Elucidating the role of ARKL1 in regulating Epstein-Barr virus (EBV) infection

by

Anup.S.Vaidya

A thesis submitted in conformity with the requirements for the degree of Master of Science
Molecular Genetics
University of Toronto

© Copyright by Anup.S.Vaidya 2016
Elucidating the role of ARKL1 in regulating Epstein-Barr virus (EBV) infection

Anup.S.Vaidya

Master of Science

Molecular Genetics
University of Toronto
2016

Abstract

ARKL1 and Arkadia were previously identified as proteins that bind to the CK2β KSSR motif. Since the Epstein-Barr virus (EBV) protein Epstein-Barr nuclear antigen 1 (EBNA1) is also known to bind to the same sequence, it was hypothesized that ARKL1 and Arkadia might play a role in regulating EBV infection. Silencing experiments demonstrated that both ARKL1 and Arkadia function to suppress EBV reactivation. ARKL1 was subsequently found to colocalize with the PRC1 protein CBX4 in distinct nuclear foci. This localization of ARKL1 to CBX4 foci was dependent upon its SIMs, as mutation of its SIMs resulted in diffused nuclear staining. In order to gain mechanistic insights into how ARKL1 regulates EBV infection, I carried out affinity purification-mass spectrometry (AP-MS) of ARKL1. ARKL1 was found to bind the transcriptional activators Jun and CREB1. This result raises the possibility that ARKL1 may function to recruit PRC1 to sites bound by Jun/CREB1, thereby repressing transcription.
Acknowledgements

First and foremost, I’d like to thank my supervisor Dr. Lori Frappier for her guidance, support and patience throughout my project. She has been a wonderful mentor who has enriched my perspective and broadened my understanding of science.

I’d also like to thank my supervisory committee members Dr. Anne-Claude Gingras and Dr. Frank Sicheri for their insights about my project. They’ve played an important role in giving this project the required direction for its betterment and ultimate success.

I also want to express my gratitude to past and present members of the Frappier lab, especially Kathy Shire and Anna Georges. Kathy has been a lab mother, and my transition into the scientific world wouldn’t have been smooth without her. Anna has been a wonderful friend throughout this captivating journey, helping me overcome the days of failed western blots with a smile. I’d also like to acknowledge the contributions of Dr. Sheila Mansouri and Dr. Jamie Baxter in helping me process data for my RT-PCR experiments and for answering my questions patiently. Finally, I’d like to thank Dr. Natasha Malik-Soni, Umama Siddiqui, Jaime Yockteng, Jasper Ho, Dr. Carlos De la Cruz Hererra, Dr. Vasu Gautam, Marios Mejdani, Chantal Magnan and Adithya Shankar for their support.

Last but not the least, I’d like to thank my parents and my sister, without whom, none of this would’ve been possible. I have no words to express my gratitude towards them for being a constant pillar of strength and encouragement.
Table of Contents

CHAPTER 1: INTRODUCTION ................................................................. 1

1.1 The Epstein-Barr Virus ..................................................................1

1.1.1 *EBV* genome and life cycle ......................................................1

1.1.2 Latent infection ...........................................................................2

1.1.3 Lytic infection ............................................................................5

1.1.4 Reactivation ................................................................................5

1.2 Epstein-Barr Nuclear Antigen 1 (EBNA1) ...........................................6

1.2.1 Functions on the *EBV* genome ..................................................7

1.2.2 Cellular effects of EBNA1 ...........................................................9

1.3 Casein Kinase 2 (CK2) ..................................................................11

1.3.1 CK2 enzymatic properties .........................................................11

1.3.2 CK2 subunits ............................................................................12

1.3.3 CK2β functions .........................................................................14

1.3.4 CK2 in viral infections ...............................................................15

1.3.5 EBNA1 binding to CK2 .............................................................16

1.4 ARKL1 ..........................................................................................16

1.5 Arkadia ..........................................................................................17

1.6 ThePolycomb Repressive Complex (PRC) ..........................................19

1.6.1 Epigenetic regulation of EBV infection .......................................20

1.7 Thesis Rationale ............................................................................21

CHAPTER 2: MATERIALS AND METHODS .............................................22
2.1 Cell Lines ................................................................. 22
2.2 Plasmids ........................................................................ 22
2.3 Determining ARKL1 Effects on SnoN ................................. 23
2.4 Assays for Lytic Reactivation of EBV upon ARKL1 or Arkadia Silencing ...................................................... 24
2.5 RNA Extraction and Quantitative PCR ................................. 24
2.6 Overexpression Experiments ............................................. 25
2.7 ARKL1 Localization to SUMO and CBX4 Foci ........................ 26
2.8 Immunofluorescence Microscopy ......................................... 26
2.9 ARKL1-EBNA1 Competition Experiment ............................... 27
2.10 ARKL1-SUMO3 Immunoprecipitation ................................ 27
2.11 Western Blotting .......................................................... 28
2.12 ARKL1 Affinity Purification Coupled with Mass Spectrometry (AP-MS) ...................................................... 28

CHAPTER 3: RESULTS .......................................................... 30
3.1 ARKL1 Silencing does not Affect SnoN Levels ...................... 30
3.2 ARKL1 or Arkadia Silencing Promotes Reactivation of EBV ................. 31
3.3 ARKL1 Overexpression Inhibits EBV Reactivation .................... 34
3.4 Effect of EBNA1 on the ARKL1-CK2β Interaction .................... 36
3.5 ARKL1 Localizes to CBX4 (PRC1) ........................................ 38
3.6 Identification of ARKL1 SIMs Important for Localization to CBX4 Foci ...................................................... 39
3.7 ARKL1 Cellular Interactors Identified by Affinity Purification-Mass Spectrometry (AP-MS) ......................... 43

CHAPTER 4: DISCUSSION AND FUTURE DIRECTIONS ..................... 47
4.1 Examining the Role of ARKL1 in SnoN Degradation ................ 47
4.2 Functional Significance of the ARKL1-CK2β Interaction ................ 47
4.3 ARKL1 Localization to CBX4 ........................................................................................................49
4.4 Identification of ARKL1 Cellular Interactors by AP-MS ..........................................................50
4.5 Future directions ................................................................................................................................55

REFERENCES ........................................................................................................................................60
CHAPTER 1: INTRODUCTION

1.1 The Epstein-Barr Virus

The Epstein-Barr virus (EBV) is a ubiquitous herpesvirus belonging to the gamma-1 subfamily of herpesviruses, lymphocryptoviridae (LCV). EBV was originally discovered in tumour biopsies of patients suffering from Burkitt’s lymphoma by Anthony Epstein, his graduate student Yvonne Barr and Bert Achong, who used electron microscopy to identify herpesvirus particles in these samples (Epstein, Achong, and Barr 1964). Since its discovery, EBV has been established as the causative agent of infectious mononucleosis (IM) or the ‘kissing disease’, transmitted through the saliva of infected individuals (Gerber et al. 1972, Paul and Bunnell 1982). Although EBV infection is largely asymptomatic, latent viral infection is known to be associated with several types of cancers including Burkitt’s lymphoma, Hodgkin’s lymphoma, post-transplant lymphoproliferative disease (PTLD), nasopharyngeal carcinoma and gastric carcinoma (Hopwood and Crawford 2000, Kuppers and Rajewsky 1998, zur Hausen et al. 2000, zur Hausen et al. 1970).

1.1.1 EBV genome and life cycle

EBV has a 184kb linear double stranded DNA genome, containing short and long unique sequences, flanked by tandem repeats. Upon infection, the terminal repeats join to form a circular episome which becomes rapidly incorporated with cellular histones (Given et al. 1979). The terminal repeats are highly variable and are signatures of clonal infection events (Kutok and Wang 2006).

Primary infection by EBV often occurs during childhood and is usually asymptomatic. When infection is delayed to adulthood, mononucleosis can result (Henle, Henle, and Diehl 1968). It is
thought that EBV initially infects the epithelial cells of the oropharynx after its transmission through the saliva. The virus subsequently gains access to the circulating B cells in the bloodstream where it establishes a life-long latent infection (Lemon et al. 1978). Entry into B cells is accomplished by interaction of viral envelope protein gp350/220 with the B cell surface molecule CD21 (Fingeroth et al. 1984). Following this interaction, CD21 aggregates on the surface and EBV is internalized (Carel et al. 1990). Infection of naïve B cells causes them to proliferate in germinal centers due to the combined action of 9 latency proteins. Most infected cells are eliminated by the host immune system. EBV subsequently downregulates expression of most of the latency proteins in order to escape immune surveillance (Laichalk et al. 2002). The viral episomes then persist in resting (latency 0) or proliferating B cells (latency I). Occasionally, EBV switches to the lytic form of infection (a process known as reactivation) where progeny EBV virions are made, enabling infection of new B cells or oral epithelial cells, thereby facilitating host to host spread of the virus through saliva (Kutok and Wang 2006, Middeldorp et al. 2003, Thorley-Lawson and Gross 2004).

1.1.2 Latent infection

There are four forms of latencies that have been identified, latency I, II and III that occur in proliferating cells, and latency 0 that occurs in resting B cells (Figure 1). Latency I, II and III involve expression of different EBV proteins whereas latency 0 is characterized by the expression of only EBV encoded non-coding RNAs (EBERs), which are also expressed in all the other forms of latency (Kuppers 2003). Infection of a naïve B cell results in latency III, which results in the proliferation and immortalization of naïve B cells into lymphoblastoid cell lines (LCLs) and is commonly seen in transplant patients with PTLD. Latency III involves expression of all six Epstein-Barr nuclear antigens (EBNA1, EBNA2, EBNA3A, EBNA3B, EBNA3C and
EBNA-LP) and three latent membrane proteins (LMP1, LMP2A and LMP2B) (Thorley-Lawson, Miyashita, and Khan 1996).

Most of the EBNA proteins are highly immunogenic and hence EBV downregulates expression of these proteins in subsequent latency programs. Latency II or the ‘default program’ is detected in nasopharyngeal carcinoma, gastric carcinoma and Hodgkin’s lymphoma. It typically involves the expression of EBNA1, LMP1, LMP2A and LMP2B, although LMP1 is not expressed in gastric carcinoma (Caldwell et al. 1998, Takada 2000). The next form of latency characterized by the expression of EBNA1 is known as latency I or the ‘EBNA1-only’ program. This form of latency is commonly observed in memory B cells of healthy individuals and in Burkitt’s lymphoma (Thorley-Lawson, Miyashita, and Khan 1996). Finally, as described earlier, latency 0 involves expression of only EBERs and is found in resting memory B cells (Thorley-Lawson 2001). In all forms of latency, the EBV DNA exists as a closed circular plasmid that is packaged in nuclesomal arrays (Given et al. 1979, Shaw, Levinger, and Carter 1979). EBV episomes are replicated once per cell cycle in S phase by the host DNA polymerase and partitioned equally amongst daughter cells. Replication is initiated at oriP and requires EBNA1 (discussed in the section on EBNA1) (Chaudhuri et al. 2001, Dhar et al. 2001).
Figure 1. The different types of EBV latency. Figure describing the different forms of latency in epithelial and B cells from Thorley-Lawson and Allday (2008).
1.1.3 Lytic infection

The spread of EBV from cell to cell within the host and to other hosts requires reactivation to the lytic form of infection. The lytic life cycle involves the ordered expression of approximately 80 EBV proteins including transcriptional transactivators, replication factors and structural proteins required for packaging viral DNA into capsids. The cascade of gene expression starts with BZLF1 (Zta) and BRLF1 (Rta), transcriptional activators, which bind multiple viral promoters to activate gene expression (Holley-Guthrie et al. 1990, Speck, Chatila, and Flemington 1997). BZLF1 also binds to the lytic origin of replication (oriLyt) to activate it (Fixman, Hayward, and Hayward 1992, Schepers et al. 1993). During lytic infection, replication is initiated at oriLyt and long concatemers of DNA attached head to tail are generated. These molecules are then cleaved into unit length genomes and packaged into virions (Tsurumi, Fujita, and Kudoh 2005, Kenney and Mertz 2014).

