Regulation of Replication and Translation During KSHV Life Cycle

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Cellular and Molecular Biology

By
Prerna Dabral

Dr. Subhash C. Verma/Dissertation Advisor
August 2019
THE GRADUATE SCHOOL

We recommend that the dissertation prepared under our supervision by

PRERNA DABRAL

Entitled

Regulation of Replication and Translation during KSHV life cycle

be accepted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Dr. Subhash Verma, Advisor

Dr. David AuCoin, Committee Member

Dr. Cyprian Rossetto, Committee Member

Dr. Paul Sumby, Committee Member

Dr. Dean Burkin, Graduate School Representative

David W. Zeh, Ph.D., Dean, Graduate School
August, 2019
**Abstract**

Kaposi’s sarcoma-associated herpesvirus (KSHV), also known as human herpesvirus-8 (HHV-8), is a γ2-lymphotropic virus of the herpesvirus family that establishes asymptomatic infections in immunocompetent hosts but causes complications including tumors in an immunocompromised individual. KSHV is associated with various malignancies like Kaposi’s Sarcoma (KS), Primary Effusion lymphoma (PEL), and a plasmablastic form of Multicentric Castleman’s Disease (MCD). Like other members of the Herpesvirus family, KSHV’s life cycle consists of two phases; the latent and the lytic phase. During latency, the virus establishes a lifelong persistent infection in the host which cannot be eliminated as the virus interferes with the immune surveillance mechanism of the host. Because of the compromised immunity of the affected individuals and various external stimuli, the virus undergoes lytic reactivation leading to the production of infectious virions and tumorigenesis.

Since latency is described as a dormant phase, virus persists in the infected host with restricted gene expression and viral replication to maintain itself in dividing tumor cells. A major latency associated protein, Latency Associated Nuclear Antigen (LANA) plays a significant role in maintaining the viral genome and regulating the growth of tumor cells. To achieve a better understanding of latency, it was essential to study how LANA interacts with cellular proteins and facilitates DNA replication. Our work determined the interaction between LANA and cellular helicases, the MCM complex in detail, where we determined the interactions of the MCM complex with LANA in different cell cycle phases. We also identified the domain of LANA that facilitated its interaction with MCM6 and was critical for latent DNA replication and maintenance of the viral genome. In addition, we demonstrated how MCMs are
absolutely required for DNA replication and persistence of the viral genome along with their association with replicating DNA.

Immune evasion is critical for maintaining viral persistence in the host for a lifetime, which also contributes in establishing a successful KSHV infection to a great extent. LANA, the master regulator of latency, has been shown to possess immunomodulatory functions. Studies described in Chapter 2, add to the existing knowledge about immune modulation during latency and highlight the role of secondary RNA structures called G-quadruplexes in regulating mRNA translation and antigen presentation of LANA. We successfully characterized the formation of RNA G-quadruplexes in LANA mRNA and their effect on the translation of LANA. The limited pool of LANA in cells leads to the inhibition of antigen presentation of LANA peptides, thereby thwarting the possibility of an immune response. We also demonstrated that a G-quadruplex binding protein; hnRNPA1, possibly destabilizes these structures and enhances the expression of LANA. Moreover, we share a brief insight into the autoregulation of protein expression by LANA through binding to the G-quadruplex forming RNA and reducing their export into the cytoplasm.

Another critical aspect of KSHV pathogenesis involves lytic DNA replication, which leads to the synthesis of viral genome and production of viral proteins with roles in enhancing and promoting virion assembly. The initiation of lytic DNA replication occurs through the cis-acting oriLyt region along with the involvement of viral and cellular proteins. In Chapter 3, we discovered the presence of G-quadruplexes in the oriLyt, which were important for the initiation of lytic DNA replication as stabilizing those G-quadruplexes led to a reduction in replication initiation and viral genome copies. In addition, we also determined the potential of RecQ, a
cellular helicase in facilitating lytic replication at these sites and how RecQ binding site in the oriLyt is crucial for lytic replication.
Acknowledgements

I would like to thank my advisor; Dr. Subhash Verma for being an exemplary mentor, whose support, guidance and encouragement has kept me going for the past five years. I admire his tenacity and assiduity towards his work and would treasure what I have learnt from him all my life. I cannot thank him enough for providing me with opportunities and encouragement throughout graduate school that has helped me become an independent scientist and identify my career goals.

I would like to extend my sincere thanks to my committee members for their valuable guidance and constructive assessment of my research work.

I am extremely grateful to Dr. Timsy Uppal for her timely assistance and inputs towards studies performed in this thesis. Special thanks to the present and previous members of my lab for their infinite support and insightful suggestions.

Finally, I would also like to extend my deepest gratitude to my family members and friends, especially my parents for their endless support and encouragement.
Introduction
Abstract
Acknowledgments
Background
References
Chapter 1: Mini-chromosome maintenance proteins cooperate with LANA during the G1/S phase of the life cycle to support viral DNA replication
Abstract
Significance
Introduction
Materials and Methods
Results
Discussion
Acknowledgements
References
Chapter 2: LANA and hnRNPA1 regulate the translation of LANA mRNA through G-quadruplex formation
Abstract
Significance
Introduction
Materials and Methods
Results
Discussion
Acknowledgements
References
Chapter 3: RecQ1 helicase unwinds the G-quadruplex structures at ori-Lyt to initiate lytic DNA replication
Abstract
Introduction
Materials and Methods
Results
Discussion
Acknowledgements
References
Conclusion
Introduction and Background

1. Herpesviruses

Herpesviridae, a family of Herpesviruses that establishes infection in humans and animals, consists of DNA viruses whose origin dates back to the ancient Greek civilization. The word herpes has been coined from the Greek word “Herpein”, which means to “creep and crawl” and refers to the infectious nature of the virus, referring to the spread of virions. These viruses infect cells where they mostly remain dormant and rely on cellular processes for DNA replication until they reactivate to produce infectious progeny virions. The virions of the Herpesviridae family primarily consist of four major components. The core consists of linear double-stranded non-segmented DNA genome that codes for over 90 genes. This core is packed in an icosahedral capsid with a diameter of approximately 100nm, which is followed by a protein matrix or tegument layer containing viral-encoded proteins. The tegument is enclosed in a final outermost layer, the lipid-containing envelope, which is acquired from the host cell membrane and contains glycoproteins essential for entry and virulence of the virus (1). Herpesviruses establish lifelong latency in their hosts but occasionally reactivate to cause severe pathologies.

Humans get infected by one or more herpesviruses in their lifetime, where the virus persist lifelong without any complications except in immunocompromised settings, where the virus leads to potentially severe outcomes including mortality. Based on the phylogenetic analysis, genome sequence and biological properties, the viruses in the Herpesviridae family have been classified into three sub-families: Alphaherpesvirinae, Betaherpesvirinae, and Gammaherpesvirinae (2, 3). These subfamilies vary with respect to the sites at which the virus establishes latency: alpha herpesviruses in neuronal cells, beta herpesviruses in epithelial cells, myeloid progenitors and/or lymphocytes, and the gamma herpesviruses in lymphocytes and
endothelial cells (4). Out of the 100 known species of herpesviruses, 8 are known to infect humans namely: herpes simplex virus 1 (HSV-1/ HHV-1), herpes simplex virus 2 (HSV-2/ HHV-2), varicella zoster virus (VZV/ HHV-3), Epstein-Barr virus (EBV/HHV-4), human cytomegalovirus (HCMV/HHV-5), human herpesvirus 6 (HHV-6), human herpesvirus 7 (HHV-7), and human herpesvirus 8 (KSHV/HHV-8) (5).

2. KSHV-Associated Diseases

Kaposi’s Sarcoma (KS) was first described in 1872 by Moritz Kaposi as “idiopathic multiple pigmented sarcoma of the skin” that was a rare vascular disease and occurred most commonly in elderly men of Eastern European, North American, Middle Eastern or Mediterranean heritage (6). The next occurrence of KS was observed between 1950 and 1970 in younger individuals in Central Africa and transplant recipient on immuno-suppressive therapies (7-9). In 1981, there was a rapid increase in the number of KS cases in homosexual men giving rise to a KS epidemic, which was later linked to AIDS pandemic, that ultimately led to the recognition of KS as an AIDS-associated malignancy. This was followed by the discovery of KSHV in 1994 by Patrick Moore and Yuan Chang in KS lesions of AIDS patients through a PCR-based technique, leading to the discovery of a new γ-herpesvirus (10). In addition to KS, KSHV has been identified as an etiological agent of Primary Effusion Lymphoma (PEL), Multicentric Castleman’s Disease (MCD) and KSHV-Inflammatory Cytokine Syndrome (KICS) and classified as an oncogenic virus.

2.1 Primary infection

In a healthy individual, primary infection of KSHV is generally asymptomatic. In some cases, clinical presentations could range from inflammation of the lymph nodes to mild flu-like symptoms, in the absence of an HIV co-infection. However, KSHV infection could lead to
various malignancies and be potentially lethal in immunocompromised individuals due to HIV infection or immunosuppressive therapies in organ transplant patients.

2.2 Kaposi’s Sarcoma (KS)

KS is an angio-proliferative tumor derived from spindle-shaped KSHV infected lymphatic-endothelial cells (11). They may appear as lesions in purple, red, or brown eruptions or discolorations on the skin, oral mucosa, lymph nodes, lungs or digestive tract (12, 13). Based on their epidemiology and clinical manifestations, KS is classified into four subtypes: Classic, Endemic, Iatrogenic and AIDS-associated.

Classic KS, which was originally described by Moritz Kaposi in 1872 in older men, is a slow disease characterized by painless purplish-red lesions on the lower extremities including legs, ankles or sole of feet. As the disease progresses, lesions could spread to the arms, hands and also the gastrointestinal tract (14). The Endemic form of KS is a more progressive and fatal subtype of KS, which usually affects lymph nodes and other organs. This type is prevalent in central and eastern Africa and commonly affects children and young adults, and manifests as nodules on feet and legs (15). “Iatrogenic” or “post-transplant” KS is an aggressive form of KS associated with organ transplant patients on immune-suppressive therapy and lesions may appear cutaneously or viscerally (16). AIDS-associated/Epidemic-associated KS is the second most common tumor associated with AIDS patients and is the most aggressive form of the disease. The extremely weakened state of the immune system in AIDS patients contributes to the risk of KS acquisition. This form displays varied clinical manifestations and lesions may appear on the skin, lymph nodes, lungs, spleen, liver and gastrointestinal tract (17, 18).
2.3 Primary Effusion Lymphoma (PEL)

PEL or Body cavity-based lymphoma (BCBL) is a rare and aggressive B cell tumor that affects immunocompromised or elderly individuals. KSHV is the etiological agent of PEL, which is most commonly found in AIDS patients often associated with a co-infection of Epstein Barr Virus (19). KSHV genomes are maintained at a relatively high copy number ranging from 50-150 viral genomes in PEL cells (19). The lymphoma develops as an effusion in body cavities including the peritoneum, pericardium, and pleurum but rarely in lymph nodes and other areas (20, 21). PEL associated tumors are extremely aggressive, progress rapidly and are highly resistant to chemotherapy with poor prognosis. The fatality rate associated with PEL is high and the mean patient survival time of six months from the time of diagnosis (22, 23).

2.4 Multicentric Castleman’s Disease (MCD)

Castleman’s disease, a rare polyclonal B-cell lymphoproliferative disorder, also known as angiofollicular lymph node hyperplasia, is characterized by angiosclerosis and vascular proliferation in the Germinal centres and can be localized (unicentric) or dispersed (multicentric) infecting many lymph nodes and spleen. KSHV is associated with the plasmablastic variant of this disease and is the etiological agent of MCD in more than 90% of the AIDS cases (24, 25). KSHV-MCD patients have high levels of viral loads in peripheral blood mononuclear cells and the treatment is hard due to the multicentric nature of the disease. Treatment with ganciclovir and antiviral therapy has shown a reduction in levels of the virus in blood (26, 27). Due to rapid progress and aggressive nature of this disease, it has a high mortality rate and patients die within 10 months of the diagnosis (28).
2.5 KSHV-inflammatory cytokine syndrome (KICS)

KICS is a recently described clinical manifestation in patients co-infected with HIV and KSHV, who display inflammatory cytokine storm and are positive for KS in skin, lymph nodes and other areas but pathologically negative for KS-MCD. This condition arises due to severe cytokine production as a result of excess viral KSHV lytic reactivation, which is evident through the detection of high viral loads in peripheral blood mononuclear cells and the elevated levels of vIL-6 (29, 30). The diagnosis of this syndrome is hard as the clinical symptoms such as high cytokine production and vIL-6 are often misled with sepsis (31).

3. KSHV Epidemiology, Pathogenesis and Current Treatments

The seroprevalence of KSHV differs substantially according to the geographic areas and sub-population. It ranges from 30 to 70% in sub-Saharan Africa where it is endemic and is one of the most common adult malignancies, 10-25% in Mediterranean areas, and is less than 10% in Asia, Northern Europe, North, Central and South America (32). The seroprevalence is markedly higher in MSM (men who have sex with men) and ranges from 25-65% in HIV-infected to 20-30% in HIV-uninfected individuals (33).

The modes of KSHV transmission vary according to its seroprevalence in the region and can occur through both sexual and non-sexual routes, with sexual transmission as the primary cause among MSMs in low seroprevalence regions. Organ transplantation is another mode of KSHV transmission where the organs from KSHV infected individual could carry the virus to the recipient (34). The exchange of bodily fluids has been identified as the most common route of KSHV transmission with sexual and non-sexual oral transmission. High levels of KSHV have been observed in saliva where the infected individuals shed virus in oropharynx following lytic
replication (35). In addition, the virus can be transmitted through other methods of fluid exchange; oral sex, blood transfusion and needle sharing (36, 37).

AIDS-associated KS incidences are reduced by 80% owing to the use of combination antiretroviral therapy (cART) as a result of the restoration of immune strength through eradication of replicating HIV. The use of cART has led to a decrease in the size of KS lesions in HIV positive individuals in addition to reducing the incidences of new KS infection. However, in resource-limited settings, where accessibility to cART is limited, AIDS-associated KS remains a major problem (38, 39). Antiviral therapy involving the use of lytic replication inhibitors; ganciclovir, foscarnet, and cidofovir, has been of limited success in KS patients. The use of ganciclovir and valganciclovir has been moderately successful in treating MCD patients. Some local treatment options such as surgical excision, radiation therapy, cryotherapy and intralesional chemotherapy (using vinblastine) have been used to treat mild cases of KS and have been successful in preventing the development of new KS lesions. Treatment of MCD currently involves the use of rituximab, a chimeric monoclonal antibody against CD20, though it exhibited side effects (32, 40, 41).

4. Genome Structure of KSHV

The genome structure of KSHV is phylogenetically similar to herpesvirus saimiri (HVS), another member of the Rhadinovirus genus (42-44). The length of the KSHV genome is approximately 160-170 kb, with the long unique region (LUR) of 140 kb that is known to code for more than 90 open reading frames (ORFs). The ORFs are named 4-75, according to their similarity with HVS genes, and the remaining 19 genes with no homology to HVS are labeled with a K prefix. The homologous genes mostly code for structural proteins or enzymes involved in viral replication and are conserved within the subfamilies of herpesviruses (45). The LUR is
flanked by terminal repeat (TR) region, which contains 801bp long repeats and does not code for any protein. The GC-content of this region is approximately 84.5% and the number of repeats varies among the isolates of KSHV, making the overall length of the KSHV genome variable (46). Few genes in LUR of KSHV encode proteins that are homologs of cellular proteins and contribute to KSHV pathogenicity. These viral genes are thought to have been picked up during virus-host co-evolution (47). The LUR also codes for non-coding RNAs like PAN RNA, and approximately 12 microRNAs (miRNA) (48-54).

5. KSHV Infection and Entry

KSHV has been shown to infect a broad range of cell types in vivo, including CD19+ peripheral blood B lymphocytes, endothelial cells, monocytes, epithelial cells, and keratinocytes, with B-cells and monocytes being the major repertoire of latent cells (55). The transcripts of many latent genes have been detected in these cells, including, ORF73 (LANA), K13 (v-FLIP), ORF72 (v-Cyclin D), K12 (Kaposin) and microRNAs (56). A small percentage i.e., 1-3% of PEL cells, undergo spontaneous reactivation, characterized by the expression of lytic genes (57). Furthermore, mesenchymal stem cells (MSCs) and CD34+ hematopoietic progenitor cells (HPCs) have been shown to be infected by KSHV in KS patients (58-60). Interestingly, KSHV has also been shown to infect T cells in vivo, but the infection does not result in the persistence of viral genome in high copies nor does it results in lytic reactivation (61).

KSHV can establish successful infections in vitro in a wider range of cells including epithelial, endothelial, epithelial, human B cells, fibroblast, CD34+ stem cell precursors of dendritic cells, and monocytes (62). Few cell lines like human dermal microvascular endothelial (HMVEC-d) cells, human foreskin fibroblast (HFF) cells, human embryonic kidney 293 cells (HEK-293), human umbilical vein endothelial (HUVEC) cells, telomerase-immortalized human
umbilical vein, and endothelial (TIVE) cells, human epithelial SLK cells, mouse fibroblasts, monkey kidney CV-1 cells, and monkey kidney epithelial VERO cells are good models to study latency as these cells maintain latent virus (63).

KSHV entry into the host cells requires the involvement of viral glycoproteins and host cells receptors, which occurs as a multi-step process. Firstly, the virus binds to cell surface receptors, which is accompanied by the initiation of signaling pathways in the host that makes the cells susceptible to infection by promoting additional steps of infection. The next step includes fusion of the viral and host plasma membrane or endocytosis, followed by disassembly of the capsid and delivery of the viral genome to the nucleus.

KSHV binds to cellular proteoglycan heparin sulfate, a glycosaminoglycan expressed commonly on cell surfaces, through glycoproteins gB (ORF8), gH (ORF22), ORF4, and gpK8.1, that mediates attachment to the host cell. This interaction promotes binding between virion glycoproteins and additional cell-specific receptors (64-67). Multiple interactions between cell surface receptors and viral glycoproteins are required to facilitate infection and the broad cell tropism of KSHV can be attributed to the ability of the virus to bind to a wide variety of receptors (68). KSHV infects B cells, macrophages and dendritic cells through binding to DC-SIGN, a cell receptor commonly used by other virus and KSHV glycoprotein gB (ORF8) has been shown to facilitate this interaction (69-72). Also, gB promotes binding of the virus to extracellular cell surface receptors, integrins α3β1, αVβ3, αVβ5, at its RGD motif and also possesses a disintegrin-like integrin-binding domain that interacts with α9β1 integrin receptor (68). EphA2, a member of the tyrosine kinase receptor family, has been reported to mediate entry of the virus through interaction with KSHV glycoproteins gH and gL (73). Additionally, xCT, a
12-transmembrane glutamate/cysteine exchange transporter, has also been shown to facilitate cell fusion and entry of KSHV (74).

Following the binding of glycoproteins to appropriate cell surface receptors of a permissive cell, the virus is endocytosed and the capsid is released into the cytosol, trafficked to the nucleus, where the viral genome is released (68, 75-77).

6. KSHV Life Cycle

KSHV infects a variety of cell types, including endothelial cells, B cells and monocytes, and establishes latency within 24 hours post infection. Like other herpesviruses, KSHV has two predominant phases in its life cycle-latent and the lytic phase. In an immunocompetent individual, latency (the default program of the virus) is characterized by the expression of a subset of viral genes and the maintenance of virus as an episome, without the production of infectious virions. However, the lytic phase of the viral life cycle is a transient, which is characterized by the expression of most the viral genes in an organized order and production of infectious virions. The tumor cells in KS, PEL, and MCD are mostly in latent state and usually, only 1-3% of the cells undergo lytic replication (78-80).

6.1. Viral Latency

Due to the absence of enzymes necessary for replication, the virus hijacks the host DNA replication machinery to facilitate its replication during latency (81). The virus maintains itself inside the host in multiple copies through tethering to the host chromosome and partitioning efficiently to daughter cells following cell division (82, 83). During latency, the virus expresses a limited number of genes owing to a transcriptionally repressive chromatin landscape that also give rise to a compact chromatin structure (84, 85).
As mentioned above, during latency, the virus expresses a limited number of proteins which aide in the persistence of the virus and help the virus evade host’s immune surveillance. These include, LANA (ORF73), vFLIP (ORF71), and vCyc (ORF72), which are transcribed from the latency locus. Other latent transcripts include, viral miRNAs and kaposins (ORFK12) (49, 86, 87). Viral interferon regulatory factor (vIRF-3, K10.5) are also expressed in latent PEL and MCD cells (88). In addition, latently-infected cells also express K1 and vIL6 at lower levels (89-92). Latency-associated nuclear protein or LANA, encoded by ORF73 is one of the most abundantly expressed proteins during latency and is considered as the master regulator of latency. LANA aides the virus in latent DNA replication through recruitment of the DNA replication proteins to the origin of latent replication, establishes contacts with host proteins to tether the viral genome to the host chromosome, and also ensures faithful partition of the viral genome to daughter cells following cell division. In addition, LANA regulates viral latency through the inhibition of lytic reactivation. During primary infection, LANA downregulates the expression of lytic genes through the recruitment of host polycomb repressive complexes (PRC1 and PRC2) at the promoter regions (93, 94). LANA also interacts with RTA (Replication and Transcriptional Activator, a lytic reactivation immediate-early (IE) gene), encoded by ORF50, both in vivo and in vitro. RTA modulates the switch between latency and the lytic cycle and is responsible for the initiation of the lytic cycle. LANA represses the promoter activity of RTA, thereby inhibiting its expression and also reducing the ability of RTA to activate its promoter (95).

6.2. Lytic Cycle

Following primary infection or the latent period, the virus can enter into the lytic cycle due to various stimuli including hypoxia and cell differentiation signals or external factors
involving the use of chemicals like sodium butyrate or valproic acid (96). The lytic phase of replication is characterized by the expression of all viral genes and replication of linear genome through rolling circle mechanism that results in long concatemeric viral DNA, which are cleaved into appropriately sized viral genomes and release of progeny virions (97-99). Lytic replication is important for KSHV pathogenesis as it facilitates the spread of the virus from one cell to another and is instrumental in KSHV-related oncogenesis. The proteins expressed during the lytic cycle possess oncogenic potential, induce cellular proliferation, and take over the immune system of the host (100-106). As mentioned earlier, the lytic cycle involves the expression of all genes of KSHV in a sequential manner categorized as IE, early (E), and late (L) lytic genes according to the sequence in which they are transcribed following the lytic cycle trigger due to external and internal stimuli (107). RTA, an IE protein, regulates the switch between latency and lytic replication and is sufficient for inducing lytic DNA replication.

The IE genes are the first set of genes expressed within 8 hours after lytic reactivation. These genes serve regulatory roles in directing the transcription of downstream genes (E and L lytic genes) to facilitate viral DNA replication and virion assembly. RTA (ORF50), is one of the earliest lytic KSHV proteins detected after reactivation and can simultaneously activate the transcription of numerous downstream viral or cellular target genes by directly binding to a sequence-specific RTA-responsive element in the viral gene promoters or by indirectly associating with cellular transcription factors (108-110). The E genes are expressed 10-24 hours post infection and perform a variety of functions such as DNA replication, expression of L genes and inhibition of antigen presentation. The L genes are expressed 48 hours post reactivation, which are comprised of viral structural proteins such as capsid and envelope glycoproteins (4, 107).
7. LANA

LANA, encoded by ORF73 of KSHV, expressed at high levels in latent cells, is a multifunctional nuclear protein with a molecular weight of 220–230 kDa (111-114). LANA is divided into three domains: the N-terminal domain, a central domain and the C-terminal domain. These domains are discovered to have different functions; the N-terminal domain includes a nuclear localization signal (NLS) and chromosome-binding site (CBS), the central domain consists of a glutamine-rich region, and the C-terminal domain consists of a leucine zipper motif and DNA binding domain, which mediates dimerization and DNA binding (115, 116). The 337 amino acids long N-terminal region is rich in serine/threonine, proline and basic residues and includes a bipartite NLS in 21-30aa and 41-47aa, where arginine/glycine residues of NLS can interact with the eukaryotic nuclear cytoplasmic receptor, importin β1 and facilitate nuclear translocation of the protein (117, 118). This domain also contains the chromatin binding site which is important for the tethering of LANA to the host chromosome through its interaction with histones (119). The 585 amino acids long central domain or the internal repeat domain (IRD) is highly polymorphic and consists of multiple repeats with hydrophilic amino acids like glutamine, glutamate, aspartic acid and leucine. The 240 amino acids long C-terminal domain containing hydrophobic and charged residues is crucial for chromatin binding, self-association of LANA, and binding to the viral DNA at the TR region. Also, LANA C-terminal amino acids 884–1089 are important for self-association of the C-terminus and binding of LANA to host chromosome (120).
8. Functions of LANA

8.1. Episome Persistence and Segregation

LANA is a multifunctional protein, which plays a vital role in the life cycle of the KSHV. KSHV exists as an episome, a closed circular extra chromosomal genome that stably maintains at a copy of around 50-100/cell in latently infected cells. In order for the virus to persist stably in the dividing tumor cells, the episomal DNA needs to replicate in synchrony with the host cellular replication in order to segregate efficiently into the daughter nuclei following mitosis. Virus lacking LANA cannot establish latent infection, which confirms the role of LANA in replication and segregation of the viral genome. Since KSHV lacks its own latent DNA replication machinery, LANA ensures that these functions are done efficiently by recruiting various cellular proteins.

As mentioned above, LUR of the KSHV is flanked by a GC-rich TR region, which also serves as the fusion site to circularize the genome. The TR region contains multiple unit length (801bp) terminal repeats, which contain LANA binding site (LBS1) and is bound by the C-terminus of LANA. Mutations in the LBS at 6, 7 and 8 residues $\text{CCC}_8$, significantly reduced the binding of LANA at the TR (121). Downstream of LBS1 is another LANA-binding site, LBS2, that binds to LANA with a lower affinity but LBS1 assists LANA in binding cooperatively to the LBS2 (122). A recent study reported the presence of a third LANA binding site, LBS3 that was earlier named as the 32 bp replication element (RE). The orientation of this site is reversed in comparison with LBS1 and LBS2. All three sites are located on the same side of the DNA where the LANA dimers interact through their hydrophobic interphase. These hydrophobic interactions stabilize the overall structure and resulting oligomers facilitate cooperative binding to lower affinity sites (123).
8.2. LANA-Host Protein Interactions

LANA interacts with a variety of host proteins to facilitate latent DNA replication and segregation of viral DNA into the daughter cells. The N-terminal 1-32 amino acid residues of LANA interact with H2A/H2B, which serves as an anchor to bind viral proteins to the host chromatin (120). LANA also binds to H2AX, an isoform of H2A, that binds to DNA breaks and enhances its phosphorylation. The phosphorylated form of H2AX binds to both LANA and TR, facilitating LANA-TR interaction, promoting episome persistence (124).

The N-terminal 5-22 amino acids, and C-terminal 936-1162 amino acids of LANA, are essential for binding to MeCP2 (methyl-CpG binding protein 2), a protein involved in transcription silencing of genes and activation of euchromatin genes. The tethering role of LANA was attributed to the fact that LANA localizes on human chromosome with the KSHV genome. LANA 1-15 amino acids are essential for binding of LANA to MeCP2 and this interaction directs LANA to the heterochromatin region (125).

DEK, a 43 kDa chromatin associated phosphoprotein, is another protein which interacts with LANA in the C–terminus. DEK facilitates the attachment of LANA C-terminus to the chromosome as the mutated form of LANA with a deletion in N-terminal residues localizes to the human and mouse chromosomes only in the presence of DEK protein (126). LANA also forms complexes with kinetochore proteins like CENP-F, a protein involved in the localization of kinetochore during mitosis, and Bub1, a kinase that serves as a spindle checkpoint and regulates the kinetochore-microtubule attachment. The interactions of LANA with these kinetochore proteins were confirmed by direct protein-protein interaction and localization assays (127). NuMA, an important nuclear mitotic apparatus protein, interacts with microtubules and dynein/dynactin and has many cell cycle regulatory functions. LANA binds to NuMA through its
C-terminal residues 840 and 963aa confirmed through immunoprecipitation assays. The two proteins co-localized in the same nuclear compartment during interphase. This dynamic interaction was lost during the beginning of mitosis but got re-established after the segregation of daughter cells (128).

