THE NANOARCHITECTURE OF THE KSHV LANA TETHER AND APPROACHES TO ITS DISRUPTION

Margaret Josephine Grant
Shelton, CT

B.S., Brown University, 2007

A Dissertation presented to the Graduate Faculty of the University of Virginia in Candidacy for the degree of Doctor of Philosophy

Department of Microbiology, Immunology, and Cancer Biology

University of Virginia
May 2018

_____

_____

_____

_____

_____

_____

ABSTRACT

Kaposi's sarcoma-associated herpesvirus (KSHV) is the causative agent of three human malignancies and presents a special concern to immunocompromised individuals. Its latency-associated nuclear antigen protein (LANA) tethers latent viral genomes to host chromatin, thereby maintaining viral infection, and as such, is an attractive target for therapeutic intervention. In an effort to better understand this protein and its tethering function, we have used super-resolution microscopy to examine LANA tethers, obtaining information that remained obscured in earlier studies using standard epifluorescence microscopy. We have determined several characteristics of these tethers, including the folding properties of the underlying viral DNA and occupancy of LANA on its available viral binding sites. Quantitative data support the prediction of a coiled-coil feature in LANA dimers, and computer modeling of a minimal LANA tether illustrates the importance of viral DNA bending and nucleosome positioning on tether structure. Preliminary data examining LANA tethers during cellular mitosis suggest a potential role for mitotic machinery in manipulating tether positioning and condensation. This work also begins to address the relative timing of host chromosome condensation and LANA tether formation. These promising early results compel further study of the mitotic LANA tether and its dynamics. Finally, proof-of-concept experiments disrupting LANA expression via CRISPR-Cas9 gene editing pave the way for future manipulation of the KSHV genome and present a potential avenue for downstream therapeutic intervention.

ACKNOWLEDGEMENTS

My journey through science began with an absolutely terrible high school chemistry instructor, whose teaching was so bad it inspired me to read the textbook on my own to figure out what was going on. In the process of teaching myself the material, I developed a genuine love for science and pursued it through college and beyond. My interest in microbiology stemmed from the recommendation that I read two popular non-fiction “biological thriller” novels by Richard Preston: The Hot Zone and The Demon in the Freezer. These introduced me to the fascinating world of highly pathogenic viruses, prompting me to read more extensive texts on emerging diseases and biodefense in my spare time. It was in these moments, reading reference texts and papers for fun after work, that I knew I had found what I wanted to study for the rest of my life. I feel incredibly fortunate to have found my passion, and I owe it all to these seemingly small and unrelated events in my past.

I absolutely cannot talk about where I am now without mentioning my family. From the beginning, my parents instilled in me a drive to succeed and to continuously pursue learning. Mom, you were always showing and teaching me new things, setting me up for academic success. Dad, your expectations were steep, but you led by example, and I knew when you said, “I’m proud of you, kiddo,” the words were never empty. Without the two of you, I would also never have had the multitude of opportunities that presented themselves over the course of my early life, all of which served to shape me as a person. I am also

grateful for the support of my sisters, who have always been there when I needed someone to lean on. Martha, I will forever remember having a “second mom” for most of my childhood, as you let me use your expensive art supplies and even brought me on dinner dates with your husband. I’ve always known I had a safe space in you. Marion, your support when I was getting my bearings in Boston was invaluable. Weekends spent with you and Melanie always put me in a better mood for the upcoming week. I will also always be grateful for your being there when I reached important milestones. I know sitting through Brown’s hour-long Baccalaureate keynote address on RNA interference was a trying experience, but I appreciate you toughing it out for me! Mary, you’ve always known the little things that will make me feel better, whether those were unexpected gifts or the right words to say when I was upset. I value you as a sister, but also as a friend.

I owe a great deal of my graduate success to the advice and support of my thesis committee members: John Bushweller, Dan Engel, and Judy White. Your guidance over the years has been invaluable. Thank you for taking the time out of your busy schedules to help me grow as a person and scientist. Dr. Bushweller, I could always count on you to challenge my point of view, pushing me to test my intellectual limits. You taught me to think critically about my science and to see problems from an interdisciplinary angle. Dr. Engel, thank you for taking on the role of my committee chair, a position that I know created extra work for you. You have instilled in me the need to fully understand the

scientific literature supporting my research and to know how my project relates to the broader field. This mindset has changed the way I look at scientific research as a whole, and I know it will serve me well when developing future projects. Dr. White, you have always been my source of calm in moments of committee meeting frustration. I could count on you to be a kind, gentle source of encouragement, while still pushing me toward success. I am also grateful for the time you spent helping me find postdoctoral opportunities in your field. Your insight led me in the right direction at a time when everything seemed like an enormous black box.

When I entered graduate school, the main piece of advice I received was that the mentor and their style mattered more to one's success than the actual project. In this respect, I lucked out immensely. Coming to graduate school from a different field, I needed all the mentoring help I could get, and both of your doors were always open to me. Dean, I have always valued your ability to think about the big picture and come up with those crazy pie-in-the-sky ideas. I am grateful for the way you cared about me, not just as a scientist, but as a person, always willing to sit down and talk about everything from job searches to my parents' health. Because of you, I will leave UVA a better scientist, broader thinker, and ultimate detective of finding passive voice in a document. Mitch, I could always rely on you to be the voice of reason in our meetings, as well as a source of cynical humor when it was most needed. I hope to take away with me even a fraction of your analytical thinking skills. Thank you for taking up so much

of your time to really sit down with me and explain concepts and programming solutions. I know it must have felt like talking to a brick wall at times, but I promise you I learned a great deal from those sessions. I know I have frustrated you both immensely at times, especially in the beginning - experiments with no proper controls, fuzzy images, forgetting important pieces of the literature, this list could be infinite... Thank you for sticking with me through all my ups and downs, calling me out when necessary, and always making sure you were there to pick me back up when I needed it. I will miss your two very different personalities and the back-and-forth banter of our weekly meetings. I strive to one day be the scientist you both are.

I could not have accomplished any of this without the help of others in the laboratory. I owe the beginnings of my training to Zachary Weiner, a graduate student in my first rotation lab, who patiently walked me through calculating serial dilutions and gently corrected my upside-down plates in the incubator. I am also grateful for Tom Ellison, a postdoc in the Kedes lab, for guiding me through nearly every single laboratory technique I needed for my graduate career. From immunostaining and DNA gels to cloning and data interpretation, your skills are unmatched, and I've always felt I learned from the best. Matt Loftus, your ability to engage in quality research while seamlessly managing a laboratory has always been impressive. Thank you for everything you do around the lab and for enabling me to do my research without worrying about stocks and resources. I also owe you for your key piece of data in our paper – your results for the BCBL-

1 TR number were absolutely the cherry on top. Aiola Stoja, thank you for your extensive help in getting that pesky 2TR sequence figured out. It was comforting to have a second pair of eyes and someone who was equally as frustrated as I was. I was always happy to see you pop around the corner with your friendly smile and bubbly personality. I owe a great deal of the CRISPR-Cas work in this dissertation to Kerry Johnson, a smart, dedicated, hilariously cynical undergraduate who never shied away from a late night or a failed experiment. I know I was supposed to be training you, but you ended up teaching me a lot as well. Thank you for being a test subject for my tentative beginnings at mentoring and for reminding me that it's always important to know how each step in a protocol works. It was a pleasure watching you grow as a scientist, and I wish you nothing but success in your bright future. Finally, I can't come away from the Kedes Lab without thanking Melissa Anderson. Sure, you walked me through Western blots and answered various lab questions, but that's not what I will remember most. I never in a million years expected to make a friend in lab, but conversations turned to Starbucks runs together, and those turned to dinner invites with you and Mary, and soon I had two very dear friends. I went through some tough times, and you ladies always had your door open. Whether it was a crazy birthday dance party or a quiet day of movies and knitting, you were always there for me when I needed it. I miss our Friday margaritas and queso and our frequent walks to Monticello. I hope you know you are only in the "laboratory

help” category of these acknowledgements because of how we met. You very firmly belong in the category of “friends”.

I don’t think anyone can make it through graduate school without the support of friends. Amber and Brett, you guys have kept me sane over the years here at UVA. Whether it was lunches at Qdoba or walks around the hospital to de-stress, I always knew I would have a good time when you were around. Thank you for being there. I also owe a great deal of gratitude to the classical music community here at UVA. The Charlottesville Symphony and the Baroque orchestra both gave me an expressive outlet away from science for a bit while nurturing my artistic side in the process. Through this community, I met three people who became the greatest of friends. Bev, Ray, and Stef, I look forward to our regular quartet sight-reading sessions. Thank you for taking me into your circle, always making me feel comfortable enough to try something just a bit outside of my comfort zone, and being supportive when I make mistakes. You all provide a welcome, nurturing getaway from the grind of science. I will dearly miss our time together.

Some ladies have a best friend. Some of us are especially lucky and have two. Krystal, I still can’t believe how fast we became friends. You’ve been there for me through my absolute worst times, comforting or commiserating, as the case called for. I will always value your advice and judgement, whether it be personal or science-related. Thank you for believing in me, being a constant cheerleader, and for answering all of my really stupid science questions that I’m

too embarrassed to ask anyone else. You are brilliant, funny, kind, and brutally honest when necessary, and you are going places, whether you know it or not. Louise, I will forever be grateful that you kept in touch when I moved away from Boston. I think our friendship has done nothing but grow since then, and I can't imagine having potentially lost all of that over some distance. You are always there when I need someone to talk to, whether it's about a life-altering situation or a cute new pair of boots I've found. Your positivity and encouragement have brought me through some dark moments, and having my day sprinkled with your conversation always makes it a little bit brighter. You are spirited, generous beyond measure, and often make me laugh out loud at my desk (sorry, Matt). I hope these years of friendship only mark the beginning.

The last thing I expected moving to Charlottesville was to find love. It wasn't even on my radar. Imagine my surprise when I met you, Michael. Your constant, unwavering support has meant the world to me. In everything - from grocery store trips and neighborhood walks to moments when you held me while I cried, completely overwhelmed – you have been my partner and confidant. I absolutely could not have completed the crazy end of this graduate school trek without you. I look forward to coming home to you every day, and the thought that I will continue to do so for the rest of my life fills me with joy and peace. I am proud to call you my fiancé and am excited to see where the rest of our journey takes us. I love you!

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

ABBREVIATIONS

aa.	amino acids
AIDS	acquired immune deficiency syndrome
APC/C	anaphase-promoting complex/cytosome
BAC	bacterial artificial chromosome
BCBL	body cavity-based lymphoma
BoHV-4	bovine herpesvirus 4
BRD	bromodomain-containing protein
bp	base pairs
Cas9	CRISPR associated protein 9
CBD	chromatin binding domain
Cdk	cyclin-dependent kinase
cGAS	cyclic guanosine monophosphate-adenosine monophosphate synthase
Ch1R1	Ch1-related 1 protein
CI	confidence interval
CRISPR	clustered regularly interspaced short palindromic repeats
CRISPRa	CRISPR activation
CRISPRi	CRISPR interference
CRISPRn	CRISPR nuclease
crRNA	CRISPR RNA
DAPI	4’,6-diamidino-2-phenylindole

DC-SIGN	Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin
DBD	DNA binding domain
DEK	Drosophila Eph kinase
DNA	deoxyribonucleic acid
DS	dyad symmetry
dsDNA	double-stranded DNA
dSTORM	direct stochastic optical reconstruction microscopy
EBNA	Epstein-Barr virus nuclear antigen
EBV	Epstein-Barr virus
EMSA	electrophoretic mobility shift assay
ERK	extra-cellular signal regulated kinase
ET	extra-terminal
FR	family of repeats
GFP	green fluorescent protein
HAART	highly active anti-retroviral therapy
HBO1	histone acetyltransferase bound to ORC 1
HCMV	human cytomegalovirus
HDAC	histone deacetylase
HHV-8	human herpesvirus 8
HIV	human immunodeficiency virus
HP1	heterochromatin protein 1

HPV	human papillomavirus
Indels	insertions/deletions
IR	internal repeat
kb	kilobases
kbp	kilobase pairs
K_D	dissociation constant
kDa	kilodaltons
KRAB	Kruppel-associated box
KSHV	Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8)
KS	Kaposi's sarcoma
LANA	latency-associated nuclear antigen
LBS	LANA binding site
LV	lentiviral vector
LV-sgRNA	lentiviral vector with sgRNA insert
MAPK	mitogen-activated protein kinase
MCM	minichromosome maintenance complex component
MeCP2	methyl-CpG-binding protein 2
MEK	mitogen-activated protein kinase kinase
MHV68	murid/murine herpesvirus 68
MOI	multiplicity of infection
mRNA	messenger RNA
MTOR	mechanistic target of rapamycin

NEBD	nuclear envelope breakdown
NHEJ	non-homologous end joining
NLS	nuclear localization signal
ORC	origin recognition complex
ORF	open reading frame
OriP	EBV origin of replication
PAM	protospacer adjacent motif
PCR	polymerase chain reaction
PEL	primary effusion lymphoma
PERV	porcine endogenous retrovirus
PML	promyelocytic leukemia
RNA	ribonucleic acid
RNA Pol II	RNA polymerase II
RRV	rhesus monkey rhadinovirus
RTA	replication transcriptional activator
SAC	spindle assembly checkpoint
sgRNA	single guide RNA
shRNA	short hairpin RNA
SMC	structural maintenance of chromosomes
TBK1	TANK-binding kinase 2
TβRII	transforming growth factor-β type II receptor
TopBP1	topoisomerase II-binding protein

TPA	tetradecanoyl phorbol acetate
TR	terminal repeat
tracrRNA	trans-activating CRISPR RNA

CHAPTER 1. GENERAL INTRODUCTION: AN OVERVIEW OF KAPOSI'S SARCOMA-ASSOCIATED HERPESVIRUS AND ITS VIRAL EPISOME-HOST CHROMATIN TETHER

1.1 Persistence of DNA tumor viruses

The recent development of Gardasil®, a vaccine against several of the human papillomaviruses, has brought to the forefront of the public mindset the threat of cancer-causing DNA viruses. Examples of such viruses in humans include the papillomaviruses and two gammaherpesviruses, Epstein-Barr virus and Kaposi's sarcoma-associated herpesvirus (1). These viruses face the challenge of maintaining their genetic presence within actively-dividing cell populations. The key to their life-long persistence within the host is the ability to tether their circular viral genomes (episomes) to host chromatin. These tethering mechanisms differ among viruses and may include both viral and cellular proteins. Later sections discuss in greater detail the particular mechanisms behind different tethers. Despite having differences in their precise execution, all viral tethers share two main purposes. The first lies in maintaining viral genetic presence in host cells by partitioning viral episomes to progeny cells during mitosis. DNA from these viruses lacks centromeres and does not integrate into the host genome, removing these two avenues of viral persistence. In addition, episomes are present in multiple copies per cell and require a system by which they can be distributed between daughter cells during cell division. Formation of a tether between viral DNA and host chromatin ensures the presence of episomes in both daughter cells following mitosis. The second purpose of viral tethers is retention of viral episomes in the nucleus during reassembly of the nuclear envelope, thereby preventing recognition of the cytoplasmic DNA by cyclic guanosine monophosphate-adenosine monophosphate synthase (cGAS).

Binding of DNA to cGAS triggers an immune response via the cGAS-STING pathway, which ultimately induces expression of interferons and inflammatory cytokines (2). Retaining the viral DNA in the nucleus prevents activation of this cascade, thereby removing this avenue of anti-viral activity from the host immune system. Both roles served by the tether result in propagation of the viral genome, ensuring continued long-term viral genetic presence in the host.

1.1.1 Papillomaviruses

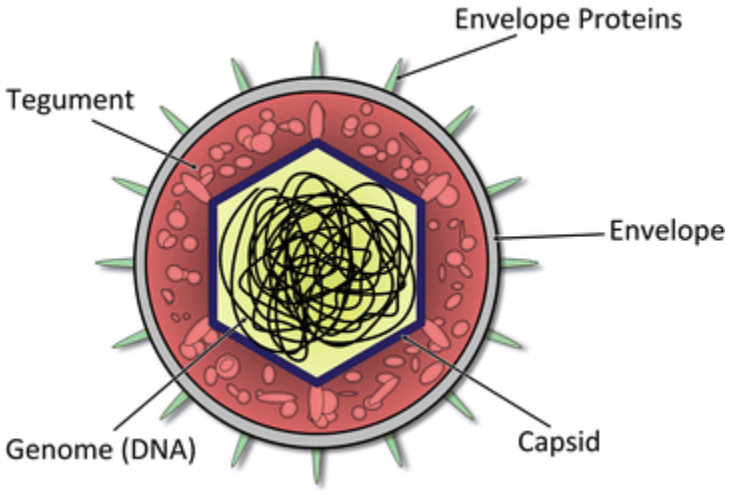
Papillomaviridae is a family of non-enveloped viruses approximately 60 nm in diameter. Papillomavirus genomes consist of approximately 8000 base pairs (bp) of double-stranded DNA (dsDNA) encoding 8 or 9 open reading frames (ORFs). There are approximately 100 defined human papillomavirus serotypes all of which infect the basal cells of mucosal or cutaneous epithelia (3). Greater than 99.7% of diagnosed cervical cancers contain HPV DNA, as do 44.1% of oropharyngeal tumors (4, 5). Papillomaviruses maintain their genetic presence by tethering their episomes to host chromosomes via the multifunctional protein E2. On the viral episome side, E2 recognizes and binds the sequences $ACCGN_4CGGT$ or $ACCN_6GGT$, which are found in the long control region (LCR) of the episome (6, 7). It is believed the host side of this interaction is an indirect one, where E2 binds to the bromo-domain protein Brd4 that, in turn, binds mitotic chromosomes. In fact, studies blocking the interaction between E2 and Brd4 both prevented E2 association with chromosomes and enhanced loss of viral genomes over time. Studies also found that Topoisomerase II-binding protein

(TopBP1) co-localized with E2 during later mitosis, perhaps suggesting that this protein is another means of E2 tethering (8). An additional element of interest in the E2 tethering mechanism is the Ch1-related 1 (Ch1R1) protein, which interacts with E2 (9). Depletion of Ch1R1 also disrupted the association of E2 with mitotic chromosomes and, as with blockade of Brd4 binding, resulted in the loss of viral genomes over time. E2 co-localized with Ch1R1 during S-phase, but not during mitosis, suggesting that Ch1R1 is required for loading E2 and viral genomes onto host chromatin during S phase (10).

1.1.2 Herpesviruses

Herpesviridae is a family of enveloped viruses with dsDNA genomes that can cause disease in humans, among other species. Each viral particle (virion) consists of a DNA-containing icosahedral capsid, an outer envelope, and a proteinaceous tegument layer between (Figure 1-1). The fully mature herpesvirus particle is 150 to 200 nm in diameter (11) and its envelope contains various glycoproteins involved in cell entry and egress. There are three sub-families of herpesviruses: alphaherpesviruses (human viruses include herpes simplex viruses 1 and 2, and varicella zoster virus), betaherpesviruses (human cytomegalovirus [HCMV], human herpesviruses 6 and 7), and gammaherpesviruses (Epstein-Barr virus [EBV] and Kaposi's sarcoma-associated herpesvirus [KSHV]). All are characterized by their ability to establish an indefinite latent infection in their hosts. This latent phase of infection is marked

Figure 1-1. Herpesvirus structure. Herpesviruses contain a double-stranded DNA core surrounded by an icosohedral capsid, a proteinaceous tegument layer, and an outer glycoprotein-studded lipid envelope.


Envelope Proteins
Tegument
Envelope
Genome (DNA)
Capsid

by minimal viral gene expression and a lack of viral spread. Environmental conditions, such as stress, can induce progression into the virus's lytic phase, triggering a cascade of increasing viral gene expression and production of new virions.

Alphaherpesviruses establish their initial infection in epithelial cells, but then spread to the nuclei of nearby sensory neurons, where they establish latency (13). Common symptoms of an active infection include mucosal sores or a vesicular rash, but more severe cases may result in encephalitis or meningitis (14). Betaherpesviruses have a broad cell tropism, infecting epithelial cells, polymorphonuclear leukocytes, and T cells (15), but they ultimately establish latency in monocytes and macrophages. Infection with the betaherpesvirus HCMV can cause pneumonia, hepatitis, and leukopenia, while human herpesviruses 6 and 7 are known for primarily causing roseola in young children. The gammaherpesviruses are divided into two genera. The *Lymphocryptovirus* genus includes EBV, which initially infects the epithelium of the nasopharynx and causes flu-like symptoms such as fever, fatigue, and sore throat. The genus *Rhadinovirus* contains only one human virus, KSHV, but encompasses additional species that infect non-human primates and mice. Both of the human gammaherpesviruses establish long-term latency in B cells.

Epstein-Barr virus is a gammaherpesvirus that infects approximately 90% of adults worldwide. It generally maintains its presence asymptomatically, however persistent infection is linked to Burkitt's lymphoma, Hodgkin's disease,

and nasopharyngeal carcinoma (16). EBV has an approximately 165 kb circular genome, which is maintained as an episome in latently-infected memory B cells (17). During latency, the viruses expresses only nine proteins, including five EBV nuclear antigens (EBNAs).Among these is EBNA1, a multi-functional protein that also maintains latent episomes in cells by tethering them to host chromatin. EBNA1 attaches to the viral genome via 24 EBNA1-binding sites, each 18 bp long, located in the viral origin of replication (OriP). Of these sites, 20 are high affinity and located within the family of repeats (FR) element, a tandem series of 30-bp repeats, and 4 are low-affinity and located in the dyad symmetry (DS) element (18, 19). The FR element is necessary for chromosome tethering, while the DS element, located approximately 1 kb away, enables replication of the viral episome. Unlike the indirect tethering to host chromosomes adopted by the papillomavirus E2, EBNA1's attachment is direct. EBNA1 binds directly to mitotic chromosomes via an AT-hook motif that binds to AT-rich sequences (20). In experiments, mutant EBNA1s lacking their host chromosome binding region were also inefficient at replicating OriP-containing plasmids, suggesting a strong relationship between the two functions (21).

Kaposi's sarcoma-associated herpesvirus (KSHV), the focus of this study, also maintains its genetic presence through a direct episome-tethering protein. This protein, called the latency-associated nuclear antigen (LANA), and its attachments will be described in later sections.

1.2 Kaposi's sarcoma

In 1872, the Hungarian physician Moritz Kaposi had five male patients exhibiting what he referred to as "idiopathic multiple pigmented sarcoma" on their lower extremities (22). A century later, this disease became a defining feature of the 1980s AIDS epidemic, manifesting as multiple purple, red, or brown pigmented plaques on skin of the arms, legs and trunk. In 1990, Beral et al. posited that the source of this disease was a sexually transmitted pathogen (23). In 1994, Chang et al. isolated DNA from these tumors and used representational difference analysis to distinguish viral DNA from that of the surrounding tissue. Therein they found sequences homologous to those found in known herpesviruses (24). The discoverers named this new virus Kaposi's sarcoma-associated herpesvirus (KSHV). It became the eighth known human herpesvirus, earning it, the ICTV designation of "human gammaherpesvirus 8", often abbreviated as HHV-8 (25). Because of its discovery method, KSHV has the distinction of being the first virus whose fragments were identified directly by polymerase chain reaction (PCR).

The most common disease from infection with KSHV is the cutaneous form of Kaposi's sarcoma (KS). Cutaneous KS manifests in three distinct phases: patchy, plaque, and nodular. Patients first present with flat, bruise-like lesions (patchy phase). These lesions often coalesce and grow in thickness (plaque phase) and eventually become increasingly raised and nodular (nodular phase). The more severe later stages of the disease present with lymphoedema, gastrointestinal lesions, respiratory distress, and fungating tissue necrosis.

Histological examination of pathological specimens indicates proliferation of spindle (elongated) cells at the sarcoma site, separated by slit-like vascular spaces (11).

There are four types of KS: classic, endemic, iatrogenic, and AIDS-associated, or epidemic (26). Each results from KSHV infection but varies most likely with the degree of immunosuppression. Classic KS typically affects elderly people of Mediterranean, Eastern European, and Middle Eastern ancestry, and is more common in men than in women. These individuals typically exhibit lesions on the lower extremities, and spread of these lesions is much rarer than in other forms of KS. Endemic, or African, KS affects younger individuals living in equatorial Africa, due to moderate levels of immunosuppression resulting from other chronic infections and malnutrition. Iatrogenic, or immunosuppression-related, KS occurs in people whose immune systems have been severely compromised, often by the drugs required following an organ transplant. Epidemic, or AIDS-related, KS is the most common type and develops in HIV-infected individuals who suffer from the most intense levels of immunosuppression. The link between HIV1-induced immunosuppression and KS has led to the classification of this tumor as an “AIDS defining” illness, marking the change in diagnosis from HIV-positive to AIDS . Treatment of HIV infection with highly active antiretroviral therapy (HAART) has decreased the number of epidemic KS cases, although patients undergoing HAART may still develop KS.

1.3 Kaposi's sarcoma-associated herpesvirus

An opportunistic pathogen, KSHV primarily presents a risk to immunocompromised patients, including transplant recipients and HIV-infected individuals. The infection rate in Northern Europe, the Americas, and Asia is less than 10%. An exception is South American Amerindians, who have a seropositivity rate of 23 to 53% (27, 28). In the Mediterranean region, infection rates are 5 to 30%, and in sub-Saharan Africa, they are highest, at times reaching 60% of the total population (29). Greater prevalence is also evident in men who have sex with men (MSM). A 2006 study showed a 24.3% prevalence rate in MSM in the Seattle and Denver areas of the United States, while data from the Multicenter AIDS Cohort Study in 2010 found 54.6% of MSM were infected (11, 30–33). KSHV is present in five genotype groups, each defined by hypervariability in the glycoprotein gene ORF-K1. These genotypes also correlate with distinct geographical areas (27, 34). KSHV DNA sequences are present in patient saliva (31) yet are often undetectable in semen, even in patients with KS (35). These findings, coupled with extensive epidemiologic data, suggest that the most common route of infection is likely salivary transmission.

In addition to causing KS, KSHV is also the etiological agent of two other human malignancies: primary effusion lymphoma (PEL), and multicentric Castleman's disease (MCD) (36, 37). PEL, also known as body cavity-based lymphoma (BCBL) is a rare non-Hodgkin's B cell lymphoma that develops in the pleural, abdominal, and pericardial spaces. It predominately infects B cells

expressing the lambda light chain and presents with no primary contiguous tumor mass, but secondary tumors can arise (38). Prognosis is poor, with a median survival time of 6 months, as standard lymphoma chemotherapies are generally ineffective (39). A single PEL cell contains between 40 and 100 copies of KSHV, and approximately 50% of these cells are co-infected with EBV (36, 40, 41). PEL cells are dependent upon the presence of KSHV and its gene products for survival, as Godfrey et al. demonstrated when they reduced KSHV copy number by knocking down expression of the genome maintenance gene LANA (42). Several clonal PEL lines are used in KSHV research, as they represent a system of stable latency and can also be induced to release infectious virus for further study.

Multicentric Castleman's Disease presents most often with mediastinal lymphadenopathy and accumulation of virus-infected lymphocytes in the splenic mantle, resulting in spenomegaly (11). This expansion of the mantle results in follicular cell lysis and also gives the disease its characteristic "onion skin" pathology (43). MCD falls into two subtypes: hyaline vascular, in which cells of the lymph nodes are hyalinized, and plasma cell, in which the interfollicular area is infiltrated by oversized, multi-nucleated plasma cells. The "multicentric" aspect of the disease arises from the fact that growth occurs at multiple sites throughout the body, rendering surgery ineffective. In rare cases, the disease can progress to hemophagocytic syndrome, a condition in which rapidly proliferating macrophages consume the red blood cells, resulting in T cell overstimulation,

pancytopenia, high free iron levels, and hepatosplenomegaly. Left untreated, the syndrome results in patient death.

HAART has proven to be somewhat effective in minimizing the symptoms of both KSHV-related primary effusion lymphoma (44) and, in combination with pegylated liposomal doxorubicin, Kaposi's sarcoma (45). The anti-proliferative effect of the mTOR inhibitor rapaymycin has proven effective in preventing lytic replication in culture (46–48). In addition, anti-viral treatments, such as cidofovir, ganciclovir, phosphonoformic acid (48), and North-methanocarbathymidine (N-MCT) (49) have successfully inhibited viral DNA synthesis and virus secretion. However, currently no KSHV-specific therapies or vaccines exist. This is of special concern in circumstances where the source of immune suppression is untreatable, such as certain auto-immune diseases. All of the KSHV-related malignancies, if left untreated, can eventually result in a slow and painful death via respiratory depression and organ failure, making research toward finding effective therapies critical.

1.4 Virus life cycle

Viral entry into cells is mediated by glycoproteins gB (ORF8), gH (ORF22), and gpK8.1A, all of which interact with heparan sulfate expressed on the cell surface (50). Additional cell surface proteins utilized in viral entry, depending on the cell type, include integrins α3β1, αVβ3, and αVβ5, cysteine transporter xCT, and the dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin, or DC-SIGN, the last of which is used primarily for entry into non-

adherent cells. Processing and transport of gH is mediated by complex formation with gL (ORF47), making it an additional part of the integral cell fusion machinery. Following binding to one of these receptors, the virus is internalized by endocytosis; binding to one of the integrin receptors activates a kinase cascade including the mitogen-activated protein kinase kinase (MEK) and extra-cellular signal regulated kinase (ERK) ½ (51) that promotes actin rearrangement for carrying the viral particle to the nucleus (52). An additional entry pathway occurs via lipid rafts on the cell surface; this pathway utilizes macropinocytosis for internalizing the viral particle. Virus internalization via integrin signaling increased dramatically between 30 and 60 minutes following infection, plateauing at 150 hours (53). The next steps appear to be highly conserved throughout the herpesvirus family. First, the virus injects its linear, double-stranded DNA through a nuclear pore. Next, a cellular DNA ligase IV and double stranded break repair enzyme mediate the circularization of the viral DNA, forming the viral “episome” by joining the tandem repeat regions that flank each end of the genome (54). Researchers observe capsids “docking” at the nuclear membrane as soon as 15 minutes post-infection, with newly delivered viral DNA content peaking at 90 minutes (50).

Viral DNA in capsids is chromatin-free, likely due to space constraints (55), but once in the nucleus, herpesvirus DNA is rapidly chromatinized. DNA from the human cytomegalovirus (hCMV) was found to be associated with histone H3 as soon as 30 minutes post-infection (56). In addition, Nitzsche et al.

observed an increase in nucleosomes on viral DNA between the early and late stages of the hCMV replicative cycle (56). Early HSV-1 chromatin predominately contains the replication-independent histone H3.3 (57) but later acquires H3.1, which is viral DNA synthesis-dependent. Histone modifications appear to require more time to occur but are clearly evident 5 days following de novo infection of endothelial cells (58). These modifications vary, reflecting one of two routes the viral DNA may take: latency or lytic replication. Section 1.5 discusses these two pathways, and their respective histone modifications.

In latent infection, promyelocytic leukemia (PML) bodies appear to form a shell surrounding the viral DNA, perhaps in an attempt to sequester viral genomes in an inactive/repressed environment (55). PML nuclear body components have been found to restrict herpesvirus infections, but the virus has evolved to express regulatory proteins that overcome this restriction. Interestingly, studies suggest that the final stages of viral gene transcription and capsid assembly in herpesviruses take place in intranuclear replication compartments developed from these PML bodies. Two separate studies on hCMV found a decrease in histone H3 at very late times post-infection (56, 59); this is consistent with the idea that viral DNA must lose its nucleosomes prior to packaging within a capsid. DNA-filled capsids bud into the inner nuclear membrane, leading to primary envelopment of the capsid. This envelope then fuses with the outer nuclear membrane, releasing the free capsid into the cytoplasm. While in the cytoplasm, the capsid acquires its tegument proteins

before budding into a glycoprotein-containing vesicle. The inversion created by invagination in this vesicle creates an envelope around the tegument (secondary envelopment) with glycoproteins studding the outside. Multiple capsids can invaginate into a single vesicle, which then fuses with the cell membrane, releasing fully tegumented and enveloped virions into the extracellular space. Figure 1-2 summarizes the complete herpesvirus life cycle.

1.5 Viral episome

The KSHV genome (Figure 1-3) is approximately 165 kbp (60), varying in the number of tandem repeats present at each end (61). Within the genome are over 90 open reading frames (ORFs 4 through 75 plus another 17 ORFs with a "K" designation), encoding various viral proteins as well as a cluster of microRNAs (50). Two regions of 801-bp (62) tandem terminal repeats (TRs) flank the genome on either side, joining together when the viral genome circularizes upon entry.

1.5.1 Viral latency

During latent infection, only a few maintenance genes are transcribed from the "latency region" of the genome (63). These include the latency-associated nuclear antigen (LANA, ORF73), a multifunction protein that acts as a transcriptional activator or repressor for a variety of genes and serves to tether the viral episome to host DNA during mitosis; v-cyclin (ORF72), a protein that

Figure 1-2. Life cycle of a herpesvirus. Diagram of the herpesviral life cycle, beginning with viral attachment and entry and followed by the release of viral DNA into the nucleus. A cascade of viral transcription results in replication and capsid assembly. Viruses then undergo two subsequent envelopment stages and are released from the cell.

Reprinted from Virus Research, Vol. 106, Issue 2, Mettenleiter TC, Budding events in herpesvirus morphogenesis, p. 167-180, Copyright (2004), with permission from Elsevier (64).

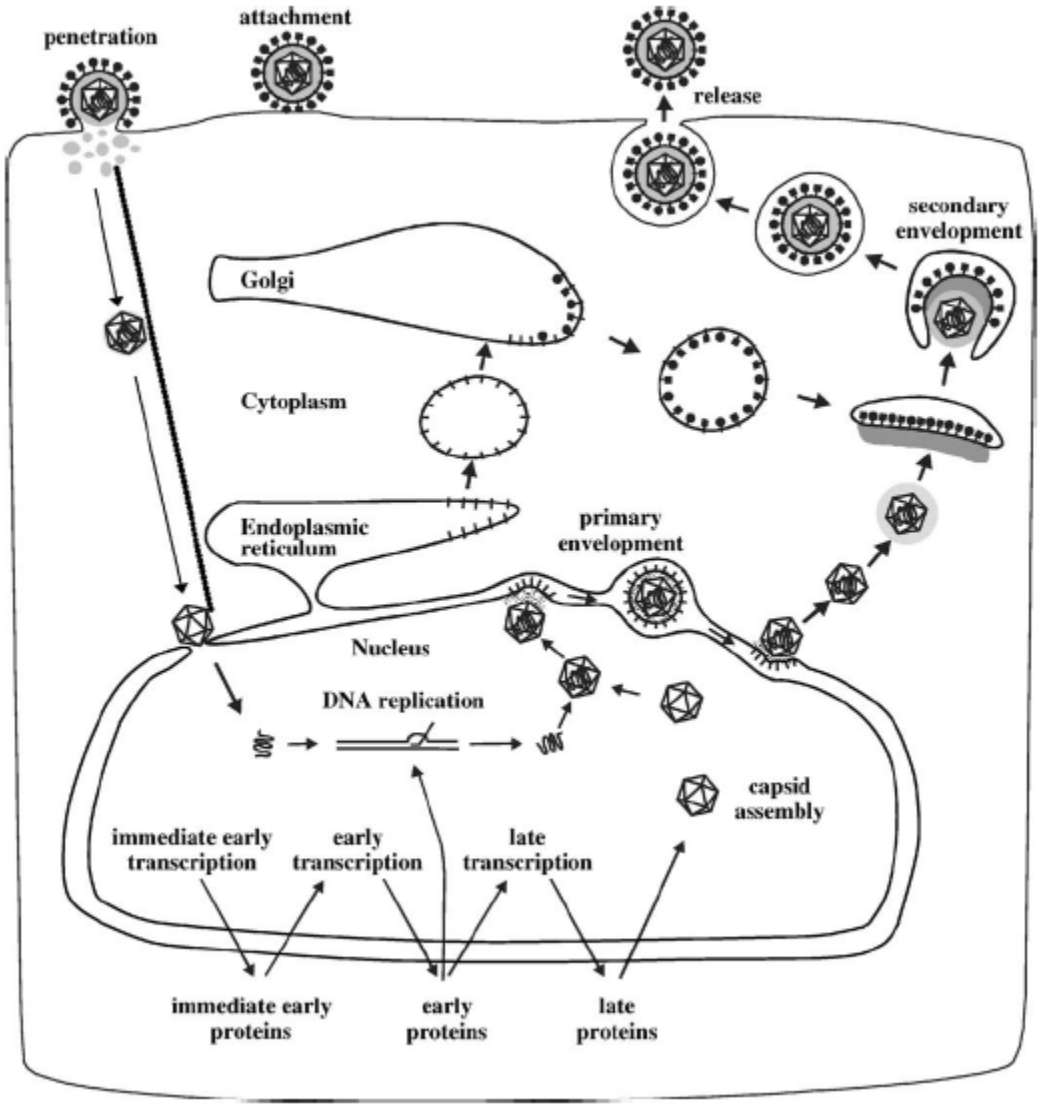
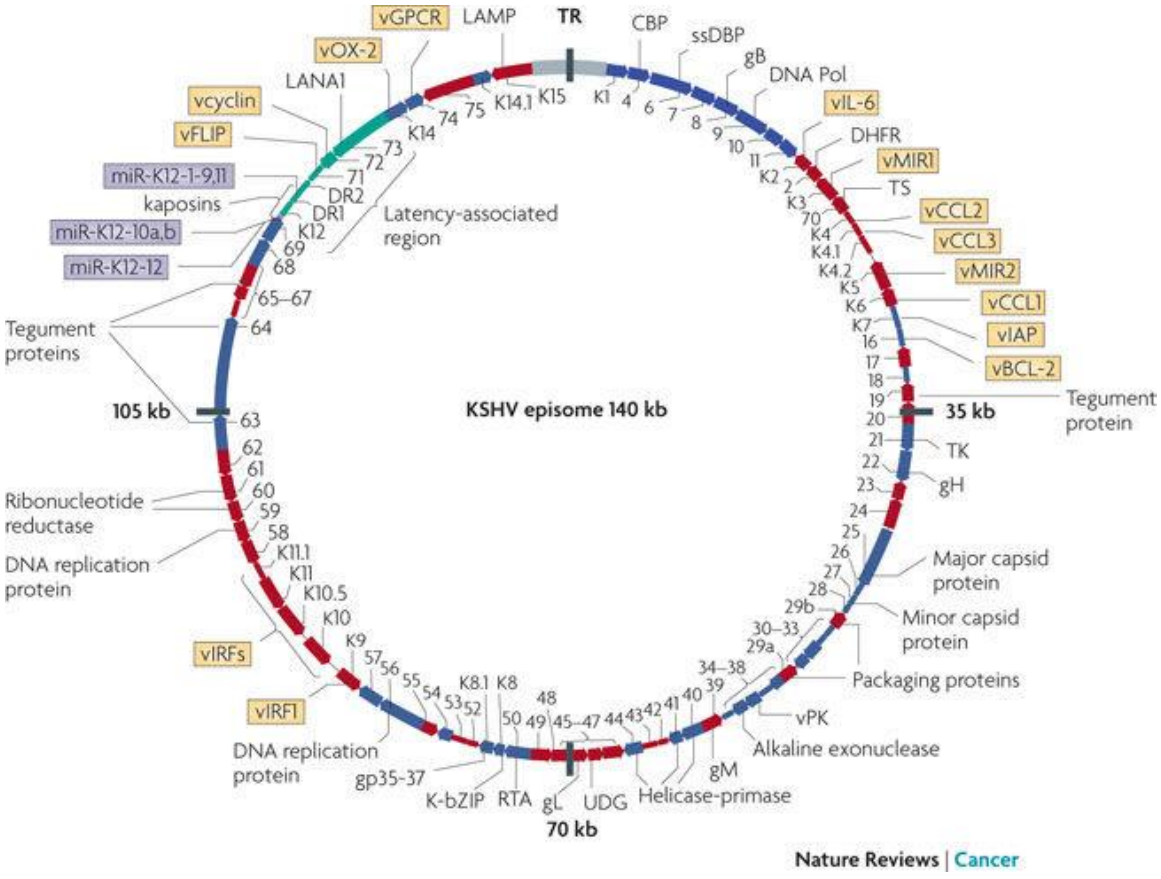

penetration
attachment
release
secondary envelopment
Golgi
Cytoplasm
Endoplasmic reticulum
primary envelopment
Nucleus
DNA replication
capsid assembly
immediate early transcription
early transcription
late transcription
immediate early proteins
early proteins
late proteins

Figure 1-3. The KSHV episome. Schematic of the KSHV episome, showing the relative locations of several key genes, as well as the long terminal repeat (TR) region that results from circularization. Note that the 140 kb size mentioned does not take into account the TR region.


vGPCR
LAMP
TR
CBP
ssDBP
gB
DNA Pol
vIL-6
DHFR
vMIR1
TS
vCCL2
vCCL3
vMIR2
vCCL1
vIAP
vBCL-2
Tegument protein
35 kb
TK
gH
Major capsid protein
Minor capsid protein
Packaging proteins
vPK
Alkaline exonuclease
gM
Helicase-primase
UDG
gL
RTA
K-bZIP
gp35-37
DNA replication protein
vIRF1
vIRFs
K10
K10.5
K11
K11.1
DNA replication protein
Ribonucleotide reductase
105 kb
Tegument proteins
miR-K12-12
miR-K12-10a,b
kaposins
miR-K12-1-9,11
vFLIP
vcyclin
LANA1
vOX-2
Latency-associated region
KSHV episome 140 kb
70 kb
Nature Reviews | Cancer

promotes growth by overcoming normal cell cycle control mechanisms (66); v-FLIP (K13), a protein that inhibits apoptosis by binding to FADD-like interferon-converting enzyme, thus encouraging tumor progression (67); and Kaposin (K12), a protein that activates MEK2, thereby increasing cytokine expression (68). This latency locus gives rise to three 3'-coterminal RNAs from the same promoter (63, 69). The first is an unspliced 5.8-kb transcript, which likely acts as the mRNA for ORF73. The second is a 5.4-kb transcript encompassing the genes for ORF71, 72, and 73, generated from splicing out 336 nucleotides from the 5' UTR of the 5.8-kb transcript. The final transcript is 1.7-kb and likely acts as the mRNA for ORF71 and 72. In addition to these three protein-encoding RNAs, 12 miRNAs are transcribed during latency. Transcription of these latent genes involves the recruitment of the host's own transcription machinery.

