**Distribution Agreement**

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Claudia Chandelle Wahoski April 10, 2022

# HIV-1 Antisense RNA Production and Latency Increases when the LEDGF/p75 and Integrase Interaction is Inhibited

by

Claudia Chandelle Wahoski

Dr. Philip Tedbury

Adviser

Biology

Dr. Philip Tedbury

Adviser

Dr. Kathleen Campbell

Committee Member

Dr. Roger Deal

Committee Member

2022

# HIV-1 Antisense RNA Production and Latency Increases when the LEDGF/p75 and Integrase Interaction is Inhibited

By

Claudia Chandelle Wahoski

Dr. Philip Tedbury

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Biology

2022

Abstract

HIV-1 Antisense RNA Production and Latency Increases when the LEDGF/p75 and Integrase Interaction is Inhibited

By Claudia Chandelle Wahoski

Human Immunodeficiency Virus type 1 (HIV-1) is a parenterally transmitted retrovirus that infects CD4+ T cells and can lead to acquired immunodeficiency syndrome. While treatment exists to manage actively replicating virus, a major barrier to curing HIV-1 is latent virus harbored in long-lasting CD4+ T cells that are transcriptionally quiescent, leading to immune evasion. However, the mechanisms of latency are incompletely understood. The HIV-1 RNA genome is reverse transcribed into DNA that is integrated into the host genome by integrase, acting in concert with host factors, including lens epithelium-derived growth factor (LEDGF)/p75. LEDGF/p75 is important for viral DNA integration into active, gene-dense regions of the host genome. Development and testing in clinical trials of new HIV-1 therapies include allosteric integrase inhibitors (ALLINIs) that target LEDGF/p75's interaction with integrase. ALLINI treatment has been shown to increase latency *in vitro*. HIV-1 antisense RNA (asRNA), an RNA transcript originating from transcription in the 3' long terminal repeat of HIV-1, has been shown to be involved in inducing and maintaining HIV-1 latency. Unpublished data from the Sarafianos lab shows that asRNA production increases in ALLINI-treated cells, specifically using the ALLINI known as BI-D. This research aims to further explore the relationship between asRNA production in ALLINI-treated cells, and the involvement in latency. We used a dual fluorescent HIV-1 reporter (Hi-Fate Tomato) to identify three different subpopulations of cells, latent, active, and uninfected cells differentiated by two fluorescent proteins using flow cytometry. We assessed the asRNA production in each subpopulation when left untreated, or treated with BI-D. We found that the Hi-Fate Tomato reporter produces detectable asRNA that can be used in subsequent asRNA research. After cell sorting, we found that treatment with BI-D does not lead to differences in asRNA in the three subpopulations of infected cells, but definitive conclusions cannot be made with the current data due to a single replicate being performed. This research provides the foundation for further work that can investigate asRNA production in ALLINI-treated cells. The implications of the results can contribute to further understanding the mechanisms of latency that can lead to curing HIV-1.

# HIV-1 Antisense RNA Production and Latency Increases when the LEDGF/p75 and Integrase Interaction is Inhibited

By

Claudia Chandelle Wahoski

Dr. Philip Tedbury

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Biology

2022

Acknowledgements

I would like to start by thanking all of the members of the Sarafianos lab, especially Dr. Stefan Sarafianos, Dr. Philip Tedbury, and Maria Cilento. Thank you Dr. Sarafianos for providing a wonderful lab environment open to undergraduate students to help them start their career in research. I have learned so many valuable skills that I will take with me throughout my career. Thank you Dr. Tedbury for being a great mentor during my time in the lab. Also, thank you for giving me the opportunity to work on the wide variety of projects, from the CRISPR imaging to HBV and finally this thesis project; they have all taught me so much. I appreciate all of the time and effort you gave to working with me, especially through the many failures and troubleshooting that came up over the years of working on these projects. Your mentorship has given me a strong foundation that I will carry with me through my graduate school career. Thank you Maria Cilento for giving me such a great introduction into research by training me as I first started, and supporting all of the work that I have done in the lab. Your great mentorship and friendship have given me so much, and I would not have survived this thesis without you (or the morning tea sessions). All of your work and time helping me with various experiments, writing, and most definitely email sending has provided so much support in my time in the lab. (I promise to not ask you to read any more emails).

Thank you to my committee members, Dr. Campbell and Dr. Deal for taking the time and effort to be on my thesis committee. I learned so much from both of your courses during my time at Emory, and you have given me the skills that I need to succeed in my career. Thank you for supporting my research, asking interesting questions, as well as supporting my journey towards graduate school. Your strong letters of recommendation were a fundamental part of my applications, and I appreciate all the time and effort you put in to writing the letters leading to a very successful application!

Lastly, I would like to thank my family, especially my mom, for supporting me throughout my life no matter what I end up doing. Thank you for listening to my many rants about the long days in lab and hard experiments. Your endless love and support (and many cards and flowers) carried me through the tough days and provided great motivation.

Table of Contents

## Introduction

*Human Immunodeficiency Virus Type 1*

As of 2020, there were about 38 million people globally infected with Human Immunodeficiency Virus Type 1 (HIV-1), and there were approximately 1.5 million new HIV-1 infections that year[1]. Worldwide, the incidence of HIV-1 infections has decreased 31% since 2010[1]. In 2019, there were around 16,000 deaths in the United States that were related to HIV-1 infections, and this emphasizes the importance of continuing HIV-1 research[2]. HIV-1 is a parenterally transmitted infection identified in humans in 1983 that can be spread through sexual contact, sharing of drug-use needles, or via mother-to-child transmission during birth or breastfeeding[3,4].

HIV-1 is an enveloped retrovirus that contains two copies of the viral single-stranded RNA genome contained within a capsid core. The virus infects cluster of differentiation 4 (CD4) positive cells of the immune system, predominantly CD4+ T cells. CD4+ T cells are an essential component of the immune system as they help regulate the innate and adaptive immune responses by producing cytokines. When HIV-1 infects a CD4+ T cell, the viral protein reverse transcriptase generates double-stranded complementary DNA (cDNA) from the single-stranded RNA template. The cDNA is then inserted into the host genome by the viral protein integrase, and the integrated cDNA is known as the provirus. Once inserted into the host genome, host transcription and translation machinery regulate the expression of the viral genes. The proteins that result from HIV-1 RNA translation need to be cleaved by the viral protein protease to produce infectious virions that bud from the host cell. Mature virions can infect other CD4+ T cells and spread the infection throughout the host.

HIV-1 infection progresses through three main stages: acute infection, chronic infection, and acquired immunodeficiency syndrome (AIDS), which results from the virus attacking the host's immune cells[5]. The acute phase of HIV-1 infection is the initial stage of infection after exposure and is characterized by the highest viral load in the blood. Patients can also show symptoms of infection such as fever, chills,

and other flu-like symptoms. The chronic phase follows the acute phase and in the absence of therapy, T cell levels rebound but viral replication continues. Patients in the chronic phase often do not show any clinical symptoms or show signs of infection. Over time and throughout the chronic phase, the number of host immune cells consistently decreases while the viral load increases. The final stage, AIDS, is the most severe phase as it leaves the host susceptible to opportunistic infections because of the depletion of CD4+ T cells that would be minor infections in an immunocompetent patient[5].

There is not a cure for HIV, but approved treatments can manage an HIV-1 infection to reduce the patient's viral load to almost undetectable levels, halting the loss of T cells, and preventing the onset of AIDS. Antiretroviral therapy (ART) is the current approved treatment for HIV-1[6]. ART consists of drugs that prevent the virus from replicating and spreading to uninfected cells, including nucleoside reverse transcriptase inhibitors, nonnucleoside reverse transcriptase inhibitors, protease inhibitors, and integrase inhibitors, among others[6]. ART only prevents the infection of new cells, but it cannot eradicate the virus from the body. This is because the integrated provirus remains in the host genome, and ART does not remove the provirus from the genome. ART maintains a low level of virus replication so that the virus is undetectable, thus nontransmissible. Research has shown that ART initiation is most effective closest to the onset of infection, as ART is less effective if started in the later stages of infection[7]. Treatment efficacy can also vary due to access to ART and the ability to adhere to the treatment requirements.

*HIV-1 Latency*

When HIV-1 infects an activated CD4+ immune cell, the virus actively replicates, produces new virions, and spreads to other cells. The host's immune system identifies and eliminates HIV-1-infected cells at the early stages of infection. However, a small number of infections occur in dormant CD4+ cells, naïve CD4+ cells, or CD4+ cells that will become memory T cells. Dormant and memory cells are transcriptionally downregulated to only perform the cellular functions necessary for survival. HIV-1 can

infect any of these types of CD4 + cells, and the virus is considered latent when it is not being actively transcribed or producing viral protein. T cells containing latent HIV-1 evade the immune system as there is no viral product to recognize. Memory T cells, including those with latent HIV-1, maintain their population via homeostatic replication, and this population maintenance contributes to the formation of a latent reservoir of HIV-1 infected cells. Memory CD4+ T cells can be reactivated upon exposure to their cognate epitope. Reactivation of memory CD4+ T cells upregulates transcription, and if those memory cells contain the HIV-1 provirus, then the viral genes are also likely to be transcribed[8,9]. This leads to a rebound of the HIV-1 infection if the host is not taking ART, but ART would be able to suppress the replication following reactivation, making ART a life-long therapy and highlighting the importance of adhering to the ART regimen throughout the patient's life[9]. The latent reservoir is thus the major barrier to curing HIV-1[9,10].

There are two proposed methods to target the latent reservoir of HIV-1 in the hopes to cure HIV: the "shock-and-kill" method, and the "block-and-lock" method. The rationale behind the "shock-and-kill" method is that reactivating the virus using latency reversal agents will give the immune system the ability to recognize and destroy infected cells[11,12]. This method is popular because there are many different types of latency reversal agents, and they can be combined for maximal effect[13,14]. Unfortunately, it has proven difficult to reactivate all latent cells as they are difficult to target and may reactivate randomly under treatment[15,16]. On the other hand, the "block-and-lock" method attempts to modify the epigenetic markings around the HIV-1 provirus to induce deep latency from which the virus cannot be reactivated, even if treatment stops[17–19]. Given that a percentage of the human genome is permanently silenced or defective retrovirus sequences, researchers are investigating how those retroviruses remain silent in order to replicate this silencing with the HIV-1 provirus[20,21]. There are studies being done to analyze the effects of various compounds in inducing deep latency to further develop the block-and-lock method[18].

*Lens Epithelium-Derived Growth Factor/p75 and Allosteric Integrase Inhibitors*

Once the HIV-1 RNA is reverse transcribed into double-stranded cDNA, the cDNA is inserted into the host genome by the viral integrase. Hundreds of host factors are reported to interact with integrase during HIV-1 infections[22,23]. Only a few of these host factors have been validated as promoting the productive and successful integration of HIV-1 DNA, and one particularly important host factor is lens epithelium derived growth factor (LEDGF)/p75.

LEDGF has two forms, p75 and p52, that result from alternative gene splicing[24,25]. Originally discovered as part of the lens epithelium development process, it has multiple roles including activating stress-related genes and double-stranded DNA break repair[26–28]. Both forms of LEDGF are host chromatin-binding factors and transcriptional co-activators that associate with gene-dense regions of the host genome as well as RNA splicing factors[24,29]. A structural domain of LEDGF/p75 known as the PWWP region gives LEDGF/p75 its chromatin interaction functions by interacting with H3K36me3 marks on nucleosomes, specifically mononucleosomes associated with euchromatin[30,31].