The Bromodomain and Extra-Terminal Domain (BET) family of proteins is involved in various cellular functions, including gene transcription, meiosis, and cell cycle regulation. LANA 75 to 777 aa and the C-terminal amino acids 982 to 1162 are important for binding to BRD4 (129). The interaction of these proteins with LANA facilitates the progression of G1/S phase of the cell cycle (130). Additionally, Brd4 co-localizes with LANA on mitotic chromosomes, thus supporting the possibility of its role in LANA-mediated tethering to the host chromosome.

8.3. Viral DNA Replication

The nuclear matrix was reported to play an important role in latent DNA replication of the virus by potentially serving as a site for initiation of replication, which kicks off by recruitment of oriP to the nuclear matrix, followed by the self assembly of the pre-RC complex consisting of ORCs, Cdt1, Cdc6 and MCMs (131, 132).

The process of DNA replication requires the presence of enzymes such as polymerase and helicases. Since LANA has no enzymatic activity of its own, it achieves these functions through the help of the cellular enzymes. TopoIIβ regulates the topology of DNA by introducing double-stranded breaks. LANA interacts with TopoIIβ through 1-32 amino acid residues of the N-terminal domain. Mutations in LANA at these amino acid residues 5-7aa and 5-15aa led to an inhibition of TR-mediated replication. LANA recruits TopoIIβ at the TR and this association is essential for the enzymatic activity of TopoIIβ on KSHV DNA (133).
The interaction between LANA and replication factor C (RFC) and Proliferating cell nuclear antigen (PCNA) is important for efficient viral replication during latency. RFC is an important enzyme that catalyzes the loading and facilitates encircling of DNA by PCNA. LANA interacts with RFC and this interaction positively regulate the activity of RFC as LANA enhances the loading of PCNA on DNA (134).

9. G-quadruplexes in Viruses

G-quadruplexes were discovered in 1962 using in vitro assays when Gellert and colleagues studied the ability of guanylic acid through X-ray diffraction (135). G-quadruplexes are non-canonical nucleic acid structures that are formed in G-rich sequences of DNA or RNA, formed by stacking of G-tetrads or G-quartet on top of each other and stabilized by the presence of K+ ions (136, 137). A G-quartet is formed by hoogsteen hydrogen-bonding between four G-residues and is stabilized by the presence of K+ ion that neutralizes the negative charge of the oxygen atoms in this secondary structure (138). RNA G-quadruplexes are more structurally stable than the DNA G-quadruplexes and are formed readily in vivo (139-141).

G-quadruplexes have been identified at multiple regions within the genome such as the origin of replication, promoter regions, coding and non-coding regions, telomeres and the sites of recombination (142). The occurrence of these structures at multiple areas in the genome leads to a modulation of biological processes, including replication, transcription, translation, alternative splicing and recombination. Though studied extensively in the past, G-quadruplexes have recently started to gain importance for their role in stomach cancer, liver cancer, colon cancer and in several neurodegenerative diseases like amyotrophic lateral sclerosis (ALS), frontotemporal dementia (FTD) and Fragile X syndrome (FXS) (143-148).
G-quadruplexes of HIV genome have been extensively studied in its both RNA and DNA-based life cycles. G-quadruplexes are present in the proviral integrated form of the virus, RNA copies of the virus, long terminal repeats (LTRs) and nef gene that codes for a viral accessory protein (149-152). The important roles of G-quadruplexes in the retrovirus life cycle have paved ways for the developments of anti-HIV drugs.

Among the herpesvirus family, G-quadruplexes have been identified in the TR regions of HSV-1, in mRNA of latent nuclear protein of EBV and the TR region of KSHV (153-155). In these viruses, G-quadruplexes were found to regulate crucial processes of the viral life cycles, like viral DNA replication, translation of viral proteins, and regulation of antigen presentation. Additionally, G-quadruplexes are also found to inhibit the chromosomal integration of HHV-6 due to their presence in the telomere regions (156).

Moreover, G-quadruplexes are found in multiple regions of the genome of Human Papillomavirus and Adeno-associated viruses, in certain regions of the genes in Hepatitis B virus and Hepatitis C virus, and non-structural proteins of influenza and severe acute respiratory syndrome coronavirus (157-162). More recently, they have been identified in genomes of Zika virus, Ebola Virus, and Marburg virus (163-165).

Since the presence of G-quadruplexes in DNA or RNA is a steric roadblock for the movement on DNA/RNA polymerase or ribosome, a number of cellular proteins have been identified that bind specifically to these secondary structures and modulate their formation through either unwinding or stabilizing these structures. RecQ1, BLM, Pif1 and FANCJ are few examples of helicases that unwind G-quadruplexes and bind to either DNA or RNA G-quadruplexes or both (166). Proteins that bind to G-quadruplexes can either stabilize or destabilized them and vary according to their substrates, function and localization, some of these
include; hnRNPA1, nucleolin, nucleophosmin, BRCA1 tumor suppressor, and ribosomal protein(167).

The identification of G-quadruplexes in genomes of several organisms and their regulatory role in biological processes that have been linked to developments of disease, has paved the way for the development of drugs that target G-quadruplexes. This has led to the discovery of G-quadruplex ligands that are small molecules which bind to G-quadruplexes with high affinity as compared to double-stranded DNA and stabilize them. Some examples of these ligands include BRACO-19, TMPyP4, PhenDC3, and Pyridostatin.

10. Overview of the Thesis

Mini-chromosome Maintenance Proteins Cooperate with LANA During the G1/S Phase of the Cell Cycle to Support Viral DNA Replication

Following primary infection of KSHV in permissive cells, there is a short phase of expression of both latent and lytic genes, following which, the virus establishes a latent infection usually within 24-48 h in those infected cells. During latency, the viral genome undergoes circularization, interacts with cellular histones to undergo higher order chromatinization, which resembles nucleosomal structures and are maintained as episomes. Latency proteins promote tumorigenesis and are crucial for the development of KSHV-associated cancers without killing the cells, whereas, lytic oncogenic proteins secrete angiogenic factors causing cell death. Latent program of the virus ensures cell growth and survival by keeping lytic replication repressed and facilitating replication and persistence of the virus in the host. Since the virus does not express any replication protein during latency, it hijacks the cellular replication machinery to perform latent DNA replication once per cell cycle in order to maintain constant copies of the viral genome in each KSHV infected PEL cell.
LANA mediates tethering of the viral genome to the host chromosome through establishing multiple contacts between the host chromatin and the viral genome and also aids in latent DNA replication by interacting with cellular proteins involved in the DNA replication. MCM complex, a member of the pre-RC machinery, accumulates at the viral origin of latent replication and aids in latent DNA replication. LANA is thought to recruit the components of pre-RC at the TR through direct interaction with them; however, the exact mechanism of how MCMs are employed at the TR was not yet known.

In this chapter, we sought to shed light on the role of MCMs in latent replication and study the LANA-MCM interaction in detail. Since a previous study from our lab identified MCMs as LANA binding partners, we first started by confirming the binding between LANA and the members of MCM complex endogenously in PEL cells and through overexpression system. Among the members of MCM complex, MCM3, MCM4 and MCM6 bound specifically to LANA and this interaction was confirmed through immunofluorescence assay in PEL cells. Interestingly, we were able to detect this binding only in the G1/S phase of the cell cycle and not in the G2/M phase, leading us to speculate that LANA is differentially post-translationally modified in different phases of the cell cycle. Next, we mapped the domains of LANA involved in binding to the MCMs through immunoprecipitation assays and found that MCM3 and MCM4 bound to the N-terminal domain of LANA whereas MCM6 bound to both the N and C terminal domains. Additionally, 1-32 amino acids region of LANA served as the minimal binding domain to MCM3 and mutation of the 14th amino acid Threonine to alanine resulted in abrogation of this binding. This led us to propose that phosphorylation of this Threonine of LANA may be differentially modified in different cell cycle phases. Similarly, we also studied the LANA C-terminal binding to MCM6 and identified a 50 amino acid region (1100-1150aa) to be critical for
this interaction and latent DNA replication. Moreover, we generated a recombinant cell line harboring KSHV BAC with a deletion in LANA specific for amino acid region 1100-1150 and these cells displayed an inhibition in nascent DNA replication as well as reduced genome copies.

In order to test the functional relevance of MCMs with regard to latent DNA replication, we performed shRNA mediated knockdown of MCMs in 293L cells and KSHV infected cells. Knockdown of MCMs led to a reduction in latent DNA replication in transient replication assays, decrease in nascent DNA synthesis and lesser copies of the viral genome. Through our chromatin immunoprecipitation assays, we showed that LANA is absolutely required for DNA replication, as MCMs could not be recruited to the origin in cells where LANA was depleted. We then went on to show that MCMs were physically present on the viral DNA during replication through isolation of proteins on nascent DNA through which we were able to demonstrate that LANA, MCMs and other replication proteins are present on the viral DNA during latent DNA replication. Through this work, we were successfully able to study the LANA-MCM binding in detail, along with the role of MCMs in latent viral DNA replication.

**LANA and hnRNP A1 regulate the translation of LANA mRNA through G-quadruplex formation**

In order to establish a successful infection inside the host, the virus employs numerous strategies to evade immune surveillance system. During virus-host co-evolution, virus has acquired a number of genes from the host and these immunomodulatory genes make up about 25% of the viral genome. These genes code for proteins that are often homologs of cellular proteins that seize control of cellular machinery in order to persist in the host for life long. Despite encountering a robust immune system in a healthy individual, the virus escapes immune detection and the host is never able to get rid of the latent virus. MHC Class I and MHC Class II
antigen presentation regulate the adaptive immune response, which is required to control KSHV-associated tumorigenesis. Both forms of antigen presentation involve the processing of antigenic peptides and presentation to CD8+ T cells (MHC Class I) or CD4+ T cells (MHC Class II). The latency program of the virus is critical for tumorigenesis and relies on the ability of the virus to stay hidden from the host, which is usually achieved by limited gene expression, making fewer KSHV-specific peptides available for antigenic processing and presentation. LANA, the most abundantly expressed protein during latency, also possess the ability to modulate immune recognition. LANA has been shown to inhibit its transcription through autoregulating its own promoter and controlling translation through inhibition of proteasomal degradation. In addition, LANA has been shown to restrict MHC Class I antigen presentation through interfering with the translocation of peptides to the endoplasmic reticulum. In this chapter, we demonstrate a regulation in LANA mRNA translation leading to an altered antigen presentation. G-quadruplexes have gained a lot of attention due to their regulatory roles in biological processes and their association with human diseases. These secondary structures have been identified in EBNA1, a homologue of LANA in EBV, where the formation of G-quadruplexes led to a reduction in the translation of mRNA and inhibition of antigen processing. Recently, we reported the formation of these structures in the terminal repeat region of KSHV that led to a reduction in DNA replication and viral genomes copies.

Upon analysis of LANA gene sequence through G-quadruplex prediction software, the QE-rich region of LANA, was found to have a high propensity for the formation of these structures. We confirmed the formation of G-quadruplexes through CD spectroscopy and Electrophoretic Mobility shift assay on wild type and scrambled LANA RNA oligo, where wild type LANA RNA oligo displayed biophysical and biochemical characteristics distinct for RNA
G-quadruplexes. Functional relevance of G-quadruplex formation in LANA was analyzed using G-quadruplex stabilizing compound, TMPyP4 in two KSHV positive PEL cells and in cells expressing LANA. We observed a reduction in the expression of LANA as compared to control cells and this reduction was found to be post-transcriptional as mRNA quantities remained unchanged between control and the treated cells. Moreover, we tested the effect of G-quadruplex stabilization using a LANA luciferase construct where the luciferase gene was placed downstream of LANA coding sequence. Not surprisingly, we observed lower luciferase units in TMPyP4 treated cells as compared to the control cells. To study the translation of G-quadruplex forming sequence of LANA exclusively, we generated a clone encompassing 250nt of LANA capable of forming G-quadruplex (G4-wild type) and a control clone (G4-disrupted), codon-optimized to code for the same amino acids but does not form G-quadruplexes. Consistent with our previous findings, we observed that G4-wild type displays lower or no protein expression as compared to G4-disrupted, aligning with our overall hypothesis that G-quadruplex formation negatively regulates translation of mRNA. Next, we went on to test the effect of G-quadruplex formation on antigen presentation and observed that stable G-quadruplexes forming cells elicit a weak CD8+ T cell immune response.

Next, we screened for proteins that bind to the G-quadruplex region of LANA and hnRNPA1, a RNA binding protein was found to be a potential binding partner. Upon analysis of the effect of hnRNPA1 on G-quadruplex formation, it was discovered that hnRNPA1 causes unwinding of RNA G-quadruplexes as it enhanced the translation of G-rich mRNA of LANA. Interestingly, we found that LANA binds to the G-rich region of its own mRNA and have inhibitory effect at higher levels of LANA. We also discovered that at higher concentrations, LANA negatively effected the nuclear to cytoplasmic export of G-rich mRNA, leading to a
reduction in translation. In summary, the studies performed in this chapter provide an insight into the role of G-quadruplexes in immune evasion strategies of the virus, in addition to the regulation of G-quadruplex formation by hnRNPA1 and LANA.

**RecQ1 helicase unwinds G-Quadruplexes at oriLyt to initiate KSHV lytic DNA replication**

In view of the biphasic lifecycle, KSHV exhibits latent and lytic replication phases. While the latent replication is aimed at maintaining the viral genome stably inside the cells and escaping host immune surveillance, lytic replication is meant for disseminating viral particles and infecting permissive cells along with promoting angiogenesis and cell proliferation. Various factors like hypoxia, oxidative stress, virus coinfection and chemical stimuli could trigger the switch of the virus from latent to lytic mode. Viral lytic replication initiates at specific region in the KSHV genome called oriLyt, which comprises of repeat regions and transcription factors sites. Like any replication process, viral lytic replication also entails the formation of a prereplication complex consisting of viral proteins that are recruited to the oriLyt. Other than viral proteins, many cellular proteins such as Topoisomerase I/II and RecQ1 helps KSHV lytic replication. The ability of RecQ1 to unfold secondary DNA structures as well as the high GC content of the origin, prompted us to investigate these regions in detail. Our results showed oriLyt to have a high propensity for the formation of secondary structures called G-quadruplexes. We used a web-based tool, QGRS mapper to analyze the formation of G-quadruplexes in the oriLyt region. Based on these results, we selected a region (wt oligo) with highest probability of G-quadruplexes formation and performed CD spectroscopy and Electrophoretic mobility assay. We used a scrambled oligo of the same length as control, which was unable to form a G-quadruplex structure. We were able to confirm the formation of these secondary structures in oriLyt, as the wild type oriLyt oligo displayed CD spectra, characteristic for G-quadruplexes.
Moreover, the mobility pattern of the oriLyt oligo also resembled G-quadruplexes as it migrated faster than the scrambled oligo of the same length. RecQ1 was shown to bind to the oriLyt region in previous study and here we sought to determine whether RecQ1 binding affects G4 unwinding and DNA replication. We analyzed the binding of RecQ1 to oriLyt through in vitro DNA affinity pull down assay, where we demonstrated specific binding of RecQ1 to the wild type oriLyt and not the scramble oligo. In addition, we performed Chromatin Immunoprecipitation assay, where we pulled down RecQ1 associated DNA fragments and observed that RecQ1 was bound specifically to the G-quadruplex forming region of oriLyt. Considering the ability of RecQ1 to bind to G-quadruplex regions, we tested the relevance of this binding by employing a RecQ1 helicase inhibitor, NMM, which showed a significant reduction in lytic replication, evidenced by a lower level of actively replicated DNA, genome copies and virions. Additionally, we depleted RecQ1 in KSHV positive cells through shRNA-based transduction, which showed a reduction in viral genome copies in RecQ1 depleted cells.

Furthermore, we showed that RTA, immediate early protein, was instrumental in recruiting RecQ1 to the G-quadruplex forming site, which pointed towards the cooperative action of viral and cellular proteins in initiating lytic replication. Since, G-quadruplexes are proven to regulate a number of biological process, we wanted to take a closer look into the functional significance of these structures. To achieve this, we first disrupted the G-quadruplex of oriLyt through site directed mutagenesis and assayed DNA replication. We observed that transfection of mutated oriLyt plasmid in KSHV positive cells, diminished replication of mutant oriLyt and also compromised the binding of RecQ1. To further investigate the role of G-quadruplexes in lytic replication, we performed Single Molecule Analysis of the Replicated DNA (SMARD), which displayed compromised initiation of lytic replication in cells treated with
ligands capable of stabilizing G-quadruplexes. Moreover, G-quadruplex stabilization through compounds in KSHV positive cells, also negatively affected late gene transcription, lytic DNA replication and progeny virion production. In conclusion, this study identifies a mechanism, where cellular and viral proteins work in conjunction to initiate lytic DNA replication at G-quadruplex forming region in the oriLyt.
References


The University of Texas Medical Branch at Galveston., Galveston (TX).


viral interferon-regulatory factor-3 is required for the survival of KSHV-infected primary

euchromatin-to-heterochromatin transition on the KSHV genome following de novo

94. Toth Z, Papp B, Brulois K, Choi YJ, Gao SJ, Jung JU. 2016. LANA-Mediated
Recruitment of Host Polycomb Repressive Complexes onto the KSHV Genome during

herpesvirus-encoded latency-associated nuclear antigen inhibits lytic replication by
78:6585-6594.

96. Davis DA, Rinderknecht AS, Zoeteweij JP, Aoki Y, Read-Connole EL, Tosato G,
Blauvelt A, Yarchoan R. 2001. Hypoxia induces lytic replication of Kaposi sarcoma-

2000. Kaposi's sarcoma-associated herpesvirus open reading frame 50/Rta protein
activates the entire viral lytic cycle in the HH-B2 primary effusion lymphoma cell line.

Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) DNA in infected cells

associated herpesvirus lytic origin (ori-Lyt)-dependent DNA replication: identification of
the ori-Lyt and association of K8 bZip protein with the origin. Journal of virology
77:5578-5588.

Human herpesvirus KSHV encodes a constitutively active G-protein-coupled receptor

101. Boshoff C, Endo Y, Collins PD, Takeuchi Y, Reeves JD, Schweickart VL, Siani MA,

ORF K9 (vIRF) is an oncogene which inhibits the interferon signaling pathway.

103. Bais C, Santomasso B, Coso O, Arvanitakis L, Raaka EG, Gutkind JS, Asch AS,
Cesaran E, Gershengorn MC, Mesri EA. 1998. G-protein-coupled receptor of
Kaposi's sarcoma-associated herpesvirus is a viral oncogene and angiogenesis activator.

104. Lee JS, Li Q, Lee JY, Lee SH, Jeong JH, Lee HR, Chang H, Zhou FC, Gao SJ,

cytokine and cytokine response pathway genes by KSHV. Science 274:1739-1744.


Mini-chromosome Maintenance Proteins Cooperate with LANA During the G1/S Phase of the Cell Cycle to Support Viral DNA Replication

Prerna Dabral, Timsy Uppal, Cyprian Rossetto and Subhash C Verma *

Journal of Virology, March 2019

*Corresponding author

Department of Microbiology and Immunology, University of Nevada, Reno School of Medicine
1664 N Virginia Street,
Reno, NV 89557

Phone: 775-682-6743, Fax: 775-327-2332

E-mail: scverma@medicine.nevada.edu

Running title: LANA recruits MCMs for latent replication
Abstract

Latency Associated Nuclear Antigen (LANA) is essential for maintaining the viral genome by regulating replication and segregation of the viral episomes. The virus maintains 50-100 episomal copies during latency and replicates in synchrony with the cellular DNA of the infected cells. Since virus lacks its own replication machinery, it utilizes the cellular proteins for replication and maintenance and LANA is shown to make many of these proteins available for replication by directly recruiting them to the viral origin of replication within the terminal repeat (TR) region. Our studies identified members of the Minichromosome maintenance (MCM) complex as potential LANA interacting proteins. Here, we show that LANA specifically interacts with the components of MCM complex, primarily during the G1/S phase of the cell cycle. MCM3 and 4 of the MCM complex specifically bound to the amino-terminal domain, while MCM6 bound to both, the amino and carboxyl terminal domains of LANA. MCM binding region in the N-terminal domain mapped to the chromatin binding domain (CBD). LANA with point mutations in the carboxyl terminal domain identified MCM6 binding domain and over expression of that domain (1100-1150aa) abolished TR replication. Introduction of peptide encompassing 1104-1123aa region of LANA reduced MCM6 association with LANA and TR replication. Moreover, a recombinant KSHV (BAC16 Δ1100-1150) with deleted 1100-1150aa domain of LANA showed reduced replication and persistence of viral genome copies as compared to the wild type BAC16. Additionally, the role of MCMs on viral replication was confirmed by depleting MCMs and assaying transient and long-term maintenance of the viral episomes. The recruitment of MCMs to the replication origins through LANA was demonstrated through chromatin immunoprecipitation and isolation of proteins on nascent replicated DNA.
(iPOND). These data clearly show the role of MCMs in latent DNA replication and potential for targeting C-terminal domain of LANA for blocking viral persistence.

**Significance**

LANA-mediated latent DNA replication is essential for an efficient maintenance of KSHV episomes in the host. During latency, virus relies on the host cellular machinery for replication, which occurs in synchrony with the cellular DNA. LANA interacts with the components of multiple cellular pathways including cellular replication machinery and recruits them to the viral origin for DNA replication. In this study, we characterize the interactions between LANA and Minichromosome maintenance (MCMs) proteins, a member of the cellular replication complex. We demonstrated a cell-cycle dependent interaction between LANA and MCMs and determined their importance on viral genome replication and maintenance through biochemical assays. In addition, we mapped a 50-amino acid region in LANA, which was capable of abrogating the association of MCM6 with LANA and blocking DNA replication. We also detected LANA along with MCM at the replication forks using a novel approach, isolation of proteins on nascent DNA (iPOND).
Introduction

Kaposi’s sarcoma associated herpesvirus (KSHV), or HHV-8, an oncogenic virus of the gamma-herpesvirus family, is associated with several malignancies including Kaposi’s sarcoma (KS), Multicentric Castleman’s Disease and Primary effusion lymphomas (PEL) in immunocompromised individuals (1, 2). Similar to other herpesviruses, the life cycle of KSHV consists of a dominant latent phase and a transient lytic phase. During latency, which is the default program of the virus, there is restricted expression of viral genes and the virus predominantly relies on the cellular proteins to sustain itself inside the host (3). In contrast, lytic or the productive phase, is characterized by the expression of viral genes in a fully orchestrated manner accompanied by the production of virions (4, 5). Among the viral genes expressed during latency is the latency associated nuclear antigen (LANA), which serves as the master regulator of latency (6). LANA, an 1162 amino acid protein, is a multifunctional nuclear protein, which plays an important role in inducing malignancies (7-9). LANA is crucial for the maintenance of KSHV episomes in the infected cells by supporting replication and segregation of the viral genome into the daughter cells (10). LANA achieves this by tethering to the host chromosome through its N-terminal chromatin binding domain (CBD) and binding to the viral genome through its C-terminal DNA binding domain (11-15).

During latency, virus replicates along with the host cellular DNA and LANA plays an important role in initiating replication at the TR origins (16, 17). The latent DNA replication occurs once per cell cycle in conjunction with cellular replication and maintains 50-100 copies of viral episomes in KSHV infected cells (18). Since only a limited number of viral genes are expressed during latency and these proteins lack any enzymatic activity including helicase/polymerase required for DNA replication, virus depends on the host’s cellular proteins
for these functions. LANA has been shown to interact with many cellular replication proteins and recruits them to the origin of replication in the terminal repeats (TR) in order to facilitate latent DNA replication (19, 20). LANA bound to the replication origin is thought to recruit replication proteins in an orchestrated manner starting with the recruitment of Origin Recognition complex (ORCs) followed by recruitment of Cdc6, Cdt1 and MCMs to form the pre-replicative (pre-RC) complex during the G1 phase of cell cycle (6, 21).

LANA has been shown to interact with Topoisomerase IIb and recruits it to the origin through its N-terminus and blocking the topoisomerase activity abrogated latent viral replication (22). LANA is shown to interact with members of the ORCs through its C-terminus during G1/S phase of the cell cycle and these interactions are essential for viral genome persistence and replication (23-25). LANA also recruits DNA polymerase clamp loader; proliferating cell nuclear antigen (PCNA) and replication factor C (RFC) to the latent origin in order to drive replication and was found to be essential for episome persistence (26).

Mini-chromosome maintenance complex or MCMs were shown to be a part of on the viral chromatin at the replication origin in the terminal repeat region but the mechanism how they are recruited to the viral genome was not clear (23-25). Our yeast-two hybrid assay identified MCMs as potential LANA interacting proteins (27). Mini-chromosome maintenance complex (MCM2-7), is a six membered replicative helicase that binds to the replication origin and melts dsDNA to initiate replication as well as acts as a helicase on elongating DNA (28-31). MCMs are localized in the nucleus and are present at a concentration of $\sim 10^5-10^6$ copies per cell (32-35). Initiation of replication requires the formation of a prereplication complex (pre-RC) at the origin of replication where components of the replication machinery are assembled sequentially. The process begins with the recruitment of origin recognition complex (ORC)
during G1 phase of the cycle, followed by the recruitment of MCM2-7 complex in an inactive state by Cdt1 (36). MCMs are phosphorylated by cellular kinases, DDK and CDK for activation and recruitment of other cellular proteins essential for replication including Cdc45 and GINS (37, 38).

Here, we show that LANA binds to MCM complex and recruits them to the origin of DNA replication, confirmed by chromatin immunoprecipitation assays as the depletion of LANA significantly reduced MCMs loading onto the origin. These proteins specifically associated during the replicative, G1/S phase of cell cycle determined by specific localization of LANA and MCMs in the nuclei of KSHV infected cells. MCMs interacted with both N and C termini of LANA and the amino acid residues between 1100-1150aa contributed to its binding to MCM6. Importantly, expression of this domain (1100-1150aa) of LANA competitively reduced MCMs binding to LANA suppressed DNA replication. Additionally, recombinant KSHV having deletion (BAC16 Δ1100-1150aa) in LANA’s C-terminal domain showed reduced replication and persistence of the viral genome as compared to the virus with wild type LANA (BAC16). We also show that depletion of MCMs blocked transient DNA replication as well as reduced copies of the latency persisting episomes. Most importantly, the association of MCMs with LANA was confirmed on replicating DNA by detecting proteins on newly synthesized DNA using a novel approach, isolation of proteins on nascent DNA (iPOND).

MATERIALS AND METHODS

Cell culture

The KSHV-negative, Burkitt lymphoma cell line, BJAB, KSHV-positive PEL cell lines, BCBL-1 and BC3 were cultured in RPMI 1640 medium supplemented with 10% bovine growth
serum, 2 mM L-glutamine, and penicillin-streptomycin (5 U/ml and 5 µg/ml, respectively). BrK.219 (generated by infecting BJAB with rKSHV.219) was obtained from Dr. Thomas Schulz (Hannover Medical School, Germany) and cultured in RPMI 1640 medium in the presence of 4.2 µg/ml puromycin (39). The human embryonic kidney cell lines, HEK293T and HEK293L were cultured in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% bovine growth serum, 2 mM L-glutamine, and penicillin-streptomycin (5 U/ml and 5 µg/ml, respectively). All cell lines were grown at 37 °C in a humidified environment with 5% CO2.

