Title
A role for host activation-induced cytidine deaminase in innate immune defense against herpesviruses

Permalink
https://escholarship.org/uc/item/15150708

Author
Bekerman, Elena

Publication Date
2013

Peer reviewed|Thesis/dissertation
A role for host activation-induced cytidine deaminase in innate immune defense against herpesviruses

By

Elena Bekerman

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Laurent Coscoy, Chair
Professor Gregory M. Barton
Professor Lin He
Professor Fenyong Liu

Spring 2013
Abstract

A role for host activation-induced cytidine deaminase in innate immune defense against herpesviruses

by

Elena Bekerman

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor Laurent Coscoy, Chair

Activation-induced cytidine deaminase (AID) is specifically induced in germinal center B cells to carry out somatic hypermutation and class-switch recombination, two processes responsible for antibody diversification. Because of its mutagenic potential, AID expression and activity are tightly regulated to minimize unwanted DNA damage. Surprisingly, AID expression has been observed ectopically during pathogenic infections. However, the function of AID outside of the germinal centers remains largely uncharacterized.

This dissertation demonstrates that infection of human primary naïve B cells with Kaposi’s sarcoma-associated herpesvirus (KSHV) rapidly induces AID expression in a cell intrinsic manner. We find that infected cells are marked for elimination by Natural Killer cells through upregulation of NKG2D ligands via the DNA-damage pathway, a pathway triggered by AID. Moreover, AID impinges directly on the viral fitness by inhibiting lytic reactivation without having a measurable effect on KSHV latency. We extend this analysis to the murine homologue of KSHV, MHV68 and find that AID mutates the viral genome at a rate that exceeds normal somatic mutation by several orders of magnitude. The tremendous mutational load accumulated by sequential passaging of MHV68 through AID-expressing cells leads to the eventual inactivation of the virus.

Importantly, we uncover two KSHV-encoded microRNAs that directly regulate AID abundance, further reinforcing the value of AID in the antiviral response. Together our findings reveal an additional role for AID in innate immune defense against herpesviruses with implications for a broader role in innate immunity to other pathogens.
Table of Contents

Abstract 1

Table of Contents i

List of Tables and Figures ii

Acknowledgments iii

Chapter 1 - Introduction
  ▪ Chapter 1.1: KSHV biology and pathogenesis 1
  ▪ Chapter 1.2: Activation-induced cytidine deaminase 5
  ▪ Chapter 1.3: Viral immune evasion 10

Chapter 2 - A role for host AID in innate immune defense against KSHV
  ▪ Introduction 13
  ▪ Results 14
  ▪ Discussion 20
  ▪ Materials and methods 22

Chapter 3 - KSHV-encoded microRNAs regulate host AID as means of immune evasion
  ▪ Introduction 25
  ▪ Results 26
  ▪ Discussion 32
  ▪ Materials and methods 33

Chapter 4 - Activation-induced cytidine deaminase inhibits MHV68 replication by hypermutating its genome
  ▪ Introduction 35
  ▪ Results 37
  ▪ Discussion 43
  ▪ Materials and methods 44

References 46

Appendix A - The impact of KSHV latent infection on host cell translation
  ▪ Introduction 56
  ▪ Results 59
  ▪ Discussion 66
  ▪ Materials and methods 68
  ▪ References 71
List of Figures and Tables

Chapter 1 - Introduction
Figure 1.1: Model of somatic hypermutation initiated by AID

Chapter 2 - A role for host AID in innate immune defense against KSHV
Figure 2.1: KSHV infection results in upregulation of AID in Primary Human Tonsillar B cells
Figure 2.2: KSHV infection leads to upregulation of NKG2D ligands
Figure 2.3: AID expression does not affect viral latency, but results in lytic reactivation defect
Figure S2.1: Additional lytic transcripts inhibited in AID-expressing BCBL-1 cells

Chapter 3 - KSHV-encoded microRNAs regulate host AID as means of immune evasion
Table 3.1: Bioinformatic Analysis of KSHV miRNAs binding to human AID
Figure 3.1: KSHV-encoded miRNAs directly target 3'UTR of hAID
Figure 3.2: miR-K12-5 and miR-K12-11 target full length AID for downregulation when expressed at physiological levels
Figure 3.3: KSHV miRNAs downregulate hAID in the context of the entire viral genome
Figure S3.1: KSHV infection does not dramatically upregulate expression of endogenous miRNA regulating AID

Chapter 4 - Activation-induced cytidine deaminase inhibits MHV68 replication by hypermutating its genome
Figure 4.1: MHV68 infection and germinal center formation in WT and IFNR-/- mice
Figure 4.2: AID-dependent mutation of the GFP reporter construct
Figure 4.3: Distribution of RGYW motifs across the MHV68 genome
Figure 4.4: Flowchart of MHV68 infection of AID-expressing cells experiment
Figure 4.5: List and distribution of AID-dependent mutations within MHV68 ORF58

Appendix A - The impact of KSHV latent infection on host cell translation
Figure A.1: Schematic representation of gradient encoding
Figure A.2: Titration of KSHV infection
Figure A.3: The translational response of HFF cells to KSHV infection
Figure A.4: Pairwise Correlation in translational profiles
Table A.1: Top 100 hits with greatest difference in translation between KSHV-infected and control samples (UV-inactivated KSHV)
Table A.2: Bioinformatic analysis of KSHV miRNAs binding to IFI6
Acknowledgements

I would like to acknowledge and express my gratitude to all the people who have guided and supported me in graduate school both personally and professionally.

First, I would like to acknowledge my graduate mentor Laurent Coscoy for his creative ideas, helpful discussions and a positive outlook. As a mentor Laurent is always guided by what is in the best interests of his students. He is one of the most approachable and caring individuals I know, and it has been an absolute pleasure being part of the Coscoy lab “family” over the last six years.

I have had the privilege to have Mark Schlissel, Britt Glaunsinger, Greg Barton, Lin He and Fenyong Liu serve on my thesis committee at some time during my graduate career. I am very thankful for their feedback and guidance in the process. I would especially like to acknowledge Greg Barton for his advice and support beyond our regular thesis meetings. Each one of these people serves an inspiration for my professional aspirations.

My past and present labmates, or Coscoy lab family members as I like to refer to them, have been absolutely instrumental in shaping me as a scientist and a friend. I owe special thanks to Andrea Pezda for her mentorship from my rotation and on. Veronica Anania has been and will remain a great colleague and a friend. Damian Trujillo and Maria Tokyuama are some of the most supportive and genuine colleagues and friends I have ever come across. Their bright ideas and technical expertise have made a tremendous contribution to my scientific accomplishments, especially in the last few years of graduate school. Trever Greene and Andrew Birnburg have been the most amazing and much needed addition to the Coscoy lab. In just a year’s time Andrew has become a friend that I feel like I’ve known for many years. His never-ending stream of ideas, thoughtfulness, and a sense of humor have made a tremendous impact on my graduate school experience. Finally, my undergraduate student Diana Jeon has offered me an absolutely invaluable help with the AID project. Her quick learning, dependability and friendly and quirky personality have made us a great team.

I would also like to acknowledge the friends I have gained in graduate school who stood by my side in times good or bad. Bettina has been one of my closest friends ever since we made a connection at the recruitment visit to Berkeley. Although, our personalities could not be more divergent in so many aspects, I know that the bond we share is one of the strongest I’ve formed in my life. Bettina is one of the most dependable and caring people I know. Lara and Brandon have been the partners in crime that helped maintain my sanity when science refused to be my friend. Patty has become my great and close friend in the last several years. Her strong convictions, unwavering support and fun demeanor have led to many fond memories of graduate school for me.

Next I would love to thank my boyfriend, Vlad for his love and support throughout the “tumultuous” years of graduate school. Although, he cannot fully appreciate the functional significance of cytidine deamination, he’s had to bear the witness to the most elaborate explanations of experimental fails and scientific frustrations. His faith and support have given me the inspiration to work harder and dreamer bigger.

And finally, I would like to thank my family (parents and Arnold) for their continued support and investment into my education. It seems like I have been a student for an eternity and I recognize that without their emotional and financial support much of this
would not have been possible. I hope that one day they will (deservedly) be able to reap the benefits of my success.
Chapter 1 - Introduction

Chapter 1.1: KSHV biology and pathogenesis

Opening remarks

Viruses belong to an extraordinarily diverse category of infectious agents and exist in virtually every ecosystem on Earth. Their absolute dependence on the living host cells for replication and the intimate co-evolution with their hosts have made the study of virology an integral part in the discovery of numerous basic biological principles. Moreover, viruses are now widely researched in clinical settings, where they serve as vectors for vaccinations and gene therapy and their gene products have important clinical applications. Advances in the understanding of viral transmission, life-cycle and anti-viral immune responses have revolutionized disease treatment and prevention. But despite such progress viral infections continue to be a leading cause of human disease. Viruses serve as etiological agents of human cancers, liver cirrhosis, immune deficiencies and other chronic diseases, many of which pose significant risk of morbidity and mortality in infected individuals. Given that knowledge, continued study of viruses and their interaction with the host is instrumental in advancing both our basic understanding of biology as well as development of new disease treatments.

Herpesviruses

Herpesviruses are a large family of animal DNA viruses that can be further subdivided into alpha, beta and gamma subfamilies. Alphaherpesviruses are neurotropic and include herpes simplex viruses, HSV1 and HSV2, and Varicella Zoster Virus (VZV), the causative agent of chicken pox. Leukotrophic betaherpesviruses include cytomegalovirus (CMV). Finally, gammaherpesviruses are lymphotropic and include two human viruses, Kaposi's sarcoma-associated herpesvirus (KSHV) and Epstein-Barr virus (EBV).

The name Herpes is derived from the Greek herpein meaning to creep and refers to the persistent, preferentially viral latent state within the host. All herpesviruses share four key components: their outer lipid bilayer envelope, inner tegument layer containing viral mRNAs and proteins, icosahedral protein capsid and the encased linear double-stranded DNA genome. Most human adults become infected with one or more species of viruses from this family and remain infected for life. Herpesviruses have co-evolved with their hosts for millions of years rendering them extremely well-adapted to life within the host. This quality makes Herpesviruses an excellent model system to investigate not only the viral adaptations but also the core immunological anti-viral host responses.

KSHV discovery and pathology

Kaposi’s sarcoma-associated herpesvirus, KSHV, also known as HHV8, is a large human double-stranded DNA virus. KSHV, along with closely related Epstein-Barr virus
(EBV) and murine homologue MHV68, belongs to a family of lymphotropic gamma-herpesviruses and exhibits strict host specificity. While largely asymptomatic in healthy individuals, KSHV is associated with a variety of malignancies in immunocompromised hosts.

KSHV was originally isolated from Kaposi’s sarcoma (KS) lesions and invariably linked to the neoplasm originally described by a Hungarian dermatologist in the nineteenth century. Named after its discoverer, KS is a papular skin and mucosal tumor characterized by spindle-shaped cells and aberrant vascular channels that give it the characteristic red hue. KS lesions generally arise as multiple distinct foci and are histologically complex bearing markers of endothelial, lymphatic and vascular origins. This heterogeneity is proposed to be a result of host cell reprogramming through viral gene expression.

KS was one of the earliest manifestations of the AIDS epidemic. The introduction of effective antiretroviral therapies in the developed world since the 1990s has led to a decrease in the incidence of KS. However, KS takes on its most aggressive form in AIDS patients and remains the leading cause of death in these and other immunosuppressed individuals. Moreover, KSHV infection is very prevalent in sub-Saharan Africa, where seropositivity rates exceed 50% and KS remains a growing public health concern.

In addition to KS, KSHV is linked to several kinds of B cell lymphoma. The primary effusion lymphoma (PEL) is a highly malignant non-Hodgkin lymphoma originating from post germinal center B cells. Unlike KS, PEL manifests itself as a pericardial effusion or tumor mass within the lymph nodes, lungs, or the gastrointestinal tract. Another KSHV-associated malignancy, the multicentric Castleman’s disease (MCD) is characterized by enlarged germinal centers with proliferating B cells and extensive vascularization. KSHV-associated germinotropic lymphoproliferative disorder (GLD) is a rare type of malignancy occurring in immunocompetent HIV-negative individuals. It occurs solely in the cases of co-infection with EBV and results in lymphadenopathy with aggregates of proliferating plasmablasts.

KSHV genome organization and life cycle

With over 80 open reading frames, 165-170 kilobase genome of KSHV consists of a central GC-low protein-coding region flanked by GC-rich terminal repeats. Terminal repeats contain binding sites for viral protein LANA and serve as the viral origin of replication. Genes conserved between KSHV and the prototypic rhadinovirus, HVS are designated as ORFs numbered from left to right. Non-homologous genes are named K1 through K15. Many of the viral genes bear homology to host genes, which is indicative of molecular piracy in the long course of co-evolution with the host.

Upon primary infection KSHV establishes a life-long persistence by preferentially entering a latent state. During latency viral genome is maintained as a multicopy chromatized episome tethered to host DNA, few genes are expressed and no virion progeny is generated. Thus, during latency infected cells survive because they are neither lysed by the egressing viral particles nor eliminated by the immune surveillance system, as the infection is mostly “invisible” to the host at this stage. Gene products detected at latency originate from a single polycistronic region and are generally limited to latency-associated
nuclear antigen (LANA) encoded by ORF73, viral cyclin D encoded by ORF72, vFLIP encoded by ORF71, Kaposin encoded by K12, and the viral microRNAs (see the next section). Of the four latent proteins, LANA is the minimally required factor for the maintenance of KSHV latency. It is a large heavily post-translationally modified protein necessary to tether viral episome to host chromosomes to ensure proper partitioning of the viral genome during cellular division. LANA serves as a common marker of latent infection and yields a punctate nuclear appearance upon immunofluorescent staining of the infected cells.

Under conditions of stress, however, the virus can reactivate to undergo lytic replicative program. Upon entry into the lytic life cycle, virtually all viral genes are activated in a temporal fashion (immediate early, early, and late); new virions are packaged and disseminated generally killing the host cell. In situ hybridization studies of KS biopsies demonstrated that the majority of tumor cells are positive for KSHV genome, however only one to three percent exist in a lytic state due to spontaneous reactivation\textsuperscript{11,12}. Reactivation can be stimulated in vitro via treatment of latently infected cells with either phorbol esters or sodium butyrate, or forced expression of transcription factor replication and transcription activator (RTA).

The immediate early genes are first to be expressed during the lytic cycle and do not require de novo protein synthesis. These proteins are responsible for activating downstream lytic genes and modulating the cell to create a more hospitable environment for replication. RTA is a classic immediate early gene and considered to be the master regulator of the latent-lytic switch for KSHV. Early genes require de novo protein synthesis, but are expressed before the viral genome replication. They play a role in nucleic acid metabolism and modulating host gene expression\textsuperscript{13}. Finally, late genes are expressed after viral replication and consist primarily of structural proteins required for virion assembly.

**KSHV microRNAs**

MicroRNAs (miRNAs) are a class of non-coding RNAs expressed by all plants and animals. They serve important functions in post-transcriptional regulation of gene expression. In 2004 EBV was the first virus to have been shown to encode and express miRNAs\textsuperscript{14}. Following that discovery at least 82 different miRNAs have since been identified in animal herpesviruses\textsuperscript{15}. KSHV encodes 12 pre-miRNAs, which can be further processed to yield at least 25 different mature miRNAs\textsuperscript{16,17}. All 12 pre-miRNAs are localized to the latency region. Specifically, miR-K1 through miR-K9 and miR-K11 belong to a cluster in the intron between K13/Orf71/vFLIP and K12/kaposin, while miR-K12-10 and miR-K12-11 reside in the coding region and the 3’UTR of K12/kaposin, respectively. All 12 miRNAs are expressed during the latency, and only two miRNAs, miR-K12-10 and miR-K12-12 increase significantly upon lytic reactivation\textsuperscript{18}. While miRNAs may play some role during the lytic life cycle, expression of viral miRNAs during latency presumably allows viruses to effectively modulate both self and host gene expression without generating additional peptides, which may put the virus at risk for immune detection.

Like cellular miRNAs, viral miRNAs are transcribed by PolII, processed into hairpin structures, and subsequently into double stranded duplexes, where one or both of the
strands eventually gets incorporated into the RISC complex. Via an imperfect base pairing these miRNAs function to inhibit protein translation and/or degrade mRNA messages. Factors reported to affect the strength and outcome of miRNA interaction with its target include AU composition, base pairing at the seed region and 3’ of miRNA, and distance to either end of the 3’UTR of the target. All of these parameters are now incorporated into various computational models used to predict miRNA targets. Interestingly, while the majority of KSHV-encoded miRNAs share no sequence homology with the host miRNAs, miR-K12-11 is an ortholog of mammalian miR-155 with perfect homology of the entire seed region.

**KSHV Transmission and laboratory models of KSHV infection**

The most well established route of infection is horizontal transmission by saliva. Evidence suggests that this is the most common route of transmission both in people from endemic regions and within the high risk groups in the Western world. The principal source of viral replication in vivo is the oropharynx, where tonsillar and lymph node tissues as well as epithelial cells harbor the virus. Shedding of the virus from these tissue deposits it directly into the saliva. Additional routes include mother-to-child vertical transmission, sexual routes, and via blood or organ transplants. While KSHV may infect a variety of cell types, data from patients exhibiting KSHV-related pathologies as well as in vivo experiments with murine γ-herpesviruses model MHV-68 suggest that latency is established primarily in the germinal center B cells and maintained long-term in the germinal center and memory B cells (see chapter 1.2 for details).

