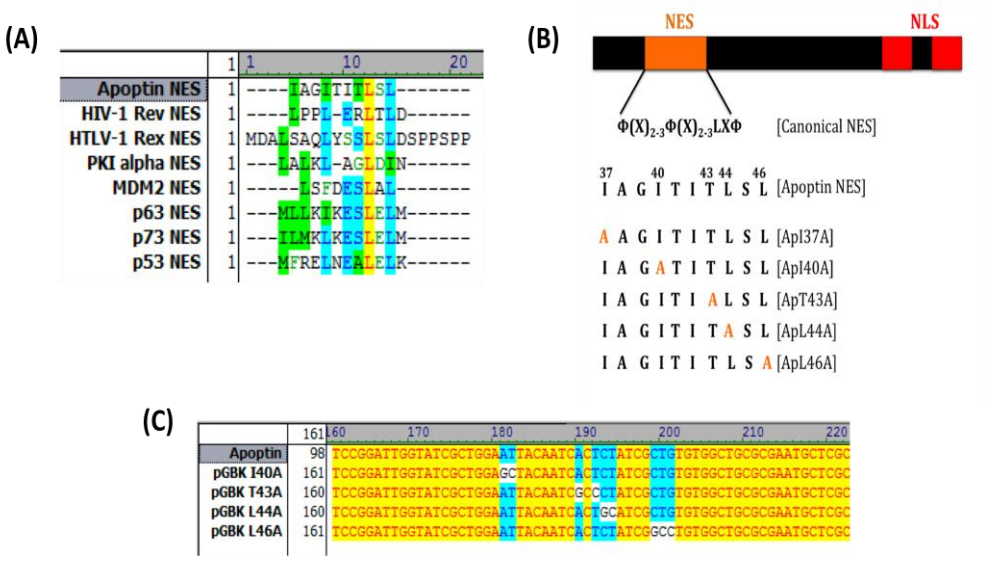

Figure 6 – NES Alignments. (A) The NES of Apoptin is shown aligned with other known canonical NES to illustrate conserved residues. (B) A graphic representation of each mutation induced in the canonical NES of Apoptin. The NES is located at the N-terminus of Apoptin while the bipartile NLS is located at the C-terminus of Apoptin. (C) Sequencing results of mutants pGBKT7-ApI40A, pGBKT7-ApT43A, pGBKT7-ApL44A, and pGBKT7-ApL46A show successful mutations when compared to WT Apoptin. ApI40A was mutated from ATT to GCT, ApT43A was mutated from ACT to GCC, ApL44A was mutated from CTA to GCA and ApL46A was mutated from CTG to GCC. All of the mutations code for alanine. Note: Each megaprimer designed for the site-directed mutagenesis had all four the original bases changed to GCC; however, alanine is an amino acid whose wobble position of its codon can be any of the four bases so the fact that ApI40A and ApL44A did not take at the wobble position does not matter.

Figure 7 – Secondary structure predictions of NESs from Apoptin, SNUPN and p53. The top line is the amino acid sequence of each NES, aligned to the highly conserved C-terminal hydrophobic residue. The two lines below each NES sequence are the secondary structure predictions from JPRED and SSPRO. For SNUPN and p53, JPRED and SSPRO predicted the correct secondary structure, as determined by crystallography or NMR, respectively (12,28). Key: E: Extended peptide, H: α helix, C/- : no prediction.