1.1.4 Reactivation

The physiological switch from latent to lytic infection in vivo is poorly understood, but in vitro reactivation can be elicited by certain chemical and biological agents including 12-O-tetradecanoylphorbol-13-acetate (TPA), sodium butyrate (NaB), calcium ionophore and immunoglobulin (Ig) (Amon and Farrell 2005). The induction of the viral lytic life cycle by these agents results in expression of the two viral immediate early proteins, BZLF1 and BRLF1, that then function to turn on expression of other lytic cycle genes (Speck, Chatila, and Flemington 1997, Tsurumi, Fujita, and Kudoh 2005).

BZLF1 is a transcriptional activator that shares similarity to the basic leucine zipper family of transcription factors (b-zip) and its expression alone is sufficient for EBV reactivation (Tsurumi, Fujita, and Kudoh 2005, Speck, Chatila, and Flemington 1997, Amon and Farrell 2005). The
BZLF1 promoter is tightly repressed under most conditions, however, it is activated in response to the agents listed above. This promoter is positively regulated by cellular transcription factors cAMP response element binding protein (CREB1) and proteins belonging to the AP-1 family, whereas its activity is negatively regulated by the Jun dimerization protein 2 (JDP2), Zinc finger E-box binding factor (ZEB) and BZLF1 SUMOylation (Liu, Liu, and Speck 1998, Murata, Isomura, et al. 2009, Murata et al. 2011). ZEB is known to bind specific sequences within the promoter for BZLF1, thereby preventing BZLF1 mediated transcriptional activation. The promoter for BZLF1 is also silenced by low levels of histone acetylation, CpG methylation and repressive epigenetic marks such as H3K27me3 and H4K20me3 (Murata et al. 2012). Following reactivation, these epigenetic modifications are reversed and the Zta promoter becomes active (Tempera and Lieberman 2014).

1.2 Epstein-Barr Nuclear Antigen 1 (EBNA1)

Epstein-Barr nuclear antigen 1 (EBNA1) is the only EBV protein required for stable maintenance of EBV episomes during latency. It is also the only EBV protein that is expressed in all forms of latency in proliferating cells. The organization of EBNA1 is shown in Figure 2 and includes 1) a central variable glycine-glycine-alanine (Gly-Ala) repeat region (amino acids 90-325), that allows EBNA1 to avoid immune detection by slowing translation (Apcher et al. 2010, Blake et al. 1997, Tellam et al. 2001); 2) two glycine-arginine (Gly-Arg) rich regions (amino acids 33-53 and 325-376), 3) a nuclear localizing sequence (NLS; amino acids 379-386), 4) the casein kinase 2 (CK2) binding site (387-394), 5) a USP7 binding site (amino acids 442-448) and 6) a DNA binding and dimerization domain (amino acids 459-607) (Frappier 2015). The significance of some of these regions shall be discussed very briefly along with EBNA1 functions below:
Figure 2. Schematic representation of the EBNA1 protein. Key structural elements are shown, including the Gly/Arg and Gly/Ala repeat regions, the CK2 binding site, the USP7 binding site and the flanking and core subcomponents of the DNA binding and dimerization domain. Adapted from (Frappier 2015).

1.2.1 Functions on the EBV genome

EBNA1 is required for the stable maintenance of EBV genomes in latency and of oriP (origin of latent replication) containing plasmids in human cells. OriP contains two major functional elements, the dyad symmetry element (DS) and the family of repeats (FR). The DS element is the origin of replication and contains four EBNA1 recognition sites. EBNA1 is involved in recruitment of the host origin recognition complex (ORC) to the DS element in EBV episomes. Within the FR element are 20 tandem copies of a 30bp region containing EBNA1 binding sites. Binding of EBNA1 to the FR element is required for mitotic segregation of EBV episomes and transcriptional activation (Reisman, Yates, and Sugden 1985, Rawlins et al. 1985).

EBV episomes are present at a low copy number in host cells and they are evenly partitioned into daughter cells during cell division in order to maintain a constant copy number. Partitioning (or segregation) involves EBNA1-mediated tethering of EBV episomes to cellular DNA during mitosis (Lupton and Levine 1985, Lee, Diamond, and Yates 1999, Krysan, Haase, and Calos

The EBNA1 regions involved in mitotic segregation, namely the N-terminal region (amino acids 65-83) and Gly-Arg region (amino acids 325-376) are also required for EBNA1’s functions in regulating transcription (Wu, Kapoor, and Frappier 2002). EBNA1 can activate transcription of EBNA and LMP latency genes from the viral Cp and LMP promoters by binding the FR element. EBNA1 also binds to the cellular protein Brd4 through amino acids 61-83 and recruits it to the FR (Lin et al. 2008). Brd4 is a cellular bromodomain containing protein that regulates transcription through its interaction with cellular histones (Dey et al. 2003, Jang et al. 2005). A previous study has shown that Brd4 is recruited by EBNA1 to the FR element (Lin et al. 2008). EBNA1 also recruits the histone chaperone proteins NAP1, TAF-1β and nucleophosmin to the FR element, and these interactions are important for EBNA1 to exert its effects on transcriptional activation (Holowaty, Zeghouf, et al. 2003, Wang and Frappier 2009, Malik-Soni and Frappier 2014, 2012). EBNA1 also recruits a complex of USP7 and GMP synthetase to the FR, which is required for deubiquitylation of histone H2B at FR, thereby facilitating transcriptional activation (Sarkari et al. 2009).

In addition to its role in transcriptional activation, previous studies showed that EBNA1 inhibits its own transcription from the Qp promoter which is used to express EBNA1 in the absence of other EBNA proteins. EBNA1 binds to recognition sequences downstream of Qp to repress its
transcription only when the FR and DS sites have been saturated by EBNA1 binding (Jones, Hayward, and Rawlins 1989, Ambinder et al. 1990, Sample, Henson, and Sample 1992). A later study demonstrated that EBNA1 does not inhibit transcription from Qp as previously believed, but actually impedes post- or cotranscriptional processing of primary transcripts (Yoshioka, Crum, and Sample 2008). This represents an additional level of regulation of its own levels through a negative feedback mechanism.

1.2.2 Cellular effects of EBNA1

EBNA1 is the only EBV protein found in all cancer types and expression of EBNA1 in EBV-negative cancer cells results in increased tumorigenicity of these cells (Humme et al. 2003). Moreover, silencing of EBNA1 in various EBV-infected cells arrests cell growth (Sheu et al. 1996, Kube et al. 1999). Based on these and other reports, an important role has been suggested for EBNA1 in EBV-induced oncogenesis, likely through its effects on altering the cellular landscape. For example, EBNA1 has been shown to increase STAT1 expression while attenuating TGFβ signalling and NF-κB activity (Wood et al. 2007, Valentine et al. 2010). EBNA1 has been implicated in increasing reactive oxygen species (ROS) and inducing an oxidative stress (Gruhne et al. 2009, Cao, Mansouri, and Frappier 2012). EBNA1 has been found to affect cellular micro RNAs that contribute to epithelial-to-mesenchymal transition while maintaining EBV latency (Mansouri et al. 2014). In addition, EBNA1 promotes tumorigenesis by inhibiting the ability of the cellular protein Nm23-H1 to suppress cell migration (Murakami et al. 2005). EBNA1 can also transactivate certain cellular genes including the genes coding for c-Jun, ATF and survivin (Wu, Kapoor, and Frappier 2002, O'Neil et al. 2008).

In addition, EBNA1 can regulate p53 levels through its effect on the cellular ubiquitin specific protease (USP7). EBNA1 interacts with USP7 through amino acids (442-448) (Holowaty, Sheng,
et al. 2003, Saridakis et al. 2005, Malik-Soni and Frappier 2012). USP7 is known to bind and regulate the levels of several cellular proteins including p53 and Mdm2, which acts as an E3 ubiquitin ligase for p53 (Li et al. 2002, Li et al. 2004). EBNA1 outcompetes these proteins for binding to the USP7 TRAF domain, resulting in lower p53 levels and increased cell survival after DNA damage (Holowaty, Zeghouf, et al. 2003, Saridakis et al. 2005).

Finally, our laboratory discovered that EBNA1 can regulate promyelocytic leukemia (PML) nuclear bodies (NBs) (Sivachandran, Sarkari, and Frappier 2008, Sivachandran, Cao, and Frappier 2010). PML NBs are nuclear foci formed on the basis of PML proteins, that regulate several cellular processes including DNA repair, cellular senescence, apoptosis and p53 activation (Pearson et al. 2000, Carbone et al. 2002, Bernardi and Pandolfi 2007). In addition to their cellular functions, PML NBs play an important role in the cellular antiviral response. As a result, several viruses encode proteins that either disrupt PML NBs or cause degradation of PML proteins (Everett 2001, Everett and Chelbi-Alix 2007). EBNA1 expression in NPC and gastric carcinomas results in reduced PML levels, and consequently reduced DNA damage repair, p53 activation and apoptosis (Sivachandran et al. 2012). The ability of EBNA1 to induce degradation of PML depends on its ability to bind to USP7 and the host CK2 kinase, both of which it recruits to PML NBs (Sivachandran, Sarkari, and Frappier 2008, Sivachandran, Cao, and Frappier 2010). CK2 was previously shown to phosphorylate PML on serine 517, which primes PML for polyubiquitylation and subsequent proteasomal degradation (Scaglioni et al. 2006). Our laboratory showed that EBNA1 increases the phosphorylation of PML proteins by CK2, thereby increasing their ubiquitylation and degradation (Sivachandran, Cao, and Frappier 2010, Cao et al. 2014).
1.3 Casein Kinase 2 (CK2)

Casein kinase 2 is an evolutionarily conserved serine/threonine kinase that is present in all eukaryotic cells. Since its identification in 1954 by Burnett and Kennedy, CK2 has been recognized as a major contributor to the cellular phosphoproteome. Indeed, more than 300 substrates have been discovered for CK2 (Bibby and Litchfield 2005). Such a large number of substrates for CK2 accounts for its pleiotropic effects in diverse cellular processes including cellular proliferation, cell-cycle regulation, apoptosis and the DNA damage response (Guerra and Issinger 1999, Homma and Homma 2005, Ahmad et al. 2008, Kang et al. 2009). CK2 is known to localize to different cellular compartments including the nucleus, ribosomes, Golgi bodies and even the plasma membrane, accounting for its participation in a plethora of cellular processes. CK2 is upregulated in several tumours, consistent with its role in promoting cell survival and proliferation (Pinna 2002, Litchfield 2003, Bibby and Litchfield 2005). In addition, some viruses are known to hijack CK2 enzymatic activities to favour viral replication (Filhol and Cochet 2009). While the complete regulation and cellular functions of CK2 are beyond the scope of this thesis, I will discuss some of the important properties of CK2 in the section below.

1.3.1 CK2 enzymatic properties

CK2 has a minimum consensus phosphorylation sequence of S/T-X-X-D/E and differs from other protein kinases in that it can utilize both ATP and GTP for phosphorylation (Pinna 1990, Pinna and Meggio 1997). The requirement for ATP or GTP is governed by cations, since ATP is preferred in the presence of Mg$^{2+}$ whereas GTP is preferred in the presence of Mn$^{2+}$ (Niefind et al. 1999). The D/E in the CK2 consensus sequence may be replaced by a phosphoserine or phosphotyrosine but not by a phosphothreonine, which allows CK2 to participate in hierarchal phosphorylation events with other kinases (Bibby and Litchfield 2005). While most CK2
substrates conform to the consensus sequence, there are some examples of efficient phosphorylation by CK2 in the absence of the minimum consensus sequence. For example, S392 (a DNA damage inducible phosphorylation site) in human p53 is effectively phosphorylated by CK2 in the absence of a consensus sequence (Keller and Lu 2002). Although traditionally viewed as a serine/threonine kinase, recent evidence suggests that CK2 is also capable of phosphorylating tyrosine residues in yeast and mammalian cells (Wilson et al. 1997, Litchfield 2003).