LEDGF/p75 was originally discovered to be involved in HIV-1 replication during studies of HIV-1 integrase multimerization[32]. Further investigation of the interaction between HIV-1 integrase and LEDGF/p75 revealed that there is a unique binding site for LEDGF/p75 in HIV-1 integrase[33]. Integrase mutagenesis studies that interfere with LEDGF/p75 binding, known as class II mutants, result in decreased integrase enzymatic activity as well as decreased stability and solubility of integrase[33]. Experiments that interfered with LEDGF/p75's ability to interact with integrase, revealed that LEDGF/p75 is involved in HIV-1 chromatin tethering[34,35]. Also, knockdown of LEDGF/p75 in HIV-1 infection models shows that LEDGF/p75 is involved in targeting the HIV-1 pre-integration complex towards transcriptional units and away from promoters of genes to ensure productive gene expression after integration[36–38]. As mentioned above, LEDGF/p75's association with RNA splicing factors could contribute to the preferential integration

of HIV-1 into highly spliced genes, as well as the association with mononucleosomes favoring active gene regions[29].

Realizing the importance of LEDGF/p75 in the productive integration of HIV-1 cDNA into the host genome, LEDGF/p75 became a promising target for HIV-1 therapeutics[39]. Integrase strand transfer inhibitors block the enzymatic function of integrase and are a component of ART, but HIV-1 has evolved resistance against many of the integrase inhibitors currently in use[40]. To inhibit integrase by interfering with host factor interactions, LEDGF/p75 inhibitors (LEDGINs), also known as allosteric integrase inhibitors (ALLINIs), were developed. These are small molecules that bind to integrase at the LEDGF-binding pocket and inhibit LEDGF/p75 binding to integrase[41–43]. ALLINIs have been shown to have a dual effect on HIV-1 replication; early-stage inhibition decreases productive integration of the HIV-1 cDNA into the host genome, and late-stage inhibition interferes with virion maturation[44,45]. Ablating the LEDGF/p75:integrase interaction has been shown to increase latency and decrease the ability to reactivate cells from latency[46,47]. This could be due to the integration of HIV-1 cDNA into silenced or gene-sparse regions of the genome[48]. The involvement of LEDGF/p75 in latency supports further characterization of HIV-1 latency and provided another approach to curing HIV-1, further promoting the "block-and-lock" strategy using ALLINIs.

*HIV-1 antisense RNA*

The HIV-1 RNA genome consists of a total of nine genes encoded in three open reading frames that result in the production of 15 proteins[49]. These proteins can be categorized into structural proteins, enzymes, envelope proteins, accessory proteins, and regulatory proteins. The main viral genes are flanked by long terminal repeats (LTRs) at the 5' and 3' end of the genome. The 5' LTR functions as the promoter of the HIV-1 genome, which regulates HIV-1 gene expression[50].

Sense RNA (sRNA) is synthesized by RNA polymerase II reading the template strand of DNA. sRNA can be referred to as messenger RNA (mRNA), and mRNA is often modified and spliced before being translated by the ribosome to produce proteins. Transcription of the HIV-1 provirus from the 5' LTR requires the viral protein Tat, a transcription factor that is one of the first viral proteins produced from expressed HIV-1 provirus, and Tat interacts with the viral trans-activation response (TAR) RNA to induce expression of the HIV-1 provirus[51]. Proviral expression results in the production of HIV-1 sRNA that is spliced and modified, as well as codes the viral proteins[52]. Full length genomic sRNA is also created and packaged in nascent virions[53].

A combination of sequencing and reverse transcription polymerase chain reaction (RT-PCR) experiments revealed the presence of an antisense RNA (asRNA) transcript in addition to the expected sRNA transcripts[54,55]. The early identification of asRNA transcripts was supported by more advanced next generation sequencing technology[56]. It is proposed that there is an ambiguous promoter located in the 3' LTR of the HIV-1 provirus that controls the expression of the antisense transcript[57,58]. The asRNA transcript is expressed at very low levels, and the predicted promoter does not appear to have the expected features of a gene promoter, such as the TATA box[58–60]. Sequencing of the asRNA transcript reveals that there is low polyadenylation of the RNA, which may be associated with the low abundance of the asRNA transcript in the cytoplasm and primary localization in the nucleus[61].

The exact mechanisms of asRNA expression and transcription are unknown, as is the function. Studies show that there is the potential translation of asRNA to yield an antisense protein (ASP)[54]. The function of ASP is not known, but there are various studies that suggest the ASP could play a role in viral entry, viral budding, and viral replication[61]. It has been shown that ASP can be associated with euchromatin in latent cells, but localizes to the plasma membrane when dormant cells are stimulated[62]. Patients infected with HIV-1 have been shown to be able to develop antibodies against ASP[63].

In an attempt to investigate the function of asRNA, studies have shown that the asRNA transcript may play a role in the epigenetic regulation of viral transcription by interacting with epigenetic silencing proteins and decreasing HIV-1 transcription from the 5' LTR[64,65]. Consistent with these findings, a number of studies have found an association between HIV-1 latency and asRNA[60,64,65]. In some latent models, depletion of asRNA can lead to reactivation from latency[66]. Other studies found that latent cells containing asRNA are less likely to be reactivated from latency when treated with a latency reversal agent[66]. The role of asRNA in latent cells is not fully understood, but the role of asRNA in the induction or maintenance of latency is well supported. It is possible that the epigenetic and silencing functions of both the asRNA transcript and the ASP could be involved. Taken together, it is possible that both the asRNA transcript itself and the ASP can play a role in HIV-1 pathogenesis.

*Dual Fluorescent HIV-1 Reporters*

Reporter viruses are common laboratory tools to study the cellular, molecular, and systemic mechanisms of viral infections. Common expression markers are enzymes such as luciferase, or fluorescent proteins such as green fluorescent protein (GFP). A single fluorescent HIV-1 reporter virus, which is most commonly used, can be used to visually verify that cells were successfully infected. In an attempt to more effectively study HIV-1 latency, HIV-1 reporters with two fluorescent proteins were developed. Dual fluorescent HIV-1 reporters contain two different fluorescent proteins controlled under two independent promoters. One variation of a dual fluorescent reporter is the Hi-Fate Tomato HIV-1 Dual Reporter (Hi-Fate Tomato), and this construct was modified from a previously published reporter to encode the tdTomato fluorescent protein under the control of the HIV-1 5' LTR promoter as well as eGFP that is controlled under the constitutively active EF1-α mammalian promoter[67].

The use of dual fluorescent reporters allows for the identification of active and latent cell populations within a single sample. Infected cells that are actively expressing HIV-1 genes will be positive

for both fluorescent proteins, whereas latent HIV-1-infected cells only express the fluorescent protein that is controlled with the constitutive promoter. These populations can be observed and quantified using flow cytometry, and the populations can be sorted using fluorescence-activated cell sorting (FACS). Dual reporters allow for the direct characterization of active versus latent cells which allows for better studies that directly investigate latent cells and the mechanisms of latency.

Dual fluorescent reporters have previously been used to investigate the efficacy of latency reversal agents, the impact of integration site on latency, the dynamics of the formation of the latent reservoir, and the impacts of abrogating the LEDGF/p75:integrase interaction[47,68–70]. For this study, the Hi-Fate Tomato reporter will be used to identify and separate latent and actively infected cell populations that will then be analyzed for the production and detection of asRNA.

*Rationale*

As previous research has connected LEDGF/p75 and asRNA to HIV-1 latency, this work seeks to investigate asRNA production as a result of treatment with the ALLINI BI-D, and how these conditions influence latency. While ALLINIs have been shown to increase latency and decrease the ability to reverse latency in HIV-1 infected cells, there has not been an established association between BI-D treatment and the production of asRNA. Preliminary data from the Sarafianos lab shows that there is an increase in asRNA when ALLINIs are used, especially with BI-D (unpublished). There has not been any association between LEDGF/p75:integrase interaction ablation and asRNA, but this study seeks to further characterize the relationship between asRNA and LEDGF/p75 depletion using BI-D.

We studied asRNA production using the dual-fluorescent Hi-Fate Tomato HIV-1 reporter and used flow cytometry and fluorescence activated cell sorting to analyze asRNA production in three subpopulations of cells. We used a modified RNAScope protocol to stain for sRNA and asRNA in cells that

were left untreated, and in cells that were treated with BI-D to analyze the effect of BI-D on HIV-1 asRNA production and if there is a correlation with HIV-1 latency.

**Materials and Methods**

*Cell Lines, HIV-1 Reporters, and Virus Constructs*

HEK 293/17 cells (American Type Culture Collection CRL-11268) were cultured and maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% Serum Plus (Sigma Aldrich), 100 U/mL penicillin, 10 µg/mL streptomycin, and 2 mM L-glutamine in a humidified incubator at 37°C with 5% carbon dioxide. HEK 293/17 cells are the 17$^{th}$ clone of a derivative of 293T cells that have the inserted simian virus 40 large T-antigen, which gives the cells increased transfection and transduction efficiency[71].

The HIV-1 reporter virus used as the wild-type reference virus was the HIV-1 NL4-3 Gag-iGFP delEnv (ARP-12455 from Dr. Benjamin Chen), referred to as wild-type HIV-1 (WT HIV-1) for this study. This reporter is a non-infectious full-length NL4-3 clone with a nonsense mutation in the envelope gene, and GFP tagged Gag protein. The use of this reporter vector allowed for visual confirmation of infection, as well as work in a biosafety level 2 setting. The dual fluorescent HIV-1 reporter is the NL4-3 Hi-Fate tdTomato EF1-α reporter (from Alon Herschhorn and modified from reference 66). For single-color flow cytometry control samples, pCMV-tdTomato (Takara) encoding the tdTomato fluorescent protein was used. The HIV-1 NL4-3 Gag-iGFP delEnv plasmid was used as a single-color control for flow cytometry.

The WT HIV-1 and Hi-Fate Tomato constructs were pseudotyped with vesicular stomatitis virus-G (VSV-G) to produce virus-like particles (VLPs) that simulate HIV-1 infections. VLP stocks were generated by transfecting a 10 cm dish containing one million HEK 293/17 cells with 5 µg of the respective reporter virus, 1.5 µg of VSV-G plasmid, and 20 µL of Xtreme-GENE HP (Roche) transfection reagent mixed in 330 µL of Opti-MEM serum-free media (Gibco). The transfection mix was incubated at room temperature for 15 minutes before pipetting dropwise onto cells. Approximately 72 hours following transfection, the supernatant was collected, filtered through 0.45 µm filters, and aliquoted into 1.7 mL Eppendorf tubes. The aliquots frozen and stored at -80°C until use.

*Allosteric Integrase Inhibitor – BI-D*

The allosteric integrase inhibitor BI-D (ALLINI, from Alan Engleman) was used. BI-D is a potent inhibitor that leads to the multimerization of integrase, decreasing LEDGF/p75 activity in early-stage infection and preventing virion maturation in late-stage infection[45,72,73]. This study focused on the early effects of ALLINIs where BI-D prevents integrase from interacting with LEDGF/p75 leading to decreased integration of the HIV-1 provirus. In all experiments where BI-D was used, the ALLINI was added at 10 µM final concentration one hour following VLP transduction. This concentration represents 10 times the published $EC_{50}$ in HEK 293 cells and allows for maximal inhibition of the integrase-LEDGF/p75 interaction with limited toxicity[73].