Given strict host specificity of KSHV, researchers working to characterize this virus have always lacked an adequate animal model. Much work on gammaherpesviruses has thus been focused on the murine homologue, MHV68. But despite significant homology between human and murine viruses, findings obtained through mouse models of infection cannot always be safely extrapolated to make conclusions about KSHV biology within a human host. Moreover, in vitro the study of KSHV has always faced a number of technical limitations. Due to the inability to efficiently infect B cells in tissue culture, host responses to KSHV have historically been studied in B cell lines derived from KSHV-infected lymphoma patients or through de novo infection of cells of endothelial origin. Differences in tissue-specific gene expression between endothelial cell lineages and B lymphocytes and epigenetic changes imprinted onto transformed B cells have prevented us from thoroughly characterizing host responses within cells that serve as primary virus reservoir in vivo. More recently, Ganem lab has developed an in vitro system that allows for efficient infection of not only immortalized B cell lines, but also a percentage of primary human B cells. This system relies on co-culture of reactivated infected endothelial cell line with cells of interest which leads to cell-to-cell spread of infectious virions. This innovation has opened up new avenues for understanding KSHV infections under the more physiologically relevant conditions. Relying on this set up has made it possible to infect primary human tonsillar B cells and uncover novel host immune responses against this pathogen detailed in the later chapters.
Chapter 1.2: Activation-induced cytidine deaminase

B cells

B cells are specialized immune cells that play a critical role in the adaptive immune response. Their main functions are to produce antibodies, to present antigens to the other immune cells and to generate immune memory responsible for the rapid response to recurring infections. B cells are formed and undergo maturation within the bone marrow. They subsequently migrate to the secondary lymphoid organs such as the spleen or the lymph node or are found circulating in the blood and the lymph system. B cell maturation involves the intricate process of B cell receptor (BCR) assembly, which in its secreted form is known as the antibody. Components necessary to make up a full BCR are two light and two heavy chains, each composed of the variable and the constant regions. Furthermore, each variable region is made up of multiple gene segments, “V”, “D” and “J” in the heavy chain and segments “V” and “J” in the light chain. The pool of different gene segments combined with junctional diversity and the variable pairing of the light and heavy chains gives rise to millions of possible mature receptors. The combinatorial diversity achieved by these means bypasses the need for the germline to specifically encode all of the individual immunoglobulin genes necessary to recognize the plethora of existing pathogens.

Germinal Centers

Germinal centers (GC) are transient compartments within the secondary lymphoid organs which form upon antigenic stimulation of B cells. Anatomical organization of these organs allows for the diverse population of partially activated B cells to interact with T cells capable of recognizing epitopes of the same antigen. At this B-T cell boarder select B cells receive further differentiation signals that trigger their migration toward the follicular dendritic cell (FDC) network where the GC reactions are initiated. Here, following a rapid proliferative burst B cells undergo somatic hypermutation (SHM) to further diversify the antibody repertoire (see next section). The resulting B cells then compete for the antigen presented by the FDCs and the help of the T-follicular helper cells. Signals stemming from these interactions determine whether the B cell goes on to form an antibody secreting factory, the plasmablast or becomes the bearer of immunological memory stored away until the next insult. An additional process that takes place within the GCs is class-switch recombination (CSR). During this reaction constant regions of the antibody-coding genes are recombined to alter the antibody isotype. Each antibody isotype has different biochemical properties and offers unique binding sites to the downstream interacting partners. The nature of the infection determines the needs for one isotype over the other. Production of B cell-stimulating cytokines by the other immune cells dictates the type of recombination. Both SHM and CSR are initiated by a single enzyme known as activation-induced cytidine deaminase, AID. Patients lacking AID exhibit hyper-IgM syndrome characterized by the dramatic decrease or complete absence of antibody isotypes other than IgM.
Somatic Hypermutation

Somatic hypermutation (SHM) is a process by which random mutations are introduced into the antibody-coding genes in an effort to improve affinity for the antigen. The overall process of mutagenesis and selection is referred to as affinity maturation. Most mutations are single nucleotide substitutions that occur between 100-200 bp from the transcriptional start site and about 2kb downstream \(^{31,32}\). Mutation frequency is highest within the V(D)J coding region, while crucial regulatory elements such as the promoter, intronic enhancer and the constant region are generally left unmutated \(^{33}\). Cytidines localized within the hotspot WRCY motifs (where W = A/T, R = A/G, and Y = C/T) are preferentially deaminated, yet other sequences are also susceptible to AID-dependent deamination \(^{34,35}\). Active transcription is a requirement for the gene to be accessible to AID, although transcription alone is not sufficient to ensure mutagenesis. Accordingly, mutational rates correlate directly with the rate of transcription \(^{36,37}\).

The process of SHM can be divided into two phases: AID-dependent deamination and error-prone repair of the deaminated motifs. Without DNA repair, replication across the mismatched U/G base pair yields a C-to-T transition mutation in one of the daughter cells. Alternatively, the mismatched deoxyuracil may be recognized and excised by uracil-N-glycosylase (UNG). UNG is a ubiquitously expressed enzyme that hydrolyzes the N-glycosidic bond of deoxyuridine in DNA. This catalysis takes place only when UNG achieves a proper conformation by associating only with deoxyuridine, and not with any of the normal deoxynucleotides \(^{38}\). Conversely, a larger fragment of the uracil-bearing strand may be excised by the mismatch repair machinery (MMR). The resulting gap can then be filled in by an error-prone polymerase generating both transition and transversion mutations \(^{39}\) (Figure 1). This mode of action clarifies why mutations are possible at the sites other than cytidines. If fact, certain types of error-prone polymerases preferentially incorporate mutations across adenine or thymine giving rise to a larger fraction of mutations at the A/T versus C/G base pairs \(^{40}\). Of note, ssDNA breaks occurring during SHM are the precursors to the dsDNA breaks required for CSR. Mismatch repair proteins such as MSH2 and MSH6 are specifically responsible for inducing double strand breaks following the excision of deoxyuracil by UNG.

AID’s substrate is ssDNA, which supports the observation that active gene transcription is required to initiate deamination. The rate of somatic hypermutation within the variable region of immunoglobulin gene during SHM is roughly six orders of magnitude greater than the rate of basal somatic mutation. However, while AID is preferentially targeted to immunoglobulin loci during SHM other regions of the genome are also susceptible to deamination, albeit to lower degree. Accordingly, AID-dependent mutations and chromosomal translocations within proto-oncogenes such as BCL-6, FAS and MYC contribute directly to malignant transformations \(^{41}\).

AID/APOBEC family of deaminases

AID belongs to a larger family of vertebrate RNA/DNA editing tissue restricted enzymes, which include APOBEC1, APOBEC2, APOBEC3G, APOBEC3A–H, and APOBEC4.
Like the rest of APOBEC family members AID contains a conserved zinc-binding motif required for conversion of a cytidine to uridine via the hydrolytic substitution of an amine group. Other zinc-dependent deaminase families widely expressed in metazoans are the cytidine deaminases (CDA), the dCMP deaminases (DCDT) or the tRNA adenosine deaminases (Tad/ADAT2). CDAs and DCDTs act on free pyrimidines in the salvage pathway, the Tad/ADAT2s edit adenosine to inosine at the anticodon of various tRNAs and are essential in bacteria, yeast and metazoans. Phylogenetics, structure and function features favor the tRNA-editing enzymes as the origin of the AID/APOBECs. This model is further supported by the observation that ADAT2 from trypanosomes can deaminate DNA.

The rise of the AID/Apopec gene family appears to have been concurrent with the emergence of the vertebrate lineage and the evolution of adaptive immunity. In primates (along with a number of other species), APOBEC3 genes have recently undergone a rapid expansion through gene duplication. Such rapid evolution is believed to be a result of selective pressure from the retroviruses and retrotransposons targeted by these genes.

Apobec1 plays a role in lipid metabolism by generating a premature stop codon, and hence, a shorter version of the ApoB mRNA with an alternative function. ApoB100 (full-length) and apoB48 (truncated) are used to transport cholesterol and triglyceride, respectively, in the blood. Apobec 2 is expressed specifically in the skeletal muscle and the heart, where it regulates TGFβ signaling necessary to specify the left-right axis in Xenopus and zebrafish embryos. Its deficiency leads to a shift in muscle fiber type, diminished body mass, and myopathy. Apobec4 is the least characterized member of the family with low sequence similarity to the other AID/APOBECs. It has been suggested to function by editing an RNA involved in spermatogenesis, however its deamination capability remains to be formally proven. Apobec3 family possesses important antiviral activity against HIV & other viruses (discussed in more detail in the next section).

Finally, AID itself was originally discovered 14 years ago using subtractive hybridization between stimulated and unstimulated lymphocytes. Its role in antibody diversification has since been a subject of extensive research. Beyond immunoglobulin genes AID has been shown induce mutations in roughly twenty-five percent of cellular genes. However, even that statistic is likely an underestimation. AID-dependent mutagenesis extent nearly doubles upon knocking out either UNG or MMR machinery, suggesting that under wild type conditions the cell is able to correct many of the deaminated bases. Curiously, AID expression has also been observed in progenitor germ cells, oocytes and early embryos suggesting a possible role development. Multiple reports now implicate AID in demethylation. This activity was shown to initiate nuclear reprogramming towards pluripotency in human somatic cells. Specifically, AID was shown to binds silent methylated OCT4 and NANOG promoters in fibroblasts, but not active demethylated promoters in ES cells.

Role of cytidine deaminases in innate immunity

For nearly a decade now, the scientific community has appreciated the role of cytidine deamination in anti-viral immunity. In a key study by Malim and colleagues APOBEC3G was found to inactivate HIV by introducing a large number of mutations in the
minus strand of viral cDNA. In addition to blocking infection by inactivating the provirus, APOBEC3G has been shown to disrupt the HIV-1 reverse transcription and impair the integration of the provirus. APOBEC3 enzymes with double catalytic domains, Apo3G, Apo3F, Apo3B, and Apo3DE, most efficiently inhibit HIV-1 replication, while those with single domains, Apo3A and Apo3C, are less effective at doing so. Beyond HIV APOBEC3 enzymes have been shown to act on other retroviruses including simian immunodeficiency virus (SIV), murine leukemia virus (MLV) and human T-cell leukemia virus type 1 (HTLV-1) and adeno-associated virus, which replicates as ssDNA. This work has since been extended to other classes of RNA viruses such as HBV and non-enveloped DNA viruses such as HPV.

Recently, in addition to its role in the humoral adaptive immunity AID has also been implicated in the innate immune defense. A group has demonstrated the ability of AID to inhibit retrotransposition of non-LTR transposon, L1 via a deamination-independent mechanism. This work proposed a role for cytoplasmic AID, where it engages assembling L1 replication complexes either co- or post-translationally. Moreover, AID has been shown to be induced in response to an infection with Abelson murine leukemia virus (Ab-MLV). This induction resulted in activation of checkpoint kinase-1 (chk1) ultimately leading to restriction in proliferation of infected cells and upregulation of NKG2D ligands to alert the innate immune system of the infection.

In parallel with these findings, there is a growing abundance of literature suggesting that AID is induced specifically upon viral infection of cells both within and outside of germinal centers. Namely, T-cell leukemia virus (HTLV-1) has been shown to activate AID in PBMCs and T-cell lines, Ab-MLV – in primary bone marrow cells, Hepatitis C virus (HCV) – in PBMCs and Raji cells, and EBV – in transformed PBMCs. Not surprisingly, there have also been reports of viral interference with deaminases. In the case of HIV, an essential protein, Vif is required to target APOBEC3G for ubiquitin-dependent degradation, thereby excluding APOBEC3G from the virions. Similarly, to counterbalance induction of AID, EBV latent protein EBNA2 acts as a dominant negative regulator of its expression. Given that AID activity is dosage dependent, this modulation by EBNA2 abrogates hypermutation phenotype when compared to EBNA2-negative cells.

**Endogenous regulation of AID**

Because of AID’s ability to induce mutations and chromosomal translocations, and potentially lead to cancer, its abundance and activity must be tightly regulated. AID’s transcription is regulated via a number of regulatory elements, which map both upstream and downstream of the open reading frame. These regions contain conserved binding sites for at least 19 different transcription factors with both activator and repressor functions. AID can also be alternatively spliced, whereby different splice products vary in their activity toward either SHM or CSR. Additionally, AID protein contains several phosphorylation sites which either promote or inhibit its function. Phosphorylation of Ser38, for example, is required for AID to associate with replication protein A, which in turns augments it binding of DNA and activity. Nuclear export and cytoplasm retention signals regulate nuclear-cytoplasmic transport of AID. Majority of AID protein is
sequestered in the cytoplasm, with only 10-15% shuttling into the nucleus where it functions 69. AID is also regulated via ubiquitination, which shortens nuclear AID half-life relative to its cytoplasmic counterpart 70. Finally, AID abundance is modulated by at least three distinct endogenous miRNAs: miR-155, miR-181b, and miR-93 71-74. Both in vitro and in vivo studies suggest that these miRNAs function to keep AID levels low in resting B cells and/or contribute to shutting down AID expression following B cell activation.

miR-155 is a small non-coding RNA processed from its precursor B cell integration cluster (Bic). Bic was originally identified as a gene activated by proviral insertions in avian leukemia virus-induced lymphomas and suspected to function through its non-coding RNA 75. Subsequent deletion of miR-155 in mice revealed an important contribution it makes to the various aspects of vertebrate immunity likely through the regulation of multitude of target genes 76. Following that discovery, two groups independently verified the role for miR-155 in direct regulation of AID. Papavasiliou group specifically ablated only the site of interaction between miR-155 and AID 74. They noted ectopic persistence of AID upon exit of B cells from the GCs. Moreover, this lack of AID regulation resulted in defective affinity maturation of the BCR. Using a similar approach Nussenzweig group concluded that miR-155 normally functions to suppress the half-life of AID mRNA 73. Lack of this activity yields a high degree of Myc-Igh translocations in their mouse model.
Chapter 1.3: Viral immune evasion

Opening remarks

Multiple arms of the host immunity mediate the responses against KSHV. Among them are type-I interferon production, NK cell cytotoxic activity, CD4+ T cell and CTL responses and neutralizing antibody production. Not surprisingly, KSHV has evolved to be well equipped in dealing with these host defenses. Over a quarter of all KSHV-encoded genes has been reported to modulate the host immune system. This ability to establish a fine equilibrium with the host by evading various attempts to prevent survival and spread of the virus allows KSHV to persist lifelong. Immune evasion strategies employed by KSHV affect both the innate and the adaptive arms of the immune system. A large fraction of these bears close homology to the host regulatory factors underscoring the virus’ success in pirating host genes for gained advantage. The mechanisms of immune evasion include interference with intercellular and intracellular defenses, inhibition of apoptosis and cell-to-cell interactions.

Evasion of the innate defenses

In regulating the innate responses virally encoded cytokines and chemokines can either compete for binding with host molecules or antagonize their respective receptors to prevent recruitment of leukocytes and to promote hematopoiesis. KSHV complement control proteins (KCPs) inhibit the complement-mediated lysis of infected cells. Host interferon (IFN) signaling is absolutely central to the first line of anti-viral defense. For that reason, KSHV encodes a whole cluster of IFN regulatory proteins to limit the expression of type I IFN as well as the downstream IFN-inducible genes. KSHV latent protein vFLIP and lytic vBcl-2 are known to prevent cellular death by apoptosis. Finally, KSHV hampers innate responses by tampering with Toll-like receptor signaling through downregulation of key components of this pathway.

Evasion of the adaptive defenses

Presentation of viral peptides on the surface MHC molecules alerts host CTLs to the infection and leads to the direct lysis of infected cells. KSHV encodes two E3 ubiquitin ligases that mark MHC as well as other important immune receptors for ubiquitin-dependent degradation. Such low MHC-expressing cells are rendered sensitive to the recognition by the NK cells, which upon activation secrete large amounts of gamma IFN and lyse infected cells. In an effort to get around this response, same viral ligases also downregulate gamma IFN receptor to limit further immune activation. Another key component of the adaptive immune defense is co-stimulation. Professional antigen-presenting cells engage with the T lymphocytes for their robust activation. KSHV inhibits this interaction through downregulation of stimulatory receptors and ligands, as well as adhesion molecules that mediate cell-to-cell interactions.
Immune evasion via KSHV miRNAs

KSHV encoded miRNAs have been the subject of great interest within the last several years in the field of virology. Researchers have taken both genome-wide and more targeted approaches to identify the genes regulated by these miRNAs. Data suggest that the KSHV miRNA cluster can regulate both viral and host gene expression. In light of their expression during latency, it is not surprising that KSHV miRNAs have been implicated in promoting viral latency. Specifically, miR-K12-9 has been shown to repress RTA, the master regulator of the lytic switch. With the exception of a couple cellular homologues, KSHV miRNAs sequences are unique in comparison with the other viral miRNAs. Interestingly, there is at least one common host target for miRNAs of several different viruses. Namely, MICB, an important stimulatory ligand upregulated during viral infections and used to trigger Natural Killer cell activation has been shown to be the target of KSHV miR-K12-7. Moreover, unique miRNAs encoded by EBV and HCMV have also been shown to target the same mRNA. It is plausible that in the cases of co-infection (co-infection of KSHV and EBV is a relatively common occurrence) these distinct miRNAs synergize to achieve a more robust inhibition of their targets.