1.3.2 CK2 subunits

CK2 typically exists as a heterotetrameric complex composed of two catalytic (CK2α and/or CK2α’) and two regulatory subunits (CK2β) (Figure 3). A CK2β dimer forms the core of the CK2 tetramer and assembly of CK2β dimers is essential for CK2 tetramer assembly (Bibby and Litchfield 2005). In mammals, although CK2α and CK2α’ are products of two different genes, they share more than 90% identity within their N-terminal 330 amino acids while having completely unrelated C-termini. On the other hand, CK2β shows no similarity to any other known protein (Pinna 2002, Litchfield 2003, Bibby and Litchfield 2005). Although largely redundant, the catalytic subunits of CK2 also have certain unique functions. For example, a knockout of CK2α’ shows defects in spermatogenesis and this cannot be overcome by CK2α (Vilk et al. 1999). Both the isoforms also show differences in cell cycle dependent subcellular localization and phosphorylation. Apart from the two catalytic isoforms of CK2, a third isoform CK2α’’ has been discovered that differs from CK2α in only the last 32 amino acids. The functional significance of this third isoform is unknown (Shi et al. 2001).

In contrast to CK2α, a single isoform of the regulatory subunit (β) of CK2 exists and it has an identical amino acid sequence in mammals and birds (Maridor et al. 1991). A CK2β knockout
has varying effects in different organisms; CK2\(\beta\) knockout mice display early embryonic lethality whereas CK2\(\beta\) homozygous mutant \textit{S.cerevisiae} are viable but display salt sensitivity (Glover 1998, Buchou et al. 2003). In addition, RNAi knockdown of CK2\(\beta\) in \textit{Caenorhabditis elegans} leads to a failure in development (Buchou et al. 2003). Targeted expression of CK2\(\alpha\) in T cells of transgenic mice leads to lymphomagenesis and mammary tumorigenesis (Seldin and Leder 1995, Landesman-Bollag et al. 2001). These results can be explained by the fact that CK2 cellular substrates include products of proto-oncogenes like c-Myc, c-Myb and c-Jun, as well as tumor suppressor proteins like BRCA1 and p53 (Bibby and Litchfield 2005).

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{ck2_structure.png}
\caption{Crystal structure of CK2. Crystal structure of the CK2 holoenzyme with catalytic (\(\alpha\)) and regulatory (\(\beta\)) subunits is shown as determined by Niefind et al. (2001) (PBD 1JWH). The purple sticks are phosphate groups (top panel). The bottom panel represents the KSSR motif that binds phosphate ions (red and orange sticks) (\textit{from Cao et.al 2014}).}
\end{figure}
1.3.3 **CK2β functions**

CK2α subunits purified from bacteria are constitutively active *in vitro* suggesting that the CK2β subunits are dispensable for CK2 activity. Nevertheless, a growing body of literature suggests that, not only does CK2β increase the stability of the catalytic subunits, but also modulates their substrate specificity. For example, CK2β can stimulate CK2 catalytic activity towards topoisomerase II and p53 while inhibiting its activity towards calmodulin (Marin, Meggio, and Pinna 1999).

However, challenges to the traditional view of CK2 as a stable tetramer arose when it was discovered through X-ray crystallography that the CK2α/CK2β interaction interface of 832Å was much smaller compared to the average interaction interface of 1722Å for permanent protein subunit interactions (Jones and Thornton 1996, Niefind et al. 2001). This observation raised the possibility of a dynamic complex that is subject to assembly and disassembly. Furthermore, the CK2β subunits are synthesized in excess of CK2α (Guerra and Issinger 1999) and live cell imaging revealed that certain pools of CK2α and CK2β do not colocalize (Stigare et al. 1993). The question then arose whether CK2β is capable of performing functions independent of its functions within the CK2 holoenzyme. The answer to this question was revealed with the discovery of A-Raf, c-Mos and Chk1 as binding partners of CK2β (independent of its binding as a part of the CK2 holoenzyme) (Hagemann et al. 1997, Chen et al. 1997). While CK2β has a stimulatory effect on the enzymatic activities of A-Raf and Chk1, its overexpression inhibits the activity of c-Mos underlying its pleiotropic effects (Hagemann et al. 1997, Chen et al. 1997, Guerra, Issinger, and Wang 2003).
1.3.4 \textit{CK2 in viral infections}

CK2 is a very important target for viruses due to its participation in diverse cellular processes. Some viruses use CK2 to phosphorylate viral proteins, while others redirect CK2 catalytic activities to different cellular substrates. By 2003, around 40 viral proteins were known to be phosphorylated by CK2 (Bibby and Litchfield 2005). Viral proteins substrates of CK2 include Rev from HIV-1 (phosphorylation allows for down-regulation of Rev) (Meggio et al. 1996), E1 and E2 proteins of BPV (phosphorylation inactivates DNA binding activity of both) (Schuck, Ruse, and Stenlund 2013), ppUL44 (phosphorylation enhances nuclear transport of ppUL44) (Alvisi et al. 2005) and EBV BZLF1 (phosphorylation modulates ability of BZLF1 to activate and repress transcription) (El-Guindy and Miller 2004). In addition, CK2 phosphorylation of the EBV protein EB2 is essential for efficient nuclear export of early and late viral mRNAs necessary for infectious virus production (Medina-Palazon et al. 2007).

Examples of viruses redirecting CK2 include the following: Adenovirus infection is known to redirect subcellular localization of the catalytic and regulatory subunits of CK2, with only CK2\(\alpha\) being associated with adenovirus pIX protein nuclear bodies and CK2\(\beta\) being found in inclusion bodies of viral structural proteins (Souquere-Besse et al. 2002). Infection with herpes simplex virus 1 leads to redistribution of CK2 to the cytoplasm where it phosphorylates the immediate early viral protein IE63 (Wadd et al. 1999). During productive infection of fibroblasts by rodent parvovirus, the viral protein NS1 mediates CK2 dependent phosphorylation of tropomyosin as well as other cytoskeletal components, leading to cytoskeletal and physiological changes that cause cell death and lysis (Corbau et al. 2000, Nuesch and Rommelaere 2007). Finally, as mentioned above, EBNA1 recruits CK2 to PML NBs resulting in increased phosphorylation and proteasomal degradation of PML proteins.
1.3.5  *EBNA1 binding to CK2*

Our laboratory identified the EBNA1 binding pocket within CK2 that is present at the interface of the two regulatory subunits of CK2 and includes a KSSR sequence in CK2β. The identification of this binding pocket was based on the finding that EBNA1 phosphorylation at S393 is essential for CK2β binding and the fact that the KSSR sequence in CK2β had been found to bind phosphates from purification buffers, as seen in the CK2 crystal structure (Figure 3). Co-IP experiments showed that the EBNA1-CK2β interaction is mediated by EBNA1 amino acids 387-394 and KSSR in CK2β. This was the first report of the KSSR sequence mediating protein-protein interactions. Since viral proteins frequently mimic cellular proteins, it was postulated that the KSSR sequence might mediate some cellular interactions of CK2β. Comparative proteomics using WT CK2β and a CK2β KSSR mutant revealed that while most CK2β interactions remained unaffected by the KSSR to AAAA mutation, a few cellular proteins bound to WT CK2β but not the KSSR mutant (Cao et al. 2014). The most prominent amongst these proteins were ARKL1 (C18orf25) and BCLAF1. Subsequent co-IP experiments confirmed that both ARKL1 and BCLAF1 bound to WT CK2β but not the CK2β KSSR mutant (Cao et al. 2014). Since my work is predominantly on ARKL1, I will discuss this protein in further detail in the next section.

1.4  **ARKL1**

ARKL1 is a 43 KDa protein that is yet to be functionally characterized. Although it is largely predicted to be an unstructured protein with no known domains, it does have two putative high affinity SUMO interacting motifs (SIMs) (Figure 4A). The gene for ARKL1 is conserved in vertebrates and is expressed in all human tissues. A recent report (published after completion of my experiments) suggests that ARKL1 may in fact be a restriction factor for influenza virus
infection. ARKL1 was found to be differentially SUMOylated upon influenza virus A infection and silencing ARKL1 using shRNA resulted in increased influenza A virus titres (Domingues et al. 2015).

Examination of the ARKL1 protein sequence reveals homology to the N-terminus of the cellular E3 ubiquitin ligase Arkadia. Arkadia, like ARKL1 was shown to bind to the CK2β KSSR motif in co-IP experiments (Cao and Frappier, unpublished data). Sequence comparison between ARKL1 and Arkadia also reveals a stretch of 25 amino acids that is rich in serine residues. This serine-rich region is similar to the EBNA1 serine-rich region important for CK2β binding, suggesting that ARKL1 may utilize this region for CK2β binding (Figure 4B). As predicted, deletion of this region in ARKL1 (amino acids 202-226) abrogated its interaction with CK2β (Cao et al. 2014).

1.5 Arkadia

Arkadia (RNFIII) is a 994 amino acid protein that contains three high affinity SIMs along with a C-terminal RING domain (Figure 4A). The RING domain allows it to function as an E3 ubiquitin ligase that targets proteins for proteasomal degradation by adding K48 linked ubiquitin chains to them (Levy et al. 2007). Arkadia was first identified in a study that showed that homozygous Arkadia knockout mice fail to develop the embryonic node, an indication of missing Nodal signaling (Episkopou et al. 2001). Arkadia was subsequently identified as a positive regulator of TGF-β signaling by targeting the negative regulators Smad7, SnoN and c-Ski for proteasomal degradation (Koinuma et al. 2003, Nagano et al. 2007, Levy et al. 2007). A later report also demonstrated that Arkadia functions in the DNA damage response pathway by recruiting xeroderma pigmentosum c (XPC), a factor involved in nucleotide excision repair (NER), to UV-damaged DNA by adding K63 linked ubiquitin to it (Poulsen et al. 2013). In
addition, Arkadia can act as a SUMO targeted ubiquitin ligase (STUbL) to bring about degradation of PML proteins in response to arsenic trioxide treatment (Erker et al. 2013).

Finally, Arkadia is known to localize to PRC1 bodies through multiple structures and function as a transcriptional repressor for around 200 cellular genes (Sun, Liu, and Hunter 2014).

**Figure 4. ARKL1 shows homology to the N-terminus of Arkadia.** A. Diagrammatic representation of Arkadia and ARKL1. B. Alignment of ARKL1 and Arkadia amino acids 47-397. The serine rich region similar to the EBNA1 serine rich region involved in CK2β binding (a) and the SUMO interacting motifs (SIMs) (b and c) are shown.
1.6 The Polycomb Repressive Complex (PRC)

The polycomb group of proteins were originally discovered in the fruitfly due to their ability to repress the Hox genes (Ringrose et al. 2003). These proteins function as transcriptional repressors that govern cell fate during development, and they also have roles in controlling cellular proliferation and neoplastic development (Haupt et al. 1991, van Kemenade et al. 2001, Visser et al. 2001, Boyer et al. 2006, Lee et al. 2006). The polycomb group proteins are classified into the polycomb repressive complex 1 (PRC1) and the polycomb repressive complex 2 (PRC2) on the basis of their assembly into distinct multimeric complexes. In humans, PRC2 is composed of EED (extraembryonic development), Ezh1 and 2 (enhancer of zeste) and Suz12 (suppressor of zeste) (Czermin et al. 2002, Muller et al. 2002), whereas PRC1 is composed of the Chromobox proteins CBX2/4/8, EDR1,2 and 3, RING1A, RING1B, BMI1, MEL18, ZNF134, YY1, SCMH1 and PHF1 (Min, Zhang, and Xu 2003, Wang, Brown, et al. 2004). The catalytic component of PRC2, Ezh2 is responsible for trimethylating histone H3K27 (and to a lesser extent H3K9), the first step in gene repression (Czermin et al. 2002). The trimethylated histones are then recognized by the chromobox domains of the CBX proteins, facilitating recruitment of PRC1 to these sites (Min, Zhang, and Xu 2003, Wang, Brown, et al. 2004). The RING1A/RING1B proteins contain a RING domain that is capable of ubiquitylating histone H2A at lysine 119, which results in chromatin compaction (Wang, Wang, et al. 2004). In addition, PRC also prevents SWI/SNF dependent remodelling of nucleosomes and binding of the transcription machinery to the promoters of repressed genes (Francis et al. 2001, Dellino et al. 2004). Apart from the above mentioned functions, PRC also recruits DNA methyltranferases to their target genes which contribute to gene repression (Vire et al. 2006, Sparmann and van Lohuizen 2006).
1.6.1 Epigenetic regulation of EBV infection