*Nucleic Acid Probes*

Nucleic acid probes were purchased from Advanced Cell Diagnostics (ACD) for RNA staining experiments to label HIV-1 sRNA and asRNA. To detect HIV-1 sRNA, an antisense Channel 2 probe targeting the HIV-1 non*Gag-Pol* region (nucleotide target range: 4988-9181) was used (ACD reference number: 317711). To detect HIV-1 asRNA, a sense Channel 1 probe targeting the *Gag-Pol* coding region (nucleotide target range: 507-4601) was used (ACD reference number: 317701). The sRNA Channel 2 probe was prepared at a 1:50 dilution in combination with the Channel 1 probe and probe diluent (ACD) for RNA staining. The channel of the probe dictates the excitation wavelength of the final amplifier (Amp4) that is used. Different combinations of probe channels and amplifiers yield different combinations of colors that can be detected using a fluorescent microscope.

*HIV-1 RNA staining*

RNAscope (ACD) and previously published multiplex immunofluorescent cell-based detection of DNA, RNA, and protein (MICDDRP) protocols were modified to optimize an RNA fluorescent *in-situ*

hybridization (FISH) protocol to stain and label HIV-1 sRNA and asRNA[74,75]. The protocol described here and used throughout the study will be referred to as the modified RNAScope procedure. All reagents, probes, and buffers were prepared at the specified temperatures and dilutions as described in the MICDDRP protocol[75]. All incubations using the HybEZ humidified oven (ACD) were at 40°C for the specified duration. HEK293/17 cells used for RNA staining were seeded on glass collagen coated coverslips (Neuvitro) additionally coated with 20 µg/mL of poly-D-lysine in a 24 well plate at 100,000 cells per well. The following day, the seeded cells were transduced with previously prepared VLPs at a dilution that resulted in approximately 50-80% of the cells transduced successfully.

At the desired time point following transduction, media from the plate was aspirated, and cells were washed twice with Dulbecco's phosphate-buffered saline (DPBS, Corning). Cells were then incubated at room temperature with 4% paraformaldehyde (PFA) for 30 minutes. The PFA was aspirated, and cells were washed twice with DPBS. After fixation, cells were then permeabilized with 500 µL of 0.1% (v/v) Tween-20 in 1x DPBS (PBS-T) and incubated at room temperature for 10 minutes. PBS-T was aspirated, and the cells were washed with DPBS twice. The coverslips were removed from the 24 well plate and immobilized on sterilized glass microscope slides by placing the back of the coverslips on a drop of clear nail polish and covered with 1x DPBS to prevent drying. A square barrier was drawn one centimeter away from each coverslip using an ImmEDGE hydrophobic barrier pen (Fisher Scientific). Immobilized coverslips were treated with 100 µL of 1:5 diluted Protease III (ACD) in DPBS and incubated in a humidified oven for 15 minutes. Protease treated cells were then submerged 10 times in 1x DPBS to wash. After washing, cells were incubated in the humidified oven for 2 hours with 100 µL of diluted HIV-1 sense and antisense RNAscope probes in ACD provided probe diluent. Cells were washed with 1:50 dilution of ACD RNAscope Wash Buffer diluted with molecular biology-grade Millipore water by submerging. Following target probe incubation and washing, cells are incubated with ACD probe amplifiers 1, 2, 3, and 4 sequentially for 30 minutes, 15 minutes, 30 minutes, and 15 minutes respectively in a humidified oven and washing with

diluted ACD Wash Buffer following each amplifier. Amp4 Alt B-FL was used for these experiments to result in Alexa 488/540 nm (excitation/emission) wavelength for sRNA, and Alto 550/580 nm (excitation/emission) for asRNA. Coverslips were washed in 1x DPBS following the final wash with ACD Wash Buffer. Samples were then stained with one drop of the ACD-provided 4',6'-diamino-2-phenylindone (DAPI) stain for one minute, then washed with 1x DPBS by submerging. Coverslips were removed from the nail polish on the initial microscope slide and were mounted onto sterilized glass microscope slides using the Prolong Gold antifade mounting medium (Thermo Fisher) with the cell-side of the coverslip in the mounting medium. To seal the coverslip and hold it in place, the edges of the coverslips were lined with clear nail polish.

If the samples were not able to be stained immediately following fixation with PFA, then the cells were dehydrated. To dehydrate the cells, 1x DPBS was replaced with 500 µL of 50% ethanol and incubated at room temperature for 5 minutes. The 50% ethanol was discarded, replaced with 500 µL of 70% ethanol, and incubated for 5 minutes at room temperature. The 70% ethanol was discarded, replaced with 500 µL of 100% ethanol, and incubated for 5 minutes at room temperature. Following the final incubation, the 100% ethanol was replaced with 500 µL of fresh 100% ethanol. Coverslips can be stored at -20°C for up to 6 months until needed for the modified RNAScope staining protocol. Once the samples were needed, the cells were rehydrated by reversing the ethanol concentration sequence and incubating for 2 minutes at room temperature until replacing the final volume of ethanol with 1x DPBS. Dehydration and rehydration steps do not impact the efficacy of RNAScope staining, nor result in significant sample loss.

*Spectral Flow Cytometry*

HEK 293/17 cells were seeded into 24 well plates without coverslips, infected with HIV-1 VLPs, and treated with BI-D as described above. Twenty-four hours following infection and treatment, cells were washed, trypsinized, and reseeded into 12 well plates, and multiple wells from the original plate were

combined into a single well to increase confluency. Cells were prepared for flow cytometry 48 hours following the initial VLP transduction and BI-D treatment.

To prepare cells for flow cytometry, media in the wells was aspirated, and cells were washed with 1x DPBS. Cells in a 12 well plate were treated with 250 µL of 0.25% trypsin-EDTA (Sigma) for two minutes. Cells were then resuspended in 1 mL of DMEM. Resuspended cells of the same treatment conditions were combined into a 15 mL centrifuge tube to increase population size. Cells were centrifuged at 100Xg for 4 minutes. Media was aspirated carefully not to disturb the cell pellet, and cells were resuspended in 1x DPBS. The centrifugation and wash steps with 1x DPBS were repeated. Cells were then fixed in 500 µL of 4% PFA and incubated at room temperature for 15 minutes. Following fixation, cells were centrifuged, PFA was aspirated, and then resuspended with 1x DPBS. The wash with 1x DPBS was repeated. Following the second round of washing, cells were filtered through a 40 µM nylon cell strainer (Corning) and collected into 5 mL 12x75 mm polystyrene culture tubes (Fisher Healthcare).

Samples were analyzed on the Cytek Aurora 4-laser (405nm-488nm-516nm-640nm) spectral flow cytometer at the Emory University Pediatrics/Winship Flow Cytometry Core. Data were unmixed in the SpectroFlo software version 2.2.0.4 (Cytek Biosciences), and sample flow cytometry data were gated and analyzed with FlowJo version 10.8.1 (Becton, Dickinson & Company).

*Fluorescence Activated Cell Sorting*

HEK 293/17 cells were seeded into 10 cm cell culture dishes at 2 million cells per dish. After 24 hours, cells were transduced with the HIV-1 dual fluorescent VLPs and treated with 10 µM BI-D as previously described. Cells were prepared for fluorescence activated cell sorting (FACS) 24 hours following the VLP transduction and BI-D treatment.

To prepare cells for FACS, the cells in the dishes were washed, trypsinized, and resuspended in media as described for spectral flow cytometry. The resuspended cells were then centrifuged at 100Xg for

4 minutes. The medium was aspirated, and the cells were resuspended in 1x DPBS. The centrifugation, aspiration, and resuspension steps were repeated twice. The final resuspension was completed with 1x DPBS with 25 µM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and 1 mM ethylenediaminetetraacetic acid (EDTA) at 4°C. Following the final resuspension, cells were filtered through a 40 µM nylon cell strainer (Corning) and collected into 5 mL 12x75 mm polystyrene culture tubes (Fisher Healthcare).

Prior to sorting, 5 mL 12x75 mm polystyrene culture tubes used for collection were filled with SerumPlus and kept at 4 °C for 24 hours. SerumPlus was removed and 1 mL of DMEM was placed in the tubes for collection, and the tubes were kept at 4 °C for the entire sorting duration. Samples were analyzed and sorted using the BD FACS Aria II SORP cell sorter at the Emory University Pediatrics/Winship Flow Cytometry Core. Once sorted, the samples were immediately moved into 15 mL centrifuge tubes and centrifuged at 100Xg for 4 minutes. The supernatant was aspirated, and the cells were resuspended in 100 µL of DMEM. 50 µL of cells were seeded onto glass coverslips coated with both poly-D-lysine and collagen (Neuvitro) and incubated at 37°C for 30 minutes. After 30 minutes, the modified RNAScope protocol was started. Following cell fixation, cells were dehydrated and stored until ready to use.

*Imaging, Image Analysis, and Statistics*

Following the modified RNAScope staining protocol, samples were imaged using the Agilent/BioTek Cytation5 fluorescent microscope in a 3x3 montage at 10x magnification. Cell nuclei were imaged using the DAPI filter cube (360/460 – excitation/emission). Sense RNA was imaged using the GFP filter cube (482/528 – excitation/emission). Antisense RNA was imaged using the TexasRed filter cube (586/647 – excitation/emission). Microscope settings were determined based on control samples and maintained for all other samples within each experiment. Images presented were adjusted for optimal

brightness and contrast, and all images were identically adjusted. The raw fluorescence values used for data calculations are not impacted by these adjustments.

Sample images were analyzed using the BioTek Gen5 software (v5.3). Total cell counts were obtained by counting the number of nuclei identified in the image montage. The identified cell nuclei created a primary mask to then identify subpopulations of cells. Cells were counted as containing sRNA if the sum of the green fluorescence intensity was above a 500,000-unit threshold within the nuclear primary mask. Cells were counted as containing asRNA if the sum of Texas Red fluorescence was above a 600,000-unit threshold within the nuclear primary mask. Cells that labelled as transduced contained sRNA or asRNA.

For rigor and reproducibility, three independent replicates were performed in the initial asRNA detection experiments and the spectral flow cytometry experiments. Time and scheduling limited fluorescence activated cell sorting and subsequent RNA staining to one replicate.

Data calculations were performed in GraphPad Prism (v9), and graphs were made in Microsoft Excel. To test for a significant difference in the percentages of sRNA and asRNA within a condition, a two-way analysis of variance (ANOVA) with Sidak's multiple comparisons was performed using Prism. To test for a significant difference in the percentages of cells in a subpopulation from fluorescence activated cell sorting, a one-way ANOVA with Tukey's multiple comparisons test was performed using Prism.

**Results**

*Hi-Fate Tomato Dual-Fluorescent HIV-1 Reporter Produces asRNA*

For this study, reporter viruses were used to investigate the relationship between HIV-1 latency and antisense RNA (asRNA), and how this relationship changes in ALLINI-treated cells. Reporter viruses are a commonly used tool to model viral infections, especially HIV-1. A widely employed reporter is NL4-3 delEnv-iGFP (referred to here as WT HIV-1), a lab adapted strain of HIV-1 that encodes GFP within of the main structural genes, *gag*[76,77]. The WT HIV-1 single fluorescent reporter was used in preliminary experiments and has shown to reliably produce detectable asRNA that increases when treated with the ALLINI BI-D (unpublished). While the GFP reporter allows for clear identification of actively infected cells, it cannot directly indicate latently infected cells. To determine if the cells are latently infected it would require the quantity of GFP to be monitored in the sample over time, and once the isolated infected cells have decreased GFP expression, they would be considered latent. The use of dual-fluorescent HIV-1 reporters provide a more reliable method of determining latent and active HIV-1 infections based on the expression of independently controlled fluorescent proteins.