Evasion of cell cycle arrest and apoptosis is an important feature of persistent viruses. In addition to viral proteins, KSHV-encoded miRNAs have now been shown to regulate these pathways. miR-K12-1 was demonstrated to target cyclin-dependent kinase inhibitor, p21. Through this action, miR-K12-1 was able to attenuate the cell cycle arrest and likely contribute not only to the viral survival but also to promoting oncogenic transformation. KSHV miR-K10 was shown to repress TWEAKR allowing for the escape from TWEAK-induced apoptosis. Moreover, miRNAs miR-K1, miR-K3, and miR-K4-3p were reported to contribute to the evasion of apoptosis by targeting effector caspase-3.

Finally, KSHV miRNAs have been shown to regulate aspects of the innate immune signaling by targeting IRAK1 and MYD88. Inhibition of these molecules led to a reduction of inflammatory cytokines. Since miRNAs generally bind their target mRNAs via imperfect base pairing, a given miRNA has the capacity to regulate tens if not hundreds of different genes. Continued effort to decipher the functions of KSHV miRNAs is likely to reveal many more host targets with critical roles in immune defense.
Figure 1.1. Model of somatic hypermutation initiated by AID (Adapted from Peled, J.U. et al. (2008). Annu. Rev. Immunol.) AID deaminates a cytidine residue, creating a uridine:guanosine (U:G) mismatch that is resolved by several pathways that may compete with one another. AID deaminates single-stranded DNA formed during transcription of both strands of the DNA (not shown). The general replication machinery can interpret the U as if it were a deoxythymidine (T). One of the daughter cells will acquire a C-to-T transition mutation. (Center) UNG can remove the uracil, leaving behind an abasic site. Short-patch base excision repair (BER) can fill the gap with error-prone polymerases, which can insert any nucleotide in place of the U, leading to transitions and transversions at G:C bases. (Right) Mismatch repair (MMR) can recognize the U:G mismatch. The U-bearing strand is excised and, at loci that undergo SHM, monoubiquitylated PCNA (proliferating cell nuclear antigen) recruits error-prone polymerases to fill the gap, leading to transition and transversion mutations at A:T bases as well as at neighboring G:C bases. (Dashed line) Long-patch BER can also be a source of mutations at A:T bases and may compete with MMR.
Chapter 2 - A role for host AID in innate immune defense against KSHV

Introduction

Multiple arms of host immunity mount responses against infectious viruses and are counteracted by numerous lines of viral defense. In particular, the humoral response is instrumental in controlling viral amplification through the action of neutralizing antibodies and antibody-dependent cell mediated cytotoxicity. Expression of activation-induced cytidine deaminase (AID) in germinal center (GC) B cells is crucial for generating high affinity antibodies during the adaptive immune response. This enzyme functions to deaminate cytidine residues within immunoglobulin genes. This activity is required for class-switch recombination (CSR) and somatic hypermutation (SHM), which together contribute to the diversification of the antibody repertoire as well as increase antibody affinity 29. While AID is preferentially targeted to immunoglobulin loci, other regions of the genome are also susceptible to deamination, albeit to lower degree 41. Accordingly, AID-dependent mutations and chromosomal translocations within proto-oncogenes such as BCL-6, FAS and MYC contribute directly to malignant transformations 91,92.

AID belongs to a larger family of tissue restricted vertebrate RNA/DNA editing enzymes, which include APOBEC1 through APOBEC4. APOBEC3 genes confer innate immunity to a wide range of retroviruses and help prevent transposition of endogenous transposable elements capable of disrupting host genome integrity 93. More recently, AID has also been implicated in similar defenses for its ability to restrict retrotransposons as well as the transforming retrovirus Ab-MLV in a mouse model 59. In parallel with these findings, a growing body of literature documents AID induction outside of GCs in response to a variety of pathogens, further supporting the notion that AID may serve dual functions in both adaptive and innate immunity 60-62,94.

In this chapter, we set out to examine AID expression in the context of an infection with a prevalent human viral pathogen, KSHV. Moreover, we explored whether AID can negatively impact KSHV fitness as an innate immune defense strategy. Our findings revealed a consistently rapid upregulation of AID within primary human B cells in response to KSHV infection. AID expression continued to rise throughout the course of the infection in our primary cell culture. Additionally, we observed upregulation of activating natural killer (NK) cell ligands of the NKG2D family on the surface of infected cells. Consistent with AID’s ability to induce DNA-damage, we found ligand induction to be dependent on the DNA-damage pathway. We also addressed KSHV latency and reactivation potential in AID expressing cells, and found that prolonged exposure to AID significantly inhibits initiation of the lytic replication program of this virus. Together our data elucidate a critical role for AID in innate immune defense against KSHV and may also suggest a key role for AID in defense against other viral infections.
Results

KSHV infection results in upregulation of AID in primary human tonsillar B cells

To assess expression of AID upon KSHV infection of primary B cells, we took advantage of a recently developed co-culture system, where B cells are infected via direct contact with reactivated iSLK.129 cells. These cells harbor latent KSHV marked by constitutive GFP expression and induction of RFP expression upon lytic reactivation \(^95\). Co-culture of primary human tonsillar cells with reactivated iSLK.129 cells gave rise to a reproducible population of infected B cells as measured by GFP expression (Figure 2.1A), with a gradual increase in the fraction of GFP-expressing cells over time. While KSHV was detectable in T cells in the tonsillar cell co-culture, T-cell infection has previously been shown to be abortive \(^96\). We therefore focused our studies on the physiologically relevant CD19+ B cell population.

KSHV infection resulted in a modest upregulation of AID within the GFP positive fraction of B cells on day 2 post-infection (Figure 2.1B). This expression continued to increase in infected cells as compared to uninfected cells during the course of the infection (Figure 2.1B). AID expression in the KSHV infected population remained upregulated as late as 7 days post-infection, which marks the limit of the tonsillar cell survival in our cytokine-free culture. To investigate whether AID upregulation was initiated at the transcript level, infected GFP positive tonsillar B cells were sorted and AID mRNA levels were determined by qPCR. As seen in Figure 2.1C, infection resulted in a significant increase of host AID transcript as compared to uninfected cells. While we cannot rule out transcript stabilization to explain this result, previous work supports the role for enhancers and silencers as the primary means of AID regulation in activated B cells \(^97\).

AID expression in GC B cells is normally stimulated by synergistic actions of IL-4 and CD40 ligand, which lead to the activation of JAK/STAT and NF-kB pathways, respectively. Given that latent infection of tonsillar B cells, as determined by GFP positive, RFP negative signal via flow cytometry, is sufficient to upregulate AID (Figure 2.1B) we set out to identify whether expression of a particular latent viral gene may be adequate to induce AID expression. KSHV latent protein vFLIP is an established constitutive activator of NF-kB raising the possibility for its role in AID upregulation \(^98\). To test this hypothesis we transiently transfected primary B cells with vFLIP, and uncovered that relative to vector control vFLIP was able to induce AID expression in the absence of other viral gene products at 48hrs post transfection (Figure 2.1D). This result demonstrates that vFLIP alone is sufficient to turn on cellular pathways responsible for AID expression. Together our results indicate that KSHV infection leads to an early transcriptional upregulation of host AID in a cell intrinsic manner, and that this expression is sustained beyond the early phase of infection.

KSHV infection leads to upregulation of NKG2D ligands via the DNA damage pathway

NKG2D ligands are a family of host-encoded proteins that enable NK cells to identify and eliminate damaged or infected cells \(^99\). It is well established that activation of the DNA
damage pathway leads to the induction of NKG2D ligand expression\textsuperscript{100}. Furthermore, previous reports have shown AID and APOBEC3G to activate the DNA damage response upon infection with Ab-MLV and HIV, respectively, resulting in NKG2D ligand induction\textsuperscript{59,101}. Upregulation of AID in KSHV infected cells led us to speculate that these cells may also upregulate NKG2D ligands. Of the eight known human NKG2D ligands, we observed increased levels of \textit{MICA}, \textit{MICB}, \textit{ULBP2} and \textit{ULBP3} transcripts by qPCR in the KSHV-infected B cell population (Figure 2.2A). The remaining four ligands either did not differ in their expression or were below the limit of detection in our assay (data not shown). Since NKG2D ligands are known to be regulated post-transcriptionally\textsuperscript{102} we wanted to ensure that the increased transcript levels correlated with higher expression of ligands at the cell surface. Using flow cytometry we measured surface expression of the two most induced transcripts, \textit{ULBP2} and \textit{MICB}. As shown in figure 2.2B, both \textit{ULBP2} and \textit{MICB} were upregulated at the surface of infected B cells. Importantly, treatment of infected cultures with the DNA damage response (DDR) inhibitors SB218078 (inhibitor of checkpoint kinase 1) or caffeine (inhibitor of ataxia telangectasia mutated kinase) resulted in diminished ligand induction within the GFP positive population, while basal ligand expression in the uninfected cells remained unchanged. Together, these results demonstrate that KSHV-infected primary human B cells, which exhibit elevated AID expression, concurrently upregulate activating NK ligands. This upregulation is at least partially dependent on the DNA damage pathway given the diminished ligand induction in presence of DDR inhibitors.

\textbf{AID expression causes a defect in lytic reactivation but does not affect viral latency}

Deamination activities of host APOBEC proteins have previously been shown to directly inactivate retroviruses and Hepatitis B virus (HBV) as a form of innate antiviral defense\textsuperscript{93}. To explore whether human AID activity can negatively impact KSHV fitness, we assessed viral latency and reactivation upon exposure to AID. To this end, we stably transduced BCBL-1 cells, a KSHV-infected cell line established from a human primary effusion lymphoma, with either AID or an empty vector control. The two cell lines exhibited similar viability and proliferation rates (data not shown). To investigate a potential impact of AID on latency, we quantified levels of LANA, a latent transcript constitutively expressed in latently infected cells, in both cell lines as late as 12 weeks post selection. We observed similar levels of LANA transcript and protein (Figures 2.3A & 2.3B), which suggested that the viral episome is equivalently maintained in cells expressing AID and in cells expressing the empty vector.

Next, we treated the two stable cell lines with sodium butyrate (NaBut) to stimulate lytic replication and compared induction of lytic transcripts. Cells recently transduced with AID showed a modest impairment of reactivation as indicated by a decrease in the mRNA expression of the master regulator of the latent-lytic switch, replication and transcription activator (RTA) relative to vector control expressing cells at 48hrs post-stimulation (Figure 2.3C). Interestingly, longer AID expression (up to 10 weeks after selection) led to a much more severe reactivation defect (Figure 2.3D). This defect was evident as early as 24 hrs post-stimulation. A similar expression pattern was observed when additional lytic transcripts were analyzed (Figure S2.1). To verify that inhibition of
lytic gene expression culminates in reduced viral titers, we quantified viral output in the supernatant of reactivated AID and control cells. While supernatant from control cell line resulted in robust infection of HFF cells as measured by LANA staining, supernatant from AID expressing cells produced dramatically fewer LANA-positive cells indicative of reduced viral titers (Figure 2.3E). Our observation that defective reactivation of KSHV occurs only upon prolonged exposure of the virus to AID suggests that in the short-term this enzyme either has a minimal impact on KSHV reactivation or that KSHV has evolved mechanisms to partially evade AID-induced damage.

**UNG2 is required in AID-expressing KSHV infected B cells to allow for robust reactivation**

A previous report looking at host proteins interacting with KSHV LANA identified Uracil DNA Glycosylase 2 (UNG2) as one of the factors directly recruited to the viral genome\(^{103}\). UNG2 is responsible for removing uracil residues generated upon either misincorporation of dUMP during replication, or deamination of cytosines. Based on these data, we hypothesized that UNG2 activity may offset the effect of AID on viral fitness. To test this possibility, we assessed the expression levels of LANA and RTA to reflect KSHV latency and reactivation capacity, respectively, upon knockdown of UNG2 via a previously validated shRNA\(^{103}\). The level of LANA remained unchanged in AID-expressing cells relative to control cells upon UNG2 knock-down (data not shown). In contrast, RTA expression upon NaBut stimulation was dramatically inhibited in AID-expressing cells transduced with UNG2 shRNA but not in control cells (Figure 2.3F). This effect was measurable as early as one week post-transduction with AID and comparable to, or even greater than, that observed after a ten week exposure to AID alone. Taken together, these results demonstrate that UNG2 expression is required in KSHV infected cells to counteract AID activities and to allow for efficient reactivation and viral output.
Figure 2.1. KSHV infection results in upregulation of AID in Primary Human Tonsillar B cells. (A) FACS analysis of KSHV infectivity of primary tonsillar B cells as measured by percent GFP positive cells (right rectangular gate) within the CD19+ population of total cells. Days post co-culture with reactivated iSLK.129 cells are indicated above. (B) FACS analysis of intracellular AID in infected GFP positive cells (red) and uninfected GFP negative cells (filled gray) as defined in (A), or unstained control (black). (C) qRT-PCR analysis of AID expression comparing sorted infected and uninfected cells. Shown are mean values ± SD from 4 independent patients. (D) FACS analysis of AID expression in primary tonsillar B cells 48hrs post transfection with vFLIP or vector control.
Figure 2.2. KSHV infection leads to upregulation of NKG2D ligands. (A) qRT-PCR analysis of the indicated NKG2D ligand transcripts from sorted infected and uninfected B cells. Error bars (SD) are derived from triplicates. Shown is one representative experiment out of three performed. (B) FACS analysis of surface MICB (top panel) and ULBP2 (bottom panel) expression in infected GFP positive cells (thick black) and uninfected GFP negative cells (filled gray), or isotype control (dashed black) on day 3 post infection. (C) Additional FACS analysis of cells treated with DNA damage inhibitors, caffeine (middle panel) and SB218078 (right panel) for the duration of infection. Infected, drug-treated samples are shown in red and uninfected drug-treated cells in thin black as compared to no-drug condition depicted same as in (B).
Figure 2.3. AID expression does not affect viral latency, but results in lytic reactivation defect. (A) qRT-PCR analysis of LANA expression in BCBL-1 cells stably transduced with AID or empty vector control at 10wks post selection. Error bars (SD) are derived from triplicates. Shown is one representative experiment out of three performed. (B) Immunofluorescent staining for LANA (green) and DAPI (blue) as nuclear marker in BCBL-1 cells transduced with AID (right) or empty vector control (left) at 10wks post selection. (C, D and F) qRT-PCR analysis of RTA expression following reactivation for 24 or 48 hrs with NaBut of BCBL-1 cells stably transduced with AID (filled gray bars) or empty vector control (dashed bars) at 1wk post selection (C). See also figure S1; 10wks post selection (D); additionally transduced with shRNA against UNG2 at 1wk post selection (F). Error bars (SD) are derived from triplicates. (E) Relative viral titers from supernatants of BCBL-1 cells described in (D). Same number of cells were reactivated for 5 days and equal volumes of supernatant used to infect WT HFF cells. Staining for LANA (green) and DAPI (blue) reflects relative infectious particles in each supernatant.
Figure S2.1. Additional lytic transcripts inhibited in AID-expressing BCBL-1 cells.
qRT-PCR analysis of K1 and K8.1 expression following reactivation for 24 or 48 hrs with sodium butyrate of BCBL-1 cells stably transduced with AID (filled gray bars) or empty vector control (dashed bars) at 10wk post selection. Error bars (SD) are derived from triplicates. Shown is one representative experiment out of three performed.

Discussion

Traditionally, AID is defined as a factor required for a robust humoral immune response given its role in SHM and CSR. Our study reveals an additional role for AID in innate immune defense against a human γ-herpesvirus, KSHV. Our data demonstrate that KSHV infection of naïve primary B cells leads to a rapid upregulation of AID in all cells harboring the virus, and that this elevated level of expression is sustained for the duration of primary cell culture. Although it is known that AID can be induced in a variety of tissues infected with viral and bacterial pathogens, the biological significance of this expression in relationship to the pathogen remains largely uncharacterized. We provide the first evidence that AID expression not only alerts the innate immune system to the pathogen through the induction of activating NK ligands, but also directly regulates viral fitness by inhibiting its ability to undergo lytic reactivation. Ectopic AID induction has puzzled the field for some time given its ability to increase the risk of oncogenic transformation. Our data demonstrate that AID function imparts a significant advantage in immune defense against a viral pathogen and thus, shed light on the value of AID expression outside of GC B cells.

Pathogens previously reported to upregulate AID accomplish this through a variety of stimuli which ultimately converge to activate NF-kB. HIV, for example, ligates CD40 receptors via virion-associated CD40L, while HTLV-1 and HCV do so via intracellular viral protein Tax and core proteins, respectively. In our case, vFLIP, latent viral gene with a known role in NF-kB activation is sufficient to induce AID expression in primary
human B cells. It would be of interest to investigate whether KSHV lacking vFLIP is partially defective or completely incapable of AID induction.

There is growing evidence for the importance of NK cells in controlling KSHV infection. KSHV-positive PEL cell lines are preferentially lysed by NK cells from healthy blood donors when compared with KSHV-negative lymphoma cell lines. Similarly, NK cells specifically eliminate KSHV-infected fibroblasts. In vivo resolution of Kaposi’s Sarcoma directly correlates with restoration of NK cell activity. We also observed an upregulation of stimulatory NK cell ligands in our primary human B cells upon KSHV infection (Fig. 2B). To counteract NK cell recognition, KSHV employs a variety of strategies to downregulate activating ligands during its latent and lytic replication programs. Viral ubiquitin ligase MIR2 and ORF54 can interfere with NK cell recognition by turning over MICA and MICB, and AICL, respectively. Additionally, miR-K12-7 has been shown to decrease levels of MICB mRNA. This multi-faceted effort to escape recognition by NK cells highlights the importance of this cell type in controlling KSHV pathogenesis.