Following entry into B cells, EBV episomes are rapidly incorporated into nucleosomes. Several studies have shown the modulation of EBV latency is in part mediated by the epigenetic modifications of histones at viral promoters. Genome-wide studies of histone modifications have shed light on the epigenetic state of EBV promoters during different forms of latency. For example, the C_p and LMP1/2 promoters show enrichment for H3K4me3 and acetylated histones (H3K9, H3K27, H4) during type III latency, when they are active. Conversely, the same marks are absent from these promoters during type I latency when these promoters are inactive. On the other hand, the Q_p promoter carries these marks during type I latency (when it is active) but not during type III latency (Gerle et al. 2007, Day et al. 2007, Tempera et al. 2010). Also, the H3K9me3 mark is present at low levels in the viral genome during type III latency when most viral latency genes are expressed but present at high levels during latency I (Day et al. 2007, Arvey, Tempera, and Lieberman 2013). Finally, the promoter for the gene for BZLF1 shows high levels of H3K27me3 during latency, a sign of facultative heterochromatin that would allow silencing of that gene, and this mark greatly decreases when the lytic cycle is induced. Consistent with this epigenetic regulation, silencing the H3K27 methyltransferase Ezh2 results in increased reactivation of EBV and the disappearance of H3K27me3 from EBV lytic cycle promoters (Murata et al. 2012, Murata and Tsurumi 2013, Tempera and Lieberman 2014).
1.7. Thesis Rationale

A goal of my thesis was to determine the functional significance of ARKL1-CK2β and Arkadia-CK2β interactions in the context of EBV infection. I present evidence suggesting that both ARKL1 and Arkadia function to suppress reactivation of EBV. During the course of my experiments, I also discovered that ARKL1 colocalizes to CBX4 nuclear foci (some of which could be PRC1 bodies). ARKL1 has two putative SIMs that resemble Arkadia SIMs and I showed that mutation of these ARKL1 putative SIMs compromised its ability to associate with SUMO and to colocalize with CBX4. Another goal of my project was to determine cellular functions of ARKL1, and how these functions are regulated. As a first step towards this goal, I identified cellular interactors of ARKL1 by affinity purification coupled to mass spectrometry. These included the Jun transcriptional activator and the chromatin regulators GRWD1 and DEK. Overall I have provided the first evidence that ARKL1 and Arkadia contribute to the regulation of EBV infection and the first information on the cellular interactors of ARKL1.
CHAPTER 2: MATERIALS AND METHODS

2.1 Cell Lines

The EBV-negative nasopharyngeal carcinoma cells CNE2Z (Sun et al. 1992) were grown in α minimum essential medium (αMEM, Gibco), the U2OS osteosarcoma cells and HEK293T human embryonic kidney cells were grown in Dulbecco’s modified essential medium (DMEM, Gibco) and the EBV-positive AGS-EBV gastric carcinoma cells (Borza et al. 2004) were grown in RPMI (Gibco) at 37°C. Each medium was supplemented with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin and 100 µl plasmacin. In addition, AGS-EBV cells were grown in the presence of 12.5% G418 in order to stably maintain EBV episomes.

2.2 Plasmids

The coding sequences for N-terminally FLAG tagged ARKL1 and ARKL1ΔS were inserted into the multiple cloning site of the pc3mvFC plasmid between the Bgl II and Not I sites. The ARKL1ΔS mutant lacking amino acids 202-226 was generated by gene synthesis and is described in detail in Cao et al, 2014. FLAG-ARKL1sim was constructed by synthesizing the ARKL1 gene with mutations converting the codons for valine (332,335,389,390), isoleucine334 and leucine392 to alanine using GeneArt Strings (Thermofisher). This ARKL1sim fragment was inserted between the NotI and BamHI sites of pc3mvFC plasmid containing FLAG-ARKL1 (by Kathy Shire). EBNA1 and EBNA1 S393A were expressed from pc3oriP plasmid constructed as previously described (Shire et al. 1999, Shire et al. 2006, Cao et al. 2014). The construction of HA-CK2β into the pcDNA5 plasmid is described in Cao et al. Plasmids expressing Myc-SUMO2 and Myc-SUMO3 (MacPherson et al. 2009), His-SUMO1 (de la Cruz-Herrera et al. 2014), Myc-USP7 (Sarkari et al. 2011) and Myc-CBX4 (Sun, Liu, and Hunter 2014) are
previously described. The plasmid expressing FLAG-DDX24 is described in Yamauchi et al. (2014).

2.3 Determining ARKL1 Effects on SnoN

CNE2Z cells on 10 cm dishes were subjected to three rounds of transfection with 125 picomoles siRNA targeting ARKL1 or Arkadia (shown in Table 1) or AllStars negative-control siRNA (Qiagen) separated by 24 hrs, using 4 µl of Lipofectamine 2000 (Thermofisher). 24 hrs after the last siRNA treatment, cells were starved for 24 hrs in αMEM containing 0.1% FBS and 25 µM SB431542 (Sigma-aldrich S4317), a potent inhibitor of the TGF-β superfamily type I activin receptor-like kinases (Inman et al. 2002), in order to allow accumulation of intracellular SnoN (Kawamura et al. 2012). Medium was replaced with fresh medium containing 5 pM TGF-β (Gibco, PHG9214) and incubated for 1 or 3 hrs. Samples were also collected prior to TGF-β addition (Bonni et al. 2001). Cells were then trypsinized (Gibco), washed twice with PBS, lysed in 9M urea, 5 mM Tris-HCl (pH 6.8) and subjected to SDS-PAGE on 10% gels.

Table 1. siRNAs Used in this Project

<table>
<thead>
<tr>
<th>siRNA</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARKL1 siRNA1</td>
<td>5’-ACGCAGUAUGUUAGCACCAGGCAA-3’</td>
</tr>
<tr>
<td>ARKL1 SIRNA2</td>
<td>5’-CAUCUGCUGGCAAUGCGCCACUCA-3’</td>
</tr>
<tr>
<td>Arkadia siRNA1</td>
<td>5’-CAUUGAGGCCCAGCUGCCAAGUGAA-3’</td>
</tr>
<tr>
<td>Arkadia siRNA2</td>
<td>5’-CAUCAAGCUUCUGCCUGCCGCAU3’</td>
</tr>
</tbody>
</table>
2.4 Assays for Lytic Reactivation of EBV upon ARKL1 or Arkadia Silencing

AGS-EBV cells on 10 cm dishes were subjected to two rounds of transfection, separated by 24 hrs, with 125 picomoles of siRNA against ARKL1, Arkadia (Table 1) or AllStars negative-control siRNA in the presence of 4 µl of Lipofectamine 2000. 24 hrs later, cells were split into two 10 cm dishes containing coverslips treated with polylsine and subjected to a third round of siRNA transfection. 24 hrs after the last round of transfection, cells were treated with a final concentration of 3 mM sodium butyrate (NaB) and 20 ng/µl tetradecanoylphorbol-13-acetate (TPA) for 24 hrs. Coverslips were subsequently transferred to 6 well dishes and fixed for immunofluorescence microscopy (section 2.8). At least 100 cells were counted on each slide for each experiment. The P-value was determined using a one tailed T-test with unequal variances.

For Western blot analyses, siRNA-treated cells in the 10 cm dishes were harvested, washed in 1 ml of PBS, split into two fractions of 500 µl and clarified by centrifugation at 135xg. One fraction was lysed in 75 µl of 9 M urea and 5 mM Tris-HCl (pH 6.8) for SDS-PAGE analysis (section 2.11) whereas the other fraction was used for RNA extraction and quantitative RT-PCR (section 2.5).

2.5 RNA Extraction and Quantitative PCR

The cell pellet was resuspended in 1ml of trizol and 200 µl of chloroform and vortexed vigorously. The trizol-chloroform mixture was then allowed to stand for 10 minutes at room temperature followed by centrifugation at 4°C for 10 minutes at 21130xg. The supernatant was transferred to a new tube, taking care not to disturb the liquid interface. 500 µl of isopropanol was added to the supernatant, vortexed and incubated at -20°C for 3 hours. The mixture was then subjected to centrifugation at 21130xg for 30 minutes. The supernatant was discarded without disturbing the pellet. The pellet was washed with 70% ethanol followed by centrifugation for 15
minutes at 4°C at 21130xg. Once again the supernatant was discarded while taking care to not disturb the pellet. The pellet was dried for 15 minutes and resuspended in RNAase-free water. Absorbance of the solution was measured and 5 µg of RNA was treated with DNAase1 for 30 minutes. The DNAase1 was inactivated by heating to 65°C for 10 minutes and 1 µg of the DNAase-treated RNA was used for cDNA synthesis. cDNA was synthesized using Superscript™ IV reverse transcriptase kit (Thermofisher) as per the manufacturer’s instructions. The synthesized cDNA was used directly for quantitative real time PCR using the SsoFast™ EvaGreen Supermix. The gene for the ribosomal protein RPL30 was used as an internal reference. The primers used for quantitative real time PCR are listed in Table 2.

**Table 2. Primers Used in this Project**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arkadia forward primer</td>
<td>5’-TCCCTAGAGTTCAGGATGG-3’</td>
</tr>
<tr>
<td>Arkadia reverse primer</td>
<td>5’-CTTAAGGAAAGCCTTCATCACTC-3’</td>
</tr>
<tr>
<td>RPL30 forward primer</td>
<td>5’-GCTTTGCATTGTGGGAAGTC-3’</td>
</tr>
<tr>
<td>RPL30 reverse primer</td>
<td>5’-CCACCATCTTCCTGCTTAG-3’</td>
</tr>
</tbody>
</table>

**2.6 Overexpression Experiments**

AGS-EBV cells grown on polylysine-treated coverslips were transfected with 1 µg of a plasmid expressing FLAG-ARKL1, FLAG-DDX24 or FLAG-ARKL1ΔS using 2 µl polyjet. 24 hrs later, a final concentration of 3 mM NaB and 20 ng/µl TPA was added. 24 hrs after induction, coverslips were rinsed with PBS and fixed and stained for immunofluorescence microscopy (see
section 2.8). At least 50 FLAG-positive and 50 FLAG-negative cells were counted on each coverslip. P value was determined using a one tailed T-test with unequal variances.

2.7 ARKL1 Localization to SUMO and CBX4 Foci

CNE2Z cells on coverslips were transfected with 0.5 µg of plasmids expressing FLAG-ARKL1, FLAG-ARKL1ΔS or FLAG-ARKL1sim and 0.5 µg of His-SUMO1, myc-SUMO2, myc-SUMO3 or myc-CBX4. 24 hrs post-transfection, cells were processed for immunofluorescence microscopy (see section 2.8).

2.8 Immunofluorescence Microscopy

Cells were washed with PBS and fixed with 3.7% formaldehyde in PBS for 20 min. Coverslips were rinsed with PBS and permeabilized with 1% Triton X-100 in PBS for 5 min. After rinsing with PBS, coverslips were blocked with 4% BSA in PBS and incubated with primary antibodies against FLAG (Bethyl A190-101A at 1:800), BZLF1 (sc-53904 at 1:400), ARKL1 (sc-84701 at 1:400), SUMO1 (sc-5308 at 1:200) and Myc (sc-40 at 1:200). Coverslips were rinsed twice with PBS and incubated in Alexa-fluor conjugated secondary antibodies goat anti-rabbit 488 (Molecular probes) and goat anti-mouse 647 (Molecular probes) for reactivation experiments. For ARKL1-SUMO/CBX4 localization experiments, donkey anti-goat/goat anti-rabbit Alexafluor 488 and donkey anti-mouse Alexafluor 555/donkey anti-rabbit 555 were used. Coverslips were rinsed three times with PBS following secondary antibody treatment and mounted using ProLong Gold antifade medium containing DAPI (Invitrogen). Images were obtained using the 40 X oil objective on a Leica inverted fluorescent microscope and processed using OpenLAB (ver.X.0) software.
2.9 ARKL1-EBNA1 Competition Experiment

HEK293T cells in 10 cm dishes were transfected with 1 µg of the pc3MVFC empty plasmid or plasmids expressing FLAG-ARKL1 or FLAG-ARKL1ΔS along with 4 µg of the pc3oriP empty plasmid, or pc3oriP plasmids expressing EBNA1 or EBNA1S393A and 1 µg of plasmid expressing HA-CK2β. 24hrs after transfection, cells were moved to 15 cm dishes. 48 hrs after transfection, cells were harvested and washed twice with PBS. The cell pellets were lysed by resuspending in a 1:4 weight by volume ratio of modified RIPA buffer (50 mM Tris HCl pH 8.0, 300 mM NaCl, 0.1% Na deoxycholate, 0.5% NP40, 2 mM EDTA) containing protease inhibitor cocktail P8340 (Sigma) and HALT phosphatase inhibitor cocktail (Thermofisher) for 30 minutes. Lysates were sonicated and clarified by centrifugation at 21130xg at 4°C. The supernatant was retained and protein concentration was measured by a Bradford assay. 2 µg of total protein at 8 µg/µl (diluted with RIPA buffer) was incubated with 30 µl of M2 anti-FLAG beads (Sigma-Aldrich A2220) (pre-washed with modified RIPA buffer) and incubated at 4°C overnight. The beads were then pelleted by centrifugation at 845xg for 3 minutes, washed with modified RIPA buffer three times, then once with wash buffer (50 mM ammonium carbonate, 75 mM KCl). Protein was eluted by boiling the beads in 2x SDS-PAGE loading buffer (65.8 mM Tris-HCl, pH 6.8, 26.3% w/v glycerol, 2.1% SDS, 0.01% bromophenol blue) for 5 min and loaded onto a 10% gel.