This study uses the Hi-Fate Tomato dual fluorescent HIV-1 reporter, which encodes the tdTomato fluorescent protein under the control of the HIV-1 5' LTR promoter, and the enhanced green fluorescent protein (eGFP) controlled under the constitutively active EF1-α mammalian promoter[67] (Figure 1). The combination of fluorescent proteins indicates infection status; double positive cells are actively infected, single positive cells for eGFP are latent, and double negative cells are uninfected. Due to the addition of the coding region of a fluorescent protein in the 3' end of the HIV-1 genome, it was necessary to determine if the Hi-Fate Tomato reporter produces detectable asRNA, and if the amount of asRNA increases when treating with BI-D.

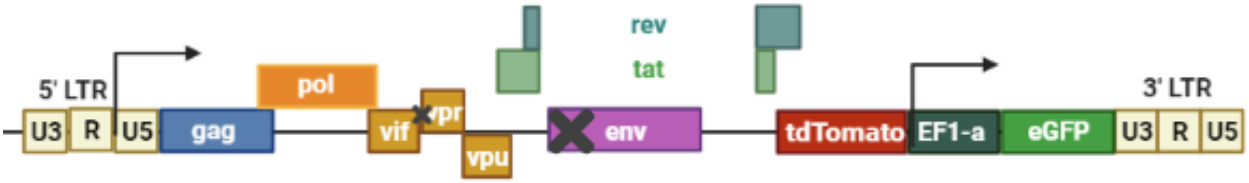



**Figure 1.** *Hi-Fate Tomato dual reporter allows for the separation of active and latent cells using independently controlled fluorescent proteins.* The Hi-Fate Tomato dual reporter is the lab-generated NL4-3 wild-type genome with two nonsense mutations, one in the *vpr* accessory protein gene, and one in the *envelope* gene, that render the reporter noninfectious. The tdTomato fluorescent protein is located in the *nef* reading frame, and the eGFP sequence directed by the EF1-α constitutive mammalian promoter follows the tdTomato sequence. The two fluorescent proteins and the internal promoter are followed by the HIV-1 3' LTR.
Figure created with Biorender and adapted from Ratnapriya, S. et al. (2021) Cell Reports (reference number 66)

The production of sense RNA (sRNA) and asRNA was determined using the modified RNAScope protocol and imaged using a Cytation5 microscope. Representative images of the sRNA (green) and asRNA (red) detected when cells transduced with WT HIV-1 (Figure 2A) or the Hi-Fate Tomato dual reporter (Figure 2C) virus-like particles (VLPs) show that similar amounts of sRNA are produced when comparing the WT HIV-1 reporter and the Hi-Fate Tomato reporter. Cells transduced with VLPs and subsequently treated with BI-D have fewer transduced cells overall, which is shown by the decreased amount of sRNA but show an increase in asRNA (Figure 2B and 2D). When the cells are transduced with WT HIV-1 VLPs without BI-D treatment, on average 51.0% of cells are successfully transduced, and when treated with BI-D the number of successfully transduced cells significantly decreases to an average of 9.7% (Figure 3A, $p = 0.008$, n = 3). When the Hi-Fate Tomato VLPs were transduced into 293/17 cells, an average of 21.8% of cells transduced successfully compared to cells treated with BI-D that produced an average of 6.2% of cells that were successfully transduced (Figure 3A, $p = 0.037$, n = 3). These data suggest that treatment with BI-D results in a similar decrease in the total number of successfully transduced cells with the Hi-Fate Tomato dual reporter compared to the HIV-1 WT reporter.

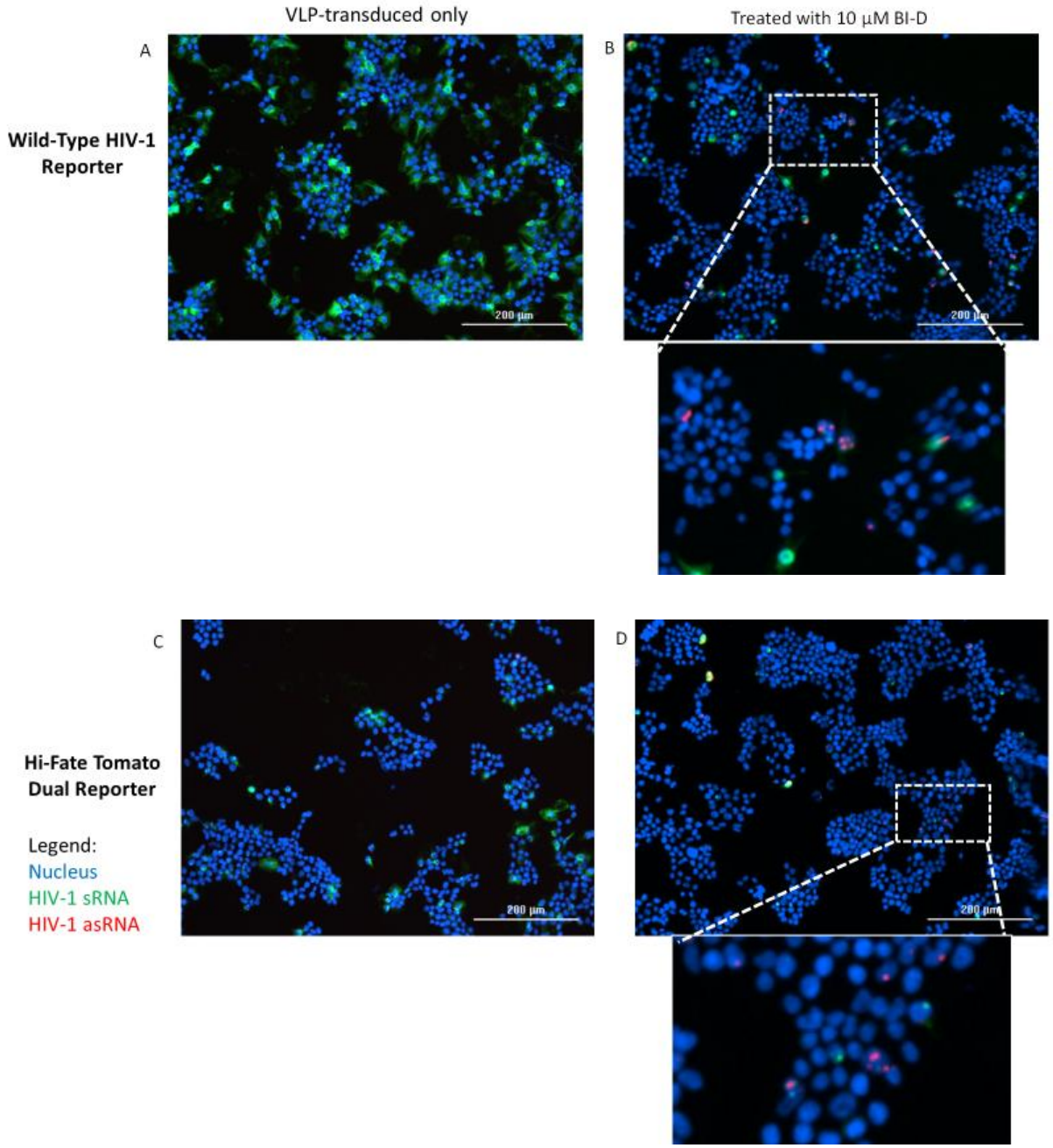



**Figure 2.** *WT HIV-1 reporter and the Hi-Fate Tomato dual reporter produce detectable asRNA.* HEK293/17 cells were transduced with VLPs and stained using a modified RNAScope protocol 48 hours after transduction to visualize HIV-1 sRNA and asRNA. Images are a single set of representative images captured at 10X magnification on a Cytation5 microscope. Images were modified for maximum brightness and contrast. Dashed white boxes highlight examples of detectable HIV-1 asRNA and are enlarged below the image to visualize asRNA.

The Sarafianos lab previously showed the WT HIV-1 reporter with BI-D treatment results in an increased percentage of cells that contain asRNA (unpublished). The coding regions of the fluorescent proteins in the Hi-Fate Tomato reporter are located near the 3' LTR, which could disrupt the production

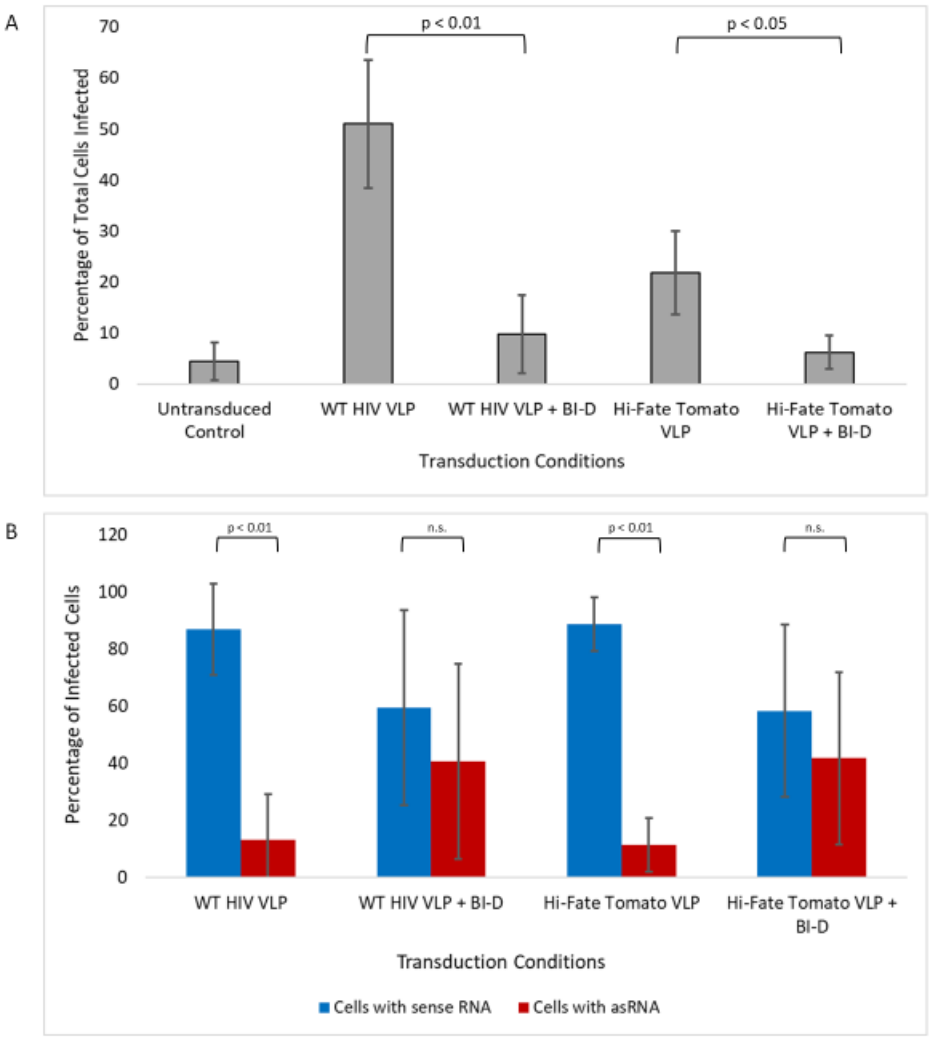



**Figure 3.** *Hi-Fate Tomato reporter produces detectable asRNA in comparable quantities, and BI-D treatment results in similar decrease in infection compared to HIV-1 Wild-Type reporter.* Cytation5 images of cells that were stained with the modified RNAScope protocol were quantified using the BioTek software. Cell counts were obtained by identifying cell nuclei, then counting an individual cell as containing sRNA or asRNA if the sum amount of green or red fluorescence was above a 500,000- or 600,000-unit threshold, respectively, within the cell nucleus. Cells labeled infected contained sRNA or asRNA. A) Percentage of the total number of cells counted, infected and uninfected, that contained sRNA or asRNA. B) Percentage of cells with sRNA and asRNA out of the cells that were infected. Error bars represent the standard deviation. A Student's two sample T-test was used to determine if there is a significant difference in the total number of cells infected (A) and two-way ANOVA with Sidak's multiple comparisons test was used to determine if there is a significant difference in the distributions of sRNA and asRNA in each condition (B). n.s – not significant ($p > 0.05$). n =3.