DNA double strand breaks that result from AID activity are known to trigger DNA-damage stress response. One of the downstream consequences of such responses is the upregulation of stimulatory NK cell ligands. One study has linked genotoxic stress associated with AID function to the induction of the activating NK ligands in cells infected with a murine retrovirus, Ab-MLV. In vivo, this response resulted in protection against Ab-MLV and improved animal survival. An analogous observation was made for APOBEC3G, where cytidine deamination enhanced NK cell recognition of HIV-infected cells via the DNA damage pathway. Consistent with these findings, we show that KSHV infected cells upregulate NKG2D ligands in a DNA-damage dependent manner, suggesting a possible role for AID in potentiating NK cell recognition of KSHV-infected cells.

Beyond this, we observe that expression of AID has a more direct adverse impact on KSHV. Our data reveal that while latency is unaffected by AID expression within the time frame examined, lytic reactivation potential is significantly suppressed. The fact that UNG2 knockdown enhances this phenotype supports the notion that the defect in lytic reactivation is due to deamination activity by AID. One possible explanation for this reactivation defect is accumulation of detrimental mutations within the virus, especially within the RTA promoter. To address this, we sub-cloned and sequenced several one kilobase-long viral genomic regions isolated from virions secreted by reactivated cells expressing AID or vector control. We observed very few mutations and their frequency did not correlate with AID expression (data not shown). While AID may not mutate KSHV genome adequately enough to inhibit reactivation, this result is not surprising as isolation of KSHV virions biased our viral sample genomes toward the less mutated productive viral particles. Because KSHV relies on numerous host components, such as transcription factors and protein synthesis machinery, to initiate and drive lytic reactivation, mutations within the host genome could have negatively impacted viral reactivation. Similar to viral genes, we did not uncover specific AID-dependent mutations within the host genes, although our sequence coverage may have been insufficient to pinpoint such regions. Alternatively, AID expression may generate a cellular environment that is deleterious to KSHV reactivation. For instance, cellular pathways activated by AID (e.g. DNA damage pathway) may affect the activity of host proteins known to contribute to RTA promoter
activation and cooperation with RTA to initiate transcription of downstream viral lytic genes. While latency is generally accepted as the default program of the virus, reactivation is thought to be important in reseeding and long-term persistence of the latent viral pool. Thus, protection against the detrimental effects of AID on lytic reactivation is of vital importance to a persistent virus like KSHV.

The substrate of AID is ssDNA, which explains its preference for actively transcribed genes. Silencing the vast majority of its genome during latency may thus be a deliberate evolutionary strategy of KSHV aimed at avoiding AID mutagenesis among other advantages. In fact, KSHV, as well as its murine homolog, MHV68, maintain latency until terminal differentiation of germinal center B cells into plasma cells. As part of this transition, B cells downregulate AID among other genes, and hence relieve the virus from potential mutational insult.

Previously characterized recruitment of UNG2 by LANA may serve as one mode of preservation of KSHV genome integrity in the presence of AID. In fact, KSHV is not the only virus employing such a strategy. HIV Vpr has been shown to counteract APOBEC3G by diminishing the incorporation of uridine through its interaction with UNG2. However, as evident from our reactivation data, even in the presence of wild type levels of UNG2, AID activity can be detrimental to KSHV in the long run. Our primary B cell infections revealed that AID expression is not transient, but sustained and increased in the course of an infection. Moreover, it is conceivable that upon exposure to stimulatory cytokines and interaction with other activated immune cells during an in vivo infection, AID expression may be augmented even further. Hence, the modulation of AID levels or activity is of vital importance to the long-term success of KSHV.

Since the discovery of its catalytic function AID has been directly implicated in mutations and chromosomal translocations responsible for tumorigenesis. In line with that knowledge, researchers who previously observed AID expression in non-germinal center B cells (e.g. hepatocytes, gastric epithelial cells) often labeled it as inappropriate. Our study supports the notion that, in fact, AID has a protective role during viral pathogenesis by marking infected cells for elimination and inhibiting viral fitness. Therefore, AID induction in the context of oncogenic viruses such as KSHV may actually limit transformation rather than serve as the culprit.

Materials and Methods

**Tissue Culture and B cell enrichment**

The request for research involving primary human tissues was reviewed by the UC Berkeley Office for Human Research Protections of the Department of Health and Human Services. It was determined that de-identified tissue to be used for research does not meet the threshold definition of "human subjects" research set forth in Federal Regulations at 45 CFR 46.102(f) allowing the study to proceed without further IRB approval. Human tonsillar tissues were obtained from the Cooperative Human Tissue Network. Tonsillar tissues were dissociated using a plunger from a 10ml syringe in tonsil media (described below), passed through a 45-um cell strainer twice and washed once in tonsil media. Mononuclear cells were isolated by centrifugation over a Histopaque (Sigma) cushion at 1400 x g for 15min.
Live cells were collected at the histopaque/medium interphase and washed with PBS, 2.5% FBS, 2mM EDTA. Cell pellet was resuspended at 1x10^7 cells/ml and cultured in medium consisting of RPMI 1640 supplemented with 15% FBS, 1% pen-strep, 1 mM sodium pyruvate, 1% nonessential amino acids (Mediatech), 2 mM L-glutamine, and 1% fungizone (Invitrogen). Prior to sorting of infected cells, B cells were enriched using EasySep® Human CD19 Positive Selection Kit (Stem Cell Technologies) following manufacturer’s instructions. Phoenix 293HEK and iSLK.129 cells were cultured in DMEM, 10% FBS and 1% penicillin and streptomycin. BCBL-1 cells were cultured in RPMI 1640, 10% FBS and 1% penicillin and streptomycin.

**KSHV reactivation and infections**

For primary tonsillar cell co-culture infections, iSLK.129 cells grown to 70% confluency were reactivated with 0.2 ug/ml doxycycline for 48hrs, tonsillar cells were added into the same flask and infection was allowed to proceed for indicated amount of time. Virus production from BCBL-1 cells was performed by reactivation with 0.3 mM sodium butyrate for 5 days. Virus-containing supernatant was filtered through .45um filter and used to infect HFFs via 2hr spinfection at 1,000 x g.

**Constructs and Reagents**

vFLIP was cloned into pMax vector (Lonza) and co-transfected with pMax-GFP in a 4:1 ratio for gating on positively transected primary cells. AID was cloned into pCruXIP retroviral vector for expression in BCBL-1 cells. Anti-UNG2 shRNA sequence was adopted from a previously published study and cloned into MSCV-miR-30 based expression system as previously described. Retroviral transductions were performed by transfecting vector of interest into a packaging cell line Phoenix along with vsv-g encoding plasmid, 48hrs later filtering resulting supernatant and spinfecting the target cells with the supernatant for 2 hrs at 1,000 x g with addition of polybrene (4ug/ml). 24hr post transduction cells were either selected with 1ug/ml puromycin or stained with anti-CD2 PE antibody (eBioscience) and sorted for positively transduced cells. All transfections of cell lines were performed using Lipofectamine2000 (Invitrogen). Primary cells were transfected using Amaxa nucleofection protocol U-015 (Lonza).

**Flow cytometry and cell sorting**

Primary cells were pre-blocked with Anti-Human CD32 Blocker (Stem Cell Technologies). For cell surface marker staining, cells were washed and stained in 1% FBS in PBS with anti-hULBP-2/5/6, anti-MICB (R&D Systems) and anti-Human CD19 PerCP-Cy5.5 (ebiosciences) antibodies. For intracellular AID staining, cells were fixed, permeabilized and stained with anti-AID antibody (Millipore) using BD Cytofix/Cytoperm Fixation/Permeabilization Solution Kit (Thermo Fisher) following manufacturer’s instructions. Cells were analyzed using LSR Fortessa cell analyzer (BD Biosciences) or sorted on MoFlo high speed sorter (DakoCytomation).

**Immunofluorescent staining**
Cells were washed and allowed to adhere onto polylysine coated glass slides. Samples were fixed/permeabilized in 50% acetone/50% acetone at -20°C. Cells were rehydrated in 3% BSA, 1% glycine in PSB, and stained with anti-LANA antibody (Advanced Biotechnologies) followed by anti-rat FITC in 3% BSA in PBS buffer in a humidified chamber. Washes were done with 4% Tween-PBS. Slides were imaged with 40x lens using a fluorescent deconvolution microscope.

Quantitative real-time PCR
RNA was extracted in Trizol (Invitrogen), treated with RQ1 DNase (Promega), and total RNA was reverse transcribed using oligo(dT)$_{15}$ primer (Integrated DNA Technologies) and SuperScriptII (Invitrogen) at 42°C for 50 minutes. cDNAs were analyzed using iTaq SYBR Green Supermix With ROX (BioRad) on an ABI7300 Real Time PCR System.

Western blot
Cells were lysed in RIPA buffer (50 mM Tris [pH 8], 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% SDS, and antiprotease). The lysates were cleared of debris by centrifugation in a microcentrifuge at 4°C and protein concentration quantified by BCA assay (Pierce). Samples were run on a 12% polyacrylamide gel and transferred to a polyvinylidene difluoride (PVDF) membrane. Blots were probed with anti-FLAG M2 (Sigma-aldrich) and anti-GAPDH (Abcam) antibodies.

Statistical Analysis
A two-tailed, paired student t-test was performed on all samples where statistical significance is indicated.
Chapter 3 - KSHV-encoded microRNAs regulate host AID as means of immune evasion

Introduction

Herpesviruses have co-evolved with their hosts for millions of years, acquiring means to evade and manipulate host immune responses. Evolutionary success of these viruses is highlighted by their life-long persistence, high prevalence and minimal pathological burden in immunocompetent hosts. In cases of immune suppression, however, these viruses can cause severe disease. Kaposi’s Sarcoma Associated Herpesvirus (KSHV) is a member of the human γ-herpesvirus family characterized by lymphotropism and strict host specificity. It is the causative agent of Kaposi’s Sarcoma, the most common form of malignancy in AIDS patients, and two lymphoproliferative disorders, primary effusion lymphoma (PEL) and multicentric Castleman’s disease (MCD)\textsuperscript{1,6}

Although KSHV can infect a variety of cell types, B cells serve as the primary reservoir of the virus \textit{in vivo}\textsuperscript{11,6}. KSHV favors establishment of latency upon infection, but may reactivate to undergo lytic replication. During latency viral genome is maintained as a multicopy chromatinized episome tethered to host DNA and no progeny virions are generated. Latent gene expression is restricted to just four protein-coding genes and a cluster of twelve miRNAs\textsuperscript{17}, all of which contribute towards promoting cell survival, segregating viral episome during mitosis and suppressing host immune responses. Upon entry into the lytic life cycle, virtually all viral genes are activated in a temporal fashion, the viral genome is replicated and progeny virions are produced, generally resulting in death of the infected cell.

Host immune system provides a broad range of protection against viral invasion including broadly acting innate defenses and highly specific adaptive immune responses. Among innate defenses are the interferon production, complement activation, secretion of inflammatory cytokines, NK-mediated cell lysis, Toll-like receptor activation, induction of apoptosis and autophagy. KSHV dedicates a large fraction of its genome to the manipulation of these immune defenses. Previous chapter delineates a novel role for AID in host innate defense against KSHV and suggests that it would be advantageous for KSHV to regulate AID activity or expression. In this chapter, we examined the ability of KSHV-encoded miRNAs to thwart AID-mediated immunity. Our data uncovered two KSHV miRNAs, K12-11 and K12-5 capable of interacting with the 3’UTR of AID and translationally repressing it. KSHV miRNAs proved to inhibit AID upon individual expression as well as in the context of the entire viral genome. Together our data reveal a novel viral immune evasion strategy employed by KSHV to counteract AID, further reinforcing the importance of this host factor in antiviral defense.
Results

KSHV-encoded miRNAs target 3’UTR of host AID

Many herpesviruses encode multiple independent effectors that target a particular host antiviral pathway in order to achieve efficient immune evasion. We thus postulated that, instead of relying solely on the host UNG2, KSHV may have evolved alternative strategies to counteract AID activity. In the absence of infection, AID expression, stability and activity are endogenously regulated via multiple mechanisms. In particular, several miRNAs, miR-155, miR-181b and miR-93 have recently been shown to regulate AID at the level of translation via interaction with its 3’UTR. Interestingly, EBV, another member of the human γ-herpesvirus family, has been shown to upregulate miR-155 expression anywhere from 20 to several thousand-fold depending on the cell type. However, we observed little to no upregulation of the three host miRNAs known to regulate AID in the day 3 KSHV infected tonsillar B cells (Figure S3.1). Thus, we instead turned our attention to the cluster of KSHV-encoded miRNAs as potential AID modulators, especially given that one of these, miR-K12-11, is a bona fide ortholog of miR-155.

Bioinformatic analysis of seventeen mature KSHV miRNAs yielded nine candidates with partial sequence complementarity to the 3’UTR of AID mRNA. Table 1 ranks the nine candidates in order of the predicted strength of interaction, and describes the extent and type of nucleotide pairing between miRNA seed region and target 3’UTR according to Bartel. To verify the validity of this prediction, we co-transfected each of these miRNAs together with a reporter construct encoding a luciferase gene upstream of the AID 3’UTR and measured luciferase activity. We observed that similar to the positive control hsa-miR-155, expression of miR-K12-11 and miR-K12-5 caused a significant decrease in luciferase signal, whereas none of the other miRNAs had a comparable effect (Figure 3.1A). This phenotype was specific to AID 3’UTR as the data were normalized to the luciferase signal in cells co-transfected with miRNAs and the control vector, which does not encode AID 3’UTR. Within the 3’UTR of AID, miR-K12-5 is predicted to bind at two adjacent positions that are located slightly upstream of the predicted miR-K12-11 binding site (Fig. 3.1B). During viral latency, all KSHV encoded miRNAs are generated from a single-stranded primary polycistronic transcript. Consequently, we examined whether co-expression of both miR-K12-11 and miR-K12-5 results in an additive effect on AID reporter downregulation. Upon co-transfection of the two miRNAs, we observed an even greater reporter inhibition relative to what is achieved with either miRNA alone (Figure 3.1C). To verify specificity of the interaction between AID and miR-K12-5 or miR-K12-11, we mutated several residues within the 3’UTR of AID predicted to interact with the seed regions of miRNAs K12-5 and K12-11. As expected, these specific mutations rescued reporter expression further validating our bioinformatics analysis (Figure 3.1D). In the case of miR-K12-5, the first of the two predicted binding sites makes the greatest contribution to reporter inhibition, which correlates with the corresponding strength of interaction score in Table 3.1.
miR-K12-5 and miR-K12-11 target full length AID for downregulation when expressed at physiological levels

To confirm our luciferase reporter results in a more physiologically relevant system, we cloned miR-K12-11 and miR-K12-5 into a short hairpin retroviral construct that would require processing by the endogenous machinery to yield functional mature miRNAs. As a negative control we generated a construct encoding an shRNA designed to target the luciferase gene. Because most cell lines do not express endogenous AID, we generated 293 HEK cells that stably express full-length AID mRNA containing the natural 5’ and 3’ UTRs. Subsequently, we transduced these AID-expressing 293 HEK cells with our miRNA constructs. A northern blot was performed on transduced cells to verify that miR-K12-5 and miR-K12-11 were expressed at physiological levels. BCBL-1 and BC-1, two cell lines derived from KSHV infected patients, served as controls. While expression levels of both miRNAs varied somewhat between the two control cell lines, which is consistent with a previous report, expression levels of these miRNAs in 293 HEK cells were comparable with at least one of the two control cell lines (Figure 3.2A). Next, we assessed the level of AID protein in the presence of either miR-K12-5 or miR-K12-11. Consistent with our reporter assay, expression of both viral miRNAs resulted in diminished levels of AID protein (Figure 3.2B). Quantitative RT-PCR analysis revealed that neither of the miRNAs destabilized the AID transcript (Figure 3.2C), suggesting that inhibition is occurring at the level of translation, as is the case with endogenous miRNAs.