**Antibodies and peptides**

The following commercial antibodies were used for this study: rat anti-LANA (Advanced Biotechnologies, Inc.), mouse anti-GAPDH (US Biological), mouse anti-Flag M2 (Sigma-Aldrich, St. Louis, MO, USA), mouse anti-Myc 9E10 (Sigma-Aldrich, St. Louis, MO, USA), rabbit polyclonal anti-MCM2, MCM5, MCM10 and PCNA (Santa Cruz Biotechnology), mouse monoclonal anti-MCM3, MCM4 and MCM7 (Santa Cruz Biotechnology). Mouse monoclonal anti-LANA hybridoma was generated at GenScript using the peptide sequence CEPQQREPPQQQREPPQQ. This peptide was also used for eluting LANA in the iPOND assay. Specific peptide, KRPRSPSSVYCQNKDTSKKVQMARLAWE (sp pep) from the C-terminal MCM6 binding residues (1104-1123aa) and the scrambled peptide, KRPRSPSSQQEPQQQEPQQQEPQQQEPQQQ (sc pep) was synthesized at GenScript. Nuclear localization signal sequences (underlined) were added for efficient translocation of these peptides into the nucleus.

**Plasmids**

MCM4 and LANA 1100-1150aa was generated by PCR amplification and cloning into a Flag-tagged vector: pA3F and myc tagged vector, EGFP myc, respectively. Integrity of the
clones was confirmed by sequencing at the Nevada Genomics Center, University of Nevada, Reno, NV. MCM3 and MCM6 constructs were obtained from Dr. Alan Diehl (Medical University of South Carolina) and Dr. John Schimenti (Cornell University), respectively. Carboxy terminal point mutants of LANA were obtained from Dr. Paul Liebermann (Wistar Institute, Philadelphia). Flag-tagged LANA, pA3F-LANA and its deletion constructs, LANA-N terminal domain (1 to 340 aa) and LANA-C terminal domain (940 to 1162 aa), Myc-tagged LANA, pA3M-LANA, GFP-NLS-Myc, GFP–LANA-N-Myc (1 to 340 aa), the truncation mutants GFP–LANA-N250-Myc (1 to 250 aa), GFP–LANA-N150-Myc (1 to 150 aa), GFP–LANA-N32-Myc (1 to 32 aa) and GFP–LANA-N33–150-Myc (33 to 150 aa) and alanine substitution mutants of LANA 1-32 were described earlier (40). The shRNA vectors for MCM3 and MCM6 were purchased from Dharmacon (GE Life Science).

**MCMs knockdown using lentiviral vectors**

The pTRIPZ lentiviral vector (Dharmacon, GE Life Sciences) containing shRNA for MCM3 and MCM6 were co-transfected with lentivirus packaging vectors, pCMV-dR8.2 and pCMV-VSVG (Addgene, Inc.) into HEK293T cells using polyethylenimine (PEI) (Polysciences, Inc.) to produce the respective lentiviral particles. Supernatants from the transfected HEK293T cells were collected for 5 days, followed by concentration of the virus by ultracentrifugation (25,000 rpm, 1.5 hr, 4 °C). The concentrated lentiviral particles were used for transducing the target cells (BCBL-1, BC3 and HEK293L) in the presence of 5µg/ml polybrene followed by selection with 1 µg/ml puromycin. The cells were treated with 1ug/ml doxycycline for at least 72 hours for the induction of knock down. The RNA interference (RNAi) efficiency was assessed by Western blot analysis with specific MCM3 and MCM6 antibodies.
Co-immunoprecipitation assays

To perform co-immunoprecipitation, cells were washed with PBS (phosphate buffered saline,) and lysed in 750ul NP-40 cell lysis buffer (1% Nonidet P-40, 50 mM Tris-HCl, pH 7.5, 150 mM NaCl and 1 mM EDTA) supplemented with protease inhibitors (1 mM phenylmethylsulfonyl fluoride, 10 µg/ml pepstatin, 10 µg/ml leupeptin and 10 µg/ml aprotinin). Cellular lysates were sonicated with a probe sonicator and centrifuged at 12,000 rpm for 10 min at 4 °C to remove cellular debris. The lysate was further treated with 100U of DNase to eliminate any cross-linking DNA. The supernatants were pre-cleared with protein A + G sepharose beads (GE Healthcare) for 30 min at 4 °C. Approximately 5% of the lysate was saved as input sample and the remaining cellular lysate was rotated overnight with specific antibodies. The immunocomplexes were captured using 50 µl protein A + G conjugated sepharose beads, which were incubated with the lysates for 2 hr at 4 °C. The beads bound immunocomplexes were collected by centrifugation at 2000 rpm for 5 min at 4 °C followed by washing thrice with 1 ml of ice-cold NP-40 buffer supplemented with protease inhibitors. The IP and input samples were boiled in 50 µl of SDS PAGE sample loading buffer for 8 min, resolved on SDS-polyacrylamide gel and transferred onto 0.45-µm nitrocellulose membranes (Bio-Rad laboratories) at 100 V for 80 min. The blots were blocked with 5% non-fat milk in TBST buffer (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% Tween 20) and incubated overnight at 4 °C with specific primary antibodies. The proteins were detected following incubation with infrared-dye-tagged secondary antibodies using Odyssey infrared scanner (LI-COR Biosciences, Lincoln, NE).

Immunofluorescence assay (IFA)

Immunofluorescence assay on KSHV-positive cells arrested in the G1/S phase was performed by growing the cells on poly-L-lysine treated coverslips for 12 hours followed by
incubation with mimosine for 120 min. The cells were air-dried for 10 min, fixed with 4% paraformaldehyde for 15 min at room temperature followed by permeabilization with 0.2% Triton X-100 in PBS for 10 min at room temperature. Cells were blocked with PBS containing 0.4% fish skin gelatin and 0.05% Triton X-100 for 30 min at room temperature. The cells were then incubated with specific primary antibodies overnight at 4°C and washed with PBS thrice before incubating with Alexa Fluor conjugated secondary antibodies (Molecular Probes) for 45 min at room temperature. Nuclear staining was performed using TO-PRO-3 (Molecular Probes) and images were captured using a confocal laser-scanning microscope (Carl Zeiss, Inc.). An immunofluorescence assay on the KSHV-positive cells, arrested in the G2/M phase, was performed after treating them with colchicine. Cells were grown on coverslips coated with poly-L-lysine for 12 hours followed by incubation with colchicine for 120 min before fixing with formaldehyde. Staining and microscopy were performed as described above.

In vitro translation and GST pull down assay

*In vitro* translation of pA3F MCM4 was performed using the Promega TNT T7 quick-coupled transcription/translation system; where 2µg of the plasmid was translated in a 50 µl reaction containing 1mM \[^{35}S\] methionine. For performing GST pull down assay, GST control LANA-N and LANA-C GST fusion proteins were expressed in *E.coli* BL21 competent cells and following induction with 1mM IPTG (isopropyl-β-d-thiogalactopyranoside), the fusion proteins were extracted using Glutathione sepharose beads. The *in vitro* translated proteins were rotated overnight with control-GST, LANA-N GST and LANA-C GST overnight in NETN binding buffer (0.1% NP-40, 20 mM Tris, 1 mM EDTA, and 100 mM NaCl) along with the protease inhibitors. Following overnight incubation, the beads were washed with NETN buffer three times and co-precipitated proteins were resolved using SDS-PAGE and detected using
autoradiography.

**Transient replication assay**

HEK293L cells in 100-mm dishes were co-transfected with 30 µg of TR-containing plasmid, p8TR with 30 µg of pA3F LANA, or with an empty vector, pA3F, as filler DNA in 293Ls, 293L shMCM3/shMCM6 Control (C) and doxycycline treated 293L shMCM3/shMCM6 Knockdown (KD). At 96 h post-transfection, cells were washed with phosphate-buffered saline followed by extraction of DNA using a modified Hirt’s lysis method, described earlier (41). Extracted DNA was dissolved in 30 µl of sterile water. Ten percent of the extracted DNA was linearized with EcoRI and the remainder with DpnI and EcoRI to remove the non-replicated DNA. The digested DNA was separated on 0.8% agarose gel followed by Southern transfer onto a Hybond N+ membrane (GE Healthcare) and hybridized with ³²P-labeled TR probes. Probes specific to the KSHV Terminal Repeat region were synthesized using a random primer labeling kit followed by purification on G-50 columns (GE Healthcare). The auto-radiographic signals were detected using a PhosphorImager, according to the manufacturer’s instructions (Molecular Dynamics, Inc.).

**Peptide transfection, co-immunoprecipitation and transient replication assay**

LANA specific (sp pep) and LANA scrambled (sc pep) peptides were transfected in the cells using Pierce protein transfection reagent according to the manufacturer’s protocol. For co-immunoprecipitation assay, pA3F LANA and MCM6 were transfected into HEK293T cells using PEI, harvested after twenty-four hours post-transfection and lysed to set up immunoprecipitation in the presence of specific or scrambled peptides. Immunoprecipitation and detection were done as described above. Replication assay was performed by transfecting the specific and scrambled peptides into 293L cells using Pierce transfection reagent for 4 h before
transfecting the plasmids, pA3F LANA and p8TR for replication assay. The replication assay was done as described above following digestion with EcoRI and DpnI.

**Generation of recombinant KSHV, Δ1100-1150 Bac16**

G-blocks (IDT) were designed to have a linear DNA fragment containing a kanamycin resistance expression cassette, an Isecl restriction enzyme site, flanked by homologous sequences to the region of LANA targeted for removal. The G-Block sequence is the following (upper case, bold sequence is the targeting sequence for the kan cassette insertion, upper case underline is the sequence required for the second round of recombination and seamless removal of the kan cassette, lower case middle sequence is the Kan^R/I-SceI sequence): 

```
CCTGCTTGCCCCACCCTGGGACCAGACCAGTCGCCCATAACTTATAACCAAGGTC
CTGGGGACTCTCTGcattattcaacaagccacagtgttctcaaaatctctgtatctgcattgcacaagatataatatcatctag
aacaataaaactgtctgccattcataaacaagtaataacaaggggtgttatgagccatattcaacgggaaacgtcttgctcagggcgcgttataaa
ctcaacatggaggttgtttctgaaacatgggcttggccttgggtggaagcgttggatgttgtttgatgtgacgaggtttgttggctcagggcgcgttataaa
ttcaacatggaggttgtttctgaaacatgggcttggccttgggtggaagcgttggatgttgtttgatgtgacgaggtttgttggctcagggcgcgttataaa
ttcaacatggaggttgtttctgaaacatgggcttggccttgggtggaagcgttggatgttgtttgatgtgacgaggtttgttggctcagggcgcgttataaa
ttcaacatggaggttgtttctgaaacatgggcttggccttgggtggaagcgttggatgttgtttgatgtgacgaggtttgttggctcagggcgcgttataaa
ttcaacatggaggttgtttctgaaacatgggcttggccttgggtggaagcgttggatgttgtttgatgtgacgaggtttgttggctcagggcgcgttataaa
ttcaacatggaggttgtttctgaaacatgggcttggccttgggtggaagcgttggatgttgtttgatgtgacgaggtttgttggctcagggcgcgttataaa
ttcaacatggaggttgtttctgaaacatgggcttggccttgggtggaagcgttggatgttgtttgatgtgacgaggtttgttggctcagggcgcgttataaa
```
competent GS1783 *E. coli* strain harboring BAC16, induced at 42°C for 15 mins. The Kan<sup>R</sup>/I-Scel-containing mutants were selected on chloramphenicol/kanamycin agar plates and correct insertional mutants were confirmed by PCR, restriction enzyme digestion (NdeI-digestion) and Southern blot analysis using LANA-specific probe. In the second red-mediated recombination step, the integrated Kan<sup>R</sup>/I-Scel cassette was removed following treatment with 1% L-arabinose and consequent arabinose-mediated I-Scel enzyme induction. The resultant kanamycin-sensitive and chloramphenicol resistant colonies were analyzed by restriction digestion and deletion in the LANA was confirmed by sequencing. BAC DNA was purified using Nucleobond Xtra BAC kit (Clontech) according to the manufacturer's instructions. BAC16 wt DNA and BAC16Δ1100-1150 DNA was transfected into 293L cells with the Metafectene Pro reagent (Biontex Laboratories GmbH, San Diego, CA) as described earlier (42). The cells containing BACmid were selected with hygromycin to obtain a pure population of cells. Both the cell lines were monitored for the maintenance of KSHV genome by genomic DNA extraction qPCR.

**IdU labeling and immunoprecipitation of replicated DNA**

MCM3 and MCM4 depleted KSHV-positive; BCBL-1 and BC3 cells were pulsed with 30 μM of IdU (Sigma-Aldrich, St. Louis, MO, USA) for 30 min and washed twice with cold PBS. Episomal DNA was extracted by the modified Hirt’s method, dissolved in 500 μl TE (10 mM Tris-HCl, 1 mM EDTA) and sonicated to get an average length of 700 bp. The samples were heat denatured at 95 °C for 5 min, incubated on ice and 10% of the extracted DNA was saved to use as input control. 50ul of 10X IP buffer (100 mm NaPO4 pH 7.0, 1.4M NaCl and 0.5% Triton X-100) was added to the IP samples, which were then incubated with 1 μg of mouse anti-IdU antibody (BD Biosciences) at room temperature with constant rotation for 1 h. Antibody bound IdU labeled DNA was precipitated using magnetic Protein A/G (GE Healthcare, Inc.)
after incubation for 30 min. The beads were washed once with 1X IP buffer (10 mm NaPO4 pH 7.0, 140 mM NaCl and 0.05% Triton X-100), resuspended in 200 µl of lysis buffer (50 mM Tris-HCl (pH8.0), 10 mM EDTA, 0.5% SDS, 0.25 mg/ml Proteinase K) and incubated overnight at 37 °C for elution. This was followed by addition of 100 µl of lysis buffer and incubation at 50 °C for 1 h. The eluted DNA was phenolized and precipitated for the quantitation of IdU labeled DNA in a semi-quantitative real-time PCR by amplifying the TR region.

**Gardella gel electrophoresis**

Gardella gels were used for assessing the episome maintenance. KSHV-positive BCBL-1 and BC3 cells depleted for MCM3 and MCM6 were loaded into the agarose gel with a lysis plug containing DNase-free Proteinase K (Sigma-Aldrich, St. Louis, MO, USA) and SDS followed by electrophoresis in a Tris-borate-EDTA buffer. The plugs were loaded on 0.8% Gardella gel and resolved at 108V for 30 hours. The gel was transferred onto a Hybond N+ membrane (GE Healthcare) and hybridized with 32P-labeled TR probes to detect KSHV episome.

**Chromatin immunoprecipitation assay (ChIP)**

Chromatin immunoprecipitation was performed as described previously (43). Approximately 4 million BC3 shControl, shLANA and BCBL-1 shControl and shLANA cells were fixed with 1% formaldehyde for 10 min at room temperature followed by the addition of glycine at a final concentration of 125 mM for 5 min to stop cross-linking. The cells were rinsed three times with ice-cold PBS and lysed in chromatin shearing buffer (Diagenode) supplemented with protease inhibitors for 10 min on ice. Chromatin was sonicated using a Bioruptor (Diagenode) to an average length of 500 to 800 bp and the lysates were centrifuged for 10 min at 13,000 rpm to remove the cell debris. The resulting supernatant was diluted four-fold with ChIP dilution buffer containing 16.7 mM Tris-HCl, pH 8.1, 167 mM NaCl and 1.2 mM EDTA with
protease inhibitors. The diluted chromatin was rotated followed by incubation overnight with either control IgG or MCM3/MCM4 and MCM6 antibodies at 4 °C. Immune complexes were collected by incubating with Protein A + G sepharose beads for 1–2h at 4 °C. The beads were collected and subsequently washed twice with low-salt buffer (0.1% SDS, 1.0% Triton X-100, 2 mM EDTA, 20 mM Tris [pH 8.1], 150 mM NaCl), and once with high-salt buffer (0.1% SDS, 1.0% Triton X-100, 2 mM EDTA, 20 mM Tris [pH 8.1], 500 mM NaCl). The beads were then washed twice with Tris-EDTA buffer and chromatin was eluted using an elution buffer (1% SDS, 0.1 M NaHCO3) and reverse cross-linked by adding 0.3 M NaCl and RNaseA at 65 °C overnight. Eluted DNA was precipitated, treated with Proteinase K at 45 °C for 2 h and purified using Phenol: Chloroform and Isoamyl alcohol. Purified DNA of the ChIP fraction and the inputs were subjected to amplification of TR with the primers (forward, 5′-GGGGGACCCCGGGCAGCGAG-3′, and reverse, 5′-GGCTCCCCAAACAGGCTCA-3′) flanking TR nucleotides 677 to 766 on an ABI StepOne plus real-time PCR machine (Applied Biosystems).

**Two-step Isolation of proteins on nascent DNA (iPOND)**

KSHV-positive cells (1.0 × 10^8 cells per sample) were incubated for 30 min with 30 µM of the thymidine analogue, EdU (5-ethynyl-2′-deoxyuridine). The cells were washed with PBS, cross-linked with 1% formaldehyde for 15 min at room temperature (RT) and quenched with 0.125 M glycine for 5 min at RT and washed three times in PBS. Cell pellets were then resuspended in 0.25% Triton-X/PBS for permeabilization and incubated for 30 min at RT. The cells were dounced 10 times in the permeabilization buffer and centrifuged at 900xg for 5 minutes. Pellets were washed once with 0.5% BSA/PBS and once with PBS using same volume used for permeabilization prior to the Click reaction.
Click reactions were performed to conjugate biotin to the EdU-labeled DNA. Cells were then subjected to the Click reaction for 2 hours by incubating the cells in 10mM sodium ascorbate, 2mM CuSO4 and photocleavable biotin-azide (Life Technologies, Inc.) at a density of $1.0 \times 10^8$ cells/5ml of Click cocktail. DMSO was added instead of biotin-azide to the negative control samples. After centrifugation at 900g for 5 minutes, the cell pellets were washed once with 0.5% BSA/PBS and once with PBS. Cells were then resuspended in lysis buffer containing 1% SDS, 50 mM Tris (pH 8.0) ($1.5 \times 10^7$ per 100 µl lysis buffer), 1 mM phenylmethylsulfonyl fluoride, 10 µg/ml pepstatin, 10 µg/ml leupeptin and 10 µg/ml aprotinin and incubated on ice for 15 min. Samples were sonicated using a microtip sonicator for 8 min at 15 W with 20-sec ON and 40-sec OFF pulses. Cell debris were removed by centrifugation at 16,100 × g for 10 min at RT and diluted 1:1 (v/v) with cold PBS containing 1 mM phenylmethylsulfonyl fluoride, 10 µg/ml pepstatin, 10 µg/ml leupeptin and 10 µg/ml aprotinin. The cell lysates were incubated overnight with 10 µg LANA antibody and the immune complexes were captured with 40 µl magnetic Protein A/G beads (pre-blocked in ssDNA) for 2 hours. The beads were washed thrice in PBS+ ssDNA and the LANA bound proteins were eluted using 25µg LANA peptide in 1X TBS at room temperature. Elution was performed twice and the eluted proteins were then incubated with 100µl Pierce Streptavidin Agarose Beads (ThermoFisher) overnight for 12-16 hours in dark. The beads were washed twice in PBS, once in 150mM NaCl and for last time in 1X PBS. Captured DNA-proteins complex were eluted under reducing conditions by boiling them in 2X SDS sample buffer for 30 min at 95°C. Protein samples were resolved on 4–20% gradient gel (Bio-Rad Laboratories) and immunoblotted with specific antibodies using the Odyssey infrared imaging system (LI-COR Biosciences, Lincoln, NE).
RESULTS

LANA associates with MCMs in KSHV-infected cells

LANA, the most abundantly expressed viral protein during latency, interacts with viral DNA and many cellular proteins including the replication complexes and recruit them to the origin of DNA replication in the terminal repeat region (23, 24, 26). MCMs, the cellular helicases, were identified as LANA interacting protein in our protein-protein and yeast-two hybrid assays (27). Chromatin immunoprecipitation assay showed the presence of MCMs on the chromatin of terminal repeat region but whether they were recruited through viral factors for replication was not known. To determine whether LANA can directly bind to MCMs and recruit it to the episomes, we assayed LANA’s interactions with MCMs through co-immunoprecipitation (co-IP) on the endogenous proteins from KSHV-positive, BCBL-1 and BC3 and control BJAB cells. The lysates were treated with DNAse to eliminate any association of proteins through DNA. Immunoprecipitation with LANA, followed by detection with specific antibodies of the MCM complex (MCM2, MCM3, MCM4, MCM5, MCM6 and MCM7), showed LANA’s interaction with MCM3, MCM4 and MCM6 (Fig. 1A, lanes 5 and 6). The specificity of this interaction was confirmed by the lack of co-precipitating MCMs from KSHV negative, BJAB cells (Fig. 1A, lane 4).

To further confirm that the association between the LANA and MCMs was not dependent on any other viral proteins, we tested their interactions in overexpression system. We performed co-immunoprecipitation from HEK293T cells transiently expressing LANA along with MCMs associating endogenously in previous assay. MCM3Flag, MCM4Flag or MCM6myc was co-transfected with a appropriate epitope tagged LANA or empty vector followed by immunoprecipitation of LANA using tag specific antibodies and detection of MCMs showed
specific co-precipitations MCM 3, 4 and 6 (Fig. 1B, panels a-c). Lack of MCMs precipitation in vector transfected cells confirmed specificity of their interactions. To further validate their interactions, we performed reverse co-immunoprecipitation using antibodies to precipitate MCMs and detecting LANA with indicated antibodies. LANA specifically coprecipitated with those MCMs (MCM3, 4 and 6) but not with the empty vectors confirming their interaction (Fig. 1C, panels a-c).
LANA co-localizes with MCMs in KSHV-positive cells during the replicative phase.

Viral DNA replication during latency is thought to occur in conjunction with the cellular DNA replication. The recruitment of MCMs to the origins is well defined, which begins in the late G1 phase of cell cycle through a sequential assembly of the components of pre-RC complex (44-49). The MCMs are converted into active helicase only during the G1/S phase of the cell cycle through phosphorylation with cellular kinases. Since the latent DNA replication occurs in synchrony with the cellular replication, we wanted to determine whether the interaction between LANA and MCMs is limited to the actively replicating phase of cell cycle or throughout the cell cycle. To this end, we performed immunofluorescence assay in replicating (G1/S) as well as in mitotic (G2/M) phase cells. To achieve these specific phases, cells were treated with mimosine (G1/S) and colchicine (G2/M) and stained with rat anti-LANA and mouse anti-MCM3, MCM4 and MCM6 antibodies. LANA showed a distinct punctate pattern in the nuclei of PEL cells (Fig. 2A), as expected.
Many of the foci of MCMs (MCM3, 4 and 6) enrichment matched to the LANA dots during the G1/S phase of the cell cycle (Fig. 2A, panels a, b and c). An enlargement of the specific foci (LANA-red and MCMs-green) showed a co-localization of LANA with these MCMs, depicted by yellow spots, during the G1/S, replicating phase of the cells cycle (Fig. 2A, panels a, b and c). In contrast, cell in the mitotic phase (colchicine treated) did not show LANA’s localization with MCMs (Fig. 2A, panels a, b and c). Nuclear staining with TO-PRO3 (blue) confirmed that these proteins were in the nuclear compartment. The differential interference contrast (DIC) images showed outline and the integrity of the cells.

The amino and carboxyl termini of LANA interact with MCMs.

In order to determine the domains of LANA responsible for interaction with MCMs, we performed co-immunoprecipitation assays in HEK293T cells transfected with LANA mutants expressing either the amino-terminal (N-1-340aa) or the carboxyl-terminal (C-940 to 1162aa) domain along with epitope tagged full length MCMs. Immunoprecipitation with anti-myc
antibody and subsequent detection of co-precipitating MCM3 and MCM4 showed its binding to the amino-terminal but not to the carboxyl terminal domain of LANA (Fig. 3A, panels a and b). Similarly, immunoprecipitation of LANA-N and LANA-C with anti-flag antibody and subsequent detection of co-precipitating MCM6 showed its binding with both termini of LANA (Fig. 3A, panel c). Specificities of these interactions were confirmed by the lack of any co-precipitating protein with empty vectors. In order to confirm that these interactions are direct we used GST pull down assay where lysates from MCM3, MCM4 and MCM6 expressing HEK293T cells were incubated with bacterially purified GST fused with LANA-N or LANA-C. Similar to the above immunoprecipitation results, MCM3 and MCM4 interacted with only the amino terminal domain of LANA, whereas, MCM6 interacted with both amino and carboxyl termini of LANA (Fig. 3A, panel d). GST alone was used as a control, which did not bind to any MCMs confirming specificity of their interaction. In addition, we confirmed direct associations of MCMs with LANA through proteins prepared from a cell free in-vitro translation system. In-vitro translated MCM4 was incubated with bacterially purified LANA truncations, LANA-N and LANA-C fused to GST. We performed in-vitro translation and interaction assay with only MCM4 because the other MCMs were untranslatable. Importantly, MCM4 bound to the amino-terminal domain of LANA similar to the above binding assays (Fig. 3A, panel e). This assay confirmed that both the amino and carboxyl terminus of LANA are capable of binding to MCMs.

In order to further map the region of LANA responsible for binding to MCMs, we performed co-immunoprecipitation assays with truncations of the amino terminus of LANA, including 1-32aa, 1-150aa, 1-250aa, 1-340aa and 33-150aa. Immunoprecipitation with anti-myc antibody and subsequent detection of MCMs identified amino acid 1-32aa as the minimal domain required for
MCM3 and MCM6. While MCM4 mapped to 1-150aa and more specifically 33-150aa region of LANA (Fig. 3B). Since LANA 1-32aa is also the chromatin binding domain, we wanted to identify the amino acid residues of this region responsible for its interaction with MCMs.

Schematic of LANA 1-32aa region with alanine mutation between 5-15aa region is depicted in figure 3C (panel a). These expression vectors were transfected into HEK293T cells followed by immunoprecipitation with anti-myc antibody to precipitate LANA 1-32aa showed amino acids 14-15 (T and G) to be critical for its association with MCM3 (Fig. 3C, panel b). Since MCMs are recruited as a hexameric complex at the origin of DNA replication, the components of MCMs
(MCM2-7) almost similar levels of LANA from both BCBL-1 and BC3 cells (Fig. 3D and E, panel a). Similarly, MCM4 co-precipitated LANA from those MCM6 depleted and control BCBL-1 and BC3 cells, suggesting for a direct association of these proteins with LANA (Fig. 3D and E, panel b).

**A 50aa region in C-terminal LANA is important for MCMs binding latent DNA replication**

Since MCM6 interacted with the C-terminal domain of LANA, we wanted to identify specific amino acid residues of the C-terminal domain responsible for interaction with MCM6. We performed MCM6 co-immunoprecipitation assay with already existing C-terminal point mutants of LANA to identify whether any of those specific amino acid residues were important for binding. Immunoprecipitation of LANA-C mutants and subsequent detection of MCM6 with anti-myc antibody showed a reduced level of co-precipitating MCM6 with few mutants (Fig. 4A). Most of these critical residues lies between 1100-1150aa region of LANA (Fig. 4A). Since MCMs are important for replication, we asked whether LANA mutant with reduced MCMs binding would have any effect on DNA replication. We addressed this by performing a transient replication assay with some of the C-terminal mutants of LANA having reduced MCM6 binding. Replication efficiencies were assayed by determining the DpnI resistant (replicated) copies of the TR plasmid, which showed a significant reduction with many mutants and almost complete abrogation with mutant, K1109/1113/1114A of LANA (Fig. 4B, DpnI resistant band intensities are presented as a bar graph). Although point mutant, M1117A showed reduced MCM6 binding but only slightly reduced replication, which suggested for additional factors contributing to MCMs association with LANA (Fig. 4). Since the amino acids at K1109/1113/1114 were important for MCM6 binding and replication, we generated a small clone of LANA (1100-1150aa) encompassing these residues for determining whether this can compete MCM6 binding
with LANA. To this end, we performed a co-immunoprecipitation assay of MCM6 with LANA in the presence of LANA 1100-1150aa, which showed a significantly reduced level of co-precipitating MCM6 (Fig. 4C). Since the levels of MCM6 were similar in both with and without LANA1100-1150aa expressing cells, we concluded that amino acid between 1100-1150 can competed their binding (Fig. 4C). Next, we wanted to determine whether disruption of MCM6 binding with LANA with this competing domain can have effect on replication. To assay that, HEK293L cells were transfected with LANA and TR containing plasmids in the presence or absence of LANA1100-1150aa expressing vector. Cells were harvested 96 hours post transfection for the extraction of DNA and digested with EcoRI to linearize the DNA and with EcoRI and DpnI to determine the replicated DNA copies. The DpnI resistant (marked by an arrow) showed a complete abrogation in the replication of TR plasmids in cells expressing LANA1100-1150aa as compared to the cells with similar amounts of empty vector (Fig. 4D, compare lane 4 with 3). The levels of transfected TR plasmids were similar in both the samples (Fig. 4D, Input lanes, 1 and 2). Cells transfected with LANA1100-1150aa showed a cross-hybridizing signal to the vector backbone of LANA1100-1150aa, but did not affect the TR plasmid signal (Fig. 4D, lane 2).