and detection of asRNA. The production of detectable asRNA by the Hi-Fate Tomato reporter and WT HIV-1 reporter needed to be determined with and without BI-D treatment. Quantification of the images of sRNA and asRNA show that 86.8% of successfully HIV-1 WT VLP-transduced cells contain sRNA, and 13.2% of cells contain asRNA (Figure 3B, $p = 0.008$). Similarly, when cells are transduced with the Hi-Fate Tomato VLP, 88.6% of transduced cells contain sRNA and 11.4% of transduced cells contain asRNA ($p = 0.006$). When cells transduced with the WT HIV-1 VLP are treated with BI-D, 59.5% of cells contain sRNA and 40.5% of cells contain asRNA ($p > 0.05$), and cells transduced with the Hi-Fate Tomato reporter and treated with BI-D result in 58.3% of transduced cells containing sRNA and 41.7% containing asRNA ($p > 0.05$). The results show that the percentages of cells with sRNA or asRNA are significantly different without BI-D, but the percentages change and become similar with BI-D treatment to show that there is a detectable increase in asRNA (Figure 3B, n = 3).

These results show that the Hi-Fate Tomato HIV-1 reporter produces detectable asRNA comparable to WT HIV-1 VLPs. Also, BI-D treatment results in decreased transduction and increased percentage of cells with asRNA. Based on these results, the Hi-Fate Tomato VLPs were used in subsequent experiments as the trends were similar, qualitatively and statistically, to the WT HIV-1 reporter.

*Spectral Flow Cytometry Analysis of Hi-Fate Tomato Populations*

The Hi-Fate Tomato dual fluorescent HIV-1 reporter produces two independently expressed fluorescent proteins (eGFP and tdTomato, as mentioned above) to identify various subpopulations of transduced cells. The use of these two fluorophores in a dual reporter has not been published, so the distribution of subpopulations after VLP transduction with and without treatment with BI-D was investigated using spectral flow cytometry.

The percentage of active, latent, and uninfected cells 48 hours after Hi-Fate Tomato VLP transduction and BI-D treatment was determined using quadrant gates, as this gating strategy was used in previous publications using dual fluorescent HIV-1 reporters[78] (Supplementary Figure 1). Transduction of HEK 293/17 cells with the Hi-Fate Tomato VLPs results in 24.9% actively infected cells, 38.4% latent cells, and 36.4% of cells are untransduced after 48 hours (Figure 4, $p > 0.05$ between all subpopulations). With BI-D treatment, the percentage of actively infected cells decreases to 7.8%, the percentage of latent cells is maintained at an average of 34.8%, and the percentage of untransduced cells increases to 57.0% (Figure 4). Comparing the percentages of cells in each subpopulation, the percentage of cells that were active and latent are significantly different ($p = 0.001$), and the percentage of cells that were latent and untransduced ($p = 0.006$) were significant. While ALLINIs have been shown to increase latency, the use of dual fluorescent HIV-1 reporters reveals that the percentage of latent cells remains similar while the percentage of untransduced cells increases[46,47].

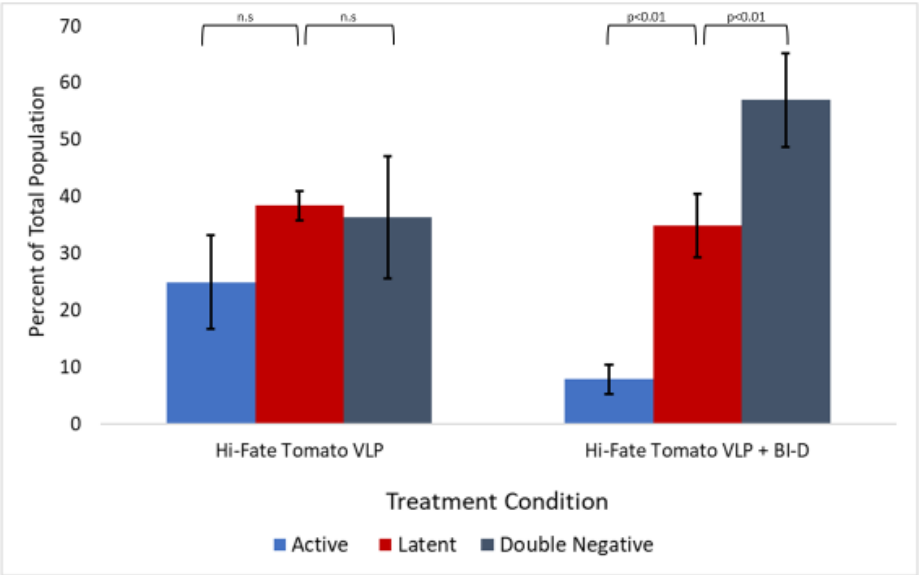



**Figure 4.** *BI-D treatment after Tomato VLP transduction results in an increase in uninfected cells and a decrease in actively infected cells.* Two samples of HEK293/17 cells were transduced with the Hi-Fate Tomato dual reporter VLP, and one sample was treated with 10 µM BI-D. The various cell populations were determined using the Cytek Aurora flow cytometer and the data were analyzed using FlowJo v10.8.1. Active infection represents eGFP and tdTomato double-positive cells. Latent cells are eGFP single-positive cells. Uninfected cells are cells that were negative for both fluorophores. Error bars represent the standard deviation of each average. A one-way analysis of variance with Tukey's multiple comparisons test was performed to determine if there is a significant difference between the active, latent, and uninfected populations in both conditions. n.s – not significant. n = 3.

*asRNA production in subpopulations of HIV-1 infected cells*

The use of the Hi-Fate Tomato dual fluorescent HIV-1 reporter allows for not only the analysis of the various subpopulations of cells within a sample using spectral flow cytometry, but also the separation of the different subpopulations using fluorescence activated cell sorting (FACS). HEK293/17 cells transduced with the Hi-Fate Tomato VLPs were separated into three subpopulations with gates drawn during cell sorting: active (both eGFP and tdTomato expressed), latent (eGFP expressed), and untransduced (no fluorescence) (Supplementary Figure 2). Following cell sorting, cells from each subpopulation were seeded onto coverslips and stained using the modified RNAScope procedure (Figure 5). The data represent the subpopulations 24 hours following VLP transduction.

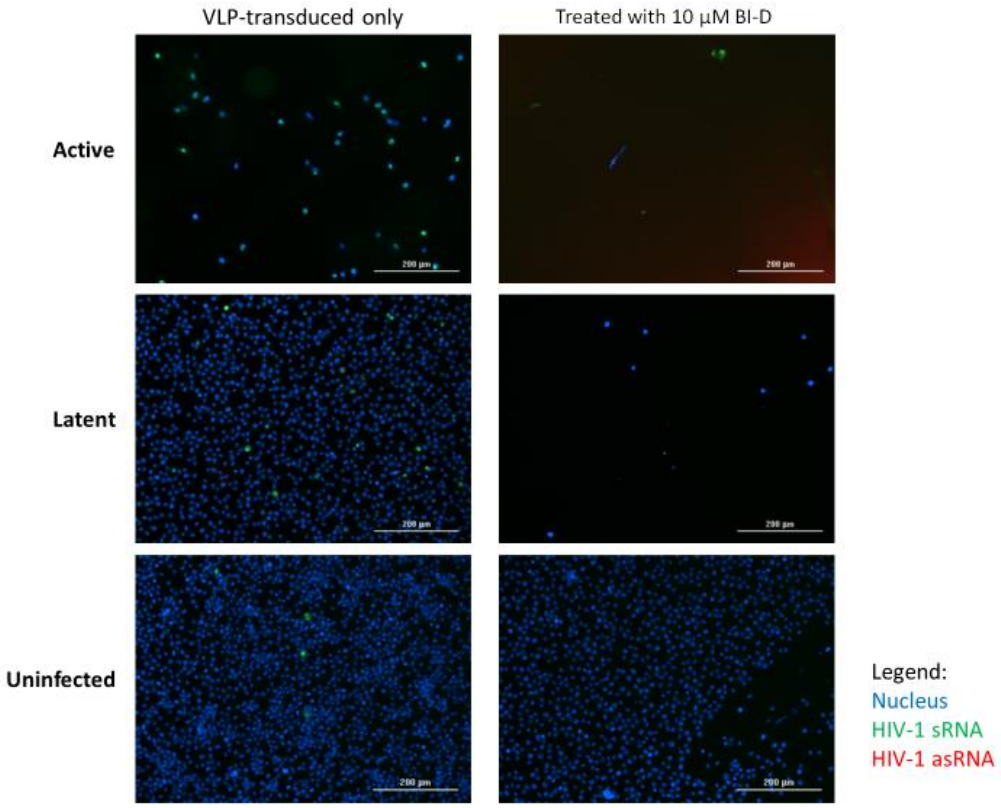



**Figure 5.** *Sorted populations of Hi-Fate Tomato transduced cells show decreased sRNA and asRNA production after BI-D treatment.* HEK293/17 cells were transduced with Hi-Fate Tomato VLPs and sorted into active, latent, and uninfected cell populations 24 hours after transduction based on fluorescent protein expression as previously described. Cells were prepared and stained according to the modified RNAScope procedure. Sparse cells were the result of decreased cells counts in the respective populations after sorting.