KSHV miRNAs mediate downregulation of AID in the context of the viral genome

To investigate the contribution of KSHV miRNAs to the regulation of AID expression in the context of the entire viral genome, we utilized a recombinant GFP-tagged bacteria artificial chromosome (BAC) encoding the entire KSHV genome (WT BAC) or one lacking the miRNA cluster from its latency region (ΔmiR BAC). While expression of these BAC’s in mammalian cells does not yield high viral titers, transfection of these constructs into cells recapitulates gene expression seen upon infection with KSHV viral particles. We transiently transfected 293 HEK cells stably expressing AID with either WT or ΔmiR BAC and compared AID levels between the resulting GFP positive cell fractions. AID protein expression was significantly lower in cells expressing full length AID mRNA that were transfected with WT BAC compared to cells transfected with ΔmiR BAC (Fig. 3.3A). However, this difference was largely mitigated in cells expressing only the AID open reading frame (AID ORF), confirming that the KSHV miRNAs predominantly target the AID mRNA via its UTR. Given that the KSHV miRNA cluster may affect global host gene expression, we observed a minor miRNA cluster-dependent decrease in AID ORF expression at early time points following transfection (24 and 48hrs) in a number of independent experiments (Fig. 3.3B). Nonetheless, full length AID was significantly more downregulated compared to AID ORF (Figure 3.3B). Together these results support the role of KSHV-encoded miRNAs in the regulation of AID protein levels via the 3’ UTR of AID.
<table>
<thead>
<tr>
<th>miRNA</th>
<th>Hit-position</th>
<th>Site Type</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-K12-11</td>
<td>1225</td>
<td>m8</td>
<td>1.252807</td>
</tr>
<tr>
<td>miR-K12-5</td>
<td>950</td>
<td>m8</td>
<td>1.212495</td>
</tr>
<tr>
<td></td>
<td>1051</td>
<td>m6</td>
<td>1.053520</td>
</tr>
<tr>
<td>miR-K12-6-5</td>
<td>717</td>
<td>m7_A1</td>
<td>1.117598</td>
</tr>
<tr>
<td>miR-K12-4-3</td>
<td>1107</td>
<td>m6</td>
<td>1.048571</td>
</tr>
<tr>
<td></td>
<td>766</td>
<td>m6</td>
<td>1.038917</td>
</tr>
<tr>
<td>miR-K12-6-3</td>
<td>1211</td>
<td>m7_M8</td>
<td>1.029981</td>
</tr>
<tr>
<td>miR-K12-10b</td>
<td>1473</td>
<td>m6</td>
<td>1.019462</td>
</tr>
<tr>
<td>miR-K12-7</td>
<td>2231</td>
<td>m6</td>
<td>1.012112</td>
</tr>
<tr>
<td>miR-K12-1</td>
<td>1989</td>
<td>m7_M8</td>
<td>.974461</td>
</tr>
<tr>
<td></td>
<td>1854</td>
<td>m7_M8</td>
<td>.966365</td>
</tr>
<tr>
<td>miR-K12-12</td>
<td>1919</td>
<td>m6</td>
<td>.968706</td>
</tr>
<tr>
<td></td>
<td>1770</td>
<td>m6</td>
<td>.963473</td>
</tr>
<tr>
<td></td>
<td>2085</td>
<td>m6</td>
<td>.960405</td>
</tr>
</tbody>
</table>

**Table 3.1. Bioinformatic Analysis of KSHV miRNAs binding to human AID.** Columns identify miRNA name, position within AID mRNA where predicted binding of seed sequence occurs, type of target site and miRNA:gene interaction score according to Bartel D.P. (2009). miRNAs are ordered according the score, from strongest to weakest predicted interaction.
Figure 3.1. KSHV-encoded miRNAs directly target 3’UTR of hAID. (A) Relative luciferase activity after co-transfection of AID 3’UTR reporter and the indicated miRNA mimic. Renilla luciferase activity is normalized to firefly luciferase activity and then normalized to the average activity of the empty control reporter. Error bars (SD) are derived from triplicates. Hsa-miR-155 serves as a positive control. Shown is one representative experiment out of three performed. (B) A schematic representation of the full length AID mRNA and the predicted binding sites for miR-K12-5 and miR-K12-11; specific 3’UTR locations noted in pentacles. (C) Relative AID 3’UTR luciferase reporter activity comparing transfection of a single miRNA, miR-K12-5 or miR-K12-11 versus both miRNAs, each at half of the original concentration. (D) Comparison of relative luciferase activity between WT and mutant 3’UTR of AID reporters. Double mut combines mutations of mut 3’UTR (950) and (1051). Left panel represents co-transfection of reporter with miR-K12-5, right panel – with miR-K12-11. Statistically significant differences are indicated (*p<.01 by two-tailed, paired student t-test).
Figure 3.2. miR-K12-5 and miR-K12-11 target full length AID for downregulation when expressed at physiological levels. (A) Northern blot analysis of total RNA isolated from either BC-1 or BCBL-1 cell lines latently infected with KSHV, or 293HEK cells stably transduced with shRNA constructs expressing individual indicated KSHV or negative control miRNAs. Probes bind to miR-K12-5, miR-K12-11 or siRNA against luciferase gene as indicated above each panel. U6 expression serves as a loading control. (B) Western blot analysis of AID protein levels in 293HEK cells stably expressing full length AID along with the indicated shRNA construct. GAPDH expression serves as a loading control. (C) qRT-PCR analysis of AID expression in 293HEK cells stably expressing full length AID along with the indicated shRNA construct. RNA is harvested at the same time as protein used for western blot analysis in (B). Error bars (SD) are derived from triplicates.
Figure 3.3. KSHV miRNAs downregulate hAID in the context of the entire viral genome. (A) FACS analysis of intracellular AID in 293HEK cells stably transduced with constructs encoding full length AID mRNA (hAID mRNA) or open reading frame of AID (hAID ORF) and transiently transfected with either WT KSHV BAC (red) or BAC lacking miRNA cluster, ΔmiR BAC(filled gray). Black histograms represent background antibody staining of 293HEK lacking AID expression. Analysis is performed at 48hrs post transfection. (B) Quantification of AID downregulation upon WT KSHV BAC transfection relative to ΔmiR BAC at 24, 48 or 72 hrs post transfection from three independent experiments. Dashed bars represent cells stably expressing hAID ORF, solid gray bars represent cells stably expressing hAID mRNA. The percentage of AID downregulation (% MFI), in cells transfected with WT KSHV BAC was calculated relative to cells transfected with the ΔmiR BAC. Shown are mean values ± SD (*p<.05 by two-tailed, paired student t-test).

Figure S3.1. KSHV infection does not dramatically upregulate expression of endogenous miRNA regulating AID. qRT-PCR analysis of miR-93, miR-155 and miR-181b expression comparing sorted infected and uninfected cells. Depicted is fold induction of miRNA in infected relative to uninfected cells. Data are normalized to the expression of miR-191. Error bars (SD) are derived from triplicates. Shown is one representative experiment out of three performed.
Infection of primary human cells with KSHV allowed us to uncover a rapid induction of AID expression in response to the virus. This discovery has contributed a novel factor to the repertoire of already known anti-KSHV responses. Given the precedents, such responses rarely go unnoticed by KSHV, a virus which co-evolved so closely with its host. Instead they are regulated through viral gene expression to minimize or even ablate the detrimental impact on the virus.

The KSHV homolog, EBV, is known to both upregulate miR-155 as well as specifically inhibit AID via its latency protein EBNA2. While we have not observed analogous activity from KSHV, we uncovered two distinct viral miRNAs, miR-K12-5 and miR-K12-11, capable of decreasing AID protein expression. Such an approach can be particularly advantageous given that miRNAs can modulate expression of AID and numerous other targets without generating peptides, which may put the virus at risk for immune detection during latency. A partial decrease in AID protein levels, achieved through expression of miR-K12-5 and miR-K12-11 is very consistent with the role of miRNAs in fine-tuning gene expression. Importantly, AID protein function is highly sensitive to its abundance given that it exhibits haploinsufficiency for antibody diversification and chromosome translocations. Additionally, AID expression levels correlate with mutation rates in cell lines. In light of these findings, even a partial decrease in AID expression achieved by miR-K12-11 and miR-K12-5 may offer significant contribution to viral defense against AID.

Quite paradoxically, it is conceivably advantageous for KSHV to not completely eliminate AID expression. Data from patients exhibiting KSHV-related pathologies as well as in vivo experiments with MHV-68 suggest that latency is established primarily in GC B cells and maintained long-term in GC and memory B cells. The highly proliferative nature of activated GC B cells affords a perfect opportunity for the expansion of KSHV-infected cellular pool. Survival of these B cells depends on AID-mediated SHM to produce optimal B cell receptor affinity. Hence, occupation of this niche would prove most suitable when AID levels are adequate to carry out antibody diversification, yet not excessively high where genotoxic stress could pose a threat to viral welfare.

In conclusion, we identify two virally-encoded miRNAs, miR-K12-11 and miR-K12-5, capable of repressing AID protein levels. This finding not only adds to the growing field of herpesvirus miRNA biology by uncovering a novel gene target for miR-K12-5 and miR-K12-11, but further reinforces the importance of AID in controlling KSHV infection. Moreover, this discovery provides the first evidence for regulation of AID via virally-encoded miRNAs. Recruitment of UNG by KSHV LANA plays an important role in protecting viral genome from deamination. It would be of interest to examine whether KSHV employs additional strategies to interfere with other host factors which participate in AID recruitment, catalytic reactions or post-deamination DNA editing.
Materials and Methods

Tissue Culture

293HEK and Phoenix 293HEK cells were cultured in DMEM, 10% FBS and 1% penicillin and streptomycin. BC-1 and BCBL-1 cells were cultured in RPMI 1640, 10% FBS and 1% penicillin and streptomycin.

Constructs and Reagents

AID was cloned into pQCXIN vector with a 3xFlag-tag for expression in 293HEK cells. Hsa-miR-155, KSHV miRNAs, and anti-Luciferase shRNA were cloned into MSCV-miR-30 based expression system as previously described. 3’ UTR of AID was cloned into dual luciferase reporter vector psiCHECK-2 (Promega). Mutant 3’UTR reporters were generated by introducing two nucleotide substitutions within the regions predicted to bind seed sequence of the corresponding miRNA. Retroviral transductions were performed by transfecting vector of interest into a packaging cell line Phoenix along with vsv-g encoding plasmid, 48hrs later filtering resulting supernatant and spinfecting the target cells with the supernatant for 2 hrs at 1,000 x g with addition of polybrene (4ug/ml). 24hr post transduction cells were selected with 600ug/ml G418. KSHV BAC constructs were a kind gift from the Gao group. All transfections of cell lines were performed using Lipofectamine2000 (Invitrogen).

Flow cytometry and cell sorting

For intracellular AID staining, cells were fixed, permeabilized and stained with anti-Flag M2 antibody (Sigma-alrich) using BD Cytofix/Cytoperm Fixation/Permeabilization Solution Kit (Thermo Fisher) following manufacturer’s instructions. Cells were analyzed using LSR Fortessa cell analyzer (BD Biosciences) or sorted on MoFlo high speed sorter (DakoCytomation).

Quantitative real-time PCR

RNA was extracted in Trizol (Invitrogen), treated with RQ1 DNase (Promega), and total RNA was reverse transcribed using oligo(dT)15 primer (Integrated DNA Technologies) and SuperScriptII (Invitrogen) at 42°C for 50 minutes. cDNAs were analyzed using iTaq SYBR Green Supermix With ROX (BioRad) on an ABI7300 Real Time PCR System. For miRNA quantification total RNA was reverse-transcribed using cDNA Synthesis kit for miRNA (OriGene). miRNA expression was assessed by quantitative RT-PCR using miRNA-specific forward and universal reverse primers.

Luciferase assay

293HEK cells were co-transfected with 2.5 pmol miRNA mimics (Qiagen) and 50ng of reporter DNA construct per well of a 96-well plate in triplicates. 24hrs later cells were lysed in Passive Lysis Buffer (Promega) and luciferase activity measured using an LMAX II Luminometer (Molecular Devices).

Northern blot
Thirty micrograms of total RNA was loaded onto 15% 8 M urea polyacrylamide gel and transferred onto Hybond-N+ Nucleic acid transfer membrane (Amersham) following electrophoresis. Probe labeling was performed using T4 polynucleotide kinase (New England Biolabs) in the presence of $[{\gamma}^{32}\text{-P}]$ATP.

**Western blot**

Cells were lysed in RIPA buffer (50 mM Tris [pH 8], 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% SDS, and antiprotease). The lysates were cleared of debris by centrifugation in a microcentrifuge at 4°C and protein concentration quantified by BCA assay (Pierce). Samples were run on a 12% polyacrylamide gel and transferred to a polyvinylidene difluoride (PVDF) membrane. Blots were probed with anti-FLAG M2 (Sigma-aldrich) and anti-GAPDH (Abcam) antibodies.

**Bioinformatic Analysis**

miRNA binding predictions against the complete *AID* transcript were performed according to Bartel's findings $^{119}$. 

**Statistical Analysis**

A two-tailed, paired student t-test was performed on all samples where statistical significance is indicated.
Chapter 4 - Activation-induced cytidine deaminase inhibits MHV68 replication by hypermutating its genome

Introduction

Gammaherpesviruses are lymphotropic viruses that exhibit very strict host specificity. As a result, the study of pathogenesis of human gammaherpesviruses, KSHV and EBV is limited to the in vitro experiments. The closest animal model of infection relies on the murine homologue, MHV68. MHV68 was originally isolated from bank voles and yellow-necked field mice in Slovakia\textsuperscript{125}. Fortuitously, MHV68 readily infects laboratory mice and establishes persistent infection similar to its human counterparts. MHV68 genome is highly homologous to both KSHV and EBV, however it does possess unique genes not present in the other viruses. Experimental evidence demonstrates that majority of these genes are dispensable for viral replication and likely contribute to the establishment of a persistent infection and the immune evasion within the specific host\textsuperscript{126,127}. However, while these genes appear divergent in their sequence from their primate counterparts, a number of them do, in fact, display functional conservation.

Although, the natural route of MHV68 infection could not be definitively established, the virus is believed to be transmitted via the respiratory route\textsuperscript{128}. In the laboratory setting, such physiologically relevant infection is achieved through intranasal inoculation. Similar to the human gammaherpesviruses, MHV68 is largely asymptomatic in healthy hosts but contributes to lymphomagenesis and eventual mortality in immunocompromised animals. Mice lacking IFN-gamma receptor succumb to a severe large-vessel panarteritis, while B cell-deficient animals develop a milder large-vessel arteritis\textsuperscript{129}.

Intranasal inoculation results in an early acute infection, during which lytically replicating virus can be found within the mucosal epithelial cells lining the lung and the nasal cavity. Days 5-8 post infection correspond to the peak in overall viral loads within the infected animal. Subsequently, the virus migrates to the draining lymph nodes and eventually to the spleen where it establishes a persistent, predominantly latent infection. Although, the mice are grossly asymptomatic at this time, they exhibit noticeable transient splenomegaly by about 2 weeks post infection and harbor the virus in as many as one percent of the total splenocytes\textsuperscript{130}. Curiously, original inoculation dose plays little role in determining the extent of latent infection within the spleen. By 3 months post infection, the frequency drops down to roughly 1 out of 10,000 cells and is maintained long-term exclusively within the B cells. These B cells appear to have undergone the germinal center reactions and established a long-lived memory phenotype.

In addition to infecting B cells and nasal epithelial cells, MHV68 can also infect macrophages and dendritic cells. Similar to KSHV, cell-free MHV68 cannot readily infect cultured B cells. However, recent advances in the KSHV field, which allow for the infection of cultured B cells via cell-to-cell spread from reactivated endothelial cells, suggest that the same may hold true for MHV68\textsuperscript{95}. Our current working hypothesis for MHV68 trafficking in vivo involves MHV68 infecting lung epithelial cells, whereby lytic reactivation produces cell-free virus or cell-to-cell spread passes the virus to the circulating macrophages and
dendritic cells. These cells, in turn, traffic to the secondary lymphoid organs and seed long-term infection of the B cells via cell-to-cell spread of the virus. An important experimental distinction between MHV68 and KSHV is that *in vitro* infection with MHV68 results in a spontaneous, robust lytic reactivation, while KSHV generally establishes latency upon infection of the permissive cell lines. Depending on the multiplicity of infection, a noticeable cytopathic effect can be observed between days 2 and 5 post infection. Hence, the viral titers of MHV68 can be experimentally determined via a standard plaque assay technique.

Growing evidence is emerging suggesting that bacterial and viral infections of the host trigger ectopic expression of AID. The significance of AID induction under these circumstances is largely unexplored. In the previous two chapters we have described the induction of AID in response to KSHV and uncovered that its expression is able to inhibit viral fitness. However, certain technical limitations have prevented us from thoroughly assessing mutagenic effect of AID during lytic reactivation of KSHV. Given that MHV68 spontaneously reactivates in tissue culture, we wanted to explore whether AID can directly mutagenize the genome of this virus as an immune defense strategy. To accomplish this we passaged the virus through AID-expressing cells and discovered that MHV68 accumulated a substantial number of genomic mutations. The rate of mutation significantly surpassed the expected rate of somatic mutation. Moreover, the acquisition of these mutations correlated with the inability of the MHV68 to produce infectious virions beyond the second passage. Together, our results demonstrate that AID is capable of directly mutating the MHV68 genome and thus, inhibiting its replication.
Results

Establishment of fluorescently-labeled MHV68 infection model

In order to better characterize gammaherpesvirus interaction with the host AID in vivo we set out to establish a trackable infection model in the lab. To this end, we inoculated WT and IFNR-/- mice with fluorescently-tagged MHV68 and assayed for the presence of infected cells at two weeks post inoculation. This virus encodes a yellow fluorescent protein (YFP) driven by the human cytomegalovirus immediate-early promoter and enhancer in a neutral locus within the viral genome. Since AID is induced primarily in the germinal center B cells, we focused our attention on the infection of this particular splenic compartment. As seen in figure 4.1, MHV68 infection induces the formation of a detectable germinal center compartment as measured by the GL-7 and Fas markers when compared to naïve mice. Moreover, consistent with the previously demonstrated role of IFN gamma in the control of infection, we observed larger germinal center compartments within these animals (Figure 4.1). Furthermore, a measurable fraction of these cells appeared positive for the fluorescently tagged virus. In particular, nearly a third of all the germinal center B cells of the infected IFNR-/- mice harbored MHV68. Together, these results establish a system where MHV68 can be easily traced within the germinal center B cells. Moreover, to more easily characterize B cell infections the percent of infected cells can be successfully augmented by infecting animals with deficient T cell responses (i.e. IFNR-/- mice).