We further wanted to determine whether the peptide encompassing these critical residues of LANA-C will have impact on disrupting LANA’s association with MCM6. To this end, we synthesized a 20 amino acid peptide (KRPRSPSSVYCQNKDTSKVKQMARLAWE) within 1104-1123aa region of LANA encompassing 1109/1113/1114 amino acid residues, which were important for MCM6 binding and DNA replication. Same size (20aa) scrambled peptide (KRPRSPSSQQEPQQEQEPQQQEQQEPPQ) was used as a control. We introduced nuclear
localization signal sequence (underlined sequence) on these peptides for efficient translocation into the nucleus.
Immunoprecipitation of LANA with anti-flag antibody showed reduced co-precipitation of MCM6 in the presence of specific LANA peptide as compared to almost similar amounts in the absence of any peptide or in the presence of scrambled (sc) peptide (Fig.5A). Relative amounts of co-precipitating MCM6 were determined based on the amounts precipitated in lane without any peptide after normalizing with their respective inputs (Fig.5A). We further wanted to determine whether the disruption of MCM6 association with LANA through this peptide would translate into reduced levels of TR replication. To this end, we performed a transient replication assay by transfecting LANA and TR plasmids in 293L cells in the absence and presence of the specific and scrambled peptides. Transfected cells were harvested 96 hours post transfection to extract DNA followed by digestion with EcoRI to linearize or with DpnI and EcoRI to determine the replicated copies after Southern hybridization. Quantification of DpnI resistant bands showed significantly reduced levels of TR replicated plasmid in cells transfected with LANA specific peptide as compared to cells transfected with no peptide or cells with scrambled peptide (Fig. [Figure 4: LANA amino acid region 1100-1150aa is essential for association with MCM6. A. HEK293T cells were transfected with flag tagged LANA-C terminus point mutants F1037A, F1044A, K1109/1113/1114A, K1138/1140/1141A, Y1021A, Y1014/1021A, Y1066A, M1117A, S1125A and myc-tagged MCM6. 36h post-transfection, cellular lysates were subjected to immunoprecipitation with anti-flag antibody to detect co-immunoprecipitated MCM6 following detection with anti-myc antibody. B. HEK293L cells were transfected with plasmids expressing full length LANA and LANA-C terminus point mutants in the presence of TR plasmid. Cells were harvested 96h post transfection; DNA was extracted using Hirt’s procedure and digested with either EcoRI or with EcoRI and DpnI. The DNA was subjected to Southern blotting and detection using TR probe. Densitometric quantification of DpnI-resistant test bands compared to respective input lanes is shown. C. HEK293T cells were transfected with MCM6-myc and full-length LANA in the presence and absence of LANA 1100-1150aa. Cellular lysates were subjected to immunoprecipitation with anti-flag antibody to detect co-immunoprecipitated MCM6 following detection with anti-myc antibody. D. HEK293L cells were transfected with plasmids expressing full length LANA, full length LANA with LANA 1100-1150aa in the presence of TR plasmid. Cells were harvested 96 hours post transfection; DNA was extracted using Hirt’s procedure and digested with either EcoRI or with EcoRI and DpnI. The DNA was subjected to Southern blotting and detection using TR probe. Arrow shows the p8TR band.]}
5B). The relative quantities of the replicated DNA were calculated in reference to the DpnI resistant band from cells with no peptides after normalizing with respective inputs (Fig. 5B).

A.

<table>
<thead>
<tr>
<th></th>
<th>Input</th>
<th>Flag IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>LANA+ MCM6</td>
<td>1.0</td>
<td>0.288</td>
</tr>
<tr>
<td>LANA+ MCM6 + sp pep</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>LANA+ MCM6 + sc pep</td>
<td></td>
<td>0.85</td>
</tr>
</tbody>
</table>

IB: Myc

B.

Figure 5: Introduction of LANA specific peptide, 1104-1123aa reduced LANA-MCM6 binding and TR mediated replication.

A. HEK293T cells were transfected with flag-tagged full-length LANA and myc-tagged full length MCM6. The cells were lysed 36h post-transfection and the lysates were incubated in the absence and presence of LANA specific (sp) and scrambled (sc) peptide. The immunoprecipitation was performed with anti-flag antibody, followed by detection of MCM6 with anti-myc antibody. Relative densities of the co-precipitating MCM6 are shown below the respective band.

B. HEK293L were transfected with plasmids expressing full length LANA and TR plasmids with and without LANA specific (sp) and scrambled (sc) peptides. Cells were harvested 96h post transfection; DNA was extracted using Hirt’s procedure and digested with either EcoRI (to linearize) or with EcoRI and DpnI. The DNA was subjected to Southern blotting and detection using TR probe. Arrow shows the TR plasmids band. Relative densities of the replicated band (indicated by arrow) normalized to DpnI resistant band from cells transfected with no peptide are shown below their respective lanes.
Recombinant KSHV with deleted 1100-1150aa LANA showed reduced replication

Our experiments so far confirmed the importance of MCMs-LANA association on KSHV latent DNA replication. In addition, amino acid residues of C-terminus LANA were important for MCM6 recruitment and DNA replication. To verify whether these C-terminal residues of LANA have effect on KSHV genome replication/maintenance, we generated a recombinant KSHV with 1100-1150aa LANA (BAC16Δ1100-1150) through homologous recombination-based two-step BAC recombineering (Fig. 6A). The BACmids, (BAC16 wt and BAC16Δ1100-1150) were transfected into 293L cells and selected with hygromycin. Transfection of both BAC16wt and BAC16Δ1100-1150 was similar, monitored by GFP signals, following 24h post transfection, but BAC16Δ1100-1150 displayed higher rate of GFP signal decrease as compared to the cells transfected with BAC16wt. We further determined the copies of viral genome retained following hygromycin selection for 6 days. Episomal DNA extracted using Hirt’s procedure from these cells showed a significantly lower KSHV copies in BAC16Δ1100-1150 cells as compared to BAC16wt cells (Fig. 6B and C). The reduction in episomal copies in BAC16Δ1100-1150 cells prompted us to evaluate whether the reduction was due to a defect in DNA replication observed in transient replication assays. To achieve this, we labeled the BAC16wt and BAC16Δ1100-1150 (Day 6 samples post hygromycin selection) with IdU, to label the replicating DNA, which was extracted through Hirt’s procedure for immunoprecipitation with anti-IdU antibody. The replicated DNA captured on antibody-protein A/G beads were eluted for quantification using qPCR with primers to amplify the terminal repeat region. As expected, we observed significantly lower amounts of newly replicated DNA in BAC16Δ1100-1150 cells as compared with BAC16wt cells (Fig. 6D).
A.}

\[
\text{K14} \quad \text{ORF73} \quad \text{ORF72}
\]

Detection of re-coding for LAMK aa 1100-1150 and insertion of kanamycin cassette by homologous recombination

\[
\text{K14} \quad \text{ORF73} \quad \text{Δ1100-1150 (+kan)} \quad \text{kanamycin} \quad \text{ORF72}
\]

Induction of lacZ enzyme and removal of kanamycin cassette by homologous recombination

\[
\text{K14} \quad \text{ORF73} \quad \Delta_{1100-1150} (-\text{kan}) \quad \text{ORF72}
\]

b.

\[
\text{Δ1100-1150}
\]

\[
\text{NdeI digest}
\]

**NdeI digest**

**ORF73 southern blot**

WT BAC16: 4902 bp

Δ1100-1150 (+kan): 5811 bp

Δ1100-1150 (-kan): 4752 bp

B.}

C.}

D.}

**Bac16wt day 3**

**Δ1100-1150 day 3**

**Bac16wt day 6**

**Δ1100-1150 day 6**

**Bac16wt**

**Δ1100-1150**

**Relative genome copies**

**Relative amounts of replicated DNA**
MCMs are required for LANA-dependent latent DNA replication.

Lower MCM6 binding to LANA-C terminal mutants and subsequent reduction in replication could also be due to a reduced binding of other cellular factors within that region. To confirm the absolute role of MCMs in KSHV DNA replication, we depleted the levels of MCM components (MCM3 or 6) in HEK293 cells and assayed DNA replication using TR containing plasmid. We transduced HEK293 cells with doxycycline inducible shRNA for depleting MCM3 or MCM6. Stably transduced cells were assayed for the depletion of MCMs after doxycycline treatment, which showed a significant reduction (shMCM3 or 6-KD) in MCM levels as compared to the untreated cells (shMCM3 or 6-C) (Fig. 7A and B, panel a). A transient replication assay was performed by transfecting these cells with LANA and TR containing plasmids. Transfected cells were harvested 96 hours post transfection to extract DNA followed by digestion with EcoRI to linearize or with DpnI and EcoRI to determine the replicated copies.
after Southern hybridization. Since LANA is crucial for TR DNA replication, cells without LANA expression did not show any replicated/DpnI resistant band (Fig. 7A and B, b panels, lanes 7, 9 and 11). HEK293L cells expressing LANA showed replicated TR copies (Fig. 7A and B, b panels, lane 8, asterisk), as expected. Cells transduced with shMCM3 or MCM6 lentivirus and treated with doxycycline (shMCM3 or 6-KD) showed no replicated TR plasmid as compared to the doxycycline untreated (shMCM3 or 6-C) cells (Fig. 7A and B, b panels, compare lane 12 with 10).
Lane 1-6 represents the amounts of TR plasmid DNA extracted from respective cells (Fig. 7A and B panel b). This confirmed that MCMs (both MCM3 and 6) are critical for LANA mediated replication of TR plasmids.

**LANA recruits MCMs at the terminal repeat region.**

We determined that MCMs were critical for latent DNA replication, and these were previously shown to be a part of the viral chromatin (23, 24), therefore, we asked whether MCMs are recruited through a viral factor, LANA or loaded onto the viral chromatin independent of LANA. To address this, we depleted LANA from the KSHV infected, BCBL-1 and BC3 cells by stably transducing shLANA lentiviral vector (50). The efficiency of LANA depletion was determined by comparing the levels of LANA with the control shRNA transduced cells, which showed a significant reduction in shLANA transduced cells (Fig. 8A and B). We determined the levels of MCMs (MCM3, 4 and 6) bound to the chromatin of KSHV genome (TR region chromatin) through chromatin immunoprecipitation and comparing with the control cells. We found that all the tested MCMs (MCM3, 4 and 6) had significantly reduced binding to the chromatin of TR region in both, BCBL-1 (Fig. 8A) and BC3 (Fig. 8B) LANA depleted...
(shLANA) cells as compared to the control shControl) cells. This confirmed that MCMs are recruited to the viral genome through LANA.

**Figure 8: LANA recruits MCMs at the TR region during latent DNA replication.**

**A. BCBL-1**

Immunoblot showing efficiency of LANA depletion in the lysates from LANA knockdown and control BCBL-1 cells.

**B. BC3**

Immunoblot showing efficiency of LANA depletion in the lysates from LANA knockdown and control BC3 cells. Chromatin immunoprecipitation assay was performed from the BCBL-1 and BC3, control and LANA depleted cells using antibodies specific for MCM3, 4 and 6. The relative copies of chromatin bound DNA was calculated by amplifying viral genome with TR specific primers and using ΔΔCt method after normalizing with respective inputs. All the experiments were performed three times in replicates and the error bars represent standard deviations of the means from three independent experiments. Statistical analysis was performed with Prism 7.0 software (Graphpad Inc.) using an unpaired t-test for the significance (p-value). Asterisks represent the P value < 0.05 (*), P value < 0.01 (**) and <0.0001 (****).
MCM depletion led to a reduction in KSHV genome replication and persistence.

Since MCMs were found to be essential for the replication of TR containing plasmids in transient replication assay, we wanted to determine whether depletion of MCMs in KSHV infected cells affects viral genome replication and persistence. To this end, we transduced BCBL-1 and BC3 cells with doxycycline inducible shMCM3 or 6 lentiviral vectors. These cells were treated with puromycin to select a pure population of transduced cells, which was further treated with doxycycline to determine the levels of MCMs depletion. Detection of MCM3 and MCM6 showed an efficient depletion in doxycycline treated (MCM3/6KD) cells as compared to the untreated (Control) cells (Fig. 9A and B). These cells were used for assaying DNA replication by pulsing them with IdU, a thymidine analog. IdU gets incorporated into the replicating DNA therefore a pulse labeling with IdU followed by immunoprecipitation with anti-IdU antibodies determined the amounts of actively replicated DNA. Detection of IdU labeled DNA by PCR amplification of specific region of the KSHV genome, TR region in this case (TR initiates replication), showed a lower level of IdU labeled DNA in MCMs (MCM3 or 6) depleted BCBL-1 (Fig. 9C) and BC3 (Fig. 9D) cells as compared to the control cells. This confirmed that MCMs play a critical role in KSHV genome replication during latency. In order to determine the effects of MCMs on KSHV genome persistence, these lentivirally transduced cells were treated with doxycycline for 2 weeks followed by assaying the levels of KSHV genome through Gardella gel analysis, which determines the native form of episomally persisting viral genome. Not surprisingly, we observed significantly reduced KSHV episomes (reduced signal intensities) in both MCM3 (Fig. 9E) and MCM6 (Fig. 9F) depleted (Kd) BCBL-1 and BC3 cells as compared to the same number of control (C) cells. This confirmed that depleted levels of MCMs
reduced KSHV genome replication, which ultimately resulted into lower copies of latently persisting KSHV genome.
LANA and MCMs are on the nascent DNA at the replication fork.

DNA replication is a fundamental process for biological inheritance, which requires the functional and physical interaction of many proteins on replicating DNA (36). During latency, KSHV genome maintains itself in the host by replicating and efficiently segregating its genome to the daughter cells (21, 51, 52). Since KSHV lacks its own set of replication proteins, it relies on the cellular proteins for these essential functions. We found LANA to be present at the replicating DNA using a recently developed technique, isolation of proteins on nascent DNA (iPOND), which enables the identification of proteins assembled at the replication fork (53). We used a modified version of this approach to determine proteins on replicating viral DNA, the two-step iPOND (54). This technique employs the use of 5-ethynyl-2’-deoxyuridine (EdU), a thymidine analog to label the replicating DNA (Fig.10A). Following a brief pulse with EdU label, cells were crosslinked to preserve the protein-DNA composition. EdU (with an alkyne functional group) labeled DNA was Clicked with biotin azide after the isolation of nuclei.
Figure 10: LANA and MCMs are part of the replication complex.

A. Schematic of two-step iPOND performed on KSHV-positive cells. 1) Approximately 100 million KSHV-positive cells were labeled with 5-ethynyl-2’deoxyuridine (EdU) for 30 minutes, harvested and washed with 1X PBS. 2) Protein and DNA were crosslinked using 1% formaldehyde for 20 min and quenched using 125mM glycine. 3) The cells were permeabilized with 0.25% Triton X-100 in PBS for 30 minutes at room temperature and 4) nuclei were isolated following centrifugation. 5) Click chemistry was performed on the nuclei, to conjugate Biotin-Azide to EdU and DMSO was used as a negative control. 6) The chromatin was sheared using sonication to generate fragments of 100-300 bp. 7) Protein-DNA complex bound to LANA was captured using monoclonal LANA antibody. 8) Protein A/G beads were used to capture antigen-antibody complexes, which were eluted from the beads using peptide specific for LANA. 9) Streptavidin beads were used to capture EdU-labeled (replicated) DNA-protein complexes. 10) Proteins bound to streptavidin beads were eluted by boiling the beads at 95°C and detected using specific antibodies.

B. Immunoblot of LANA, MCM6 and PCNA obtained through two-step iPOND using respective antibodies from KSHV positive, BCBL-1 and BrK.219 cells. 100 million cells were labeled with EdU, permeabilized and Click reaction was performed using biotin azide and DMSO. The cells were lysed, sonicated and proteins were pulled down using LANA antibody. The proteins associated with DNA were pulled down using streptavidin beads and eluted by boiling in the loading buffer.
DNA was fragmented and the protein-DNA complex was subjected to an enrichment of viral DNA by an immunoprecipitation with anti-LANA antibody since LANA binds to the KSHV genome (13). LANA-DNA complex was eluted from the Sepharose beads by competing it off with LANA peptide (same peptide used for generating the antibody). Eluted complex was then subjected to streptavidin pull down to precipitate the EdU labeled (replicated) and biotin Clicked DNA. Proteins associated with replicated DNA were determined by immunoblot. Immunoprecipitation of LANA enriched viral genome and biotin pull down isolated the replicated DNA, thus allowed the identification of proteins present on the viral replication fork. We labeled the KSHV positive cells, BCBL-1 and BrK.219 cells with EdU for 30 minutes followed by cross-linking and permeabilization for Clicking with biotin-azide or DMSO, control. An immunoprecipitation with anti-LANA antibody followed by elution showed an efficient precipitation of LANA, as expected, and a LANA binding protein, PCNA (Fig. 10B, eluted lanes). Further precipitation of replicated DNA (biotin labeled) with streptavidin from the eluate showed LANA and PCNA, which are part of the replication complex (Fig. 10B, Streptavidin pull down lanes). Detection of MCM6, a representative component of the MCM complex confirmed MCMs to be a part of the replicating DNA. Samples Clicked with DMSO did not show any detectable level of these proteins, confirmed the specificity of streptavidin pull down of replicated/biotinylated DNA (Fig. 10B, Streptavidin pull down lanes).

DISCUSSION

KSHV, the etiological agent of Kaposi Sarcoma, Primary Effusion Lymphoma and Multicentric Castleman’s Disease, establishes life-long latency in the host and expresses a limited subset of genes during latency (1, 2, 5). LANA is the most predominantly expressed protein during latency, which is essential for maintenance and persistence of the virus inside the
host cells (55-57). LANA plays an important role in the persistence by tethering viral genome to the host chromosome through simultaneous interaction with the terminal repeat region of viral genome and the host chromatin (11, 13). In addition, LANA is also essential for the duplication of viral genome as it is involved in the recruitment of cellular replication factors at the latent origin of viral genome replication (23, 24).

MCMs were identified as a potential LANA binding partner in our previous study (27). Additionally, MCMs were shown to be a part of the viral chromatin during latent replication and siRNA depletion of MCM5 reduced the replication of 2xTR plasmid (23). In addition, MCM3 along with ORC2 was found to accumulate at multiple sites on the viral genome as a part of the pre-RC complex (25). However, the mechanism how MCMs were getting recruited to the viral genome was not known. In this study, we determined that LANA specifically binds to MCM and recruit them to the latent origin of the terminal repeat region. MCM complex, a hexameric component comprising of MCM2-7 binds to the replication origin in a stepwise manner and unwinds the DNA to initiate replication (58-60). We identified three components of MCM complex, MCM3, MCM4 and MCM6, directly binding to LANA in various biochemical assays.

Further, we wanted to know whether MCMs associated with LANA throughout the cell cycle as LANA remains bound to the chromatin during both G1/S and mitotic phase of the cells cycle but MCMs are only loaded during the replicative phase of the cell cycle (12, 58). Not surprisingly, MCMs associated with LANA only during the G1/S (replicative) phase of the cell cycle, which indicated that either LANA stabilizes the loading of MCMs to the origin, or differential post-translational modifications of LANA help in the recruitment of MCMs only during the G1/S phase. Our attempts to determine differential post-translational modification on
LANA during the G1/S and mitotic phase has not been successful yet but mutation at Threonine at 14\textsuperscript{th} amino acid to alanine in LANA disrupted the binding of MCM3, which bound to the minimal 1-32aa region of LANA. This may suggest that specific amino acid residues and post-translational modifications including phosphorylation may contribute to its binding to replication proteins.

Interestingly, one component of the MCM complex, MCM6 bound to both the amino- and carboxyl-terminal domain of LANA, which could be due to the fact that both termini of LANA (amino and carboxyl) can associate together (61). This association possibly helps in bringing the viral genome, bound to the carboxyl terminus of LANA to the host chromosome at the site of amino terminal bound LANA for efficient replication. Previous studies have shown that the association of viral genome to the host chromosome is required as LANA mutants defective in chromosome binding (5-15aa alanine substitution mutants) do not support efficient DNA replication (62). Importantly, the binding of MCM6 in the carboxyl-terminus of LANA was also important as expression of a small region of LANA spanning the MCM6 binding residues (1100-1150aa) disrupted its association with LANA. Additionally, over-expression of this MCM6 interacting domain in a replication assays disrupted replication of TR containing plasmid.

To further investigate the relevance of this region in LANA-MCM6 binding and TR mediated replication, we used a custom peptide of LANA-C region, 1104-1123aa, in our assays. As expected, overexpression of LANA specific peptide reduced MCM6 binding to LANA and significantly reduced the TR mediated replication as compared to the control peptide. Furthermore, our recombinant KSHV-infected cell line with a deletion of 50aa region in the C-terminal domain of LANA (\(\Delta\)1100-1150aa) showed a rapid loss of viral episomes as compared to
the cells with wild type LANA. The rapid loss of viral genome from cells containing recombinant KSHV with deleted Δ1100-1150 amino acids in the C-terminal domain was attributed to a significantly reduced replication of the viral genome. These studies suggested that targeting C-terminal domain of LANA could provide a potential target for blocking KSHV latent DNA replication. A recent study also identified 1138-1140aa of LANA as the key residues essential for its interaction with BRD proteins, which play an important role in viral genome persistence and viral DNA replication (63). Since the main residues for MCM6 binding (1104-1123aa) are adjacent to the BRD interacting residues (1138-1140), it can be postulated that targeting 1100-1150aa could disrupt the binding of multiple proteins and thus significantly reduce the persistence of viral genome from the latently infected cells.

The functional relevance of MCMs, which acts as replicative helicases, on KSHV replication was also tested in transient replication assays as well as in genome persistence after depleting them from the KSHV infected cells. Since MCMs are important helicases required for unwinding the DNA during replication, their functions were not compensated by other cellular proteins, as MCMs depleted cells showed reduced replication of TR plasmid. MCMs depletion from the KSHV infected, BC3 and BCBL-1 cells also reduced the persisting copies of the viral genome possibly due to a reduced replication over successive rounds cell divisions, which were detected by the incorporation of thymidine analog in replicating cells.

Although the association of MCM5 was shown at viral chromatin but the mechanism of its recruitment was not known. Our chromatin immunoprecipitation assay from KSHV infected cells with depleted levels of LANA, which directly binds to MCMs, showed significantly reduced MCMs at the chromatin of TR region. This confirmed that MCMs are recruited at the TR through their binding with LANA. The involvement of MCMs in replicating the viral
genome was confirmed by detecting MCMs on the replicated viral DNA using a new approach, isolation of proteins on nascent DNA (iPOND). Since iPOND detects the proteins on any replicating DNA, we modified the approach to determine proteins on virally replicated DNA. We introduced an additional step to immunoprecipitate LANA, which specifically binds to KSHV genome, to preferentially enrich the LANA bound viral DNA. The replicated DNA, which was clicked with biotin, was isolated by selectively precipitating with streptavidin and the proteins bound to these replicated DNA was assayed. Importantly, we detected MCMs at the replicating DNA along with already known proteins, which are the components of cellular replication machinery. This confirmed that MCMs are involved in replicating latent viral genome and LANA directly binds to MCM3, MCM4 and MCM6 to recruit them to the replication origin on viral genome.

Acknowledgements

We thank Dr. Alan Diehl (Medical University of South Carolina) and Dr. John Schimenti (Cornell university), for MCM constructs. We also thank Dr. Paul Liebermann (Wistar Institute) for carboxyl terminal point mutants of LANA. This work was supported from the National Institute of Health (CA174459 and AI105000).
References


63. **Domsic JF, Chen HS, Lu F, Marmorstein R, Lieberman PM.** 2013. Molecular basis for oligomeric-DNA binding and episome maintenance by KSHV LANA. PLoS Pathog **9:**e1003672.
LANA and hnRNP A1 regulate the translation of LANA mRNA through G-quadruplex formation

Prerna Dabral, Jay Babu and Subhash C. Verma *

(under revision)

*Corresponding author

Department of Microbiology and Immunology, University of Nevada, Reno School of Medicine
1664 N. Virginia Street,
Reno, NV 89557
Phone: 775-682-6743, Fax : 775-327-2332
E-mail: scverma@medicine.nevada.edu
Abstract

During the latent phase, Kaposi’s Sarcoma-associated HerpesVirus (KSHV) maintains itself inside the host by escaping the host immune surveillance mechanism through restricted protein expression. Latency-Associated Nuclear Antigen (LANA), the most abundantly expressed protein, is essential for viral persistence as it plays important roles in latent viral DNA replication and efficient segregation of the viral genome to daughter cells following the cell division. LANA evades immune detection by maintaining its protein level below a threshold required for detection by the host immune system but sufficient to persist in the host cell, by auto-regulating its expression through activation of its promoters and inhibiting antigen presentation by interacting with proteins of MHC class I and MHC class II pathways. Here, we propose a link between regulation of protein expression and restricted immune recognition through formation of G-quadruplex in G-rich region of mRNA of LANA. We show that the formation of these stable structures in LANA mRNA inhibits its translation and deters antigen presentation, which was supported by treatment of cells with TMPyP4, a G-quadruplex stabilizing ligand. We also identify heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) as a protein that interacts with the G-rich sequence of LANA mRNA and unwinds these stable secondary structures leading to increase in its translation.
Significance

LANA, the most abundantly expressed protein during latency, is a multi-functional protein, which is absolutely required for the persistence of the virus in the host cell. Even though the functions of LANA in aiding pathogenesis of the virus have been extensively studied but the mechanism of how LANA escapes host’s immune surveillance is not fully understood. In this study, we shed light on the auto-regulatory role of LANA to modulate its expression and immune evasion through formation of G-quadruplexes in its mRNA. We used G-quadruplex stabilizing ligand to define the inhibition in LANA expression and presentation on the cell surface through MHC class I. We defined the auto regulatory role of LANA and identified a cellular RNA helicase, hnRNP A1 in regulating the translation of LANA mRNA. This interaction of hnRNP A1 with LANA mRNA could be further exploited for controlling KSHV latency.
Introduction

KSHV, the etiological agent of Kaposi’s Sarcoma (KS), Primary Effusion Lymphoma (PEL) and Multicentric Castleman’s Disease (MCD), displays two distinct phases in its life cycle: a predominant latent phase and a short productive lytic phase (1-4). Like other herpesviruses, KSHV establishes a persistent latent infection through the expression of a limited number of proteins, among that, some of them play immunomodulatory roles to prevent recognition by the host immune surveillance system (5).

Latency-associated nuclear antigen (LANA), encoded by OR73, is the most predominantly expressed protein during latency and is essential for viral replication and maintenance of the viral genome through segregation of viral episomes to the daughter cells following mitosis (6-8). LANA achieves this by establishing simultaneous contacts with the host chromosome and the viral genome with the involvement of many cellular proteins (9-11). One of the modes through which LANA aids in the pathogenesis is by helping the virus to evade host immune surveillance through inhibition of multiple pathways (12). LANA has been shown to avoid recognition by CD8 (+) T cells through inhibition of major histocompatibility complex class I (MHC-I) presentation by interfering with translocation of peptides generated by proteasome to the endoplasmic reticulum (13). Previous studies have demonstrated the role of LANA in inhibition of CD4 (+) T cell responses by downregulating the major histocompatibility complex class II (MHC-II) genes through blocking CIITA transcription and interaction with Regulatory factor X (RFX) complex (14, 15).