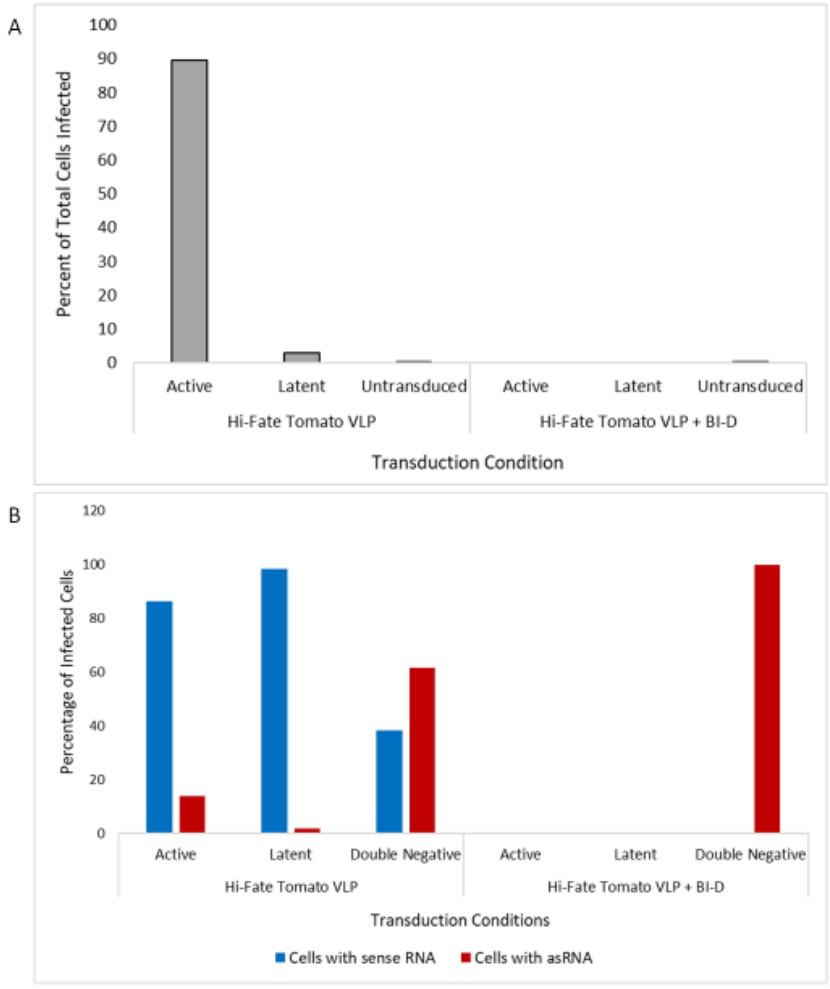



**Figure 6.** *BI-D treatment results in decreased infection of the Hi-Fate Tomato VLPs, actively infected, and latent cells contain mostly HIV-1 sRNA.* Following VLP transduction, cell sorting, RNA staining, and imaging on the Cytation5 microscope, cells with sRNA and asRNA were counted using the BioTek software. Cell counts were obtained by identifying cell nuclei, then counting an individual cell as containing sRNA or asRNA if the sum amount of green or red fluorescence was above a 500,000 or 300,000 unit threshold, respectively, within the cell nucleus. Cells labeled infected contained sRNA or asRNA. A) Percentage of the total number of cells counted, infected and uninfected, that contained sRNA or asRNA in each sorted population when untreated or treated with BI-D. B) Percentage of cells with sRNA and asRNA out of the cells that were infected in each sorted population from two treatment groups, with and without BI-D. N =1.

The percentage of detectable sRNA or asRNA was determined by quantifying fluorescence in the images, and any cell that had detectable RNA was considered transduced. Cells that were transduced with the Hi-Fate Tomato VLP resulted in 89.5% of cells sorted based on fluorescence as actively infected (eGFP and tdTomato double positive), compared to 3.0% of cells sorted as latent (eGFP single positive), and 0.4%

cells sorted as untransduced (no fluorescence) that were then shown to have detectable sRNA or asRNA (Figure 6A). When cells were treated with BI-D: 0% of sorted active cells, 0% of sorted latent cells, and 0.05% of sorted untransduced cells have detectable sense or asRNA (Figure 6A). 86.2% and 13.7% of the double positive, actively infected cells contain sense and asRNA, respectively (Figure 6B). The cells sorted into the latent population showed 98.3% of the transduced cells containing sense RNA and 1.7% of cells contain asRNA (Figure 6B). Approximately 38.5% of the cells sorted as untransduced but shown to express HIV-1 RNA contain sense RNA, and 61.5% of cells contain asRNA. When treated with BI-D, the active and latent populations did not produce cells that contained sRNA or asRNA, as well as resulted in a low number of cells that were seeded on the coverslips. For example, the active population did not have any cells seeded on the coverslip, and the latent population contained 35 cells total. BI-D treatment results 100% of the infected untransduced cells containing asRNA, which is due to only 3 cells containing any detectable RNA, all asRNA (Figure 6B). These data suggest that treatment with BI-D decreases successful transduction by limiting sense and asRNA production compared to cells that were not treated with BI-D. The distribution of sRNA and asRNA within the infected cells are not representative due to the low population numbers sorted by the flow cytometry gates as well as the shorter time following VLP transduction.

**Discussion**

HIV-1 continues to infect millions of people each year, and the progression of an HIV-1 infection into AIDS is a major public health issue. While treatments exist to manage viral replication below detectable levels and decrease transmission of an established infection, a major barrier to curing HIV-1 is dormant provirus harbored in the latent reservoir. HIV-1 asRNA can be transcribed from an unknown promoter in the 3' LTR of the HIV-1 genome, and asRNA has been shown to be present in latent cells as well as to decrease the reactivation of infected cells from latency[57,58,60,64,65]. Therapies to manage and eradicate the HIV-1 reservoir have been developed with two main strategies: inducing deep latency or reactivating cells from latency. One approach to treating HIV-1 by inducing deep latency is using ALLINIs that work to decrease integration into active gene regions and interfere with virion maturation[44,45,48]. Unpublished data from the Sarafianos lab reveal that there is an increase in asRNA production when the LEDGF/p75:integrase interaction is ablated in LEDGF/p75 knockout cells and or by treatment with ALLINIs. This study aimed to define how asRNA production is related to ablating the LEDGF/p75:integrase interaction using the ALLINI BI-D, and connect the relationship to how BI-D-induced asRNA is related to HIV-1 latency.

We used the dual fluorescent HIV reporter Hi-Fate Tomato with independently controlled fluorescent proteins to better understand HIV-1 latency. While dual reporters have been used for other investigations of HIV-1 latency, this reporter has not been previously used to investigate HIV-1 asRNA. The first aim of this study was to determine if the Hi-Fate Tomato dual reporter produces detectable asRNA with the RNAScope probes available as the fluorescent protein coding regions are located in the 3' end of the HIV-1 genome and would be part of the antisense transcript, and to determine if treatment with BI-D leads to similar decreases in HIV-1 infection compared to WT HIV-1. Previous use of the WT HIV-1 reporter (NL4-3 Gag-iGFP delEnv) revealed an increase in asRNA production and a decrease in active infection when cells are treated with BI-D (unpublished). When we transduced 293/17 cells with WT HIV-1 VLPs and

treated with BI-D were stained for HIV-1 sRNA and asRNA, there was an observable decrease in sRNA, indicating a decrease in overall infection, and there was an increase in the percentage of infected cells that contained asRNA. The cells transduced with Hi-Fate Tomato VLPs had a similar response to BI-D treatment with decreased transduction and increased asRNA production in transduced cells confirmed that this reporter can confidently be used for asRNA investigations as well as the impact of BI-D on HIV-1 infection.

The search for HIV-1 therapeutics has led to researchers looking to target the LEDGF/p75:integrase interaction using ALLINIs, especially in response to emerging resistance to integrase active site mutations[39,40]. LEDGF/p75 is a host chromatin-binding factor that directs HIV-1 complementary DNA into active gene regions[36–38]. The ALLINI BI-D was discovered as an inhibitor of HIV-1 integration at early-infection stages, and it has been shown that BI-D, like other ALLINIs, potently interferes with late-stage viral maturation to inhibit viral budding[44,45]. Interference with the LEDGF/p75:integrase interaction with ALLINIs was previously shown to increase transcription near CpG islands in DNA sequences that marked with DNA methylation, leading to repressed expression, and since LEDGF/p75 is a chromatin mark reading protein, the lack of LEDGF/p75 interacting with the HIV-1 preintegration complex leads to increased HIV-1 cDNA integration outside of active gene regions, which increases the presence of latent cells under ALLINI treatment[24,29,36–38,47]. The mechanism behind the observed decrease in the ability to reactivate latent provirus treated with ALLINIs has not been fully investigated, but as asRNA is shown to be involved in latency and asRNA was increased with BI-D treatment, we suspect believe that asRNA is also involved in LEDGF/p75 ablation and HIV-1 latency.

Production of HIV-1 asRNA has not previously been related to ALLINI treatment, specifically with BI-D, and while asRNA is proposed to play a role in inducing and maintaining latency, cells driven into latency using BI-D and the presence of asRNA has not been previously investigated[60,65,66]. To address this gap in the relationship, this study looked at asRNA production in the various subpopulations of cells

(actively infected, latently infected, not infected), to try to correlate presence of asRNA with different phenotypes. Treating cells with BI-D led to fewer actively infected cells and increased the proportion of cells that were labeled as uninfected determined by spectral flow cytometry. Surprisingly, the number of latent cells remained the same when cells were treated with BI-D. In contrast, other studies using a different dual-fluorescent HIV-1 reporter published an increase in the latent population when treated with another ALLINI, but did not report the changes in the uninfected population[47]. This could suggest that the phenomenon seen in this study is unique to BI-D, and that the latent cells seen in other studies where LEDGF/p75:integrase is ablated could be due to the cells not being productively infected at all, or the entire provirus being silenced. It is also possible that the differences in conclusions are due to differences in data analysis between the current study and the study from the Debyser group.

Due to time and scheduling constraints, the HI-Fate Tomato transduced cells were sorted 24 hours after transduction and BI-D treatment. All other experiments we performed analyzing asRNA production and subpopulation proportions were completed 48 hours following transduction and BI-D treatment. Previous and current members of the Sarafianos lab have conflicting results that suggest asRNA production peaks 24 hours or 48 hours after VLP transduction, so we thought that 24 hours after transduction would be sufficient. Early HIV-1 research found that HIV-1 short RNAs are present between 12- and 16-hours post-infection, and full length HIV-1 RNA is present around 24 hours post-infection[79]. That study did not follow RNA production after 24 hours, so it is unknown if those time points represent the peak of RNA production after infection. Another study looked at the timeline of HIV-1 protein production, and they determined that early HIV-1 proteins, such as Rev and Tat, are present in 50% of cells 14-30 hours after infection, and the remaining late HIV-1 proteins, such as Gag, are present an additional 1-5 hours following early protein detection[80]. This means that full protein expression is present in around 50% of cells 30-33 hours post-infection[80]. The data from our cell sorting experiments suggest that 48 hours following transduction allows for optimal RNA presence, as well as allows for enough time

for the reporter fluorescent proteins to be expressed. As the fluorescent proteins in the Hi-Fate Tomato reporter were placed in the *nef* gene region which is translated in later timepoints after HIV-1 infection, it is likely that there was not enough time given for the fluorescent proteins to be completely expressed. Compared to the spectral flow cytometry experiments performed 48 hours post-infection, there were far fewer cells that were sorted as actively infected, and a large increase in cells identified as untransduced despite more cells being analyzed overall (cell counts not shown). The differences in cell counts are most likely due to sorting the samples after only 24 hours post-infection. This conclusion is further supported by RNAScope images that were taken of cells that were fixed 24 hours after cell sorting, totaling 48 hours post-infection (Supplementary Figures 3 and 4). There is considerable sRNA production in cells that were sorted into the untransduced population when the cells were fixed 48 hours after transduction, suggesting that measuring sRNA and asRNA production after 24 hours does not accurately represent the optimal RNA levels in each subpopulation due to incorrect subpopulation assignments. Subsequent studies will sort the cells into subpopulations 48 hours post-VLP infection as this better represents the RNA levels expected, as well as allows for accurate sorting of the cells into their respective populations based on fluorescent protein expression. Definitive conclusions about the presence of asRNA in subpopulations of transduced cells or if asRNA levels change with BI-D treatment cannot be made with the current data.