The untranscribed portions of latent viral DNA acquire suppressive marks, including widespread H3K27me3 (70, 71), which is absent from the latency locus and terminal repeats (72). Instead, these two specific regions contain H3K9/K14-ac and H3K4me3 modifications – marks associated with active chromatin and transcriptional activation (70–75). The chromatin status of the TRs, however, remains unclear. In 2004, Sakakibara et al. found that immunoprecipitated heterochromatin protein 1 (HP1), after cross-linking, was associated with TR DNA (76). As HP1 binds to H3K9me2/3 markers and is a characteristic of heterochromatin, this is in direct contrast to the earlier suggestions of an active TR chromatin conformation.

1.5.2 Lytic reactivation

Lytic reactivation stems from activation of ORF50, the region encoding the replication and transcriptional activator (RTA), the “switch” protein that has the capacity to initiate the lytic cascade of virally encoded proteins necessary for productive (lytic) infection (77). During latency, the RTA promoter has both H3K27me3 and H3K4me3 markers, indicating bivalent chromatin (70). As the virus transitions from latent to lytic phase, the H3K27me3 marker is slowly lost from the ORF50 promoter, while H3K4me3 and AcH3 increase. RTA has also been found to interact with the SWI/SNF chromatin remodeling complexes, which are essential for transcription initiation; RTA mutants with no SWI/SNF interaction were found to have reduced promoter activity (78).

While only 1 to 3% of PEL cells spontaneously progress to lytic phase in culture, they are inducible by treatment with histone deacetylase (HDAC) inhibitors, such as valproic acid and sodium butyrate, as well as tetradecanoyl phorbol acetate (TPA), which initiates reactivation via stimulation of the MAPK/ERK pathway and by demethylating the ORF50 promoter (79, 80). TPA in particular can induce RTA expression in as little as four hours (50). TPA specifically was also found to induce histone deacetylases by enhancing the DNA binding activity of transcriptions factors (81). Ye et al. found that the ORF50 promoter contained a binding site for Sp1, which is required for activation of promoters by sodium butyrate (82). In addition, the promoters for two lytic genes

(v-Ox2 and v-GPCR) were found to contain 3 putative Zebra transactivator-like binding sites, which are known to be induced by TPA (69).

In addition to losing suppressive marks, lytic genes begin to acquire activating histone modifications, including hyperacetylation (55). These genes are expressed in a cascade of three stages: immediate early, early, and late genes. Immediate early genes are mostly involved in counteracting host defense mechanisms and in producing the virus's own replication machinery. Other genes are implicated in the removal of histones from nascent DNA before being encapsidated. Late-stage KSHV genes are those involved in the final stages of virion production, including those encoding the structural proteins necessary for viral assembly. Transcription of late genes requires viral DNA replication (synthesis).

In 2017, Chen et al. noted the formation of viral transcriptional "factories" upon lytic activation (83). These factories were multilobular spaces with low or no staining for host DNA. Around these spaces was an outer ring of LANA positioned around an inner ring of RNA Polymerase II (RNA Pol II), with RTA RNAs found within. The study also found that single-stranded viral DNA co-localized with RNA Pol II, indicating RNA Pol II was recruited to the sites of viral DNA replication. Interestingly, they also noted proteasomal degradation of cellular RNA Pol II during lytic reactivation, which resulted in downregulation of several cellular housekeeping genes. This full-on "hijacking" of the cellular RNA

Pol II system ensures the prioritization of viral DNA replication at the expense of cellular maintenance.

1.5.3 Terminal repeat region

The terminal repeats are GC-rich, with 84.5% GC content, compared to the 53.5% found in the central unique region of the genome (62). Distribution of the TRs is uneven between the two ends, but one viral episome contains between 35 and 60 in total (61, 84). The total number of TRs per episome is inconsistent among different KSHV strains but may be consistent within a single strain. Studies of various PEL lines, primary tumors, and KS biopsies have shown that while several strains of KSHV are monoclonal with respect to their number of TRs per episome (e.g. BCBL-1), others are biclonal (e.g. BC-3) or oligoclonal (84, 85). The extensive variation in TR number suggests that KSHV can be successfully packaged in capsids with at least a 20% variation in genome size (84). Circularization by way of these TRs allows for rolling-circle replication of the genome. In addition to being the site of circularization, the TRs play an important role in episome persistence. Within each TR are three binding sites by which the genome is tethered to host chromatin during viral latency via LANA. This is especially critical during mitosis, when the partitioning of viral episomes is vital for infection of progeny cells and breakdown of the nuclear membrane renders episomal DNA susceptible to detection by cGAS. Shrestha and Sugden demonstrated that 2 TRs were necessary for maintenance of plasmids within cells and that plasmids with fewer than 16 TRs showed recombination,

generating constructs with greater numbers of TRs (86), They reasoned that a larger number of TRs provided the plasmid with a selective advantage by increasing its capacity for supporting DNA replication, establishment, and long-term maintenance, all functions supported by the TR region.

The terminal repeat regions also each contain a full copy of the episome's origin of replication, at nucleotides 509 through 587 (87). The minimal required element for replication consists of this region plus a 62-base pair upstream element. This renders each terminal repeat self-sufficient for replication of viral DNA.

1.6 Latency-associated nuclear antigen (LANA)

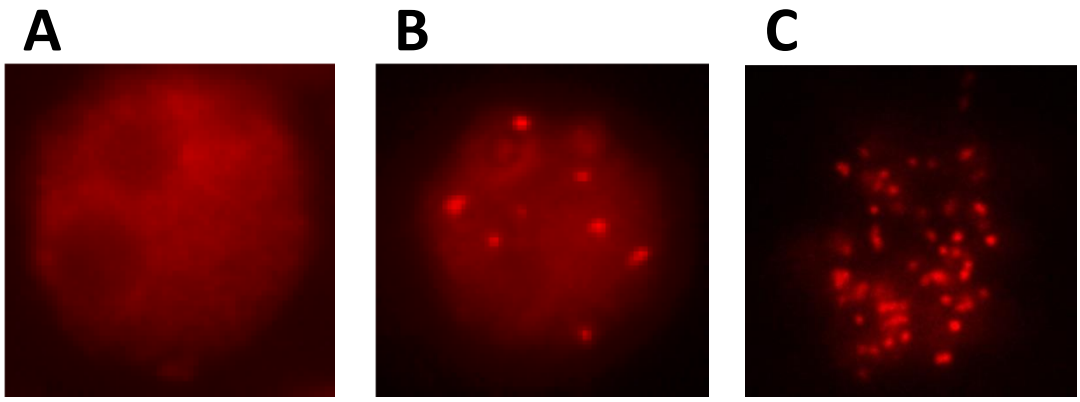
1.6.1 LANA functionality

The latency-associated nuclear antigen, or LANA, is one of the few viral proteins produced during KSHV latency and has multiple functions. One of its main roles is the tethering of viral episomes via its C-terminus to host chromatin via its N-terminus. In this way, it ensures maintenance of the KSHV genome through sequential cell divisions without requiring DNA integration into the host genome (88–90). In agreement with this function, LANA associates with heterochromatin during interphase and with condensed chromosomes during mitosis (88, 91, 92). Later sections discuss the connections mediating this function in further detail. LANA is generally diffuse and nuclear via immunofluorescence in the absence of DNA containing the viral LANA binding

sites (LBSs) and punctate in the presence of LBSs (Figure 1-4) (88, 93). Ballestas et al. (94) found that DNase treatment, but not RNase treatment, eliminated characteristic punctate nuclear LANA staining, thus indicating that the presence of episomal DNA mediates this focal concentration. While the N-terminal binding site alone coats chromosomes broadly, C-terminal residues 996 through 1139 create peri-centromeric and -telomeric foci even in the absence of TRs (95).

Another major function of LANA is as a transcriptional regulator of viral genes. In particular, LANA represses ORF50 lytic reactivation by interacting with the transcription factor Sp1 and binding directly to the ORF50 promoter (96). Removal of LANA enables Sp1 interaction with the promoter, leading to transcription of ORF50 and RTA induction, driving the progression of the lytic cycle. After RTA induction, LANA disassociates from the episome within 4 hours (97). Deletion of LANA from the system altogether results in an increase in lytic gene expression, both before and after induction, and increases the number of infectious virions produced (98). In addition to regulating ORF50, LANA affects episome replication by recruiting MCM3, MCM5, ORC, and HBO1 (a histone acetyltransferase) to the TR region, thereby contributing to DNA replication (73, 74). In fact, Stedman et al. found that ORC2 only bound to TRs in the presence of LANA (74). Methylation at arginine-20 (mediated by PRMT1) increased the degree of LANA occupancy on episomal DNA and reduced gene expression in

Figure 1-4. Punctate LANA tethers are only visible in the presence of TRs. **(A-B)** Epifluorescence images show BJAB cells transfected with LANA-expressing plasmid plus either **(A)** a plasmid containing zero TRs, resulting in diffuse, non-punctate LANA staining, or **(B)** a plasmid containing 8 TRs, resulting in punctate LANA staining. **(C)** A BJAB cell infected with virus containing the KSHV BAC16 also shows punctate LANA staining.


A
B
C

those regions (97). One study found that LANA also binds to LBS-like sequences in the host genome and associates with H3K4me3-rich regions (99). In this way, it may play a role in host gene expression.

LANA also indirectly affects transcription of host genes through association with methyltransferases. For example, LANA binds the transforming growth factor-β type II receptor (TβRII) promoter, resulting in its methylation (73). This prevents transcription and encourages tumor development, as TβRII inhibits growth and promotes apoptosis. In this way, LANA also acts as a promoter of tumor growth. Other examples of this function include LANA binding directly to the tumor suppressor p53 (100) and binding to GSK-3β (101), which is a negative regulator of β-catenin; this interaction disturbs the Wnt/β-catenin cascade and drives cells into S phase.

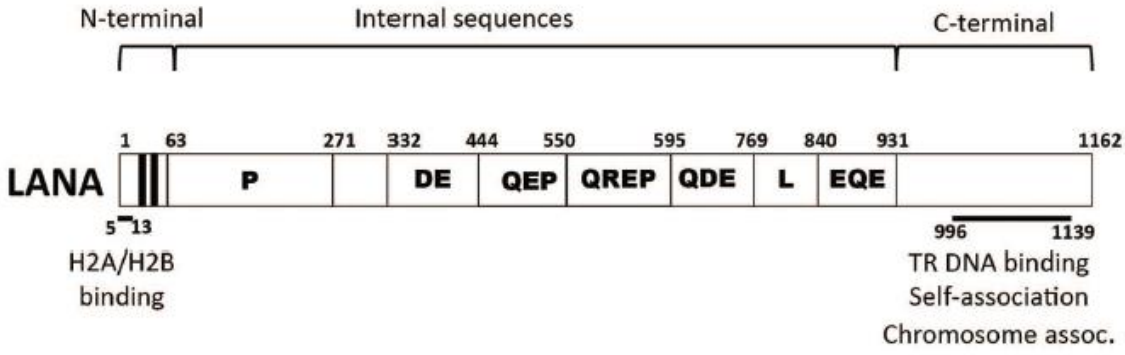
1.6.2 LANA structure

LANA is 135-kDa, consisting of approximately 1162 amino acids. Its exact length depends greatly upon the length of its internal repeat (IR) region and varies by KSHV strain. It has a proline-rich N-terminus, characterized by its host chromatin binding domain at amino acids 5 to 13. The IR region can be divided into three sections: aa. 340-431 with a DEED or DEEED pattern; aa. 440-756 exhibiting patterns QQQEP, QQREP, and QQQDE; and aa. 760-931 containing QEQELEE or QELEVEE (leucine zipper) motifs (102). Research to date has shown that the IR region has roles in short- and long-term episome persistence (residues 889-949) (103) and replication (residues 329-931) (103, 104). The

LANA C-terminus contains a high content of basic residues (87) and the viral episome terminal repeat binding domain at aa 996-1139 (105). Both the N- and C-termini contain a nuclear localization signal (NLS), ensuring that the protein can perform its various nuclear functions (106, 107). The C-terminus also contains a secondary host chromatin binding domain that forms a cooperative association with the primary N-terminal binding site, although the mechanism of this cooperativity is unclear (108). In fact, when the N-terminal binding site is mutated, function of the C-terminal secondary host chromatin binding domain is lost (109), indicating that it does not possess standalone binding function. The primary function of the C-terminus is making contact with the viral DNA via the 3 LBSs found in each TR. Grundhoff et al. (110) found specifically that residues 1037 through 1164 were essential for LANA-mediated replication of full-length TR regions. Structures of the N- and C-termini will be discussed in detail in the following sections. Figure 1-5 presents a schematic of the LANA protein and its various domains.

Several studies have suggested that truncated isoforms of LANA are also important. Toptan et al. identified several downstream alternative translation initiation sites, most in the first IR region of the gene (111). These isoforms lack the N-terminal NLS and localize to the cytoplasm, despite the presence of the C-terminal NLS. In addition, they cannot bind chromatin, but rather localize with mitochondrion-containing structures and cytoplasmic ribosomal structures. Zhang

Figure 1-5. Schematic diagram of the LANA protein. A diagram showing the various regions of the LANA protein, including its proline-rich region (P), the internal repeat regions (DE, QEP, QREP, QDE, L, EQE), and the N-terminal nuclear localization signal (indicated by the vertical black bars). Horizontal black bars indicate the locations of the N-terminal histone binding region and the C-terminal region involved in TR DNA binding, LANA self-association, and chromosome association.


N-terminal
Internal sequences
C-terminal
1
63
271
332
444
550
595
769
840
931
1162
LANA
P
DE
QEP
QREP
QDE
L
EQE
5
13
996
1139
H2A/H2B
binding
TR DNA binding
Self-association
Chromosome assoc.

et al. showed these cytoplasmic LANA isoforms bind cGAS, inhibiting cGAS-dependent phosphorylation of the serine/threonine-protein kinase TBK1 and Interferon Regulatory Factor 3 (which induce expression of type I interferons that induce expression of interferon-stimulated genes), and thereby inhibiting cGAS-mediated restriction of lytic replication (113). In 2004, Canham et al. found a LANA isoform that was truncated by 76 amino acids at its C-terminus due to an additional, non-canonical poly-A tail causing a premature stop codon (114). This protein localized to the nucleus, but did not associate with endogenous LANA, localize to nuclear heterochromatin, or bind to TRs, and its function is unknown.

1.6.3 LANA N-terminus and binding to host chromatin

Primary contact of host chromatin is made via the N-terminus. Woodard et al. (115) showed that phosphorylation of residues serine-10, serine-13, and threonine-14 was necessary for effective binding of the N-terminus to host chromatin. In 2002, Krithivas et al. (116) concluded that LANA contacts the host chromatin via two host proteins: methyl-CpG-binding protein 2 (MeCP2), involved in transcriptional silencing of genes in CpG-methylated regions, and the Drosophila Eph kinase (DEK), responsible for inducing supercoiling in circular DNA. Their studies showed MeCP2 binding to the N-terminal chromatin binding domain (CBD) of LANA and DEK binding via C-terminus residues 986 through 1043. In 2006, however, it was found that LANA's N-terminus makes contact with the host DNA directly via a pocket between the host genome's histone H2A and H2B (117). Residues 1 to 23 form a hairpin, stabilized by hydrogen bonds, that

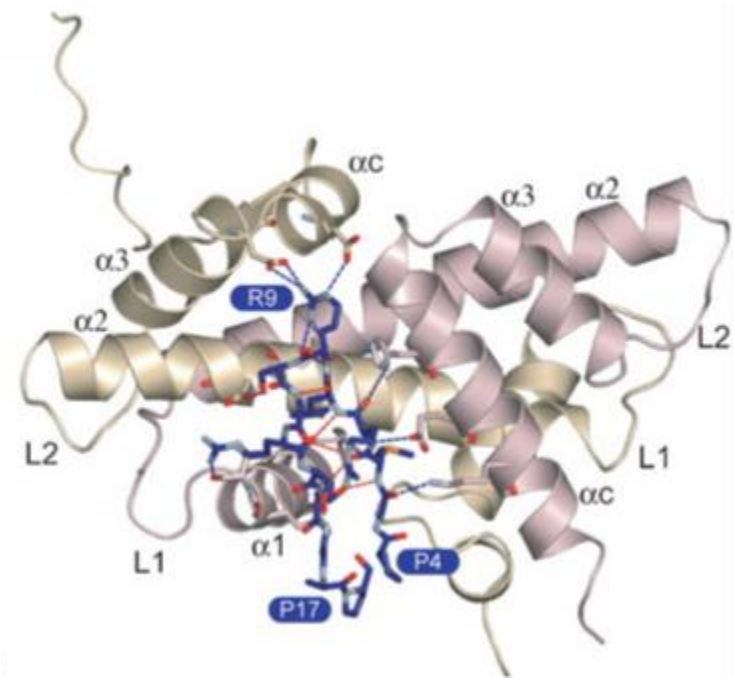
rests in the acidic pocket between helices α1 and αC of H2B, taking the place of the H4 tail (Figure 1-6). The K_D of this interaction has been calculated at 184 +/- 28.5 nM (118). This binding of LANA to the nucleosomal surface promotes self-association among host DNA fibers and formation of a compact 30-nm chromatin conformation (119).

Matsumura et al. (77) accounted for these conflicting models by suggesting a two-step process, in which nucleosome binding through the N-terminal CBD of LANA facilitates subsequent interactions via cellular co-factors, such as MeCP2 and DEK. They further characterized these interactions by mapping binding of MeCP2's methyl-CpG-binding domain, transcription repression domain, and co-repressor-interacting domain with the C-terminal DNA-binding domain of LANA, modulated by a 7-amino acid N-terminal motif in the CBD. They posited that this association with MeCP2 might allow KSHV a method in which to alter expression of both host cell and viral genes.

LANA has also been found to bind to Brd4, a factor in cell growth and RNA Pol II transcription (120). The LANA C-terminal DNA-binding domain (aa. 982-1162) binds to the extraterminal domain of Brd4 (aa. 471-730). There is an additional minor binding site on LANA at amino acids 475 to 777. The functional significance of this binding seems to relate to cyclin-dependent kinase (Cdk) 2 expression. LANA activates the Cdk2 promoter, while binding from Brd4 inhibits the activation. LANA and Brd4 were found to co-localize at mitotic chromosomes,

Figure 1-6. LANA N terminus binding to host chromatin. The LANA N-terminus (blue and red) wedges between helices α1 and αC of H2B (pink) on the surface of H2A (tan), thus mediating host chromosome attachment.

From Science, The Nucleosomal Surface as a Docking Station for Kaposi's Sarcoma Herpesvirus LANA, Barbera A et al., Copyright (2006) (117). Reprinted with permission from AAAS.


αC
α3
α2
α3
R9
α2
L2
L2
L1
αC
α1
L1
P4
P17

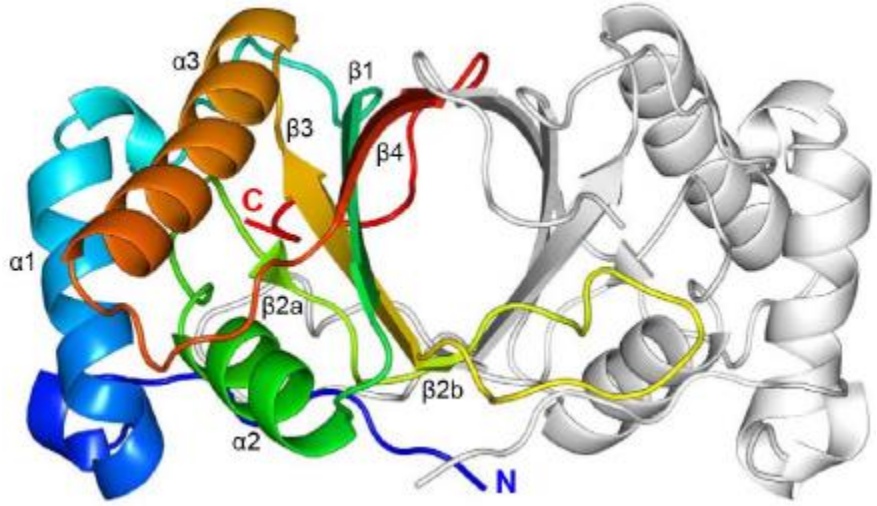
as one would predict from known Brd4 expression patterns.

While tethering to mitotic host chromatin ensures that viral episomes will remain undegraded and nuclear following mitosis, it does not guarantee equal partitioning. In fact, in 2017, Chiu et al. followed viral DNA tagged with a LacO/I system and found that episomes partitioned to daughter cells in an unequal fashion (121). Interestingly, they also observed the clustering of viral genomes in latently-infected PEL cells, determined by the wide spread of fluorescent signals seen coming from KSHV episomes targeted by fluorescence in-situ hybridization. It was suggested that LANA N-termini may indiscriminately attach to histone-wrapped DNA, including histones found on another episome rather than those on nearby host chromatin. This provides one method by which genomes might form the observed clusters.

1.6.4 LANA C-terminus and binding to episome

LANA in its role as tether functions as a dimer. In 2013, three independent studies determined the crystal structure of the C-terminus of this dimer, finding that each monomer contributes 4 beta strands to an intermolecular beta-barrel (122–124). Hydrophobic amino acids project into the core and form a hydrophobic cluster. The beta-barrel is flanked by alpha helices 2 and 3; the alpha helix 1 rests on these and doesn't contact the beta barrel (Figure 1-7). Oligomerization of LANA's C-terminal fragment occurs via the alpha 1 and 3 helices, forming tetrameric and pentameric rings (93). The ability of LANA C-

Figure 1-7. Structure of the LANA C terminus. The crystal structure of LANA's C-terminal domain as it forms a homodimer. Shown is the central beta-barrel flanked by alpha helices 2 (green) and 3 (orange), while alpha helix 1 (blue) rests on these and does not contact the beta-barrel.


α3
β1
β3
β4
C
α1
β2a
β2b
α2
N

terminal dimers to oligomerize was shown to be essential for DNA binding, replication, and episome maintenance, especially the N-terminal-most portion of the LANA DBD, which wraps around the minor groove of DNA. Dimerization of the LANA C-terminus leads to a “basic top” region, which was found to interact via amino acids 1125 – 1129 with the extra-terminal (ET) domain of BRD proteins.

LANA dimers bind to the viral episome via three LANA binding sites (LBS1, 2 and 3) in each 803-bp terminal repeat (61, 125–127). In 1999, Cotter et al. determined that radiolabeled cosmids spanning the end regions of the KSHV genome could be immunoprecipitated with an anti-LANA Ab, suggesting that LANA bound to sequences within the TR region (89). Garber et al. narrowed the binding region down to a 20 base pair sequence (126), of which 14 base pairs are conserved between LBS1 and 2 and 13 between LBS1 and 3 (125). The core recognition sequence of the binding site is located from positions G5 to G10 of the 20 bp binding region and contains a cytosine to guanine substitution at position 8 between LBS1 and both LBS 2 and 3. One research group suggested that this substitution is the main cause in the affinity difference between the sites (125). The LBSs are arranged with LBS1 in the middle and LBS2 and 3 flanking on either side. LBS3 is on the opposite DNA strand as LBS1 and 2 and in the opposite orientation, placing all three LBSs on the same surface of the helical DNA and allowing bound LANA dimers to interact with each other. Binding between LBS1 and 2 is cooperative, as demonstrated by a decreased affinity for

LBS2 in the absence of LBS1 (127). In addition, binding of KSHV LANA to a sequence containing the combinations LBS1-1 or LBS2-2 was greatly decreased over a sequence containing LBS1-2 (93). The central LBS (LBS1) has the highest affinity for LANA, at a K_D of 72 nM, compared with that of LBS2 (6.2 µM) and LBS3 (7.0 µM) (87, 125). KSHV LANA binding to the LBSs is an endothermic reaction, but with a favorable ΔG. The ability of each LANA dimer to carry out its function in replication and transcription is directly proportional to the affinity for its binding site – i.e. LANA bound at LBS1 played a stronger role in biological functions than LANA bound at LBS2 (127). Episomal DNA persisted over replication cycles only in transfections containing complete TRs. This indicates that full unit-length terminal repeats are necessary for replication (94). In addition, Shrestha et al. found that the ~800 bp spacing between TRs was necessary for their function (86).

The centers of the binding sites are 22 base pairs apart center to center, placing them on the same surface of the DNA (B-DNA has approximately 10.5 base pairs per turn). This allows for direct protein-protein interactions between the LANA dimers bound at each site. Ponnusamy et al. determined that this LANA tetramer interface forms an angle of 102 degrees between the two dimers, with a maximum rotation angle between dimers of 24 degrees (93). They also suggested that residue alanine 1121 at the tetramer interface was important for LANA affinity to LBSs, implicating a role of tetramerization in LANA function.

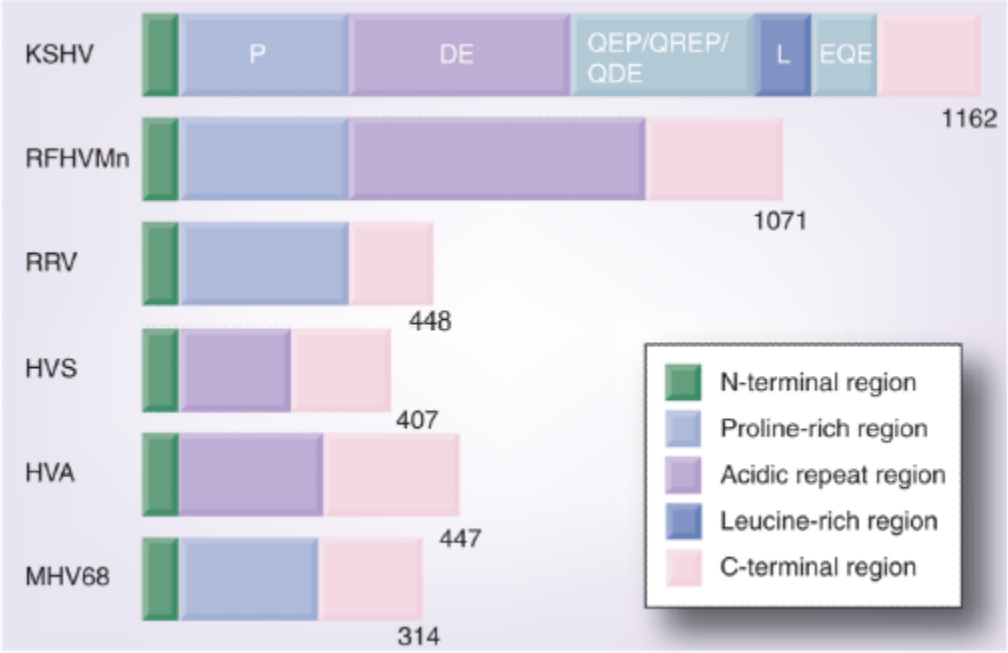
Binding of the LANA C-terminus to LBS1 is asymmetric, which is different from that of the binding seen in EBV's EBNA1 tethering protein and HPV's E2 tethering protein (125). The binding of LANA to the viral episome also creates a bend in the DNA, directed toward the major groove (128). Binding of a single LANA dimer to LBS1 causes a bend of 57 degrees, while cooperative binding of both LBS1 and LBS2 creates a bend of 110 degrees; additional binding to LBS3 confers an even more extensive bend to the DNA (125). It is likely that this distortion plays a role in recruitment of ORC and initiation of DNA replication (74).

In addition to their crystal structure of the LBS-bound LANA C-terminus, Hellert et al. also found a right-handed spiral structure made up of LANA dimers, with their positive lysine side chains in the center of the spiral (125). This positively-charged core was able to bind DNA in a non-sequence-specific manner, demonstrating yet another potential route for LANA oligomerization.

1.6.5 LANA internal repeat region

The acidic IR (internal repeat) region of LANA is the least conserved. In fact, the IR region of the ORF73-encoded protein in different gammaherpesviruses can vary greatly in length, and in some cases it is completely absent from the protein, as in murine herpesvirus 68 (MHV68), rhesus monkey rhadinovirus (RRV), and bovine herpesvirus 4 (BoHV-4) (1). Figure 1-8 compares the ORF73-encoded proteins of several different

Figure 1-8. A comparison of the domains of six different gammaherpesvirus LANA homologs. Schematics of six different gammaherpesvirus ORF73-encoded proteins. Note the variety in protein length, as well as the lack of an acidic repeat region in both rhesus monkey rhadinovirus (RRV) and murine herpesvirus 68 (MHV68) homologs.


KSHV
P
DE
QEP/QREP/
QDE
L
EQE
1162
RFHVMn
1071
RRV
448
HVS
407
HVA
447
MHV68
314
N-terminal region
Proline-rich region
Acidic repeat region
Leucine-rich region
C-terminal region

gammaherpesviruses, demonstrating differences in length and domain content among them. The function of the IR region has not been as widely studied as that of its termini. In 2007, two independent studies concluded that the IR region played a role in immune evasion, antagonizing peptide translation and proteosomal degradation, thereby inhibiting antigen presentation (130, 131). This is similar to findings on the internal repeat region of EBNA1 of EBV; in this system, the interspersed alanine residues in the repeats create a hydrophobic surface that may interfere with hydrophobic interactions between a proteasome subunit and the protease (132). In KSHV LANA, the key region responsible for immune evasion was found to be the junction between the second and third IR regions, regardless of whether it was located 3' or 5' of the reporter gene (131). In 1999, Gao et al. determined that the IR region of LANA varies in length depending upon the viral strain, giving DNA lengths between 2876 and 4023 bp and protein sizes between 200 and 258 kDa (133). This polymorphism was generally consistent within each KSHV isolate, including from multiple lesions at different sites on the same patient. The length of the LANA IR region was stable long-term in culture and did not correlate with EBV coinfection, number of copies of the viral genome per cell, or ability of the virus to induce via TPA treatment. The IR region of LANA may also play a role in its lytic reactivation properties, as the MHV68 LANA homolog, which lacks the IR region, is unable to repress the ORF50 promoter (134). In the same, study, however, KSHV and MHV68 LANAs

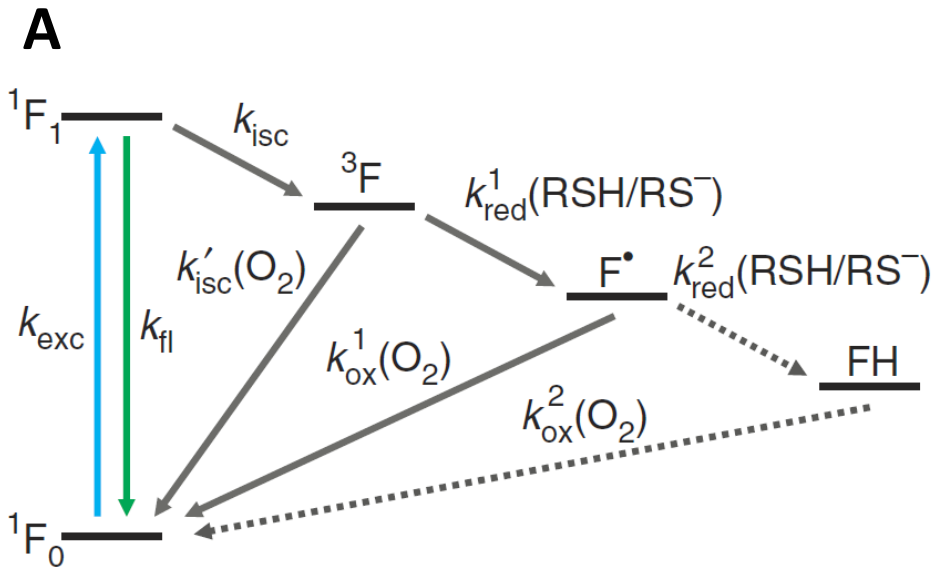
were comparable in their ability to tether episomes, perhaps implying that the IR region is less critical for this function.

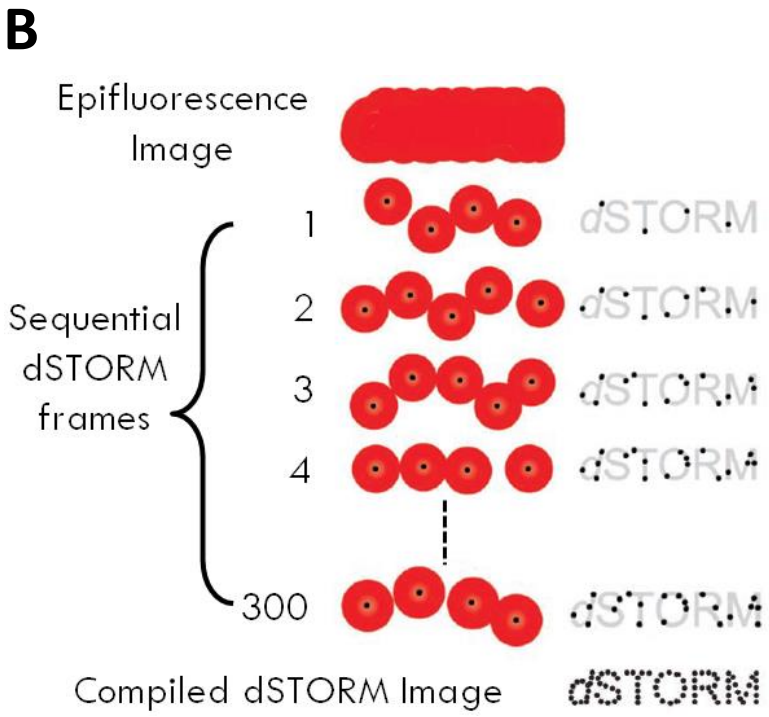
1.7 Direct stochastic optical reconstruction microscopy (dSTORM)

Distinguishing between two neighboring terminal repeat regions from each other requires a resolution of 60-90 nm. The resolution of conventional light microscopy is limited to roughly one-half the wavelength of light being used (Abbe limit). For the visible light spectrum (roughly 400 to 700 nm), this results in a resolution limit of, at best, 200 nm. Electron microscopy has the necessary resolution but is lacking in three-dimensional information as well as the contrast needed to distinguish individual molecules.

As a result, this study implements the super-resolution imaging technique known as direct stochastic optical reconstruction microscopy, or dSTORM. Upon excitation with a wavelength-specific laser, fluorophores are excited into their higher-energy singlet state (135). From there, the fluorophore can either remain in the singlet state, continuing the cycle of excitation and emission until it is bleached, or undergo intersystem crossing, in which the spin of one of its electrons flips, and the molecule crosses over into a dark (non-fluorescing) triplet state (Figure 1-9A). The fluorophore then interacts with an electron-donating primary thiol present in the mounting medium, creating a dye radical that can only be returned to the ground singlet state upon interaction with molecular oxygen (135). An oxygen-scavenging mounting medium (such as a mixture of

Figure 1-9. Principles behind direct stochastic optical reconstruction microscopy (dSTORM). (A) A diagram showing the energy levels available to a fluorescent dye molecule. In dSTORM, the majority of fluorophores are trapped in the three right-most (triplet state) energy levels via a reducing, oxygen-scavenging environment. This minimizes the number of fluorescent-capable fluorophores in the left-most singlet states, thus enabling single molecule detection. **(B)** A schematic illustrating the acquisition of 300 sequential dSTORM frames, the results of which are then compiled to produce an image with enhanced resolution over the epifluoresence image.


A
$^1F_1$
k_{isc}
^3F
$k^1_{red}(RSH/RS^-)$
$F^\bullet$
$k^2_{red}(RSH/RS^-)$
$k'_{isc}(O_2)$
k_{exc}
k_{fl}
$k^1_{ox}(O_2)$
FH
$k^2_{ox}(O_2)$
$^1F_0$


B
Epifluorescence Image
Sequential dSTORM frames
1
2
3
4
300
dSTORM
Compiled dSTORM Image

glucose, glucose oxidase, and catalase) traps most of the available oxygen, making this return to the fluorescent singlet state a rare event. Even in the presence of molecular oxygen, some fluorophore radicals are remarkably stable, and can have a lifetime of several hours (136). Dempsey et al. found that upon conversion to the dark state, cyanine dyes in particular (such as the Alexa Fluor 647 used in our studies) form a mass spectrometry compound with the sum molecular weight of both the dye radical and the primary thiol present in the mounting medium (137). They thus determined that a chemical reaction driven by the activating laser light results in the formation of a non-fluorescent thiol-dye adduct, wherein the thiol is attached to the centrally-located polymethine bridge of the dye molecule.

In this system, only a few molecules at a time are present in the photo-active singlet state, while most are trapped in the triplet state and are inactive. Wavelength-specific laser stimulation causes excitation for these few photo-active molecules, allowing for fluorescence to occur. The result is a system in which only a few well-separated molecules fluoresce at a time. The point-spread function of each fluorescent “flash” can be fitted to a 2D Gaussian function, thereby approximating the center of the fluorescence and, thus, the point source of the emission. Multiple frames of these intermittent “flashes” can be superimposed, reconstructing the layout at sub-diffraction resolution in two dimensions (Figure 1-9B). Localization precision is proportional to the number of fluorescent photons collected and is generally within the 10-20 nm range. This is

a 10-fold improvement over the conventional microscopy limit and well within the requirements to distinguish between two adjacent terminal repeats.

While this method offers improved resolution in two dimensions, it does not address the issue of a third dimension. This problem is solved with the introduction of astigmatism-based 3D localization (138). A cylindrical lens placed in the path of the acquisition beam creates an elliptical distortion in the point-spread function in the x or y axis. Both the axis and the extent of the distortion depend upon the source's position on the z axis. A standard curve is experimentally derived using cylindrical beads. Sample data are then back-fitted to the curve to determine the z axis position of individual point sources.

CHAPTER 2. SUPER-RESOLUTION MICROSCOPY REVEALS STRUCTURAL MECHANISMS DRIVING THE NANOARCHITECTURE OF A VIRAL CHROMATIN TETHER

This chapter is modified from: Grant MJ, Loftus MS, Stoja AP, Kedes DH, Smith MM. Superresolution Microscopy Reveals Structural Mechanisms Driving the Nanoarchitecture of a Viral Chromatin Tether. PNAS. 2018 Apr 2. PMID: 29610353 | https://doi.org/10.1073/pnas.1721638115

2.1 ABSTRACT

By tethering their circular genomes (episomes) to host chromatin, DNA tumor viruses ensure retention and segregation of their genetic material during cell divisions. Despite functional genetic and crystallographic studies, there is little information addressing the three-dimensional structure of these tethers in cells, issues critical for understanding persistent infection by these viruses. Here, we have applied direct stochastic optical reconstruction microscopy (dSTORM) to establish the nanoarchitecture of tethers within cells latently infected with the oncogenic human pathogen, Kaposi's sarcoma-associated herpesvirus (KSHV). Each KSHV tether comprises a series of homodimers of the latency-associated nuclear antigen (LANA) that bind with their C-termini to the tandem array of episomal terminal repeats (TRs) and with their N-termini to host chromatin. Super-resolution imaging revealed that individual KSHV tethers possess similar overall dimensions and, in aggregate, fold to occupy the volume of a prolate ellipsoid. Using plasmids with increasing numbers of TRs, we found that tethers display polymer power-law scaling behavior with a scaling exponent characteristic of active chromatin. For plasmids containing a two-TR tether, we determined the size, separation, and relative orientation of two distinct clusters of bound LANA, each corresponding to a single TR. From these data, we have generated a three-dimensional model of the episomal half of the tether that integrates and extends previously established findings from epifluorescent, crystallographic, and epigenetic approaches. Our findings also validate the use of

dSTORM in establishing novel structural insights into the physical basis of molecular connections linking host and pathogen genomes.

2.2 INTRODUCTION

Initial epifluorescence studies identified LANA as nuclear punctae recognized by serum antibodies from KSHV infected patients (29, 30, 139, 140). Ten years later, Adang et al. used a combination of flow cytometry and qPCR to demonstrate a direct proportionality between the number of LANA punctae and the amount of viral DNA, leading to the conclusion that each nuclear dot represented a single viral episome (141). Previous studies examining these punctae showed association of LANA with mitotic chromosomes at the resolution of epifluorescence microscopy (88, 89). Kelley-Clarke et al. later suggested preferential localization of these LANA punctae to centromeric and telomeric regions on metaphase chromosomes (142). Such studies have provided a solid foundation for further inquiry into the nature of this tethering mechanism.

Many fundamental features of full-length tethers in cells have remained elusive due to the resolution constraints of epifluorescence microscopy. While X-ray crystallography has resolved the structures of an N-terminal 23 amino acid LANA peptide bound to nucleosomes, and a C-terminal 139 amino acid peptide in complex with LBS1 (117, 125), questions remain as to the architectural features of a full-length tether. It is unknown how TR chromatin folds and whether this is consistent among tethers, both within and across cell lines. The close

packing of nucleosomes and presence of heterochromatin protein 1 (HP1) at TRs support heterochromatin conformation, whereas epigenetic, replication, and transcription studies suggest active chromatin (71, 72, 76, 127, 143, 144). Questions also remain as to how LANA is distributed across the TRs, whether it attains full occupancy on all three LBSs, and whether supernumerary LANA dimers form large complexes at each TR to entrap episomal DNA loops (128). There is also no information indicating whether the number of TRs bound by LANA reaches a plateau. LANA dimers might bind to a maximum number of TRs in a plasmid or episome, regardless of the number of total TRs available. Finally, it is not clear how all of these elements influence tether folding and the ability to effectively bind to both viral and host genomes. To address these questions, we applied super-resolution microscopy to study KSHV infected cells with the goal of generating architectural data on full-length KSHV tethers.