G-quadruplexes are secondary DNA/RNA structures formed in nucleic acid sequences rich in guanine residues. G-quartets, which are formed by Hoogsteen hydrogen bonding between four guanine residues, stack on top of each other to form G-quadruplexes. These structures have
been reported to play important roles in biological process like DNA replication, transcription, alternative splicing and translation (16-20). The formation of G-quadruplexes structures is regulated by various proteins, which either stabilize or unwind these structures (21). G-quadruplexes have recently gained importance as potential drug targets against pathogenic viruses due to their occurrence in the genomes of many viruses and their regulatory role in viral life cycle. The presence of G-quadruplex structures have been reported in the central polypurine tract in the integrase gene of Human Immunodeficiency virus (HIV), which promotes recombination by facilitating the dimerization of RNA templates and template switching (22). In addition, the formation of these structures has been reported in the U3 region of the long terminal repeats (LTR) in the Sp1 and NF-kB binding sites, which negatively regulates the transcriptional activation of HIV (23). Also, the Nef region of HIV forms these secondary structures, which regulates the HIV infection by inhibiting the expression of Nef gene (24). A recent study reported the presence of G-quadruplexes in the preS2/S gene promoter of Hepatitis B virus, which positively regulated its transcription and the disruption of these structures led to a reduction in virion production (25). Hepatitis C virus has also been shown to form G-quadruplexes and its stabilization led to an inhibition in viral replication and protein expression (26). Among herpesviruses, G-quadruplex are present in the inverted repeat region of Herpes Simplex Virus (HSV-1) and the stabilization of these structures led to a reduction in HSV-1 infection (27). These G-quadruplex are enriched in the replication compartments following HSV-1 infection, which possibly regulates the viral life cycle by controlling viral replication (28). Epstein-Barr nuclear antigen (EBNA1) protein of Epstein-Barr virus (EBV) has been reported to preferentially bind to the G-quadruplex forming RNA and promotes EBNA-1 dependent recruitment of origin recognition complex (ORC) to the origin in addition to attachment of
EBNA1 to the metaphase chromosomes (29). A recent study also reported the formation of G-quadruplexes in EBNA mRNA, which serves as steric blocks for the movement of ribosomes to inhibit the translation of EBNA1 mRNA (30). Our lab has reported the formation of G-quadruplexes in the terminal repeat region of KSHV and their stabilization led to a reduction in latent viral DNA replication (31). These studies substantiate the vital role of G-quadruplex regulatory structures in the pathogenesis of viruses.

G-quadruplexes are thermodynamically stable structures whose formation is tightly regulated by the proteins that promote stabilization or unwinding of these structures in order to stop their hindrance in RNA or DNA metabolism (32, 33). A number of such proteins are already identified but the mechanism of how and when these proteins are recruited at definite sites is still unclear. Few helicases, including, DHX9 (DEAH Box helicase 9), RHAU, a RNA helicase of DEAH box family, DDX1 (DEAD-box RNA helicase 1) and DDX21, have been reported to unwind RNA G-quadruplexes (34-37). Another family of proteins, which have implicated in regulating G-quadruplex formation, are the heterogenous nuclear ribonucleoproteins (hnRNPs). These are the RNA binding family of proteins that regulate gene expression by performing a variety of functions including, mRNA splicing, mRNA stability and trafficking and modulating protein translation (38). A member of the hnRNP family, hnRNP A1 has been reported to enhance translation of RON receptor tyrosine kinase by binding to the G-quadruplexes formed in the 5’ UTR region (39). Another study has reported the interaction of hnRNP H/F with the G-quadruplex at the 3’ end of p53 mRNA that enhanced its expression during DNA damage (40). Similarly, hnRNP A2 positively regulated the translation of FMR1 (fragile X mental retardation) mRNA by destabilizing the G-quadruplex structure formed by the (CGG)_n repeat, which is
responsible for the silenced gene product of FMR1 gene that results in fragile X mental syndrome (41).

In this study, we report the formation of G-quadruplexes in the mRNA of LANA, through biophysical and biochemical methods. The stabilization of these structures with TMPyP4, a G-quadruplex stabilizing ligand inhibited the translation of LANA that resulted in a reduced protein expression in KSHV infected as well as LANA expressing cells. We studied the regulatory role of G-quadruplexes through a G-quadruplex containing LANA sequence (G4 wild type) along with a clone containing codon modified sequence of G4 wild type that has disrupted G-quadruplexes forming ability. Furthermore, we also demonstrated the regulatory role of G-quadruplex in inhibiting endogenous antigen presentation to CD8 (+) T cells owing to a reduced translation of LANA mRNA. Finally, we identified cellular and viral proteins involved in regulation of G-quadruplex formation. We provide evidence towards the role of hnRNP A1 in enhancing the translation of LANA and also demonstrate the ability of LANA to regulate its translation.

MATERIALS AND METHODS

Cell lines, plasmids and reagents

The KSHV-positive BCBL-1 cell line was grown in RPMI 1640 supplemented with 10% FBS, 2mM glutamine, 5U/ml penicillin and 5µg/ml streptomycin. BJAB-L-YFP cell line was grown in RPMI 1640 supplemented with 10% FBS, 2mM glutamine, 5U/ml penicillin and 5µg/ml streptomycin and 0.5µg/ml puromycin. HEK293T and HEK293L cells were grown in DMEM supplemented with 8% bovine growth serum, 2mM glutamine, 5U/ml penicillin and 5µg/ml streptomycin(55). B3Z, T cell hybridoma cell line, a kind gift from Dr. Nilabh Shastri (John Hopkins University) and Dr. Charles L. Stentman (Dartmouth University), were grown in
RPMI 1640 supplemented with 10% FBS, 2mM glutamine, 1mM pyruvate, 50\(\mu\)M 2-mercaptoethanol, 100U/ml penicillin and 100\(\mu\)g/ml streptomycin. H-2K\(^b\)–expressing HEK293 (HEK293KbC2), a gift from Jonathan Yewdell (NIH), were cultured in DMEM supplemented with 8% bovine growth serum, 2mM glutamine, 5U/ml penicillin and 5\(\mu\)g/ml streptomycin.

pA3F-LANA, pLVX-YFP-Flag and pLVX-L-YFP-Flag were described earlier (56, 57). hnRNP A1 was generated by PCR amplification and cloning into a Flag-tagged vector: pA3F. LANA luciferase was generated by sub cloning LANA in pGL3 basic vector. pA3F-G4 wild type clone (containing G-quadruplex forming sequence of LANA) and pA3F-G4 disrupted clones (codon optimized to disrupt the formation of G-quadruplexes but coding for the same amino acids) were commercially synthesized from Genscript (Genscript, Inc.). pA3F-G4 wild type ova, pA3F-G4 disrupted ova and pA3F-LANA ova were generated by sub cloning the sequence coding for SIINFEKL peptide in pA3F-G4 wild type, pA3F-G4 disrupted and pA3F-LANA, respectively. The integrity of the clones was confirmed by DNA sequencing performed at the Nevada Genomics Center, University of Nevada, Reno, NV.

Antibodies

The following commercial antibodies were used for this study: mouse anti-GAPDH (US Biological), mouse anti-Flag M2 (Sigma-Aldrich, St. Louis, MO, USA), hnRNP A1 (Santa Cruz Biotechnology), Lamin A/C (Santa Cruz Biotechnology). Mouse monoclonal anti-LANA hybridoma was generated at Genscript (Genscript, Inc.).

CD spectroscopy

CD spectroscopy was performed on Aviv Biomedical spectrometer, where circularly polarized UV light was used to record CD spectra at 25\(^\circ\)C in a series of progressive scans from 320nm to 200nm in a quartz cuvette of 1mm path length. The RNA oligonucleotides, LANA-wt
(5’- UGGAAGAGCAGGAAGAGCAGGAGUUAGAGGA-3’) and LANA-scrambled (5’-UAACCGAUGAUAUGAGUCAGAUAUAUAAGCA-3’) used to record CD spectra were resuspended at a final concentration of 5µM in sodium cacodylate buffer (10mM, pH 7.4) with 100mM KCl.

**Electrophoretic mobility shift assay**

Electrophoretic mobility shift assay (EMSA) was performed as described before (31). Briefly, wild type LANA (5’-UGGAAGAGCAGGAAGAGCAGGAGUUAGAGGA-3’) and scrambled LANA (5’-UAACCGAUGAUAUGAGUCAGAUAUAUAAGCA-3’) RNA oligonucleotides were labeled with radioactive $^{32}$P using T4 polynucleotide kinase (New England Biolabs). The oligonucleotides were resuspended at a final concentration of 2µM in 10mM sodium cacodylate buffer with 100mM KCl and resolved on a 15% non-denaturing polyacrylamide gel. The oligos were resolved in 1X TBE buffer with 100mM KCl either in the presence or absence of antisense oligos. The gels were dried using Gelair gel dryer (Bio-Rad Inc.) and autoradiography was performed using a phosphorimager (GE Healthcare Life Sciences).

**In vitro translation and RNA quantification**

*In vitro* translation of pA3F-G4 wild type and G4 disrupted was performed using the Promega TNT T7 quick-coupled transcription/translation system; where 2µg of the plasmid was translated in a 50µl reaction containing 1mm methionine at 30°C for 2 hours. The translated product was resolved using SDS-PAGE and proteins were detected using anti-flag antibody. RNA levels in the *in vitro* translated samples were determined by extracting the RNA from the translation mix using TRIzol reagent (Thermo Fisher Scientific) as per the manufacturer’s protocol. cDNA was synthesized using high-capacity RNA-to-cDNA kit (Applied Biosystems
Inc.) as per the manufacturer’s protocol and quantified using gene specific primers: G4-wt-forward; 5’-GAGCAGGAATTAGAGGAGGT-3’, G4-wt-reverse; 5’-CTGCTCCTGCTCTTCCA-3’ and G4-disrupted-forward; 5’-AGTAGAAAGAACAAGAGC-3’, G4-disrupted-reverse; 5’-CTACTTTCTCTAATTCTTG-3’)

RNA quantification and protein expression following compound treatment

Cells were treated with 10µM TMPyP4 for 24 hours, harvested and lysed in NP-40 cell lysis buffer (1% Nonidet P-40, 50mM Tris-HCl, pH 7.5, 150mM NaCl and 1mM EDTA) supplemented with protease inhibitors (1mM phenylmethylsulfonyl fluoride, 10µg/ml pepstatin, 10µg/ml leupeptin and 10µg/ml aprotinin). Cells treated with DMSO were used as a control and the cellular lysates were resolved on SDS-PAGE, proteins were detected using specific antibodies. To analyze the effect of TMPyP4 treatment on transcription, cells were treated with compound for 24 hours following which total RNA was extracted using Illustra RNAspin Mini Kit (GE Healthcare) according to the manufacturer’s protocol. cDNAs were synthesized using high-capacity RNA-to-cDNA kit (Applied Biosystems Inc.) as per the manufacturer’s protocol and quantified using gene specific primers. The Ct values were normalized to the housekeeping gene, GAPDH (forward, 5’-CAGCAAGAGCACAAGAGGAAGA-3’; reverse, 5’-TTGATGGTACATGACAAGGTGCGG-3’). All the reactions were run in triplicates.

RNA Crosslinking immunoprecipitation assay

RNA crosslinking immunoprecipitation assay was performed as described before (58). Briefly, 10 million cells were harvested, fixed in 1% formaldehyde and quenched by the addition of 125mM glycine, following which the cells were washed and the cells pellet was resuspended in cell lysis buffer [5mM Pipes (KOH) pH 8.0, 85mM KCl, 0.5% NP-40, fresh protease inhibitors and RNAseOUT]. After incubation for 10 minutes on ice, the nuclei were isolated
following centrifugation and resuspended in Diagenode chromatin shearing buffer D (with protease inhibitors and RNaseOUT). The nuclei were sheared to an average chromatin size of 200-400bp using BioRuptor (Diagenode, Inc.) following which the cell debris were pelleted and the sheared chromatin was diluted with three volumes of Dilution buffer (1.2mM EDTA, 16.7mM Tris pH 8.0, protease inhibitors and RNaseOUT). The diluted chromatin was incubated with specific antibodies overnight and protein A/G magnetic sepharose beads were added for two hours to capture the immune complexes. The beads were washed with low salt buffer (0.1% SDS, 1% Triton-X100, 2mM EDTA, 20mM Tris pH 8.0, 150mM NaCl) twice and once with high salt buffer (0.1% SDS, 1% Triton-X100, 2mM EDTA, 20mM Tris pH 8.0, 500mM NaCl). The beads were finally washed with 1X TE and resuspended in 150µl elution buffer (1% SDS, 100mM NaHCO₃) with gentle vortexing for 15 minutes to elute the complex. The samples were reverse cross-linked overnight at 65°C using 0.3M NaCl. Next day, the samples were treated with proteinase K at 45°C for 2 hours followed by DNase treatment at 37°C for 30mins. TRIzol was added to the samples, incubated for 5 minutes followed by addition of chloroform and incubation at room temperature for 15 minutes. The samples were centrifuged and RNA in the aqueous phase was precipitated by the addition of 1X isopropanol and ammonium acetate. The samples were allowed to precipitate at -80°C for atleast 30 minutes following which the samples were centrifuged and the pellets were washed twice with 75% ethanol. Finally, the pellet was resuspended in 30µl nuclease free water.

**In vitro RNA pull-down assay**

Wild type LANA RNA oligo (5’-UGGAAGAGCAGGAAGAGCAGGAGUUAGAGGA-3’) and scrambled LANA RNA oligo (5’-UAACCGAUGAUAUGAGUCAGAUAUAUAGCA-3’) were biotinylated at the 3’end by
Pierce 3’ RNA biotinylation kit according to the instruction manual. To prepare the cellular lysate, approximately 20 million KSHV-positive BCBL-1 cells were harvested and lysed in 1% NP-40 lysis buffer with protease inhibitors and RNaseOUT. The samples were sonicated and centrifuged to remove cellular debris. The purified biotinylated RNA oligos were incubated with the lysate for three hours at 4°C. Pierce Streptavidin agarose beads (Thermo Fisher Scientific) were then added to the samples to pull down the proteins bound to the biotinylated RNA oligos following incubation with the beads for two hours. The beads were washed thrice with 1% NP-40 lysis buffer following which they were loaded onto a 9% SDS-PAGE, transferred on nitrocellulose membrane and probed with specific antibodies. The IP samples were also subjected to LC/MS analysis at Mitch Hitchcock Nevada Proteomics Center, University of Nevada, Reno.

**Dual Luciferase assay**

LANA was sub cloned in pGL3 basic vector with luciferase gene downstream of the LANA coding region. HEK293T cells were transfected with 250ng pGL3-LANA and 10ng Renilla luciferase expressing plasmid, pRRLSV40. 24 hours post transfection; the cells were harvested, lysed in 1% NP-40 lysis buffer and centrifuged to remove cell debris. The firefly reporter was measured by adding Luciferase assay Reagent II (LARII) (Promega, Inc.) to a 96-well plate containing supernatant from the transfected cells and quantified by the luminometer. The reaction was stopped and Renilla luciferase activity was measured by addition of Stop and Glo reagent and quantified using the luminometer. The relative luciferase levels were calculated after normalization of firefly luciferase with Renilla luciferase to account for the transfection discrepancies.
Flow cytometry

YFP and L-YFP plasmids were transfected in HEK293T cells as described previously. Four hours post transfection the cells were treated with DMSO or TMPyP4. 24 hours post treatment; the cells were trypsinized, washed with PBS, and then fixed in PBS supplemented with 2% paraformaldehyde for 15 min at room temperature. Finally, the cells were washed once with PBS and the cell pellet was resuspended in PBS. GFP-positive cells were detected using a FACS Calibur flow cytometer equipped with Cell Quest Pro software and analyzed using FlowJo software. The fluorescence intensity was expressed as a percentage of total number of cells and the reduction in florescence intensity due to TMPyP4 treatment was calculated for both YFP and LYFP transfected cells.

Antigen presentation assay

HEK293Kbc2 cells were transfected with LANA-ova, G4 wild type-ova and G4 disrupted-ova plasmids and harvested 24 hours post transfection. The cells were incubated with T cell hybridoma, B3Z cells for 18 hours at effector to target ratios of 1:2, 1:1 and 1:0.5. The cells were incubated with buffer containing 0.125% NP-40, 9mM MgCl₂, 100mM β-mercaptoethanol and 5mM ortho-nitrophenyl-β-galactoside (ONPG) for 4 hours at 37°C. The absorbance was measured at 450nm using EMax Absorbance Microplate Reader (Molecular Devices, USA) and read using SoftMax Pro v. 6.4 software.

RNA subcellular isolation

293L cells were transfected with G4 wild type or G4 disrupted along with different concentrations of LANA. Cells were harvested 24 hours post transfection; total and cytoplasmic RNA extraction was performed using Active motif RNA subcellular isolation kit according to the manufacturer’s instructions. cDNA was synthesized using ABI cDNA synthesis kit and G4
wild type/G4 disrupted RNA was quantified using gene specific primers. Normalization of cytoplasmic RNA was performed using total RNA content and the RNA content of the cytoplasmic fractions was compared between the samples.

**Table 1: List of oligonucleotides and primers used in the study**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>LANA wild type RNA oligo</td>
<td>UGGAAGAGCAGGAAGAGCAGGAGUGAAGAGGA</td>
</tr>
<tr>
<td>LANA scrambled RNA oligo</td>
<td>AAACAGGUCUCCGGAAGAUGAUG</td>
</tr>
<tr>
<td>G4 wild type LANA Forward</td>
<td>GAGCAGGAATTAGAGGAGGT</td>
</tr>
<tr>
<td>G4 wild type LANA Reverse</td>
<td>CTGCTCTGCTCTTCCCA</td>
</tr>
<tr>
<td>G4 disrupted LANA Forward</td>
<td>AGTGAAGAAACAAGAGC</td>
</tr>
<tr>
<td>G4 disrupted LANA Reverse</td>
<td>CTACCTCCTAATTCTTG</td>
</tr>
<tr>
<td>AS1 G4 wild type LANA</td>
<td>TTCCCTCTCTTCCAC</td>
</tr>
<tr>
<td>AS2 G4 wild type LANA</td>
<td>CTCCCTCAACTCCTG</td>
</tr>
<tr>
<td>G4 Disrupted-CLIP Forward</td>
<td>GAATTAGAAGAAGTAAGAAGAAGAAGAAGAGCAG</td>
</tr>
<tr>
<td>G4 Disrupted-CLIP Reverse</td>
<td>TTGTCGTCATCGTCTTTTGATGCA</td>
</tr>
<tr>
<td>G4 Wild type-CLIP Forward</td>
<td>TGCAGGAGAGAGCAGGAATTAG</td>
</tr>
<tr>
<td>G4 Wild type-CLIP Reverse</td>
<td>TGTCGTCATCGTCTTTTGATGCA</td>
</tr>
<tr>
<td>7sk forward</td>
<td>ATCTGTACCCACATTGATCG</td>
</tr>
<tr>
<td>7sk reverse</td>
<td>GCAGCATCTACTGATACCC</td>
</tr>
<tr>
<td>LANA 1-32 Forward</td>
<td>TTGAATTCTATGGCGCCCCCGGGGA</td>
</tr>
<tr>
<td>LANA 1-32 Reverse</td>
<td>TTTGGATCCTCTTTCCGGAGACCTGT</td>
</tr>
</tbody>
</table>

**Statistical analysis**

P-values were calculated by two-tailed t test using GraphPad Inc. (Prism 8) software for statistical significance. Asterisks represent the P value < 0.05 (*), P value < 0.01 (**) and <0.001 (***) , where ns denotes non-significant.

**RESULTS**

**LANA mRNA forms G-quadruplexes**

G-quadruplexes are formed by hoogsteen hydrogen bonding between G-quartets and is stabilized by the presence of K+ ions. Due to their roles in modulating vital biological processes,
these structures have recently gained importance as potential drug targets. EBV and KSHV, both members of the Herpesvirus family are known to share functional similarities. Of our particular interest is the expression of latent proteins and immune evasion strategies during latency. Expression of EBNA1, the functional homolog of KSHV LANA in EBV is regulated by the presence of G-quadruplexes in its mRNA, which led us to determine the effects of G-quadruplexes on LANA expression (30). First of all, we analyzed the LANA mRNA sequence for the presence of G-quadruplex using a web-based tool, QGRS mapper (http://bioinformatics.ramapo.edu/QGRS/index.php) that predicts the formation of these structures (42). The sequence coding for LANA mRNA was imported into the software and the probability of G-quadruplex formation, represented as G-scores (high value implies a higher likelihood of G-quadruplexes formation) were determined in LANA mRNA shown by yellow highlighted region (Fig. 1A). Based on the G-scores, a region in the QE rich domain of LANA mRNA, showed a high likelihood of G-quadruplex formation (Fig. 3A, panel a). G-quartets are building blocks of G-quadruplexes, which are formed by Hoogsten hydrogen bonding between four guanine residues as illustrated by Fig. 1B. These co-planar G-quartets stack upon each other to form a G-quadruplex and this interaction is stabilized by the presence of N+ or K+ ions (Fig. 1B).

Due to their complex structure and folding, G-quadruplexes possess unique biophysical and biochemical properties. Circular dichroism (CD) spectroscopy has been used to study secondary structures of nucleic acids, and expresses positive and negative CD signals at specific wavelengths depending upon the structure of G-quadruplexes (43). To validate the presence of G-quadruplexes in LANA mRNA, we performed CD spectroscopy on wild type LANA RNA oligo and scrambled RNA oligo. We observed a CD spectra pattern with maximum absorption at
260 nm and a negative minimum at 240 nm for wild type LANA oligo, a characteristic pattern of RNA G-quadruplexes (Fig. 1C, panel a). In contrast, the scrambled LANA oligo, did not show such pattern, validating the absence of G-quadruplexes (Fig. 1C, panel b). Furthermore, the presence of G-quadruplexes in the LANA mRNA was also confirmed by electrophoretic mobility shift assay (EMSA). The wild type and scrambled LANA RNA oligos were labeled with $^{32}$P and resolved on a native polyacrylamide gel. Antisense oligos (AS1 and AS2) complementary to LANA wild type RNA oligo, added in molar excess, were used in the indicated lanes to confirm the specificity of the mobility shift by G-quadruplex forming sequence.
oligos were resolved on a native gel, which showed a shift in the mobility of wild type LANA RNA oligo in the presence of specific antisense oligos as compared to the wild type LANA oligo alone (Fig. 1D, compare lane 3 with lane 1). This shift is attributed to the disruption of G-quadruplex structure formation by antisense oligos as they bind to the complementary G-rich region on LANA oligo, thus making it unavailable for the formation of G-quadruplexes. These results validated the presence of G-quadruplex structures in LANA mRNA.

**G-quadruplex stabilizing compound inhibits the translation of LANA mRNA**

G-quadruplex formation has been shown to influence a number of biological processes including the translation of mRNA. Since the translation of EBNA1, a homolog of LANA, was severely impaired due to the stabilization of G-quadruplexes, we asked whether the G-quadruplex in LANA mRNA affected its translation. To address this, we used TMPyP4, a widely used compound that stabilized the formation of G-quadruplexes and treated KSHV positive, BC3 and BCBL-1 cells to analyze the effects of G-quadruplex formation on endogenous expression of LANA. Upon treatment with the compound for 24 hours, we observed a reduction in the endogenous levels of LANA in both the BC3 and BCBL-1 cells, compared to control DMSO-treated cells (Fig. 2A, panel a, compare lanes 1 and 3 with lanes 2 and 4, respectively). Since the G-quadruplex stabilizing compounds are known to act on G-quadruplexes formed on both DNA and RNA, we wanted to ensure that the transcription of LANA mRNA was not playing a role in reducing the levels of LANA expression. Hence, we extracted RNA from TMPyP4 or DMSO-treated BC3 and BCBL-1 cells and quantified the levels of LANA mRNA by qPCR using gene-specific primers. The relative mRNA levels of LANA, normalized to the GAPDH, showed similar levels in both the control and TMPyP4-treated cells (Fig. 2A, panel b). Next, we wanted to confirm that the reduction in endogenous LANA expression in KSHV-infected cell lines was
not because of any other viral latent proteins, we determined the effect of TMPyP4 on exogenously expressing LANA cells. We utilized LANA expressing, BJAB and HEK293T cells to treat with TMPyP4 and assayed LANA expression. BJABs stably expressing LANA (L-YFP-Flag) and the HEK293T cells transfected with pA3F-LANA were treated with DMSO or TMPyP4 for 24 hours before assaying LANA levels using anti-flag antibody. Similar to the above observations, we detected a reduction in the expression of LANA in TMPyP4-treated cells as compared to control DMSO-treated cells (Fig. 2A, panels c and e, compare lanes 1 and 2). We also ensured that TMPyP4 treatment didn’t lead to any disparity in LANA mRNA levels due to compound treatment (Fig. 2A, panels d and f). We also analyzed the reduction in LANA expression quantitatively by detecting the levels of LANA fused YFP signals through flow cytometry after treating the cells with TMPyP4 or DMSO. Cells transfected with YFP alone were used as control in our experiments. Reduction in fluorescence intensity representing the levels of YFP fused LANA protein showed a significant reduction following TMPyP4 treatment as compared to the control, YFP expressing cells (Fig. 2B). Furthermore, we determined the effects of TMPyP4 on LANA translation in a luciferase assay where the luciferase gene was cloned downstream in frame with LANA gene (Fig. 2C, panel a). HEK293L cells transfected with pGL3-LANA were treated with TMPyP4 for 24 hours, lysed and luminescence was quantified after incubation with a specific substrate. We observed luminescence reduction in LANA-luciferase expressing cells treated with TMPyP4 as compared to the control DMSO-treated cells (Fig. 2C). As expected, levels of mRNA in TMPyP4 or DMSO-treated cells were similar (Fig. 2C, panel b). This supported our hypothesis that stable G-quadruplexes are formed on LANA mRNA and stabilization of these structures on mRNA led to a reduction in translation possibly, because of a steric roadblock for the movement of translation machinery.
Fig. 2. Stabilization of G-quadruplex formation by TMPyP4 inhibited LANA translation

A. (a) Immunoblot depicting endogenous protein levels in BC3 and BCBL-1 cells following treatment with TMPyP4. BC3 and BCBL-1 cells were treated with 10µm TMPyP4 for 24 hours following, which cells were harvested, lysed in NP-40 lysis buffer and the lysates were resolved on SDS-PAGE for the detection of LANA using anti-LANA antibody. Cells treated with DMSO were used as a negative control. Anti-GAPDH antibody was used to ensure equal loading of the proteins. (b) Relative mRNA levels of LANA in BC3 and BCBL-1 cells following treatment with 10µm TMPyP4 for 24 hours. RNA was extracted from compound treated cells, cDNA was synthesized and quantified using primers specific for LANA. The values were normalized against GAPDH. DMSO treated cells were used as a negative control. (c) Immunoblot depicting endogenous protein levels in BJAB-LYFP cells following treatment with TMPyP4. BJAB-LYFP cells were treated with 10µm TMPyP4 for 24 hours and the cellular lysates were resolved on SDS-PAGE and proteins were detected using anti-Flag antibody. Cells treated with DMSO were used as a negative control. Anti-GAPDH antibody was used to ensure equal loading of the proteins. (d) Relative mRNA levels of LANA in BJAB-LYFP cells following treatment with 10µm TMPyP4 for 24 hours. RNA was extracted from compound treated cells, cDNA was synthesized and quantified using primers specific for LANA. The values were normalized against GAPDH. DMSO treated cells were used as a negative control. (e) Immunoblot depicting the levels of overexpressed protein, LYFP, following treatment with TMPyP4. HEK293T cells were transfected with LYFP, treated with 10µm TMPyP4 and harvested 24 hours post treatment. The cellular lysates were resolved on SDS-PAGE and proteins were detected using anti-Flag antibody. Cells treated with DMSO were used as a negative control. Anti-GAPDH antibody was used to ensure equal loading of the proteins. (f) Relative mRNA levels of LANA in HEK293T cells transfected with, following treatment with 10µm TMPyP4 for 24 hours. RNA was extracted from compound treated cells, cDNA was synthesized and quantified using primers specific for LANA. The values were normalized against GAPDH. DMSO treated cells were used as a negative control.