We chose to also investigate the untransduced population of cells in both treatment conditions in the case that HIV-1 cDNA was integrated into a region that silenced both the HIV-1 promoter and the EF1-α promoter, thus indicating deep latency. This would suggest that other studies using dual reporters that only analyze single-fluorescent cells as the latent population are potentially missing a portion of the latent cells that would have been classified as uninfected. Although low cell counts make our results largely uninterpretable, it is interesting to see that there is a higher percentage of cells with asRNA in the untransduced subpopulation. If these data are consistent through future replicates, then the initial

hypothesis that asRNA is associated with deep latency is supported. The data from the early FACS experiments merits further investigation and the ability to modify the methods used currently.

If the HIV-1 cDNA integrated near a promoter that allowed for transcription from the 3' LTR of the HIV-1 genome, then we would expect an increase in asRNA production. The experiments in our study and future experiments will potentially reveal the asRNA production in different subpopulations of cells. Given that the percentage of latent cells remained the same between untreated and BI-D treated cells, we would expect the asRNA production to be increased in all subpopulations of BI-D treated cells, but to a lesser degree in the untransduced subpopulation. Future experiments will investigate this more thoroughly, as well as investigate the results of treating each subpopulation with a latency reversal agent to assess HIV-1 reactivation. Treating the sorted subpopulations of transduced cells with latency reversal agents will reveal if the double negative population (untransduced) contained provirus that is completely silenced due to integration in gene-sparse and inactive DNA regions. The asRNA levels in each of the subpopulations could be related to the ability to reactivate the cells as well as how the levels of asRNA are related in untreated cells and BI-D treated cells. The current cell sorting results provide a foundation for these future experiments.

While dual fluorescent reporters are versatile in HIV-1 latency research, these reporters for modeling HIV-1 infection have some limitations. For example, the constitutive mammalian promoter used for HIV-1 independent expression of a fluorophore can lead to variation in the expression of the fluorophore that is detected by flow cytometry, and weak expression of a constitutive promoter can lead to the underestimation of the number of latent cells in a population[78,81]. This underestimation of latent cells is the main reason for including double negative (untransduced) cells in the analyzed sorted populations. Also, certain fluorescent proteins can be more difficult to detect as they are not as bright as other fluorescent proteins, especially following fixation with paraformaldehyde as performed in the spectral flow cytometry experiments. Assessment of these limitations have led to the continuous

modifications of the dual reporters, especially the testing of combinations of constitutive promoters and various fluorescent proteins[67]. The modifications made to create the Hi-Fate Tomato dual reporter made by the Herschhorn lab were chosen to optimize detection of the fluorophores and enhance flow cytometry sensitivity. Also, it is important to note that although the Hi-Fate Tomato dual reporter produces asRNA, the functionality of the asRNA from this reporter cannot be validated. As asRNA has been shown to increase latency, and translation of the asRNA producing the antisense protein (ASP) could potentially play a role in inducing and maintaining latency, the insertion of fluorescent proteins in the asRNA region could disrupt function of the asRNA transcript and would certainly prevent ASP translation[60,61,64,65]. As the Hi-Fate tomato reporter responds similarly to BI-D treatment, both in terms of infection efficiency and asRNA production, we concluded that the asRNA in the Hi-Fate Tomato reporter likely retains its function in the induction and maintenance of latency in this context, which maintains the validity of using a dual-reporter to study asRNA.

This study used HEK293/17 cells as the model cell line for all experiments. HEK293/17 cells are a derivative of 293T cells that have increased transfection efficiency[71]. Early experiments investigating a live-cell imaging technique to image HIV-1 asRNA used 293/17 cells (unpublished). During the transition period between projects, cell type was not immediately brought into question as these cells were acceptable to use in preliminary experiments. Previous experiments with HIV-1 infection and testing with BI-D has been completed using HEK293 cells[73]. Also, the use of an adherent cell line allows for ease of imaging and execution of the modified RNAScope protocol without the need for additional steps. However, as it was necessary to treat the 293/17 cells as suspension cells following sorting, there is no longer an advantage for using an adherent cell line. Future experiments would use more physiologically relevant cells, such as Jurkat T cells or primary T cells that more accurately represent the cell type that HIV-1 would infect in a natural infection. These cells were not used for the present study as they are suspension cells, which introduces added difficulties when preparing the cells for RNA staining.

In conclusion, this study revealed that the Hi-Fate Tomato dual-fluorescent HIV-1 reporter produces detectable asRNA that can be analyzed to investigate asRNA production in response to BI-D treatment, and the involvement of both BI-D and asRNA in HIV-1 latency. The current results cannot support reliable conclusions about the asRNA production in the three subpopulations of Hi-Fate Tomato transduced cells, nor can an association between asRNA, BI-D, and HIV-1 latency can be made at this time. The early results showing an increased percentage of infected cells containing asRNA in the untransduced cell population is consistent with the rationale for this study and warrants further investigation. These studies provide the foundation for future experiments that will further analyze asRNA production in subpopulations of cells, followed by reactivation experiments. The implications of these studies will provide further insight into the mechanisms of HIV-1 latency, which can contribute to the development of an HIV-1 cure.

**Acknowledgements**

Thank you to the Emory University School of Medicine Department of Pediatrics, the Atlanta Pediatric Research Alliance, and the Center for AIDS Research at Emory University (P30AI050409) for supporting and funding for this research.

Thank you to the Emory University Pediatrics/Winship Flow Cytometry Core and the Core members for training me to use the Cytek Aurora Flow Cytometer, helping with data analysis, and for helping run the FACS samples.

The HIV-1 NL4-3 Gag-iGFP delEnv Non-infectious Molecular Clone (ARP-12455) was obtained through the NIH HIV Reagent Program, Division of AIDS, NIAID, NIH and contributed by Dr. Benjamin Chen.

Thank you to Dr. Alon Herschhorn for gifting the Sarafianos lab with the Hi-Fate Tomato HIV-1 dual fluorescent reporter.

Thank you to Dr. Alan Engleman for gifting us BI-D for these experiments.

**References**


1. U S Department of Health & Human Services. Global Statistics | HIV.gov. 1 https://www.hiv.gov/hiv-basics/overview/data-and-trends/global-statistics (2021).
2. Centers of Disease Control. Basic Statistics | HIV Basics | HIV/AIDS | CDC. 1 https://www.cdc.gov/hiv/basics/statistics.html (2021).
3. CDC. Ways HIV Can Be Transmitted | HIV Transmission | HIV Basics | HIV/AIDS | CDC. (2020).
4. Barré-Sinoussi, F. *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science (80-. ).* **220**, 868–871 (1983).
5. Division of HIV/AIDS Prevention. About HIV/AIDS | HIV Basics | HIV/AIDS | CDC. *Centers for Disease Control and Prevention* https://www.cdc.gov/hiv/basics/whatishiv.html (2021).
6. United States Food and Drug Administration. FDA-Approved HIV Medicines. *aidsinfo.nih.gov* 8–11 https://hivinfo.nih.gov/understanding-hiv/fact-sheets/fda-approved-hiv-medicines (2014).
7. Kelley, C. F. *et al.* Incomplete peripheral CD4+ cell count restoration in HIV-infected patients receiving long-term antiretroviral treatment. *Clin. Infect. Dis.* **48**, 787–794 (2009).
8. Chun, T. W. *et al.* In vivo fate of HIV-1-infected T cells: Quantitative analysis of the transition to stable latency. *Nat. Med.* **1**, 1284–1290 (1995).
9. Finzi, D. *et al.* Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science (80-. ).* **278**, 1295–1300 (1997).
10. Blankson, J. N., Persaud, D. & Siliciano, R. F. The challenge of viral reservoirs in hiv-1 infection. *Annual Review of Medicine* vol. 53 557–593 (2002).
11. Archin, N. M., Sung, J. M., Garrido, C., Soriano-Sarabia, N. & Margolis, D. M. Eradicating HIV-1 infection: Seeking to clear a persistent pathogen. *Nature Reviews Microbiology* vol. 12 750–764 (2014).
12. Sengupta, S. & Siliciano, R. F. Targeting the Latent Reservoir for HIV-1. *Immunity* vol. 48 872–895

(2018).

13. Yang, H. C. *et al.* Small-molecule screening using a human primary cell model of HIV latency identifies compounds that reverse latency without cellular activation. *J. Clin. Invest.* **119**, 3473–3486 (2009).

14. Lehrman, G. *et al.* Depletion of latent HIV-1 infection in vivo: A proof-of-concept study. *Lancet* **366**, 549–555 (2005).

15. Cillo, A. R. *et al.* Quantification of HIV-1 latency reversal in resting CD4+ T cells from patients on suppressive antiretroviral therapy. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 7078–7083 (2014).

16. Ho, Y. C. *et al.* Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell* **155**, 540 (2013).

17. Kessing, C. F. *et al.* In Vivo Suppression of HIV Rebound by Didehydro-Cortistatin A, a "Block-and-Lock" Strategy for HIV-1 Treatment. *Cell Rep.* **21**, 600–611 (2017).

18. Moranguinho, I. & Valente, S. T. Block-and-lock: New horizons for a cure for hiv-1. *Viruses* vol. 12 (2020).

19. Elsheikh, M. M., Tang, Y., Li, D. & Jiang, G. Deep latency: A new insight into a functional HIV cure. *EBioMedicine* vol. 45 624–629 (2019).

20. Griffiths, D. J. Endogenous retroviruses in the human genome sequence. *Genome Biology* vol. 2 (2001).

21. Rowe, H. M. & Trono, D. Dynamic control of endogenous retroviruses during development. *Virology* vol. 411 273–287 (2011).

22. Ako-Adjei, D. *et al.* HIV-1, Human Interaction database: Current status and new features. *Nucleic Acids Res.* **43**, D566–D570 (2015).

23. HIV-1 Interactions Database. *Natl. Libr. Med.*

24. Ge, H., Si, Y. & Roeder, R. G. Isolation of cDNAs encoding novel transcription coactivators p52 and

p75 reveals an alternate regulatory mechanism of transcriptional activation. *EMBO J.* **17**, 6723–6729 (1998).

25. Singh, D. P., Kimura, A., Chylack, L. T. & Shinohara, T. Lens epithelium-derived growth factor (LEDGF/p75) and p52 are derived from a single gene by alternative splicing. *Gene* **242**, 265–273 (2000).

26. Singh, D. P. *et al.* Lens epithelium-derived growth factor: Effects on growth and survival of lens epithelial cells, keratinocytes, and fibroblasts. *Biochem. Biophys. Res. Commun.* **267**, 373–381 (2000).

27. Shinohara, T., Singh, D. P. & Fatma, N. LEDGF, a survival factor, activates stress-related genes. *Progress in Retinal and Eye Research* vol. 21 341–358 (2002).

28. Daugaard, M. *et al.* LEDGF (p75) promotes DNA-end resection and homologous recombination. *Nat. Struct. Mol. Biol.* **19**, 803–810 (2012).

29. Singh, P. K. *et al.* LEDGF/p75 interacts with mRNA splicing factors and targets HIV-1 integration to highly spliced genes. *Genes Dev.* **29**, 2287–2297 (2015).

30. Eidahl, J. O. *et al.* Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. *Nucleic Acids Res.* **41**, 3924–3936 (2013).

31. Van Nuland, R. *et al.* Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain. *Epigenetics Chromatin* **6**, 12 (2013).