2.3 MATERIALS AND METHODS

2.3.1 Cell lines

BCBL-1 cells have been in the laboratory of D.H.K. who was a co-author on the study first describing the line (145). BJAB cells were a gift from Don Ganem (Novartis Institutes for Biomedical Research) and are described by Menezes et al. (146). SLKp/Caki-1p (147) cells were a gift from Adam Grundhoff (Heinrich Pette Institute, Leibniz Institute for Experimental Virology, Hamburg, Germany) and Don Ganem and are described by Grundhoff and Ganem (148). SLKp cells were selected despite their inclusion in the ICLAC database of

commonly misidentified cell lines because of their utility as an epithelial, non-lymphocytic line and their stable infection with BCBL-1-derived virus. iSLK-BAC16 cells were a gift from Rolf Renne (University of Florida) and are described by Brulois et al. (149). COS-7 cells were a gift from Jim Casanova (University of Virginia). BJAB and BCBL-1 cells had typical lymphocyte non-adherent morphology and were CD45 positive by immunofluorescence staining. BCBL-1 cells were syndecan-1 (CD138) positive by immunostaining and flow cytometry (141). BCBL-1, SLKp, and iSLK-BAC16 cells were LANA positive by immunofluorescence staining; BJAB and COS-7 cells were LANA negative.

Adherent cell lines were maintained in Dulbecco's modified Eagle's medium (Gibco) supplemented with 10% fetal bovine serum (Gibco). They were harvested from culture dishes with 0.05% trypsin after a wash with 1X phosphate buffered saline. Non-adherent cell lines were maintained in RPMI 1640 (Gibco) supplemented with 10% fetal bovine serum, 20 mM HEPES, 2 mM L-glutamine, 0.05 mM beta-mercaptoethanol, and 200 µg/mL sodium bicarbonate. All experimental conditions for imaging took place in identical media but lacking phenol red. All cell lines were maintained at 37^{o}C in 5% CO_2 and tested negative for mycoplasma.

2.3.2 Plasmids

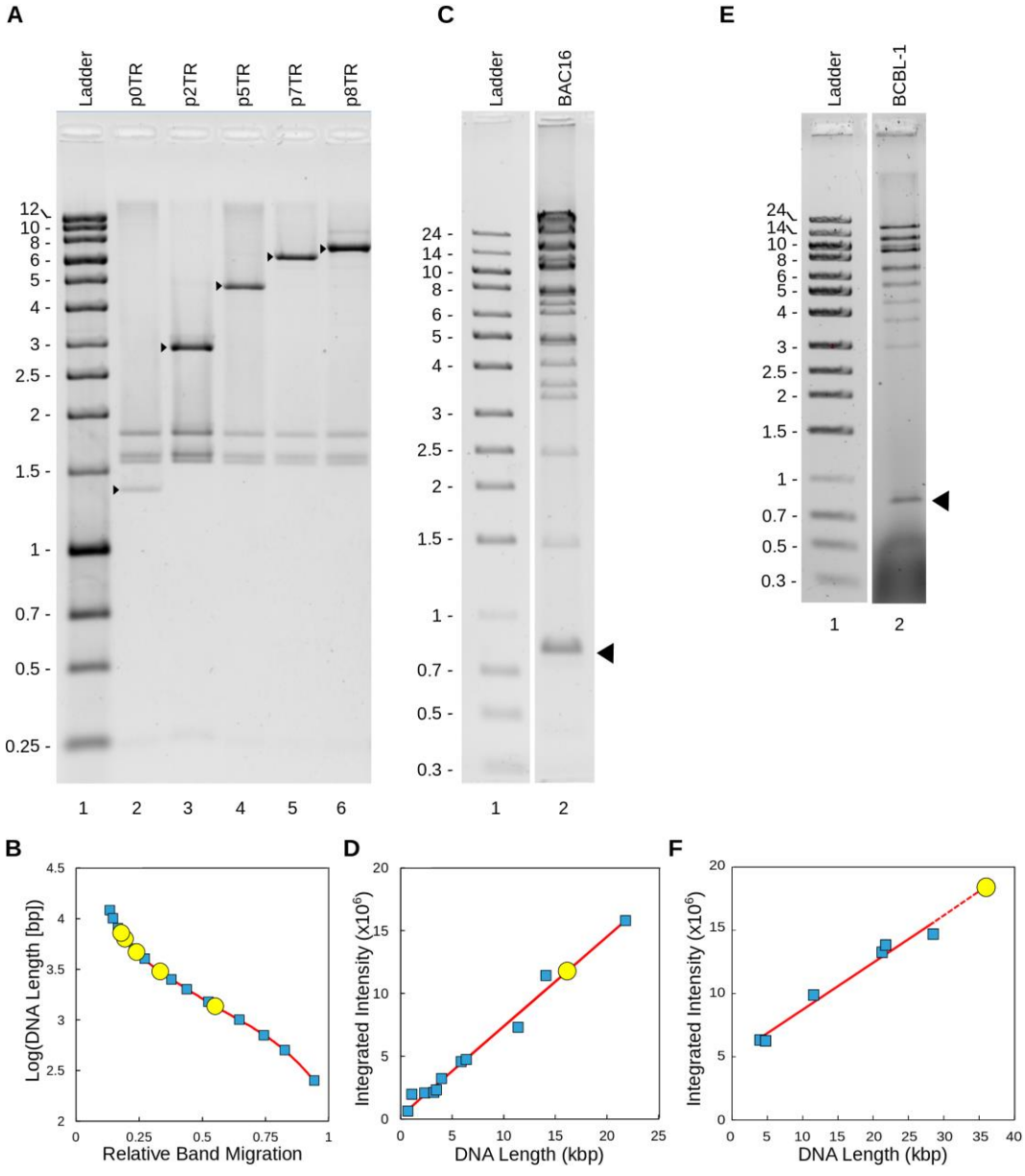
The pcDNA3-LANA plasmid was a gift from Rolf Renne (University of Florida) and is described by Renne et al. (150). BAC16 was a gift from Jae Jung (University of Southern California) and is described by Brulois et al. (149). The

p8TR plasmid is the 8TR-containing Z6 cosmid subcloned into modified pREP9, as described by Ballestas et al. (88). The plasmid, produced by the lab of Kenneth Kaye (Harvard Medical School), was a gift from Paul Lieberman (The Wistar Institute). Plasmids containing 0, 2, 5, and 7 terminal repeats were created by transforming DH5α E. coli (Thermo Fisher Scientific) with the p8TR plasmid and screening by PstI restriction endonuclease digestion for colonies that showed recombination within the TR region and gave rise to a lower number of TRs. Insert size was confirmed by gel electrophoresis, as shown in Figure 2-1A and B. The plasmids isolated from these E. coli were subsequently grown up in SURE E. coli (Agilent Technologies) to preserve fidelity of the new TR number.

2.3.3 Determining TR number in BAC16 and BCBL-1 virus

Linear viral DNA was obtained from BCBL-1 cells using the method described by Lagunoff et al. (61). BCBL-1-derived KSHV DNA and BAC16 DNA were digested with NotI + SpeI. The resulting digests were run on an agarose gel, which was then stained with ethidium bromide. We then plotted band intensity as a function of DNA length for each band present in a single copy per genome. We used this line to calculate the amount of DNA represented by the 801-bp band. For BAC16, this band gave a result of 16.2 kbp, which accounts for 20 TRs Figure 2-1C and D. Because the BAC16 is circular, one TR is not accounted for by this method, yielding a final calculation of 21 TRs. For the BCBL-1-derived viral DNA shown in Figure 2-1E and F, the 801-bp band was

Figure 2-1. Number of TRs in each pNTR plasmid and BAC16. (A) Plasmids with variable (*N*) numbers of TRs (pNTR) were digested with PstI, which cuts on either side of the TR-containing insert. Length of the released fragment (arrowheads) allowed determination of TR number. Lane 1, ladder (kbp), lane 2, plasmid lacking TRs, lanes 3-6, plasmids with 2, 5, 7, and 8 TRs, respectively. **(B)** Relative mobilities of DNA standards (blue squares) were used to construct a best-fit curve for the log of DNA length as a function of migration (red line). Lengths of PstI fragments from plasmids with variable numbers of TRs (yellow circles) were determined from the standard curve based on their migration. **(C)** Lane 1, ladder (kbp). Lane 2, NotI + SpeI digest of BAC16 isolating the 801-bp TR fragments (arrowhead). **(D)** The number of TRs within BAC16 was determined from the intensity of ethidium bromide staining. Fragments present once per BAC (blue squares) were used to determine a best-fit line (red) for band intensity as a function of DNA length (R^2 = 0.98). Intensity of the TR-containing fragment (yellow circle) was then mapped to this standard curve, yielding a calculation of 21 TRs per BAC16 genome. **(E)** Lane 1, ladder (kbp). Lane 2, NotI + SpeI digest of BCBL-1-derived KSHV DNA, isolating the 801-bp TR fragments (arrowhead). **(F)** The number of TRs in BCBL-1-derived virus was determined as in D and was calculated to be 45 (R^2 = 0.97) for the experiment shown. This experiment was performed in triplicate, yielding a final result of 43 +/- 6 TRs per BCBL-1 episome.


A
Ladder
p0TR
p2TR
p5TR
p7TR
p8TR
12
10
8
6
5
4
3
2.5
2
1.5
1
0.7
0.5
0.25
1
2
3
4
5
6
C
Ladder
BAC16
24
14
10
8
6
5
4
3
2.5
2
1.5
1
0.7
0.5
0.3
1
2
E
Ladder
BCBL-1
24
14
10
8
6
5
4
3
2.5
2
1.5
1
0.7
0.5
0.3
1
2
B
Log(DNA Length [bp])
Relative Band Migration
D
Integrated Intensity (x10^6)
DNA Length (kbp)
F
Integrated Intensity (x10^6)
DNA Length (kbp)

representative of 36.0 kbp, which accounts for 45 TRs. Because this DNA is linear, two TRs are not accounted for by this method, yielding a final result of 47 TRs. For BCBL-1-derived KSHV, this experiment was performed in triplicate, and the resulting number of TRs was 43 +/-6.

2.3.4 Transfections

$1x10^7$ BJAB cells were transfected with 1µg pNTR plasmid + 1µg pcDNA3-LANA plasmid using Amaxa Nucleofector (Lonza) with Solution V on program T-020. Cells were harvested for analysis at 48 hours post-transfection.

2.3.5 Infection of BJABs with BAC16 virus

Infection of BJABs was performed as described by Plaisance-Bonstaff (151). In brief, iSLK-BAC16 cells were seeded at $3x10^5$/mL in a 6-well plate. After 24 hours, they were induced with doxycycline (1 µg/mL) and sodium butyrate (1 mM). 48 hours later, the media was replaced with 3 mLs RPMI containing BJAB cells at a density of $3x10^5$/mL. Four days later, BJABs were removed from the iSLK-BAC16 cells and cultured for 14 days with hygromycin selection (100 µg/mL) before staining. A separate, identically treated well of iSLK-BAC16 cells without BJABs was cultured for 7 days, at which point all induced iSLK-BAC16s were found to be dead by Trypan blue exclusion.

2.3.6 Antibodies

The anti-LANA antibody LN53 (rat, monoclonal) (152) was conjugated using an Alexa Fluor® 647 succinimidyl ester kit (Thermo Fisher Scientific). The approximate number of dye molecules per antibody (three) was calculated according to the manufacturer's instructions.

2.3.7 dSTORM sample preparation

Cells were fixed in 4% paraformaldehyde for 10 minutes. After a 20-minute rat IgG block at 4 µg/mL, Alexa Fluor®-conjugated LN53 antibody was applied at a concentration of 1:400 in BD Perm/Wash (BD Biosciences) and incubated for 40 minutes. Cells were washed in Perm/Wash and DAPI was then applied at 0.5 µg/mL for 15 minutes. Coverslips (Cat. #72204-01, Electron Microscopy Sciences) were cleaned with alkaline peroxide. Their centers were masked off and TetraSpeck 0.1 µm beads (Thermo Fisher Scientific) were applied to the rest of the surface; these were attached to the slide at 100^{o}C for 30 min. The center masks were removed and the coverslips were coated in poly-L-lysine for 1 hour. The poly-L-lysine solution was aspirated and the coverslips were air-dried overnight. Cells were deposited (Cytospin 2, Shandon) onto the centers of these prepared coverslips for 4 minutes at 500 rpm. Samples for dSTORM analysis were mounted in a solution of 0.22 mg/mL glucose oxidase, 40 µg/mL catalase, 0.14M beta-mercaptoethanol, and 0.55M glucose.

2.3.8 Instrumentation

dSTORM data were collected using custom instrumentation based on an Olympus IX81-ZDC2 inverted microscope (Olympus America Inc., Melville, NY, USA) configured in a standard geometry for localization microscopy (153, 154). The illumination beam from a 643 nm diode laser (DL-640-100, CrystaLaser, Reno, NV, USA), coupled to a laser clean-up filter (ZET640/20X, Chroma Technology Corp., Rockingham, VT, USA) and a 5X Galilean beam expander (BEO5M, Thorlabs, Inc., Newton, NJ, USA), was focused on the backplane of a 60X, 1.2NA, water-immersion objective (UPSLAPO60XW, Olympus) by means of a f+350 mm lens (Thorlabs) to yield an illumination profile with a $1/e^2$ radius of 34 µm. Ground state depletion and dSTORM readout of Alexa Fluor® 647 dye molecules was achieved with a laser intensity of 1.9 kW/cm^2 as measured at the sample. Single molecule emissions were collected through the same objective using a filter set comprising a dichroic mirror (Di01-R635, Semrock, Rochester, NY, USA) and emission filter (FF01-692/40, Semrock). The image was expanded using f+75 mm and f+150 mm lenses (Thorlabs) arranged as a ~2X telescope and acquired with an EMCCD camera (iXonEM DU-897D-CS0-BV, Andor Technology PLC, South Windsor, CT, USA) to yield an image pixel size of 0.134 µm (155). Astigmatism for axial localization was introduced using an f+1000 mm cylindrical lens (LJ1516RM-A, Thorlabs) placed in the acquisition path (138).

2.3.9 dSTORM data acquisition

dSTORM images were acquired using a water-immersion objective to minimize axial compression of the data for LANA tethers located far from the coverslip (156). Because of the sensitivity of water-immersion lenses to the tilt of the coverslip (157), axial localization was calibrated for each sample by imaging point-source fluorescent beads attached to the same coverslip as the samples being studied. Images of the astigmatic point spread function were acquired every 50 nm in the axial dimension, spanning 2 µm across the focal plane, controlled by a 3-axis nanopositioning stage (Nano-LPS100, Mad City Labs Inc., Madison, WI, USA). dSTORM images were collected at a frame rate of 32.4 Hz, for a total of 15,000 to 25,000 frames, using an EM gain of 100 and pre-amplifier gain of 5X on the iXon camera. Transmitted light images were collected every 200 frames to provide subsequent correction for lateral drift during analysis, while axial positioning was stabilized by the ZDC2 instrumentation of the microscope. An epifluorescence image of DAPI staining was saved after the dSTORM data acquisition and the few nuclei exhibiting chromosome condensation were excluded from further analysis.

2.3.10 dSTORM data processing

dSTORM images were processed using a pipeline of custom python modules, implementing standard single molecule localization algorithms (153). In-house software tools utilized numpy (version 1.11.2) and scipy (version 0.18.1) scientific computing libraries (158, 159). Candidate dye emissions were

identified by size and photon threshold criteria and localized within an 11 x 11 pixel region of interest. Each point spread function (PSF) was evaluated by fitting with a two-dimensional Gaussian ellipse by least squares to estimate parameters for the x and y center coordinates, amplitude, major and minor axes radii, and an offset. The lateral localization precision was calculated for each emission as described by Thompson et al. (155). For each emission, photon yield was calculated as the product of the Gaussian amplitude and area, while the background pixel noise was estimated as the standard deviation of intensities from the perimeter pixels of the region of interest. The axial location of a molecule was determined using its major and minor axis radii from the 2D Gaussian fit to interrogate the standard curve established from the point source fluorescent beads (138, 160) and its axial localization precision was calculated as described by DeSantis et al. (160). The transmitted light images obtained through dSTORM acquisition were used to determine lateral image drift using a sub-pixel cross-correlation algorithm (161) and dSTORM molecule coordinates were corrected by interpolation.

2.3.11 Determination of emission counts, R_g, and statistical treatments

Emission for tethers with 2, 5, 7, 8, and 21 TRs were acquired under identical conditions and quantified using frames 1000 to 20,000 of the data sets. These conditions could not be extend to larger tethers such as KSHV episomes in BCBL-1 and SLKp with over 40 TRs, due to the longer times required to achieve ground state depletion and spatial separation of single emissions.

LANA tethers were characterized by their radius of gyration, calculated as the root mean squared vectors of the emissions using the standard representation,

$$R_g^2 = \frac{1}{N}\sum_{i=1}^{N}(r_i - \overline{r})^2$$

where r_i is the location of individual emissions and $\overline{r}$ is the centroid of all the emissions in the tether. The coordinates of $\overline{r}$ were calculated as the average of the coordinates for all emissions weighted by the inverse of their respective lateral and axial localization precision.

Summary and comparative statistics were based on nonparametric approaches. The 95% confidence intervals for median estimates were calculated by bootstrap resampling. The similarity of the BCBL-1 and SLKp tethers were evaluated by Mood's Median test ($\alpha = 0.05$). Power curve analysis by bootstrap resampling of the BCBL-1 (N=65) and SLKp (N=34) populations revealed that, with a statistical power of 0.8, Mood's Median was expected to reject the null hypothesis of equality if the SLKp tether median R_g values were $\leq$ 68 nm (20%) smaller or $\geq$ 64 nm (19%) larger than BCBL-1.

2.3.12 LANA tether analysis

The software tools used to analyze and visualize the LANA tethers were supported by the additional Python libraries scikit-learn (version 0.17.1) (162), statsmodels (version 0.6.1) (163), Geometric Tools (version wm5) (164),

matplotlib (version 1.5.3) (165), and mayavi.mlab (version 4.5.1) (166). The dSTORM signals for individual LANA tethers were isolated from the data sets and further processed for analysis. Tethers were chosen to be well separated to avoid potential overlap within a subset of episome clusters (121). The axial localization precision for each emission was re-evaluated by calculating the individual axial position 1000 times using values of the major and minor PSF widths randomly drawn from their respective error distributions. The larger of this Monte Carlo estimate and the original analytical calculation was saved as the working axial localization precision and emissions with values greater than +/- 60 nm were discarded.

To determine aggregate parameters individual LANA tethers were uniformly oriented by geometry-based principle component analysis and aligned at their centers of mass. The dimensions of the aggregate structures along the X, Y, and Z axes were estimated as the coordinates spanning quantile probabilities from 2.5% to 97.5% of the distribution along each axis. Solid 3D representations of tether emissions are depicted by spheres with a radius equal to their median axial localization precision. Similarly, tether images are rendered as Gaussian spots scaled to a width that is twice the median axial localization precision.

Individual 2TR LANA tethers were analyzed further by dividing their emissions into two groups using k-means clustering. The midpoint between the centers of the two clusters was translated to the zero origin and the complete tether was then oriented to place the two cluster centers along the x-axis on

either side of the origin. The tether was then rotated about the x-axis to achieve the greatest variance in the y-dimension. Vectors fit through the long dimensions of the two clusters defined a ”V” geometry, and each tether was then reflected horizontally or vertically or both such that the wider separation was oriented towards the positive y-axis and greatest tilt from vertical was oriented towards the positive x-axis. The complete set of 2TR LANA tethers were then superimposed at their origins to evaluate the parameters of the aggregate structure.

2.3.13 Modeling 2TR LANA tethers

The software written to model 2TR LANA tethers utilized additional python modules from the programming libraries Biopython (version 1.68) (167, 168) and PeptideBuilder (version 1.0.1) (169), and the application PyMOL (version 1.7.6) (170).

The X-ray crystal structure of a dimer of LANA C-terminal DNA binding domains in complex with LBS1 DNA suggests that the upstream LANA peptides may enter on either side of the episomal DNA (125) (PDB 4UZB). Therefore, in modeling LANA dimers, we placed the LN53 antibody epitopes on the N-terminal side of a coiled-coil domain which is then connected by a domain of unknown structure to the dimer of C-terminal DNA binding domains. This model is similar to a more general outline proposed previously (125). We modeled domains of unknown structure based on psi and phi angles from their predicted secondary structure. These were done solely to provide plausible constraints on the

placement of the EQEQ epitopes relative to the LBS, and with the exception of the coiled-coil model (171) and C-terminal DNA binding domains with known crystal structures, the models are not expected to match actual LANA protein folding.

Complete sets of 2TR models were constructed in which the TR1 and TR2 tetrasomes were arranged over 9 translational positions each. The critical parameter for fitting the models to the dSTORM data is the phase of the tetranucleosome in TR2 because it is located between the two sets of LBS DNA sequences. While the phase of the tetranucleosome in TR1 is presumably important for the connection to flanking chromatin, its phase does not affect the structural relationship between the LANA dimers in the two TRs. Therefore, for the 2-dimer models, we fit each of the 28 p2TR datasets with the 9 chromatin models that combined TR1 phase 3 with TR2 phases 1 through 9. For the 3-dimer models, we fit the same 28 p2TR datasets with the 6 chromatin models that combined TR1 phase 6 with TR2 phases 4 through 9.

2TR models were fit to the dSTORM data using a limited memory approximation of the Broyden-Fletcher-Goldfarb-Shanno algorithm for bound constrained minimization (172) as implemented in SciPy. Each of the two clusters of dSTORM emissions for individual tethers was further divided into either two or three clusters by k-means to yield either four or six centroids per tether, for the 2-dimer and three-dimer models respectively. To begin optimization, the center of mass of the epitope binding sites in a model was

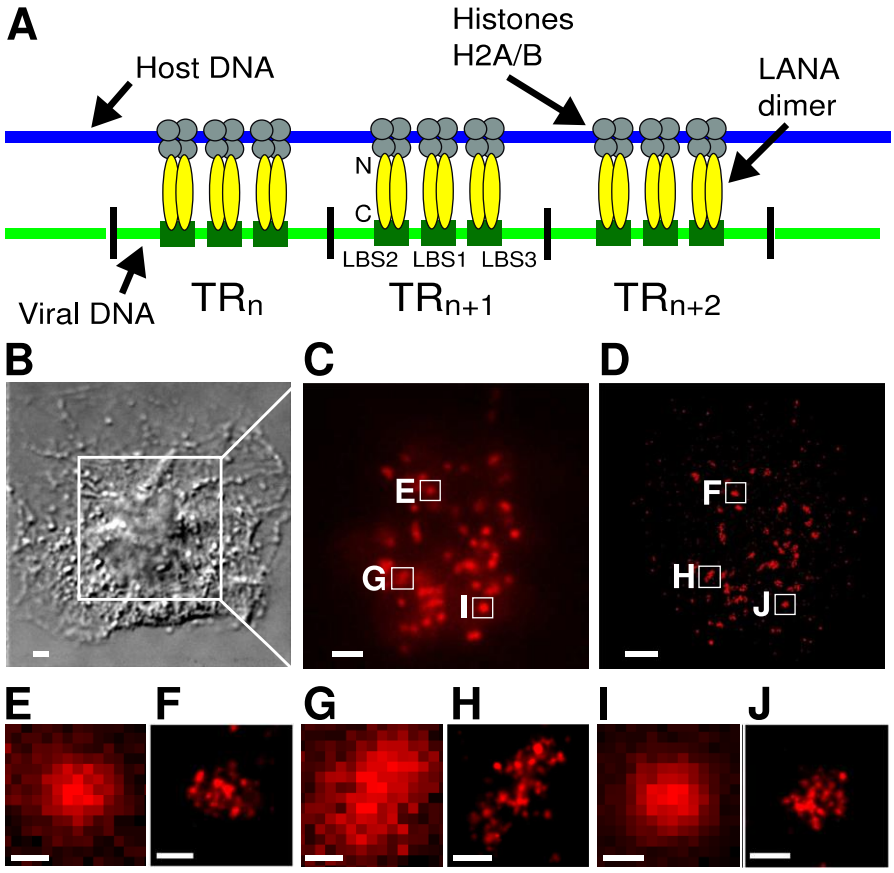
aligned with the global center of mass of the cluster centroids in the dSTORM data. The initial rotational orientation of the model was randomized. Eighteen parameters were then varied to minimize the Euclidean distance between the model epitope sites and their respective dSTORM cluster centers. For fitting, the model was allowed three dimensions of rotation and three dimensions of translation. The four or six epitope sites in the models were treated as linear vectors of a fixed 51 nm in length originating at the LBS sites and they were allowed to rotate freely but constrained to a maximum 45^{o} tilt, or less, from perpendicular to the DNA. The best fit for each model was then evaluated for its ability to account for the dSTORM emissions. The Gaussian probability volumes of the dSTORM emissions were scaled by 2.2 times their lateral and axial localization precision. Similarly, the model epitope site probability volumes were scaled by 2.2 times the median axial and lateral localizations of their associated dSTORM emissions. The fit was then scored for the fraction of the dSTORM ellipsoids intersecting with the epitope ellipsoids of the model.

2.4 RESULTS

2.4.1 The nanoscale dimensions of LANA tethers are uniform and consistent across different cell types

In principle, since LANA binds both host and episome DNA (Figure 2-2A), the packaging of TR chromatin could be determined primarily by the

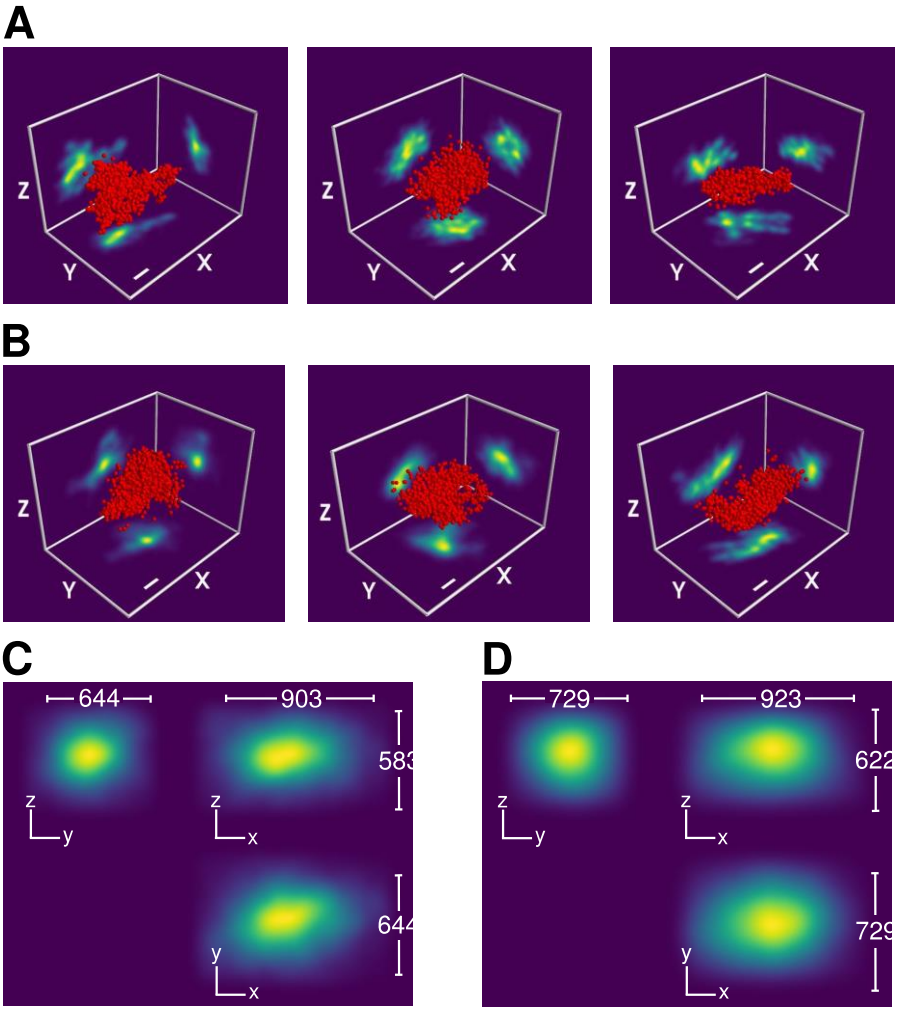
Figure 2-2. dSTORM offers improved image resolution of LANA tethers. (A) Schematic depiction of the anatomy of a KSHV LANA tether. LANA dimers (shown in yellow) bind via their N termini to histones H2A/B (grey). The C termini bind with sequence specificity to three LANA binding sites (LBSs) located on each terminal repeat (TR). **(B)** An anti-LANA stained BCBL-1 cell imaged by differential interference contrast microscopy. **(C)** The boxed region in **(B)** viewed by conventional epifluorescence microscopy. **(D)** The same field of view in **(C)** imaged by dSTORM microscopy. (Scale bars for **B-D** = 2 μm). **(E, G, and I)** The three LANA tethers boxed in **(C)** are enlarged for better visualization. **(F, H, and J)** The three LANA tethers boxed in **(D)** are enlarged for comparison with their paired standard epifluorescence images **(E, G, and I)**. (Scale bars for **E-J** = 250 nm).


A
Host DNA
Histones
H2A/B
LANA
dimer
N
C
LBS2 LBS1 LBS3
Viral DNA
TRn
TRn+1
TRn+2
B
C
D
E
G
I
F
H
J
E
F
G
H
I
J

state of the host chromatin, by features intrinsic to the viral episome, or both. To distinguish among these alternatives, we examined the tether morphology of the same viral isolate in two different cell types. We found that dSTORM resolved each epifluorescent LANA dot into a cluster of individual emissions from fluorophore-labeled anti-LANA mAbs (Figure 2-2B-J). We interpret these emissions as blinks within a Gaussian localization volume from each conjugated dye molecule present at approximately three dye molecules per mAb (see Section 2.3.6), up to six mAbs bound per LANA dimer (see discussion of coiled-coil domain, below), and up to three LANA dimers per TR. Since all the episomes from the same virus have the same number of TRs (84, 85), and each TR is 801 bp, we asked whether the polymer architectures revealed by dSTORM would be similarly consistent across multiple tethers. We examined 65 individual LANA tethers from BCBL-1 cells (a KSHV-positive PEL line) and 34 from SLKp/Caki-1p cells, a human epithelial line exogenously infected with KSHV derived from BCBL-1 (147, 148). We found that while each tether presented a unique shape, their overall dimensions were similar in the two cell lines (Figure 2-3A and B).

To quantify this similarity, we aligned clusters along their three principal axes and calculated their dimensions and radii of gyration, R_g (see Section 2.3.11) (Figure 2-3C and D). The median R_g for tethers in BCBL1 cells was 341 nm, 95% CI [309-353] and in SLKp cells it was 368 nm, 95% CI [323, 405]. There was no statistical difference between the R_g of BCBL-1 and SLKp tethers

Figure 2-3. KSHV LANA tethers have consistent dimensions within and between cell types. (A) Three-dimensional projections of three representative LANA tethers from asynchronous BCBL-1 cells. Red spheres depict the Cartesian coordinates of individual photon emissions, with their radii scaled to the median axial localization precision of 30 nm. **(B)** Three representative LANA tethers from asynchronous SLKp cells, shown as in **(A)**. (Scale bars for **A and B** = 250 nm.) **(C)** LANA tethers from BCBL-1 cells (N=65) were oriented by principal axes and aligned at their centroids. The compilation is rendered as Gaussian emissions with 30 nm localization precision and presented from three architectural viewpoints, indicated by the X, Y, and Z directional axes. Median dimensions are shown by the bars. **(D)** LANA tethers from SLKp cells (N=34) superimposed and rendered as in **(C)**.

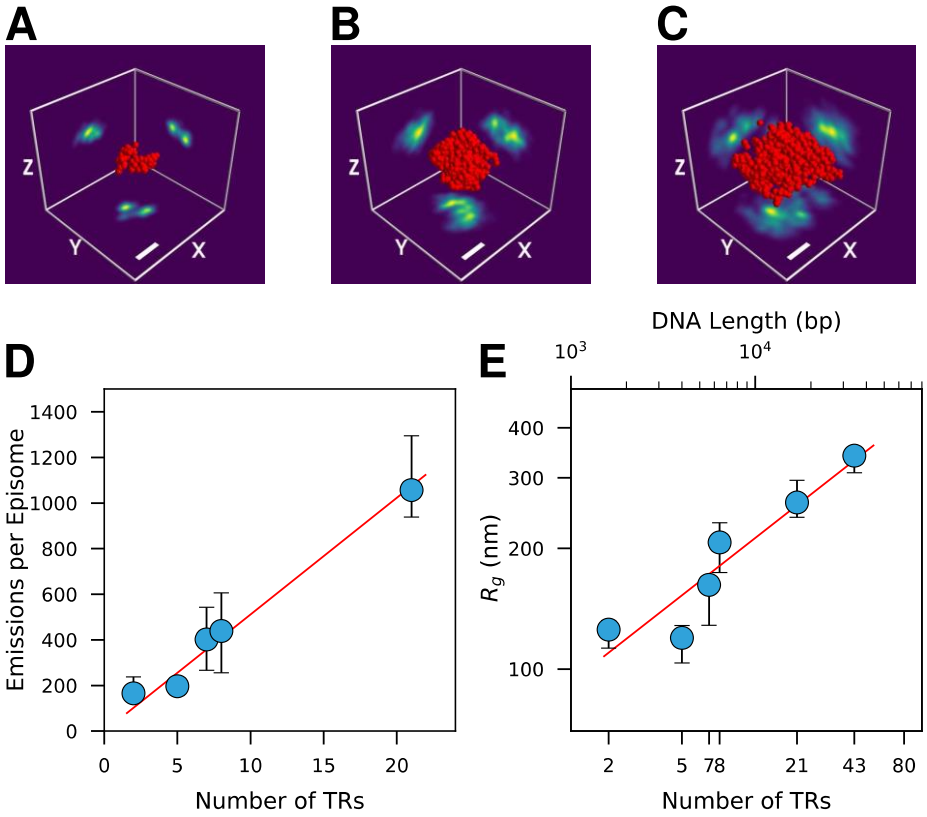

A
Z
Y
X
B
Z
Y
X
C
644
903
583
644
z
y
z
x
y
x
D
729
923
622
729
z
y
z
x
y
x

(P = 0.26, Mood's Median Test). The similarity of these parameters in a single viral strain in two different cell types shows that the polymer folding of KSHV LANA tethers is likely to be intrinsic to the viral episome itself, rather than being a result of the cellular milieu.

2.4.2 LANA occupancy scales with the number of TRs

We postulated that if the tethers visualized by dSTORM reflect the underlying TR architecture, then their structural parameters should change systematically with different numbers of TR elements. To challenge that hypothesis, we examined synthetic tethers with different numbers of TRs. We co-transfected BJAB cells, a KSHV-negative human B cell line, with two plasmids, one encoding LANA and the other containing 2, 5, 7, or 8 tandem TRs arranged head-to-tail as in the native episome (Figure 2-1A and B). In addition, we infected BJAB cells with BAC16 virus that we determined contains approximately 21 TRs (Figure 2-1C and D). Transfection with a LANA-encoding plasmid and a plasmid lacking any functional TR sequence (p0TR) resulted only in diffuse nuclear staining by epifluorescence, consistent with earlier work (88, 107). In contrast, individual tethers in cells harboring plasmids with 2, 8, or 21 TRs showed a progressive increase in size (Figure 2-4A-C). Accompanying this increase was a corresponding linear rise in the number of dSTORM fluorophore emissions and,

Figure 2-4. LANA tethers show polymer-like scaling as a function of the number of TRs. (A-C) The 3D-projections of representative LANA tethers from a cell transfected with p2TR **(A)**, p8TR **(B)**, or BAC16 with 21 TRs **(C)** are shown at identical scales as described in Figure 2-3. (Scale bar for **A-C** = 250 nm). **(D)** The linear relationship between the median number of emissions and the number of TRs per tether is shown for p2TR (N=28), p5TR (N=39), p7TR (N=13), p8TR (N=17), and BAC16 (N=32) (R^2 = 0.99). **(E)** The log of median R_g data for the tethers in **(D)** and BCBL-1 KSHV episomes (43 TRs, N=65) are shown as a function of the log of the tether DNA lengths (top axis), or the log of the TR number (bottom axis). The regression (solid red line) through tethers with known numbers of TRs (blue circles) was used to calculate the scaling exponent c = 0.36 ± 0.07 (R^2 = 0.94). Error bars represent 95% CIs.


A
B
C
Z
Y
X
D
Emissions per Episome
0
200
400
600
800
1000
1200
1400
0
5
10
15
20
Number of TRs
E
DNA Length (bp)
10^3
10^4
R_g (nm)
100
200
300
400
2
5
78
21
43
80
Number of TRs

therefore, anti-LANA Ab binding (Figure 2-4D).

Our observation that an increase in TR number results in a proportional increase in LANA binding argues against a plateau for TR occupancy per plasmid or episome, at least over the range of 1-21 TRs.

2.4.3 Tether polymer conformation has the characteristics of active chromatin

At present, there is conflicting evidence regarding the conformation of TR chromatin, which exhibits features of both repressed and active conformations. On one hand, TR nucleosomes are tightly packed in a typically “closed” conformation (143). Furthermore, LANA recruits the histone H3 methyltransferase SUV39H1 and establishes heterochromatin protein 1 (HP1) binding to the TR chromatin (76). In contrast, in support of an active conformation, TR chromatin is marked by histone hyperacetylation (71, 72, 143) associated with bromodomain protein BRD4 (120, 123), and the TRs have both promoter and origin of replication functions (127, 144).

To address the conformational state directly, we took advantage of the fact that the radius of gyration, R_g, of a chromatin polymer follows the power-law relationship, $R_g \propto L^c$ where L is the DNA length in bp and c is the scaling exponent. This scaling parameter is a sensitive measure of canonical active, inactive, and repressed states of mammalian chromatin (173). By plotting R_g against L we determined the power-law scaling exponent for the LANA-bound TR

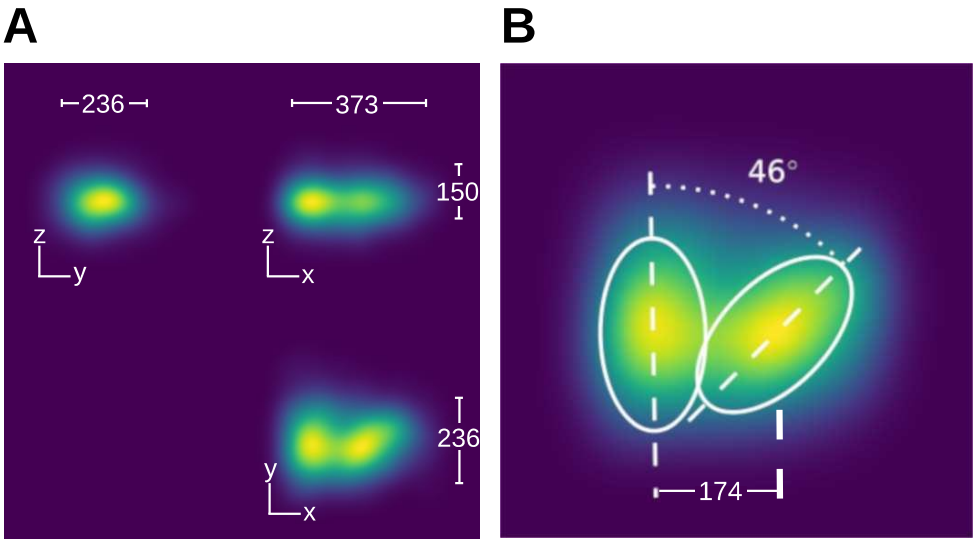
chromatin to be c = 0.36 +/- 0.07 (Figure 2-4E). This exponent approximated the value of 0.37 +/- 0.02 found by Boettiger et al. for active chromatin and is distinct from values for inactive (c = 0.30 +/- 0.02) or repressed (c = 0.22 +/- 0.02) chromatin (173). We obtained similar results for TR plasmid scaling in COS-7 cells (Figure 2-11A, located at the end of this chapter). This finding clarifies the long-standing paradox surrounding the chromatin folding state of the TR region.

2.4.4 Adjacent TRs are arranged with a specific asymmetric architecture

Closer analysis of the dSTORM emissions from individual p2TR tethers revealed that they frequently gave rise to two distinct, resolvable clusters (Figure 2-4A) and alignment of the complete set of 28 p2TR data (see Section 2.3.12) reinforced this separation (Figure 2-5A).

Individual clusters had an R_g of 90 nm, 95% CI [82, 96], which is similar to the R_g of 86 nm predicted by extrapolation of the power-law function to a single TR (Figure 2-4E). Thus, we interpreted the two clusters as the signals emanating from two adjacent TRs. We obtained similar results for p2TR dSTORM images acquired in COS-7 cells (Figure 2-11B). The p2TR tethers, individually or as an aligned composite, showed striking asymmetry. They comprised flat pairs of prominently elongated ellipsoids whose centroids were spaced apart by 174 nm, 95% CI [152, 185] (Figure 2-5B). Moreover, the long axes of the two clusters intersected to form a 46 degree angle (Figure 2-5B), demonstrating a specific

Figure 2-5. p2TR tethers show two distinct clusters of LANA Ab that form a 46° angle between their long axes. (A) Compilation of emission data from 28 p2TR tethers; emissions are rendered as in Figure 2-3. Median dimensions (nm) of the combined data are shown. **(B)** Approximately 2X magnified view of the Y-X projection from **A** demonstrating the 46° angle formed between the two clusters.