2.10 ARKL1-SUMO3 Immunoprecipitation

HEK293T cells in 10 cm dishes were transfected with 2.5 µg of the pc3MVFC empty plasmid or plasmids expressing FLAG-ARKL1 or FLAG-ARKL1ΔS, and with 2.5 µg of plasmid expressing myc-SUMO3 using 2 µl of polyjet per µg of DNA. Cells were moved to 15cm dishes 24hrs post
transfection, harvested 48 hrs post-transfection and processed for a FLAG-IP as described in section 2.9. Samples were subjected to SDS-PAGE using a 15% gel.

2.11 Western Blotting

Proteins were separated by SDS-PAGE and transferred to a nitrocellulose membrane for 1hr at a 100V. The nitrocellulose membrane was blocked in 4% skim milk in PBS containing 0.1% tween-20 (PBS-T) for 1 hr then incubated with the following antibodies: actin (Santa Cruz, sc-1615 at 1:1000), ARKL1 (Santa Cruz, sc-84702 at 1:200), BZLF1 (Santa Cruz, sc-53904 at 1:5000), BMRF1 (Millipore, MAB-8186 at 1:5000), FLAG (Bethyl, A190-101A at 1:20,000), Myc (Santa Cruz, sc-789 at 1:2000) and SnoN (Santa Cruz, sc-9141 at 1:200). Membranes were washed with PBS-T three times for 5 min each, then incubated for 1 hr with secondary antibodies; horse radish peroxide conjugated donkey anti-goat, goat anti-rabbit and goat anti-mouse. Membranes were subsequently washed three times with PBS-T for 10 min each and developed using the Western blotting luminol reagent (Santa Cruz, sc-2048) or the Amersham ECL reagent (GE healthcare) where indicated.

2.12 ARKL1 Affinity Purification Coupled with Mass Spectrometry (AP-MS)

HEK293T cells in three 10 cm dishes were transfected with 5 µg of the pc3MVFC empty plasmid or plasmid expressing FLAG-ARKL1 or FLAG-ARKL1sim using polyjet as described above. Cells were transferred to 15 cm dishes 24 hrs post-transfection and harvested 48 hrs post-transfection. The cell pellet was washed twice with ice cold PBS then resuspended in ice cold modified RIPA buffer (50 mM Tris HCl pH 8.0, 300 mM NaCl, 0.1% Na deoxycholate, 0.5% NP40, 2 mM EDTA) containing protease inhibitor cocktail P8340 (Sigma) and HALT phosphatase inhibitor cocktail (Thermofisher) in a 1:4 weight:volume ratio. The samples were incubated for 30 mins on ice, sonicated and subjected to centrifugation at 21130xg for 30 mins,
then the supernatant was transferred to a new tube. A Bradford assay was used to measure protein concentration and 3 mg of total protein at 8 µg/µl was added to 30 µl of M2 anti-FLAG beads (Sigma-Aldrich A2220) (pre washed with modified RIPA buffer). Samples were mixed at 4°C for 4 hrs and collected by centrifugation at 845xg for 3 min. The beads were washed three times with modified RIPA buffer and once with wash buffer (50 mM ammonium carbonate, 75 mM KCl). Elution was performed by adding freshly prepared elution buffer (500 µl of ammonium hydroxide in 5 ml HPLC, pH 11-12) to the beads and mixing at 4°C for 15 min. The beads were pelleted and the supernatant was transferred to a new tube. The elution step was repeated twice and elutions were pooled into a single tube. The tube containing the eluted proteins was subjected to centrifugation at 21130xg to pellet the residual beads. The supernatant was transferred to a new tube and lyophilized overnight. 100 µl HPLC grade water was added and the tubes were lyophilized again until the liquid had completely evaporated. Dried protein was resuspended in 50 mM ammonium bicarbonate containing proteomics grade trypsin (Sigma-Aldrich T6567) and incubated overnight at 37°C. The next day, 2.5 µl of proteomics grade trypsin was added to the solution and reincubated for 2 hrs at 37°C. The samples were then lyophilized until all the liquid had evaporated. The resultant tryptic peptides were detected by tandem liquid chromatography-mass spectrometry (LC-MS/MS) using a LTQ orbitrap system (Thermo Finnigan) and identified using Mascot software (Matrix science, United Kingdom). LC-MS/MS was performed in Jack Greenblatt’s lab.
CHAPTER 3: RESULTS

3.1 ARKL1 Silencing does not Affect SnoN Levels

Since ARKL1 shows sequence homology to the N-terminus of Arkadia without possessing the RING domain required for E3 ubiquitin ligase activity, I postulated that ARKL1 might function as a competitive inhibitor of Arkadia by binding to its substrates and preventing their interaction with Arkadia. Arkadia functions as a positive regulator of TGF-β by ubiquitinating the SnoN repressor and targeting it for proteasomal degradation upon stimulation with TGF-β, thereby allowing Smad2/3/4 dependent transcriptional activation of TGF-β responsive gene (Koinuma et al. 2003, Levy et al. 2007). To investigate if ARKL1 regulates SnoN levels, CNE2Z cells were treated with siRNA against ARKL1 or a negative-control siRNA and stimulated with TGF-β. Samples were collected at various time points post stimulation and blotted for SnoN. ARKL1-silenced cells showed a decrease in SnoN upon TGF-β stimulation comparable to cells treated with a siRNA control. On the other hand, cells treated with Arkadia siRNA showed a basal increase in SnoN levels, consistent with previous reports of Arkadia functioning as an E3 ubiquitin ligase for SnoN (Figure 5). Although Arkadia silencing could not be verified due to lack of an antibody that detects endogenous Arkadia, the ability of this siRNA to down regulate Arkadia mRNA levels was verified in later experiments by quantitative RT-PCR (see Figure 6C). Based on this result, I conclude that ARKL1 does not regulate SnoN levels upon TGF-β stimulation.
Figure 5. ARKL1 silencing does not affect SnoN levels. CNE2Z cells were treated with a siRNA against ARKL1, Arkadia or a negative-control siRNA and induced with TGFβ for 3hrs. Samples were collected at the indicated time points and lysed in 9M urea. Western blotting was performed using antibodies against SnoN, ARKL1 and actin.

3.2 ARKL1 or Arkadia Silencing Promotes Reactivation of EBV

ARKL1 and Arkadia bind to the same site in CK2 that EBNA1 uses to impact EBV infection, including promoting EBV lytic infection. Therefore, I wanted to determine if ARKL1 or Arkadia affected the mode of EBV infection. To investigate if ARKL1 or Arkadia have a role in reactivating EBV to the lytic cycle, a gastric carcinoma cell line containing latent EBV episomes (AGS-EBV cells) was treated with two different siRNAs against either ARKL1 or Arkadia. This cell line was chosen because it is highly permissive for EBV reactivation. Silencing of ARKL1 or Arkadia resulted in increased expression of the EBV lytic cycle proteins BZLF1 and BMRF1 compared to a control siRNA, as assessed by western blotting (Figure 6A). BZLF1 is the immediate early EBV lytic cycle protein that functions to turn on the expression of other lytic cycle genes, including the gene for the viral polymerase processivity factor BMRF1. As
described earlier, the percentage of cells showing spontaneous reactivation is typically low and hence it is common to treat EBV containing cells with NaB/TPA in order to reverse the repressive epigenetic modifications at the promoters of lytic cycle genes. ARKL1 or Arkadia silencing using either siRNA resulted in greater EBV reactivation even in the presence of NaB/TPA (Figure 6B). Arkadia silencing was verified by quantitative RT-PCR (Figure 6C).

To determine if the increase in BZLF1 upon ARKL1 or Arkadia silencing is due to an increase in the number of cells entering the lytic cycle, AGS-EBV cells on coverslips were silenced for ARKL1 or Arkadia and immunofluorescence microscopy was performed by staining for BZLF1. Either ARKL1 or Arkadia silencing resulted in an increase in the number of cells expressing BZLF1 both in the presence and absence of NaB/TPA (Figure 7). While ARKL1 siRNA1 caused an increase in the number of cells expressing BZLF1 both in the presence and absence of NaB/TPA, siRNA2 caused an increase in the number of cells showing reactivation only upon induction. This could be related to the very low levels of BZLF1 expression, which might have been missed while counting BZLF1-positive cells in the siRNA2 silenced sample.
Figure 6. ARKL1 or Arkadia silencing induces reactivation of EBV. AGS-EBV cells were treated with siRNA against ARKL1, Arkadia or AllStars siRNA control and treated with A) DMSO (uninduced) or B) NaB/TPA (induced). Silencing using two different siRNAs is shown. Equal amounts of cell lysates were loaded for each condition and probed for ARKL1, BZLF1, BMRF1 and actin. Note that twice as much total cell lysate was loaded for uninduced samples as compared to induced samples and developed using the more sensitive Amersham ECL prime reagents (GE healthcare). Band quantification for BZLF1 is shown beside uninduced samples and for BMRF1 is shown for uninduced and induced samples. C) Arkadia transcripts quantified.
using RT-PCR are shown for uninduced and induced samples. RT-PCR was performed in duplicate. Average values are shown with standard deviation.

![Graph showing percentage of BZLF1 positive cells](image)

**Figure 7. ARKL1 or Arkadia silencing increases the percentage of AGS-EBV cells reactivating to the lytic cycle.** AGS-EBV cells on coverslips were treated with siRNA against ARKL1, Arkadia or AllStars negative-control siRNA and harvested for immunofluorescence microscopy 24hrs after treatment with DMSO (uninduced) or NaB/TPA(induced). Cells were stained using an antibody against BZLF1. Averages and standard deviations from three independent experiments are shown, where at least 100 cells were counted for each condition.