Mutation of a reporter gene by mAID in vitro

Prior to examining the effects of AID on MHV68 in a complex in vivo system, we wanted to assess the ability of this enzyme to act on MHV68 fitness in vitro. To do so we cloned either the wild type AID (WT) or the dominant negative (DN) H56R/E58Q catalytic mutant into retroviral vectors for stable expression in cell lines. NIH3T3 cells readily infectable with MHV68 were stably transduced with either empty vector control or the two AID constructs. To test the efficacy of AID activity in these cells we generated a reporter GFP construct. This GFP construct was specifically mutated to introduce a single nucleotide substitution responsible for a premature stop codon within the open reading frame. Moreover, this premature stop codon was placed in the context of an RGYW hot-spot motif for AID. Following stable expression of AID or control constructs, the cells were transduced and selected for the presence of the GFP reporter. As seen in figure 4.2, at two weeks post selection there is a detectable appearance of GFP-positive cells in AID-expressing cells. On the contrary, the vector control and DN AID-expressing cells produce very few GFP-positive cells that either correspond to a small number of false-positive cells or spontaneous reversion of the mutation in a small fraction of the population (Figure 4.2A).

As an additional control, we transduced the same three cell lines with WT GFP and assayed for the loss of GFP signal to infer the ability of AID to inactivate this gene. Analogous to the reversion experiment, we observed a more rapid loss of GFP positive cells in AID-expressing cells relative to the control over time (Figure 4.2B). The loss of this signal
was likely a result of mutagenesis since the cells were kept under G418 selection to ensure maintenance of the GFP-encoding sequence. Together these results demonstrate that stable expression of AID in NIH3T3 cells supports proper AID-function as measured by our reporter.

**Deamination hot spots are evenly distributed throughout the mouse genome**

Although, the specific sequence motifs are not an absolute requirement for AID target selection, a number of groups have determined a preferential motif over-represented in the sampled sequences. To ensure that genomic region selected for further mutational analysis harbors sufficient number of hotspot motifs we performed a bioinformatic analysis of the distribution of RGYW motif throughout the genome. As seen in figure 4.3, fractioning the entire MHV68 genome in 1kb fragments revealed a fairly uniform distribution of RGYW-containing sequences throughout the genome. Each 1kb fragment contained on average 28 distinct hot spot motifs. To select a genomic region for mutational analysis we chose the 1 kb fragment between nucleotides 77,000 and 78,000, which spans the coding region for ORF58. This gene encodes a type 3 transmembrane protein shown to interact with gp48 and participate in actin-dependent viral spread. Sequence analysis for this gene revealed 28 distinct hot spot motifs and no complex repetitive sequence that could impede its amplification.

**Exposure to AID largely inactivated the virus after just two rounds of replication**

To establish whether MHV68 is susceptible to AID-induced mutagenesis similar to our reporter construct or is somehow protected from such mutations we performed a series of infections (Figure 4.4). We infected either AID or the two control cell lines (DN AID and empty vector control) with MHV68 at a low MOI. After a two hour infection, the cells were thoroughly washed and allowed to produce lytic MHV68. All three cell lines exhibited qualitatively similar cytopathic effect around day 4-5 post infection. The resulting viral supernatants were harvested and titered using a standard plaque assay technique (data not shown). The infection was then repeated using a fresh batch of cells stably expressing AID or vector control by inoculating the cells with an equal quantity of MHV68 from the first harvest. All three cell lines appeared to continue to experience cytopathic effect and to produce the virus. However, the plaque assay for the second generation of virus repeatedly yielded extremely low titers for MHV68 which originated from AID-expressing but not the control cells.

Although, we originally intended to carry out at least four passages of MHV68 through our stable cell lines, we instead collected the second generation viral supernatant for mutational analysis. 1 kb fragment spanning ORF58 was successively amplified and T/A cloned into plasmid to be amplified for sequence analysis. Fifteen to twenty clones encoding ORF58 originating from the virus passaged through either WT AID or control cells were sequenced. Only one of the twenty control clones possessed a single mutation (Figure 4.5 A). On the other hand, more than half the clones from AID sample contained at least one mutation (Figure 4.5 A). Upon closer examination of the distribution of hot spot motifs within the sequenced region, it became apparent that majority of acquired mutations fell
outside of the specific 4-bp motif. Moreover, the mutations preferentially occurred at the T/A rather than C/G base pairs.

Susceptibility to AID mutagenesis is known to correlate well with the transcriptional status of the gene. Hence, we expect there to be a discrepancy between mutational load within the MHV68 genome during latency and lytic replication. To compare the susceptibility to mutation of lytically replicating MHV68 to latent viral genome, we stably expressed the same constructs within the S11e cell line\textsuperscript{133}, which is latently infected with MHV68. In contrast to the data obtained with lytically reactivated virus, we uncovered no mutations in the latent MHV68 genome in either the control or AID expressing cells (data not shown). Together these data reveal that transcriptionally active MHV68 genome is highly susceptible to AID mutagenesis, while latency may serve as effective means of protecting genomic integrity in presence of AID.

Figure 4.1. MHV68 infection and germinal center formation in WT and IFNR\textsuperscript{-/-} mice. WT and IFNR\textsuperscript{-/-} were infected with MHV68-YFP intranasally. At 2 weeks post infection whole spleens were harvested and stained for germinal centers (GC) by flow cytometry. Depicted are total live splenocytes gated on CD19 positive B cells. Rectangular boxes delineate GC cells (GL-7 high, Fas high) in the top panel. Percent infected cells (YFP+) within the germinal centers is depicted by the gates drawn in blue (bottom panel).
Figure 4.2. AID-dependent mutation of the GFP reporter construct. NIH3T3 cells were stably transduced with either empty vector control, dominant negative AID (DN) or wild type AID (WT), the subsequently stably transduced with either mutant GFP encoding a premature stop codon (A) or WT GFP (B). 14 days post selection the appearance of GFP positive cells (pink gate, panel A) or disappearance of GFP positive cells (pink gate, panel B) were assessed by flow cytometry. The relevant population percentages are noted in the upper right-hand corner.
Figure 4.3. Distribution of RGYW motifs across the MHV68 genome. Bioinformatic analysis of the distribution of AID hot-spot motif (RGYW) along the length of MHV68 genome. Each bar represents a genomic region of 1kb and plots the occurrence of the hotspot motifs within each bin. The average number of motifs per 1kb bin in the above analysis is 28 with a standard deviation of 7.
AID-expressing (or vector control) cells were infected with MHV68 at an MOI of 0.1. Upon entry MHV68 spontaneously entered lytic replication cycle; resulting viral supernatant was harvested and titered. A fresh batch of AID-expressing cells was infected and resulting viral supernatant harvested for gDNA analysis. Specific fragment of MHV68 genome was amplified, subcloned into TA cloning vector pCR2.1 and transformed into bacteria. Individual clones were selected for sequencing.
Figure 4.5 List and distribution of AID-dependent mutations within MHV68 ORF58. 950-BP long fragment of ORF58 of MHV68 was subcloned and sequenced following 2 rounds of amplification in AID-expressing NIH3T3 cells. Green boxes indicated hot-spot RGYW motifs; yellow boxes indicate mutated nucleotides observed once in the sequenced population, light blue boxes indicate mutated nucleotides observed in two separate clones.

### Discussion

In this study we have confirmed a previously made observation that MHV68 preferentially infects germinal center B cells during early establishment of latency. While offering particular advantages such as the rapid proliferation, this unique compartment poses its challenges for viral replication. Namely expression of AID presents a potential risk of destabilizing the viral genome and inactivating the virus. To test this hypothesis, we examined the ability of AID to restrict MHV68 infectivity in an *in vitro* assay. We verified that the exposure of MHV68 to AID during lytic replication resulted in an accumulation of a high number of mutations within the viral genome. This in turn, culminated in a severe
defect in the production of infectious virus after just two rounds of viral replication. On the contrary, latent viral genome appeared protected from the analogous mutagenesis.

It is well established that AID targets ssDNA which becomes accessible during active transcription of dsDNA. During latency gammaherpesviral genomes are maintained as largely chromatinized, inaccessible circular episomes\textsuperscript{134}. It is thus, reasonable to postulate that MHV68 genomic DNA conformation in the predominantly transcriptionally silent latent state serves as effective means of protection against cytidine-deamination. Upon lytic reactivation, however, the viral genome is likely more vulnerable to mutagenesis. Our data have shown that MHV68 does not employ special means that effectively protect its DNA from the action of AID. Moreover, during lytic replication viral genome is very highly transcribed arguing that viral DNA may actually be a preferential target for deamination in addition to the host immunoglobulin genes. Nonetheless, it was quite surprising to see a defect in infectious viral titers as early as after two rounds of replication in our \textit{in vitro} system. Admittedly, the level of AID expression in our stable cell lines may be elevated relative to the endogenous AID expression during germinal center reactions. It would thus be of interest to generate cell lines with varying degrees of AID expression and correlate them to the extent of inhibition of viral replication.

A number of groups have made the observation that exit from viral latency for MHV68 and KSHV coincides with the differentiation of the infected B cell. Specifically, MHV68 protein M2 was shown to be responsible for plasma B cell differentiation\textsuperscript{135}. This developmental transition signifies the completion of the BCR maturation process during the germinal center reaction and converts B cells into antibody secreting cells. At this stage, key factors in antibody diversification process including AID are shut down. Hence, it is conceivable that MHV68 M2 effector protein stimulates downregulation of AID as one of the events required to establish a favorable milieu for viral reactivation.

Having established an AID-dependent phenotype in vitro, it would now be of interest to characterize MHV68-AID interaction \textit{in vivo} or \textit{ex vivo}. We are currently setting up a system for infectioning B cells with MHV68 in culture. Using a sequential infection model detailed in the introduction we hope to be able to infect primary mouse splenocytes. This system will allow us to ask whether infection with MHV68 triggers AID upregulation in naïve B cells prior to germinal center formation similar to what we observed with KSHV. Moreover, we would be able to probe cellular signaling pathways responsible for the induction of AID using various knock-out mice with defects in immune sensing and signaling. Additionally, we would be able to investigate which viral effectors are required to trigger AID upregulation via MHV68 BAC mutagenesis.

\textbf{Materials and Methods}

\textbf{Mice and viruses}

C57BL/6 mice were obtained from The Jackson Laboratory (Bar Harbor, ME). Experiments were conducted with 6- to 10-wk-old mice in accordance with institutional guidelines for animal care and use. MHV68 was obtained from the American Type Culture Collection (VR-1465) and propagated on 3T12 cells. Viral titers were determined in a standard plaque assay on NIH 3T3 cells. Viral titers reflect the titer of the stock at the time
of the infection (i.e., after freeze-thawing). Virus containing supernatants were cleared of cellular debris by low-speed centrifugation for 20 min. Virus was subsequently concentrated by centrifugation at 24,000 × g for 2 h and purified over a 30% sucrose cushion by centrifugation at 85,000 × g. Pelleted virus was resuspended in media and stored at −80°C. NIH 3T3 and phoenix cells were grown in DMEM supplemented with 10% FBS and 100 units/ml penicillin and 100 μg/ml streptomycin. Mice were anesthetized with isoflurane and inoculated intranasally with 5 × 10^4 PFU per mouse in 20-μl volume.

**Organ harvest**
Spleens were harvested at day 14 to 16 post infection. Splenocytes were isolated by dissociating the spleen and passing through a 40 μm cell strainer. Cells were then pelleted, resuspended in red blood cell lysis buffer (150 mM NH₄Cl, 10 mM KHCO₃, 0.1 mM EDTA), and incubated at room temperature for 5 minutes. Cells were then pelleted, washed and resuspended in RPMI medium with 10% FBS and Pen/Strep before counting.

**Flow cytometry**
Splenocytes were stained in PBS, 1% FBS using anti-CD19 PE Cy5.5 conjugated antibody (eBiosciences), anti-GL7 APC-conjugated antibody (eBiosciences), anti-Fas PE-Cy7 conjugated (BD biosciences). Infected cells were fixed in PBS, 1% PFA prior to flowcytometric analysis. Data were collected on an LSR II flow cytometer (BD Biosciences) and analyzed using FlowJo software (Tree Star).

**Constructs and Retroviral transduction**
The coding region of AID was amplified from primary mouse splenocytes stimulated in culture with LPS and IL-4 to upregulate AID. DN H56R/E58Q AID was generated via sawing PCR by designing internal primers with the desired nucleotide substitution. Mutant GFP gene was generated by introducing a premature stop codon aa position 107. AID sequences were cloned into pCru-X-IP, GFP-coding sequences were cloned into pQC-X-IN. Retroviruses were produced by co-transfecting Phoenix cells with the vector of interest plus VSV-G and harvesting virus containing supernatant 48hrs later. NIH3T3 were transduced in growth phase with filtered (.45um) supernatant by 2hr spinfection at 1,000x g. Stable cells lines were selected with either 600μg/ml G418 or 1μg/ml puromycin. Mutational analysis was performed by PCR amplifying 1kb region between positions 77,000 and 78,000 in the MHV68 genome using Taq polymerase which leaves T/A overhangs. The resulting fragment was then cloned into pCR2.1 using T/A cloning kit (Invitrogen, #K2020-20).

**Isolation of viral genomic DNA**
To obtain genomic DNA from viral particles, virus was concentrated and purified over a sucrose cushion as described earlier. The viral pellet was resuspended in PBS. An equal volume of 2× proteinase K buffer (20 mM Tris-Cl pH 8, 10 mM EDTA, 1% SDS, 100 μg/ml proteinase K) was added, and the mixture was incubated overnight at 50–60°C. Viral genomic DNA was extracted twice with phenol–chloroform and once with chloroform only. The DNA was precipitated with an equal volume of isopropanol, washed once with 70%
ethanol, and resuspended in 10 mM Tris·Cl pH 8. The DNA was treated with RNase A at 25°C for 30 min and then again extracted with phenol–chlooroform, precipitated with isopropanol, washed with ethanol, and resuspended in 10 mM Tris·Cl pH 8. Viral DNA from latent virus was copurified with genomic DNA from S11e cells, which are persistently but not exclusively latently infected with MHV68. S11e cells were lysed in proteinase K buffer and genomic DNA extracted as described above for DNA from viral particles.

References


Tam, W., Ben-Yehuda, D. & Hayward, W. S. bic, a novel gene activated by proviral insertions in avian leukosis virus-induced lymphomas, is likely to function through its noncoding RNA. *Molecular and cellular biology* **17**, 1490-1502 (1997).


Verma, S. C. *et al.* Latency-associated nuclear antigen of Kaposi’s sarcoma-associated herpesvirus recruits uracil DNA glycosylase 2 at the terminal repeats and is important for


Appendix A - The impact of KSHV latent infection on host cell translation

Note: The following study was performed in collaboration with Jamie G. Bates from Patrick O. Brown laboratory at Stanford University. Materials and procedures used were adopted from an article published in an open access journal (doi:10.1371/journal.pone.0037108)

Introduction

Kaposi’s Sarcoma-associated herpesvirus (KSHV) is a prevalent human oncogenic virus belonging to the gamma-herpesvirus family. It is known to cause persistent infections, which result in serious malignant transformations in immunocompromised individuals. The most common disease associated with KSHV is Kaposi’s sarcoma, a highly vascularized cancer originating from the endothelial tissues. Endothelial cells infected by KSHV are characterized by spindle-shaped morphology and harbor mostly latent virus. Latency, a state preferentially adopted by KSHV, supports minimal viral gene transcription sufficient to ensure host cell survival, partitioning of the viral genome during cell divisions and evasion of numerous host immune responses. The genomic region containing latent genes, or latency region, is limited to four protein-coding genes and a cluster of miRNAs. Despite their limited number, these viral genes are able to accomplish a great deal of host modulation due to their ability to interact with multiple different host factors or affect signaling networks with wide-ranging downstream consequences.

KSHV has been shown to induce endothelial cell reprogramming upon infection. Specifically, infection of blood vessel epithelial cells (BEC) induces expression of lymphatic markers such as PROX1, while infection of lymphatic endothelial cells (LEC) stimulates the converse by inducing a more BEC-like phenotype\(^1\)-\(^3\). As a result of these and other modulations KSHV alters normal cellular biology by promoting angiogenesis. Specific processes that contribute to increased angiogenesis are enhanced migration and invasion, accelerated growth of vasculature even in low growth factor environment, secretion of angiogenesis promoting factors or their cognate receptors\(^2\),\(^4\)-\(^8\). More recently, latent infection has also been shown to alter cellular metabolic pathways revealing that lipogenesis is essential for survival of latently infected cells\(^9\).

Specific signaling pathways shown to be regulated by KSHV infection include MAPK, JAK–STAT, Notch, HIF and Wnt, which in turn regulate immune responses, cellular growth, differentiation and apoptosis\(^10\).

Immune responses as well as clinical therapies aimed at controlling viral infections are generally targeted against aspects of lytic viral life-cycle. However, controlling latent viral maintenance holds the most promise of eradicating a persistent viral infection. Thus, it is of paramount value to extensively characterize cellular changes that occur during latent infections. Traditionally studies aimed at understanding host cell reprogramming by latent KSHV have taken the approach of examining changes in mRNA expression levels using microarray or deep-sequencing platforms. While such approaches have been instrumental in understanding virus-host interactions, they do have limitations. Such methods can readily uncover genes transcriptionally suppressed or activated by the virus, or transcripts whose stability is affected by the viral factors. However, host expression...
profiles can also be regulated at the level of translation. A key biological phenomenon which falls within that category is miRNA-based regulation of gene expression. These small non-coding RNAs exert their effects through imperfect base-pairing with their targets. While in some instances this association leads to transcript degradation similar to what occurs in the case with siRNAs, more often the effect is exerted by a block in translation. In addition to miRNAs controlling expression, viral proteins may alter the expression of translational machinery, which would in turn regulate cellular protein output. Previous screens designed to characterize global changes in host gene expression during viral infection have likely overlooked numerous important changes that occur at the level of translation.