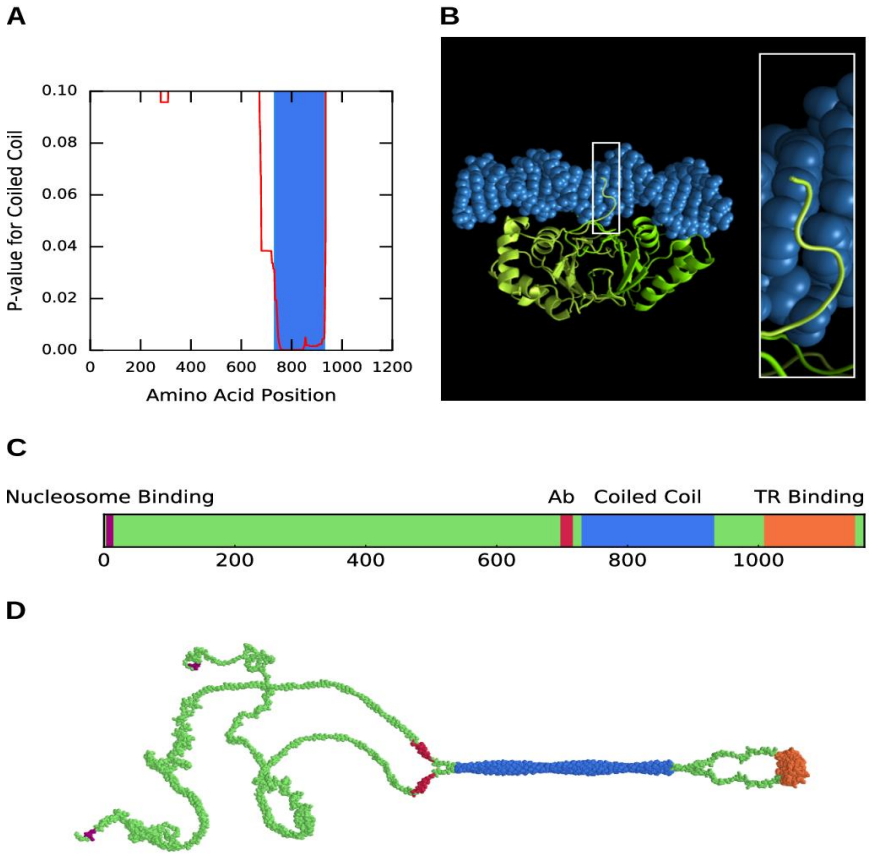

A
236
373
150
z
y
z
x
236
y
x
B
46°
174

differential placement of the individual TRs within the p2TR structure. These parameters provided strong constraints on the underlying folding of the p2TR chromatin tether.

2.4.5 2TR images support the prediction of a LANA coiled-coil domain

To begin to interpret the 2TR tether images at the molecular level, we evaluated the structure of LANA dimers N-terminal to the LBS binding domains, for which previous X-ray crystal structures are known (123–125). Examination of the primary protein sequence of LANA with Paircoil2 (174) revealed a segment strongly predicted to form a coiled-coil (Figure 2-6A; $P < 0.025$). To test for the presence of the coiled-coil, we measured the number of single molecule emissions detected by dSTORM for LANA tethers stained with LN53-A647, which recognizes the tetrapeptide epitope EQEQ (111). There are 22 such epitopes within the LANA protein sequence and this number of mAbs would yield over 18,000 dSTORM emissions per tether. However, 19 of the most C-terminal of these epitopes would be embedded within the putative coiled-coil region of LANA, occluding mAb recognition. The remaining three epitopes just N-terminal to the coiled-coil would yield on the order of only 2,500 emissions per tether. Indeed, we routinely recovered fewer than 2,000 emissions per tether, consistent with only 2 to 3 functional epitopes per LANA protein, providing strong empiric data for the predicted coiled-coil. Furthermore, the positioning of these available epitopes at the N-terminal end of a rigid coiled-coil domain is supported

Figure 2-6. Modeling LANA dimers: the LANA protein. (A) PairCoil2 (174) predicts a coiled-coil domain in LANA. The position of the domain is shown as a plot of probability versus amino acid number (LANA sequence GenBank: U75698.1). **(B)** The structure of a LANA dimer bound to LBS1, adapted from (125). Shown is the wrapping of LANA around the DNA to the side opposite LBS binding (inset). **(C)** Diagram of the various domains of LANA, including the nucleosome binding domain (purple), LN53 Ab binding sites outside the coiled coil region (red), coiled-coil region (blue), and TR binding domain (orange). LN53 Ab sites C-terminal to the red region were discounted from the TR modeling due to their being buried within the coiled-coil. **(D)** Structural rendering of LANA, showing the features and color scheme as in **C**.

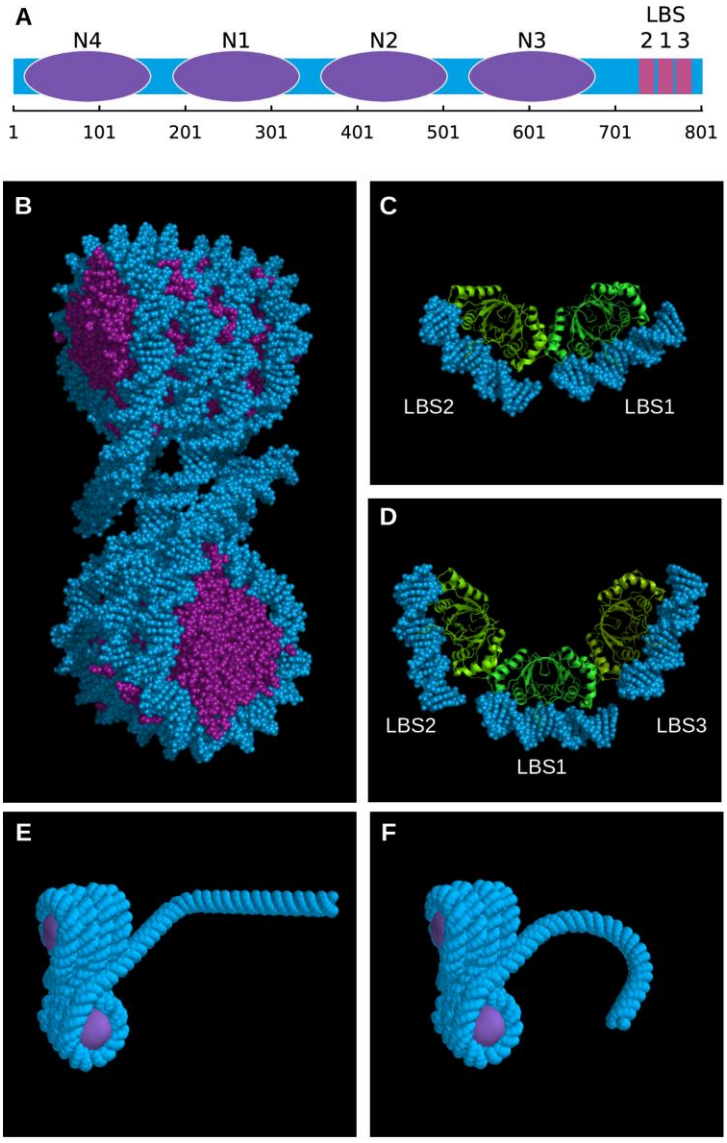

A
B
0.10
0.08
0.06
0.04
0.02
0.00
P-value for Coiled Coil
0
200
400
600
800
1000
1200
Amino Acid Position
C
Nucleosome Binding
Ab
Coiled Coil
TR Binding
0
200
400
600
800
1000
D

by the separation between the two emissions clusters in the 2TR images. We generated a molecular model of LANA dimers incorporating these features (Figure 2-6).

2.4.6 LANA-imposed DNA bending and nucleosome translational positioning are predicted to direct TR chromatin folding

Although the DNA of each TR is known to be bent by LANA dimer binding (93, 125, 128), and to be occupied by four nucleosomes (143), the constraints these factors impose on tether architecture are currently unknown. To address this question, we varied the parameters of DNA bending and nucleosome positioning in a set of molecular models (Figure 2-7). We modeled the 113 bp of nucleosome-free DNA that contains the three LBS sites in two modes. In the first, both LBS1 and LBS2 are occupied by LANA dimers, while LBS3 is unoccupied, inducing a 110 degree bend in LBS DNA (Figure 2-7C), as previously determined by Wong and Wilson (128). In the second mode, LBS1, LBS2, and LBS3 are all occupied by LANA dimers, inducing a bend of approximately 180 degrees (Figure 2-7D), as predicted by Hellert et al. (125). For the remaining 688 bp of TR DNA occupied by four nucleosomes, we adapted the crystal structure of a synthetic tetrasome (PDB 1ZBB) (175) because of its similarity to the spacing of the nucleosomes mapped to the TR (143) (Figure 2-7B). We digitally joined the tetrasome DNA to the LBS DNA, bent at either 110 (Figure 2-7E) or 180 degrees (Figure 2-7F). We varied the position of the joint by a single base pair at a

Figure 2-7. Modeling LANA dimers: the TR DNA. (A) Schematic diagram of a single 801-bp TR showing localization of four nucleosomes (purple ovals) and three LBSs (magenta rectangles) on the DNA (blue). **(B)** Structure of DNA (blue) wrapping around four nucleosomes (purple), adapted from the tetrasome crystal structure (PDB 1ZBB). **(C)** Structure of two LANA dimers (green) binding to two adjacent LBSs (blue) of a single TR, showing the 110 degree DNA bend formed by protein binding. **(D)** Structure of three LANA dimers bound to all three LBSs in a TR, demonstrating the nearly 180 degree bend in the bound DNA. **(E)** Model of the DNA (blue) in a single TR, including the wrapping of four nucleosomes (purple), showing the 110 degree bend formed when only two adjacent LBSs are bound by LANA. **(F)** Model as in **E** but showing the 180 degree bend resulting from all three LBSs bound by LANA.

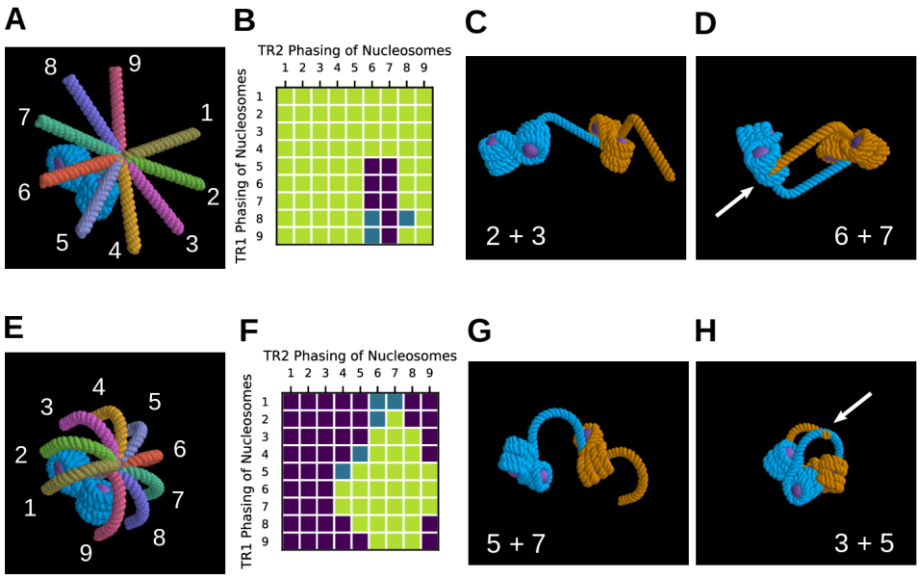

A
N4
N1
N2
N3
LBS
2 1 3
1
101
201
301
401
501
601
701
801
B
C
LBS2
LBS1
D
LBS2
LBS3
LBS1
E
F

time over a 9 bp span, designated phases 1-9, to model the effect of nucleosome translational positioning on tether architecture through an entire turn of a DNA helix (Figure 2-8A and E). Each base pair change in the nucleosome phasing changes the trajectory of the chromatin fiber by approximately 36 degrees (1/10 of a full turn). To extend the model to include two adjacent TRs, we joined pairs of these single TR models at each of the 9 nucleosome phases to produce the complete set of 81 2TR models (9 phases for each tetrasome) (Figure 2-8B and F), keeping the repeat length at 801 bp between LBS1 sites. For models with occupancy of only LBS1 and LBS2, we eliminated 11 of the 81 combinations from further consideration based on steric interference, either between two tetrasomes or the trajectory of connecting DNA (shown in purple and blue, Figure 2-8B). For the models with full occupancy of three LBS sites, 50 models were ruled out (shown in purple and blue, Figure 2-8F). Examples of 2TR phase combinations with or without steric hindrance are depicted in Figure 2-8. The impact of both LANA-dependent DNA bending and nucleosome translational positioning on the construction of these 2TR models, and the ability of these parameters to eliminate certain tetrasome configurations, demonstrates the importance of these two factors for tether architecture.

2.4.7 A simple molecular model captures major properties of 2TR LANA tethers

Remarkably, our 2TR models captured key features of the dSTORM

Figure 2-8. Modeling 2TRs: nucleosome phasing. (A) A DNA-wrapped tetrasome (blue) displaying the rotational effects of the nine theoretically possible nucleosome phases on the position of the LBS-containing DNA fragment emerging from a single TR in which only two LBSs are bound by LANA (numbered 1 through 9). **(B)** A chart depicting the combinations of nucleosome phasing for two sequential TRs (TR1 and TR2), in which only two LBSs are bound by LANA. The nucleosome phasing results in “disallowed” arrangements (purple squares), wherein the elements face steric hindrance, “permitted” arrangements (green squares), in which there is no steric hindrance, and arrangements with likely interference (blue squares). **(C-D)** An example of a “permitted” phasing combination from B (TR1 = phase 2, TR2 = phase 3) is shown in **C**, in contrast to a “disallowed” phase (TR1 = phase 6, TR2 = phase 7) presented in **D**. Regions of steric hindrance are marked with white arrows. **(E)** The nine theoretically possible phases of the LBS-containing DNA fragment emerging from a TR in which all three LBSs are bound by LANA. **(F)** A chart depicting nucleosome phasing combinations for two sequential TRs in which all three LBSs are bound by LANA. Colors are as in **B**. **(G-H)** An example of a "permitted" phasing combination from **F** (TR1 = phase 5, TR2 = phase 7) is shown in **G**, in contrast to a "disallowed" phase (TR1 = phase 3, TR2 = phase 5) presented in **H**.

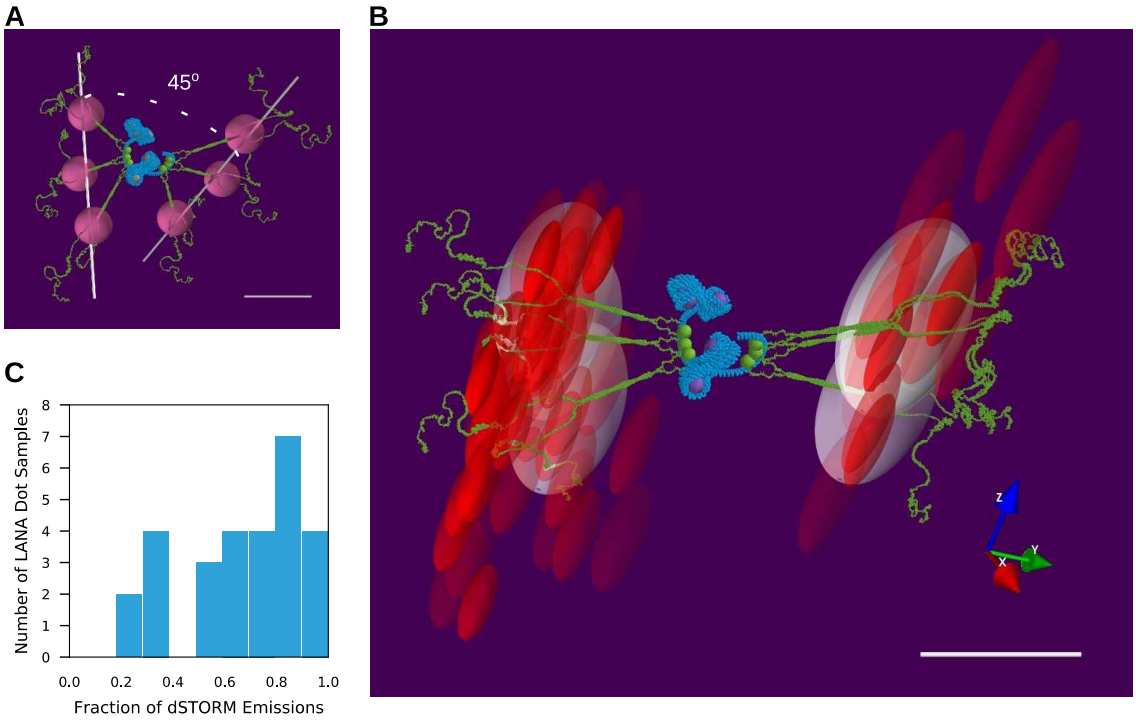

A
B
C
D
8
9
7
1
6
2
5
4
3
TR2 Phasing of Nucleosomes
1 2 3 4 5 6 7 8 9
TR1 Phasing of Nucleosomes
1 2 3 4 5 6 7 8 9
2 + 3
6 + 7
E
F
G
H
4
3
5
2
6
1
7
9
8
TR2 Phasing of Nucleosomes
1 2 3 4 5 6 7 8 9
TR1 Phasing of Nucleosomes
1 2 3 4 5 6 7 8 9
5 + 7
3 + 5

images. In particular, the model with full LANA occupancy of all three LBSs on both TRs and the second tetrasome joined to the DNA at “phase 8” mimicked features of the image (Figure 2-9A, and see also Figure 2-8E). The addition of three LANA dimers to each TR of the 2TR model produced an elongated arrangement of epitope binding sites and predicted the localization of their associated dye emissions. Furthermore, the long axes joining the epitope binding sites on each TR give rise to approximately a 45 degree angle between them by the combination of tetranucleosome positioning and LBS DNA bending. The distance between the centers of the two sets of epitopes was approximately 151 nm. Both the angle and spacing of emissions in the model approximated those of the experimental data (Figure 2-9B).

We tested the ability of our chromatin models to account for the dSTORM data of individual tethers by optimizing the fit of the modeled epitope sites (dye locations) to the experimental p2TR emissions data. We then determined the best fit models by scoring each by the fraction of dSTORM emissions that they captured. One example of such a fit is shown in Figure 2-9B.

In a pair-wise comparison of models with LANA dimers at all three LBS sites versus models with LANA dimers only at LBS1 and LBS2, the full occupancy models were significantly better at accounting for the 28 p2TR datasets ($P < 3.0 \times 10^{-12}$, Paired T-test). Overall, the median fraction of dSTORM emissions accounted for by the optimal full occupancy models was 0.72, 95% CI

Figure 2-9. dSTORM analysis of p2TR tethers is consistent with a model featuring full LBS and nucleosome occupancy on two sequential TRs. (A) Model depicting two TRs (blue), each containing four nucleosomes (grey) and bound by three LANA dimers (green) occupying LBSs. Predicted coiled-coil portions of LANA dimers extend outward at approximately right angles to the LBSs. Magenta spheres depict the position of a 13 nm-long anti-LANA mAb. Two white lines represent the best linear fit connecting three LANA dimers for each TR, forming a 45° angle. **(B)** Data from one p2TR example aligned with the model in **(A)**. Red ellipsoids depict the probability volume of individual fluorophore emissions, scaled by their lateral and axial localization precision. White ellipsoids depict the model-predicted probability volumes scaled by data-derived localization precision. **(C)** Histogram displaying the distribution of the fraction of observed emissions whose localization precision volumes intersect with the model-predicted volume. (Scale bars = 50 nm)

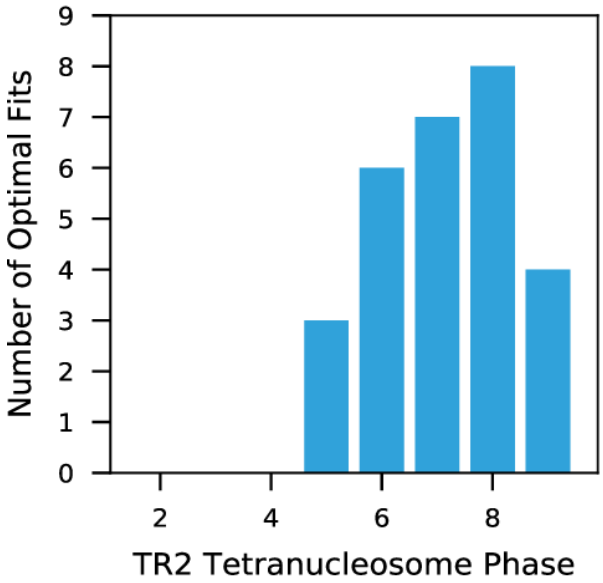

A
45°
B
C
Number of LANA Dot Samples
0
1
2
3
4
5
6
7
8
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of dSTORM Emissions
Z
Y
X

[0.61, 0.83] (Figure 2-9C) for the p2TR structures assembled in BJAB cells. We carried out the same analysis of p2TR LANA tethers expressed in COS-7 cells and found the median fraction of emissions covered to be 0.78, 95% CI [0.66, 0.94] (Figure 2-11C -D). Interestingly, the optimal phasing between TRs was not the same for all 2TR tethers examined. Rather, this parameter varied from tether to tether, with phases 5-9 each contributing to the set of best fit models (Figure 2-10). This range of phases suggests a basis for the shape variability seen among the individual LANA tethers comprised of larger numbers of TRs (Figure 2-3).

2.5 DISCUSSION

Until now, LANA tethers have appeared as poorly defined nuclear punctae via epifluorescence microscopy. The localization data from this study offer new insights into the architecture of the tethers linking KSHV episomes to human chromatin. First, we found that the folding of tether chromatin was intrinsic to the viral episome and independent of the cellular environment. Consistent shapes and emission counts were measured for comparable TR episomes within the B cell-derived lines BCBL-1 and BJAB, the non-lymphocytic epithelial cell line SLKp, and even the monkey kidney fibroblast cell line COS-7. Second, regardless of the number, TRs were proportionally occupied by LANA dimers. This argues for a relatively uniform chromatin environment across the terminal repeats wherein LBS sites are equally accessible to LANA binding in all TRs.

Figure 2-10. Multiple tetranucleosome translational positions are compatible with the p2TR dSTORM data in BJAB cells. The frequency of the occurrence of an optimal fit is plotted as a function of the tetranucleosome phasing of TR2 in models using 3 LANA dimers per TR.


Number of Optimal Fits
0
1
2
3
4
5
6
7
8
9
2
4
6
8
TR2 Tetranucleosome Phase

Third, we provided the first physical proof that TR chromatin folds with parameters characteristic of active chromatin, likely driven by nucleosome-free regions in the TRs, helping to clarify conflicting evidence pertaining to the chromatin state (71, 72, 76, 127, 143, 144). Fourth, the ability of dSTORM to resolve single TR units within 2TR clusters demonstrates that the episomal half of LANA dimers must be well ordered. If LANA dimers radiated randomly from the LBSs or were highly unstructured in their episomal half, we would be unable to separate discrete clusters in the analyses of plasmids containing 2 TRs. Finally, our data argue that all three LBS sites in individual TRs are likely to be fully occupied by LANA dimers. The discrete dimensions and positional asymmetry of the dye emission clusters in the 2TR data are not well accounted for if only two LBS sites are occupied, because the axis through the epitopes of each TR is shortened and there is less bending of the LBS DNA. Furthermore, models with three LANA dimers per TR consistently performed better than models with only two LANA dimers per TR in explaining dSTORM data for 26 independent 2TR datasets. The linear relationship between LANA dye emissions and TR number suggests that this is true for complete KSHV episomes as well. We identified several key features of tether structure by modeling the 2TR data. First, the modeling revealed that the internal repeats of LANA likely comprise a coiled-coil domain for the LANA dimer. This domain was first predicted by in silico analysis, and then supported by quantitative analysis of LN53 mAb binding, and masks all but the two to three N-terminal most mAb binding sites. Second,

the modeling highlighted an impact of translational positioning of the nucleosomes for each TR. Moving each tetranucleosome along the nine possible phases of the DNA helix alters the trajectory of the output TR DNA. This, in turn, greatly impacts the relative positions of the LBSs on adjacent TRs and, hence, their associated anti-LANA fluorophore emissions. The architecture driven by nucleosome positioning and LBS DNA bending positions the LANA N-terminal histone binding domains in a way that facilitates exploration of the surrounding environment, promoting their primary function of tethering to host chromatin.

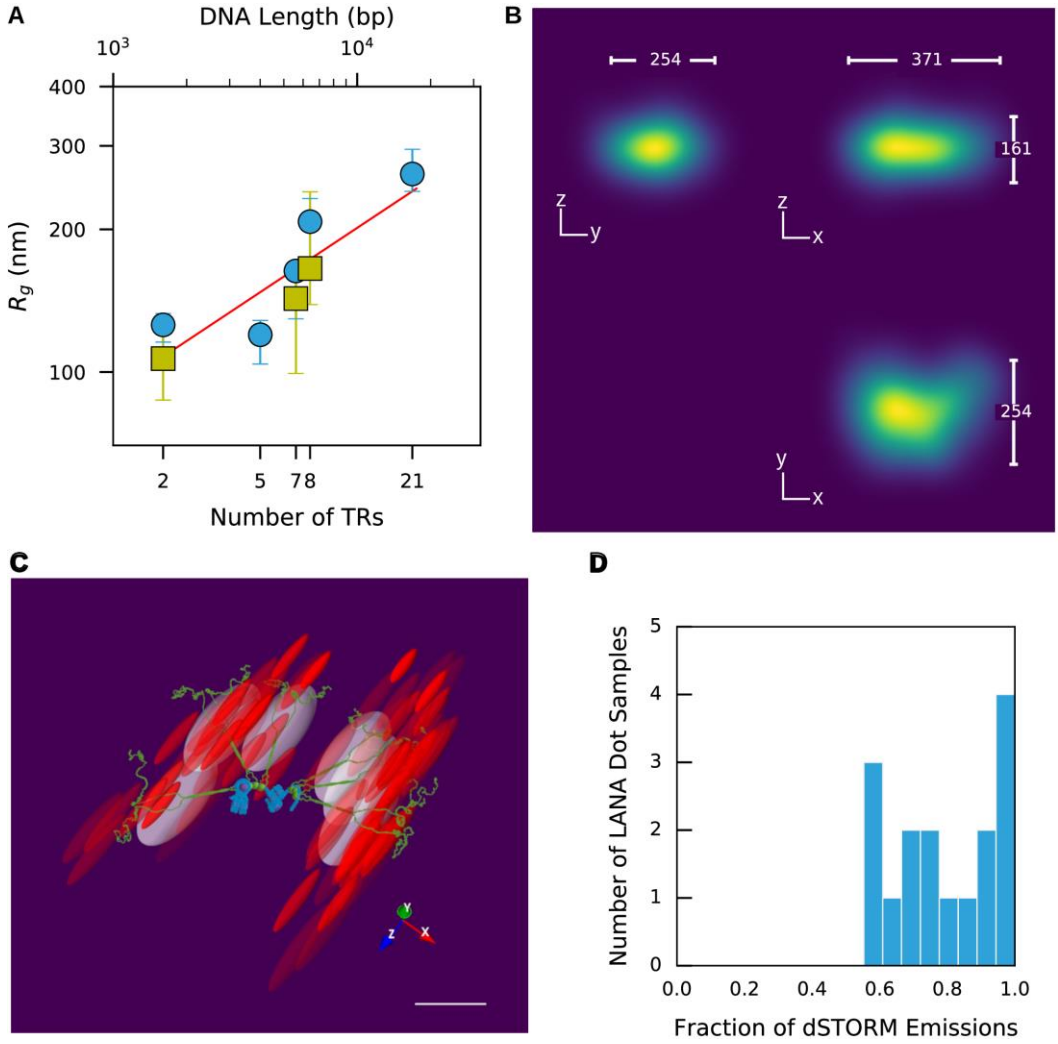
While the 2TR model describes most of the tethers well, a few datasets are exceptions mainly due to their having relatively larger R_g values. Others have proposed large multi-LANA dimer complexes based on crystallography of the C-terminal binding domain (123–125), and it is possible these 2TR exceptions reflect such assemblies. Further, our current model has some limitations. For example, it is unlikely that the fixed packing of tetranucleosomes we adapted from the crystallographic structure reflects the full range of conformations adapted by TR nucleosomes in vivo. Unlike the KSHV tether histones, the histones present in the tetranucleosome crystal structure lacked any posttranslational modifications. The presence of such modifications might alter the tetrasome structure, potentially loosening or tightening the histone packing, and thereby impact our model and its fit to our data.

Recently, Chiu et al. described the clustering of a subset of KSHV episomes (121). In our study, we focused on individual TR structures at single

molecule localization precision (10-30 nm) and excluded any ambiguous structures from our analyses. We occasionally noted single bright LANA dots by epifluorescence but most often dSTORM resolved these into distinct LANA tethers. In these cases the N-termini of the LANA tethers might have also clustered episomes as Chiu et al. suggest.

Our current working model captures major features of the KSHV tether, detailing the likely stoichiometry and relative positions of its major molecular components emanating from the TR region of the viral genome within the cellular milieu. The model also provides a platform for future experimentation that could include determining the nanoarchitecture of tethers with greater numbers of TRs. This approach has broad applicability to a wide variety of persistent viral pathogens, thus, potentially contributing to our ability to target them therapeutically.

Figure 2-11. Analysis of LANA tethers in COS-7 cells. (A) Scaling of R_g with increasing number of TRs in COS-7 cells supports the results seen in BJAB cells. A plot of R_g vs TR number for p2TR, p7TR, and p8TR in COS-7 (yellow squares), shown in comparison with the BJAB data (blue circles). The combined data produce a scaling exponent of c = 0.35 ± 0.07, R^2 = 0.89. **(B)** COS-7 p2TR data shows two clusters of emissions for each tether, similar to that seen in BJAB cells. Compilation of emission data from 16 p2TR tethers in COS-7 cells; emissions are rendered as in Figure 2-3. Median dimensions (nm) of the combined data are shown. **(C)** Data from one p2TR example aligned with a model similar to that shown in Figure 2-9A. Red ellipsoids depict the probability volume of individual fluorophore emissions, scaled by their lateral and axial localization precision. White ellipsoids depict the model-predicted probability volumes scaled by data-derived localization precision. Scale bar = 50 nm. **(D)** Histogram displaying the distribution of the fraction of observed emissions whose localization precision volumes intersect with the model-predicted volume.
Median overlap fraction = 0.78, 95% CI [0.66, 0.94].


A
DNA Length (bp)
10^3
10^4
400
300
200
100
R_g (nm)
2
5
78
21
Number of TRs
B
254
371
161
254
z
y
z
x
y
x
C
D
Number of LANA Dot Samples
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of dSTORM Emissions

CHAPTER 3. EXAMINING THE PROPERTIES OF THE KSHV TETHER DURING CELLULAR MITOSIS USING SUPER-RESOLUTION MICROSCOPY

3.1 ABSTRACT

The critical moment for an episomal tether is during cellular mitosis. During this time, the tether must retain viral DNA with the host chromatin both to ensure faithful partitioning of the episomes between daughter cells and to prevent activation of the cGAS-STING pathway by cytoplasmic DNA following breakdown of the nuclear envelope. In this study, we investigated the properties of the LANA tether during mitosis. We determined that tethers during mitosis had an approximately four-fold lower volume than those examined during interphase. We also found that tethers in the adherent cell line SLKp shared a common directionality with each other and also aligned with the metaphase plate. Taken together, these two findings suggest a strong influence of the cellular mitotic machinery on the behavior of the LANA tether. Finally, we demonstrated that LANA tethers are located on the outside of mitotic chromosomes in BCBL-1 cells, rather than being embedded within the condensed chromatid arms. This finding agrees with the hypothesis that during mitosis, cellular chromatin condenses prior to attachment of viral episomes to host cell histones, rather than occurring afterward. These preliminary experiments pave the way for future work applying the resolution of dSTORM to answering key questions about the mitotic KSHV LANA tether.

3.2 INTRODUCTION

3.2.1 The mitotic switch

Mitosis is the stage in the cell cycle wherein a cell with duplicated DNA from S phase undergoes division into two identical daughter cells. The word “mitosis”, from the Greek for “warp thread”, refers to the thread-like appearance of the chromosomes when they condense during the process (176). The main trigger for the start of mitosis is the activation of the protein kinase cyclin B-Cyclin-dependent kinase (Cdk) 1, a protein that is primarily cytoplasmic during interphase, but which localizes to the nucleus during mitosis (177). In fact, nuclear import of Cyclin B-Cdk1, along with cell rounding, is one of the first events in the entire mitotic process (178, 179). Cyclin B-Cdk1 is activated on a timescale of approximately 30 minutes and plays many roles throughout the mitotic process, including promoting nuclear envelope breakdown (NEBD) and regulating the timing of the various mitotic stages (prophase, metaphase, anaphase, and telophase – detailed below) (179, 180).

3.2.2 Prophase

One of the most essential components of mitosis is the compaction of the DNA that was duplicated during S phase. The measurement of the degree to which mitotic DNA is compacted compared to interphase DNA varies from 2- to 1000-fold, depending upon the assessment (181, 182). A family of proteins called the structural maintenance of chromosome (SMC) proteins is responsible for DNA compaction. During prophase, the pentameric SMC complex condensin II

forms a ring-like structure that can entrap strands of DNA and pull them together, resulting in axial compression (183–185). Condensin II is found on chromosomes from early prophase through the end of mitosis and is activated by phosphorylation by cyclin B-Cdk1 in the early stages of prophase (186–188). The serine/threonine kinase Aurora B also phosphorylates, and thereby regulates, condensin complexes, and is required for association of condensin with mitotic chromatin (189–191). Additional chromatin condensation occurs when histone H4 tails bind to an acidic patch on another nearby nucleosome. In interphase, this interaction is prevented by H4K16 acetylation, but during mitosis, Aurora B recruits a deacetylase, thereby triggering this additional method of condensation (192, 193). Another SMC complex, cohesin, plays the role of holding sister chromatids together, joined at a point called the centromere, which plays an important role in the later stages of cell division (194).

In addition to condensing its chromatin, a mitotic cell must also reorganize its microtubule network into a structure called the mitotic spindle, which enables the sister chromatids to be pulled apart. The kinase Aurora A is important in forming the mitotic spindle by regulating the movement of microtubules (195). This network begins with two microtubule nucleation points, called centrosomes. Microtubules assemble between the two centrosomes, which then move to opposite ends of the cell (182). Each microtubule in this network finds and connects to a pair of chromatids via the complex of proteins at its centromere called the kinetochore. At this point, the spindle assembly checkpoint (SAC)

detects any kinetochores that are not correctly attached to microtubules and prevents the degradation of cyclin B until the last kinetochore is attached to the microtubules (196–199). Aurora B kinase also has a role here in correcting erroneous microtubule attachments to kinetochores (200).

The timing of the mitotic stages is highly dependent upon ubiquitin-mediated proteolysis of the various protein kinases and phosphatases involved, thereby insuring the correct balance for each part of the mitotic process. Ubiquitin-mediated proteolysis requires two enzymes – an E2 enzyme that carries the ubiquitin and an E3 ubiquitin ligase to bind E2 and the substrate together (201). The E3 ligase involved in mitosis is the anaphase-promoting complex/cytosome (APC/C). It is kept inactive by the SAC until late prophase when it is phosphorylated, and thereby activated, by cyclin B-Cdk1 (202, 203). Around the same time, cyclin B-Cdk1 also phosphorylates the nuclear lamins, the filamentous proteins providing nuclear structure, resulting in nuclear envelope breakdown (NEBD) (178, 179). The SMC protein condensin I, which remains cytoplasmic during interphase, is imported into the nucleus following NEBD (186, 187). There it binds to chromosomes and performs both lateral and additional axial compression of chromosome arms (184).

3.2.3 Metaphase

When all kinetochores are properly attached to microtubules, the SAC is turned off, cyclin B is degraded, and metaphase begins (180). Sister chromatids

line up along the cell equator, and the microtubules attached to kinetochores ensure that sister chromatids are connected to the spindle poles on either end of the cell (182).

3.2.4 Anaphase

APC/C is the protein switch that causes the shift from metaphase to anaphase, as it begins to degrade cyclin B. At this stage, Aurora B moves from centromeres to microtubules in order to coordinate cytokinesis (200). Ubiquitin-mediated proteolysis activates separase, a protease that cleaves the cohesin rings holding sister chromatids together (204). Breakdown of cohesin, along with the departure of condensin I from chromosomes, allows sister chromatids to separate from each other (182, 186). Kinetochore microtubules shorten, drawing the sister chromatids toward opposite spindle poles, during the period known as anaphase A (182). This is followed by anaphase B, during which microtubules pull the spindle poles apart, further separating sister chromatids. Interestingly, the highest level of chromatin compaction is seen here, during the segregation period of anaphase (205).

3.2.5 Telophase and cytokinesis

Once chromosomes are positioned at the two spindle poles, the mitotic spindle begins to disassemble. At this stage, cyclin B-Cdk1 levels are so low that the cell is driven toward exiting mitosis. The balance of phosphatases and kinases once again shifts, promoting a return to the cell's interphase architecture.

Aurora B is removed from chromatin, allowing for decondensation to occur (206). Phosphatases dephosphorylate the lamins at the cell poles, allowing the nuclear membrane to re-form around the two separate sets of chromosomes. The cell membrane then pinches in at its equator, which bisects the cell into two equally-sized daughter cells in the process known as cytokinesis. At this stage, both daughter cells return to G1 phase and are ready to begin the cell cycle once again.

3.2.6 Role of LANA during mitosis

From a viral perspective, the main role of LANA during mitosis is its function of tethering viral episomes to host chromosomes (89). This ensures faithful partitioning of episomes to daughter cells (90, 207) and also serves to retain viral DNA in the nucleus following cytokinesis. Early studies showed association of LANA with mitotic chromosomes (88, 91, 92), and in 2007, Kelley-Clarke et al. suggested preferential localization of LANA punctae to centromeric and telomeric regions of metaphase chromosomes (95).

LANA also directly interacts with the cellular mitotic machinery. It upregulates the expression of survivin, a member of the chromosomal passenger complex, which regulates the localization and expression of the crucial Aurora B kinase (208–210). Deletion of survivin causes defects in chromosome segregation and cytokinesis (211), making it an essential part in mitotic regulation. LANA also directly recruits the Aurora B kinase, which induces phosphorylation of survivin’s tyrosine 34 (212). This modification alters the

activity of histone acetyltransferases and deacetylases, leading to increased viral copy number and a boost in KSHV replication in latently-infected B cells. In addition, LANA interferes with the G2/M checkpoint of the cell cycle, blocking p53-mediated apoptosis and relieving nocodazole-induced cell cycle arrest (213).

3.2.7 Effects of mitosis on the TR region

In 2004, Stedman et al. examined the effects of the cell cycle on LANA occupation of the KSHV episome's TR region. They found that the amount of LANA bound to TRs was reduced 3-fold in cells arrested at G2/M phase with colchicine compared to that seen in an asynchronous population (143). The discovery that cellular levels of LANA were reduced in the colchicine-arrested cells served as an explanation for this decrease, but from a viral tethering perspective, the decrease in LANA binding to TRs during G2/M phase is perhaps puzzling, as this is the time when a strong, functional tether is most important.

3.3 MATERIALS AND METHODS

3.3.1 Isolating Mitotic Populations

Mitotic adherent SLKp cells were isolated by mitotic shake-off. Mitotic BCBL-1 cells were isolated by treatment with 0.05 µg/mL colcemid for 24 hours before harvest, yielding approximately 33% mitotic cells. In experiments involving mitotic spreads, cells were washed 1x in PBS, spun 5 minutes at 600 rpm, resuspended in 7.5 mL 0.075M KCl at 37 degrees C for 1 hour, and fixed in 4% paraformaldehyde. Spreads were created by dropping cells from a height of approximately 30 cm onto poly-lysine-coated coverslips. During imaging, mitotic

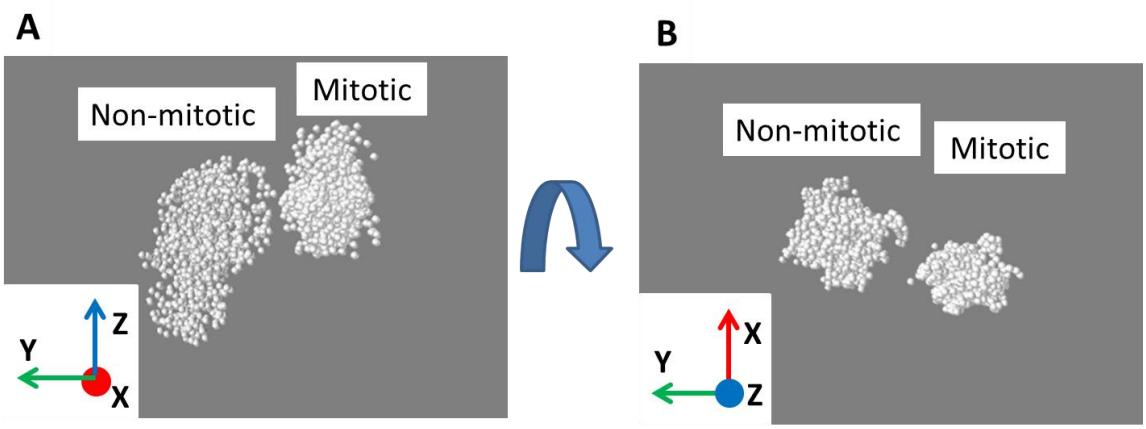
cells were distinguished from non-mitotic cells by their positive staining for anti-phospho-histone 3, a marker for late G2 to M phase.