B. Mean reduction in fluorescent intensity following TMPyP4 treatment. HEK-293L cells were transfected with YFP and LYFP, treated with 10µm TMPyP4 or DMSO for 24 hours and fixed with paraformaldehyde. Flow cytometry was performed on the cells to quantify the number of GFP positive cells. Reduction in fluorescence was determined by comparing the fluorescence in TMPyP4 treated cells as compared to the DMSO treated cells. LYFP expressing cells showed a quantifiable reduction in fluorescence following TMPyP4 treatment.

C. (a) Schematic illustration of LANA luciferase plasmid. LANA was subcloned in pGL3 basic vector upstream of the luciferase gene in frame to make a LANA-luciferase fusion protein. (b) Luciferase assay showing reduction in luciferase levels following TMPyP4 treatment in HEK293L cells. Cells were transfected with LANA-pGL3 and treated with TMPyP4 or DMSO for 24 hours. These cells were used for the extraction of RNA and luciferase assay. Relative mRNA levels were expressed relative to the GAPDH. Luciferase levels were significantly reduced in TMPyP4 treated cells despite of similar RNA levels.

Region of LANA mRNA with G-quadruplex sites negatively regulated translation

After having established that G-quadruplex stabilization inhibits LANA translation, we wanted to selectively analyze the G-rich sequence of LANA mRNA for its translation in the absence of G-quadruplex stabilizing compounds. To achieve this, we generated 250 nt long QE
region of LANA (857-915aa) in pA3F vector (G4 wild type) encompassing the G-rich residues that forms G-quadruplex (Fig. 3A, panels a and b). G4 disrupted LANA construct, where the G-residues were mutated (marked in red) so it could no longer form a G-quadruplex, but still code for the same amino acids as the G4 wild type (Fig. 3A, panel c). In order to assess the effect of G-quadruplex on translation of LANA mRNA, we first performed an in vitro translation assay using G4 wild type and G4 disrupted constructs. The resulting protein following TNT T7 transcription/ translation was resolved on SDS-PAGE and detected using anti-flag antibody. Comparison of the protein bands showed significantly reduced levels in G4 wild type as compared to the G4 disrupted clone, suggesting that the presence of G-quadruplexes in LANA mRNA modulated its translation (Fig. 3B, panel a). It can be argued that the differences in the translation levels could be due to a difference in transcription from two different constructs. To address this, we isolated RNA from in vitro transcription/translation reaction for G4 wild type and G4 disrupted and quantified them through qPCR using gene-specific primers, which showed almost similar levels of RNA in both G4 wild type and G4 disrupted samples (Fig. 3B, panel b).

To analyze the effect of G-quadruplexes on the expression in human cells, we transfected these constructs in HEK293T cells and detected the expression of these proteins 24 hours post-transfection following SDS-PAGE and anti-Flag detection.

Interestingly, we did not detect any expression in cells transfected with G4 wild type, as compared to cells transfected with G4 disrupted, which displayed a band specific for G4 disrupted protein (Fig. 3C, panel a lanes 2 and 3). This suggested that G4 wild type mRNA formed stable G-quadruplex structures, which abrogated the translation completely as the levels of transcribed mRNA were almost similar (Fig. 3C, panel b). To further confirm that the reduction in expression was specifically due to the formation of G-quadruplex structures, we
disrupted these structures by anti-sense oligonucleotides, tested in gel shift assay, and assayed protein expression. We used anti-sense non-specific oligonucleotides, which were unable to disrupt the G4 structure, as a control. In vitro transcription/translation on G4 wild type, G4 wild type with specific anti-sense oligos; Sp-As (AS1: 5’- TTCCTGCTCTCCAC-3’; AS2: 5’- CTCTCTAACTCCTG-3’), and G4 wild type with non-specific oligos; nSp-As (NS1: 5’- GCCACCGAACAACCCC-3’; NS2: 5’- CACTAGCCCCCCCCC-3’), and resolving them on SDS-PAGE and detection with flag antibody, showed a significant increase in expression of G4-
wild type in the presence of anti-sense oligos, as compared to an overall lower expression when G4 wild type was translated alone or in the presence of non-specific antisense oligos (Fig.3D, panel a). The difference in the mRNA levels was statistically insignificant under these conditions (Fig.3B panel b, 3C panel b and 3D panel b). These observations confirmed that G-quadruplex formation inhibits the translation of G-rich LANA mRNA.

G-quadruplexes modulate antigen presentation

LANA has been shown to evade the host’s immune surveillance by inhibiting MHC Class I and Class II antigen presentation. Additionally, the presence of G-quadruplexes regulates its expression, therefore we wanted to investigate how G-quadruplexes of LANA mRNA affect immunomodulation. To this end, we cloned ovalbumin epitope (SIINFEKL) downstream of G4 wild type, G4 disrupted and full-length LANA, to generate ovalbumin fusion constructs and transfected into an antigen presenting cells (HEK293Kbc2). Cells expressing G4 wild type or G4 disrupted fused with ovalbumin epitope were harvested 24 hours post transfection and incubated with different ratios of T cell hybridoma, B3Z cells in a 96-well plate, at 37\(^{0}\)C for 18 hours. 150\(\mu l\) medium was replaced with a solution containing 5mm ONPG, 0.5% NP40 and incubated at 37\(^{0}\)C for 4 hours, following which the \(\beta\)-galactosidase activity was quantified by measuring the OD at 450nm. Upon comparing relative levels of \(\beta\)-galactosidase between G4 wild type ova and G4 disrupted ova-transfected cells, we found a higher activity in G4 disrupted ova-transfected cells as compared to G4 wild type clone (Fig.4B, panel a). This confirmed that G-quadruplex formation in the G4 wild type mRNA inhibited its translation leading to a reduced expression and a lower surface antigen presentation on HEK293Kbc2 cells. Similarly, the \(\beta\)-galactosidase activity of B3Z cells incubated with TMPyP4 treated full length LANA-ova transfected cells was lower than DMSO-treated cells (Fig.4C, panel a).
A. Transfection of G4 wild type-SIIN/G4 disrupted-SIIN or Full length LANA-SIIN (treated with DMSO/TMPyP4)

H-2Kb expressing HEK293 cells
B3Z cells (T cells hybridoma expressing TCR)

β-galactosidase

Measure β-galactosidase levels

B.  

a. 

Absorbance (450 nm)

B3Z: 293Kbc2 ratio

G4 wild type  
G4 disrupted

1:2 1:1 1:0.5

b. 

Ovalbumin

Relative mRNA levels

G4 wild type  
G4 disrupted

C.  

a. 

Absorbance (450 nm)

B3Z: 293Kbc2 ratio

DMSO  
TMPyP4

1:2 1:1 1:0.5

b. 

Ovalbumin

Relative mRNA levels

LANA  
DMSO  
TMPyP4
This confirmed that G-quadruplex formation in LANA mRNA leads to reduced translation, hence a lower surface expression and antigen presentation. We ensured to check the mRNA levels, which were uniform in all the samples negating any effect on antigen presentation due to the altered transcription (Fig. 4B, panel b and 4C, panel b).

RNA G-quadruplexes interact with cellular proteins

Since the G-quadruplexes are known to interact with a number of proteins that help to open or stabilize the secondary structures, we wanted to identify proteins that interacted with LANA G4 mRNA. To achieve this, we performed biotinylated RNA pull-down assay, where wild type LANA RNA oligo with G4 site and a scrambled oligo with disrupted G4 were biotinylated at the 3’ end using Pierce RNA biotinylation kit. The biotinylated RNA was purified
and incubated with cellular extract from KSHV-positive BCBL-1 cells for three hours, following which streptavidin beads were added for two hours, to capture the proteins associated with the RNA. The beads were washed stringently to remove any non-specific proteins before subjecting them for mass spectrometry analysis (Table 2).

Table 2: List of proteins bound to the LANA RNA with G-quadruplexes

<table>
<thead>
<tr>
<th>Name of the protein</th>
<th>Accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterogeneous nuclear ribonucleoproteins A2/B1</td>
<td>ROA2_HUMAN</td>
</tr>
<tr>
<td>Epididymis secretory sperm binding protein</td>
<td>A0A024RB53_HUMAN</td>
</tr>
<tr>
<td>Heterogeneous nuclear ribonucleoprotein A3</td>
<td>P51991</td>
</tr>
<tr>
<td>Heterogeneous nuclear ribonucleoprotein A/B</td>
<td>D6R9P3_HUMAN</td>
</tr>
<tr>
<td>Heterogeneous nuclear ribonucleoprotein A1</td>
<td>P04256</td>
</tr>
<tr>
<td>Histone H2A</td>
<td>A0A024R017_HUMAN</td>
</tr>
<tr>
<td>Cell growth-inhibiting protein 34</td>
<td>Q08ES8_HUMAN</td>
</tr>
<tr>
<td>40S ribosomal protein S16</td>
<td>RS16_HUMAN</td>
</tr>
<tr>
<td>ATP synthase subunit gamma, mitochondrial</td>
<td>ATPG_HUMAN</td>
</tr>
<tr>
<td>Nucleoside diphosphate kinase A</td>
<td>E7ERL0_HUMAN</td>
</tr>
<tr>
<td>Heterogeneous nuclear ribonucleoprotein H3</td>
<td>P31942</td>
</tr>
<tr>
<td>Propionyl-CoA carboxylase beta chain, mitochondria</td>
<td>P05166</td>
</tr>
<tr>
<td>Coiled-coil domain-containing protein</td>
<td>A2RUR9</td>
</tr>
<tr>
<td>Four and a half LIM domains protein 3</td>
<td>Q13643</td>
</tr>
<tr>
<td>Elongation factor 1-beta</td>
<td>P24534</td>
</tr>
<tr>
<td>Serine/arginine-rich splicing factor 1</td>
<td>Q07955</td>
</tr>
<tr>
<td>Poly(rC)-binding protein 1</td>
<td>O19048</td>
</tr>
<tr>
<td>Calcium-activated chloride channel regulator 2</td>
<td>Q8BG22</td>
</tr>
<tr>
<td>60S ribosomal protein L11-1</td>
<td>A8WQ43</td>
</tr>
<tr>
<td>Heterogeneous nuclear ribonucleoprotein C</td>
<td>G3V9R8</td>
</tr>
<tr>
<td>RNA-binding motif protein, X chromosome</td>
<td>Q9WV02</td>
</tr>
<tr>
<td>TAR DNA-binding protein 43</td>
<td>TARDBP</td>
</tr>
<tr>
<td>Heterogeneous nuclear ribonucleoprotein K</td>
<td>O19049</td>
</tr>
<tr>
<td>Elongation factor 1-beta</td>
<td>P24534</td>
</tr>
<tr>
<td>Serine/arginine-rich splicing factor 3</td>
<td>P84104</td>
</tr>
<tr>
<td>Succinate--CoA ligase [ADP-forming] subunit beta</td>
<td>Q482S1</td>
</tr>
<tr>
<td>Small nuclear ribonucleoprotein Sm D1</td>
<td>Q4R5F6</td>
</tr>
<tr>
<td>THAP domain-containing protein 10</td>
<td>Q5NVVM3</td>
</tr>
<tr>
<td>Interleukin enhancer-binding factor 2</td>
<td>Q5RFJ1</td>
</tr>
<tr>
<td>30S ribosomal protein S8</td>
<td>Q06FM9</td>
</tr>
<tr>
<td>Guanine deaminase</td>
<td>Q9WTT6</td>
</tr>
<tr>
<td>Glycine--tRNA ligase</td>
<td>Q8L785</td>
</tr>
<tr>
<td>Adenosine deaminase</td>
<td>C6DH28</td>
</tr>
<tr>
<td>Quinolinate synthase A</td>
<td>A5VZW1</td>
</tr>
</tbody>
</table>
Our assay identified a large number of proteins enriched in wild type LANA RNA pull down samples in contrast to the scrambled RNA sample. Amongst these, we found the hnRNP family of proteins specifically associating with the wild type RNA oligo as compared to the scrambled oligo. We confirmed this interaction through immunoblot after resolving the proteins pulled down with biotinylated RNA oligo, which specifically came down with oligo having wild type G4 sites (Fig. 5A). Further, we wanted to determine whether hnRNP A1 binds to the RNA with G4 sites in vivo. To this end, we performed a cross-linking RNA immunoprecipitation assay on HEK239T cells transfected with G4 wild type or G4 disrupted constructs. Cells were harvested 24 hours post transfection and the cell lysates were incubated with hnRNP A1 antibody to capture RNA associated with hnRNP A1. The RNA-protein immune complexes were reverse crosslinked and the bound RNA was quantified using primers specific for G4 wild type or G4 disrupted mRNA and the relative bindings were calculated relative to the input samples. Our data showed that hnRNP A1 specifically binds to G4 wild type mRNA as compared to the G4 disrupted mRNA thus confirming our previous observation of biotinylated RNA pull down that hnRNP A1 binds to G-quadruplex forming mRNA (Fig. 5B, panel a). To ensure the specificity of hnRNP A1 binding, we quantified it’s binding on a known, 7sk mRNA, which
showed similar binding in both G4 wild type and G4 disrupted samples (Fig. 5B, panel b). This confirmed that difference in the hnRNP A1 binding to mRNA is dependent on G4 sites. The next step was to analyze whether hnRNP A1 bound on LANA mRNA in KSHV-infected cells. To this end, we performed crosslinking RNA immunoprecipitation assay with hnRNP A1 on BCBL-1 cells.

We compared the binding of hnRNP A1 on LANA mRNA with and without stabilizing the G-quadruplexes by treating with TMPyP4 or DMSO, respectively. Comparison of relative hnRNP A1 binding to the G-rich region in BCBL-1 cells treated with DMSO or TMPyP4 showed higher affinity of hnRNP A1 to the DMSO-treated as compared to TMPyP4-treated cells (Fig. 5C, panel a). This suggested that hnRNP A1 selectively binds to the G-quadruplexes and stabilization with ligands, TMPyP4, reduces its binding. We also analyzed the binding of hnRNP A1 to a region on LANA mRNA (1-32aa), without G-quadruplexes structures, which showed significantly reduced binding confirming specificity of hnRNP A1 binding to the G-quadruplexes (Fig. 5C, panel b). hnRNP A1, has been shown to regulate a number of biological processes by unfolding of G-quadruplex structures (44, 45). Since hnRNP A1 selectively bound to the G-rich region of LANA mRNA, we wanted to analyze whether hnRNP A1 expression affected the translation of LANA. We first analyzed the effect of hnRNP A1 expression on full length LANA through luciferase assay. To do this, we transfected HEK293L cells LANA luciferase constructs in the presence of different concentrations of hnRNP A1. The cells were harvested 24 hours post transfection, lysed and the lysates were incubated with luciferase substrate for determining the levels of LANA fused luciferase. Our data showed a gradual increase in the luciferase levels, higher expression of proteins with increasing concentrations of hnRNP A1 (Fig. 5D). Furthermore, we confirmed the role of hnRNP A1 on enhancing the
expression of G-quadruplex containing mRNA by using G4 wild type and G4 disrupted clones. HEK293T cells were transfected with either pA3F-G4 wild type or pA3F-G4 disrupted along with increasing concentrations of hnRNP A1 and the lysates were resolved on SDS-PAGE 24 hours post-transfection.
Detection of LANA G4 wild type protein with anti-flag antibody progressively enhanced from no detectable expression to a significant expression with an increasing conc. of hnRNP A1 (Fig. 5E, IB:Flag). In contrast, hnRNP A1 expression had no effect on the G4 disrupted protein expression, which was evident by a similar level of G4 disrupted protein expression with varying concentration of hnRNP A1 (Fig. 5E panel b). This corroborated with our previous results and led us to conclude that hnRNP A1 binds selectively to G-quadruplex forming mRNA of LANA, unfolds these secondary structures and facilitates the translation of LANA mRNA.
LANA inhibits the nuclear export of G-rich mRNA by selective binding

A number of studies have reported the auto regulatory mechanisms of LANA through which LANA controls its own expression and antigen presentation. Here, we show that the expression of LANA is controlled by G-quadruplexes through the involvement of cellular proteins. We speculate that LANA regulates its expression by associating with G4 sites. We confirmed this by testing the binding of LANA to G4 wild type mRNA and G4 disrupted mRNA through RNA cross linking immunoprecipitation assay. HEK293L cells were transfected with pA3F-G4 wild type or pA3F-G4 disrupted along with full length LANA and the cells were harvested 24 hours post-transfection for chromatin isolation (Fig. 6A, panel a). The LANA bound chromatin was immunoprecipitated using anti-LANA antibody and the RNA associated with the chromatin was purified for quantitation using iTaq SYBR Green One step kit and gene-specific primers. Our data normalized to respective input samples, demonstrated specific binding between LANA and G4 wild type mRNA as compared to G4 disrupted mRNA thus confirming the specificity of LANA binding to the G-quadruplex forming RNA (Fig. 6A, panel b). Next, we wanted to confirm our findings that LANA binds to the G-rich region of LANA mRNA in KSHV positive cells. We achieved this by treating the cells with G4 stabilizing compound, TMPyP4. 24 hours post treatment, the cells were fixed with formaldehyde and the chromatin was sheared to immunoprecipitate with anti-LANA antibody. LANA bound RNA in the immune complex was purified for quantitation using One-step iTaq SYBR kit with primers specific to G-rich region. Our data showed that LANA binds to the G-rich of LANA mRNA in KSHV positive cells treated with TMPyP4 as compared to the control, DMSO treated cells, suggesting that LANA preferentially binds to the stabilized G-quadruplexes (Fig. 6B, panel a). The quantification of 1-32 aa region of LANA mRNA, which does not form G-quadruplexes, was
used as a control to confirm the specificity of LANA’s binding to the G-quadruplex region of the LANA mRNA (Fig. 6B, panel b).

We further wanted to explore how G-quadruplexes maintained the required level of LANA expression. To do this, we overexpressed LANA at varying concentrations along with LANA clones in frame with luciferase (pGL3-LANA). The cells were harvested 24 hours post transfection, lysed and the luciferase activity was measured. Comparison of relative luciferase from cells transfected with different LANA concentrations, we found an interesting pattern in luciferase expression with LANA concentrations. Luciferase units, indirectly representing the levels of LANA expression through G-quadruplexes gradually increased with increasing LANA concentration, however it started to decrease in higher LANA expressing cells (Fig. 6C). This led us to conclude that LANA has a self-regulatory role in controlling its own expression and our findings prove that hnRNP A1 and LANA modulate LANA’s expression through interaction with the G-quadruplex on LANA mRNA. Since our results show that transfection of pA3F-G4 wild type clone in HEK293T cells does not show any protein expression despite of similar mRNA levels to pA3F-G4 disrupted clone, which expresses an expected protein, we hypothesized that LANA could be blocking the mRNA export for translation through interacting with the G-quadruplexes. To prove this, we transfected HEK293L cells with G4 wild type or G4 disrupted clones along with LANA (at different concentrations) following which cells were harvested and the nuclear and cytoplasmic RNA were extracted using Active Motif nuclear/cytoplasmic RNA isolation kit. We also extracted total RNA from a portion of cells used for nuclear/cytoplasmic RNA extraction. cDNA was synthesized and quantified using gene specific primers. We calculated the amounts of mRNA exported to the cytoplasm (to determine
the amounts available for translation) relative to the total mRNA following normalization with housekeeping gene, GAPDH.
Comparison of the cytoplasmic mRNA levels of G4 wild type in cells transfected with different concentrations of LANA showed a slight increase in the cytoplasmic G4 wild type mRNA levels at lower LANA concentration (Fig. 6D, panel b). However, at higher LANA concentration, we observed a decline in the cytoplasmic levels of G4 wild type mRNA thus lesser availability for translation (Fig. 6D panel b). On the contrary, there was no significant difference in the cytoplasmic RNA levels of G4 disrupted mRNA in cells transfected with different amounts of LANA (Fig. 6D, panels c and d) This confirmed that LANA retained G4 wild type mRNA possibly through binding with the G-quadruplexes in the mRNA.
Discussion

KSHV establishes persistent latent infection in host cells and employs various mechanisms to avoid immune recognition. During latency, the virus expresses only a limited number of proteins of which LANA is the most abundantly expressed protein that ensures efficient viral replication and distribution of viral genome to the daughter cells (46). LANA has been shown to auto regulate its expression by regulating its promoter as well as inhibition of peptide synthesis and proteasomal degradation through two central region domains (CR2 and CR3; aa 442–768 and aa 769–920 respectively) of LANA (47-50). Additionally, CR1 (aa 330-442) domain of LANA has been shown to play a role in restricted MHC class I antigen presentation, through reduced translocation of peptides to the ER (13, 51). We have previously shown that LANA interacts with Regulatory Factor X proteins and disrupts the association of CIITA for MHC II gene expression (14). These studies substantiate the role of LANA in regulating its expression and restricting recognition by the host’s immune system; however, there has been no direct evidence demonstrating a link between these two processes. In this manuscript, we report the formation of G-quadruplexes in the G-rich regions of LANA mRNA and the immunological role of these structures in restricted immune recognition.

G-quadruplexes have gained recognition as regulatory nucleic acid structures that influence a number of biological processes. Many members of the Herpesviridae family have been shown to form G-quadruplexes that regulate viral gene expression, viral DNA replication and virion production (52, 53). Moreover, G-quadruplexes have gained recognition as immunomodulatory structures in few gammaherpesviruses, such as EBV and AIHV (Alcelaphine herpesvirus 1), where these structures abrogate protein synthesis and restrict antigen presentation thereby helping in immune evasion (30, 54). Upon analysis of LANA
mRNA sequence through QGRS mapper, a G-quadruplex motif predicting software, we found regions of LANA to form potentially stable G-quadruplexes. CD spectroscopy and EMSA further confirmed the presence of these secondary structures. We demonstrated the functional significance of G-quadruplexes formation through treatment with G-quadruplex stabilizing ligand, TMPyP4, which resulted in the reduction of LANA mRNA translation despite of similar levels of mRNA synthesis. This was also confirmed by treating the HEK293T cells expressing LANA, with G-quadruplex stabilizing ligand and determining the levels of translated LANA through a western blot. Stabilizing the G-quadruplexes though TMPyP4 of LANA mRNA fused to YFP showed a quantifiable reduction in LANA-YFP expression by flow cytometry. Additionally, the fusion of LANA coding sequence upstream of luciferase reporter gene demonstrated that stabilizing the G-quadruplexes of LANA mRNA reduces the expression of LANA-luciferase gene possibly by blocking the movement of translational machinery through these secondary structures. All these assays convincingly demonstrate the role of G-quadruplexes in controlling LANA expression. Further, we tested the specific role of these G-quadruplexes by generating a small clone of LANA peptide from the region of LANA gene containing G-quadruplexes. We also generated G-quadruplexes disrupted clone but with same aa sequence, through codon optimization, to ensure that the peptide sequence did not affect translation efficiency. Almost none to very little expression in G-quadruplex clone, despite equal mRNA levels, confirmed that regulatory role of these secondary structures. Disruption of these secondary structures with anti-sense oligos further substantiated the role of G-quadruplexes in regulating mRNA translation.

LANA has been shown to modulate antigen presentation by interacting with the components of antigen presentation pathway to limit its surface presentation (13, 14).
Interestingly, we found a link between the antigen presentation and G-quadruplex formation in LANA mRNA, as the antigen presentation of LANA was downregulated upon stabilizing the G-quadruplex structures in KSHV-positive cells along with highly inhibited antigen presentation in cells expressing the G-rich region of LANA. This led us to the conclusion that G-quadruplex formation inhibits antigen presentation through a reduction in the mRNA translation. G-quadruplex formation is shown to be regulated (stabilized/destabilized) through the involvement of a number of cellular proteins (21, 32). Our mass spectrometry results revealed hnRNP A1 as a highly specific interacting partner of the G-quadruplex forming mRNA. Previous studies have found hnRNP A1 to be a G-quadruplex destabilizing protein, which was similar to its effect on increasing the expression of G-quadruplex wild type clone (44, 45). This may suggest that hnRNP A1 increased the expression by selectively binding and potentially destabilizing the G-quadruplexes. Interestingly, we also discovered LANA to be binding at the G-quadruplex region of its mRNA and modulating its translation. Importantly, binding of LANA to the G-quadruplex region on its RNA inhibited the export of those mRNAs to the cytoplasm as LANA primarily localizes in the nucleus. Based on these observations, we propose a model in which the G-quadruplex formation on various regions of LANA including the 857-917aa (a sub-region of CR3 domain) controls LANA mRNA translation and expression of LANA protein (Fig.7). Expression of LANA is regulated by the relative amounts of LANA, and in cells where LANA levels are lower, relative binding of hnRNP A1 to LANA mRNA is higher, which leads to its export to the cytoplasm and destabilization of the G-quadruplexes. The translation machinery of the cytoplasm translates LANA mRNA leading to the formation of mature protein or Defective Ribosomal Proteins (DRiPs). The mature LANA proteins translocate to the nucleus, thus increasing the overall levels of LANA in the nucleus. When the concentration of LANA is
LANA is required for maintaining the viral genome in latent cells and its expression is modulated by the level of LANA in the cells. At lower level of LANA, hnRNP A1 binds to the G-quadruplex region of LANA mRNA and exports it to the cytoplasm following destabilization of the G-quadruplex structures for translation. Once a sufficient or higher level of LANA is synthesized and translocated to the nucleus, it binds to the G-quadruplex region of the LANA mRNA and retains them in the nucleus to control an excess synthesis of LANA protein. G-quadruplex mediated control of LANA mRNA translation ensures optimal expression without being overly producing mature or Defective Ribosomal Proteins (DriPs), which is presented through MHC class I molecules, to escape the host’s immune surveillance system.
sufficient enough in the nucleus, LANA binds to the G-quadruplex forming region of its mRNA in relatively higher copies than hnRNP A1, thereby inhibiting mRNA export into the cytoplasm, thus making it unavailable for translation and hence lower levels of protein is available for antigen presentation.

Acknowledgements

This work was supported by the National Institute of Health (CA174459 and AI105000). PD was partially supported by the Reno Cancer Foundation. We thank Dr. Matthew Tucker and Yftah Tal-Gan, Department of Chemistry, University of Nevada, Reno for helping with CD spectroscopy. We also acknowledge Dr. Sally Dupre for her help with Flow Cytometry.
References


55. **Thakker S, Strahan RC, Scurry AN, Uppal T, Verma SC.** 2017. KSHV LANA upregulates the expression of epidermal growth factor like domain 7 to promote angiogenesis. Oncotarget 9:1210-1228.