Accordingly, a more realistic view of cellular proteome can be achieved by looking specifically at mRNAs engaged in translation by enriching the mRNA fraction associated with multiple ribosomes known as polysomes. The distinction between actively translated mRNA and its inactive counterpart can be made since polysome associated fraction is significantly heavier compared to mRNA sequestered in the messenger ribonucleoprotein (mRNP) particles or monosome-associated mRNA. Using such an approach we performed a genome-wide polysome profiling assay called Gradient Encoding to better understand the changes in the translational profiles of cells infected with KSHV. This technique relies on DNA microarray analysis using differentially labeled mRNAs within sequential fractions of a linear sucrose gradient as the source of the signal (Figure A.1). Gradient Encoding has been demonstrated to yield an accurate and reproducible ranking of the positions of mRNAs in the gradient. This allows sensitive detection of changes in the average number of ribosomes per mRNA, from which we infer the relative changes in translation rate of each mRNA from one condition to another. In a pilot experiment, we have been able to assess translational status of roughly 2,000 human genes and uncovered host IFI6 as a potential target for translational regulation by KSHV. Existing data implicating IFI6 in antiviral immunity along with our polysome experiment make IFI6 an attractive candidate for further characterization. It would be of value to understand the mechanism by which KSHV may regulate IFI6 translation as well as to better characterize the biological function of IFI6 in the context of KSHV infection.
Figure A.1. Schematic representation of gradient encoding. (Adapted from doi:10.1371/journal.pone.0037108.g001) Gradient encoding relies on separation of mRNAs in a sucrose gradient by virtue of the number of tethered ribosomes. mRNAs from each fraction are sequentially labeled with varying Cy5(red)/Cy3(green) ratio and analyzed by microarray. Each mRNA’s ratio of red/green represents its average position within the gradient.
Results

Comparison of KSHV-infected and uninfected cells yields extensive differences in translational profiles

To identify genes translationally regulated during latent KSHV infection, we prepared large quantities of cell-free viral lysate by reactivating BCBL-1 cell line, which is derived from KSHV-infected lymphoma patient. Following purification and concentration of the viral supernatant we performed a titration assay to determine the ratios of HFF cells to viral supernatant required to achieve 100% infection rate. As evident from the immunofluorescent images in Figure A.2A, we were able to obtain 100% infection efficiency with HFF cells as determined by staining for the viral latent protein LANA. We then scaled up the infection to obtain sufficient material for gradient encoding, passaged the infected and uninfected cells for 5 days post infection to ensure establishment of a robust latent infection and harvested the cell lysates. Lysates in duplicates were subsequently fractionated by sedimentation through linear sucrose gradients. As seen in figure A.3, we observed a stark shift of ribosomes from the polysomes to the 80S peak within the infected cells. Such shift is highly indicative of very significant changes occurring at the level of translation and is likely to yield a very extensive list of potential targets whose translation is affected. While such data could hold some value, we suspected that the cause behind the observed shift might reflect more of a cell cycle arrest phenotype than a latent-infection phenotype attributable to specific viral effectors.

Concentrating large quantities of infectious virus by centrifugation inevitably concentrates the cellular and viral debris generated during the lytic infection. This undesirable component of the supernatant may exert some level of toxicity capable of triggering biological responses within cultured cell which may affect translation. To better control for this caveat, we prepared identical stocks of infectious concentrated virus and UV-inactivated non-infectious virus for the next round of testing. Under these conditions both the infectious and the negative control lysates contain comparable quantities of cell debris. Using immunofluorescent assay, we confirmed that UV-treated virus gave rise to a largely non-productive infection of HFF cells as compared to the untreated virus (Figure A2.B).

We then repeated gradient fractionation using duplicate KSHV-infected samples and triplicate UV-treated KSHV-infected samples. Gradient fractions were then encoded, whereby the mRNA from successive fractions was labeled with increasing ratios of Cy5 to Cy3. mRNAs derived from fractions in the lighter portion of the gradient therefore have a lower Cy5 to Cy3 ratio, whereas those deeper in the gradient have a higher Cy5 to Cy3 ratio. The resulting ratios were quantitatively measured for each mRNA species by hybridization to the DNA microarray. Resulting expression levels were expressed as log2 ratios and infected vs UV-treated samples were compared and ranked using SAM scores. The top one hundred hits are depicted in Table 4.1. SAM (Significance Analysis of Microarrays) is a statistical technique for finding significant genes in a set of microarray experiments. It was originally proposed by Tusher, et al. 13. SAM score reflects the mean difference between the two samples (infected vs uninfected) divided by the standard
deviation plus a "factor" which helps to prevent false high scores due to mostly low standard deviation rather than high difference in mean.

Moreover, we carried out a pairwise two class analysis to ask which spots on the array were statistically different between the UV-treated and untreated KSHV-infected samples. The numbers used in the analysis are log2 values of the Cy5/Cy3 ratios which reflect position within the gradient (higher number means deeper in the gradient, and hence bound to more ribosomes). The data were mean centered and plotted. Replicate samples displayed good correlation as evident from figure A.4. Conversely, the correlation was much lower across samples when KSHV-infected samples were compared to UV-treated KSHV samples.

Because infected HFF cells proliferated much slower compared to the negative control cells, mRNA collected from the infected sample was the limiting factor in the analysis. Robust signal from both the control and KSHV-infected samples was collected for roughly 2,000 spots and would need to be repeated on a larger scale to achieve a better coverage. Nonetheless, SAM analysis yielded a number of real outliers, which upon further investigation may prove to be truly transcriptionally regulated by the viral effectors.

**IFI6 candidate from the screen has known roles in antiviral immune response**

One of the interesting candidates with diminished translation profile in the KSHV-infected sample versus the control was the gene called interferon alpha-inducible gene 6 (IFI6), also known as G1P3. It scored as the 36th out of the top 100 hits from the microarray with a SAM score of 1.672, meaning that on average its translation appeared to be 1.672 times suppressed relative to the control (Table 4.1). IFI6 is a poorly characterized gene to date. Emerging data is beginning to suggest that IFI6 encodes a low molecular weight mitochondrial protein that may stabilize mitochondrial function and oppose apoptosis\(^\text{14}\). As an interferon-stimulated gene, it is reported to influence the innate immune response to IFNs.

**Bioinformatic analysis of KSHV miRNA binding to IFI6.**

Given our data from the previous chapters on the translational regulation of AID by KSHV-encoded miRNAs, we became curious whether our array candidate IFI6 may also be regulated by the miRNAs from this cluster. To address this question we performed a bioinformatics analysis aligning all known KSHV miRNAs against the full length transcript of *IFI6*. Table A.2 depicts the top scoring hits from this analysis. A total of 4 viral miRNAs fell below our 25kCal/mol folding energy cutoff. They were predicted to bind both to the untranslated and the coding sequences of *IFI6* in one or more regions. While miRNAs most commonly affect expression of their targets through the interaction with the 3’ UTR, it may be worthwhile to follow up on all three modes of binding. A luciferase reporter encoding the relevant regions of IFI6 would be most suitable to verify the predicted interactions initially. Any of the scoring hits can then be further validated by expressing a full length *IFI6* mRNA and assaying its protein levels in the presence of KSHV miRNAs.
Figure A.2. Titration of KSHV infection. A) Immunofluorescent staining of HHF cells infected with KSHV-containing supernatant harvested from reactivated BCBL-1 cells. DAPI (blue) stain marks the nuclei and reflects the total number of cells present in the field of view. LANA (green) staining marks the infected cells to aid in calculating the percent infected cells from total number of cells stained. B) Immunofluorescent staining of HHF cells infected with equal volumes of untreated KSHV-containing supernatant harvested from reactivated BCBL-1 cells or supernatant treated with UV radiation to inactivate the virus. As in (A) DAPI stain marks the nuclei, LANA (green) marks the infected cells.
Figure A.3. The translational response of HFF cells to KSHV infection. Depicted are OD260 traces (arbitrary units, AU) of sucrose gradients (top of gradient is on the left) from either uninfected (bottom) or KSHV-infected (top) lysates. Alternating gray and white columns represent collected fractions. Specific ribosomal fractions are labeled above.
Figure A.4. **Pairwise Correlation in translational profiles.** Raw Log2 Cy5/Cy3 ratios for all array-analyzed points are plotted in a pairwise manner. 3 biological replicated of KSHV-infected samples are compared to two biological replicates of control UV-treated KSHV-infected samples. Greater Pearson R values reflect greater correlation. Outliers falling away from the straight line represent genes with most different translation profiles.
<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Name</th>
<th>probe.ID</th>
<th>Fold Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>H9R024176</td>
<td>Tubulin, beta</td>
<td>TUBB</td>
<td>-1.44746</td>
</tr>
<tr>
<td>H9R030700</td>
<td>Eukaryotic translation initiation factor 4A, isoform 1</td>
<td>EIF4A1</td>
<td>0.30503</td>
</tr>
<tr>
<td>H9R028398</td>
<td>Tubulin, beta 2C</td>
<td>TUBB2C</td>
<td>-0.70921</td>
</tr>
<tr>
<td>H9R031777</td>
<td>Annexin A2</td>
<td>ANXA2</td>
<td>0.25348</td>
</tr>
<tr>
<td>H9R032007</td>
<td>Heat shock 70kDa protein 8</td>
<td>HSPA8</td>
<td>-1.4146</td>
</tr>
<tr>
<td>H9R024724</td>
<td>FUS (involved in t(12;16) in malignant lymphoma)</td>
<td>FUS</td>
<td>0.867324</td>
</tr>
<tr>
<td>H9R013424</td>
<td>Peptidylprolyl isomerase B (cyclophilin B)</td>
<td>PPIB</td>
<td>0.53878</td>
</tr>
<tr>
<td>H9R029742</td>
<td>Source</td>
<td>s4d-14295043</td>
<td>2.12358</td>
</tr>
<tr>
<td>H9R022933</td>
<td>JAM A/C</td>
<td>ALCAM</td>
<td>0.31420</td>
</tr>
<tr>
<td>H9R028129</td>
<td>Annexin A2 pseudogene 1</td>
<td>ANXA2P1</td>
<td>0.24576</td>
</tr>
<tr>
<td>H9R030975</td>
<td>Ribosomal protein, large, P0</td>
<td>RPLP0</td>
<td>0.69158</td>
</tr>
<tr>
<td>H9R027451</td>
<td>Ribosomal protein L8</td>
<td>RPL8</td>
<td>1.17687</td>
</tr>
<tr>
<td>H9R030511</td>
<td>60S ribosomal protein L39</td>
<td>RPS3A</td>
<td>0.93319</td>
</tr>
<tr>
<td>H9R011606</td>
<td>Inhibitor of DNA binding 3, dominant negative helix-loop-loop</td>
<td>IKBKCI</td>
<td>0.25591</td>
</tr>
<tr>
<td>H9R028421</td>
<td>Annexin A2 pseudogene 2</td>
<td>ANXA2P2</td>
<td>0.05820</td>
</tr>
<tr>
<td>H9R031963</td>
<td>Ribosomal protein L7</td>
<td>RPL7</td>
<td>-1.12711</td>
</tr>
<tr>
<td>H9R031958</td>
<td>Source</td>
<td>s4d-14296689</td>
<td>1.577248</td>
</tr>
<tr>
<td>H9R031440</td>
<td>Source</td>
<td>s4d-1430428</td>
<td>0.374447</td>
</tr>
<tr>
<td>H9R033324</td>
<td>Mysin phosphate-ribo interacting protein</td>
<td>MRP</td>
<td>0.42367</td>
</tr>
<tr>
<td>H9R030211</td>
<td>Source</td>
<td>s4d-1429937</td>
<td>-0.67727</td>
</tr>
<tr>
<td>H9R018884</td>
<td>Spermaticg protein POZ</td>
<td>PSPO</td>
<td>2.37346</td>
</tr>
<tr>
<td>H9R037313</td>
<td>G protein-coupled receptor 161</td>
<td>GPR161</td>
<td>1.33382</td>
</tr>
<tr>
<td>H9R049987</td>
<td>M2 macrophage nuclear ribonucleoprotein D3, polypeptide 18kDa</td>
<td>MRNAD3</td>
<td>1.92365</td>
</tr>
<tr>
<td>H9R031224</td>
<td>PTP alpha (gene sequence 28)</td>
<td>PTPMA</td>
<td>-0.73161</td>
</tr>
<tr>
<td>H9R010842</td>
<td>G protein-coupled receptor 124</td>
<td>GPR124</td>
<td>0.51321</td>
</tr>
<tr>
<td>H9R030237</td>
<td>E3 ubiquitin-protein ligase (ubiquitin conjugation protein 1, 130kDa)</td>
<td>UBE3A</td>
<td>-0.13632</td>
</tr>
<tr>
<td>H9R040667</td>
<td>Proteasome (prosome, macropain) subunit, beta type</td>
<td>PSMB7</td>
<td>-1.17575</td>
</tr>
<tr>
<td>H9R016297</td>
<td>Zinc finger protein</td>
<td>ZNF623</td>
<td>-0.19967</td>
</tr>
<tr>
<td>H9R011509</td>
<td>Cortactin</td>
<td>CTN</td>
<td>0.01034</td>
</tr>
<tr>
<td>H9R029394</td>
<td>C11orf51</td>
<td>C11orf51</td>
<td>-0.07298</td>
</tr>
<tr>
<td>H9R029533</td>
<td>Source</td>
<td>s4d-1420121</td>
<td>-0.49957</td>
</tr>
<tr>
<td>H9R030525</td>
<td>Source</td>
<td>s4d-1407007</td>
<td>0.867615</td>
</tr>
<tr>
<td>H9R029714</td>
<td>Thrombomodulin</td>
<td>THBD</td>
<td>1.46894</td>
</tr>
<tr>
<td>H9R019757</td>
<td>beta-2-microglobulin</td>
<td>B2M</td>
<td>-0.24352</td>
</tr>
<tr>
<td>H9R025218</td>
<td>Interferon, alpha-inducible protein (clone IFI1-16)</td>
<td>GIP3</td>
<td>0.326901</td>
</tr>
<tr>
<td>H9R011515</td>
<td>Chromosome 10 open reading frame 56</td>
<td>C10orf56</td>
<td>-1.53038</td>
</tr>
<tr>
<td>H9R007085</td>
<td>Microtubule-actin crosslinking factor 1</td>
<td>MACF1</td>
<td>0.136204</td>
</tr>
<tr>
<td>H9R034375</td>
<td>Myosin, light polypeptide 6, alkali, smooth muscle actin</td>
<td>MYL6</td>
<td>1.54695</td>
</tr>
<tr>
<td>H9R003313</td>
<td>Caldesmon 1</td>
<td>CSTN1</td>
<td>0.344177</td>
</tr>
<tr>
<td>H9R050307</td>
<td>Source</td>
<td>s4d-1474210</td>
<td>1.64605</td>
</tr>
<tr>
<td>H9R006611</td>
<td>Dynamin, light chain, TIO1 type 1</td>
<td>DYNLT1</td>
<td>-0.51305</td>
</tr>
<tr>
<td>H9R033499</td>
<td>Eukaryotic translation elongation factor</td>
<td>EEF1D</td>
<td>-0.07577</td>
</tr>
<tr>
<td>H9R010124</td>
<td>Golgi autografting, golgin subfamily a, 3</td>
<td>GOLGA3</td>
<td>-0.30526</td>
</tr>
<tr>
<td>H9R039398</td>
<td>Ikr-1-like 3, myosin (C. elegans)</td>
<td>FERTL3</td>
<td>-1.012</td>
</tr>
<tr>
<td>H9R029311</td>
<td>Family with sequence similarity 50, member A</td>
<td>FAM50A</td>
<td>0.815705</td>
</tr>
<tr>
<td>H9R030561</td>
<td>Annexin A/C</td>
<td>ALN</td>
<td>0.779957</td>
</tr>
<tr>
<td>H9R009182</td>
<td>Pak1</td>
<td>PAK1</td>
<td>-0.39158</td>
</tr>
<tr>
<td>H9R018414</td>
<td>Peroxidase</td>
<td>PRDX</td>
<td>0.24286</td>
</tr>
<tr>
<td>H9R013917</td>
<td>Rhasa-rhms domain-containing protein</td>
<td>RANBP2</td>
<td>-0.56866</td>
</tr>
<tr>
<td>H9R030246</td>
<td>Acidic (leucine-rich) nuclear phosphoprotein 32 family</td>
<td>ARP2/3B</td>
<td>0.57128</td>
</tr>
<tr>
<td>H9R014457</td>
<td>Immediate early response 3</td>
<td>ERI3</td>
<td>0.822</td>
</tr>
<tr>
<td>H9R012863</td>
<td>Exocyst complex component 7</td>
<td>EXOC7</td>
<td>0.143697</td>
</tr>
<tr>
<td>H9R016836</td>
<td>Transcription factor 20 (AR1)</td>
<td>TCF20</td>
<td>-1.51592</td>
</tr>
<tr>
<td>H9R031800</td>
<td>Chromatin remodeling homolog 6</td>
<td>CHD6</td>
<td>-1.25181</td>
</tr>
<tr>
<td>H9R029474</td>
<td>Statmin-like 3</td>
<td>STMN3</td>
<td>-0.20482</td>
</tr>
<tr>
<td>H9R008229</td>
<td>Arp2/3 compnent 25</td>
<td>ANKRD25</td>
<td>-0.4026</td>
</tr>
<tr>
<td>H9R032436</td>
<td>Acidic (leucine-rich) nuclear phosphoprotein 32 family</td>
<td>ARP2/3</td>
<td>-0.52558</td>
</tr>
<tr>
<td>H9R014457</td>
<td>Immediate early response 3</td>
<td>ERI3</td>
<td>-0.922</td>
</tr>
<tr>
<td>H9R013098</td>
<td>Exocyst complex component 7</td>
<td>EXOC7</td>
<td>0.143697</td>
</tr>
<tr>
<td>H9R016836</td>
<td>Transcription factor 20 (AR1)</td>
<td>TCF20</td>
<td>-1.51592</td>
</tr>
<tr>
<td>H9R031800</td>
<td>Chromatin remodeling homolog 6</td>
<td>CHD6</td>
<td>-1.25181</td>
</tr>
<tr>
<td>H9R029474</td>
<td>Statmin-like 3</td>
<td>STMN3</td>
<td>-0.20482</td>
</tr>
<tr>
<td>H9R008229</td>
<td>Arp2/3 compnent 25</td>
<td>ANKRD25</td>
<td>-0.4026</td>
</tr>
<tr>
<td>H9R032436</td>
<td>Acidic (leucine-rich) nuclear phosphoprotein 32 family</td>
<td>ARP2/3</td>
<td>-0.52558</td>
</tr>
<tr>
<td>H9R014457</td>
<td>Immediate early response 3</td>
<td>ERI3</td>
<td>-0.922</td>
</tr>
<tr>
<td>H9R013098</td>
<td>Exocyst complex component 7</td>
<td>EXOC7</td>
<td>0.143697</td>
</tr>
<tr>
<td>H9R016836</td>
<td>Transcription factor 20 (AR1)</td>
<td>TCF20</td>
<td>-1.51592</td>
</tr>
<tr>
<td>H9R031800</td>
<td>Chromatin remodeling homolog 6</td>
<td>CHD6</td>
<td>-1.25181</td>
</tr>
<tr>
<td>H9R029474</td>
<td>Statmin-like 3</td>
<td>STMN3</td>
<td>-0.20482</td>
</tr>
<tr>
<td>H9R008229</td>
<td>Arp2/3 compnent 25</td>
<td>ANKRD25</td>
<td>-0.4026</td>
</tr>
<tr>
<td>H9R032436</td>
<td>Acidic (leucine-rich) nuclear phosphoprotein 32 family</td>
<td>ARP2/3</td>
<td>-0.52558</td>
</tr>
<tr>
<td>H9R014457</td>
<td>Immediate early response 3</td>
<td>ERI3</td>
<td>-0.922</td>
</tr>
<tr>
<td>H9R013098</td>
<td>Exocyst complex component 7</td>
<td>EXOC7</td>
<td>0.143697</td>
</tr>
<tr>
<td>H9R016836</td>
<td>Transcription factor 20 (AR1)</td>
<td>TCF20</td>
<td>-1.51592</td>
</tr>
<tr>
<td>H9R031800</td>
<td>Chromatin remodeling homolog 6</td>
<td>CHD6</td>
<td>-1.25181</td>
</tr>
</tbody>
</table>
Table A.1. Top 100 hits with greatest difference in translation between KSHV-infected and control samples (UV-irradiated KSHV). Table Columns can be interpreted as follows: 1) Probe.ID is a HEEBO identifier and is unique for each probe. A given gene “name” or “Gene.symbol” may have several different corresponding probes. 2) Name - full gene name. 3) Gene symbol. 4) ks1 - raw log2 Cy5/Cy3 ratios for KSHV-infected sample. 5) - 8) are the KSHV-infected or UV-irradiated KSHV-infected replicates. All values are mean centered, thus the “mc” prefix indicates values relative to ks1. 9) SAM score reflects the difference in means/standard deviation. 10) FDR - false discovery rate (% of time the outcome is falsely identified). 11) fold change. 12) q-value.
<table>
<thead>
<tr>
<th>miRNA Name &amp; Sequence</th>
<th>miRNA-Target Alignment</th>
<th>Folding Energy (Kcal/mol)</th>
<th>Targeted Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>kshv-miR-K12-10a TAGTGTTGTCCCCCGAGTGCC</td>
<td>---------------CGGGGGGCCTAGTGCCGACCAGCTG</td>
<td>-33.4</td>
<td>ORF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kshv-miR-K12-8-3p CUAGGCUCUGAGAGAGCA</td>
<td>C-TTCTCTCCTCCAAG-CTCAGTAG</td>
<td>-27.6</td>
<td>5'UTR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kshv-miR-K12-12-3p UGGGGGAGGGGUGCUGUUGA</td>
<td>GCTCCGAGGAACTCTCTTCTGCT</td>
<td>-25.9</td>
<td>3'UTR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kshv-miR-K12-9-3p CUGGGCUAUACGCACUGCUUGUA</td>
<td>GCAAGCTAAGAAAGTGCCTCGAG</td>
<td>-25.5</td>
<td>ORF</td>
</tr>
</tbody>
</table>