3.4 RESULTS

3.4.1 LANA tethers condense along with condensed host chromatin

Because herpesviruses are chromatinized like their host DNA, we asked whether the episomal tether also undergoes condensation during mitosis. We measured the median R_g of 53 LANA tethers in mitotic SLKp cells and 31 tethers in non-mitotic cells. Non-mitotic tethers had a median R_g of 322 +/- 62.8 nm, while mitotic tethers were smaller, with an R_g of 204 +/- 47.3 nm (p = 1.0 x 10^{-14}), indicating an approximately 4-fold reduction in tether volume during mitosis (Figure 3-1). This suggests that the episomal DNA undergoes condensation along with the host chromatin during mitosis.

Figure 3-1. LANA tethers in SLKp cells undergo condensation during mitosis. **(A)** One representative tether from a non-mitotic SLKp cell and one from a mitotic cell, both viewed along their shortest axes (x axes), as determined by principle component analysis. Note the visible difference in size between the two tethers, with the non-mitotic example being larger than the mitotic example. **(B)** The same two tethers as in **(A)**, with the view rotated by 90 degrees down their longest axes (z axes). The difference in size between the two tethers remains the same from this viewpoint. **(C)** The mean calculated R_g values of 31 tethers in non-mitotic SLKp cells and 53 tethers in mitotic cells. The difference is statistically significant ($P = 1.0 \times 10^{-14}$) and represents an approximately four-fold difference in volume between the two sets of tethers.



C

Condition	R_g	N
Non-mitotic	322 +/- 62.8 nm	31
Mitotic	204 +/- 47.3 nm	53

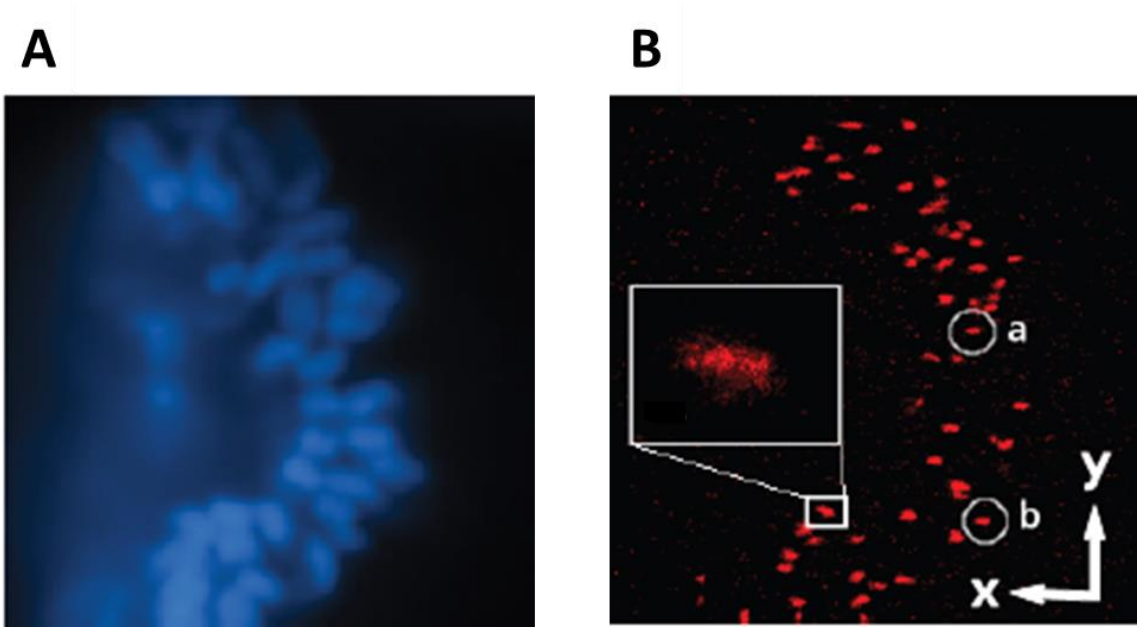
3.4.2 LANA tethers orient with respect to the metaphase plate during mitosis

Upon further examination of mitotic LANA tethers in SLKp cells, we noticed a common directionality among tethers within a single cell. That is, their long axes were oriented with respect to the metaphase axis - parallel to each other and pointing toward the center of the prophase chromatin “donut”. Figure 3-2 shows this phenomenon in one representative SLKp cell. This suggests that the cellular machinery dictating the directionality of condensing host chromatin also has an effect on episomal DNA and, therefore, the overall tether.

3.4.3 LANA tethers in BCBL-1 cells are located on the outside of condensed chromosomes

While tethering to mitotic host chromatin is clearly crucial during metaphase, it is unknown whether tethers are formed prior to or after condensation of the host DNA. If attachment occurred prior to condensation, we would expect to see multiple LANA tethers embedded within the condensed host chromatin. However, if condensation preceded attachment, we would expect to see the tethers exclusively on the outer surface of condensed chromatids. Previous work of others has extensively shown association of LANA tethers with mitotic chromosomes (88, 91, 92), but Szekely et al. in particular noticed that

Figure 3-2. LANA tethers in mitotic SLKp cells orient together with respect to the direction of the metaphase axis. (A) DAPI staining (blue) of the condensed chromosomes in an SLKp cell in prophase. The cell is positioned such that half of the prophase "donut" shape formed by the chromosomes is coming toward the viewer while the other half is oriented away (into the page). **(B)** LANA staining (red) of the same cell as in **(A)**. The LANA tethers exhibit a common directionality, with their long axes oriented parallel to each other and pointing toward the center of the cell. Circled tethers (a) and (b) are two examples clearly showing parallel orientation. The inset shows a third example tether, magnified to more clearly demonstrate how its long axis is facing toward the center of the "donut" formed by the prophase chromosomes.

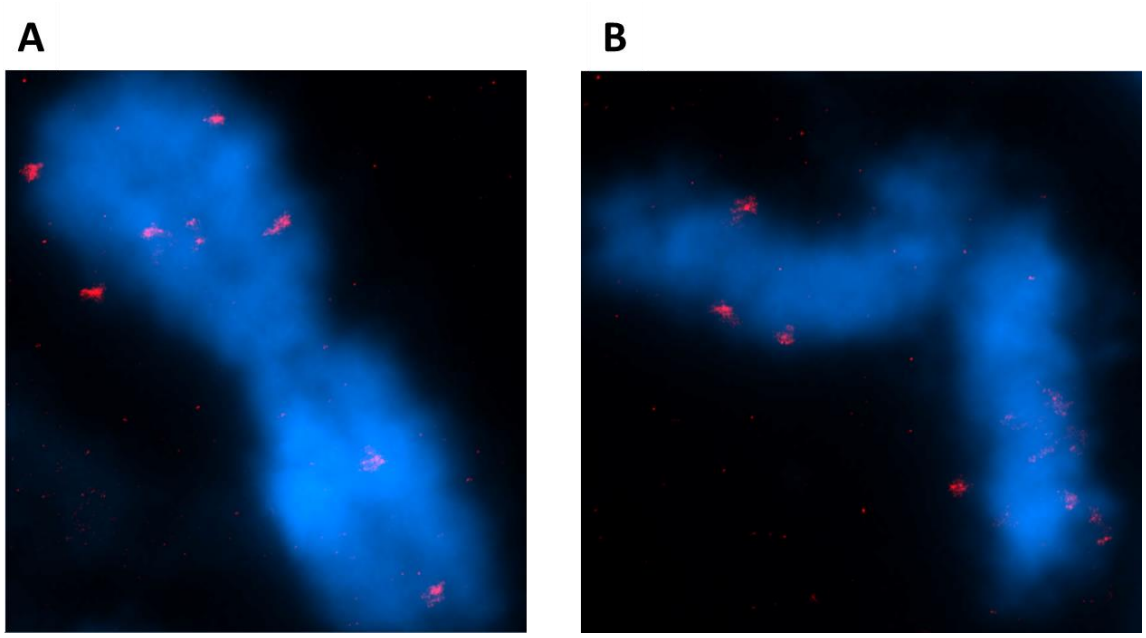

A
B
a
b
y
x

tethers seemed exclusively located on the outside of condensed chromatin arms and not imbedded within (92). To test this localization for ourselves, we examined the positioning of LANA tethers on the condensed chromatids of a BCBL-1 mitotic spread. Overlaying the LANA tethers on the DAPI staining revealed no tethers embedded within the host DNA (Figure 3-3); rather, tethers were constrained to the outer surface of the chromatids. While not definitive, this agrees with a model in which host DNA condensation occurs prior to LANA-host chromatin attachment.

3.5 DISCUSSION

The data presented in this chapter are a set of preliminary experiments delving into the nature of the KSHV LANA tether during the crucial mitotic phase of the cell cycle. Here we show that mitotic tethers in SLKp cells have an approximately four-fold lower volume than tethers during interphase. This seems to suggest that the cellular machinery causing chromatin condensation also has an effect on KSHV chromatin. Future experiments might examine volume differences between mitotic and non-mitotic tethers in different KSHV-positive cell lines to determine whether this effect is universal and whether the extent of the compression is cell type-dependent. We show that mitotic tethers in SLKp cells have a common directionality with respect to each other and to the metaphase plate. This suggests that the cellular forces acting to localize condensed chromatids to the metaphase plate are also directly affecting LANA

Figure 3-3. LANA tethers are located on the outside of mitotic BCBL-1 chromosomes. (A and B) Two pairs of condensed sister chromatids, stained by DAPI (blue) and viewed by epifluorescence microscopy, with superimposed LANA staining (red) acquired by dSTORM. In both examples, LANA tethers appear to be preferentially located on the outside of the condensed chromosomes and are not embedded within, suggesting that condensation occurred prior to tether attachment.


A
B

tethers. One follow-up experiment might involve looking at the directionality of tethers at the various mitotic stages to ascertain whether tethers are, indeed, “following” the structure provided by the cellular mitotic machinery throughout the entire process. More extensive long-term experiments might include the disruption of various proteins required for successful mitosis to determine which of these are influencing this behavior of LANA tethers. Here we also present data suggesting that LANA tethers in BCBL-1 cells are located on the outside of mitotic chromosomes and are not embedded within. This agrees with a timeline in which cellular chromatin condenses prior to episomal attachment, rather than the reverse. While we do not claim to have answered this major question definitively, our data certainly agree with this model, proposed by Szekely et al. (92). One potential criticism of our study might be that some tethers could simply be embedded in chromosomes in such a way that they are inaccessible to our antibody for staining. A direct fusion of LANA with a fluorescent protein amenable to super-resolution microscopy would solve this problem and enable us to get a more definitive answer on whether additional LANA tethers are hiding amidst the condensed cellular chromatin. Overall, our results suggest marked changes to the LANA tether during mitosis and pave the way for future studies examining mitotic tethers with super-resolution microscopy.

CHAPTER 4. CRISPR-Cas9 GENOMIC EDITING OF THE KAPOSI'S SARCOMA-ASSOCIATED HERPESVIRUS LATENCY GENE LANA

4.1 ABSTRACT

Extensive research has demonstrated the key role of LANA in maintaining KSHV episomes within dividing cell populations. As such, it is a natural target for disrupting viral latency and is an attractive prospect for disease intervention. With the recent advent of CRISPR-Cas genome editing, disruption of a protein-encoding gene has become a relatively streamlined process, requiring only a plasmid containing the selected nuclease protein and a plasmid encoding the appropriate RNA sequence for targeting to DNA. Basic cloning, often using a Type IIS restriction enzyme that cleaves outside of its recognition sequence, allows for directional insertion of the desired target sequence. Although CRISPR-Cas9 has been previously used in herpesviruses (214, 215), there are, to date, no publications of its application to KSHV specifically. Therefore, in an effort to disrupt the LANA gene and create a non-functional protein, we have applied the CRISPR-Cas technique to LANA. We have targeted the LANA gene in the stably LANA-expressing B cell line BJAB-LANA. We show successful clearance of visible LANA protein by immunofluorescence. In addition, we show by DNA sequencing the localized insertions/deletions in the LANA sequence that are characteristic of non-homologous end joining, the main DNA repair mechanism following CRISPR-Cas editing. Our successful disruption of LANA in this cell line sets the stage for targeting it in cell lines containing full latent KSHV episomes, thus rendering it a potential tool in the elimination of latent KSHV from cells.

4.2 INTRODUCTION

4.2.1 History of CRISPR-Cas9

In 1987, Japanese researchers found a series of short direct repeats interspaced with short unique sequences in *E. coli* genomes (216). These clusters of repeated sequences were later referred to as clustered regularly interspaced palindromic repeats, or CRISPRs. They were later detected in several different bacteria and archaea (217) and were eventually implicated in DNA repair (218, 219). The key to their importance came in 2005, when three groups determined that the unique spacer sequences came from plasmids or viruses (220–222). An additional piece of the puzzle also came in 2005, with the determination that genes located near the CRISPR locus encoded proteins with putative nuclease domains, later called the CRISPR-associated (Cas) proteins (220, 222–224). In 2006, it was proposed that the CRISPR locus acted as an adaptive defense system using antisense RNAs as signatures of past viral or phage invaders (225). Barrangou et al. then provided the first experimental evidence that the CRISPR-Cas system was directly involved in the adaptive immunity of bacteria, demonstrating that following viral infection, bacteria directly integrated viral genomic sequences into the CRISPR sites (226). Removing or adding different viral sequences to the CRISPR locus changed the bacterial cell's resistance to certain phages. The CRISPR-Cas system was first harnessed as an easily manipulated gene editing tool in 2013 by Feng Zhang at the Broad Institute of MIT and Harvard (227).

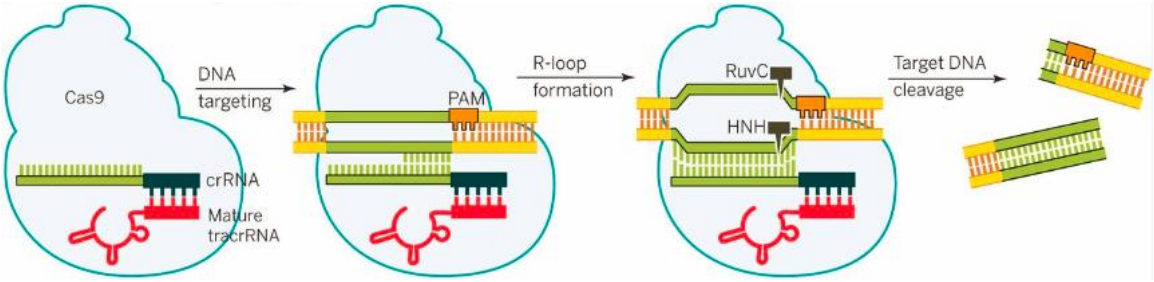
4.2.2 Mechanism of CRISPR-Cas9

The contents of a functional CRISPR-Cas locus are threefold: the CRISPR array of repeats with interspersed DNA-targeting sequences encoding a CRISPR RNA (crRNA), an upstream sequence encoding the trans-activating crRNA (tracrRNA), and an operon of the Cas protein-encoding gene(s) (228). A single crRNA is 20 bp in length and is complementary to the desired target sequence. The tracrRNA forms a secondary structure that is recognized and bound by the Cas protein(s). In gene editing experiments, the crRNA and tracrRNA are often combined via a linker to form a single guide RNA (sgRNA) (Figure 4-1). The Cas protein(s) contain nuclease domains with which to cut the target DNA sequence.

During CRISPR-Cas gene editing, the crRNA binds the target sequence. The linked tracrRNA is then bound by the Cas protein (Figure 4-1), which creates a double-stranded break in the DNA. This DNA break is repaired through a mechanism called non-homologous end joining (NHEJ), wherein the DNA is subjected to a process of random chew-back and fill-in, creating insertions/deletions (indels). These indels frequently give rise to frameshift mutations, which introduce premature stop codons and, presumably, render the encoded protein product truncated and/or non-functional. NHEJ repairs approximately 75% of cuts within two hours, with full repair complete in 24 hours (229). This timeline, combined with the efficiency of

Figure 4-1. The mechanism of CRISPR-Cas9 gene editing. Diagram of a single guide RNA (sgRNA), showing both the crRNA sequence component that guides the sgRNA to its target sequence and the tracrRNA component, recognized by and bound to the Cas9 protein. The Cas9 protein is directed toward the gene of interest. Note the presence of the required protospacer adjacent motif (PAM) sequence, which is the starting site of DNA strand separation. The crRNA binds via sequence complementarity and the two strands of the target sequence separate, forming the “R-loop”. Recognition of the PAM sequence triggers the cleavage activity of Cas9’s RuvC and HNH domains, generating a double-stranded break at the target site.

From Science, Genome editing. The new frontier of genome engineering with CRISPR-Cas9, Doudna JA and Charpentier E, Copyright (2014) (228). Reprinted with permission from AAAS.


Cas9
DNA
targeting
crRNA
Mature
tracrRNA
PAM
R-loop
formation
RuvC
HNH
Target DNA
cleavage

CRISPR-Cas cutting, can result in short turnaround time for gene interruption.

The most commonly used CRISPR-Cas system uses the singly-acting protein Cas9, which uses an HNH cleavage domain to cut the DNA strand complementary to the crRNA and a RuvC-like cleavage domain to cut the opposite strand. The mutation of either cleavage domain results in nickase activity rather than the creation of a double-stranded break, which is often exploited experimentally to reduce off-target cleavage (228). Mutation of both sites results in a targeted DNA binding protein with no cleavage activity; this is often used in conjuction with a transcriptional repressor or activator to alter gene transcription at the target site. Binding of Cas proteins to the target DNA sequence requires the presence of a short protospacer adjacent motif (PAM) sequence; for Cas9, the required sequence is NGG. The PAM sequence is also the starting site of strand separation, with recognition of the PAM sequence triggering Cas9 activity (230).

4.2.3 Targeting a viral gene

Targeting a viral gene with CRISPR-Cas9 presents a unique challenge compared with targeting a cellular gene. While a cellular gene is present in only two copies per cell, there are anywhere from 40 to 100 herpesviral episomes in a single primary effusion lymphoma cell (36, 40). While to date there are no publications directly targeting CRISPR-Cas9 to a gene in KSHV, there is precedent for multi-copy targeting. In 2015, Yang et al. targeted porcine endogenous retroviruses (PERVs) in a porcine kidney cell line (231). Using

droplet digital PCR, they estimated the number of PERVS at 62 copies per cell and managed to successfully target 37% of these by day 17 of CRISPR-Cas9 treatment. In 2016, a research group in the Netherlands successfully targeted EBNA1 in the EBV-positive SNU719 cell line, a line that had previously been shown to contain approximately 800 EBV copies per cell (215, 232) . In 2015, Wang and Quake also successfully targeted the EBV genome in a Burkitt's lymphoma cell line (214). They demonstrated both decreased cell proliferation and increased apoptosis after treatment with multiple guide RNAs. They also showed a 1.4kb deletion between two cuts in EBNA1 that successfully underwent repair ligation within 6 days.

In addition to being present in multiple copies per cell, KSHV genomes are mostly in a state of repressed chromatin, possessing H3K27me3 markings (70, 71). When Van Diemen et al. tried to target HSV-1 during quiescence, they were unsuccessful, citing previous work showing that the HSV-1 genome is kept in a tightly repressed state that is unresponsive to reactivation stimuli (215). If the KSHV genome is as tightly repressed as that of HSV-1, the sgRNA and/or the Cas9 protein might be unable to access the target sequence and successfully make a double-stranded break.

4.2.4 Previous attempts at disrupting LANA

One of the difficulties of targeting KSHV's LANA in particular is the extraordinarily long half-life of the protein. Chen et al. estimated the LANA half-life in KSHV-positive BC-3 cells at approximately 28 hours, while De Leon

Vazquez et al. later found its half-life in BCBL-1 cells to range from 9 to 12 hours (104, 233). A half-life this long makes interpretation of CRISPR-Cas9 targeting results more complicated, as earlier time points might still show active protein present in the cells, while the gene itself has already been successfully altered and is therefore not producing any new functional protein. Previous studies on the disruption of LANA have averted this half-life issue by using a KSHV BAC with the LANA gene deleted altogether (90, 98). Ye et al. found that disruption of LANA led to the loss of episomes in cells and rendered them virus-free after 2 weeks in culture (90). In a similar LANA knockdown experiment using short hairpin RNA (shRNA), Godfrey et al. found that KSHV copy number decreased in the absence of LANA, eventually reaching a plateau where a stable number of KSHV copies remained (42). Both results are consistent with LANA's role in maintaining viral episomes in cells by tethering them to host chromatin. One additional discovery by Godfrey et al. was the reliance of certain PEL lines upon the presence of LANA (and hence, presumably, KSHV) for survival. In particular, they found that two PEL lines (BCP-1 and JSC-1) showed either growth inhibition or apoptosis after treatment with shRNA against LANA (42). This information could make targeting LANA by CRISPR-Cas9 in KSHV-positive PEL lines more difficult, as there would be no reliable way to determine successful gene disruption or loss of protein expression in a population of dead cells. In addition, any cell death observed could be due to either toxicity of the CRISPR system or to KSHV loss, making interpretation complicated.

In 2008, Li et al. determined the effects of LANA deletion of the expression of lytic genes (98). They found that LANA disruption increased the expression of lytic genes, consistent with LANA's role as a transcriptional repressor for the lytic portions of the KSHV genome. Interestingly, they also found an increased number of infectious virions produced compared with cells containing LANA; this supports their findings of increased lytic expression but seems to contradict the earlier work that led to loss of virus from cells. As our current experiments only use CRISPR-Cas genome editing in the context of a stably LANA-expressing cell line, and not in a cell line containing full-length KSHV episomes, we cannot draw our own conclusions about episome loss or increased/decreased viral load as a result of LANA disruption.

4.3 MATERIALS AND METHODS

4.3.1 Cell culture

The BJAB-LANA cell line was a gift from Kenneth Kaye at Harvard Medical School. They established this line by transfecting a LANA-encoding plasmid into BJAB cells, a KSHV-negative B cell line, creating a line that stably expressed LANA (103, 146). BJAB-LANA cells were maintained in RPMI 1640 (Gibco) supplemented with 10% fetal bovine serum, 20 mM HEPES, 2 mM L-glutamine, 0.05 mM β-mercaptoethanol, 200 µg/mL sodium bicarbonate, 100 units/mL of penicillin and 100 µg/mL of streptomycin. The adherent 293T cell line developed from human embryonic kidney cells was obtained from the American Type Culture Collection. These cells were maintained in Dulbecco's modified

Eagle's medium (Gibco) supplemented with 10% fetal bovine serum (Gibco), 100 units/mL of penicillin and 100 µg/mL of streptomycin. They were harvested from culture dishes with 0.05% trypsin after a wash with 1X phosphate buffered saline.

4.3.2 sgRNAs

Three sgRNAs targeting the 5' end of LANA were chosen using the computer software at crispr.mit.edu (developed by the Zhang Lab, MIT) and were labeled AB, CD, and EF. Three sgRNAs targeting a region 1 kb downstream of AB, CD and EF were similarly chosen and labeled αβ, γδ, and εζ. Figure 4-2 illustrates the location of each guide RNA on the LANA gene. Oligos (Table 4-1) were obtained from Eurofins Scientific Testing. We used the original insert between the BsmBI sites (the sgRNA placeholder) as a "Control" scrambled guide RNA that had no target within the LANA sequence.

4.3.3 Annealing and phosphorylation

The forward and reverse oligo for each sgRNA were annealed (A+B, C+D, E+F, α+β, γ+δ, and ε+ζ) and phosphorylated using T4 polynucleotide kinase to create double-stranded DNA fragments at a concentration of 100 µM. The protocol for annealing and phosphorylation is described by Pham et al. (234).

4.3.4 Creation of lentiviral backbone and ligation

The lentivirus backbone, P413 (LV-sgRNA-BsmBI-WTCas9) (Figure 4-3), was a gift from Mazhar Adli at the University of Virginia. The vector contained both ampicillin and puromycin resistance markers. Five µg of vector was cut with

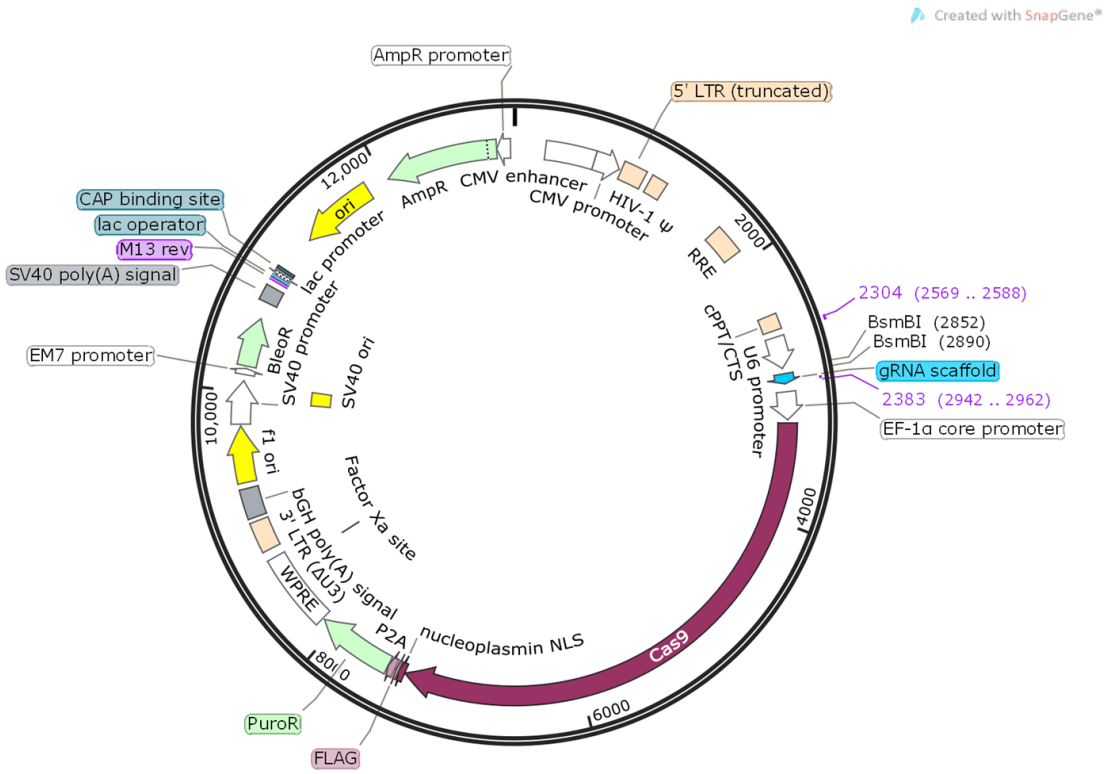
Figure 4-2. Locations of each guide RNA within the LANA sequence. The DNA sequence of LANA, in its native right-to-left orientation, with the sequence of each guide RNA highlighted according to the color legend on the bottom. The three N-terminal guides are located approximately 1 kbp away from the more central guides. Note that the sequence of guide εζ occurs five separate times within the internal repeat region. Figure created by Kerry Johnson and used with permission.

TTATGTCATTTCCTGTGGAGAGTCCCCAGGACCTTGGTTTTCCCCTGGCTGGGTTAATGGCAGGGGCTTTTTAA
ACTTAACTATGGAAGATTGTAGGTTTCCTGCCAGGGGGTGACTAGCTTCCCAGGCTAGGCGGGCCATTTGTAC
TTTCTTACTTGTGTCTTTGTTCTGACAATACACATATACACAATAAGTTATGGGCGACTGGTCTGGTCCAGGGT
GGGGCAAGCAGGACACGGGGCCTGCCTTTACTCCTCCAAACTGGAAGGCCTGAGATAATTTTTTAAGTCCGTA
TGGGTCATTGCCCCAAAAAATCACTGCAAACTTCCATTGACACTTTGGATCTCGTCTTCCATCCTTTCCCAAAA
AGCGTCTATAAAAGATGTGTTGTGGCCTAGCTTTCGCAGGACAATCATCTATCTGTCTGTAAGGGACCGGTGG
TTGTTGGTATCTTGGATGTGGCTTTTTTGGGTGGGTAACTGGAACGCGCCTCATACGAACTCCAGGTCTGTGGG
GTGGTGATGTTCTGAGTACATAGCGGTATTCGCGAGATGGGCCAGGTTGTGGGTCATCGTCTGGTGTATTATC
TCCTGGTGGGCTACTGGCAATTTGTTCATGTGTGCTAACAACAGGGTAATCCACTTCCATTTCGTCCTCGGATG
ACGACCCGTGCAAGATTATGGGCTCTTCCACCGTCTCCTGCTCCTGCTGTTCCACCCCCTGCTGCTCCTGCTCT
TCCACCTCCTCTAACTCCTGCTGCTCCTGCTCTTCCACCTCCTCTAACTCCTGCTCTTCCTGCTCTTCCACCTCCT
CTAACTCCTGCTCTTCCTGCTCTTCCACCTCCTCTAACTCCTGCTCCTCCTGCTCCTCCTGCTCCTGCTCTTGCTC
CTCCACCTCCTCTAATTCCTGCTCTTCCTGCTCCTGCTCTTGCTCTTCCACCTCCTGCTCTTGCTCTTCCACCTCC
TGCTCCTCTAACTCCTGCTCCTGCTCCTCTAACTCCTGCTCCTGCTCCTCTAACTCCTGCTCCTGCTCCTCTAAC
TCCTGCTCCTGCTCCTCTAACTCCTGCTCCTGCTCCTCTAACTCCTGCTCCTGCTCCTCTAACTCCTGCTCCTGC
TCCTCTAACTCCTGCTCCTGCTCCTCTAACTCCTGCTCCTGATCCTCTAACTCCTGCTCCTGCTCCTCTAACTCC
TGCTCCTGCTCCTCCTGCTGCTCCTGCTCCTCCTGCTGCTCCTGTTCATCCTGCTGCTGCTGCTCATCCTGCTGC
TGCTGCTCATCCTGCTGCTGCTGCTCATCCTGCTGCTGCTGCTCATCCTGCTGCTGCTGCTCATCCTGCTGCTGC
TCATCCTGCTGCTCCTGCTCATCCTGCTGCTCCTGCTCATCCTGCTGCTCCTGCTCATCCTGCTGCTGCTCATCC
TGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTCATCC
TGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTCATCC
TGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTCATCC
TGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGTGGCTCC
CGCTGCTGTGGCTCCCGCTGCTGTGGCTCCCGCTGCTGTGGCTCCCGCTGCTGTGGCTCCCGCTGCTGTGGCTC
CCGCTGCTGTGGCTCCCGCTGCTGGGGCTCCCGCTGCTGTGGCTCCCGCTGCTGTGGCTCCTGCTGCTGTGGCT
CCTGCTGCTGTGGCTCCTGCTGCTGTGGCTCCTGCTGCTGTGGCTCCTGCTGCTGTGGCTCCTGCTGCTGTGGCT
CCTGCTGCTGTGGCTCCTGCTGCTGTGGCTCCTGCTGTTGTGGCTCCTGCTGTTGTGGCTCCTGCAGGGGCTCC
TGCTGCTGTGGCTCCTGCTGTTGTGGCTCCTGCAGGGGCTCCTGCTGCTGTGGCTCCTGCTGCTGTGGCTCCTG
CTGTTGTGGCTCCTGCAGGGGCTCCTGCTGCTGTGGCTCCTGCTGCTGTGGCTCCTGCTGTTGTGGCTCCTGCT
GCTGTTGTGAACTTTGGATGCTCAACGTTTTGTTTCCATCGCCCCCGTCCTCCTCGTCCTCCTTCTTGTCCTCCT
CCTCGTCATCCTCCTCGTCCTCATTGTCCTCATCATCGTCATCCTCCTCGTCCTCCTCCTCCTCGTCCTCCTCCTC
GTCCTCCTCCTCGTCCTCCTCCTCGTCATCCTCCTCGTCATCCTCCTCGTCATCCTCCTCGTCATCCTCCTCGTCA
TCCTCCTCGTCATCCTCCTCGTCATCCTCCTCGTCATCCTCCTCGTCATCCTCCTCGTCATCCTCCTCGTCATCCT
CCTCGTCCTCCTCATCTGTCTCCTGCTCCTCCTCATCATCCTTATTGTCATTGTCATCCTTGTCAACCTGACTTTC
CTTGCTAATCTCGTTGTCCCCATTATCCTCGCCAGCCTGATTATTTTCGGAACATTCTTTTTCATTCTTGGATGC
TTCTTCTGCAATCTCCGCAAGGAGCACCAACATGGCTGTGTCATCACCCCAGGATCCCTCAGACGGGGATGAT
GATCCTATGGAGATGGGAGATGTAGGCGGTTGGCGTGGCGGAGTATCGCCATCGCTGGATGATCCCACGTAG
ATCGGGGACTCTGTGGCCCATGGGGGGTACACACTACGGTTGGCGAAGTCACATCTAGGGGGAGAGACTGGG
GGCGACTGACATATTGGGTTTAGTGTAGAGGGACCTTGGGGGGACGATAGCCTTCTTTTTCTCAGGCTACGCA
GGGTAGACGGAGCTAAAGAGTCTGGTGACGACTTGGAGGGAGGCTCGGGTGGAGGAGTCGTGGGTGAGTGTG
GAGGTGTAGTCTGCTGCGAGGGTGGCGGACGCATAGGTGTTGAAGAGTCTGGCCTTCCTGTAGGACTTGAAA
GCGGTGGCCTTTGAGAAGACTCTGGAGACTGCGTGGGTGGCAATGCAGGAGATGGAGAATGAGTATCCGTGG
TCCCCGGAGACACAGGATGGGATGGAGGGATTGGGGAGGAAGACGTGGTTACGGGGGGTAAGAGTGCCGGT
GGAGGTAAAGGTGTTGCGGGAGCGGGTGAAGGAATGGGAGCCACCGGTAAAGTAGGACTAGACACAAATGC
TGGCAGCCCGGATGTGAACACTGTGGGACTTCCAGGTATAGGCAAGGTGTGGGGTCCACATTCCCGGCCGTCG
ATGGAGTCGGCGACATGCTTCCTTCGCGGTTGTAGATGTAGGTCATCGCCAAGGTCACATCTTTCCGGAGACC
TGTTTCGTTTCCTACAACTTCCTCTCGTTAAGGGCGCGCCGGTGCTCCGTCCCGACCTCAGGCGCATTCCCGGG
GGCGCCAT

Table 4-1. Sequences of the oligos used to form all guide RNAs in this study. Table created by Kerry Johnson and used with permission.

sgRNA Oligo	Oligo Sequence
A	CACCGTAGATGTAGGTCATCGCCA
B	AAACTGGCGATGACCTACATCTAC
C	CACCGGGACTTCCTGGTATAGGCA
D	AAACTGCCTATACCAGGAAGTCCC
E	CACCGGTGAAGGAATGGGAGCCAC
F	AAACGTGGCTCCCATTCCTTCACC
α	CACCGTGACGATGATGAGGACAATG
β	AAACCATTGTCCTCATCATCGTCAC
γ	CACCGTGACGAGGAGGAGGACGAGG
δ	AAACCCTCGTCCTCCTCCTCGTCAC
ε	CACCGGGATGACGAGGAGGATGACG
ζ	AAACCGTCATCCTCCTCGTCATCCC

Figure 4-3. Map of the CRISPR-Cas9 lentiviral construct used in this study. Lentiviral vector P413 was a gift from Mazhar Adli at the University of Virginia and conveniently integrates both the Cas9 protein and a single guide RNA sequence. Important components to note include the U6 promoter for the guide RNA, BsmBI sites for easy crRNA sequence insertion, remainder of the sgRNA scaffold, Cas9-encoding gene, nuclear localization signal (NLS), and puromycin resistance gene. Map created using SnapGene Viewer (GSL Biotech).



P413 LV-sgRNA(BsmBI-WTCas9)
13,026 bp

25 units of the restriction endonuclease BsmBI to open the vector for sgRNA insertion. Directional cloning is made possible by the nature of the BsmBI recognition sequence and its distal cut site.

Each sgRNA was ligated into P413 as outlined in Pham et al. (234). The ligation reaction was left at room temperature for approximately 1 hour and then was incubated at 65°C for 10 minutes to inactivate the ligase.

4.3.5 Transformation and mini/maxi prep

The ligation was transformed into Stbl3 E. coli (Thermo Fisher Scientific) using 5 μL of the ligation and 50 μL of the Stbl3s. Stbl3 cells were chosen as they have good fidelity for retaining repeat regions such as those found in lentiviral backbones. The transformation was then placed on ice for 30 minutes after which it underwent a heat shock at 42°C for 30 seconds. It was then left at room temperature for 2 minutes after which 950 μL of LB was added. The mixture was placed in a shaker set to 37°C for an hour and was then plated onto ampicillin plates and left to grow overnight in an incubator set to 37°C. Colonies were chosen and each lentiviral vector with sgRNA insert (LV-sgRNA) was then mini prepped following manufacturer's instructions (QIAGEN).

4.3.6 Cloning confirmation assays

The LV-sgRNA mini preps were cut with BsmBI at 55°C for 1 hour. This was done to ensure the sgRNA insert had correctly ligated into the backbone. If the insertion was successful, the BsmBI restriction should not cut anywhere as

the BsmBI restriction sites would have been destroyed in the initial vector digest. Plasmids with successful insertions were then maxi prepped following manufacturer's instructions (QIAGEN). All of the LV-sgRNAs were then sent for Sanger sequencing using the U6 primer to determine whether they contained the correct insert.

4.3.7 Lentivirus production

The lentiviruses were created using 3 plasmids: the LV-sgRNA; psPAX, a plasmid encoding the gag-pol packing proteins (gifted to us by Mazhar Adli at the University of Virginia); and pCMB-VSV-G, a plasmid encoding the VSV-g glycoprotein (gifted to us from Bob Weinberg, Addgene plasmid #8454), at a ratio of 4:3:1, totaling 30 μg of DNA per 10 cm dish. The original P413 lentiviral vector received from the Adli lab prior to cutting with BsmBI was used as a control "scrambled" gRNA, as the sequence in the insert region would not direct the system to anywhere on the LANA sequence. The DNA was transfected into 293T cells using Lipofectamine 2000 (Thermo Fisher Scientific; amounts specified by the manufacturer) to produce the completed lentiviruses. Lipofectamine-containing media was removed after 16 hours and replaced with fresh media.

On day 2 after transfection, the lentiviruses were harvested from the media and new media was put back on the 293T cells. On day 3 after transfection, both the cells and media were harvested and stored at 4°C. On day 4 the cells/supernatants were centrifuged at 5000 RPM for 5 minutes to clear the cells and debris. The lentiviruses were then concentrated via ultracentrifugation

at 25,000 RPM for 2 hours and were resuspended in 500 μL of RPMI without FBS. The viruses were then quantified using a p24 ELISA assay (Lenti-X p24 Rapid Titer Kit, Takara Clontech).

4.3.8 Spinfection and puromycin selection

The lentiviruses were transduced into BJAB-LANA cells at an MOI of 10. Polybrene was added at a concentration of 8 μg/mL to mediate the charge difference between the viruses and the cells. All conditions were centrifuged at 3,000 RPM for 30 minutes at 32°C. After 48 hours, puromycin was added at a concentration of 4 μg/mL to select for the cells that had received and integrated the lentiviral constructs.

4.3.9 Immunostaining

The samples were prepared for immunofluorescence microscopy and immunostained for LANA. The cells were fixed in 4% paraformaldehyde for 10 minutes. After a 20-minute rat IgG block at 4mg/mL, Alexa Fluor® 488-conjugated anti-LANA LN53 antibody (152) was applied at a concentration of 1:1000 in saponin-based BD PermWash (BD Biosciences) and incubated for 40 minutes. Cells were washed in PermWash and DAPI was then applied at 0.5 mg/mL for 15 minutes. Cells were deposited onto slides using a Shandon Cytospin 2 for 4 minutes at 500 rpm and mounted in Fluoro-Gel (Electron Microscopy Sciences).

4.3.10 Fluorescent microscopy

Microscopy was performed using a Nikon Eclipse TE2000-E fluorescent

microscope. DAPI was viewed with a 90 ms exposure time. LANA staining was viewed using the FITC (fluorescein isothiocyanate) channel with a 500 ms exposure time.

4.3.11 Whole cell PCR

For each PCR sample, approximately 100,000 cells were spun down and resuspended in distilled water. The cells were then boiled at 100°C for 10 minutes. 2 μL (1.6 units) of Proteinase K was added to the cells along with buffer and distilled water. The cells were then incubated at 50°C for 1 hour, boiled again at 100°C for 10 minutes, and placed on ice for 10 minutes. Samples were then spun at 13,000 RPM for 1 minute and the supernatants were collected. Each 50 uL PCR reaction used 5uL of this supernatant as a template. The PCR reaction was run using the program displayed in Table 4-2 using Taq polymerase (New England Biosciences) and the primers shown in Table 4-2.

4.3.12 Sequencing

PCR reactions were run on an agarose gel and the appropriate bands were cut out and purified (Wizard SV Gel and PCR Clean-Up System; Promega). Samples were sent to GenScript or GeneWiz for Sanger sequencing using primer CC-PCR-AF-Rev.