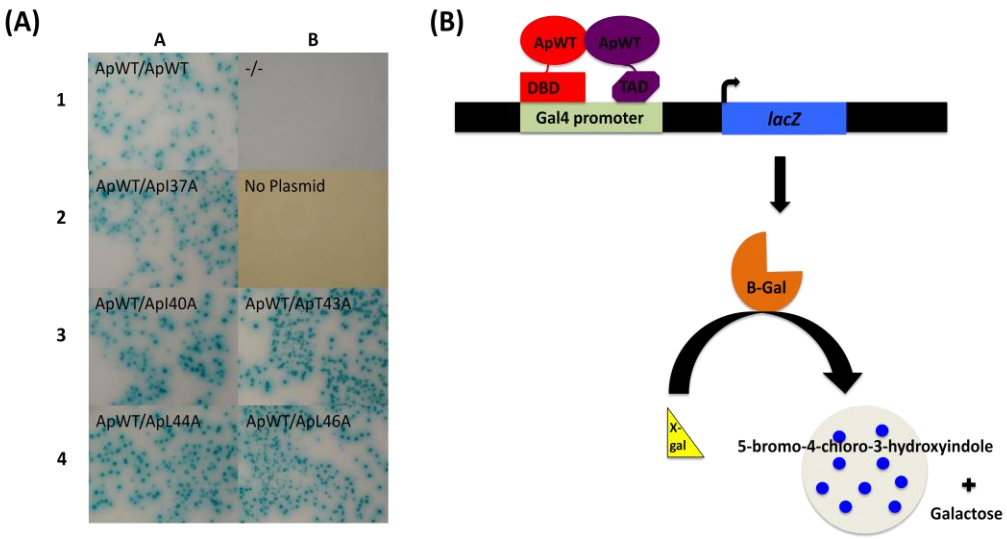

Figure 8 – All Apoptin mutants retain multimerization in a yeast 2-hybrid assay. (A) Y190 yeast were transformed with pACT2-ApWT/pGBKT7-ApWT (positive control, panel 1A), pACT2-no insert/pGBKT7-no insert (negative control, panel 1B), no plasmid (negative control, panel 2B), pACT2-ApWT/pGBKT7-ApI37A (panel 2A), pACT2-ApWT/pGBKT7-ApI40A (panel 3A), pACT2-ApWT/pGBKT7-T43A (panel 3B), pACT2-ApWT/pGBKT7-L44A (panel 4A) and pACT2-ApWT/pGBKT7-L46A (panel 4B). The yeast were lifted off the plates with 3mm Whatman paper, cracked open in liquid nitrogen, and soaked in X-gal for 20 hours at 30° C. (B) A graphic representation of when protein-protein interaction occurs between ApWT and Ap mutant (and/or ApWT). When the two halves of the transcription factor (DBD and TAD) are brought into close proximity, they turn on the Gal4 promoter thereby transcribing *lacZ* and producing β-galactosidase. When given X-gal, its cleaved by B-galactosidase into galactose and a blue colony stain, 5-bromo-4-chloro-3-hydroxyindole.