58. Rossetto CC, Pari G. 2012. KSHV PAN RNA associates with demethylases UTX and JMJD3 to activate lytic replication through a physical interaction with the virus genome. PLoS Pathog 8:e1002680.
RecQ1 helicase unwinds G-Quadruplexes at oriLyt to initiate KSHV lytic DNA replication

Prerna Dabral, Timsy Uppal and Subhash Verma

Abstract

KSHV utilizes distinct origins of DNA replication during the latent and lytic phases of its life cycle. Lytic DNA replication requires the participation of cellular and viral proteins and initiates from a specific region in the KSHV genome, called oriLyt. oriLyt region contains multiple domains and binding sites for viral and cellular proteins. We discovered the presence of multiple G-quadruplex (G4s) sites in the oriLyt region of the KSHV. G-quadruplexes are secondary structures in nucleic acid sequences that are considered as regulators of biological process like replication, transcription, translation and replication initiation. Herpesviruses have been reported to have a high density of PQS (putative quadruplex formation sites) in their genome including in the regulatory regions, which controls replication and transcription. The binding of RecQ1 to the oriLyt region of KSHV and its ability to unwind the G4 structures led us to speculate that G-quadruplex sites play an important role in lytic DNA replication. In this study, we confirmed the formation of stable G-quadruplexes through biochemical and biophysical assays. We also demonstrated the enrichment of RecQ1 at the G4 sites of the oriLyt. Functional significance of RecQ1 mediated unwinding of G4 sites were confirmed through inhibition and depletion of RecQ1 activity and proteins levels. Moreover, detection of replication initiation through single molecule analysis of the replicated DNA (SMARD) approach demonstrated that G4 stabilization lead to a subdued initiation of replication from oriLyt.
Introduction

Kaposi Sarcoma associated Herpesvirus (KSV) or HHV-8, a gammaherpervirus is the etiological agent of Kaposi Sarcoma, Primary effusion lymphoma, Multicentric Castleman’s disease and KSHV inflammatory syndrome (KICS) (1-4). Like other viruses of the Herpesviridae family, the life cycle of KSHV is biphasic, characterized by a default latent phase and a short and transient lytic phase. The latent phase of viral life cycle is characterized by a restricted gene expression and latent DNA replication, which occurs once per cycle facilitated by the latent associated nuclear antigen (LANA) after recruiting the cellular replication machinery to the origin of latent DNA replication (5). The lytic phase of infection involves orchestrated expression of Immediate early (IE), early (E) and late (L) genes with the replication of viral genome to produce infectious virions following packaging into capsids and teguments (6). Lytic replication initiates at the origin of lytic DNA replication (oriLyt), during which the virus uses viral and cellular proteins to replicate its genome presumably through a rolling circle mechanism. There are two copies of oriLyt in KSHV genome, namely oriLyt-L and oriLyt-R, present between open reading frames (ORFs) K4.2 and K5 and between K12 and ORF71, respectively (7, 8). The two oriLyt regions are 1.7Kb in length and consist a highly similar 1.1Kb long core sequence along with 600bp GC repeats, which are present in the form of 20bp and 30bp tandem repeats. The core sequence contains C/EBP motifs that are required for lytic DNA replication as they are involved in binding to the K8 protein of KSHV and an AT palindromic sequence involved in replication process. In addition, the core sequence also contains an RTA dependent promoter, which includes RTA binding site, RRE (RTA response element) and TATA box consensus sequence (9, 10). Some proteins that are essential for lytic DNA replication are a part of the core replication machinery and conserved among herpesviruses and include ORF 9
(DNA polymerase), ORF 6 (single-stranded binding protein), ORF 59 (processivity factor), ORF44 (helicase), PRF56 (primase), and ORF 40/41 (primase-associated factor) (11-13). ORF 50 or RTA and K8 are essential for lytic DNA replication and have been shown to form pre-replication complex with the core proteins following recruitment to the oriLyt region (9, 14).

In addition to the viral proteins, many cellular proteins were identified to bind to the oriLyt region, including (Topo) I and II, MSH2/6, RecQL, DNA-PK, poly(ADP-ribose) polymerase 1 (PARP-1), Ku autoantigens and SAF-A (scaffold attachment factor A). These proteins are involved in a number of cellular processes such as replication, recombination and repair and are shown to be present in the viral replication foci during lytic replication.

G-quadruplexes or G4 structures are thermodynamically stable non-canonical structures formed in DNA/RNA sequences carrying four runs of at least three guanines separated by few bases (15). The building block of a G-quadruplexes is a guanine tetrad, which is formed by hoogsteen hydrogen bonding between four guanines giving rise to a planar structure that stack on top of each other through π-π interactions. The presence of monovalent cations such as K+ ions and Na+ ions reduces repulsion between negatively charged oxygen atoms thereby promoting the formation of G-quadruplexes (16). These structures have been identified in eukaryotes, prokaryotes and viruses and implicated in regulation of various biological processes including replication, transcription and translation (17). The presence of G-quadruplexes has been reported in a number of human viruses including Human Immunodeficiency Virus (HIV), Hepatitis B virus, Hepatitis C virus, Ebola Virus, Zika virus, Epstein Barr Virus (EBV) and KSHV (18, 19). Recently, G-quadruplexes have gained importance as regulators of replication initiation from the origins (20). These structures have been identified in approximately 90% of the origins in human and mouse and about 70% of the origins in Drosophila and replication initiation sites are
reported to be few hundred base pairs downstream of G4 sites (21). G-quadruplex forming regions have been identified as potential factors determining the location of replication initiation and their presence have been demonstrated in early replication sites (22, 23). Moreover, association of G-quadruplexes with protein involved in their unwinding or resolving negatively supercoiled duplex DNA, has led to the hypothesis that unfolding of these secondary structures located in the origin facilitates the recruitment of many replication factors such as ORCs, which have been shown to have high affinity for these G4 sites (22, 24). The presence of G-quadruplexes in the genome could be detrimental, as they serve as roadblocks to the movement of cellular enzymes involved in replication or transcription. Therefore, the cells employ ways to resolve these structures to ensure precise functioning of the biological processes.

Helicases are cellular motor proteins that separate two annealed strands of nucleic acids using energy derived from nucleoside triphosphate hydrolysis (25-28). Their role in unwinding G4 structures was initially reported in late 1990s, where SV-40 helicase, Bloom syndrome helicase and Saccharomyces cerevisiae proteins, Sgs1p helicase and SEP1 protein were reported to resolve the G4 structures (29-32). RecQ belongs to a conserved family of DNA helicases that are called guardians of the genome due to their roles in DNA replication, recombination and repair (33, 34). The human RecQ helicases include RECQ1, WRN, BLM, RECQ4, and RECQ5 and Bloom, Werner and Rothmund-Thompson syndrome and are associated with mutations in BLM, WRN and RecQ4, respectively (35). BLM helicase encoded by the BLM gene has been shown to unwind the G4 structures and its G4 unwinding activity has been implicated in promoting telomere replication and genome integrity (36-39). Another member of the RecQ family, WRN helicase, has been shown to reduce genome instability due to its ability to resolve (CGG) repeats associated with fragile X syndrome in addition to promoting efficient replication at the telomeric
sequences (40-42). Though there has not been any direct evidence of a helicase unwinding activity on G4s in the KSHV oriLyt, a recent report showed that RecQL4, a member of the RecQ family important for replication initiation and fork progression, binds to the G4 structures (43, 44).

RecQ1, another member of the RecQ family, is the most abundant RecQ protein in the cells and was upregulated in EBV transformed B cells (45, 46). RecQ1 has been shown to repress chromosome instability, reduce DNA damage and promote initiation at the paused replication forks (47-49). In addition to being linked to replication initiation at the origin, RECQ1 was reported to bind to the oriLyt region of KSHV and EBV (50, 51). Our lab had previously reported the formation of G-quadruplexes throughout the genome and its impact on latent DNA replication prompted us to determine the role of oriLyt G-quadruplex on lytic DNA replication (52). In this study we demonstrated the formation of these secondary structures in the oriLyt region through biophysical and biochemical assays. Moreover, we investigated the binding of RecQ1 to the G4 forming region of the oriLyt and substantiate the indispensable role of RecQ1 in KSHV lytic replication through inhibition of RecQ1 helicase activity and depletion of RecQ1. Additionally, we evaluated the relevance of G-quadruplex formation in lytic replication through their disruption by site-direction mutagenesis of the G-residues. We also used ligand to stabilize the formation of these structures to determine its effect on initiation of lytic replication through SMARD.

MATERIALS AND METHODS

Cell lines, plasmids and antibodies

The KSHV-positive BCBL-1/BC3 cell line was grown in RPMI 1640 supplemented with 10% FBS, 2mM glutamine, 5U/ml penicillin and 5ug/ml streptomycin. BJAB-L-YFP cell line
was grown in RPMI 1640 supplemented with 10% FBS, 2mM glutamine, 5U/ml penicillin and 5µg/ml streptomycin and 0.5µg/ml puromycin. iSLKtet-RTABAC16-WT, and iSLKtet-RTA-Bac16-RTASTOP cells were maintained in DMEM supplemented with 10% Tet-Free Fetal Bovine Serum with additional 600µg/mL hygromycin B, 400µg/mL G418, 1µg/mL puromycin. iSLK.219 cells with recombinant KSHV BACs were induced by doxycycline. The following commercial antibodies were used for this study: rabbit anti-RecQ1 (Bethyl laboratories, Inc.) mouse anti-GAPDH (US Biological), rabbit anti-RTA (custom synthesized at GenScript, Inc.). Site directed mutagenesis was performed in wild type 8088sc (8088wt) plasmid carrying the KSHV oriLyt-L sequence using QuikChange Lightning Site-Directed mutagenesis kit according to the manufacturer’s protocol. The integrity of the resulting mutated plasmid i.e 8088mut was confirmed by sequencing at the Nevada Genomics Center, University of Nevada, Reno, NV.

CD spectroscopy

CD spectroscopy was performed on oriLyt wild type and scrambled oligos using Aviv Biomedical spectrometer, where circularly polarized UV light was used to record CD spectra in a series of progressive scans from 320nm to 200nm in a quartz cuvette of 1mm path length at 25°C. The wild type and scrambled oriLyt DNA oligonucleotides were used at a final concentration of 5µM in sodium cacodylate buffer (10mM, pH 7.4) with 100mM KCl. G-quadruplexes were allowed to form in the oligos by initial denaturation at 95°C for 5 minutes followed by slow cooling at room temperature.

Electrophoretic mobility shift assay

Electrophoretic mobility shift assay (EMSA) was performed as described before (52). Briefly, wild type oriLyt and scrambled oriLyt DNA oligonucleotides were labeled with radioactive ³²P using Terminal deoxynucleotidyl transferase, TdT (New England Biolabs) and
resuspended at a final concentration of 2µM in 10mM sodium cacodylate buffer with 100mM KCl. Following denaturation and slow cooling, the oligos were resolved on a 15% native polyacrylamide gel in 1X TBE buffer with 100mM KCl, gels were dried using Gelair gel dryer (Bio-Rad Inc.) and autoradiography was performed using a phosphorimager (GE Healthcare Life Sciences). For G4 disruption, antisense oligos were incubated with the labeled oligos before denaturation.

**In vitro DNA pull-down assay**

Wild type oriLyt oligo (5’-CACGGGGTTCGGTGCGGGGGGGGCTAG-3’) and scrambled oligo (5’-ACTATATTGTTCAATAACTATATTATAAATCT-3’) were biotinylated at the 3’end Terminal deoxynucleotide transferase, TdT (New England Biolabs). The cellular lysate was prepared from approximately 20 million KSHV-positive BCBL-1 cells, which were lytically induced for 24 hours using sodium butyrate (NaB) and 12-o-tetradecanoylphorbol-13-acetate (TPA), harvested and lysed in 1% NP-40 lysis buffer with protease inhibitors. The samples were sonicated and centrifuged at 4°C to remove cellular debris. The purified biotinylated DNA oligos were incubated with the lysate for three hours at 4°C and Pierce Streptavidin agarose beads (Thermo Fisher Scientific) were added to the samples for two hours to pull down the proteins bound to the biotinylated DNA. The beads were washed thrice with 1% NP-40 lysis buffer following which they were loaded onto a 9% SDS-PAGE, transferred on nitrocellulose membrane and probed with RecQ1 antibody.

**Chromatin immunoprecipitation assay (ChIP)**

Chromatin immunoprecipitation was performed using iDeal ChIP qPCR kit (Diagenode) according to the manufacturers instructions. Briefly, KSHV positive cells were induced for lytic reactivation for 24 hours following which approximately 8 million cells were fixed with 1%
formaldehyde for 10 min at room temperature and cross-linking was stopped by the addition of glycine at a final concentration of 125 mM for 5 min. The cells were washed two times with ice-cold PBS and nuclei were extracted following cell lysis using kit supplied buffers. Cells were finally lysed in chromatin shearing buffer supplemented with protease inhibitors for 10 min on ice and chromatin was sonicated using a Bioruptor (Diagenode) to an average length of 200 to 300 bp. The lysates were centrifuged at 16,000g for 10 minutes to remove the cell debris and the supernatant was incubated with specific antibody and magnetic beads overnight at 4°C. Following day, the beads were collected and subsequently washed at 4°C for 5 minutes each with wash buffers. The beads bound DNA was subjected to Proteinase K digestion and the DNA was finally eluted from the beads by incubation at 100°C for 15 minutes. Purified ChIP DNA samples and the inputs were subjected to amplification with specific primers using ABI StepOne plus real-time PCR machine (Applied Biosystems).

**Transfection of oriLyt plasmids and transient replication assay**

Approximately 20 million BCBL-1 cells were transfected with 30µg of oriLyt plasmids (either 8088wt or 8088mut) using Neon transfection system (Thermo Fisher Scientific). The cells were allowed to recover for 24 hours and induced for lytic reactivation (24 hours for Chromatin Immunoprecipitation assay and 48 hours for replication assay). For replication assay, cells were washed with phosphate-buffered saline followed by extraction of DNA using a modified Hirt’s lysis method, described earlier (53, 54). A fraction of DNA was linearized with EcoRI and the remainder with DpnI and EcoRI to remove the non-replicated DNA. Digested DNA was resolved and transferred to a Nylon membrane followed hybridization with 32P-labeled 8088sc probes. The auto-radiographic signals were detected using a PhosphorImager, according to the manufacturer’s instructions (Molecular Dynamics, Inc.).
Single Molecule Analysis of the Replicated DNA (SMARD)

Single Molecule Analysis of Replicated DNA (SMARD) was used to analyze the effect of G-quadruplex formation on the initiation of lytic DNA replication as described previously (55) (56). BCBL-1 cells were treated with G4 stabilizing compound PhenDC-3 and TMPyP4 at a concentration of 10μM and induced lytically for 24 hours. The cells were sequentially labeled with 30 μM 5-iodo-2’-deoxyuridine (IdU) (Sigma-Aldrich) at 37°C for 4 h, washed with PBS and labeled with second nucleotide analogue, 5’-chboro-2’-deoxyuridine (CldU) at 30 μM (Sigma-Aldrich) for 4 h. The cells were finally washed and resuspended in PBS (1 × 10^6 cells per ml) and molten 1% InCert agarose (Lonza Rockland, Inc., Rockland, ME, USA) in PBS (1:1 vol/vol). Agarose gel plugs embedded with labeled cells were made by solidification in a cold plastic mold on ice for 30 min. The plugs were incubated in lysis buffer (1% n-lauroylsarcosine, 0.5 M EDTA and 20 mg/ml proteinase K) for at least 72 h at 50°C, following which the proteinase K was removed by washing the plugs with 1X TE and phenylmethanesulfonyl fluoride (PMSF) (Sigma-Aldrich). Next, the KSHV genome was linearized using Pmel (New England Biolabs Inc.) where the plugs were first washed with pre-digestion buffer (10 mM MgCl₂ and 10 mM Tris-HCl (pH 8.0) before digesting with 70 units of Pmel at 37°C overnight. The Pmel digested gel plugs were rinsed twice with TE, casted onto a 0.7% SeaPlaque GTG agarose gel (Lonza Rockland, Inc.) and resolved by PFGE for 36 hours. KSHV genome was detected using Southern blotting and hybridization with a ^32P-labeled KSHV TR- specific probe. Band specific for KSHV genome was excised from the gel and DNA was extracted using GELase treatment (Epicentre Biotechnologies, Madison, Wisconsin, 1 unit per 50 μl of agarose suspension) that digests the agarose releasing the DNA into suspension. Following this, the DNA was stretched on 3-aminopropyltriethoxysilane (Sigma-Aldrich) coated glass slides and
denatured in alkaline denaturing buffer (0.1N NaOH/70% Ethanol, and 0.1% β-mercaptoethanol) and fixed with 0.5% glutaraldehyde. The DNA was hybridized overnight with biotinylated probes and the next day, slides were rinsed in 2x SSC (1x SSC) 1% SDS, washed in 40% formamide solution containing 2x SSC at 45°C for 5 min and rinsed in 2x SSC-0.1% IGEPAL CA-630. The slides were washed 4 times with 4x SSC-0.1% IGEPAL CA-630, followed by blocking with 1% BSA for 20 min. Next, the slides were treated with NeutrAvidin Alexa Fluor 350 (Life Technologies Inc.) and biotinylated anti-avidin antibodies (Vector Laboratories, Inc.) for 20 min each, after washing with PBS containing 0.03% IGEPAL CA-630 in between. Slides were further treated with NeutrAvidin Alexa Fluor 350 for 20 min and washed as mentioned above. Following this, the slides were incubated with mouse anti-IdU monoclonal antibody, mouse anti-CldU monoclonal antibody and biotinylated anti-avidin D for 1 h. After washing, the slides were incubated with NeutrAvidin Alexa Fluor 350 and fluorescent secondary antibodies against mouse and rabbit i.e Alexa Fluor 488 and Alexa Fluor 594 (Invitrogen Molecular Probes) for 1 h. The slides were washed again and coverslips were mounted with ProLong gold anti-fade reagent (Life Technologies Inc.), following which fluorescence microscopy was performed.

**RecQ1 knockdown using lentiviral vectors**

The pTRIPZ lentiviral vector (Dharmacon, GE Life Sciences) containing shRNA for RecQ1 was co-transfected with lentivirus packaging vectors, pCMV-dR8.2 and pCMV-VSVG (Addgene, Inc.) into Lenti-X 293T (Takara Bio) cells using polyethylenimine (PEI) (Polysciences, Inc.) to generate the lentiviral particles. Lentivirus containing supernatants from the transfected cells were collected for 5 days, followed by concentration of the virus by ultracentrifugation (25,000 rpm, 1.5 hr, 4°C). The concentrated virus was used for transducing the target cells, BCBL-1 in the presence of 5µg/ml polybrene followed by selection with 1 µg/ml
puromycin. The cells were treated with 1µg/ml doxycycline for at least 72 hours for the induction of knock down. The RNA interference (RNAi) efficiency was assessed by Western blot analysis using RecQ1 antibody.

**IdU labeling and immunoprecipitation of replicated DNA**

IdU labeling was performed as described previously (54). Briefly, cells were pulsed with 30 µM of IdU for 30 min, washed twice with cold PBS and the episomal DNA was extracted by a modified Hirt’s method. The samples were dissolved in 500 µl TE (10 mM Tris-HCl, 1 mM EDTA), sonicated to get an average length of 700 bp and heat denatured at 95 °C for 5 min. 10% of the extracted DNA was used as input control and 1 µg of mouse of anti-IdU antibody was added to samples and incubated at room temperature with constant rotation for 1 h. Magnetic Protein A/G Antibody were used to pull down the bound IdU labeled DNA and the beads were washed once with 1X IP buffer (10 mm NaPO4 pH 7.0, 140 mM NaCl and 0.05% Triton X-100), resuspended in 200 µl of lysis buffer (50 mM Tris-HCl (pH8.0), 10 mM EDTA, 0.5% SDS, 0.25 mg/ml Proteinase K) and incubated overnight at 37 °C for elution. This was followed by addition of 100 µl of lysis buffer and incubation at 50 °C for 1 h. The eluted DNA was phenolized and precipitated for the quantitation of IdU labeled DNA in a real-time PCR by amplifying the oriLyt region.

**KSHV Virion purification**

KSHV virions were purified as described previously (57). Briefly, 15 million cells were induced with 0.3M NaB and 20 ng/mL TPA for 96 hours, culture supernatant was cleared by centrifugation, and filtered through a 0.45-µm filter to remove cell debris. Virus was concentrated by ultracentrifugation (25,000 rpm for 2h at 4°C), resuspended in serum free RPMI and 50 µl of virus supernatant diluted with 250 µl of 1X PBS was digested with DNase-I (5 U/50
µl of supernatant) at 37°C for 1 h. DNase-I was heat inactivated at 70°C for 10 min, supernatants were mixed with equal volume of lysis buffer (0.1 mg/ml of proteinase K in water) and incubated at 50°C for 1 h. Proteinase K was heat-inactivated at 75°C for 20 min, DNA was purified using PCI (Phenol: Chloroform: Isoamyl alcohol), precipitated with ethanol at -20°C and resuspended in sterile milliQ water. The viral DNA was amplified with qPCR using primers specific for LANA and virus quantities were calculated based on the LANA standards. All the reactions were run in triplicates.

**Quantitation of KSHV RNA**

Total RNA was isolated from the cells using Illustra RNAspin Mini Kit (GE Healthcare) according to the manufacturer’s protocol. cDNAs were generated using high-capacity RNA-to-cDNA kit (Applied Biosystems Inc.) as per the manufacturer’s protocol and quantified using specific primers. The Ct values were normalized to the housekeeping gene, GAPDH. All the reactions were run in triplicates.

**Determination of viral genome copies**

Five million KSHV positive cells were induced for lytic reactivation for 24 hours with or without compound treatment. Total genomic DNA was isolated using PureLink™ Genomic DNA Purification Kit according to the manufacturer’s instructions and quantified using primers specific for the DNA sequence of RTA. The Ct values were normalized to the housekeeping gene, GAPDH. All the reactions were run in triplicates.

**Table 1: List of primers used in the study**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF8 Forward</td>
<td>ACGGTCCAGTACTTCCACACAA</td>
</tr>
<tr>
<td>ORF8 Reverse</td>
<td>TCTCTTTCTGCGGCTGGCA</td>
</tr>
<tr>
<td>ORF59 Forward</td>
<td>TGTGTAAGTCCCGGTTGGTT</td>
</tr>
<tr>
<td>ORF59 Reverse</td>
<td>TCCGGTATAGAATCGGGAACCT</td>
</tr>
<tr>
<td>ORF65 Forward</td>
<td>TGGATCATGACTACGCTCACCA</td>
</tr>
<tr>
<td>ORF65 Reverse</td>
<td>CCATCCTCCTCAGATAGGCCTCATAA</td>
</tr>
<tr>
<td>ORF72 Forward</td>
<td>AGCTCAGAAGCCTCAGCCTATTT</td>
</tr>
<tr>
<td>ORF72 Reverse</td>
<td>AAGTGACGTCCGTCGCTAAGACT</td>
</tr>
<tr>
<td>GAPDH Forward</td>
<td>CAGCAAGAGCACAAGAGGAAGA</td>
</tr>
<tr>
<td>GAPDH Reverse</td>
<td>TTGATGGTACATGACAAGGTGCGG</td>
</tr>
<tr>
<td>oriLy-L Forward</td>
<td>GCTTCCAGCGACGGAAAATAAG</td>
</tr>
<tr>
<td>oriLy-L Reverse</td>
<td>ACGGAGAATAGACCAGCCTGAA</td>
</tr>
<tr>
<td>oriLyt R1 Forward</td>
<td>CCCCCAACTGTATTCAACCCTCCT</td>
</tr>
<tr>
<td>oriLyt R1 Reverse</td>
<td>CACCCAGAGCGAGATCCAGGC</td>
</tr>
<tr>
<td>oriLyt R2 Forward</td>
<td>CCACCGCTGGCGCCCGCCCAGAA</td>
</tr>
<tr>
<td>oriLyt R2 Reverse</td>
<td>TCGTTACGGGTAATCCCAAGAG</td>
</tr>
<tr>
<td>oriLyt R-Os Forward</td>
<td>CTTTAATCCCTAGGCAGCTGAG</td>
</tr>
<tr>
<td>oriLyt R-Os Reverse</td>
<td>GGTATTCCCAGCTGGTTGC</td>
</tr>
<tr>
<td>Genomic 50(RTA)-Forward</td>
<td>AAGGTTCGACTGGACAACGTCA</td>
</tr>
<tr>
<td>Genomic 50(RTA)-Reverse</td>
<td>CCTGGACGTGGTGCGTGATAAA</td>
</tr>
</tbody>
</table>

**Statistical analysis**

P-values were calculated by two-tailed t test using GraphPad Inc. (Prism 8) software for statistical significance. Asterisks represent the P value < 0.05 (*), P value < 0.01 (**) and <0.001 (***) , where ns denotes non-significant

**RESULTS**

**oriLyt DNA formed stable G4 structures**

DNA G-quadruplexes are secondary structures formed in G-rich stretches of DNA that have been implicated in a number of biological processes. These structures have been reported in the latent nuclear protein of EBV, terminal repeat region of KSHV, viral gene promoters as well as in the oriLyt region of Human Cytomegalovirus (HCMV) (52, 58, 59). Given the important role of the G4 structures in viral life cycle, we were interested in investigating their role on lytic DNA replication of KSHV. Firstly, we analyzed the oriLyt sequence of KSHV for the presence
of G-quadruplex using a web-based tool, QGRS mapper that predicts the formation of these structures (http://bioinformatics.ramapo.edu/QGRS/index.php)(60). G-scores, predictor of the G-quadruplexes formation, were determined for the oriLyt region of KSHV and the region with the high G-scores are represented by yellow highlighted region (Fig. 1A). The oriLyt sequence with the highest G-score was located between the AT rich region and the RRE element. Since, G4
structures are unique with respect to their structure and folding (Fig. 1B), they display distinct biophysical and biochemical properties.

We validated the formation of G-quadruplexes by performing CD spectroscopy of an oligo containing the wild type oriLyt G4 site or a scrambled oligo. Oligo with specific G4 site showed a spectral pattern with a maximum at 260 nm and a negative minimum at 240 nm for
Fig. 2. Biophysical and biochemical analysis confirmed the presence of G-quadruplexes in the oriLyt region.
A. Circular Dichroism spectral analysis of (a) oriLyt wild type oligonucleotide having a high G-score and (b) Scrambled DNA oligonucleotide which did not form G4s was used as a negative control. The oligos were scanned for a wavelength range of 320 nm to 200 nm and molar ellipticity was plotted on the Y-axis with wavelength on the X-axis.
B. (a) Schematic of step involved in Native PAGE analysis of wild type oriLyt or scrambled oriLyt. (b) Electrophoretic mobility shift assay (EMSA) performed in the presence of K+ ions on wild type oriLyt and scrambled oligos labeled with $^{32}$P and resolved on a native polyacrylamide gel. Antisense oligos (AS1: GCCACCGAACAACCCC and AS2: CACTAGCCCCCCC) were complementary to wild type oriLyt oligo, added in molar excess, were used in the indicated lanes to confirm that the mobility shift was caused by G-quadruplex forming sequence.

wild type oriLyt oligo (Fig. 2A panel a). This pattern is specific for G-quadruplexes and was not observed for scrambled oligo (Fig. 2A panel b). We further confirmed the G-quadruplex formation on the oriLyt oligo in an electrophoretic mobility shift assay. The wild type and scrambled oriLyt DNA oligos were labeled with $^{32}$P using Terminal deoxynucleotidyl transferase and resolved in the presence of K+ ions on a 15% native gel following which autoradiography was used to determine the mobility of these oligos (Fig. 2A panel a). Our results showed that due to the formation of condensed G-quadruplex structures, wild type oriLyt G4 forming DNA oligo migrated faster than the scrambled oligo (Fig. 2A panel b). To corroborate that the increased mobility of the wild type oriLyt oligo was caused by the formation of these secondary structures, we used antisense oligos to disrupt the formation of G-quadruplexes. The antisense oligos (AS1: GCCACCGAACAACCCC and AS2: CACTAGCCCCCCC) were complementary to the G-rich region of wild type oriLyt oligo. Following the incubation of these antisense oligos with G4 oriLyt oligo and resolution on a native gel, we observed a shift in the mobility of wild type oriLyt oligo in the presence of specific antisense oligos as compared to the wild type oriLyt oligo alone (Fig. 2A panel b, compare lane 3 with lane 1). Based on these findings, we concluded that the fast migration of oriLyt G4 oligos on native gel is due to the formation of G-quadruplexes as
the disruption of these structures by antisense oligos led to a significant shift in the mobility of the wild type oligo. These results validated that oriLyt forms G-quadruplexes.