4.3.13 TA cloning and sequencing

DNA extracted from the gel-purified bands was also used as insert in a TA cloning reaction (TOPO TA Cloning Kit, Invitrogen) at an insert:vector ratio of 3:1.

Table 4-2. PCR program and primers used for whole-cell PCR. PCR steps marked in bold italics were repeated for 35 cycles. Table created by Kerry Johnson and used with permission.

Temperature (in ^{o}C)	**Time (MM:SS)**
95	10:00
95	***00:30***
53	***00:30***
68	***01:00***
68	05:00
Primer Name	**Sequence**
CC-PCR-Fwd	GGGGAGGAAGACGTGGTTAC
CC-PCR-AF-Rev	CGCCCTTAACGAGAGGAAG

The ligation was transformed into DH5α E coli (Thermo Fisher Scientific) and colonies were grown and mini prepped (Qiagen). Mini prep DNA was sent for sequencing using the M13 forward sequencing primer for the TA vector.

4.4 RESULTS

4.4.1 CRISPR-Cas9 targeting of LANA results in a decrease in LANA immunostaining in a LANA-stable BJAB cell line

We first sought to determine whether targeting of LANA by CRISPR-Cas9 would decrease the amount of LANA within BJAB-LANA cells. We designed 3 guide RNAs within 275 bp of the LANA translational start site; guide RNA AB initially failed to ligate into the vector and so was not included in this early experiment. In addition, we had a control guide RNA with a scrambled sequence designed not to cut within the LANA gene. Approximately 2 months after the initial transduction with lentiviruses, cells were immunostained for LANA and examined by epifluorescence microscopy.

Staining controls for this experiment are presented in Figure 4-4. BJAB cells that never contained the LANA plasmid stained negative for LANA, resulting in no background staining (Figure 4-4A). BJAB-LANA cells that did not receive the CRISPR-Cas9 system had strong diffuse nuclear LANA staining, with nearly 100% of cells staining positive for LANA (Figure 4-4B). Experimental staining results are presented in Figure 4-5. BJAB-LANA cells which had received the CRISPR-Cas9 system with the Control (scrambled) guide RNA (Figure 4-5A)

Figure 4-4. Staining controls for the CRISPR-Cas9 editing of LANA. (A) Two representative fields of both the DAPI and LANA staining in BJAB cells lacking the LANA plasmid, showing the absence of any background LANA staining. **(B)** Two representative fields of both DAPI and LANA staining in untreated LANA-containing BJAB-LANA cells, showing positive LANA staining. Figure created by Kerry Johnson and used with permission.

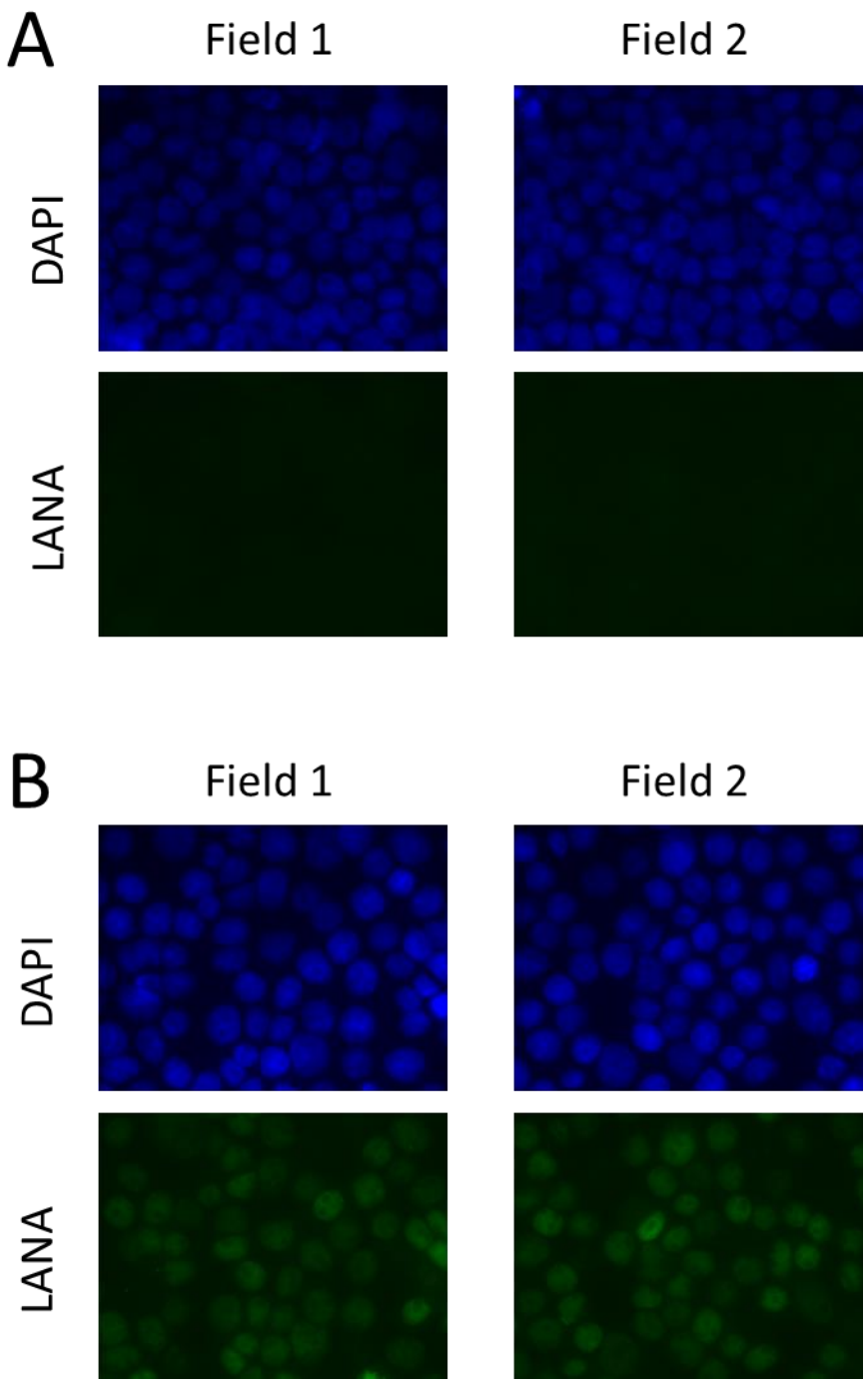
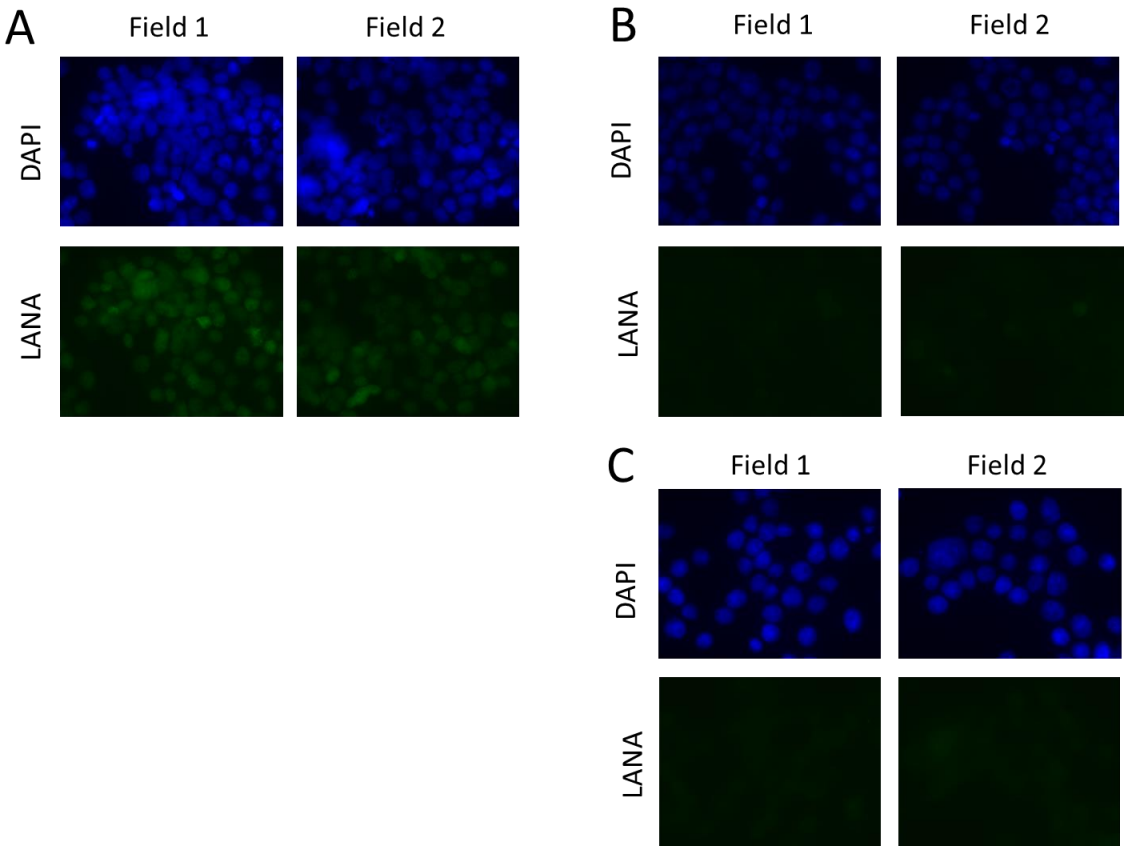

A
Field 1
Field 2
DAPI
LANA
B
Field 1
Field 2
DAPI
LANA

Figure 4-5. CRISPR-Cas9 targeting of a single site in LANA in BJAB-LANA cells decreases the amount of LANA visible by immunofluorescence. (A) BJAB-LANA cells treated with the Control (scrambled) guide RNA, showing nearly 100% of cells expressing the LANA protein. **(B-C)** BJAB-LANA cells treated with the CD **(B)** or EF **(C)** anti-LANA guide RNAs, showing a marked decrease in the amount of visible LANA staining compared to **(A)**. Figure created by Kerry Johnson and used with permission.


A
Field 1
Field 2
DAPI
LANA
B
Field 1
Field 2
DAPI
LANA
C
Field 1
Field 2
DAPI
LANA

showed the same staining pattern and intensity as cells which did not receive the CRISPR-Cas9 system. However, cells receiving the system with either guide RNA CD (Figure 4-5B) or guide RNA EF (Figure 4-5C) showed a marked decrease in the amount of LANA staining compared to these two controls. In fact, these samples showed a nearly complete loss of LANA staining. This suggested that the LANA knockdown was successful and due exclusively to the specific targeting of LANA DNA by the CRISPR guide RNAs.

4.4.2 Sequencing across the Cas9-induced cut shows the expected mixture of insertions/deletions

To verify that CRISPR-Cas9 cutting of the LANA gene had occurred in conjunction with the expected random indels, we used PCR to amplify the section of the gene encompassing the targeted sites and sequenced the resulting products. Because these PCR products represented a pool of multiple indels from multiple different LANA plasmids, we expected to see a clean consensus sequence up until the region of the cut, at which point we expected to see messy/unclear sequence resulting from the multiple overlapping indel sequences.

Figure 4-6 summarizes the results of this sequencing. Figure 4-6A shows the sequence obtained from treatment with the Control guide RNA (top) as compared with the sequence obtained from treatment with the CD guide RNA (bottom); the highlighted yellow/orange region marks the location of guide RNA

Figure 4-6. Bulk sequencing results show editing of the LANA gene with CRISPR-Cas9. (A and B) Sequencing results for the region of LANA targeted by guide RNAs CD **(A)** or EF **(B)** as compared against the Control (scrambled) guide RNA. The Control sample shows clean LANA sequence, while both the CD and EF samples show clean, reinforced sequence up until the point of the expected cut, followed by a mixture of sequence signals. sgRNA CD is highlighted in orange; sgRNA EF is highlighted in yellow.

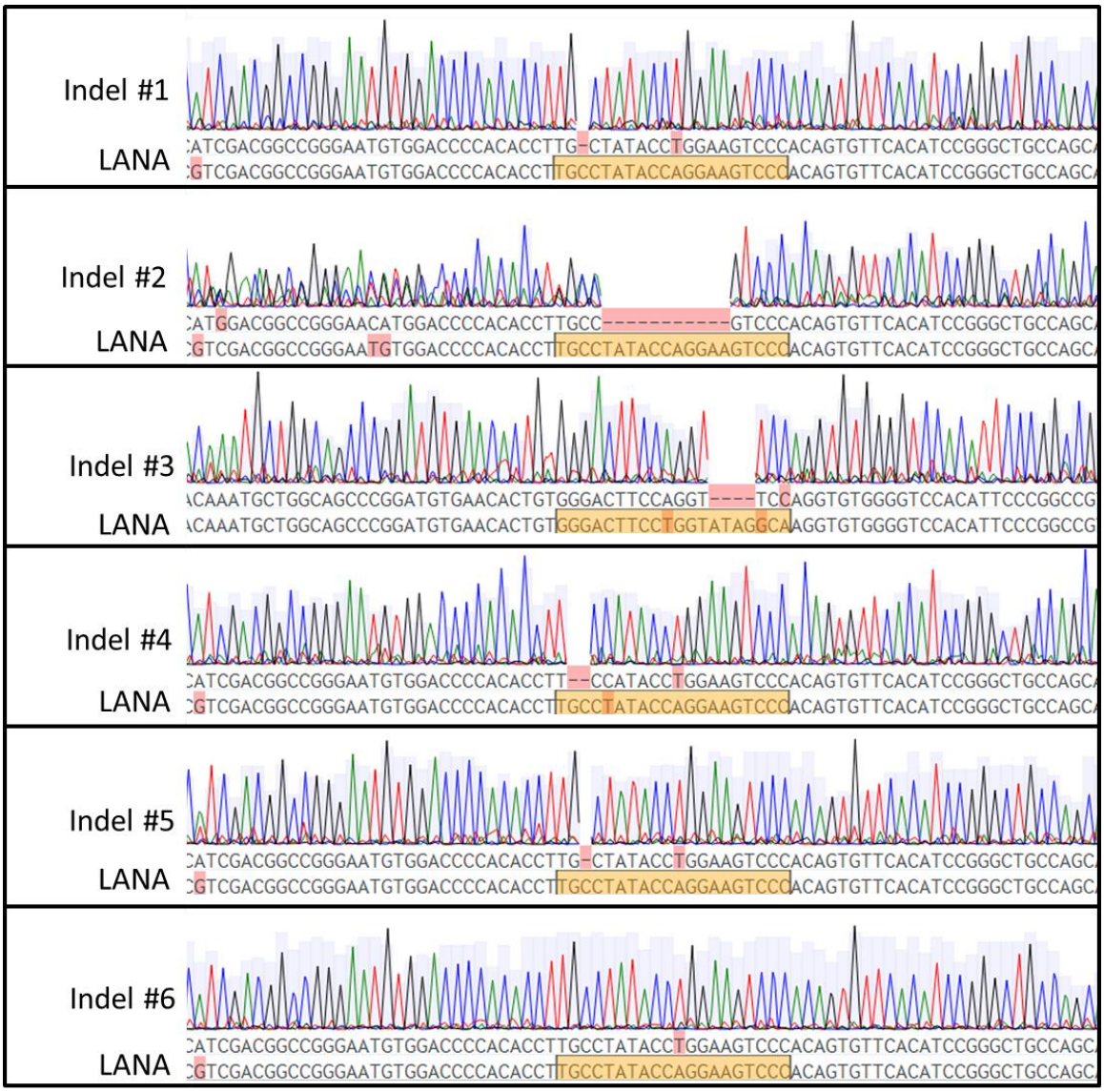

A
Control (scrambled) sgRNA
TGTGGACCCCACACCTTGCCTATACCTGGAAGTCCCACAGTGTTCACATCCGGGCTGCCAGCATTTGTGTCTAGTCCT
CD anti-LANA sgRNA
TGTGGACCCCACACCTTGCC-ATACCTGGAAGTCCCACAGTGTTCACATCCGGGCTGCCAGCATTTGTGTCTAGTCCT
B
Control (scrambled) sgRNA
TGCCAGCATTTGTGTCTAGTCCTACTTTACCGGTGGCTCCCATTCCTTCACCCGCTCCCGCAACACCTTTACCTCCAC
EF anti-LANA sgRNA
TGCCAGCATTTGTGTCTAGTCCTACTTTACCGGTG----CCATTCTTCCTCCCGCTCCTCCAACACCTTTACTTCCAC

CD. While the Control sequence remains clear and reinforced for the entire span, the sequence from CD-treated cells shows a one base pair deletion, followed by a mixture of sequence signals. Figure 4-6B shows a similar result for guide RNA EF, with an even more extreme 4 bp deletion and several mismatched base pairs within the remaining unreinforced sequence. These results verified that LANA gene editing had occurred at the predicted location of the guide RNA. Furthermore, it appears that different plasmids within the sample were edited differently, leading to a lack of strong reinforced consensus sequence in the results.

To visualize the different individual indel edits undergone by the various LANA plasmids, we took the pool of PCR products from the CD-treated cells and performed TA cloning, in which only a single PCR product (the result of a single plasmid's indel) would ligate into each TA vector. These ligations were transformed into bacteria and single colonies were picked, each containing a TA vector with a single indel pattern. Extracted DNA was sent for sequencing using the TA vector primer.

The resulting sequence patterns for all 6 colonies are shown in comparison to unedited LANA sequence in Figure 4-7. Overall deletions ranging from 1 to 11 base pairs can be seen, along with nearby base pair substitutions. This size indel is similar to that noted by Yang et al. in their CRISPR-Cas9 targeting of porcine endogenous retroviruses (231). They saw indels ranging

Figure 4-7. Sequencing results showing 6 independent indels created by CRISPR-Cas9 cutting of LANA. Shown are 6 indels, each within one box, created by the CRISPR-Cas9 targeting of LANA with sgRNA CD, the sequence of which is highlighted in orange. “LANA” denotes native LANA sequence, while the trace and underlying base calls denote sequencing results for the indels. Remarkably, there are different degrees of editing visible, spanning from an 11-bp deletion (Indel #2) to no editing/a perfectly conserved edit (Indel #6). Note that the sequence for Indel #3 is the reverse complement of the other examples; also note the single incorrect nucleotide (our reference sequence has an A where the sequencing results find a T) within the highlighted guide RNA region.


Indel #1
LANA
Indel #2
LANA
Indel #3
LANA
Indel #4
LANA
Indel #5
LANA
Indel #6
LANA

from 1 to 148 bp, with 80% of indels being smaller than 9 bp. Of note in our data is Indel #6, in the last box, which either did not undergo any gene editing or experienced an indel equivalent to its original sequence. These results demonstrated the wide variety of indels experienced by the individual LANA plasmids within a pool of CRISPR-Cas9-treated cells. The sequence changes observed explain the loss of LANA immunostaining in these cells, as any gene products would either be truncated or severely altered in their amino acid sequence.

4.4.3 Targeting 2 or more sites within the LANA gene did not show increased LANA clearance over single cuts

In 2014, Wang and Quake found that increasing the number of CRISPR-targeted cuts in DNA had the greatest impact on cell proliferation and apoptosis (214). Van Diemen et al. also noted increased herpesviral clearance from cells using two guide RNAs instead of just one (215). To determine whether we would see improved LANA clearance with two guide RNAs over one, we generated three additional guides (αβ, γδ, and εζ) targeting a region of LANA approximately 1 kb downstream of guides AB, CD, and EF (Figure 4-2). We had nine separate conditions, each containing one guide from AB, CD, and EF and one guide from αβ, γδ, and εζ. In addition, we had one condition using all six guide RNAs together and one condition just using the εζ guide RNA; since εζ targets the repeat region of LANA and cuts in 5 separate places, we reasoned that it might result in higher LANA clearance than the other conditions.

The immunofluorescence results of our experiments 36 days after spinfection with more than one guide RNA are shown in Figure 4-8. In all conditions, there is a decreased amount of LANA compared with the Control (scrambled guide RNA) in Figure 4-8A. Quantification of the number of LANA-positive cells per condition is displayed in Table 4-3. While all conditions contain fewer LANA-positive cells than the Control sample, they do not show a higher LANA clearance than the singly-targeted examples in Figure 4-5. This may be due to the difference in timing of the analysis – the doubly-targeted samples were fixed and immunostained 36 days after spinfection of the CRISPR-Cas system, whereas the singly-targeted samples were analyzed 2 months after spinfection. This allowed nearly an additional 30 days for the single-guide CRISPR-Cas system to continue to target and disrupt the LANA-encoding gene, possibly increasing the visible effects.

While double cuts did not show greater LANA clearance than single cuts in comparing the two experiments, we did find that in the multiple-cuts experiment, those conditions with greater than 2 cut sites (εζ alone and all guides combined) (Figure 4-8E) had a higher LANA clearance than those with only 2 cut sites, in agreement with Wang and Quake's observation that all 7 of their guides combined had the greatest overall effect on cell proliferation and apoptosis (214). We also noted that conditions with the highest LANA clearance were those

Figure 4-8. CRISPR-Cas9 targeting two or more sites within LANA simultaneously decreases the amount of visible LANA protein. (A) DAPI and LANA staining for cells in the Control (scrambled) guide condition, showing positive LANA staining. **(B, C, and D)** DAPI and LANA staining for cells treated with guides AB **(B)**, CD **(C)**, and EF **(D)** in combination with the guides to the middle section of LANA (αβ, γδ, and εζ). LANA staining shows variable results, with some conditions exhibiting a greater loss of LANA than others. **(E)** Staining results for the conditions with guides cutting at more than two sites simultaneously. Note the nearly complete absence of LANA staining in the εζ condition, which cuts five times in the LANA internal repeat region. Figure created by Kerry Johnson and used with permission.

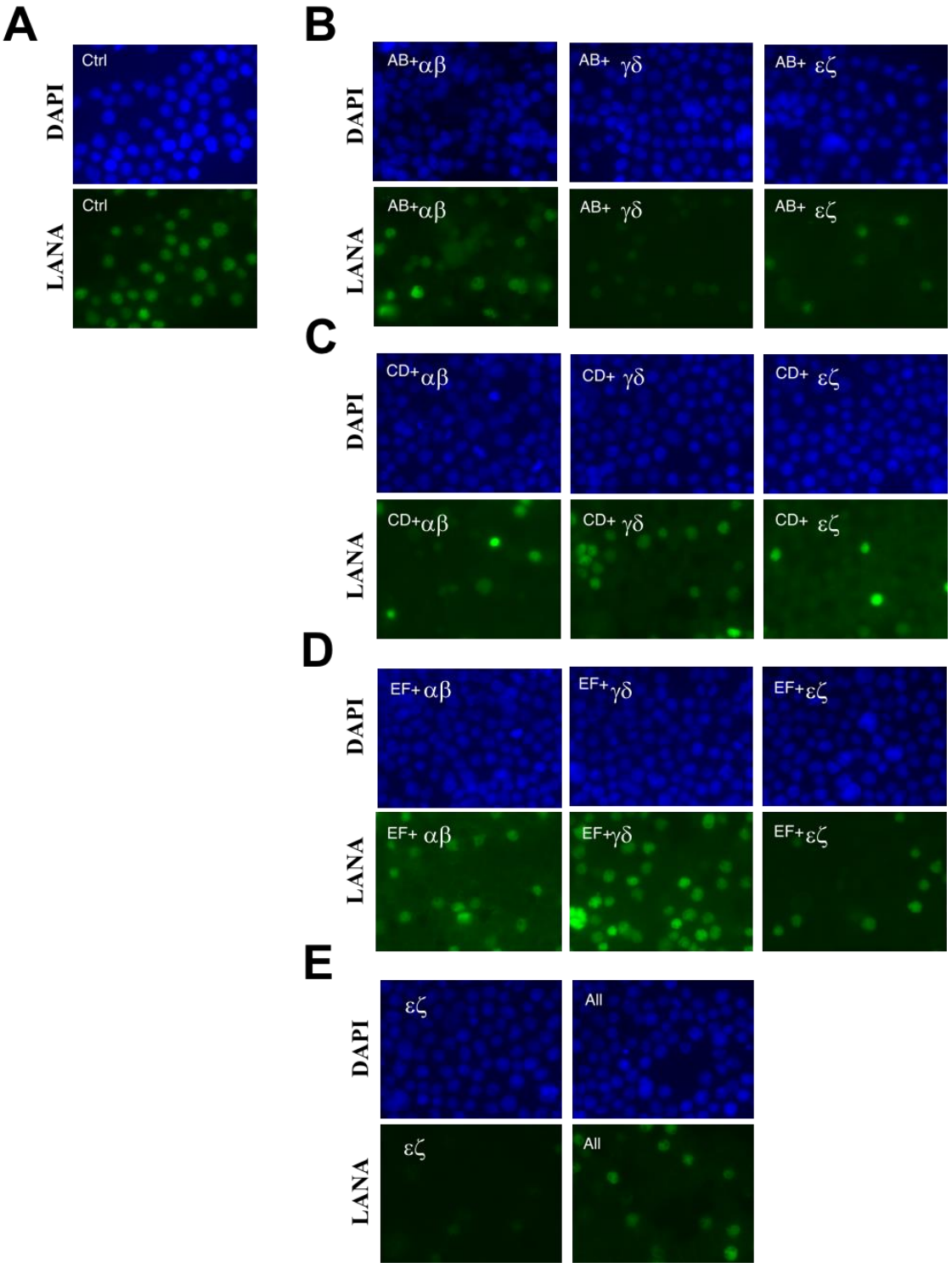

A
DAPI
Ctrl
LANA
Ctrl
B
DAPI
AB+αβ
AB+ γδ
AB+ εζ
LANA
AB+αβ
AB+ γδ
AB+ εζ
C
DAPI
CD+αβ
CD+ γδ
CD+ εζ
LANA
CD+αβ
CD+ γδ
CD+ εζ
D
DAPI
EF+αβ
EF+γδ
EF+εζ
LANA
EF+ αβ
EF+γδ
EF+εζ
E
DAPI
εζ
All
LANA
εζ
All

Table 4-3. The percentage of LANA-positive cells in each CRISPR-Cas9 condition, in order of decreasing effectiveness at LANA clearance. Table created by Kerry Johnson and used with permission.

Experimental Condition	% LANA Positive Cells
Ctrl (“scrambled”)	85%
εζ	2%
CD+εζ	6%
AB+εζ	10%
All lentiviruses	15%
EF+εζ	17%
CD+γδ	20%
AB+γδ	20%
EF+αβ	22%
CD+αβ	22%
AB+αβ	31%
EF+γδ	43%

containing guide RNA εζ, which cuts in 5 places in the internal repeat section of LANA. Again, this seems to corroborate the finding that multiple cuts are more efficient at gene disruption than only one or two cuts.

4.4.4 Targeting a repeated sequence in the LANA internal repeat region shows the greatest LANA clearance by immunostaining

In selecting the set of 3' LANA cut sites, we obtained a guide RNA (εζ) that recognized 5 identical sites in the internal repeat region of LANA. Strikingly, the 5 conditions with the greatest LANA clearance were those utilizing this εζ guide RNA, as shown in Table 4-3. This led us to conclude that targeting the repeat region of the gene was the most effective method of eliminating LANA expression using CRISPR-Cas9, even moreso than having 2 independent cuts spaced farther apart on the gene.

4.5 DISCUSSION

While herpesviruses have previously been targeted by CRISPR-Cas, to date there have been no published applications to KSHV in particular. Here we show the successful targeting and cleavage of the KSHV tethering protein gene LANA by CRISPR-Cas9 as found in the LANA-stable, KSHV-negative B cell line BJAB-LANA. We show nearly complete clearance of the LANA protein at 2 months after application of the CRISPR-Cas system using single guide RNAs. We confirm the presence of these cuts by seeing their various indel repairs through sequencing across the affected region of the gene. While targeting two independent cut sites in the gene did not seem to result in higher protein

clearance than targeting a single cut site, this may have been due to the difference in the timing of analysis for these samples. The singly-cut samples were immunostained almost 30 days after the doubly-cut samples, giving the CRISPR-Cas system more time to target, cleave, and repair additional copies of LANA. Conditions targeting more than 2 cut sites simultaneously showed much higher LANA clearance than those targeting only 2 sites, suggesting that introducing multiple disruptions is more effective at introducing protein truncations. This was especially true of conditions containing guide RNA εζ, which targeted 5 sites in the internal repeat section of LANA, suggesting that targeting repeats might be the most effective way of disrupting a gene. Our increased protein clearance resulting from targeting multiple sites at once agrees with previous findings by Wang and Quake, who noted that the condition using all 7 of their guide RNAs against EBV simultaneously led to the greatest decrease in cell proliferation and greatest increase in apoptosis (214).

One potential problem in using CRISPR-Cas to disrupt a gene is the generation of escape mutants, seen by Van Diemen et al. in their experiments (215). This condition arises when the NHEJ resolution following DNA cleavage results in a conservative, in-frame mutation that does not generate a premature stop codon. In these cases, the gene is not only transcribed normally, and likely produces a functional protein, it has also rendered itself immune to any further targeting by the guide RNA due to its sequence change. Of course, a repair with a perfectly conserved indel is also possible and would allow for future targeting of

the gene by the CRISPR-Cas system. In such a case, as in Figure 4-7 Indel #6, it is impossible to know whether the sequence of that particular LANA plasmid was unedited or experienced a perfectly conserved indel.

Here, we demonstrate the successful CRISPR-Cas targeting of a multi-copy viral gene in cells. While the experiments laid out here did not include cells harboring the full KSHV episome, it is easy to see how this work, in principle, could be applied to such a condition. One potential issue in targeting a latent viral episome was demonstrated by Van Diemen et al. in their attempts to target HSV-1 during quiescence (215). Their CRISPR-Cas approach to this virus was ineffective, and they posited that this was due to the tightly repressed state of the viral genome, which had previously been unresponsive to any reactivation stimuli. The same issue could arise with KSHV, of which the majority of genes are repressed during the latent phase of the virus. This would necessitate targeting those genes that are freely transcribed during latency, such as LANA.

With the available lentiviral route for introducing the CRISPR-Cas9 system into cells, it is not a far reach to see potential therapeutic possibilities for latent tumor viruses in infected individuals. One of the largest hindrances to therapeutic use of CRISPR-Cas is the potential for off-target effects – binding and cleavage of untargeted genes due to identical, or even close, matches to the guide sequence. Kuscu et al. found that such off-target effects were enriched in areas of the genome with an open chromatin configuration (235). They also found a much lower rate of indels in off-target locations than at target sites and

reiterated the importance of the nearby PAM sequence in successful binding, a factor that makes off-target cleavage less likely. In the same year, Wu et al. found that a seed match triggered binding of Cas9, but that more extensive base pairing was necessary for actual DNA cleavage (236). For that reason it has been suggested that shorter guide RNA sequences might result in fewer off-target effects due to there being fewer potential partial base pair matches possible. Van Diemen et al., in their work on multiple different herpesviruses, showed no off-target effects for their top 9 predicted sgRNAs (215). Thus, while off-target cleavage is important to consider when targeting genes in living organisms, the literature suggests it is rare, perhaps suggesting that CRISPR-Cas poses less of a treatment risk to the host genome than initially expected.

CHAPTER 5. GENERAL SUMMARY, DISCUSSION, AND FUTURE DIRECTIONS

5.1 SUMMARY

To date, research on the KSHV LANA tether has been limited by the constraints of conventional light microscopy. In the first part of this work, we overcame this limit by applying the super-resolution technique of dSTORM to the study of the LANA tether. We determined that the size and overall shape of tethers is consistent within a single cell line. Interestingly, we also determined consistency among tethers in two drastically different cell lines harboring the same KSHV strain. This demonstrates that the folding of a LANA tether depends upon the virus itself and not upon the cellular environment. We showed an increase in the size of tethers with increasing number of TRs, suggesting that LANA binding does not reach a "plateau," at least up to and including systems with 21 TRs. More importantly, we showed that this increase is consistent with the folding behavior of a polymer, with a scaling constant suggestive of active chromatin. This finding clarifies the conflicting literature on the chromatin folding state of the TR region.

We determined that each of our 2TR tethers was resolvable into two distinct clusters of emissions, something that would never have been noticed without the use of super-resolution microscopy. Upon using the protein folding prediction program PairCoil2, we found a coiled-coil domain through the central region of the LANA dimer, which would mask all but three of the binding epitopes for our antibody. This was consistent with our emission counts from the 2TR dSTORM data, which were much lower than would be expected for all 22 binding

epitopes exposed. Upon creating a model of a 2TR tether, we demonstrated the importance of nucleosome phase in the relative positions of the two TRs and, hence, the relative positions of the two clusters of antibody emissions. This information, combined with the extreme DNA bend necessary to accommodate our data, demonstrated the full LANA occupancy of all three LBSs per TR. A model in which only two LBSs are occupied did not fit our data due to a reduced DNA bend from the binding of only two LANA dimers per LBS.

Preliminary work studying the LANA tether in mitosis showed a condensation of the tether during condensation of cellular chromatin. This suggests that TR chromatin is responding to the same condensation stimuli as the host cell. We found that the LANA tethers in a single cell seemed to orient themselves with respect to the metaphase axis. Again, this suggests that the stimuli influencing the condensation of cellular DNA are also influencing the motion and directionality of the TR DNA. The localization of LANA tethers on the outside of condensed chromatid arms, rather than buried within them, implies that chromatin condensation occurs prior to LANA tether attachment during mitosis rather than afterward.

We also performed a proof-of-principle study on clearing LANA from cells using CRISPR-Cas9 genome editing. We successfully cleared LANA from a KSHV-negative, LANA-positive cell line using single guide RNAs. We showed the presence of various indels at the predicted cut sites, demonstrating successful cuts and subsequent gene editing. While two simultaneous cuts in two

different locations on the LANA gene did not show increased protein clearance over a single cut, this may have been due to differences in post-treatment observation times. Finally, we determined that targeting a repeated sequence in the internal repeat region of LANA showed the greatest extent of LANA clearance out of all multiple-cut conditions. This suggests the utility of targeting repeat sequences as an efficient method of gene destruction, a concept that might be applied to other viruses and have therapeutic potential downstream.

5.2 DISCUSSION AND FUTURE DIRECTIONS

5.2.1 Determining the path of viral DNA in a LANA tether

Our current model of a 2TR tether draws information about the underlying DNA and its trajectory from previously obtained crystal structures of a tetranucleosome and the C terminus of LANA bound to LBS1 (125, 175). Our model also demonstrates the impact of changes in nucleosome phasing to the relative positions of tether-bound LANA proteins. Ultimately, however, we are making assumptions about the trajectory of TR DNA in a tether, as our LN53 antibody emissions are separated from the DNA by each LANA dimer's long coiled-coil region. To determine the DNA trajectory, and to test our model, we propose studying 2TR tethers incorporating a LANA with a C-terminal V5 tag, described by Toptan et al. (111). Having this tag would allow us to locate the LANA C termini (rather than the more distal internal repeat regions) and, hence, the positions of the LBSs themselves. If we did this for tethers with different numbers of TRs and calculated R_g values for each, we could once again determine the packing state of this underlying chromatin, as laid out by Boettiger et al. (173). We would expect to confirm our findings that the region folds as active chromatin. However, it is possible that the V5 tag on LANA could interfere somehow with how the chromatin folds, possibly preventing the binding of other proteins often found in the region, such as ORC.

One additional avenue of further study is in exploring the contributions of TR histone acetylation to the "active" folding conformation of the LANA tether.

While the region certainly contains such histone modifications (71, 72, 143), it is unclear how much of the folding is driven by these marks versus how much is caused by being “open” due to the positions of nucleosomes and the binding of additional proteins, such as ORC. We propose treatment of systems containing different numbers of TRs with histone deacetylases to remove such activating marks from the region, followed by calculation of R_g and determination of the scaling exponent indicative of folding state. Both native KSHV and the BAC16 could face problems in this experiment, as a variety of histone modifications are used to control whether the viral episome is in a latent or lytic state. The application of histone modifying proteins could disrupt this delicate balance and potentially modify viral gene expression patterns. As such, this experiment would likely need to be restricted to systems using only TR-containing plasmids.

5.2.2 Localizing the N terminus of LANA in a 2TR tether

Our current model of a 2TR tether focuses on what is going on between LANA’s C terminus and the site of LN53 binding (the N terminus of its internal repeat region). It currently provides no information on the positions of the associated N termini of the proteins. To determine this, we propose a two-color dSTORM experiment, looking at the relative positions of the LN53 antibody and the commercially available 4C11 antibody (Novus Biologicals), which recognizes LANA’s N terminus. One might predict that the more N-terminal antibody would generate a larger R_g value than the internal LN53 antibody, as it is positioned farther away from the site of LBS binding. Alternatively, in place of the LN53

antibody, one could use the LANA construct containing a V5-tagged C terminus and compare that R_g value with that of the 4C11 antibody, expecting to see an even more extreme difference than with LN53. One might repeat this experiment using a truncated LANA construct, missing its entire internal repeat region, with a C-terminal V5 tag, described by Garber et al. (126). This should create less drastic difference between the R_g values obtained for the V5 tag and the 4C11 antibody.

It is possible, however, that beyond the coiled-coil region of the LANA dimer, the remainder of the protein is flexible, allowing it to fold back on itself. In fact, a model by Chiu et al. shows LANA N termini looping back and attaching to the histones of the episome itself, rather than branching outward to contact host histones (121). If this is the case, we might not see the expected difference in R_g values between the two antibodies, or between 4C11 and the V5 tag.

One idealistic experiment for future consideration would involve “tracing” the LANA protein using all three markers (the 4C11 antibody, the LN53 antibody, and the V5 tag) simultaneously, in a three-color dSTORM experiment. Such a study presents plenty of technical problems, including obtaining clear channel separation and finding appropriate fiduciary markers for calibrating the z-axis that would be visible in all three channels. It would, however, offer the ultimate look at three different locations on LANA and allow comparison of their relative positions relative to the LBS DNA, effectively tracing the trajectory of the proteins.

5.2.3 Developing a tool to analyze tethers with any number of TRs

Our experiments analyzing the R_g of tethers with varying number of TRs have been restricted by the availability of TR-containing plasmids. Recombination events in DH5α E. coli serendipitously created plasmids with 0, 2, 5, and 7 TRs, complementing our 8TR construct and BCBL-1 cells with wild-type virus. Use of the BAC16 then provided an intermediate point, at approximately 21 TRs. While we successfully demonstrated an increase in R_g with increasing TR number using these tools, it would, of course, be beneficial to fill in the data for every point along the way. The Smith Lab has been working on developing a system in which any number of 801-bp TRs could be arranged tandemly using a fundamental single-TR building block. This system consists of native sequence containing all three LBSs surrounded by the sequence used to determine the crystal structure of a tetranucleosome (175), to a total of 801 bp. Compatible ends on this sequence allow for ligation of multiple units in tandem, to create a plasmid construct containing any number of desired TRs. This system provides a potential tool for filling in the gaps in our data, but it also opens the door to modeling 1TR or 3TR tethers in the future. (Modeling tethers with even greater numbers of TRs becomes less and less feasible, as it becomes nearly impossible to determine the order in which the TRs are connected.) One possible hindrance in working with a 1TR system is the difficulty in distinguishing its very few antibody-associated fluorophore emissions over background. This could

especially become a problem when using the 4C11 antibody, which has only a single recognition site per LANA protein, unlike the LN53 antibody.

Another set of possible experiments using this system of TR building blocks could involve altering the DNA sequence surrounding the LBSs, either in nature or in length, to see the effects on the relative positioning of bound LANA proteins. Changing the nature of the sequence itself might render it more or less favorable as a site for wrapping around nucleosomes. A change in the number of nucleosomes between LBS regions would have an impact on the distance between bound LANA proteins, affecting the relative locations of their associated antibodies. Similarly, changing the length of the sequence between LBS regions could have a similar effect, either decreasing the space available for nucleosome wrapping or creating space for an additional one to move in. Determining the "tipping point" for a change in either direction would be interesting from both a tether standpoint and a chromatin biology standpoint.

5.2.4 Perturbing the components of a 2TR LANA tether

To date, our experiments have used wild-type LANA and wild-type LBSs, but in order to understand how the different components of the system affect its structure, we propose modifying its various elements. One of the key discoveries using our 2TR system was the full LANA occupancy of all three LBSs. This induces a DNA bend of nearly 180 degrees and affects the trajectory of the bound LANA proteins, impacting the locations of their associated antibodies. We propose experiments on a 2TR system where either LBS2 or LBS3 (or both) has

been mutated to abolish LANA binding. In fact, a 2TR plasmid containing both mutations in both TRs has already been generated by Matt Loftus in the Kedes Laboratory. The singular loss of either LBS2 or LBS3 binding should decrease the DNA bend to approximately 110 degrees, while loss of both should decrease it even further, to approximately 57 degrees (237). Either condition should directly impact the relative trajectories of bound LANA proteins and their associated antibody emissions, causing a deviation from our current model. Determining the parameters for a model that best fits these new data might serve to confirm or contradict the bend angles determined by Wong and Wilson (237). To further test our determined bend angles, we could repeat the methodologies of Wong and Wilson, using electrophoretic mobility shift assays (EMSA). These bending assays implement probes where the binding positions are systematically altered with respect to the ends of the fragment, changing their migration rate within the gel. The various migration rates are then used to calculate the bend created by protein binding to the DNA. Agreement of bend angles between the two methodologies (modeling following dSTORM and EMSA) would serve to validate our model(s) of 2TR tethers.