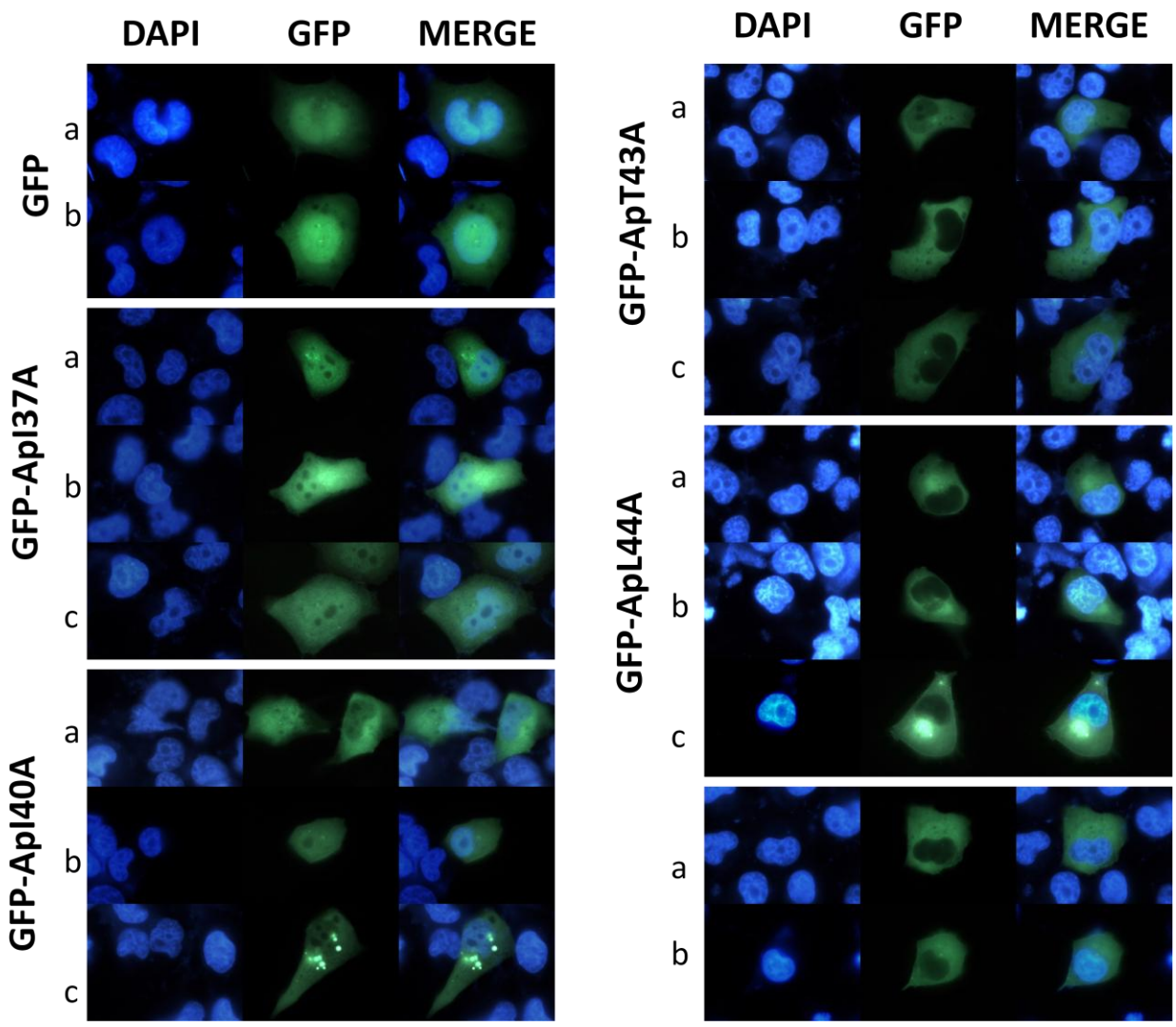

Figure 9 – Subcellular localization of mutated Apoptin NES. Constructs used are Apoptin with mutated NESs and are truncated to lack the NLS, and are fused to the C-terminus of GFP.

Figure 10 – Comparison of hydrophobic residues in SNUPN and Apoptin NES. This figure illustrates the known hydrophobic residues (in red) of SNUPN's NES that interact CRM1. Hydrophobic residues of Apoptin's NES likely to interact with CRM are shown in blue. It seems plausible, based on this alignment, that perhaps I33 of Apoptin also could interact with CRM1 like M1 of SNUPN.