Table A.2. Bioinfomatic analysis of KSHV miRNAs binding to IFI6. First column depicts KSHV-encoded miRNA name and mature miRNA sequence. Second column displays the details of the predicted base-pairing, where the target mRNA sequence is on the top and miRNA is on the bottom. Third column lists the folding energy for the predicted interaction below the 25 (Kcal/mol) cut-off. Last columns places the predicted miRNA interaction within the context of the full length mRNA transcript by specifying whether predicted binding occurs within the 5’ or the 3’ untranslated regions (UTR) or the open-reading frame (ORF).

Discussion

A growing abundance of literature now supports the notion that measuring the mRNA expression levels provides an incomplete picture of the cellular phenotype under investigation because transcriptome does not accurately represent the proteome. Hence, to better understand specific cellular changes occurring upon infection with KSHV a more global proteomic approach must be taken. In this appendix we summarize the preliminary data comparing translational profiles of KSHV-infected and uninfected HFF cells. Following 5 day infection with KSHV, these cells establish a robust latent infection mimicking the natural progression of KSHV infection in vivo. Our data demonstrate that gradient fractionation coupled with the microarray analysis is an effective strategy to characterize many actively translated cellular transcripts. We reveal numerous candidates with inhibited translation, among which IFI6 hold particular promise given its suspected role in innate immunity.

A previous report looking at transcriptional reprogramming in endothelial, fibroblast, and B cells early upon KSHV infection revealed changes in the expression in
nearly two percent of all genes examined (~22,000 total)\textsuperscript{16}. The study noted cell-specific effects, documenting that only a fraction of affected genes was common to all the cell lines, while the majority was unique to one or two of the three cell lines. Although, KSHV preferentially establishes latency upon infection, during the initial hours of infection there is a detectable burst of lytic gene transcription. It is thus, not very surprising that host gene expression analysis performed at such early time point is likely to yield numerous affected host targets. Moreover, some transcriptional changes occurring soon post-infection may be a reflection of the viral sensing and cellular activation by the pattern recognition receptors rather than a deliberate reprogramming by the viral products. Therefore, upon establishment of the physiologically relevant long-term latency many of such targets may return to normal levels, while other previously unaffected genes could in turn change in their expression. Our approach allowed infected cells to reach as state of equilibrium, where changes in translational profiles of infected cells were likely a result of latent viral gene/miRNA activity.

One of the interesting hits uncovered in our screen is IFI6. While the protein remains poorly characterized, several reports generated evidence in support of its role in immunity, especially in the context of viral infections. One of the studies documenting the role of IFI6 in immunity conducted a large screen testing the ability of over 380 human interferon-stimulated genes (ISGs) to inhibit the replication of several important animal viruses\textsuperscript{17}. IFI6 appeared to exhibit a targeted antiviral activity by inhibiting replication of the yellow fever virus (YFV). Another study demonstrated that IFI6 expression upon interferon treatment of hepatitis C patients was also shown to correlate well with early virological response and successful control of viremia\textsuperscript{18}. IFI6’s induction by the interferon signaling along with the preliminary evidence for its role in restricting viral replication make it an attractive candidate for further investigation into its function in KSHV pathogenesis and immunity in general.

Additional correlative evidence for the role of IFI6 in immunity comes from a study examining genetic risk factors for lupus\textsuperscript{19}. In this study, researchers find that the risk variant MECP2/IRAK1 gene correlates with hypomethylation status of several genes including IFI6. Hypomethylation is known to correlate with increased gene transcription. Hence, the correlation of hyperactivity of IFI6 and increased risk of autoimmune disorder like lupus implies the role of IFI6 in mediating immune responses. Altogether, the documented reports of IFI6 expression provide support for a critical role in immunity with particular emphasis on antiviral response.

If translational inhibition of IFI6 by KSHV reproduces upon future analysis, we can suspect that IFI6 plays a significant part in antiviral defense against KSHV as precedents suggest. Devoting genomic space to encode a factor which may restrict expression of a particular host gene signifies this gene’s efficacy in hampering the viral fitness. It would be curious to explore whether bioinformatics predictions for KSHV-encoded miRNAs targeting of IFI6 are corroborated by the experimental evidence. Furthermore, it should be established whether this potential inhibition occurs via the direct predicted interaction or via a different intermediary.
Materials and Methods

Tissue Culture

HFF cells were cultured in DMEM, 10% FBS and 1% penicillin and streptomycin. BCBL-1 cells were cultured in RPMI 1640, 10% FBS and 1% penicillin and streptomycin.

Viral Production and Infection

To produce KSHV BCBL-1 cells were grown to 500,000 cells/ml confluence and reactivated with 0.3 mM sodium butyrate for 5 days. Virus-containing supernatant was filtered through .45um filter and used to infect HFFs via 2hr spinfection at 1,000 x g. To generate UV-inactivated virus, viral supernatant was placed into uncovered tissue culture dish and exposed to UV light in the tissue culture hood for 30 minutes. To prevent media evaporation, virus containing dish was placed into a larger water-containing dish for the duration of treatment.

Preparation of lysates for Gradient Encoding for adherent cells

At least 10 x 10^6 cells were used per sample per gradient. Cells were washed twice in ice-cold TKM buffer (20mM Tris pH8,140mM KCl, 5mM MgCl2 in RNase-free water) + cycloheximide (100ug/ml). Following the second wash, buffer was aspirated and cells scraped gently into eppendorf tubes and spun at 500rpm on benchtop centrifuge at 4C. Excess wash media was removed and cells resuspended at a final density of ~ 30 – 40 x10^6 cells/mL in ice-cold TKM + cycloheximide buffer with the following added fresh (0.1% Brij 58, 0.1% Na Deoxycholate, 0.22 mg/mL heparin,0.5mM DTT, RNase and protease inhibitors). Lysates were incubated on ice for 15 min flicking tube occasionally and spun for 5min at 9000 rpm in eppendorfs to clear nuclei from lysate at 4C. Supernatants were transferred to fresh tubes and OD was measured to ensure collection of at least 150 ug RNA in no more than 500 uL for each gradient. Samples were snap frozen and stored at ~80C for gradient analysis.

Sucrose Gradient Preparation

Sucrose gradients were prepared using the Gradient Master (Biocomp) according to the manufacturer’s suggestions. Five percent and 60% (w/v) sucrose solutions were prepared by dissolving sucrose in Gradient Buffer (20 mM Tris-HCl pH 8.0, 140 mM KCl, 5 mM MgCl2, 0.5 mM DTT, 0.1 mg/ml cycloheximide) at room temperature. The 60% solution was dispensed into an SW41 ultracentrifuge tube through a cannula underneath the 5% solution. Using an 11-step program (Biocomp, SW41 SHORT SUCR 5–50 11), the two solutions were mixed on the Gradient Master to form a linear gradient. After preparation, gradients were chilled for 30 min at 4C.

Sucrose Gradient Velocity Sedimentation

Immediately before centrifugation ~3.75 A260 units of lysate were transferred to the surface of the gradient after an equal volume was removed from the top of the gradient. Gradients were centrifuged at 41,000 rpm (RCFave =207,000) for 70 min at 4C using a SW41 rotor and then stored at 4C until fractionation. The Gradient Station (Biocomp)
trumpet tip was lowered into the ultracentrifuge tube at a rate of 0.17 mm per second. Fractions (12 drops each, ~550 ml) were collected into a 96-well plate containing 600 ml of lysis solution with b-mercaptoethanol added) using a fraction collector (Foxy Jr.). The absorbance of the gradient at 260 nm was measured during fractionation using a UV6 system (Teledyne-Isco).

**Gradient Encoding**

*In vitro* transcribed mRNAs with 25 nt polyA tails derived from the Methanococcus jannaschii genome were added to each encoded fraction at 100pg per mRNA such that each fraction contained 4-to-6 control mRNAs recognized by 16-to-24 unique probes on the MEEBO DNA microarray. 20 fractions were separated into pools A & B, and the RNA was harvested as per the invitrogen Purelink Mini RNA purification kit for liquid samples (cat.# 12183018A) with the exception that the lysis buffer:lysatE:EtOH were in a 1:1:2 stoichiometry before loading onto the column. 1 ug of purified RNA was amplified using the Amino Allyl MessageAmp II aRNA kit and labeled with Cy5 for pool A and Cy3 for pool B.

**DNA Microarray Production and Prehybridization Processing**

MEEBO oligonucleotide microarrays were printed on epoxysilane-coated glass (Schott Nexterion E) by the Stanford Functional Genomic Facility. The MEEBO microarrays contain ~39,000 70-mer oligonucleotide probes, representing ~30,000 unique genes. A detailed description of this probe set can be found at http://www.microarray.org/sfgf/meebo.do. Prior to hybridization, slides were first incubated in a humidity chamber (Sigma Cat# H6644) containing 0.56SSC (16SSC = 150 mM NaCl, 15 mM sodium citrate [pH 7.0]) for 30 min at room temperature. Slides were snap-dried at 70–80C on an inverted heat block. The free epoxysilane groups were blocked by incubation with 1 M Tris-HCl (pH 9.0), 100 mM ethanolamine, and 0.1% SDS for 20 min at 50uC. Slides were washed twice for 1 min each with 400 ml of water, and then dried by centrifugation. Slides were used the same day.

**DNA Microarray Sample Preparation, Hybridization, and Washing**

Poly-adenylated RNAs were amplified in the presence of aminoallyl-UTP with Amino Allyl MessageAmp II aRNA kit (Ambion Cat# 1753). For mRNA expression experiments, universal human reference RNA was used as an internal standard to enable reliable comparison of relative transcript levels in multiple samples. Amplified RNA (5–10 mg) was fluorescently labeled with NHS-monoester Cy5 or Cy3 (GE HealthSciences Cat# RPN5661). Dye-labeled RNA was fragmented (Ambion Cat# 8740), then diluted into in a 50-ml solution containing 36 SSC, 25 mM Hepes-NaOH (pH 7.0), 20 mg of human Cot-1 DNA (Invitrogen Cat# 15279011), 20 mg of poly(A) RNA (Sigma Cat# P9403), 25 mg of yeast tRNA (Invitrogen Cat# 15401029), and 0.3% SDS. The sample was incubated at 70uC for 5 min, spun at 14,000 rpm for 10 min in a microcentrifuge, then hybridized at 65uC using the MAUI hybridization system (BioMicro) for 12–16 h.

For gradient encoding experiments, each fraction was divided into two pools, “A” and “B”. Amplified RNA from pools A was fluorescently labeled with NHS-monoester Cy5,
and RNA from pools B was fluorescently labeled with NHSmonoester Cy3. Amplified RNA from pools A and B were comparatively hybridized to a DNA microarray to obtain each mRNA’s average sedimentation within the gradient.

Following hybridization, microarrays were washed in a series of four solutions containing 400 ml of 26 SSC with 0.05% SDS, 20058 SSC, 16SSC, and 0.26SSC, respectively. The first wash was performed for 5 min at 65uC. The subsequent washes were performed at room temperature for 2 min each. Following the last wash, the microarrays were dried by centrifugation in a low-ozone environment (<5 ppb) to prevent destruction of Cy dyes. Once dry, the microarrays were kept in a low-ozone environment during storage and scanning.

**Scanning and Data Processing**

Scanning and Data Processing were performed as previously described. The Cy5/Cy3 ratio for the encoded samples were related to ribosome number as previously described, briefly below. Log2(Cy5/Cy3) ratios from the oligos representing the *in vitro* transcribed doping controls were normalized by linear regression to the expected ratios of doping controls from each fraction, allowing empirical Cy5/Cy3 ratios to be expressed as a function of fraction number. The mid-point of each fraction was expressed in terms of time and related to the Log of the midpoint of each A260 peak expressed in time – allowing fraction number to be expressed in terms of ribosome peaks, and thus Cy5/Cy3 ratios to be related to average ribosome numbers. For SAM, unpaired two-class t-tests were performed with default settings with mean centering applied to all arrays (R-package samr; [http://cran.rproject.org/web/packages/samr/index.html](http://cran.rproject.org/web/packages/samr/index.html)). Only non-control microarray features that passed quality filtering in all replicates were used in the analyses. GO term and KEGG pathway analysis was performed using Gene Trail software, which can be found at [http://genetrail.bioinf.uni-sb.de/](http://genetrail.bioinf.uni-sb.de/) using overrepresentation analysis with the set of all genes for which translation data were acquired as reference. Settings were all default using the FDR adjustment for multiple testing.

**Target Screening**

We used a publicly available search engine, RNA22 ([http://cbcsrv.watson.ibm.com/rna22.html](http://cbcsrv.watson.ibm.com/rna22.html)), to obtain putative binding information between miRNAs and the target. For sites predicted by RNA22, we considered heteroduplexes with maximum folding energy of -25 Kcal/mol and minimum 14 paired bases.
References