The other main component in our 2TR system is the LANA protein itself. We have already proposed experiments involving a truncated version of LANA, lacking its internal repeat region. One consideration in using this modified protein is how the lack of an internal repeat region might affect the formation of tethers. One study demonstrated a deficiency in episome persistence with the

removal of internal LANA sequence (238), suggesting that the tethers formed by such a mutated protein might not be equivalent to those incorporating wild-type LANA. As such, we must acknowledge that any tethers we analyze using such a modified protein might exhibit different behaviors or have a different structure from wild-type tethers. One additional avenue of LANA modification might be to alter, but not eliminate, the section of internal repeats forming the coiled-coil. The amino acid composition of this region could potentially be modified (for example, by changing the pattern of hydrophobic and hydrophilic residues) such that it can no longer form a coiled-coil structure when it dimerizes. This might be done via point mutations or even by substituting the coiled-coil-forming region of the LANA gene with different sequence entirely. It would be interesting to then look at the dSTORM data and see if this changes the relative positions of LN53 antibody emissions. For example, if a coiled-coil is no longer formed, the internal regions of both LANA proteins might become longer and more flexible, placing the LN53 epitopes farther away from the LBS DNA and possibly in unpredictable locations. One potential side effect of preventing coiled-coil formation is that it might completely abolish LANA dimer formation altogether. Alternatively, LANA's dimerization domain could be purposefully mutated to observe the effects. If LANA can no longer dimerize, it will probably not bind to LBSs the same way. It is possible that only one LANA might bind to each LBS, effectively halving the number of antibody-associated emissions we observe. Binding of only one LANA per LBS instead of a dimer would also likely alter the level of TR DNA

bending, affecting the relative locations of antibody emissions for two adjacent TRs and, in turn, requiring changes to our current model in order to accommodate the data.

5.2.5 Further exploring the role of LANA during cellular mitosis

Nearly 20 years ago, researchers first described the association of LANA with condensed chromosomes during mitosis (88, 92). Our preliminary work looking at LANA tethers in mitotic BCBL-1 cells supports their findings, but it also demands further investigation. We show that tethers are located on the outside of condensed chromosomes and are not embedded within them, suggesting a course of events in which condensation occurs prior to attachment. One possible criticism of our study is that we may have simply missed embedded tethers due to antibody inaccessibility. To address this problem, we propose a set of experiments using a Dendra2-tagged LANA, which does not depend upon such accessibility. Applying this method to the spread chromosomes of BCBL-1 cells would provide a snapshot of tether localization at the time of fixation. In addition, we propose to use two-color dSTORM and localize the remainder of the KSHV episome using either FISH or immunostaining for H3K27me3, a prevalent marker on the episome (70, 71).

A system using a Dendra2-tagged LANA also has the potential for use in live-cell imaging, allowing us to track LANA tethers through the various stages of mitosis and see how they respond to the stimuli of the mitotic machinery. We showed that tethers in SLKp cells both condense during mitosis and orient

themselves with respect to the metaphase plate. Tracking tethers through mitosis might allow us to see when condensation occurs and whether this coincides with chromosome condensation. We also propose a set of experiments determining which components of mitosis are crucial for both tether condensation and orientation. Knockdown of proteins such as Condensin II (which pulls DNA strands together), Aurora B kinase (which allows Condensin II to associate with chromatin), or Aurora A (which controls microtubule movements) by siRNA or CRISPR-Cas might shed light on whether these proteins influence the behavior of tethers. Chromatin immunoprecipitation assays performed on LANA during mitosis might also provide some interesting information, perhaps showing whether any of the mitotic machinery proteins directly interact with LANA itself.

One must exercise caution when examining a mitotic system involving LANA binding to TRs. Stedman et al. found that the amount of LANA bound to TRs was reduced three-fold in cells arrested at G2/M phase by colchicine compared to that seen in an asynchronous population (143). As a result, all mitotic studies should be approached carefully, with the understanding that LANA binding may not perfectly reflect that seen during other stages of the cell cycle. Reduced LANA binding may equate to incomplete occupancy of all three LBSs, therefore changing the bend angle of the underlying DNA and the positions of the associated emissions. Alternatively, reduced binding may indicate that some TRs are fully occupied by LANA while others are not bound by LANA at all, and

this distribution may be consistent across tethers or vary greatly among them. Thus, our plot showing an increase in tether size with the number of TRs might not look the same in a mitotic system. In addition, we might no longer be able to superimpose multiple 2TR tethers and see separation, as different tethers could have different LANA occupancies, providing each with its own DNA bend and set of localized emissions.

5.2.6 Determining the role of LANA during lytic reactivation of KSHV

Our studies thus far have examined LANA during KSHV latency, but questions remain as to its localization and role during lytic reactivation. Chen et al. recently published work showing the formation of large, circular "transcriptional factories" within reactivated cells (83). Their images show rings of LANA surrounding rings of RNA polymerase, all in a complex with viral RNA. We performed some preliminary work looking at cells containing reactivated BAC16 48 hours after induction. While standard epifluorescence showed the same pattern of rings seen by Chen et al., the resolution of dSTORM showed that these ring-like factories are actually spiral-shaped curls. This raised the question of how LANA makes the transition from the speckled tether pattern of latency to these much larger circular structures during reactivation. To address this, we propose a timecourse following LANA in the hours after induction. This plan requires modification of the BAC16, removing the constitutively active GFP gene and inserting a Dendra2 tag on LANA for live tracking. Another question this finding raised is the orientation of LANA in these spiral structures. We can

take advantage of the resolution of dSTORM and its ability to collect data in two color channels to distinguish between a C-terminal V5 tag on LANA and a more N-terminal antibody or tag. Options for this second tag include the LN53 antibody, which binds N-terminal to the coiled-coil region; the 4C11 antibody, which binds at LANA's N terminus; and a built-in N-terminal Dendra2 tag. We anticipate that comparing the locations of two distal tags or antibodies could give us information on the orientation of LANA molecules within the spiral – for example, whether the C-terminus is facing into the spiral or away from it. Determining orientation could lead to future experiments on the possible functions of LANA in the spiral complex. One additional question of interest is whether LANA is still bound to TRs during lytic reactivation. Using a tagged LANA or bound antibody in conjunction with a FISH probe to the regions surrounding the TRs could provide information on whether the two are still in close proximity after lytic induction. If association between LANA and TRs is lost, a timecourse experiment might provide information as to when this happens in the course of lytic reactivation.

5.2.7 Examining the EBNA1 tether of EBV

Our novel application of super-resolution microscopy to the study of the KSHV LANA tether would also be well-suited to similar tethers, especially that of EBV's EBNA1. In KSHV, all of the LBS-containing TRs are tandemly arranged. In EBV, there are two separate binding regions, one with four EBNA1 binding sites and one with 20, separated from each other by approximately 1 kb. While

each of the native KSHV tethers appears as a single ellipsoid of antibody emissions, the separation in EBV might result in two individual, uneven clusters of emissions separated by a short distance. In this case, it might be difficult to determine which two clusters belong to the same episome. This aside, examination of an EBNA1 tether via super-resolution microscopy would be quite interesting. In our study, the nearly 180 degree bend caused by three adjacent LBSs influenced the trajectory of the bound LANA and, hence, the location of the associated emissions. EBNA1 binding to its dyad symmetry element also causes DNA bending (239). Binding to one high-affinity site induced a bend of 55 to 76 degrees, while binding to two adjacent sites caused a bend of 88 to 112 degrees (240). Super-resolution analysis of an EBNA1 tether might allow for construction of an EBNA1 binding model, with the extent of DNA bending greatly influencing the locations of emissions, similar to our 2TR system with LANA.

While these EBNA1 tether experiments seem analogous to those performed on the LANA tether, the system of EBV infection may pose some difficulty. Preliminary work looking at the EBNA1 staining pattern in EBV-infected cells via epifluorescence microscopy did not show distinct the nuclear punctae like those seen when staining for LANA. Rather, the EBNA1 staining formed a more diffuse pattern of very tiny speckles, much smaller than LANA punctae and greater in number. This poses two potential issues for dSTORM analysis. The first is the difficulty in distinguishing one speckle from another. Since they are so numerous, they are located in much closer proximity to each other than LANA

punctae. This may be an effect of the sheer number of EBV episomes in infected cell lines (one such line was found to have approximately 800 copies of EBV per cell (232)) and could possibly be dealt with by creating an artificial dual-transfection system like the one we employed in studying LANA tethers. Providing the KSHV-negative cell line BJAB with both a TR-containing plasmid and a LANA-encoding plasmid resulted in the formation of fewer tethers than wild-type viral infection. More importantly, the number of tethers was titratable by altering the quantity of TR-containing plasmids used in the transfection. Applying a similar technique to EBNA1 tethers might dial back the number of observed speckles such that they are separated enough to be resolved one from another. The second potential issue in analyzing EBNA1 speckles is resolving them from any background staining or non-tether-associated EBNA1 staining. Because the speckles we observed by epifluorescence were so tiny and faint, they might simply be indistinguishable from their surroundings, making their analysis impossible, possibly even in an artificial dual-transfection system.

5.2.8 CRISPR-Cas9 genome editing in a KSHV-positive cell line

While our preliminary study using CRISPR-Cas9 to knock out LANA in BJAB-LANA cells was successful, the ultimate goal would be to eliminate LANA from a KSHV-positive cell line. Godfrey et al. demonstrated that knockdown of LANA using shRNA led to growth inhibition or apoptosis in some KSHV-positive PEL lines (42), suggesting therapeutic potential for a system of successful LANA elimination. Initial attempts in our laboratory of removing LANA from BCBL-1

cells using a lentiviral CRISPR-Cas9 system had inconclusive results. The majority of cells died following puromycin selection, despite the high transduction efficiency suggested by a fluorescent reporter lentivirus applied in parallel. While it is tempting to argue that the high level of cell death was due to successful elimination of LANA, that seems improbable, as cells treated with a control guide RNA not directed toward LANA also showed significant death. In order to definitively say death was not related to loss of LANA, it might be a good idea to start by disrupting a non-essential KSHV gene or to use a cell line that definitely does not depend upon its KSHV infection for survival (e.g. SLKp). The timing of cell death (in the days immediately following puromycin selection) suggested a possible hypersensitivity of BCBL-1 cells to puromycin. While a kill curve was used to obtain the “optimal” dose, surviving cells still never seemed to fully recover from the shock of selection. Hygromycin, on the other hand, seems to show a better pattern of survival following selection, as seen in unrelated experiments. One possible tactic might be to switch out the puromycin resistance gene in the lentiviral construct for a hygromycin resistance gene and see if that improves cell survival rates following the selection procedure. This might allow the experiment to progress long enough to assess LANA protein levels and their effects on the cells.

It would be interesting to see the fate of the KSHV episomes themselves upon successful disruption of LANA in KSHV-positive cells. Ye et al. found that LANA deletion led to loss of episomes, ultimately resulting in virus-free cells (90),

while similar experiments by Godfrey et al. found that the KSHV copy number only reached a low, but non-zero, plateau (42). To determine which of these patterns occurs in our own system, we propose isolating KSHV episomes via Hirt extraction, followed by Gardella gel analysis and Southern blotting for quantification. Complete elimination of episomes, as seen by Ye et al., suggests a potential role of LANA-targeted CRISPR-Cas in treating KSHV-positive individuals. However, if our results reflect that of Godfrey et al., where episomes remained after treatment, further research would need to be done in order to determine how episomes were successfully maintaining in a presumably LANA-free environment. One possible cause for this phenomenon would be the lingering presence of residual LANA, perhaps due to an in-frame, conservative indel within the LANA sequence, rendering it immune to recognition by the guide RNA. This could be overcome with the simultaneous application of multiple guide RNAs to different regions of the LANA gene, increasing the chances of successful gene disruption.

5.2.9 Extension of our CRISPR-Cas work to the EBNA1 protein

In our work cutting multiple sites in LANA with CRISPR-Cas, we had the most successful knockout using the εζ guide RNA directed at the internal repeat region. We propose performing a similar experiment targeting the internal repeats of EBV's analogous EBNA1 protein. Because targeting multiple sites in the repeat region of LANA was the most successful avenue of gene disruption in our study, it would be interesting to see the effects of a similar study on EBNA1.

Kerry Johnson, the student who performed the experiments involving the εζ guide RNA, used the crispr.mit.edu site (Zhang Lab, MIT) to generate a possible guide RNA for the glycine-alanine internal repeats of EBNA1 (Figure 5-1). This guide targets ten separate occurrences of the same sequence, highlighted in yellow, giving it great potential for the multi-site disruption of EBNA1. Downstream experiments could examine the fate of EBV episomes following successful elimination of EBNA1. Having enough surviving cells to analyze might pose a challenge for such a study, as CRISPR-Cas9 experiments targeting EBNA1 by Wang and Quake showed increased cell death via apoptosis (214).

5.2.10 Applications of CRISPRi and CRISPRa to the study of KSHV

While classic CRISPR-Cas genome editing relies upon the cutting of double-stranded DNA and creation of indels to create stop codons, thereby inactivating a gene, two more recently developed forms of CRISPR only involve the guidance of transcriptional modulators to the gene of interest. CRISPR interference (CRISPRi) and CRISPR activation (CRISPRa) utilize a modified Cas9 protein lacking in endonuclease activity to simply recruit transcriptional repressors or activators, respectively, to transcriptional start sites. Unlike classic CRISPR-Cas, which makes relatively permanent sequence changes, these systems provide a reversible way of modulating gene expression in a gene-specific fashion. Combining this utility with the added flexibility of an inducible

Figure 5-1. Targeting the internal repeats of EBNA1. The DNA sequence of EBNA1, with a suggested guide RNA targeting the protein's internal repeat region highlighted in yellow. Note that this potential guide RNA targets 10 separate places in the EBNA1 sequence. Figure created by Kerry Johnson and used with permission.

ATGTCTGACGAGGGGCCAGGTACAGGACCTGGAAATGGCCTAGGAGAGAAGGGAGACACATCTGGACCAGA
AGGCTCCGGCGGCAGTGGACCTCAAAGAAGAGGGGGTGATAACCATGGACGAGGACGGGGAAGAGGACGAG
GACGAGGAGGCGGAAGACCAGGAGCCCCGGGCGGCTCAGGATCAGGGCCAAGACATAGAGATGGTGTCCGG
AGACCCCAAAAACGTCCAAGTTGCATTGGCTGCAAAGGGACCCACGGTGGAACAGGAGCAGGAGCAGGAGC
GGGAGGGGCAGGAGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGC
AGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGGAGG
GGCAGGAGCAGGAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGG
AGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGGAGCAGG
AGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGGAGC
AGGAGGGGCAGGAGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGGGCAGG
AGGGGCAGGAGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAGCAGG
AGGGGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGG
AGCAGGAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGG
GGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGCAGGAGGGGCAGGAGCAGG
AGGTGGAGGCCGGGGTCGAGGAGGCAGTGGAGGCCGGGGTCGAGGAGGTAGTGGAGGCCGGGGTCGAGGAG
GTAGTGGAGGCCGCCGGGGTAGAGGACGTGAAAGAGCCAGGGGGGGAAGTCGTGAAAGAGCCAGGGGGAG
AGGTCGTGGACGTGGAGAAAAGAGGCCCAGGAGTCCCAGTAGTCAGTCATCATCATCCGGGTCTCCACCGCG
CAGGCCCCCTCCAGGTAGAAGGCCATTTTTCCACCCTGTAGGGGAAGCCGATTATTTTGAATACCACCAAGAA
GGTGGCCCAGATGGTGAGCCTGACGTGCCCCCGGGAGCGATAGAGCAGGGCCCCGCAGATGACCCAGGAGA
AGGCCCAAGCACTGGACCCCGGGGTCAGGGTGATGGAGGCAGGCGCAAAAAAGGAGGGTGGTTTGGAAAGC
ATCGTGGTCAAGGAGGTTCCAACCCGAAATTTGAGAACATTGCAGAAGGTTTAAGAGCTCTCCTGGCTAGGA
GTCACGTAGAAAGGACTACCGACGAAGGAACTTGGGTCGCCGGTGTGTTCGTATATGGAGGTAGTAAGACCT
CCCTTTACAACCTAAGGCGAGGAACTGCCCTTGCTATTCCACAATGTCGTCTTACACCATTGAGTCGTCTCCCC
TTTGGAATGGCCCCTGGACCCGGCCCACAACCTGGCCCGCTAAGGGAGTCCATTGTCTGTTATTTCATGGTCTT
TTTACAAACTCATATATTTGCTGAGGTTTTGAAGGATGCGATTAAGGACCTTGTTATGACAAAGCCCGCTCCTA
CCTGCAATATCAGGGTGACTGTGTGCAGCTTTGACGATGGAGTAGATTTGCCTCCCTGGTTTCCACCTATGGTG
GAAGGGGCTGCCGCGGAGGGTGATGACGGAGATGACGGAGATGAAGGAGGTGATGGAGATGAGGGTGAGGA
AGGGCAGGAGTGA

EBNA1 Target Sites with Example Guide RNA

promoter and the genome integration capability of a lentivector could result in a powerful tool for analyzing the functions of any number of genes, host or viral, by simply adding different guide RNAs.

Appendix 1 details our laboratory's attempts at creating a stably integrated, cumate-inducible CRISPRi system in BCBL-1 cells. In principle, guide RNAs for any conceivable gene could later be added to these cells via transfection or transduction, followed by cumate induction to turn on expression of the repressive Cas9 fusion protein. In the days following induction, various tests might be run to evaluate the effects of the repressed gene. Should we desire to turn off the repression, we could simply remove cumate from the media and allow the cells to resume normal gene expression. A similar approach could be taken with CRISPRa, allowing for a cell line in which we could analyze the effects of gene overexpression. Provided they each had their own selection marker and induction compound, both constructs could potentially even be integrated into the same cells, resulting in a stable line for studying the phenotype of any gene's knockout or overexpression, or perhaps the effects of experiencing both conditions back-to-back.

Our efforts to create a stable CRISPRi cell line were generally unsuccessful (see Appendix 1), but if this system were to become a reality in the future, it would provide a strong tool for studying latent viruses, including KSHV. Previous KSHV gene knockout studies involved directly altering the gene in one of the KSHV BACs, which can be difficult to manipulate and often have unwanted

characteristics, such as constitutively active fluorescent proteins. Systems utilizing BACs can also be viewed as “artificial” when compared with existing KSHV-positive patient-derived cell lines. An integrated CRISPRi/CRISPRa system in, for example, BCBL-1 cells would provide an arena for gene manipulation in the environment of a stable latent infection with wild-type KSHV.

APPENDIX 1: DEVELOPMENT OF AN INDUCIBLE CRISPR INTERFERENCE SYSTEM IN THE KSHV-POSITIVE CELL LINE BCBL-1

The traditional CRISPR-Cas9 method of gene knockdown relies on the function of an endonuclease and the resulting indels to alter a gene and produce premature stop codons. However, this method (referred to as CRISPRn, for CRISPR nuclease) can have unwanted results. Because the resulting indels are random, they can potentially introduce gain-of-function mutations in addition to the desired stop codons. There is also the possibility of a silent mutation, in which the protein produced is wild-type, but the genomic sequence is altered such that it can no longer be targeted by the guide RNA used. In fact, Mandegar et al. found that 30 to 50% of mutated alleles in their CRISPRn system contained in-frame indels (241). To remove the unpredictability of indels from the equation, researchers developed a related method, utilizing the targeting capability of the Cas9 protein. The method, known as CRISPR interference (CRISPRi), uses a Cas9 protein with two point mutations in its RuvC-like and NHN nuclease domains (242). This protein, dCas9, lacks endonuclease activity but can still be targeted to a particular region of the genome through the use of guide RNAs. Research has taken advantage of this system to target transcriptional activators or repressors to the transcriptional start sites of genes as a method of controlling gene expression. A study by Mandegar et al. utilized a fusion between dCas9 and a Kruppel-associated box (KRAB) domain to suppress gene expression (241). The KRAB domain recruits histone deacetylases and chromatin remodeling complexes to the targeted region (243). It also enhances localized deposition of H3K9me3 and heterochromatin protein 1, silencing gene

expression in the area. Mandegar et al. demonstrated an increased loss of gene expression using this dCas9-KRAB fusion CRISPRi system (greater than 99% knockdown) over the traditional CRISPRn system (a 60 to 70% knockdown). This compelled us to create a stable BCBL-1 cell line in which KRAB-dCas9 was present and only required addition of a selected guide RNA to target a gene of interest. Since the KRAB-mediated repression is reversible, unlike CRISPRn cutting, this cell line would present a convenient alternative method of gene knockdown for studies on gene function in KSHV.

While Mandegar et al. used a gene trap approach of integrating their CRISPR system into the host genome, we chose to employ the PiggyBac transposon-mediated system of gene transfer used by Yang et al. in their targeting of porcine endogenous retroviruses (231). The PiggyBac system, named for its discovery in mutant baculovirus strains, uses terminal inverted repeats to “cut and paste” an inserted gene into TTAA nucleotide elements (244). Importantly for our purposes, this system can easily accommodate inserts up to 14 kbp in size without a significant reduction to its transposon activity (our desired KRAB-dCas9 fusion is well within this limit, at approximately 4.4 kbp).

In addition to requiring that our KRAB-dCas9 system be integrated into the host genome, we also desired an inducible system in which the fusion protein expression could be turned on and off. For this reason, we chose to employ the PBQM812A-1PB-Cuo-MCS-IRES-GFP-EF1α-CymR-Puro Inducible cDNA Cloning and Expression Vector from System Biosciences, Inc. (referred to herein

as PBV) (Figure A1-1). This vector places an inserted gene of interest under a non-toxic cumate-inducible switch and integrates the construct into the host genome. It also contains a constitutively active puromycin selection marker and GFP reporter gene activated by the cumate switch. We cut this vector with NheI and BstBI to allow for directional insertion. Our insert was the transcriptional repressor Kruppel-associated box (KRAB) attached to the N terminus of a nuclease-inactivated dCas9. This insert was obtained by extracting the KRAB-dCas9 fragment from the vector pHR-SFFV-KRAB-dCas9-P2A-mCherry (a gift from Jonathan Weissman; Addgene plasmid #60954) (245) via PCR using the high-fidelity Q5 polymerase (New England Biosciences) and the PCR program in Table A1-1. NheI and BstBI sites were added to the respective ends of the PCR primers, shown in Table A1-1. The PCR product was cut sequentially with NheI and BstBI, ligated into the PBV to form KRAB-dCas9-PBV, and used to transform Stbl2 E. coli (Thermo Fisher Scientific).

KRAB-dCas9-PBV was transfected into BCBL-1 cells along with the PiggyBac Transposase Vector (System Biosciences, Inc.) required for integration using ExtremeGENE HP (Sigma Aldrich), according to the manufacturer's guidelines. Two days after transfection, cells were put under 4 μg/mL puromycin selection. Selection lasted for seven days, at which point the media was replaced with fresh, puromycin-free media and surviving cells were allowed to recover and grow out. After several days of outgrowth, a subset of these cells

Figure A1-1. Map of the cumate-inducible PiggyBac vector used in this study. The cumate-inducible PiggyBac vector was obtained from System Biosciences. Important features include the cumate switch site and constitutively active CymR repressor gene, GFP reporter gene, puromycin resistance gene, and core insulator sequences (image sourced from https://www.systembio.com/products/gene-expression-systems/piggybac-transposon/pb-cuo-mcs-ires-gfp-ef1-alpha-cymr-puro/ (246)).

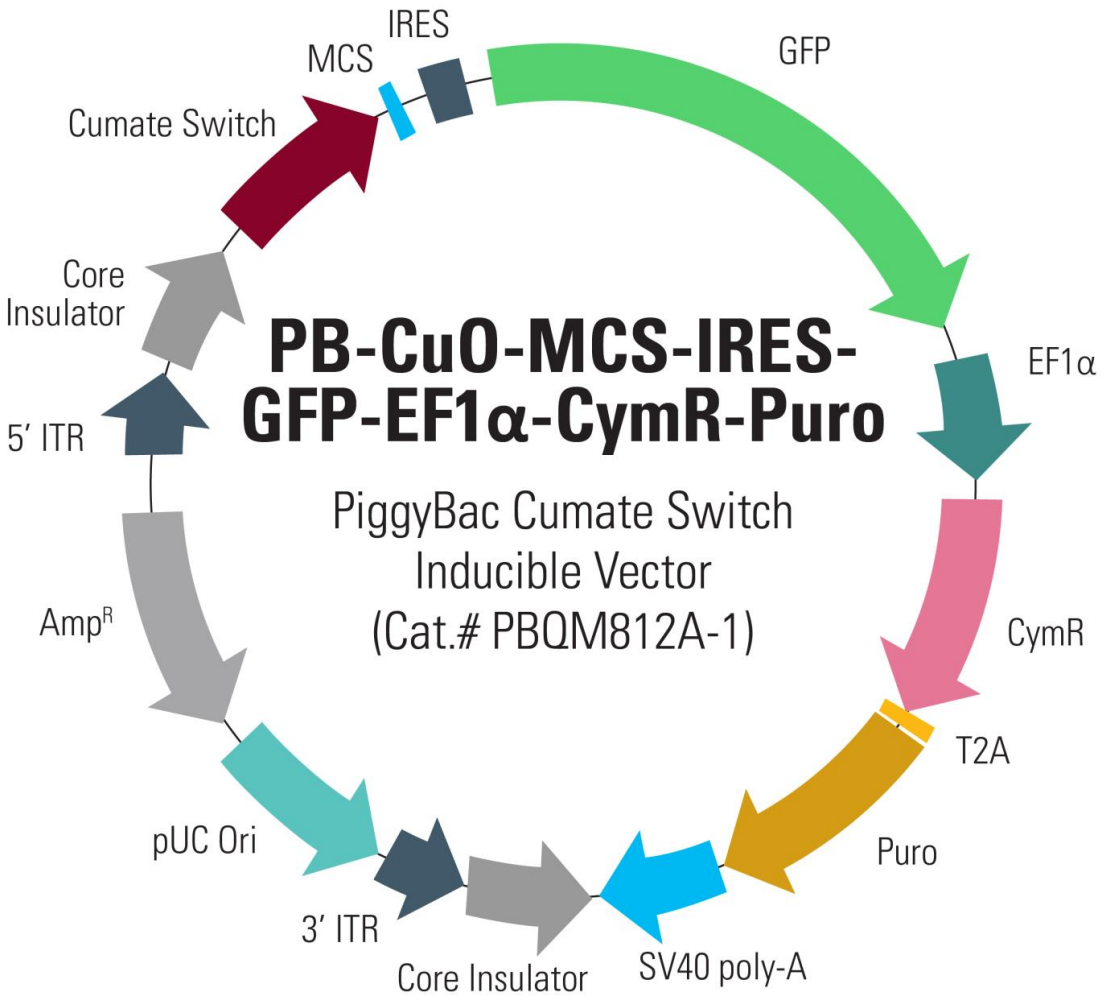

IRES
MCS
GFP
Cumate Switch
Core
Insulator
EF1α
PB-CuO-MCS-IRES-
GFP-EF1α-CymR-Puro
5′ ITR
PiggyBac Cumate Switch
Inducible Vector
(Cat.# PBQM812A-1)
CymR
T2A
Amp^R
Puro
pUC Ori
3′ ITR
SV40 poly-A
Core Insulator

Table A1-1. PCR program and primers used to extract KRAB-dCas9 fragment. PCR steps marked in bold italics were repeated for 35 cycles.

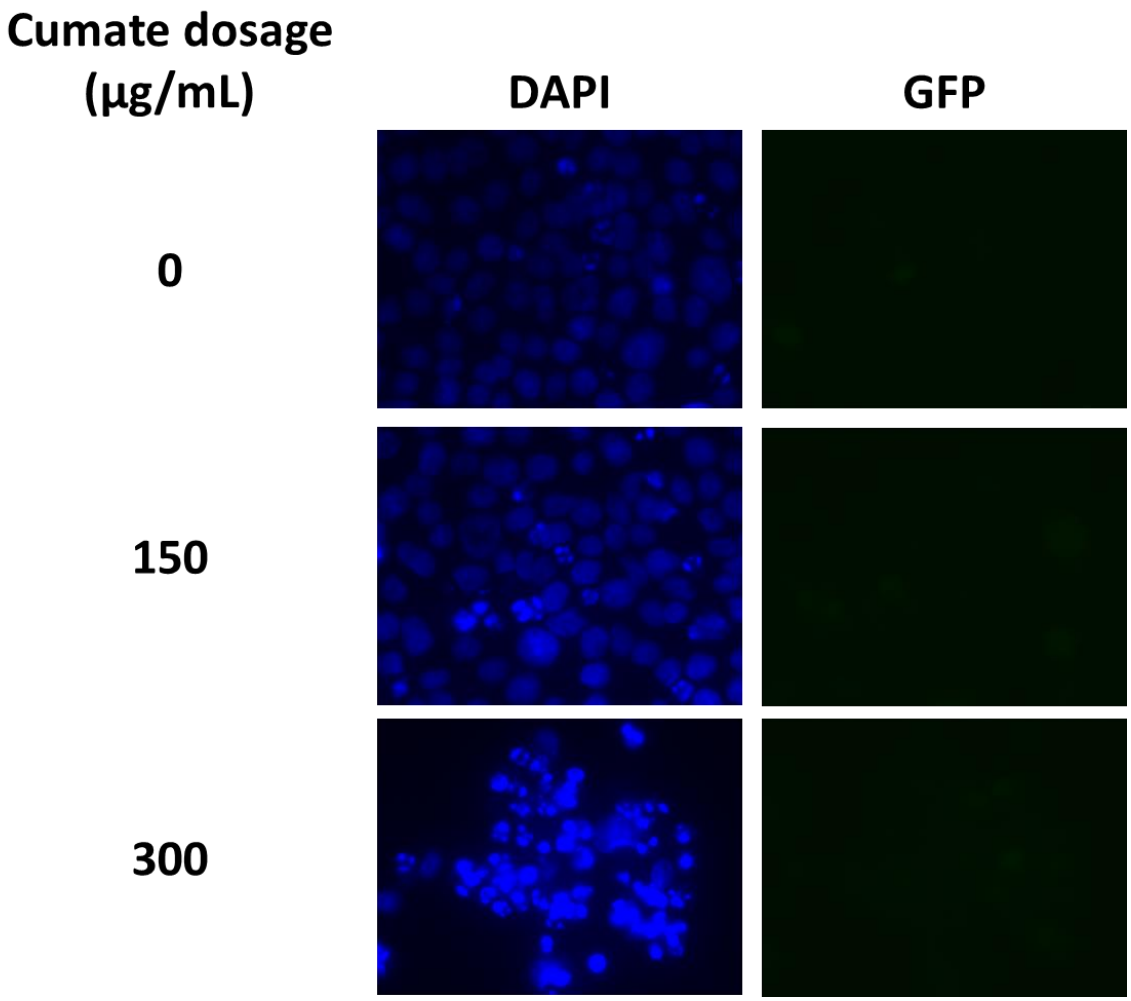
Temperature (in ^{o}C)	**Time (MM:SS)**
98	00:30
98	***00:10***
69	***00:30***
72	***02:30***
72	02:00
Primer Name	**Sequence**
PB-NheI-Fwd	GCTCGCTGTTGCTAGCCTTCTGCTTCCCGAGCTCTA
PB-BstBI-Rev	GCTCGCTGTTTTCGAAAAGCGCATGAACTCCTTGAT

was placed under a cumate titration, ranging from 30 µg/mL to 300 µg/mL per manufacturer's instructions, for 72 hours to induce expression of both the KRAB-dCas9 and GFP. Cumate-treated cells were fixed with 4% paraformaldehyde, stained with DAPI, deposited on slides (Cytospin, Shandon), and evaluated for GFP expression by fluorescence microscopy.

Figure A1-2 shows representative fields from the transfected, cumate-induced BCBL-1 cells. Cells induced with two different doses of cumate (150 and 300 µg/mL) showed no marked difference in GFP expression from untreated cells, indicating the cumate induction of the reporter gene was unsuccessful. Also of note is the increase in cell death with increasing doses of cumate. At 300 µg/mL, the maximum dose recommended by the manufacturer, nearly all cells were found dead by DAPI staining. A similar cumate titration was performed on BCBL-1 cells that had not been transfected with the KRAB-dCas9-PBV, and the results looked identical, including the pattern of increasing cell death (data not shown). This indicated that the cumate was likely toxic to BCBL-1 cells.

To ensure that the PBV insert had successfully integrated into the host genome, we performed whole cell PCR on our transfected cells over a region of the Cas9 gene. DNA gel electrophoresis showed the expected product band in transfected cells, while control cells (not transfected with the KRAB-dCas9-PBV) did not show this band, indicating specificity of the PCR product (data not

Figure A1-2. Effects of cumate titration on BCBL-1 cells transfected with KRAB-dCas9-PBV. Epifluorescence microscopy of DAPI staining and GFP expression following a cumate titration on transfected BCBL-1 cells. Two separate doses of cumate (150 µg/mL and 300 µg/mL) failed to induce GFP expression, resulting in fields that are indistinguishable from that of the control condition (0 µg/mL cumate). Raising the dosage of cumate was associated with an increase in cell death, as seen by the increasing incidence of nuclear blebbing visible in the DAPI staining.


Cumate dosage
(µg/mL)
DAPI
GFP
0
150
300

shown). This indicated to us that the Cas9-containing insert had successfully integrated into the genome via the PiggyBac system but was failing to reactivate with cumate.

In preparation for the successful creation of a cumate-inducible KRAB-dCas9 cell line, we prepared a system for later introducing guide RNAs into the cells. We chose to use the lentiviral vector pLenti SpBsmBI sgRNA Hygro (a gift from Rene Maehr; Addgene plasmid # 62205) (247) due to its BsmBI insertion site, allowing for directional cloning with only a single restriction enzyme. Oligos were designed using crispr.mit.edu (Zhang lab, MIT), focusing on the region +/-150 base pairs around the transcriptional start site (the optimal region for CRISPRi targeting (241)) of the LANA gene. Five different guide RNAs were designed in this region, labeled iAB, iCD, iEF, iGH, and iIJ. Oligos were ordered, annealed, and ligated into the vector as described in Sections 4.3.2-3. The oligo product forming iEF repeatedly failed to ligate into the vector and so did not undergo any later steps. The resulting successful ligation products were used to transform Stbl3 E. coli (Thermo Fisher Scientific). Lentivirus was produced and quantified using the procedure outlined in Section 4.3.7.

Despite the lack of a functioning cumate-inducible system, BCBL-1 cells successfully transfected with the KRAB-dCas9-PBV construct were infected with these guide RNA lentiviruses (for detailed methods, see Section 4.3.8) and later placed under hygromycin selection at 1.2 mg/mL for 7 days to obtain transduced cells. No experiments were performed on these cells, as the cumate switch

failed to turn on the GFP reporter gene and, likely, the KRAB-dCas9 insert as well, rendering the presence of a guide RNA irrelevant.

Should future experiments lead to successful induction of the GFP and KRAB-dCas9 in BCBL-1 cells, this system would provide a useful tool for the cumate-controlled CRISPRi knockdown of any gene of interest. Only a guide RNA would need to be provided for targeting to the appropriate gene. This could mark a novel alternative to other knockdown methods in the study of viral gene functions.

REFERENCES

1. Feeney KM, Parish JL (2009) Targeting mitotic chromosomes: a conserved mechanism to ensure viral genome persistence. *Proc R Soc B Biol Sci* 276(1662):1535–1544.

2. Chen Q, Sun L, Chen ZJ (2016) Regulation and function of the cGAS-STING pathway of cytosolic DNA sensing. *Nat Immunol* 17(10):1142–1149.

3. de Villiers E-M, Fauquet C, Broker TR, Bernard H-U, zur Hausen H (2004) Classification of papillomaviruses. *Virology* 324(1):17–27.

4. Doorbar J (2006) Molecular biology of human papillomavirus infection and cervical cancer. *Clin Sci Lond Engl 1979* 110(5):525–541.

5. Chaturvedi AK, et al. (2011) Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol Off J Am Soc Clin Oncol* 29(32):4294–4301.

6. Dell G, et al. (2003) Comparison of the structure and DNA-binding properties of the E2 proteins from an oncogenic and a non-oncogenic human papillomavirus. *J Mol Biol* 334(5):979–991.

7. McBride AA (2013) The Papillomavirus E2 proteins. *Virology* 445(1):57–79.

8. Donaldson MM, Boner W, Morgan IM (2007) TopBP1 regulates human papillomavirus type 16 E2 interaction with chromatin. *J Virol* 81(8):4338–4342.

9. Parish JL, Bean AM, Park RB, Androphy EJ (2006) ChlR1 is required for loading papillomavirus E2 onto mitotic chromosomes and viral genome maintenance. *Mol Cell* 24(6):867–876.

10. Feeney KM, Saade A, Okrasa K, Parish JL (2011) In vivo analysis of the cell cycle dependent association of the bovine papillomavirus E2 protein and ChlR1. *Virology* 414(1):1–9.

11. Fukumoto H (2011) Pathology of Kaposi's sarcoma-associated herpesvirus infection. *Front Microbiol* 2. doi:10.3389/fmicb.2011.00175.

12. Bloom DC, Dhummakupt A (2016) The Herpes Simplex Viruses. *Neurotropic Viral Infections: Volume 2: Neurotropic Retroviruses, DNA Viruses, Immunity and Transmission*, ed Reiss CS (Springer International Publishing, Cham), pp 111–134.

13. Adler B, Sattler C, Adler H (2017) Herpesviruses and Their Host Cells: A Successful Liaison. *Trends Microbiol* 25(3):229–241.

14. Grinde B (2013) Herpesviruses: latency and reactivation – viral strategies and host response. *J Oral Microbiol* 5. doi:10.3402/jom.v5i0.22766.

15. Jenkins FJ, Hoffman LJ (2000) Overview of Herpesviruses. *Infectious Causes of Cancer: Targets for Intervention*, ed Goedert JJ (Humana Press, Totowa, NJ), pp 33–49.

16. Crawford DH (2001) Biology and disease associations of Epstein-Barr virus. *Philos Trans R Soc Lond B Biol Sci* 356(1408):461–473.

17. Thorley-Lawson DA (2005) EBV the prototypical human tumor virus--just how bad is it? *J Allergy Clin Immunol* 116(2):251–261; quiz 262.

18. Reisman D, Yates J, Sugden B (1985) A putative origin of replication of plasmids derived from Epstein-Barr virus is composed of two cis-acting components. *Mol Cell Biol* 5(8):1822–1832.

19. Deakyne JS, Malecka KA, Messick TE, Lieberman PM (2017) Structural and Functional Basis for an EBNA1 Hexameric Ring in Epstein-Barr Virus Episome Maintenance. *J Virol* 91(19). doi:10.1128/JVI.01046-17.

20. Sears J, et al. (2004) The amino terminus of Epstein-Barr Virus (EBV) nuclear antigen 1 contains AT hooks that facilitate the replication and partitioning of latent EBV genomes by tethering them to cellular chromosomes. *J Virol* 78(21):11487–11505.

21. Kanda T, Otter M, Wahl GM (2001) Coupling of Mitotic Chromosome Tethering and Replication Competence in Epstein-Barr Virus-Based Plasmids. *Mol Cell Biol* 21(10):3576–3588.

22. Sternbach G, Varon J (1995) Moritz Kaposi: idiopathic pigmented sarcoma of the skin. *J Emerg Med* 13(5):671–674.

23. Beral V, Peterman TA, Berkelman RL, Jaffe HW (1990) Kaposi's sarcoma among persons with AIDS: a sexually transmitted infection? *Lancet Lond Engl* 335(8682):123–128.

24. Chang Y, et al. (1994) Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* 266(5192):1865–1869.

25. International Committee on Taxonomy of Viruses (ICTV) *Int Comm Taxon Viruses ICTV*. Available at: https://talk.ictvonline.org/taxonomy/p/taxonomy-history?taxnode_id=20160940 [Accessed February 20, 2018].

26. What Is Kaposi Sarcoma? Available at: https://www.cancer.org/cancer/kaposi-sarcoma/about/what-is-kaposi-sarcoma.html [Accessed September 5, 2017].

27. Kazanji M, et al. (2005) Serological and molecular evidence that human herpesvirus 8 is endemic among Amerindians in French Guiana. *J Infect Dis* 192(9):1525–1529.

28. Biggar RJ, Whitby D, Marshall V, Linhares AC, Black F (2000) Human Herpesvirus 8 in Brazilian Amerindians: A Hyperendemic Population with a New Subtype. *J Infect Dis* 181(5):1562–1568.

29. Gao SJ, et al. (1996) KSHV antibodies among Americans, Italians and Ugandans with and without Kaposi's sarcoma. *Nat Med* 2(8):925–928.

30. Kedes DH, et al. (1996) The seroepidemiology of human herpesvirus 8 (Kaposi's sarcoma-associated herpesvirus): distribution of infection in KS risk groups and evidence for sexual transmission. *Nat Med* 2(8):918–924.

31. Casper C (2006) HIV serodiscordant sex partners and the prevalence of human herpesvirus 8 infection among HIV negative men who have sex with men: baseline data from the EXPLORE Study. *Sex Transm Infect* 82(3):229–235.

32. Chatlynne LG, et al. (1998) Detection and Titration of Human Herpesvirus-8–Specific Antibodies in Sera From Blood Donors, Acquired Immunodeficiency Syndrome Patients, and Kaposi's Sarcoma Patients Using a Whole Virus Enzyme-Linked Immunosorbent Assay. *Blood* 92(1):53–58.

33. Labo N, Miley W, Kesler M, Jacobson L, Whitby D (2010) Current epidemiology of KSHV infection in American men who have sex with men. *Infect Agent Cancer* 5(Suppl 1):A13–A13.

34. Zong J-C, et al. (1999) High-Level Variability in the ORF-K1 Membrane Protein Gene at the Left End of the Kaposi's Sarcoma-Associated Herpesvirus Genome Defines Four Major Virus Subtypes and Multiple

Variants or Clades in Different Human Populations. *J Virol* 73(5):4156–4170.