*P<0.05, **P<0.0001 (One tailed t-test with unequal variances)

### 3.3 ARKL1 Overexpression Inhibits EBV Reactivation

One way to verify that the results of my silencing experiments are really due to effects on ARKL1 or Arkadia (as opposed to off target effects) is to determine if overexpression has the opposite effect. To determine if ARKL1 overexpression inhibits EBV reactivation, FLAG-ARKL1 was overexpressed in AGS-EBV cells and cells were induced for 24hrs by treatment with NaB/TPA. Immunofluorescence microscopy performed for BZLF1 revealed that FLAG-
positive cells showed a lower percentage of EBV reactivation as compared to FLAG-negative untransfected cells within the same slide (Figure 8A). To elucidate if CK2β binding was necessary for ARKL1 to suppress reactivation, an ARKL1 mutant lacking the serine rich region involved in CK2β binding (FLAG-ARKL1ΔS) was overexpressed alongside FLAG-ARKL1 and the number of BZLF1-positive cells were identified by immunofluorescence microscopy (following induction with NaB/TPA) and counted. The data for this experiment is expressed as the ratio of BZLF1-positive FLAG-positive cells to the BZLF1-positive FLAG-negative cells in the same sample. A ratio greater than 1 indicates exogenous protein expression increases reactivation whereas a ratio less than 1 indicates exogenous protein expression suppresses reactivation. As demonstrated in Figure 8B, overexpression of FLAG-ARKL1 resulted in a reactivation ratio less than 1 whereas overexpression of another FLAG-tagged protein that is unrelated to EBV infection (FLAG-DDX24) showed a ratio greater than 1. FLAG-DDX24 was used as a comparison because other lab members have previously observed that overexpression of any exogenous protein can induce EBV reactivation to some degree. Similar to WT ARKL1, FLAG-ARKL1ΔS also exhibited a reactivation ratio less than 1. Importantly, overexpression of either ARKL1 or ARKL1ΔS caused a decrease (3-4 fold) in EBV reactivation relative to the nonspecific protein control (DDX24). Taken together, these data indicate that ARKL1 suppresses EBV reactivation and that CK2β binding may not be necessary for it to suppress reactivation. Note that I also tried to overexpress FLAG-Arkadia in AGS-EBV cells to determine if Arkadia overexpression inhibits EBV reactivation. However, Arkadia is an unstable protein due to its ability to autoubiquitylate and target itself for proteasomal degradation (Xia et al. 2013, Erker et al. 2013). As a result, very few FLAG-Arkadia positive cells could be detected; too few to draw a meaningful conclusion on effects of Arkadia overexpression on EBV reactivation.
3.4 Effect of EBNA1 on the ARKL1-CK2β Interaction

Next, I wanted to determine if EBNA1 can outcompete ARKL1 for CK2β binding, since both of these proteins bind to the same KSSR site on CK2β. To this end, either FLAG-ARKL1 or FLAG-ARKL1ΔS (or empty plasmid) was overexpressed with HA-CK2β in the presence or absence of EBNA1, and a FLAG IP was performed to monitor the recovery of HA-CK2β with FLAG-ARKL1. In this experiment, HEK293T cells were transfected with four times the amount of plasmid expressing EBNA1 as compared to the plasmid expressing FLAG-ARKL1 in an attempt to have a higher expression level of EBNA1. The results of this experiment demonstrate that the recovery of HA-CK2β with FLAG-ARKL1 was unaffected by the presence of EBNA1 (Figure 9). Therefore, I was unable to demonstrate that EBNA1 outcompetes ARKL1 for binding to CK2β.
Figure 8. ARKL1 overexpression inhibits EBV reactivation. A) AGS-EBV cells on coverslips were transfected with FLAG-ARKL1, induced with NaB/TPA for 24hrs, fixed and stained for BZLF1 and FLAG. The percentage of FLAG-positive and FLAG-negative cells expressing BZLF1 was then determined. Averages and standard deviations from three independent experiments are shown, where at least 50 FLAG-positive and FLAG-negative cells were counted. P=0.06 B) AGS-EBV cells on coverslips were transfected with either FLAG-DDX24, FLAG-ARKL1 or FLAG-ARKL1ΔS. Cells were induced with NaB/TPA for 24 hrs and fixed and stained for BZLF1 and FLAG. Graphs represent averages and standard deviations of two independent experiments where at least 50 FLAG-positive and FLAG-negative cells were counted.
Figure 9. EBNA1 does not outcompete ARKL1 for binding to CK2β. HEK293T cells were transfected with plasmids expressing FLAG-ARKL1 and HA-CK2β in the presence or absence of a plasmid expressing EBNA1. Cells were harvested and IP was performed using M2 anti-FLAG beads. Samples were loaded on a 10% gel and probed for FLAG, HA, EBNA1 and actin (loading control).

3.5 ARKL1 Localizes to CBX4 (PRC1)

Work from the laboratory of Tony Hunter demonstrated that multiple Arkadia sequences coordinate its localization to CBX4 foci, which was used as a marker of polycomb bodies. Based on this finding, I asked if ARKL1 can also localize to CBX4 foci. To this end, FLAG-ARKL1 was overexpressed along with Myc-CBX4 in both U2OS (osteosarcoma) cells and CNE2Z (nasopharyngeal carcinoma) cells. U2OS cells were chosen because they are commonly used to study polycomb foci, whereas CNE2Z cells were chosen because of their relevance in the context of EBV infection. Following overexpression, these cells were stained for FLAG and Myc for visualization by immunofluorescence microscopy. FLAG-ARKL1 was found to localize to Myc-CBX4 in discrete nuclear foci (Figure 10).
3.6 Identification of ARKL1 SIMs Important for Localization to CBX4 Foci

ARKL1 has two putative SIMs (amino acids 332-336 and 389-393), one of which is a high affinity VIDLT type SIM also found in Arkadia (Figure 11A). Since Arkadia SIMs are involved in mediating its interactions with PML proteins and polycomb bodies (Erker et al. 2013, Sun, Liu, and Hunter 2014), I hypothesized that the ARKL1 SIMs might be important for some of its functions, including its ability to affect EBV reactivation. Hydrophobic core residues of SIMs are important for SUMO binding (Jardin, Horn, and Sticht 2015), and therefore I mutated the valine, isoleucine and leucine residues of the ARKL1 putative SIMs to alanine (Figure 11A) (hereby referred to as ARKL1sim). To assess whether the mutation of ARKL1 SIMs affect its ability to associate with SUMO, I overexpressed either WT FLAG-ARKL1, FLAG-ARKL1ΔS or FLAG-ARKL1sim with His-SUMO1, Myc-SUMO2 or Myc-SUMO3 in CNE2Z cells and stained them for FLAG, SUMO1 and Myc. Immunofluorescence microscopy revealed that, while both WT FLAG-ARKL1 and FLAG-ARKL1ΔS localized to SUMO1, SUMO2 and SUMO3 in
distinct nuclear foci, FLAG-ARKL1sim showed diffused nuclear localization without forming nuclear foci (Figure 11B). Another way I tested the effect of mutating the ARKL1 putative SIMs on SUMO binding was to transfecf HEK293T cells with either WT FLAG-ARKL1 or FLAG-ARKL1sim along with Myc-SUMO3 and perform a FLAG-IP. The SUMO3 monomer was recovered with WT FLAG-ARKL1 but not with FLAG-ARKL1sim suggesting that mutation of the ARKL1 SIMs prevents binding to SUMO3 (Figure 11C). The data as a whole indicate that the sequences mutated in ARKL1sim are functional SIMs.

Since CBX4 is known to be SUMOylated (Merrill et al. 2010), I postulated that ARKL1 localization to CBX4 may be dependent upon its SIMs. Consistent with its failure to localize to SUMO, FLAG-ARKL1sim also failed to localize to Myc-CBX4 in distinct nuclear foci, unlike WT FLAG-ARKL1 and FLAG-ARKL1ΔS (Figure 12). This suggests that ARKL1 interacts with CBX4 (or associated proteins) through SUMO-SIM interactions.
Figure 11. Mutation of SIMs in ARKL1 disrupts SUMO binding. A) Schematic representation of ARKL1 structure. The sequence of ARKL1 SUMO interacting motifs (SIMs) and corresponding mutations in the ARKL1sim mutant are shown in boxes. B) FLAG-ARKL1, FLAG-ARKL1ΔS or FLAG-ARKL1sim was overexpressed with His-SUMO1, Myc-SUMO2 or Myc-SUMO3 and cells were harvested 24 hours post transfection and processed for immunofluorescence microscopy by staining for ARKL1 and SUMO1 or Myc. C) FLAG-ARKL1 or FLAG-ARKL1sim was overexpressed with Myc-SUMO3 and a FLAG IP was performed. The blot was probed for FLAG and Myc and developed using the Amersham ECL reagent (GE healthcare).

Figure 12. ARKL1sim does not localize to CBX4. FLAG-ARKL1, FLAG-ARKL1ΔS or FLAG-ARKL1sim was overexpressed with Myc-CBX4 in CNE2Z cells. Cells were processed for immunofluorescence microscopy by staining for FLAG and Myc 24hrs post transfection.
3.7 ARKL1 Cellular Interactors Identified by Affinity Purification-Mass Spectrometry (AP-MS)

One of my objectives when I started this project was determining ARKL1 functions by identifying its cellular interactors. To this end, I created an adenovirus expression vector expressing C-terminally SPA tagged ARKL1 and used it to transduce AGS cells. This was done in order to maintain ARKL1 expression at physiologically relevant levels. Infected cells were harvested, a FLAG IP was performed and protein-protein interactions determined by AP-MS. A large fraction of the interactors identified in this experiment were proteins frequently identified as contaminants in mass spectrometry experiments as established by comparison with the contaminant repository of affinity purification (CRAPome) (Mellacheruvu et al. 2013). In addition, the known ARKL1 interactor CK2β was not recovered in this experiment. Based on this, I hypothesized that either the placement or the bulkiness of the SPA tag might interfere with some biologically relevant ARKL1 protein-protein interactions. Further support for this hypothesis was provided by the observation that, unlike N-terminally FLAG tagged ARKL1, C-terminally SPA tagged ARKL1 was unable to suppress EBV reactivation (data not shown). Therefore, the AP-MS experiment was repeated in duplicate by transient transfection of N-terminally FLAG tagged WT ARKL1 and FLAG-ARKL1sim followed by a FLAG IP. The proteins that showed at least one peptide in each replicate for both WT ARKL1 and ARKL1sim but no peptides with the empty plasmid control are summarized in Table 3. Also shown are the frequency with which each protein was recovered as a contaminant in mass spectrometry experiments in general (CRAPome) and the fold increase of the average spectral count of each protein over the average spectral count recovery in the CRAPome. No protein interactions were disrupted by the mutation of ARKL1 SIMs, however the interaction with USP7 was enhanced.
relative to recovery with WT ARKL1. The presence of the known ARKL1 interactor CK2β lends credence to the validity of this data. The transcription factors Jun and CREB1 were also recovered, the former belonging to the bZip family of proteins like BZLF1 (Angel and Karin 1991, Bartsch et al. 1998). Both these proteins have never been identified in previous mass spectrometry experiments using a FLAG IP. The other proteins that were identified in less than 25% of mass spectrometry experiments using a FLAG IP include the chromatin regulator proteins GRWD1 and DEK (Alexiadis et al. 2000, Higa et al. 2006), the prostrate cancer metastasis susceptibility protein GNL3 (Lee et al. 2015) and the DEAD box helicase DDX56. The other proteins that were recovered but are unlikely to be relevant interactors include the PSME3 proteasome activator sub-unit and the SRP9/14 proteins involved in the formation of stress granules (Berger et al. 2014). Finally, the deubiquitinating enzyme USP7 was also found to interact with ARKL1, which was confirmed in a coIP experiment (Figure 13). Since USP7 is known to stabilize several proteins by cleaving ubiquitin attached to them, it will be interesting to determine if USP7 also stabilizes ARKL1.
Table 3. Table Showing Interactors of ARKL1 Identified by Mass Spectrometry