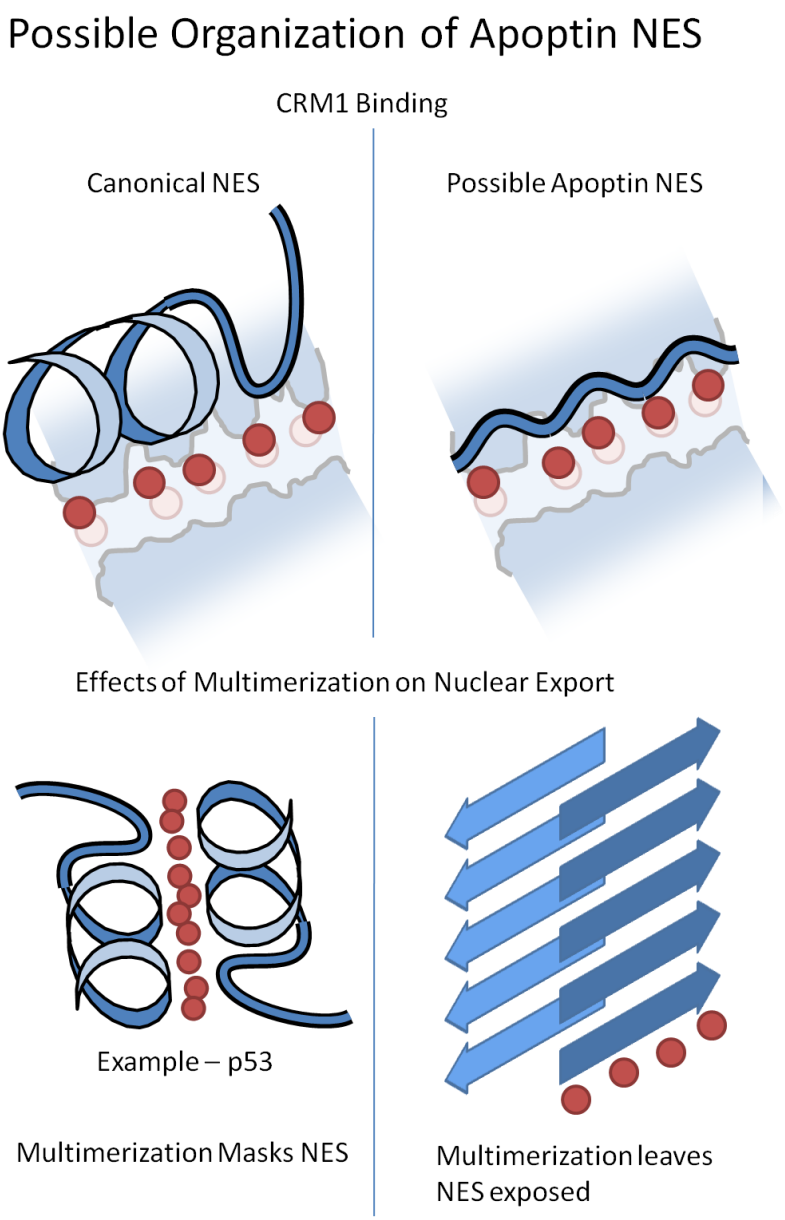

Figure 11 – Possible alternative secondary structures and multimerization mechanisms of Apoptin. (Left, top) Diagram of a canonical α-helix NES, like SNUPN, binding to CRM1 (12). (Left, bottom) Diagram showing how multimerization of α-helical NESs can mask them, as occurs during p53 tetramerization (28). (Right, top) Speculative Apoptin NES structure based on secondary structure prediction programs. (Right, bottom) Diagram showing how an extended peptide conformation of Apoptin's NES could allow for both multimerizations and nuclear export, if it multimerizes similar to amyloid fibrils (29).

References

1. Danen-Van Oorschot, A. A. A. M. *et al.*, Apoptin induces apoptosis in human transformed and malignant cells but not in normal cells. *Proc. Natl. Acad. Sci.* **94**, 5843-5847 (1997).

2. Heilman, D. W., Teodoro, J. G. & Green, M. R., Apoptin nucleocytoplasmic shuttling is required for cell type-specific localization, apoptosis, and recruitment of the anaphase-promoting complex/cyclosome to PML bodies. *Journal of Virology* **80** (15), 7535-7545 (2006).

3. Lee, M. S. *et al.*, Production of chicken anemia virus (CAV) VP1 and VP2 protein expressed by recombinant Escherichia coli. *Process Biochemistry* (2008).

4. Peters, M. A., Jackson, D. C., Crabb, B. S. & Browning, G. F., Chicken anemia virus VP2 is a novel dual specificiity protein phosphatase. *J. Biol. Chem.* **277** (42), 39566-39573 (2002).

5. He, C. Q. *et al.*, Indentification of chicken anemia virus putative intergenotype recombinants. *Virology* **366** (1), 1-7 (2007).

6. Danen-van Oorschot, A. A. A. *et al. J. Biol. Chem.* **278** (30), 27729-27736 (2003).

7. Zhuang, S. M. *et al.*, Apoptin, a portein derived from the chicken anemia virus, induces apoptosis p53-independent apoptosis in human osterosarcoma cells. *Cancer Research* **55**, 486-489 (1995).

8. Leliveld, S. R. *et al.*, Apoptin induces tumor-specific apoptosis as a globular multimer. *J. Biol. Chem* **278** (11), 9042-9051 (2003).

9. Wadia, J. S., Wagner, M. V., Ezhevsky, S. A. & Dowdy, S. F., Apoptin/VP3 contains a concentration-dependent nuclear localization signal (NLS), not a tumoriogenic selective NLS. *Journal of Virology* **78**, 6077-6078 (2004).