35. Diamond C, et al. (1997) Absence of Detectable Human Herpesvirus 8 in the Semen of Human Immunodeficiency Virus-Infected Men without Kaposi's Sarcoma. *J Infect Dis* 176(3):775–777.

36. Cesarman E, Chang Y, Moore PS, Said JW, Knowles DM (1995) Kaposi's Sarcoma–Associated Herpesvirus-Like DNA Sequences in AIDS-Related Body-Cavity–Based Lymphomas. *N Engl J Med* 332(18):1186–1191.

37. Soulier J, et al. (1995) Kaposi's sarcoma-associated herpesvirus-like DNA sequences in multicentric Castleman's disease [see comments]. *Blood* 86(4):1276–1280.

38. Hassman LM, Ellison TJ, Kedes DH (2011) KSHV infects a subset of human tonsillar B cells, driving proliferation and plasmablast differentiation. *J Clin Invest* 121(2):752–768.

39. Chen Y-B, Rahemtullah A, Hochberg E (2007) Primary Effusion Lymphoma. *The Oncologist* 12(5):569–576.

40. Uppal T, Banerjee S, Sun Z, Verma S, Robertson E (2014) KSHV LANA—The Master Regulator of KSHV Latency. *Viruses* 6(12):4961–4998.

41. Dittmer DP, Damania B (2013) Kaposi sarcoma associated herpesvirus pathogenesis (KSHV)—an update. *Viral Pathog Vaccines* 3(3):238–244.

42. Godfrey A, Anderson J, Papanastasiou A, Takeuchi Y, Boshoff C (2005) Inhibiting primary effusion lymphoma by lentiviral vectors encoding short hairpin RNA. *Blood* 105(6):2510–2518.

43. Kim B, Jeon YK, Kim CW (2009) Kaposi Sarcoma Herpes Virus-associated Hemophagocytic Syndrome Complicated by Multicentric Castleman Disease and Kaposi Sarcoma in a HIV-negative Immunocompetent Patient: An Autopsy Case. *J Korean Med Sci* 24(5):970.

44. Sprinz E (2004) Successful treatment of AIDS-related Castleman's disease following the administration of highly active antiretroviral therapy (HAART). *Ann Oncol* 15(2):356-a-358.

45. Martín-Carbonero L, et al. (2008) Long-Term Prognosis of HIV-Infected Patients with Kaposi Sarcoma Treated with Pegylated Liposomal Doxorubicin. *Clin Infect Dis* 47(3):410–417.

46. Sin S-H, Fakhari FD, Dittmer DP (2010) The Viral Latency-Associated Nuclear Antigen Augments the B-Cell Response to Antigen *In Vivo*. *J Virol* 84(20):10653–10660.

47. Nichols LA, Adang LA, Kedes DH (2011) Rapamycin Blocks Production of KSHV/HHV8: Insights into the Anti-Tumor Activity of an Immunosuppressant Drug. *PLoS ONE* 6(1):e14535.

48. Medveczky MM, Horvath E, Lund T, Medveczky PG (1997) In vitro antiviral drug sensitivity of the Kaposi's sarcoma-associated herpesvirus. *Aids* 11(11):1327–1332.

49. Zhu W, et al. (2005) Potent Antiviral Activity of North-Methanocarbathymidine against Kaposi's Sarcoma-Associated Herpesvirus. *Antimicrob Agents Chemother* 49(12):4965–4973.

50. Chakraborty S, Veettil MV, Chandran B (2012) Kaposi's Sarcoma Associated Herpesvirus Entry into Target Cells. *Front Microbiol* 3. doi:10.3389/fmicb.2012.00006.

51. Sharma-Walia N, et al. (2005) ERK1/2 and MEK1/2 Induced by Kaposi's Sarcoma-Associated Herpesvirus (Human Herpesvirus 8) Early during Infection of Target Cells Are Essential for Expression of Viral Genes and for Establishment of Infection. *J Virol* 79(16):10308–10329.

52. Raghu H, Sharma-Walia N, Veettil MV, Sadagopan S, Chandran B (2009) Kaposi's Sarcoma-Associated Herpesvirus Utilizes an Actin Polymerization-Dependent Macropinocytic Pathway To Enter Human

Dermal Microvascular Endothelial and Human Umbilical Vein Endothelial Cells. *J Virol* 83(10):4895–4911.

53. Krishnan HH, et al. (2004) Concurrent Expression of Latent and a Limited Number of Lytic Genes with Immune Modulation and Antiapoptotic Function by Kaposi's Sarcoma-Associated Herpesvirus Early during Infection of Primary Endothelial and Fibroblast Cells and Subsequent Decline of Lytic Gene Expression. *J Virol* 78(7):3601–3620.

54. Muylaert I, Elias P (2007) Knockdown of DNA Ligase IV/XRCC4 by RNA Interference Inhibits Herpes Simplex Virus Type I DNA Replication. *J Biol Chem* 282(15):10865–10872.

55. Paulus C, Nitzsche A, Nevels M (2010) Chromatinisation of herpesvirus genomes. *Rev Med Virol* 20(1):34–50.

56. Nitzsche A, Paulus C, Nevels M (2008) Temporal Dynamics of Cytomegalovirus Chromatin Assembly in Productively Infected Human Cells. *J Virol* 82(22):11167–11180.

57. Placek BJ, et al. (2009) The histone variant H3.3 regulates gene expression during lytic infection with herpes simplex virus type 1. *J Virol* 83(3):1416–1421.

58. Günther T, Grundhoff A (2010) The Epigenetic Landscape of Latent Kaposi Sarcoma-Associated Herpesvirus Genomes. *PLoS Pathog* 6(6):e1000935.

59. Groves IJ, Reeves MB, Sinclair JH (2009) Lytic infection of permissive cells with human cytomegalovirus is regulated by an intrinsic “pre-immediate-early” repression of viral gene expression mediated by histone post-translational modification. *J Gen Virol* 90(Pt 10):2364–2374.

60. Renne R, Lagunoff M, Zhong W, Ganem D (1996) The size and conformation of Kaposi’s sarcoma-associated herpesvirus (human herpesvirus 8) DNA in infected cells and virions. *J Virol* 70(11):8151–8154.

61. Lagunoff M, Ganem D (1997) The structure and coding organization of the genomic termini of Kaposi’s sarcoma-associated herpesvirus. *Virology* 236(1):147–154.

62. Russo JJ, et al. (1996) Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8). *Proc Natl Acad Sci U S A* 93(25):14862–14867.

63. Dittmer D, et al. (1998) A cluster of latently expressed genes in Kaposi’s sarcoma-associated herpesvirus. *J Virol* 72(10):8309–8315.

64. Mettenleiter TC (2004) Budding events in herpesvirus morphogenesis. *Virus Res* 106(2):167–180.

65. Mesri EA, Cesarman E, Boshoff C (2010) Kaposi’s sarcoma and its associated herpesvirus. *Nat Rev Cancer* 10(10):707–719.

66. Ojala PM, et al. (1999) Kaposi's sarcoma-associated herpesvirus-encoded v-cyclin triggers apoptosis in cells with high levels of cyclin-dependent kinase 6. *Cancer Res* 59(19):4984–4989.

67. An J, Sun Y, Sun R, Rettig MB (2003) Kaposi's sarcoma-associated herpesvirus encoded vFLIP induces cellular IL-6 expression: the role of the NF-κB and JNK/AP1 pathways. *Oncogene* 22(22):3371–3385.

68. McCormick C (2005) The Kaposin B Protein of KSHV Activates the p38/MK2 Pathway and Stabilizes Cytokine mRNAs. *Science* 307(5710):739–741.

69. Talbot SJ, Weiss RA, Kellam P, Boshoff C (1999) Transcriptional analysis of human herpesvirus-8 open reading frames 71, 72, 73, K14, and 74 in a primary effusion lymphoma cell line. *Virology* 257(1):84–94.

70. Toth Z, et al. (2010) Epigenetic Analysis of KSHV Latent and Lytic Genomes. *PLoS Pathog* 6(7):e1001013.

71. Sun R, et al. (2017) Epigenetic Landscape of Kaposi's Sarcoma-Associated Herpesvirus Genome in Classic Kaposi's Sarcoma Tissues. *PLOS Pathog* 13(1):e1006167.

72. Günther T, Grundhoff A (2010) The Epigenetic Landscape of Latent Kaposi Sarcoma-Associated Herpesvirus Genomes. *PLoS Pathog* 6(6):e1000935.

73. Pantry SN, Medveczky PG (2009) Epigenetic regulation of Kaposi's sarcoma-associated herpesvirus replication. *Semin Cancer Biol* 19(3):153–157.

74. Stedman W, Deng Z, Lu F, Lieberman PM (2004) ORC, MCM, and Histone Hyperacetylation at the Kaposi's Sarcoma-Associated Herpesvirus Latent Replication Origin. *J Virol* 78(22):12566–12575.

75. Hilton IB, et al. (2013) The Open Chromatin Landscape of Kaposi's Sarcoma-Associated Herpesvirus. *J Virol* 87(21):11831–11842.

76. Sakakibara S, et al. (2004) Accumulation of heterochromatin components on the terminal repeat sequence of Kaposi's sarcoma-associated herpesvirus mediated by the latency-associated nuclear antigen. *J Virol* 78(14):7299–7310.

77. Matsumura S, Fujita Y, Gomez E, Tanese N, Wilson AC (2005) Activation of the Kaposi's Sarcoma-Associated Herpesvirus Major Latency Locus by the Lytic Switch Protein RTA (ORF50). *J Virol* 79(13):8493–8505.

78. Gwack Y, et al. (2003) Principal role of TRAP/mediator and SWI/SNF complexes in Kaposi's sarcoma-associated herpesvirus RTA-mediated lytic reactivation. *Mol Cell Biol* 23(6):2055–2067.

79. Cohen A (2006) An essential role of ERK signalling in TPA-induced reactivation of Kaposi's sarcoma-associated herpesvirus. *J Gen Virol* 87(4):795–802.

80. Chen J, et al. (2001) Activation of latent Kaposi's sarcoma-associated herpesvirus by demethylation of the promoter of the lytic transactivator. *Proc Natl Acad Sci U S A* 98(7):4119–4124.

81. Wang SE, et al. (2003) Role of CCAAT/Enhancer-Binding Protein Alpha (C/EBP) in Activation of the Kaposi's Sarcoma-Associated Herpesvirus (KSHV) Lytic-Cycle Replication-Associated Protein (RAP) Promoter in Cooperation with the KSHV Replication and Transcription Activator (RTA) and RAP. *J Virol* 77(1):600–623.

82. Ye J, Shedd D, Miller G (2005) An Sp1 response element in the Kaposi's sarcoma-associated herpesvirus open reading frame 50 promoter mediates lytic cycle induction by butyrate. *J Virol* 79(3):1397–1408.

83. Chen CP, et al. (2017) Kaposi's Sarcoma-Associated Herpesvirus Hijacks RNA Polymerase II To Create a Viral Transcriptional Factory. *J Virol* 91(11). doi:10.1128/JVI.02491-16.

84. Judde J-G (2000) Monoclonality or Oligoclonality of Human Herpesvirus 8 Terminal Repeat Sequences in Kaposi's Sarcoma and Other Diseases. *J Natl Cancer Inst* 92(9):729–736.

85. Boulanger E, et al. (2005) Mono/oligoclonal pattern of Kaposi Sarcoma-associated herpesvirus (KSHV/HHV-8) episomes in primary effusion lymphoma cells. *Int J Cancer* 115(4):511–518.

86. Shrestha P, Sugden B (2014) Identification of Properties of the Kaposi's Sarcoma-Associated Herpesvirus Latent Origin of Replication That Are Essential for the Efficient Establishment and Maintenance of Intact Plasmids. *J Virol* 88(15):8490–8503.

87. Hu J, Garber AC, Renne R (2002) The Latency-Associated Nuclear Antigen of Kaposi's Sarcoma-Associated Herpesvirus Supports Latent DNA Replication in Dividing Cells. *J Virol* 76(22):11677–11687.

88. Ballestas ME, Chatis PA, Kaye KM (1999) Efficient Persistence of Extrachromosomal KSHV DNA Mediated by Latency-Associated Nuclear Antigen. *Science* 284(5414):641–644.

89. Cotter MA, Robertson ES (1999) The Latency-Associated Nuclear Antigen Tethers the Kaposi's Sarcoma-Associated Herpesvirus Genome to Host Chromosomes in Body Cavity-Based Lymphoma Cells. *Virology* 264(2):254–264.

90. Ye F-C, et al. (2004) Disruption of Kaposi's Sarcoma-Associated Herpesvirus Latent Nuclear Antigen Leads to Abortive Episome Persistence. *J Virol* 78(20):11121–11129.

91. Szekely L, et al. (1998) Restricted expression of Epstein-Barr virus (EBV)-encoded, growth transformation-associated antigens in an EBV- and human herpesvirus type 8-carrying body cavity lymphoma line. *J Gen Virol* 79(6):1445–1452.

92. Szekely L, et al. (1999) Human herpesvirus-8-encoded LNA-1 accumulates in heterochromatin- associated nuclear bodies. *J Gen Virol* 80(11):2889–2900.

93. Ponnusamy R, et al. (2015) KSHV but not MHV-68 LANA induces a strong bend upon binding to terminal repeat viral DNA. *Nucleic Acids Res*:gkv987.

94. Ballestas ME, Kaye KM (2001) Kaposi's Sarcoma-Associated Herpesvirus Latency-Associated Nuclear Antigen 1 Mediates Episome Persistence through cis-Acting Terminal Repeat (TR) Sequence and Specifically Binds TR DNA. *J Virol* 75(7):3250–3258.

95. Kelley-Clarke B, Ballestas ME, Komatsu T, Kaye KM (2007) Kaposi's sarcoma herpesvirus C-terminal LANA concentrates at pericentromeric and peri-telomeric regions of a subset of mitotic chromosomes. *Virology* 357(2):149–157.

96. Lu F, Day L, Gao S-J, Lieberman PM (2006) Acetylation of the latency-associated nuclear antigen regulates repression of Kaposi's sarcoma-associated herpesvirus lytic transcription. *J Virol* 80(11):5273–5282.

97. Campbell M, et al. (2012) Protein Arginine Methyltransferase 1-directed Methylation of Kaposi Sarcoma-associated Herpesvirus Latency-associated Nuclear Antigen. *J Biol Chem* 287(8):5806–5818.

98. Li Q, Zhou F, Ye F, Gao S-J (2008) Genetic Disruption of KSHV Major Latent Nuclear Antigen LANA Enhances Viral Lytic 2 Transcriptional Program. *Virology* 379(2):234.

99. Mercier A, Arias C, Madrid AS, Holdorf MM, Ganem D (2014) Site-Specific Association with Host and Viral Chromatin by Kaposi's Sarcoma-Associated Herpesvirus LANA and Its Reversal during Lytic Reactivation. *J Virol* 88(12):6762–6777.

100. Friborg J, Kong W, Hottiger MO, Nabel GJ (1999) p53 inhibition by the LANA protein of KSHV protects against cell death. *Nature* 402(6764):889–894.

101. Fujimuro M, et al. (2003) A novel viral mechanism for dysregulation of beta-catenin in Kaposi's sarcoma-associated herpesvirus latency. *Nat Med* 9(3):300–306.

102. Alkharsah KR, Schulz TF (2012) A role for the internal repeat of the Kaposi's sarcoma-associated herpesvirus latent nuclear antigen in the persistence of an episomal viral genome. *J Virol* 86(3):1883–1887.

103. De Leon Vazquez E, Kaye KM (2011) The Internal Kaposi's Sarcoma-Associated Herpesvirus LANA Regions Exert a Critical Role on Episome Persistence. *J Virol* 85(15):7622–7633.

104. De Leon Vazquez E, Carey VJ, Kaye KM (2013) Identification of Kaposi's Sarcoma-Associated Herpesvirus LANA Regions Important for Episome Segregation, Replication, and Persistence. *J Virol* 87(22):12270–12283.

105. Komatsu T, Ballestas ME, Barbera AJ, Kaye KM (2002) The KSHV latency-associated nuclear antigen: a multifunctional protein. *Front Biosci* 7:d726–730.

106. Schwam DR, Luciano RL, Mahajan SS, Wong L, Wilson AC (2000) Carboxy Terminus of Human Herpesvirus 8 Latency-Associated Nuclear Antigen Mediates Dimerization, Transcriptional Repression, and Targeting to Nuclear Bodies. *J Virol* 74(18):8532–8540.

107. Piolot T, Tramier M, Coppey M, Nicolas J-C, Marechal V (2001) Close but Distinct Regions of Human Herpesvirus 8 Latency-Associated Nuclear Antigen 1 Are Responsible for Nuclear Targeting and Binding to Human Mitotic Chromosomes. *J Virol* 75(8):3948–3959.

108. Kelley-Clarke B, De Leon-Vazquez E, Slain K, Barbera AJ, Kaye KM (2009) Role of Kaposi's Sarcoma-Associated Herpesvirus C-Terminal

LANA Chromosome Binding in Episome Persistence. *J Virol* 83(9):4326–4337.

109. Barbera AJ, Chodaparambil JV, Kelley-Clarke B, Luger K, Kaye KM (2006a) Kaposi's sarcoma-associated herpesvirus LANA hitches a ride on the chromosome. *Cell Cycle* 5(10):1048–1052.

110. Grundhoff A, Ganem D (2003) The Latency-Associated Nuclear Antigen of Kaposi's Sarcoma-Associated Herpesvirus Permits Replication of Terminal Repeat-Containing Plasmids. *J Virol* 77(4):2779–2783.

111. Toptan T, Fonseca L, Kwun HJ, Chang Y, Moore PS (2013) Complex Alternative Cytoplasmic Protein Isoforms of the Kaposi's Sarcoma-Associated Herpesvirus Latency-Associated Nuclear Antigen 1 Generated through Noncanonical Translation Initiation. *J Virol* 87(5):2744–2755.

112. Juillard F, Tan M, Li S, Kaye KM (2016) Kaposi's Sarcoma Herpesvirus Genome Persistence. *Front Microbiol* 7. doi:10.3389/fmicb.2016.01149.

113. Zhang G, et al. (2016) Cytoplasmic isoforms of Kaposi sarcoma herpesvirus LANA recruit and antagonize the innate immune DNA sensor cGAS. *Proc Natl Acad Sci U S A* 113(8):E1034-1043.

114. Canham M (2004) A naturally occurring C-terminal truncated isoform of the latent nuclear antigen of Kaposi's sarcoma-associated herpesvirus does not associate with viral episomal DNA. *J Gen Virol* 85(6):1363–1369.

115. Woodard C, et al. (2012) Phosphorylation of the Chromatin Binding Domain of KSHV LANA. *PLoS Pathog* 8(10):e1002972.

116. Krithivas A, Fujimuro M, Weidner M, Young DB, Hayward SD (2002) Protein Interactions Targeting the Latency-Associated Nuclear Antigen of Kaposi's Sarcoma-Associated Herpesvirus to Cell Chromosomes. *J Virol* 76(22):11596–11604.

117. Barbera AJ, et al. (2006) The Nucleosomal Surface as a Docking Station for Kaposi's Sarcoma Herpesvirus LANA. *Science* 311(5762):856–861.

118. Beauchemin C, Moerke NJ, Faloon P, Kaye KM (2014) Assay Development and High-Throughput Screening for Inhibitors of Kaposi's Sarcoma–Associated Herpesvirus N-Terminal Latency-Associated Nuclear Antigen Binding to Nucleosomes. *J Biomol Screen* 19(6):947–958.

119. Chodaparambil JV, et al. (2007) A charged and contoured surface on the nucleosome regulates chromatin compaction. *Nat Struct Mol Biol* 14(11):1105–1107.

120. You J, et al. (2006) Kaposi's Sarcoma-Associated Herpesvirus Latency-Associated Nuclear Antigen Interacts with Bromodomain Protein Brd4 on Host Mitotic Chromosomes. *J Virol* 80(18):8909–8919.

121. Chiu Y-F, Sugden AU, Fox K, Hayes M, Sugden B (2017) Kaposi's sarcoma-associated herpesvirus stably clusters its genomes across

generations to maintain itself extrachromosomally. *J Cell Biol* 216(9):2745–2758.

122. Correia B, et al. (2013) Crystal Structure of the Gamma-2 Herpesvirus LANA DNA Binding Domain Identifies Charged Surface Residues Which Impact Viral Latency. *PLoS Pathog* 9(10):e1003673.

123. Domsic JF, Chen H-S, Lu F, Marmorstein R, Lieberman PM (2013) Molecular Basis for Oligomeric-DNA Binding and Episome Maintenance by KSHV LANA. *PLoS Pathog* 9(10):e1003672.

124. Hellert J, et al. (2013) A structural basis for BRD2/4-mediated host chromatin interaction and oligomer assembly of Kaposi sarcoma-associated herpesvirus and murine gammaherpesvirus LANA proteins. *PLoS Pathog* 9(10):e1003640.

125. Hellert J, et al. (2015) The 3D structure of Kaposi sarcoma herpesvirus LANA C-terminal domain bound to DNA. *Proc Natl Acad Sci* 112(21):6694–6699.

126. Garber AC, Shu MA, Hu J, Renne R (2001) DNA Binding and Modulation of Gene Expression by the Latency-Associated Nuclear Antigen of Kaposi's Sarcoma-Associated Herpesvirus. *J Virol* 75(17):7882–7892.

127. Garber AC (2002) Latency-associated Nuclear Antigen (LANA) Cooperatively Binds to Two Sites within the Terminal Repeat, and Both

Sites Contribute to the Ability of LANA to Suppress Transcription and to Facilitate DNA Replication. *J Biol Chem* 277(30):27401–27411.

128. Wong L-Y, Wilson AC (2005) Kaposi's Sarcoma-Associated Herpesvirus Latency-Associated Nuclear Antigen Induces a Strong Bend on Binding to Terminal Repeat DNA. *J Virol* 79(21):13829–13836.

129. Ballestas ME, Kaye KM (2011) The latency-associated nuclear antigen, a multifunctional protein central to Kaposi's sarcoma-associated herpesvirus latency. *Future Microbiol* 6(12):1399–1413.

130. Zaldumbide A, Ossevoort M, Wiertz EJHJ, Hoeben RC (2007) In cis inhibition of antigen processing by the latency-associated nuclear antigen I of Kaposi sarcoma herpes virus. *Mol Immunol* 44(6):1352–1360.

131. Kwun HJ, et al. (2007) Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 mimics Epstein-Barr virus EBNA1 immune evasion through central repeat domain effects on protein processing. *J Virol* 81(15):8225–8235.

132. Levitskaya J, Sharipo A, Leonchiks A, Ciechanover A, Masucci MG (1997) Inhibition of ubiquitin/proteasome-dependent protein degradation by the Gly-Ala repeat domain of the Epstein–Barr virus nuclear antigen 1. *Proc Natl Acad Sci* 94(23):12616–12621.

133. Gao S-J, et al. (1999) Molecular polymorphism of Kaposi's sarcoma—associated herpesvirus (human herpesvirus 8) latent nuclear antigen: evidence for a large repertoire of viral genotypes and dual infection with different viral genotypes. *J Infect Dis* 180(5):1466–1476.

134. Gupta A, Oldenburg DG, Salinas E, White DW, Forrest JC (2017) MHV68 Expressing KSHV LANA Reveals Both Functional Conservation and Divergence in LANA Homologs. *J Virol*. doi:10.1128/JVI.00992-17.

135. van de Linde S, et al. (2011) Direct stochastic optical reconstruction microscopy with standard fluorescent probes. *Nat Protoc* 6(7):991–1009.

136. van de Linde S, et al. (2011) Photoinduced formation of reversible dye radicals and their impact on super-resolution imaging. *Photochem Photobiol Sci Off J Eur Photochem Assoc Eur Soc Photobiol* 10(4):499–506.

137. Dempsey GT, et al. (2009) Photoswitching Mechanism of Cyanine Dyes. *J Am Chem Soc* 131(51):18192.

138. Huang B, Wang W, Bates M, Zhuang X (2008) Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy. *Science* 319(5864):810–813.

139. Moore PS, et al. (1996) Primary characterization of a herpesvirus agent associated with Kaposi's sarcomae. *J Virol* 70(1):549–558.

140. Simpson GR, et al. (1996) Prevalence of Kaposi's sarcoma associated herpesvirus infection measured by antibodies to recombinant capsid protein and latent immunofluorescence antigen. *The Lancet* 348(9035):1133–1138.

141. Adang LA, Parsons CH, Kedes DH (2006) Asynchronous Progression through the Lytic Cascade and Variations in Intracellular Viral Loads Revealed by High-Throughput Single-Cell Analysis of Kaposi's Sarcoma-Associated Herpesvirus Infection. *J Virol* 80(20):10073–10082.

142. Kelley-Clarke B, Ballestas ME, Komatsu T, Kaye KM (2007) Kaposi's sarcoma herpesvirus C-terminal LANA concentrates at pericentromeric and peri-telomeric regions of a subset of mitotic chromosomes. *Virology* 357(2):149–157.

143. Stedman W, Deng Z, Lu F, Lieberman PM (2004) ORC, MCM, and Histone Hyperacetylation at the Kaposi's Sarcoma-Associated Herpesvirus Latent Replication Origin. *J Virol* 78(22):12566–12575.

144. Watanabe A, et al. (2007) A novel KRAB-Zinc finger protein interacts with latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus and activates transcription via terminal repeat sequences. *Virus Genes* 34(2):127–136.

145. Renne R, et al. (1996) Lytic growth of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) in culture. *Nat Med* 2(3):342–346.

146. Menezes J, Leibold W, Klein G, Clements G (1975) Establishment and characterization of an Epstein-Barr virus (EBC)-negative lymphoblastoid B cell line (BJA-B) from an exceptional, EBV-genome-negative African Burkitt's lymphoma. *Biomed Publiee Pour AAICIG* 22(4):276–284.

147. Stürzl M, Gaus D, Dirks WG, Ganem D, Jochmann R (2013) Kaposi's sarcoma-derived cell line SLK is not of endothelial origin, but is a contaminant from a known renal carcinoma cell line. *Int J Cancer* 132(8):1954–1958.

148. Grundhoff A, Ganem D (2004) Inefficient establishment of KSHV latency suggests an additional role for continued lytic replication in Kaposi sarcoma pathogenesis. *J Clin Invest* 113(1):124–136.

149. Brulois KF, et al. (2012) Construction and Manipulation of a New Kaposi's Sarcoma-Associated Herpesvirus Bacterial Artificial Chromosome Clone. *J Virol* 86(18):9708–9720.

150. Renne R, et al. (2001) Modulation of Cellular and Viral Gene Expression by the Latency-Associated Nuclear Antigen of Kaposi's Sarcoma-Associated Herpesvirus. *J Virol* 75(1):458–468.

151. Plaisance-Bonstaff K (2012) Analysis of Kaposi's Sarcoma-Associated Herpesvirus miRNAs' Role in the Maintenance of Viral Latency. Dissertation (University of Florida).

152. Kellam P, et al. (1999) Characterization of monoclonal antibodies raised against the latent nuclear antigen of human herpesvirus 8. *J Virol* 73(6):5149–5155.

153. Gould TJ, Verkhusha VV, Hess ST (2009) Imaging biological structures with fluorescence photoactivation localization microscopy. *Nat Protoc* 4(3):291–308.

154. Hess ST, Girirajan TPK, Mason MD (2006) Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy. *Biophys J* 91(11):4258–4272.

155. Thompson RE, Larson DR, Webb WW (2002) Precise nanometer localization analysis for individual fluorescent probes. *Biophys J* 82(5):2775–2783.

156. Hell S, Reiner G, Cremer C, Stelzer EHK (1993) Aberrations in confocal fluorescence microscopy induced by mismatches in refractive index. *J Microsc* 169(3):391–405.

157. Arimoto R, Murray JM (2004) A common aberration with water-immersion objective lenses. *J Microsc* 216(Pt 1):49–51.

158. Jones E, Oliphant T, Peterson P (2001) *SciPy: Open source scientific tools for Python* Available at: http://www.scipy.org/.

159. van der Walt S, Colbert SC, Varoquaux G (2011) The NumPy Array: A Structure for Efficient Numerical Computation. *Comput Sci Eng* 13(2):22–30.

160. DeSantis MC, Zareh SK, Li X, Blankenship RE, Wang YM (2012) Single-image axial localization precision analysis for individual fluorophores. *Opt Express* 20(3):3057–3065.

161. Guizar-Sicairos M, Thurman ST, Fienup JR (2008) Efficient subpixel image registration algorithms. *Opt Lett* 33(2):156–158.

162. Pedregosa F, et al. (2011) Scikit-learn: Machine Learning in Python. *J Mach Learn Res* 12(Oct):2825–2830.

163. Perktold J, Seabold S, Taylor J, statsmodels contributors (2017) *Statsmodels* Available at: http://statsmodels.sourceforge.net/.

164. Schneider PJ, Eberly DH (2003) *Geometric Tools for Computer Graphics* (Morgan Kaufmann, San Francisco, CA, USA).

165. Hunter JD (2007) Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* 9(3):90–95.

166. Ramachandran P, Varoquaux G (2011) Mayavi: 3D Visualization of Scientific Data. *Comput Sci Eng* 13(2):40–51.

167. Cock PJA, et al. (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25(11):1422–1423.

168. Hamelryck T, Manderick B (2003) PDB file parser and structure class implemented in Python. *Bioinformatics* 19(17):2308–2310.

169. Tien MZ, Sydykova DK, Meyer AG, Wilke CO (2013) PeptideBuilder: A simple Python library to generate model peptides. *PeerJ* 1:e80.

170. Schrödinger, LLC (2015) *The PyMOL Molecular Graphics System, Version 1.7*.

171. Wood CW, et al. (2014) CCBuilder: an interactive web-based tool for building, designing and assessing coiled-coil protein assemblies. *Bioinformatics* 30(21):3029–3035.

172. Byrd R, Lu P, Nocedal J, Zhu C (1995) A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM J Sci Comput* 16(5):1190–1208.

173. Boettiger AN, et al. (2016) Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* 529(7586):418–422.

174. McDonnell AV, Jiang T, Keating AE, Berger B (2006) Paircoil2: improved prediction of coiled coils from sequence. *Bioinformatics* 22(3):356–358.

175. Schalch T, Duda S, Sargent DF, Richmond TJ (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* 436(7047):138–141.

176. Mitosis | Define Mitosis at Dictionary.com Available at: http://www.dictionary.com/browse/mitosis [Accessed October 19, 2017].

177. Nurse P (1990) Universal control mechanism regulating onset of M-phase. *Nature* 344(6266):503–508.

178. Gavet O, Pines J (2010) Activation of cyclin B1-Cdk1 synchronizes events in the nucleus and the cytoplasm at mitosis. *J Cell Biol* 189(2):247–259.

179. Gavet O, Pines J (2010) Progressive activation of CyclinB1-Cdk1 coordinates entry to mitosis. *Dev Cell* 18(4):533–543.

180. Wieser S, Pines J (2015) The biochemistry of mitosis. *Cold Spring Harb Perspect Biol* 7(3):a015776.

181. Antonin W, Neumann H (2016) Chromosome condensation and decondensation during mitosis. *Curr Opin Cell Biol* 40:15–22.

182. O'Connor CM, Adams JU (2010) *Essentials of Cell Biology* (NPG Education, Cambridge, MA).

183. Hirano T (2016) Condensin-Based Chromosome Organization from Bacteria to Vertebrates. *Cell* 164(5):847–857.

184. Green LC, et al. (2012) Contrasting roles of condensin I and condensin II in mitotic chromosome formation. *J Cell Sci* 125(Pt 6):1591–1604.

185. Shintomi K, Hirano T (2011) The relative ratio of condensin I to II determines chromosome shapes. *Genes Dev* 25(14):1464–1469.

186. Piazza I, Haering CH, Rutkowska A (2013) Condensin: crafting the chromosome landscape. *Chromosoma* 122(3):175–190.

187. Gerlich D, Hirota T, Koch B, Peters J-M, Ellenberg J (2006) Condensin I stabilizes chromosomes mechanically through a dynamic interaction in live cells. *Curr Biol CB* 16(4):333–344.

188. Abe S, et al. (2011) The initial phase of chromosome condensation requires Cdk1-mediated phosphorylation of the CAP-D3 subunit of condensin II. *Genes Dev* 25(8):863–874.

189. Petersen J, Hagan IM (2003) S. pombe aurora kinase/survivin is required for chromosome condensation and the spindle checkpoint attachment response. *Curr Biol CB* 13(7):590–597.

190. Nakazawa N, et al. (2008) Dissection of the essential steps for condensin accumulation at kinetochores and rDNAs during fission yeast mitosis. *J Cell Biol* 180(6):1115–1131.

191. Thadani R, Uhlmann F, Heeger S (2012) Condensin, Chromatin Crossbarring and Chromosome Condensation. *Curr Biol* 22(23):R1012–R1021.

192. Robinson PJJ, et al. (2008) 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. *J Mol Biol* 381(4):816–825.

193. Shogren-Knaak M, et al. (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311(5762):844–847.

194. Peters J-M, Nishiyama T (2012) Sister chromatid cohesion. *Cold Spring Harb Perspect Biol* 4(11). doi:10.1101/cshperspect.a011130.

195. Reber S, Hyman AA (2015) Emergent Properties of the Metaphase Spindle. *Cold Spring Harb Perspect Biol* 7(7):a015784.

196. Hoyt MA, Totis L, Roberts BT (1991) S. cerevisiae genes required for cell cycle arrest in response to loss of microtubule function. *Cell* 66(3):507–517.

197. Li R, Murray AW (1991) Feedback control of mitosis in budding yeast. *Cell* 66(3):519–531.

198. Li X, Nicklas RB (1995) Mitotic forces control a cell-cycle checkpoint. *Nature* 373(6515):630–632.

199. Rieder CL, Cole RW, Khodjakov A, Sluder G (1995) The checkpoint delaying anaphase in response to chromosome monoorientation is mediated by an inhibitory signal produced by unattached kinetochores. *J Cell Biol* 130(4):941–948.

200. D'Avino PP, Giansanti MG, Petronczki M (2015) Cytokinesis in animal cells. *Cold Spring Harb Perspect Biol* 7(4):a015834.

201. Hershko A, Ciechanover A (1998) The ubiquitin system. *Annu Rev Biochem* 67:425–479.

202. Geley S, et al. (2001) Anaphase-promoting complex/cyclosome-dependent proteolysis of human cyclin A starts at the beginning of mitosis and is not subject to the spindle assembly checkpoint. *J Cell Biol* 153(1):137–148.

203. Patra D, Wang SX, Kumagai A, Dunphy WG (1999) The xenopus Suc1/Cks protein promotes the phosphorylation of G(2)/M regulators. *J Biol Chem* 274(52):36839–36842.

204. Siomos MF, et al. (2001) Separase is required for chromosome segregation during meiosis I in Caenorhabditis elegans. *Curr Biol CB* 11(23):1825–1835.

205. Mora-Bermúdez F, Gerlich D, Ellenberg J (2007) Maximal chromosome compaction occurs by axial shortening in anaphase and depends on Aurora kinase. *Nat Cell Biol* 9(7):822–831.

206. Ramadan K, et al. (2007) Cdc48/p97 promotes reformation of the nucleus by extracting the kinase Aurora B from chromatin. *Nature* 450(7173):1258–1262.

207. Komatsu T, Ballestas ME, Barbera AJ, Kelley-Clarke B, Kaye KM (2004) KSHV LANA1 binds DNA as an oligomer and residues N-terminal to the oligomerization domain are essential for DNA binding, replication, and episome persistence. *Virology* 319(2):225–236.

208. Lu J, et al. (2009) Latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus (KSHV) upregulates survivin expression in KSHV-Associated B-lymphoma cells and contributes to their proliferation. *J Virol* 83(14):7129–7141.

209. Lens SMA, Vader G, Medema RH (2006) The case for Survivin as mitotic regulator. *Curr Opin Cell Biol* 18(6):616–622.

210. Mita AC, Mita MM, Nawrocki ST, Giles FJ (2008) Survivin: key regulator of mitosis and apoptosis and novel target for cancer therapeutics. *Clin Cancer Res Off J Am Assoc Cancer Res* 14(16):5000–5005.

211. Wheatley SP, McNeish IA (2005) Survivin: a protein with dual roles in mitosis and apoptosis. *Int Rev Cytol* 247:35–88.

212. Lu J, et al. (2014) Kaposi's sarcoma-associated herpesvirus-encoded LANA contributes to viral latent replication by activating phosphorylation of survivin. *J Virol* 88(8):4204–4217.

213. Kumar A, et al. (2014) Kaposi sarcoma herpes virus latency associated nuclear antigen protein release the G2/M cell cycle blocks by modulating ATM/ATR mediated checkpoint pathway. *PloS One* 9(6):e100228.

214. Wang J, Quake SR (2014) RNA-guided endonuclease provides a therapeutic strategy to cure latent herpesviridae infection. *Proc Natl Acad Sci* 111(36):13157–13162.

215. van Diemen FR, et al. (2016) CRISPR/Cas9-Mediated Genome Editing of Herpesviruses Limits Productive and Latent Infections. *PLOS Pathog* 12(6):e1005701.

216. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A (1987) Nucleotide sequence of the iap gene, responsible for alkaline phosphatase

isozyme conversion in Escherichia coli, and identification of the gene product. *J Bacteriol* 169(12):5429–5433.

217. Mojica FJ, Díez-Villaseñor C, Soria E, Juez G (2000) Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol* 36(1):244–246.

218. Makarova KS, Aravind L, Grishin NV, Rogozin IB, Koonin EV (2002) A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res* 30(2):482–496.

219. Guy CP, Majerník AI, Chong JPJ, Bolt EL (2004) A novel nuclease-ATPase (Nar71) from archaea is part of a proposed thermophilic DNA repair system. *Nucleic Acids Res* 32(21):6176–6186.

220. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiol Read Engl* 151(Pt 8):2551–2561.

221. Mojica FJM, Diez-Villasenor C, Garcia-Martinez J, Soria E (2005) Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *J Mol Evol* 60(2):174–182.

222. Pourcel C, Salvignol G, Vergnaud G (2005) CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA,

and provide additional tools for evolutionary studies. *Microbiol Read Engl* 151(Pt 3):653–663.

223. Jansen R, Embden JDA van, Gaastra W, Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 43(6):1565–1575.

224. Haft DH, Selengut J, Mongodin EF, Nelson KE (2005) A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput Biol* 1(6):e60.

225. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV (2006) A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct* 1:7.

226. Barrangou R, et al. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819):1709–1712.

227. Cong L, et al. (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* 339(6121):819–823.

228. Doudna JA, Charpentier E (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346(6213):1258096.

229. Riballo E, et al. (2004) A pathway of double-strand break rejoining dependent upon ATM, Artemis, and proteins locating to gamma-H2AX foci. *Mol Cell* 16(5):715–724.

230. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507(7490):62–67.

231. Yang L, et al. (2015) Genome-wide inactivation of porcine endogenous retroviruses (PERVs). *Science* 350(6264):1101–1104.

232. Kim DN, et al. (2013) Characterization of naturally Epstein-Barr virus-infected gastric carcinoma cell line YCCEL1. *J Gen Virol* 94(Pt 3):497–506.

233. Chen W, Dittmer DP (2011) Ribosomal protein S6 interacts with the latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus. *J Virol* 85(18):9495–9505.

234. Pham H, Kearns NA, Maehr R (2016) Transcriptional Regulation with CRISPR/Cas9 Effectors in Mammalian Cells. *Methods Mol Biol Clifton NJ* 1358:43–57.

235. Kuscu C, Arslan S, Singh R, Thorpe J, Adli M (2014) Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol* 32(7):677–683.

236. Wu X, et al. (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol* 32(7):670–676.

237. Wong L-Y, Wilson AC (2005) Kaposi's Sarcoma-Associated Herpesvirus Latency-Associated Nuclear Antigen Induces a Strong Bend on Binding to Terminal Repeat DNA. *J Virol* 79(21):13829–13836.

238. De Leon Vazquez E, Kaye KM (2011) The Internal Kaposi's Sarcoma-Associated Herpesvirus LANA Regions Exert a Critical Role on Episome Persistence. *J Virol* 85(15):7622–7633.

239. Frappier L, O'Donnell M (1992) EBNA1 distorts oriP, the Epstein-Barr virus latent replication origin. *J Virol* 66(3):1786–1790.

240. Bashaw JM, Yates JL (2001) Replication from oriP of Epstein-Barr Virus Requires Exact Spacing of Two Bound Dimers of EBNA1 Which Bend DNA. *J Virol* 75(22):10603–10611.

241. Mandegar MA, et al. (2016) CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. *Cell Stem Cell* 18(4):541–553.

242. Larson MH, et al. (2013) CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc* 8(11):2180–2196.

243. Lupo A, et al. (2013) KRAB-Zinc Finger Proteins: A Repressor Family Displaying Multiple Biological Functions. *Curr Genomics* 14(4):268.

244. Wilson MH, Coates CJ, George AL (2007) PiggyBac Transposon-mediated Gene Transfer in Human Cells. *Mol Ther* 15(1):139–145.

245. Gilbert LA, et al. (2014) Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* 159(3):647–661.

246. PB-Cuo-MCS-IRES-GFP-EF1α-CymR-Puro Inducible cDNA Cloning and Expression Vector *Syst Biosci*. Available at: https://www.systembio.com/products/gene-expression-systems/piggybac-transposon/pb-cuo-mcs-ires-gfp-ef1-alpha-cymr-puro/ [Accessed March 17, 2018].

247. Pham H, Kearns NA, Maehr R (2016) Transcriptional Regulation with CRISPR/Cas9 Effectors in Mammalian Cells. *Post-Transcriptional Gene Regulation*, ed Dassi E (Springer New York, New York, NY), pp 43–57.