<table>
<thead>
<tr>
<th>Protein identified</th>
<th>Empty plasmid Total spectral counts</th>
<th>ARKL1 Total spectral counts</th>
<th>Fold increase over AVG SC from CRAPome (AVG of both replicates)</th>
<th>ARKL1 sim Total spectral counts</th>
<th>Fold increase over AVG SC from CRAPome (AVG of both replicates)</th>
<th>Number of experiments (Found/total) from CRAPome using FLAG purification</th>
<th>Average SC from CRAPome</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1orf25/Arkadia like 1</td>
<td>0 0 510 642</td>
<td>N/A</td>
<td>576 655</td>
<td>N/A</td>
<td>0/156</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CSNK2A1/Casnine Kinase 2, Alpha 1 Polypeptide</td>
<td>0 0 309 208</td>
<td>89.1</td>
<td>285 192</td>
<td>82.2</td>
<td>31/156</td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td>CSNK2B/Casnine Kinase 2, Beta Polypeptide</td>
<td>0 0 251 104</td>
<td>110.9</td>
<td>176 103</td>
<td>87.2</td>
<td>5/156</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>CSNK2A2/Casnine Kinase 2, Alpha Prime Polypeptide</td>
<td>0 0 197 106</td>
<td>101</td>
<td>165 115</td>
<td>93.3</td>
<td>7/156</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>GNL3/Guanine Nucleotide Binding Protein-Like 3 (Nucleolar)</td>
<td>0 0 21 39</td>
<td>5.5</td>
<td>47 32</td>
<td>7.3</td>
<td>13/156</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>HSPH1/Heat Shock 105kDa/110kDa Protein 1</td>
<td>0 0 19 46</td>
<td>5.4</td>
<td>37 62</td>
<td>8.25</td>
<td>60/156</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>SRP9/Signal Recognition Particle 9kDa</td>
<td>0 0 16 11</td>
<td>7.9</td>
<td>10 13</td>
<td>6.8</td>
<td>3/156</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>PSME3/Proteasome activator subunit 3</td>
<td>0 0 15 42</td>
<td>6.5</td>
<td>27 59</td>
<td>9.8</td>
<td>38/156</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>GRWD1/Glutamate-Rich WD Repeat Containing 1</td>
<td>0 0 8 22</td>
<td>5.6</td>
<td>15 29</td>
<td>8.1</td>
<td>8/156</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>SRP14/Signal Recognition Particle 14kDa</td>
<td>0 0 8 17</td>
<td>4.8</td>
<td>19 19</td>
<td>7.3</td>
<td>3/156</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>JUN/jun proto-oncogene</td>
<td>0 0 7 9</td>
<td>N/A</td>
<td>11 11</td>
<td>N/A</td>
<td>0/156</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>DEK/DEK proto-oncogene</td>
<td>0 0 5 9</td>
<td>2.5</td>
<td>4 8</td>
<td>2.1</td>
<td>14/156</td>
<td>2.8</td>
<td></td>
</tr>
<tr>
<td>LUC7L3/LUC7-like 3 pre-mRNA splicing factor</td>
<td>0 0 3 13</td>
<td>1.2</td>
<td>13 14</td>
<td>2.1</td>
<td>81/156</td>
<td>6.6</td>
<td></td>
</tr>
<tr>
<td>RCN2/Reticulocalbin 2</td>
<td>0 0 2 7</td>
<td>1.1</td>
<td>4 6</td>
<td>1.25</td>
<td>89/156</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>DDX56/Dead box helicase 56</td>
<td>0 0 2 4</td>
<td>1.6</td>
<td>6 7</td>
<td>3.4</td>
<td>8/156</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>RBM39/ RNA binding motif protein 39</td>
<td>0 0 1 8</td>
<td>0.4</td>
<td>2 5</td>
<td>0.3</td>
<td>81/156</td>
<td>11.1</td>
<td></td>
</tr>
<tr>
<td>USP7/Ubiquitin-Specific-processing Protease 7</td>
<td>0 0 1 7</td>
<td>1</td>
<td>10 27</td>
<td>4.5</td>
<td>32/156</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>CREB1/CAMP Responsive Element Binding Protein 1</td>
<td>0 0 1 3</td>
<td>N/A</td>
<td>4 4</td>
<td>N/A</td>
<td>0/156</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

* The known interactor CK2β is highlighted
Figure 13. ARKL1 interacts with USP7. FLAG-ARKL1 was overexpressed with Myc-USP7 in HEK293T cells and a FLAG IP was performed. The Western blot was probed for FLAG, Myc and actin (loading control).
CHAPTER 4: DISCUSSION AND FUTURE DIRECTIONS

4.1 Examining the Role of ARKL1 in SnoN Degradation

ARKL1 is a partial paralogue of Arkadia and may be a product of a genome duplication event during evolution (Sun, Liu, and Hunter 2014). Closer examination of the ARKL1 sequence reveals that it lacks the Arkadia RING domain necessary for E3 ligase activity, raising the possibility that ARKL1 may function as a competitive inhibitor for binding to Arkadia substrates. Alternatively, ARKL1 may function in concert with other E3 ligases to target cellular proteins for proteasomal degradation. To examine these possibilities, ARKL1 or Arkadia was silenced in CNE2Z cells, cells were stimulated with TGFβ and SnoN levels were examined. SnoN is a negative regulator of TGFβ signalling that is ubiquitylated by Arkadia in response to TGFβ stimulation. While Arkadia silencing resulted in greater basal SnoN levels and lesser SnoN degradation upon TGFβ stimulation, ARKL1 silencing had no effect on basal SnoN levels or SnoN degradation upon TGFβ stimulation compared to the negative siRNA control. Therefore, unlike Arkadia, it seems that ARKL1 may not participate in TGFβ signalling. Consistent with this result, a study published during my experiments also found that ARKL1 had no effect on stimulating TGFβ responsive genes (Sun, Liu, and Hunter 2014).

4.2 Functional Significance of the ARKL1-CK2β Interaction

The CK2β KSSR motif was originally identified due to its ability to interact with the EBV protein EBNA1. Based on this, I postulated that ARKL1 and Arkadia might play a role in modulating EBV infection. Silencing of ARKL1 and Arkadia resulted in increased EBV reactivation, suggesting that both of these proteins might function to maintain EBV latency. It is worth mentioning that ARKL1 silencing using two different siRNAs resulted in varying degrees of reduction in Arkadia transcript levels. Whether this is a primary effect of nonspecific siRNA
targeting or a secondary effect of ARKL1 silencing remains unclear. If the latter is true, it raises the interesting possibility of ARKL1 regulating Arkadia transcription. Previous data from our laboratory indicated that CK2β silencing also results in increased EBV reactivation (Xinliu Li and Lori Frappier, unpublished). Based on this finding, I hypothesized that CK2β and ARKL1/Arkadia might function together to suppress EBV reactivation.

Overexpression of ARKL1 in AGS-EBV cells resulted in decreased EBV reactivation compared to untransfected cells on the same slide, while it is known from previous work in our lab that overexpression of FLAG-tagged proteins in general tends to increase EBV reactivation compared to untransfected cells on the same slide. Consistent with previous results, I found that overexpression of FLAG-DDX24 (which is not known to be related to EBV infection) increased EBV reactivation. To determine if the suppression of EBV reactivation by ARKL1 involves binding to CK2β, I also overexpressed the ARKL1ΔS mutant that was previously shown to be impaired for CK2β binding in AGS-EBV cells. Like WT ARKL1, ARKL1ΔS also suppressed EBV reactivation suggesting that ARKL1 may not require binding to CK2β in order to exert its effects on EBV infection, at least under conditions of overexpression. While the significance of the ARKL1-CK2β interaction remains unclear, ARKL1 may itself be a substrate of CK2, adding to the ever expanding list of CK2 substrates. In support of this possibility, there are 22 serine and 9 threonine residues that are predicted to be CK2 phosphorylation sites in ARKL1 (Kinase Phos 2.0 software). Alternatively, ARKL1 may act as a scaffold protein to bridge the interaction between CK2 and some of its substrates.

Since viral proteins frequently outcompete cellular proteins in their interactions, it is possible that EBNA1 may outcompete ARKL1 and Arkadia for CK2β binding. However, I was unable to demonstrate this possibility in a coIP experiment. The caveat of this experiment was that the IP
may not be truly reflective of EBV infection in vivo in terms of the relative expression of EBNA1 and ARKL1 within the cell. Alternatively, it is possible that EBNA1 may not have evolved to outcompete ARKL1 and/or Arkadia for binding to the KSSR motif but to recruit CK2 to the same cellular targets as ARKL1 and/or Arkadia in a manner beneficial for EBV infection.

4.3 ARKL1 Localization to CBX4

During the course of my experiments, I also found that ARKL1 localizes to the PRC1 component CBX4 in distinct nuclear foci. While only some of these foci corresponded to PML NBs (data not shown), the published literature suggests that others may be polycomb bodies (Mao, Zhang, and Spector 2011, Pirrotta and Li 2012, Sun, Liu, and Hunter 2014). The PRC is known to suppress lytic EBV infection by adding the H3K27me3 and other repressive epigenetic marks to lytic promoters. However, little is known about how the PRC is recruited to EBV episomes in vivo. It is plausible that a cellular factor may be responsible for PRC recruitment to EBV episomes. It would be interesting to explore the possibility that ARKL1 might be this factor. Thus far, I have not determined if ARKL1 actually interacts with PRC1 or other forms of CBX4.

CBX4 is an E3 SUMO ligase that is capable of SUMOyinating itself. Therefore, I rationalized that ARKL1 localization to CBX4 might be dependent upon non-covalent SUMO-SIM interactions. Examination of the ARKL1 sequence revealed two putative SIMs, one of which is predicted to be a high affinity VIDLT type SIM present in other SUMO binding proteins including Arkadia. Mutation of these putative SIMs resulted in disruption of ARKL1 foci and a pan nuclear staining, thereby compromising the ability of this mutant to associate with SUMO1, 2 and 3 foci. As expected, this ARKL1 mutant also failed to localize to CBX4. Work from Tony Hunter’s lab demonstrated that mutation of multiple Arkadia sequences including its SIMs results in similar disruption of Arkadia foci, although this mutant was still able to IP with CBX4. Much like
Arkadia, the ARKL1sim mutant was also able to IP with CBX4 (data not shown). Furthermore, Arkadia is known to function as a transcriptional repressor of approximately 200 cellular genes, possibly through its association with the PRC (Sun, Liu, and Hunter 2014). It is possible that ARKL1 may affect transcription of specific genes by playing a role in recruitment of transcriptional activators/repressors.

### 4.4 Identification of ARKL1 Cellular Interactors by AP-MS

Finally, in order to determine its cellular interactors, I performed AP-MS with ARKL1. An earlier attempt using a C-terminally SPA tagged ARKL1 (expressed at low levels from an adenovirus) only yielded nonspecific interactions, based on comparison with the CRAPome (Mellacheruvu et al. 2013). One reason for the lack of interactors might be that the ARKL1 C-terminus is important for ARKL1 interactions, and the presence of the bulky SPA tag might interfere with these interactions. In support of this theory, I found that, unlike the N-terminally FLAG-tagged ARKL1, overexpression of the C-terminally SPA-tagged ARKL1 did not repress EBV reactivation. Therefore, the AP-MS experiment was repeated with N-terminally FLAG tagged ARKL1, and the ARKL1sim mutant was also included, to determine which ARKL1 interactions were dependent upon its SIMs. The top interactors recovered included CK2α, CK2β and CK2α′. These interactions validate my data because our previous studies showed that ARKL1 interacts with CK2 tetramers through CK2β (Cao et al. 2014). No interactions were disrupted by mutations of the ARKL1 SIMs, indicating that SUMO binding is not critical for any of the interactions we detected in this assay.

Among the specific interactions identified were Jun and CREB1, members of the AP-1 (activator protein) family of transcription factors. Like BZLF1, the Jun subfamily of proteins (c-Jun, JunB and JunD) contain a basic leucine zipper domain and are capable of binding DNA sequences
called TPA response elements to trigger transcription of their target genes (Eferl and Wagner 2003). External stimuli such as growth factors and cytokines can trigger the mitogen-activated protein kinase (MAPK) cascades, which eventually leads to the phosphorylation of different transcription factors and activation of the genes for Jun and Fos (Hess, Angel, and Schorpp-Kistner 2004). Members of the Jun and Fos subfamilies are capable of forming homo and heterodimers using the leucine zipper region and bind to DNA through their basic regions (Kerppola and Curran 1991). In fact, FosB was also recovered with ARKL1 sim in both replicates and with WT ARKL1 in one replicate. A recent study also indicates that Jun-Fos, like BZLF1, can activate transcription through binding methylated CpG islands (Gustems et al. 2014). This is significant in light of the observations that the BZLF1 promoter is highly methylated during latent EBV infection and is activated by high levels of c-Jun in AGS cells (Feng et al. 2007). Like Jun, CREB1 also contains a basic leucine zipper domain and is known to activate the Zp promoter (Adamson et al. 2000). Combining these results with the localization of ARKL1 to CBX4, it is possible to form a working hypothesis about how ARKL1 suppresses EBV reactivation (Figure 14). ARKL1 might function to recruit the PRC to EBV lytic genes bound by Jun/CREB1, thereby interfering with transcriptional activation. Further experiments would help shed light on the functional consequences of the ARKL1-Jun and ARKL1-CREB1 interactions.
Figure 14. Model of how ARKL1 could suppress EBV reactivation. A) During latent infection, Jun binds to the ZII site within the Zp promoter of EBV to activate BZLF1 expression. ARKL1 may bind to the ZII bound Jun and recruit PRC1 to these sites. PRC1 recruitment would suppress EBV reactivation by ubiquitylation of H2AK119. B) During lytic infection, the ARKL1-Jun interaction may be disrupted, which would allow Jun to activate transcription from the Zp promoter. Adapted from Murata, Sato, et al. (2009).
Also enriched in my mass spectrometry experiment was the recently identified nucleolar protein nucleostemin (GNL3), which has been recognized as an important factor in protecting stem cells and progenitor cells from replication induced fork damage (Meng et al. 2013). Interestingly, GNL3 regulates recruitment of PML IV to DNA and SUMOylation of telomeric repeat binding factor (TRF1), thereby protecting against DNA damage (Hsu, Lin, and Tsai 2012). Considering the ability of ARKL1 to be modified by and associate with SUMO, it might be prudent to explore the role of SUMOylation in the ARKL1-GNL3 connection.