10. Poon, I. K. *et al.*, Apoptin nuclear accumulation is modulated by a CRM1-recongnized nuclear export signat that is active in normal but not tumor cells. *Cancer Res* **16** (65), 7059-7064 (2005).

11. Alvisi, G., Poon, I. K. H. & Jans, D. A., Tumor-specific nuclear targeting: Promises for anti-cancer therapy? *Drug Resistance Updates* **9**, 40-50 (2006).

12. Dong, X. *et al.*, Structural basis for leucine-rich nuclear export signal recongnition by CRM1. *Nature* (2009, In press).

13. la Cour, T. *et al.*, Analysis and prediction of leucine-rich nuclear export signals. *Protein Engineering, Design and Selection* **17** (6), 527-536 (2004).

14. Zhang, Y. H. *et al.*, Activation of the tumor-sepcific death effector apoptin and its kinase by an N-terminal determinant of simian virus 40 large T antigen. *Journal of Virology* **78** (15), 9965-9976 (2004).

15. Poon, I. K. H. *et al.*, A tumor cell-specific nuclear targeting signal within chicken anemia virus

VP3/Apoptin. *Journal of Virology* **79** (2), 1339-1341 (2005).

16. Tavassoli, M., Guelen, L., Luxon, A. & Gaken, J., Apoptin: Specific killer of tumor cells? *Apoptosis* **10**, 717-724 (2005).

17. Hickman, E. S., Moroni, M. C. & Helin, K., The role of p53 and pBR in apoptosis and cancer. *Current Opinion in Genetics & Development* (12), 60-66 (2002).

18. Coffey, M. C., Strong, J. E., Forsyth, P. A. & Lee, P. W. K., Retrovirus therapy of tumors with activated ras pathway. *Science* **282**, 1332-1334 (1998).

19. Harper, J. W., Burton, J. L. & Solomon, M. J., The anaphase promoting comples: it's not just for mitosis any more. *Genes & Development* (16), 2179-2206 (2002).

20. Thornton, B. R. & Toczyski, D. P., Precise destruction: an emergin picture of the APC. *Genes & Dev.* **20**, 3069-3078 (2006).

21. Koyama, A. H. *et al.*, Role of virus-induced apoptosis in a host defense mechanism against virus infection. *J. Med. Invest.* **45**, 37-45 (1998).

22. Jiang, X. & Wang, X., Cytochrome C-mediated apoptosis. *Annu. Rev. Biochem* (73), 87-106 (2004).

23. Wang, S. & El-Diery, W. S., TRAIL and apoptosis induction by TNF-family death receptors. *Oncogene* **22**, 8628-8633 (2003).

24. Teodoro, J. G., Heilman, D. W., Parker, A. E. & Green, M. R., The viral protein apoptin associates with the anaphase-promoting complex and induced G2/M arrest and apoptosis in the absence of p53. *Genes & Dev.* **18**, 1952-1957 (2004).

25. DeConti, D. & Medeiros, A., Uncoupling the multimerization and nuclear export activities of CAV VP3 via site-directed mutagensis. *Major Qualifying Project, Worcester Polytechnic Institute* (2008).

26. Clontech, Yeast Protocol Handbook **Protocol No. PT3024-1** (Version No. PR742227) (2008).

27. Kudo, N. *et al.*, Molecular cloning and cell cycle-dependent expression of mammalian CRM1, a protein involved in nuclear export of proteins. *The Journal of Biological Chemistry* **272** (47), 29742-29751 (1997).

28. Clore, G. M. *et al.*, High-resoltion structure of the oligomerization domain of p53 by multidimensional NMR. *Science* **265** (5170), 386-391 (1994).

29. Nelson, R. *et al.*, Structure of the cross-B spine of amyloid-like fibrils. *Nature* **435** (7043), 773-778 (2005).

30. Heilman, D. W., Green, M. R. & Teodoro, J. G., The anaphase promoting comples: a critical target for viral proteins and anti-cancer drugs. *Cell Cyle* **4** (4), 560-586 (2005).