**RecQ1 bound to G4 forming site on oriLyt**

In light of the role of G-quadruplexes in cellular metabolism, it can be assumed that G4 structures interact with a variety of cellular proteins that either support the formation or destabilize these structures, which is an area of extensive study (17). Interestingly, a study identified the involvement of RecQ1, a cellular helicase, in supporting KSHV lytic DNA replication by binding to the oriLyt region, an interaction facilitated by other viral proteins. RecQ family of proteins has been shown to facilitate DNA replication through their G4 resolving ability (61). This led us to investigate the G4 binding ability of RecQ1 with respect to oriLyt DNA. To achieve this, we first performed a DNA affinity pull down assay, where wild type oriLyt oligo carrying a G4 site and a scrambled oligo were biotinylated at the 3’ end using Biotin-11-UTP and Terminal deoxynucleotidyl transferase (TdT) followed by purification of biotinylated DNA and incubation with cellular extract from KSHV positive BCBL-1 induced for 36 hours. After incubating the biotinylated DNA and the cell lysates together for three hours, streptavidin beads were added to capture the proteins associated with the DNA (Fig. 3A panel a). To eliminate the non-specifically bound proteins, beads were washed stringently, following which the proteins were resolved on SDS-PAGE. Upon detection using RecQ1 antibody, we observed that RecQ1 specifically pulled down with wild type oriLyt oligo and not the scrambled oligo (Fig. 3A panel b).

Next, we wanted to validate our findings *in vivo*, so we performed Chromatin immunoprecipitation assay using RecQ1 antibody on BCBL-1 cells that had been induced for lytic reactivation for 36 hours. Three sets of primers were used to quantify relative
A.  

- Wild type oriLyt

  - Biotinylation using Biotin-11-dUTP
  - Denaturation, slow cooling in the presence of K⁺ ions
  - Incubate together
  - Streptavidin beads
  - Resolve proteins on SDS-PAGE

B.  

- C/EBP (1-8) - AT palindrome

  - RRE - TATA

  - oriLyt transcript

  - R1 - R2

  - RecQ1 ChIP

  - % Input

  - R1
  - R2
  - R-Os
RecQ1 enrichment of ChIP DNA; R1 specific for region that carries a G4 site, R2 specific for region that was previously identified to bind oriLyt but does not possess a G4 site and R-Os, specific for a region outside the RecQ binding area of oriLyt (50) (Fig. 3B panel a). Upon comparison, we observed that RecQ1 selectively bound to R1 more than R2 and R-Os. Collectively, these findings revealed binding of RecQ1 to the G4 site in oriLyt of KSHV (Fig. 3B panel b).

**RTA recruited RecQ1 to the oriLyt region**

RTA plays an indispensable role in lytic DNA replication and interacts with other viral proteins to perform a variety of functions in replication initiation. RTA binds at the RTA Response Elements (RRE) in the oriLyt region and activates the promoter of oriLyt to initiate transcription. In addition, it has been shown to interact with proteins of the prereplication complex and recruits them to the origin to facilitate lytic DNA replication (14). There has been evidence that RecQ1 binds to RTA directly and its binding to the oriLyt region is compromised in the absence of the RRE domain. We wanted to investigate the binding of RTA and RecQ1 with respect to G4 forming sites in the oriLyt. To achieve this, we performed a chromatin immunoprecipitation assay with anti-RTA antibody on 24 hour induced BCBL-1 cells and quantified RTA enrichments using the three primers indicated on the schematic (Fig. 3B panel...
a). Not surprisingly, we found that RTA enriched the region with RRE but so much the region lacking G4 site (Fig. 4A panel a). Next, we analyzed the relative enrichment of RecQ1 at these regions in the absence of RTA. To this end, we performed a chromatin immunoprecipitation assay with RecQ1 in iSLKTet-RTABAC16-WT and iSLKTet-RTA-Bac16-RTASTOP cells. We found that a relative binding of RecQ1 at R1 decreased in cells lacking RTA as compared to the wild type cells (Fig. 4A panel b).

---

**Fig. 4. RTA facilitated the binding of RecQ1 at oriLyt**

A. (a) ChIP assay was performed using lytically induced BCBL-1 cells, where the cells were harvested, crosslinked and chromatin bound DNA was pulled down using RTA antibody to determine the relative enrichment of RTA bound regions of oriLyt using specific primers (R1, R2 and R-Os) (b) ChIP assay was performed using lytically induced iSLK-Tet-RTABAC16-WT and iSLKTet-RTA-Bac16-RTASTOP cells where the cells were harvested, crosslinked and chromatin bound DNA was pulled down using RecQ1 antibody to determine the relative binding of RecQ1 on oriLyt using primers specific for R1.
RecQ1 was required for lytic DNA replication

Given the presence of RecQ1 at oriLyt and in the viral replication compartments, we hypothesized RecQ1 to play an important role in lytic DNA replication. To this end, we blocked RecQ1’s activity by treating the KSHV positive cells with N-methyl mesoporphyrin IX (NMM), an inhibitor of RecQ1 helicase activity, and assayed lytic DNA replication.

First, we analyzed the amount on newly replicated DNA in treated cells through IdU immunoprecipitation assay, where KSHV positive cells were treated with NMM or DMSO and induced for lytic reactivation for 24 hours before labeling the cells with IdU and purification of newly replicated nascent DNA with anti-IdU antibody. The amounts of newly replicated DNA were quantified using oriLyt specific primers. Upon comparison, we observed that the amount of newly replicated DNA was significantly reduced in cells treated with NMM as compared to control cells treated with DMSO (Fig. 5A panel a). Since RecQ1 inhibition was negatively affecting active lytic replication, we wanted to analyze its effect on viral lytic gene expression and genome copies. This was tested by quantifying the transcripts of immediate early and late genes i.e ORF59 and ORF 8, respectively. NMM treated cells showed a reduction in late gene transcription without affecting the transcription of latent gene (Fig. 5A, panel b, c and d). We also determined the copies of replicated genome following NMM treatment, which was significantly reduced in BCBL-1 cells treated with NMM as compared to the control, DMSO treated (Fig. 5A panel e). Having established that blocking RecQ1 activity inhibited lytic DNA replication, we wanted to further confirm whether that led to a reduction in virion production. As expected, we observed a decrease in the number of virions produced from BCBL-1 cells treated with NMM as compared to the control cells (Fig. 5A panel f). Similarly, we also confirmed the importance of RecQ1 in lytic DNA replication through depletion of RecQ1 through shRNA.
Fig. 5. RecQ1 was essential for lytic DNA replication

A. BCBL-1 cells were treated with DMSO or 2μM NMM and induced for lytic reactivation (a) 24 hours post lytic induction, cells were labelled with IdU, genomic DNA was extracted, newly replicated DNA was pulled down anti-IdU antibody and the DNA was quantified using primers specific for oriLyt region. (b) 24 hours post lytic induction, mRNA was extracted, cDNA was synthesized and quantified using latent gene, ORF72 specific primers (c) 24 hours post lytic induction, mRNA was extracted, cDNA was synthesized and quantified using immediate. early lytic gene, ORF59 specific primers. (d) 48 hours post lytic induction, mRNA was extracted, cDNA was synthesized and quantified using late lytic gene, ORF8 specific primers. (e) 24 hours post lytic induction, genomic DNA was extracted and viral genome copies were quantified using primers specific for genomic DNA sequence of RTA. (f) 96 hours post lytic induction, virus was concentrated from the supernatant of induced cells and quantified using ORF73 primers.
BCBL-1 cells stably transduced with inducible shRecQ1 lentivirus were assayed for the depletion of RecQ1 post doxycycline treatment, which showed a significant reduction in RecQ1 levels (shRecQ1-Kd) as compared to control cells (shRecQ1-C) (Fig. 5B panel a). Quantification of viral genome copies in shRecQ1-C and shRecQ1-Kd BCBL-1 cells following lytic reactivation showed significantly reduced copies in cells depleted RecQ1 (Fig. 5B panel b).

**G4 disruption inhibited RecQ1 binding and lytic replication**

The presence of G4 sites in oriLyt as well as binding of RecQ1 to these sites, prompted us to evaluate the functional relevance of these sites in the genome. G-quadruplex formation is highly dependent on the occurrence of four repeats of at least three G- residues separated by bases. We performed site-directed mutagenesis in oriLyt on a plasmid to disrupt the G4 sites (8088sc), by changing the G-rich sequence from

\[
\text{ACGGGGTTTGTTCGGTGGC} \underline{GGGG} \underline{GGGGGGGGG} \quad (8088\text{wt})
\]

to

\[
\text{ACGGGGTTTGTTCGGTGGCAATAAGGGGGGGG} \quad (8088\text{mut})
\]

This reduced the propensity of stable G-quadruplex formation indicated by a drop in G-score. We tested these clones for their efficiency to undergo lytic DNA replication by performing transient replication assay by transfecting BCBL-1 cells with 8088wt or 8088mut plasmids and inducing the cells for lytic reactivation for 48 hours. DNA was extracted from these cells and digested with EcoRI to linearize or with DpnI and EcoRI to determine the replicated copies after Southern hybridization. Upon quantifying DpnI resistant/replicated DNA band, we observed a lower lytic DNA replication of plasmid with mutated G4 site (8088mut) as compared to the cells transfected with 8088wt plasmid (Fig. 6B panel a).
A.

1. Transfection with 8088sc
2. Induction with TPA/NaB for 24 hours

B.

**BCBL-1**

- **a.**
  - EcoRI
  - EcoRI+Dpnl

**b.**

Relative RecQ1 binding to orLytRI

- 8088 wt
- 8088 mut

**Significance:** ****
In our previous experiments, we showed that RecQ1 specifically binds to oriLyt G4 site as compared to other regions in oriLyt. Since, mutation of G4 sites deterred the formation of G4 structures (determined by the G-scores), we wanted to test the effect of G4 disruption on RecQ1 binding. We achieved this by performing Chromatin Immunoprecipitation assay on BCBL-1 cells transfected with 8088wt or 8088mut using RecQ1 antibody. BCBL-1 cells were transfected with 8088wt or 8088mut plasmids and 24 hours post transfection, the cells were induced for lytic reactivation for 24 hours. DNA extracted from the RecQ1 ChIP was amplified using primers specific for G4 region (R1). We observed a reduced binding of RecQ1 to R1 in cells transfected with 8088mut as compared to cells transfected with 8088wt, indicating that RecQ1 selectively binds to the G4 sites and this binding is compromised by the disruption of G4 sites (Fig. 6B panel b).

**G-quadruplex stabilization reduced initiation of lytic replication**

G-quadruplex are naturally occurring secondary structures in nucleic acids, which are destabilized by the cellular proteins. However, there are many G-quadruplex ligands, which can bind and selectively stabilize these G-quadruplexes. G4 sites have been associated with replication initiation and a recent report clearly establishes a link between G4 sites and origin
activity in mammalian genome (61, 62). This prompted us to study the effect of G-quadruplex stabilization on lytic DNA replication in KSHV.

Although the detailed mechanism of herpesviral genome replication is not fully understood but the initiation is proposed to initiate through bidirectional theta-type replication (63). We analyzed the effect of G-quadruplex stabilizing ligands on replication initiation through Single Molecule Analysis of the Replicated DNA (SMARD) assay, a technique which has been used for studying the replication events in EBV and KSHV (64, 65). This involves multiple steps including sequential labeling of cells with halogenated nucleotides, IdU and CldU, embedding the cells in agarose plugs, lysis of cells and digestion and linearization of DNA. The plugs containing DNA were resolved on Pulse field Gel electrophoresis and the KSHV specific band was excised to isolate linear KSHV genome following gelase digestion. This was followed by stretching theDNA on a positively charged glass slides and fluorescent in-situ hybridization (FISH) with KSHV specific probes. The labeled DNA was finally detected by monoclonal antibodies against IdU and CldU and fluorescently labeled secondary antibodies, whereas the biotinylated probes were detected by fluorescence conjugated avidin. The replication initiation sites are determined by the detection of IdU (red) label flanked by the CldU (green) label on both sides. DNA molecules were arranged based on the increasing length of the IdU label from the oriLyt-L. The molecules signifying a bidirectional initiation site showed a progressively increasing red signal surrounded by the green signal, which indicated bidirectional movement of the replication fork. Upon analysis of labeled DNA molecules from DMSO or TMPyP4 treated BCBL-1 cells, we observed a reduction in the number of molecules with initiation at oriLyt-L characterized by red signal flanked by green signal (Fig. 7). Thus, we concluded that G-quadruplex stabilization leads to reduction in lytic DNA replication initiation.
Fig. 7. G-quadruplex stabilization reduced initiation of lytic replication at oriLyt-L.

BCBL-1 cells were treated with DMSO or TMPyP4, induced for lytic reactivation for 24 hours and SMARD was performed on them. KSHV genome was detected by blue FISH probe signals Red and green labels represented IdU (first label) and CldU (second label) labels, respectively following detection by immunostaining. Yellow arrows indicate replication forks originating from the origin i.e oriLyt L and green labels flanking red showed bidirectional movement of the replication fork.
G-quadruplex stabilization inhibited lytic DNA replication

To evaluate the role of G4 stabilization on virus lytic replication, we used two G-quadruplex stabilizing compounds PhenDC-3 and TMPyP4 to treat BCBL-1 cells along with DMSO as a negative control and PAA, a known inhibitor of lytic replication as a positive control.

G4 stabilization inhibited initiation of lytic DNA replication, which is important for transcription of late lytic genes. Therefore, we wanted to analyze the effect of G-quadruplex formation on lytic gene transcription. This was tested by quantifying the transcripts of lytic immediate early and late genes i.e ORF59 and ORF65, respectively. Late gene expression was significantly reduced after the treatment of cells with both G4 stabilizing compounds (PhenDC-3 and TMPyP4) and PAA as compared to the control cells with no significant effect on early and latent gene transcription (Fig. 8A compare panel a, b and c). Next, we tested the effects of G4 stabilization on lytic DNA synthesis through quantifying the relative viral genome copies in BCBL-1 cells treated with these compounds or DMSO. Following induction of lytic reactivation for 24 hours, genomic DNA was isolated and viral genome copies were quantified using genome specific primers. Importantly, we observed a reduction in viral genome copies in BCBL-1 cells treated with Phen DC-3 and TMPyP4 along with cells treated with a known inhibitor of lytic DNA replication, PAA in comparison with DMSO treated cells (Fig. 8A panel d). This led us to conclude that stabilization of G-quadruplexes in the cells lead to a reduction in lytic DNA replication. We further went on to assay the amounts of virion produced following the stabilization of G-quadruplexes, which showed a significant reduction in cells treated with PhenDC-3 or TMPyP4 as compared to the DMSO control. Treatment with PAA, as expected,
Fig. 8. G-quadruplex stabilization inhibited lytic DNA replication.

A. BCBL-1 cells were treated with DMSO/10μM Phen-DC3/10μM TMPyP4/0.5mM PAA and induced for lytic reactivation. (a) 24 hours post lytic induction, mRNA was extracted, cDNA was synthesized and quantified using latent gene, ORF72 specific primers (b) 24 hours post lytic induction, mRNA was extracted, cDNA was synthesized and quantified using immediate early lytic gene, ORF59 specific primers (c) 48 hours post lytic induction, mRNA was extracted, cDNA was synthesized and quantified using late lytic gene, ORF65 specific primers. (d) 24 hours post lytic induction, genomic DNA was extracted and viral genome copies were quantified using primers specific for genomic DNA sequence of RTA. (e) 96 hours post lytic induction, virus was concentrated from the supernatant of induced cells and quantified using ORF73 primers.
completely abolished virion production. All these observations confirmed that stabilization of G4 structures in KSHV genome inhibits lytic DNA replication.

**Discussion**

KSHV, a member of the gammaherpesvirus family, maintains a biphasic life cycle consisting of latent and lytic phases. While latency ensures lifetime persistence of the virus in the host, lytic DNA replication is essential to maintain viral reservoirs in the infected cells, dissemination of the virus and promotion of tumorigenesis through the expression of oncogenic gene products (6).

Latent DNA replication occurs in synchrony with the cell cycle and the duplicated genome is segregated into the daughter cells. LANA is responsible for these functions in addition to latent DNA replication where it recruits cellular DNA replication proteins to the origin of latent replication, oriP (66). Latency to lytic reactivation switch can be triggered be various stimuli such as hypoxia, oxidative stress, viral coinfection and chemicals (67). RTA is indispensable for lytic reactivation and is responsible for the switch of viral life cycle from latency to lytic through activation of its promoter and promoters of other viral genes, facilitating viral DNA replication through binding to the origin, acting as a components of the prereplication complex and recruitment of cellular and viral factor to the origin, all of which are required for lytic DNA replication (14). This mode of replication differs from the latent mode in many ways as it originates from a distinct origin of replication, oriLyt, requires functioning of multiple viral proteins and results in the production of thousands of viral genome copies. Apart from the viral proteins, a number of cellular proteins are shown to help in lytic DNA replication through binding to the oriLyt region including topoisomerases (Topo) I and II, RecQ1, poly (ADP-ribose) polymerase I (PARP-1), DNA-PK and scaffold attachment factor A (SAF-A) (50).
G-quadruplexes are regulatory structures formed in G-rich DNA/RNA sequences that have been reported to play significant roles in biological processes. These structures have been identified in viral genomes of many viruses including HIV, EBV, HPV, HCMV and KSHV (59, 68). Owing to a high GC content in their genomes, herpesviruses have high propensity to form G4 structures. Many studies have reported high putative quadruplex sequences (PQS) in herpesviral genomes (18, 59, 69). The regulation of DNA replication by G4 structures has been well studied but their role in origin firing or initiation of replication is relatively unexplored (61, 70, 71). A recent study provides strong evidence regarding the substantial role of G4s in origin activity (62).

RecQ1, a cellular helicase, is a member of the RecQ family of helicases, which are known for their roles in recombination, repair and replication. In addition, other members of the RecQ family such as BLM and WRN helicase are shown to unwind G4 structures and facilitate telomeric replication (61). RecQ1 is shown to promote genome stability through its role in DNA damage response and in facilitating restarting of the paused replication forks (72). Additionally, it has been proposed to have a role in origin activity at lytic DNA replication origin through binding to the oriLyt region (50, 51, 71). We wanted to confirm whether the association of RecQ1 at oriLyt was through the G-quadruplexes and further explore the relevance of this association in KSHV life cycle.

We analyzed the DNA sequence of oriLyt through a web-based tool for determining the potential G-quadruplex site. Our analysis revealed the formation of a strong G-quadruplexes in oriLyt DNA, which was confirmed through CD spectroscopy and EMSA. The involvement of RecQ1 in the prereplication complex and in the viral replication centers through binding at oriLyt has shed light on the origin-based activity of RecQ1. G4 unwinding activity being another
widely studied function of RecQ1 could potentially explain its engagement at oriLyt. As an extension to previous studies on RecQ1 in KSHV, we found RecQ1 to be selectively enriched at the G4 forming site in oriLyt as compared to other regions, which was also confirmed in biotinylated DNA pulldowns and Chromatin Immunoprecipitation assay. Moreover, the role of RTA in RecQ1-oriLyt binding was also substantiated by a reduced pull down of RecQ1 at oriLyt in RTA deleted cell lines. Further evidence on the role of RecQ1 in lytic replication was shown by a reduction in genome copies of the virus in RecQ1 depleted cells. More interestingly, the inhibition of helicase activity of RecQ1 through the use of NMM led to reduction in actively replicated DNA, genome copies, virion production and late gene expression. Additionally, we demonstrated the importance of the G-residues in G4 structures on lytic replication through mutation of these residues in oriLyt, which showed a significant reduction in lytic replication and reduced RecQ1 binding to the oriLyt region.

Importantly, we evaluated the effects of G4 stabilization on lytic replication through the use of compounds like PhenDC-3 and TMPyP4, which stabilizes these G-quadruplexes. Of our particular interest was the observation through SMARD, which demonstrated that stabilization of G-quadruplexes compromised the initiation of replication at oriLyt. These results also corroborated in the subsequent assays where G4 stabilization led to reduction in viral genome copies, late lytic gene expression and progeny virion production. Taken together, our findings confirmed that G-quadruplex formation in the oriLyt region regulate the initiation of lytic DNA replication in KSHV and RecQ1 binding at the G4 sites promotes replication possibly through unwinding these secondary structures. However, additional studies are needed to define the detailed mechanism of lytic DNA replication at oriLyt and RecQ1 mediated unwinding of G-quadruplexes for replication initiation.
Acknowledgements

We thank Dr. Jae Jung (University of Southern California) for iSLK-Bac16-TetRTA and iSLK-Bac16-ΔRTA cell lines. We also thank Dr. Sudha Sharma (Howard University) for RecQ1 antibody. We would also like to acknowledge Dr. Matthew Tucker (University of Nevada, Reno) for help with CD spectroscopy.
References


Concluding remarks and future directions

Twenty-five years since the discovery of KSHV, concerted research from all over the world has provided insights into KSHV biology but detailed mechanism of viral persistence and immune evasion still remains unanswered. None of the currently available treatments have been successful in eliminating KSHV infection due to a complex latency program. cART (combined antiretroviral therapy) has shown promise in HIV-infected KS patients, but these have been associated with immune system related complications and some of the cases have resulted in KS, despite of having no HIV load. Antiviral therapy, though successful, is not adequate to eliminate latent reservoir of the virus from the body. Hence, there is a need to have a complete understanding of KSHV pathobiology in order to design effective therapeutics against the virus. The work performed in this dissertation delineates few approaches employed by the virus in its life cycle i.e. DNA replication and immune evasion during latency and lytic DNA replication.

KSHV establishes latency following an acute or primary infection, during which the virus expresses a limited number of proteins and persists in the host for its lifetime by escaping immune surveillance. Since the virus lacks its own proteins required for latent replication, it hijacks the cellular replication machinery for replicating the viral genome. Latent DNA replication is essential for the maintenance of viral episomes in the host cells. LANA, the master regulator of latency, has been shown to help this process by establishing multiple points of contact between the viral and host genomes through direct interactions with cellular proteins. Members of the MCM complex, a six membered cellular helicase, were identified as LANA interacting proteins in our lab. We pursued this interaction and characterized the binding between MCMs and LANA through co-immunoprecipitation assays from KSHV infected cells. In addition to determining the specificity of this interaction with respect to replicative and mitotic
phase of the cell cycle, we were also able to elucidate a minimal binding domain of LANA required to interact with MCMs. This binding domain was absolutely essential for mediating interaction with one of the members of the MCM complex, MCM6, and was required for latent DNA replication and maintenance of the viral genome. We were also able to depict the therapeutic potential of this domain through overexpressing a peptide specific for the minimal binding domain that competed with full length LANA for binding to MCM6 and resulted in reduced latent DNA replication. Moreover, we were able to validate the importance of MCMs in latent DNA replication. We also showed that MCMs could only be recruited at the viral latent origin with the help of LANA as cells depleted for LANA showed significantly reduced MCMs loading. Finally, we showed that MCMs are required for active replication of latent DNA and both LANA and MCMs were assembled at the replication forks.

Apart from latent DNA replication, the other mechanism through which virus persists inside the host is immune evasion. The virus achieves this by encoding proteins that modulate various pathways involved in immune recognition. LANA was earlier shown to regulate its expression and evade antigen presentation, here through our studies, we proposed a novel scheme through which LANA avoids antigen presentation. We studied the formation of secondary nucleic acid structures called G-quadruplexes in LANA mRNA, where we first confirmed their formation through bioinformatics, biophysical and biochemical analysis. We were able to demonstrate the inhibitory effect of these structures through treatment of KSHV positive cells with G-quadruplex stabilizing ligand, TMPyP4, which led to a reduction in translation of LANA mRNA in addition to inhibition of antigen presentation, characterized by a weak CD8+ T cell response. We also studied the formation of these structures in detail through designing expression clones coding for G-quadruplex forming sequence and validated our
previous findings as these clones displayed compromised translation and reduced antigen presentation which corroborated with our hypothesis. We also studied the regulation of G-quadruplexes formation by cellular RNA binding protein, hnRNPA1, which played antagonistic role in the formation of G-quadruplexes through binding to G-rich LANA mRNA and increasing translation of G-quadruplex forming mRNA. In contrast, we observed that LANA regulated its expression through inhibiting the export of G-rich RNA from the nucleus to the cytoplasm at higher LANA concentration in cells.

After having studied the regulatory effect of G-quadruplex formation in LANA mRNA, we analyzed other regions of KSHV genome for G-quadruplex formation and discovered oriLyt region of KSHV to have a high propensity for G-quadruplex formation. oriLyt region is the site for initiation of lytic DNA replication, where cellular and viral proteins cooperate to initiate replication. RecQ1, a cellular helicase capable of unwinding the G-quadruplexes was shown to bind to oriLyt in affinity binding assay. We confirmed the formation of G-quadruplexes in oriLyt as well as demonstrated the specific binding of RecQ1 to the G-quadruplex forming sequence of oriLyt. We showed RecQ1 to be important for lytic DNA replication through chemical inhibition of its helicase activity and depletion of RecQ1 in KSHV positive cells. In addition, we also demonstrated the inhibitory effects of G-quadruplex formation on initiation of lytic replication, active DNA replication and virus production.

The study of various mechanisms involved in latency and lytic reactivation, furthered our understanding of KSHV biology thus contributing towards the development of therapeutic targets to block latency and lytic reactivation. Through our findings, we were able to discover new potential targets to suppress the latent and lytic infections of KSHV. Peptides specific for LANA-MCM6 binding domain i.e. 1100-1150 amino acids of LANA can compete with the full
length LANA for its binding to MCM6, leading to a reduction in replication of viral latent genome and virus persistence in the infected cell. Moreover, we can also reduce the expression of LANA, the master regulator of latency, through stabilization of G-quadruplexes in its mRNA, by introducing G-quadruplex stabilizing compounds. We can further stabilize the formation of these structures in LANA mRNA, by introducing a peptide that competes with the binding of hnRNPA1 to G-rich LANA mRNA. Since LANA is absolutely essential for latent DNA replication and persistence of the viral genome and LANA depletion results in a loss of viral episomes, we could target latent infection of the virus by targeting the expression of LANA. Lytic infection of the virus could be suppressed by impeding the initiation of lytic DNA replication at oriLyt region of KSHV, through stabilizing the G-quadruplexes formed in this region. Though the solitary use of these approaches might not be enough to eradicate KSHV infection, application of the above-mentioned approaches along with the currently available treatment options might be propitious, when used in conjunction.
In order to use the above-mentioned methods to treat KSHV infection, further research needs to be performed to improvise the design of peptides, that can compete with LANA and MCM6 binding or hnRNPA1 and LANA mRNA binding, in human cells. This research should be aimed at achieving efficient delivery of these peptide into the target cells, maximizing their half-lives by making them resistant to proteases in cells and eliminating any cytotoxic effects associated with them. As we seek to suppress the expression of LANA and subdue initiation of lytic DNA replication through G-quadruplex stabilization, extensive research on the biology of G-quadruplexes is required as they are highly complex and dynamic structures, a fact that has limited discovery of new ligands that stabilize these structures. Another limitation associated with the G4 ligands is their specificity. Since G-quadruplexes are also formed in the cellular genome, there is a need to ensure that these ligands are selective for targeting the viral G-quadruplexes and not the ones in the host genome or double stranded DNA. In addition to stabilizing G-quadruplexes, we also look forward to identifying cellular proteins involved in the regulation of G-quadruplexes formed in LANA mRNA and oriLyt. Nucleolin, a cellular protein, has been recently shown to facilitate inhibition of latent protein (EBNA1) expression and antigen presentation through binding directly to the G-quadruplex forming region of EBNA1 mRNA. Therefore, we want to investigate the effect of nucleolin on LANA mRNA G-quadruplexes, through knockdown and stable expression of nucleolin in KSHV positive cells lines. By studying LANA expression and antigen presentation in these cell lines, we can investigate the role of
nucleolin in KSHV latency. In addition, through our experiments, RecQ1 binds to the G-quadruplex forming site in oriLyt and potentially initiates DNA replication, further experiments aimed at confirming this observation and elucidating the mechanism of RecQ1 unwinding at oriLyt need to be performed for discovering a novel target for inhibiting lytic DNA replication. In this direction, we would like to study the mechanistic details by generating a recombinant KSHV genome using BAC homologous recombineering approach to disrupt G-quadruplex formation in the oriLyt region. Real-time monitoring of G-quadruplexes using fluorescent small molecule probes would also be helpful to investigate the dynamics of G-quadruplex formation.