Also recovered in this experiment were the chromatin regulator proteins GRWD1 and DEK. GRWD1 is known to bind histones and facilitate loading of MCM onto chromatin (Sugimoto et al. 2015). It is also known to act as a substrate-specific adaptor for the CUL4-DDB1 E3 ubiquitin ligases (Higa et al. 2006). On the other hand, DEK is believed to be an anti-apoptotic protein that is involved in inducing positive supercoils in cellular DNA (Waldmann et al. 2002). Although most of its functions remain unknown, it is upregulated during viral infections. For example, DEK is upregulated during human papilloma virus 16 (HPV16) infection, by the action of the HPV E6 and E7 proteins (Wise-Draper et al. 2005). Intriguingly, DEK helps target the Kaposi’s sarcoma herpes virus latency-associated nuclear antigen (LANA) to cellular chromosomes (Krithivas et al. 2002). LANA is the KSHV functional equivalent of EBV EBNA1.

Finally, the deubiquitylating enzyme USP7, that is known to stabilize several cellular proteins, was also recovered. More USP7 peptides were recovered with the ARKL1sim mutant, suggesting that, while mutation of ARKL1 SIMs prevents its localization to specific nuclear bodies, it allows increased interaction with some of its protein interaction partners. The ARKL1-USP7 interaction was further verified in a co-IP experiment. Based on its known functions, it is possible that USP7 functions as a regulator of intracellular ARKL1 levels. In addition, USP7 is
known to deubiquitylate histone H2B, which results in gene silencing in both flies and mammals (van der Knaap et al. 2005, Sarkari et al. 2009). Interestingly, USP7 binds components of PRC1 and functions to reverse ubiquitylation of K119 on histone H2A that is catalysed by the RING1B component of PRC1 (Sanchez et al. 2007, Luo et al. 2015). In further studies, it would be interesting to explore the possible interplay between ARKL1, USP7 and PRC1 in mediating gene silencing. The other proteins that were recovered include HSPH1, SRP9/14, PSME3, LUC7L3, RCN, DDX56 and RBM39, all of which are proteins that are frequently identified in AP-MS experiments. SRP9/14 and PSME3 are proteins involved in the stress response and could be associated with ARKL1 overexpression induced stress. Therefore, most of these proteins are unlikely to be functional ARKL1 interactors.

It is worth mentioning that CBX4 was not recovered as an interactor of ARKL1 in the AP-MS experiments. This could represent a general lack of amenability of CBX4 peptides for mass spectrometry. In support of this argument is the fact that CBX4 has been recovered in only 8/411 mass spectrometry experiments, with an average of one spectral count per experiment. Alternatively, ARKL1 may interact with CBX4 transiently or indirectly. Finally, it is also possible that the PRC may not be bound by ARKL1, but the presence of SIMs allows ARKL1 to localize to highly SUMOylated foci, including CBX4 foci and PML NBs (shown by Jen Cao). However it is worth noting that I did recover the RING1 component of PRC1 in at least one replicate for both WT ARKL1 and ARKL1sim. RING1 was not shown in Table 3, since it was not consistently recovered with ARKL1.

In conclusion, I have identified several interactors of ARKL1 through AP-MS. Future research should be directed at determining the functional consequences of these protein-protein interactions. This will help reveal the cellular functions of ARKL1 and its regulation.
4.5 Future directions

My studies have identified ARKL1 as a protein that suppresses EBV reactivation. A study published during the course of my experiments demonstrated that influenza A virus infection causes increased ARKL1 SUMOylation and silencing of ARKL1 results in increased influenza A virus titres (Domingues et al. 2015). These observations raise the possibility that there is an interplay between ARKL1 SUMOylation and its ability to function as a restriction factor for influenza A virus infection. Therefore, it would be interesting to explore the possibility that ARKL1 SUMOylation might modulate its role in EBV infection. Thus, to determine if ARKL1 is differentially SUMOylated in EBV infected cells, I would transf ect AGS and AGS-EBV cells with plasmids expressing myc tagged SUMO1, SUMO2 and SUMO3 and perform an IP for endogenous ARKL1 in the presence and absence of lytic induction by NaB/TPA treatment. Following SDS-PAGE and Western blotting, I will probe for ARKL1. The presence of higher molecular weight ARKL1 bands might indicate the presence of ARKL1 SUMOylation and comparison between AGS cells and AGS-EBV cells would allow determination of effects of EBV infection upon ARKL1 SUMOylation. To confirm these results, I will also express His-tagged SUMO1, SUMO2 and SUMO3 in AGS and AGS-EBV cells along with FLAG-tagged ARKL1 and perform a His purification (de la Cruz-Herrera et al. 2014) after treatment of these cells by DMSO or NaB/TPA. As the His purification is performed under denaturing conditions, this would confirm that the detected higher molecular weight SUMO band is ARKL1 modified covalently by SUMO. If ARKL1 is found to be differentially SUMOylated in EBV infected cells, I would make a mutant of ARKL1 that lacks the consensus SUMOylation sequences based on the GPS-SUMO prediction software (Zhao et al. 2014), and test if overexpression of this
mutant could suppress EBV reactivation comparable to WT ARKL1 in AGS-EBV cells using immunofluorescence microscopy for BZLF1.

ARKL1 was also found to localize to the polycomb protein CBX4 in distinct nuclear foci (some of which could be PRC1 bodies). This localization was disrupted when ARKL1 SIMs were mutated, resulting in pan nuclear staining. Since the ARKL1sim mutant presumably does not localize to polycomb bodies which are known to suppress EBV reactivation, I would like to test the ability of this ARKL1sim mutant to suppress EBV reactivation. Briefly, I would transfect AGS-EBV cells with a plasmid expressing wild type FLAG-ARKL1 or FLAG-ARKL1sim and harvest these cells for immunofluorescence microscopy 24 hrs post-transfection. I would then stain these cells for BZLF1 and FLAG and count FLAG-positive and FLAG-negative BZLF1-positive cells. An inability of the ARKL1sim mutant to suppress EBV reactivation would indicate that ARKL1 SIMs are important for it to suppress EBV reactivation.

PRC components like Ezh2 have been shown to localize by chromatin immunoprecipitation (ChIP) to specific EBV lytic promoters during latent EBV infection. This colocalization correlates with the presence of the H3K27me3 repressive epigenetic mark on the promoters for EBV lytic genes (Murata et al. 2012, Murata and Tsurumi 2013). Since ARKL1 localizes to the PRC1 component CBX4 in distinct nuclear bodies, I would perform ChIP using an antibody against endogenous ARKL1 in AGS-EBV cells, followed by quantitative PCR for the promoters of BZLF1, BRLF1,BMRF1 and a region within the EBNA1 open reading frame (negative control), in the presence and absence of lytic induction with TPA/NaB. BRLF1 is an immediate-early EBV protein that functions with BZLF1 to turn on the expression of EBV lytic cycle genes (Holley-Guthrie et al. 1990). This would determine if ARKL1 localizes to specific EBV lytic promoters. Next, I would silence ARKL1 in AGS-EBV cells and perform ChIP using an
antibody for the PRC2 component Ezh2 (Murata et al. 2012) and the PRC1 component RING1 (Basu et al. 2014). This would be followed by quantitative PCR using primers for the promoters of BZLF1, BRLF1 and BMRF1. If the association of both the PRC components with EBV lytic promoters is found to be reduced upon ARKL1 silencing, it would indicate that ARKL1 might play a role in recruiting or maintaining PRC upon these promoters. If the association of PRC2 with these promoters remains unaffected whereas PRC1 association decreases, it would indicate ARKL1 affects PRC1 recruitment to EBV promoters without affecting PRC2. I would also repeat these ChIP experiments using an antibody against the H3K27me3 and H2AK119Ub as the PRC is involved in the deposition of these marks on the promoters of the genes to be repressed (Wang, Wang, et al. 2004, Sparmann and van Lohuizen 2006). All in all, these experiments would help elucidate the role of ARKL1 in shaping the epigenetic landscape on EBV lytic cycle promoters and its contribution to the maintenance of EBV latency.

In order to further confirm the identity of the ARKL1 nuclear foci, I would also transiently overexpress the PRC1 component RING1 with ARKL1 and visualize them by immunofluorescence microscopy. RING1 would be chosen for this experiment because AP-MS of ARKL1 resulted in recovery of RING1 in one replicate with WT ARKL1 and both replicates with ARKL1sim. If ARKL1 is found to localize to RING1 foci, it would further validate the localization of ARKL1 to PRC and indicate that ARKL1 may not require CBX4 for its PRC1 recruitment.

The CRISPR/Cas9 system has allowed large scale genome editing at multiple locations in the mammalian genome (Cong et al. 2013). In the long run, I propose using the CRISPR/Cas9 system to engineer AGS-EBV cells in which the ARKL1 genomic locus has been edited to create various ARKL1 mutants including the ARKL1sim mutant. Such an approach would allow
examining the effects of various ARKL1 mutants on EBV reactivation while avoiding the caveats of overexpression. First, I would start by using CRISPR/Cas9 to delete the gene for ARKL1 to recapitulate the effects of ARKL1 silencing on EBV reactivation (positive control). Next, I would make mutants of ARKL1 that lack functional SIMs, the consensus SUMOylation sequences or the CK2β binding sequence to examine their effects on EBV reactivation. As an example, if ARKL1 SIMs are required for its effects on EBV reactivation, mutating these SIMs would result in increased reactivation compared to cells with WT ARKL1. Using this technology to create different ARKL1 mutants would not only confer the above advantage, but would also allow us to explore the differences in localization of different ARKL1 mutants on the EBV genome.

Finally, I carried out AP-MS of ARKL1 to identify its cellular interactors. The bZIP protein family members Jun and CREB1 were identified as ARKL1 interactors in this experiment. I propose confirming these interactions by coIP experiments in AGS and AGS-EBV cells in the presence and absence of NaB/TPA treatment. The BZLF1 promoter (Zp) contains three cis-acting elements referred to as ZI, ZII and ZIII that are bound by transcriptional activators (Murata, Isomura, et al. 2009). Jun and CREB1 are known to activate BZLF1 transcription by binding to the ZII element (Murata, Sato, et al. 2009, Murata et al. 2011). To test whether ARKL1 interferes with the ability of Jun and CREB1 to activate transcription from Zp, I would co-transfect AGS cells with a plasmid expressing the Zp linked luciferase reporter and siRNA against ARKL1. If ARKL1 silencing increases luciferase activity compared to a control siRNA as predicted, I would co-transfect AGS cells with a plasmid expressing the luciferase reporter under the control of the Zp promoter in which the ZII site is mutated, and siRNA against ARKL1 (Murata et al. 2011). Since ARKL1 was found to bind to Jun and CREB1 in my AP-MS
experiment, I expect ARKL1 to act through the ZII site. Therefore, ARKL1 silencing in AGS cells transfected with the Zp linked luciferase reporter with a mutated ZII site would not result in increased luciferase activity compared to cells treated with a control siRNA. To confirm the results of my silencing experiments, I would co-transfect AGS cells with the above mentioned luciferase reporters and a plasmid expressing ARKL1. I predict ARKL1 would suppress luciferase activity only in the reporter with the WT Zp but not in the reporter with the mutated ZII site. This assay would help validate the hypothesis that ARKL1 might bind to Jun/CREB1 and prevent it from activating transcription from EBV lytic promoters, through recruitment of PRC to these sites.
REFERENCES


Corbau, R., V. Duverger, J. Rommelaere, and J. P. Nuesch. 2000. "Regulation of MVM NS1 by protein kinase C: impact of mutagenesis at consensus phosphorylation sites on