32. Cherepanov, P. *et al.* HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J. Biol. Chem.* **278**, 372–381 (2003).

33. Busschots, K. *et al.* Identification of the LEDGF/p75 Binding Site in HIV-1 Integrase. *J. Mol. Biol.* **365**, 1480–1492 (2007).

34. Vanegas, M. *et al.* Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering. *J. Cell Sci.* **118**, 1733–1743 (2005).

35. Lapaillerie, D. *et al.* Modulation of the intrinsic chromatin binding property of HIV-1 integrase by LEDGF/p75. *Nucleic Acids Res.* **49**, 11241–11256 (2021).

36. Shun, M. C. *et al.* LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev.* **21**, 1767–1778 (2007).

37. Schröder, A. R. W. *et al.* HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**, 521–529 (2002).

38. Engelman, A. N. & Singh, P. K. Cellular and molecular mechanisms of HIV-1 integration targeting. *Cellular and Molecular Life Sciences* vol. 75 2491–2507 (2018).

39. Christ, F. & Debyser, Z. The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy. *Virology* vol. 435 102–109 (2013).

40. Malet, I. *et al.* Mutations associated with failure of raltegravir treatment affect integrase sensitivity to the inhibitor in vitro. *Antimicrob. Agents Chemother.* **52**, 1351–1358 (2008).

41. Jurado, K. A. & Engelman, A. Multimodal mechanism of action of allosteric HIV-1 integrase inhibitors. *Expert reviews in molecular medicine* vol. 15 e14 (2013).

42. Christ, F. *et al.* Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat. Chem. Biol.* **6**, 442–448 (2010).

43. Kessl, J. J. *et al.* Multimode, cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. *J. Biol. Chem.* **287**, 16801–16811 (2012).

44. Christ, F. *et al.* Small-Molecule Inhibitors of the LEDGF/p75 Binding Site of Integrase Block HIV Replication and Modulate Integrase Multimerization. *Antimicrob. Agents Chemother.* **56**, 4365 (2012).

45. Le Rouzic, E. *et al.* Dual inhibition of HIV-1 replication by integrase-LEDGF allosteric inhibitors is predominant at the post-integration stage. *Retrovirology* **10**, 144 (2013).

46. Vansant, G. *et al.* Impact of LEDGIN treatment during virus production on residual HIV-1

transcription. *Retrovirology* **16**, 1–17 (2019).

47. Vranckx, L. S. *et al.* LEDGIN-mediated Inhibition of Integrase-LEDGF/p75 Interaction Reduces Reactivation of Residual Latent HIV. *EBioMedicine* **8**, 248–264 (2016).

48. Ciuffi, A. *et al.* A role for LEDGF/p75 in targeting HIV DNA integration. *Nat. Med.* **11**, 1287–1289 (2005).

49. Seitz, R. Human Immunodeficiency Virus (HIV). *Transfus. Med. Hemotherapy* **43**, 203–222 (2016).

50. Roebuck, K. A. & Saifuddin, M. Regulation of HIV-1 transcription. *Gene Expression* vol. 8 67–84 (1999).

51. Berkhout, B., Silverman, R. H. & Jeang, K. T. Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell* **59**, 273–282 (1989).

52. Purcell, D. F. & Martin, M. A. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J. Virol.* **67**, 6365–6378 (1993).

53. Sundquist, W. I. & Kräusslich, H. G. HIV-1 assembly, budding, and maturation. *Cold Spring Harbor Perspectives in Medicine* vol. 2 (2012).

54. Miller, R. H. Human immunodeficiency virus may encode a novel proten on the genomic DNA plus strand. *Science (80-. ).* **239**, 1420–1422 (1988).

55. Landry, S. *et al.* Detection, characterization and regulation of antisense transcripts in HIV-1. *Retrovirology* **4**, 71 (2007).

56. Schopman, N. C. T. *et al.* Deep sequencing of virus-infected cells reveals HIV-encoded small RNAs. *Nucleic Acids Res.* **40**, 414 (2012).

57. Ludwig, L. B. *et al.* Human Immunodeficiency Virus-Type 1 LTR DNA contains an intrinsic gene producing antisense RNA and protein products. *Retrovirology* **3**, 80 (2006).

58. Bentley, K., Deacon, N., Sonza, S., Zeichner, S. & Churchill, M. Mutational analysis of the HIV-1 LTR as a promoter of negative sense transcription. *Arch. Virol.* **149**, 2277–2294 (2004).

59. Michael,', N. L. *et al.* Negative-strand RNA transcripts are produced in human immunodeficiency virus type 1-infected cells and patients by a novel promoter downregulated by Tat. *J. Virol.* **68**, 979 (1994).

60. Kobayashi-Ishihara, M. *et al.* HIV-1-encoded antisense RNA suppresses viral replication for a prolonged period. *Retrovirology* **9**, 38 (2012).

61. Gholizadeh, Z., Iqbal, M. S., Li, R. & Romerio, F. The HIV-1 Antisense Gene ASP: The New Kid on the Block. *Vaccines* **9**, (2021).

62. Affram, Y. *et al.* The HIV-1 Antisense Protein ASP Is a Transmembrane Protein of the Cell Surface and an Integral Protein of the Viral Envelope. *J. Virol.* **93**, (2019).

63. Vanhée-Brossollet, C. *et al.* A natural antisense RNA derived from the HIV-1 env gene encodes a protein which is recognized by circulating antibodies of HIV+ individuals. *Virology* **206**, 196–202 (1995).

64. Zapata, J. C. *et al.* The Human Immunodeficiency Virus 1 ASP RNA promotes viral latency by recruiting the Polycomb Repressor Complex 2 and promoting nucleosome assembly. *Virology* **506**, 34–44 (2017).

65. Saayman, S. *et al.* An HIV-Encoded antisense long noncoding rna epigenetically regulates viral transcription. *Mol. Ther.* **22**, 1164–1175 (2014).

66. Kobayashi-Ishihara, M. *et al.* HIV LTR-driven antisense RNA by itself has regulatory function and may curtail virus reactivation from latency. *Front. Microbiol.* **9**, 1066 (2018).

67. Ratnapriya, S. *et al.* Intra- and extra-cellular environments contribute to the fate of HIV-1 infection. *Cell Rep.* **36**, 109622 (2021).

68. Battivelli, E. *et al.* Distinct chromatin functional states correlate with HIV latency reactivation in infected primary CD4+ T cells. *Elife* **7**, (2018).

69. Hashemi, F. B. *et al.* HIV Provirus Stably Reproduces Parental Latent and Induced Transcription

Phenotypes Regardless of the Chromosomal Integration Site. *J. Virol.* **90**, 5302 (2016).

70. Chavez, L., Calvanese, V. & Verdin, E. HIV Latency Is Established Directly and Early in Both Resting and Activated Primary CD4 T Cells. *PLoS Pathog.* **11**, (2015).

71. Pear, W. S., Nolan, G. P., Scott, M. L. & Baltimore, D. Production of high-titer helper-free retroviruses by transient transfection. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 8392–8396 (1993).

72. Van Bel, N. *et al.* The Allosteric HIV-1 Integrase Inhibitor BI-D Affects Virion Maturation but Does Not Influence Packaging of a Functional RNA Genome. *PLoS One* **9**, (2014).

73. Feng, L. *et al.* The Competitive Interplay between Allosteric HIV-1 Integrase Inhibitor BI/D and LEDGF/p75 during the Early Stage of HIV-1 Replication Adversely Affects Inhibitor Potency. *ACS Chem. Biol.* **11**, 1313–1321 (2016).

74. Wang, F. *et al.* RNAscope: A Novel in Situ RNA Analysis Platform for Formalin-Fixed, Paraffin-Embedded Tissues. *J. Mol. Diagn.* **14**, 22 (2012).

75. Shah, R. *et al.* Single-cell multiplexed fluorescence imaging to visualize viral nucleic acids and proteins and monitor hiv, htlv, hbv, hcv, zika virus, and influenza infection. *J. Vis. Exp.* **2020**, (2020).

76. Hübner, W. *et al.* Sequence of Human Immunodeficiency Virus Type 1 (HIV-1) Gag Localization and Oligomerization Monitored with Live Confocal Imaging of a Replication-Competent, Fluorescently Tagged HIV-1. *J. Virol.* **81**, 12596–12607 (2007).

77. Zhang, H. *et al.* Novel Single-Cell-Level Phenotypic Assay for Residual Drug Susceptibility and Reduced Replication Capacity of Drug-Resistant Human Immunodeficiency Virus Type 1. *J. Virol.* **78**, 1718–1729 (2004).

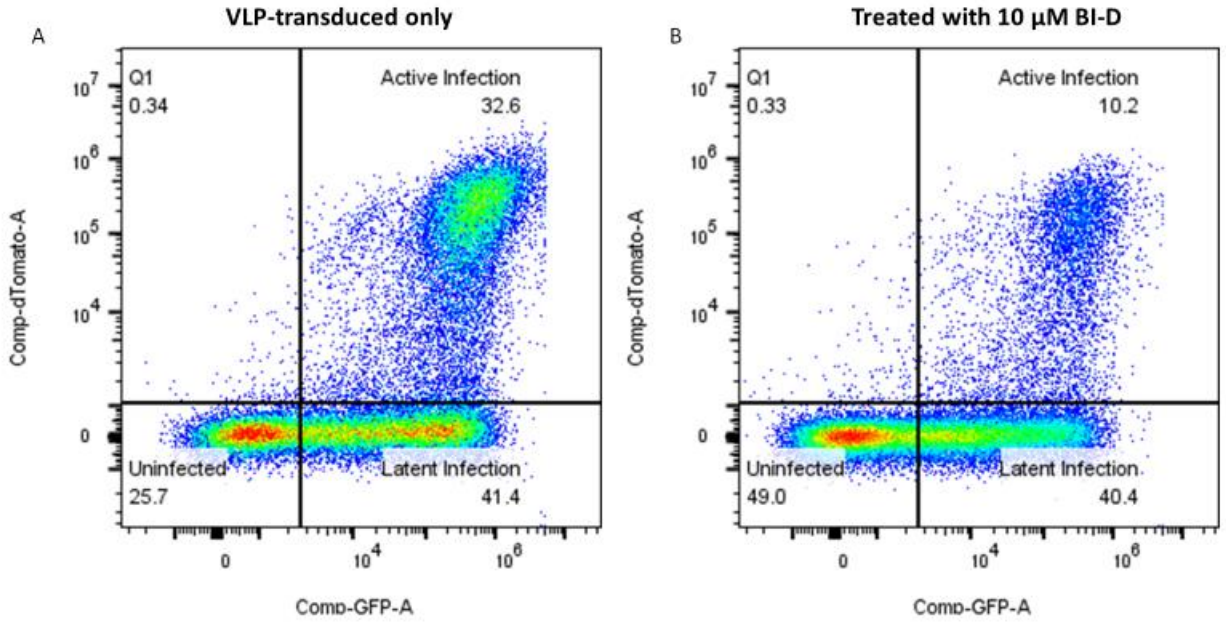
78. Dahabieh, M. S., Ooms, M., Simon, V. & Sadowski, I. A Doubly Fluorescent HIV-1 Reporter Shows that the Majority of Integrated HIV-1 Is Latent Shortly after Infection. *J. Virol.* **87**, 4716–4727 (2013).

79. Kim, S. Y., Byrn, R., Groopman, J. & Baltimore, D. Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. *J. Virol.* **63**, 3708–3713 (1989).

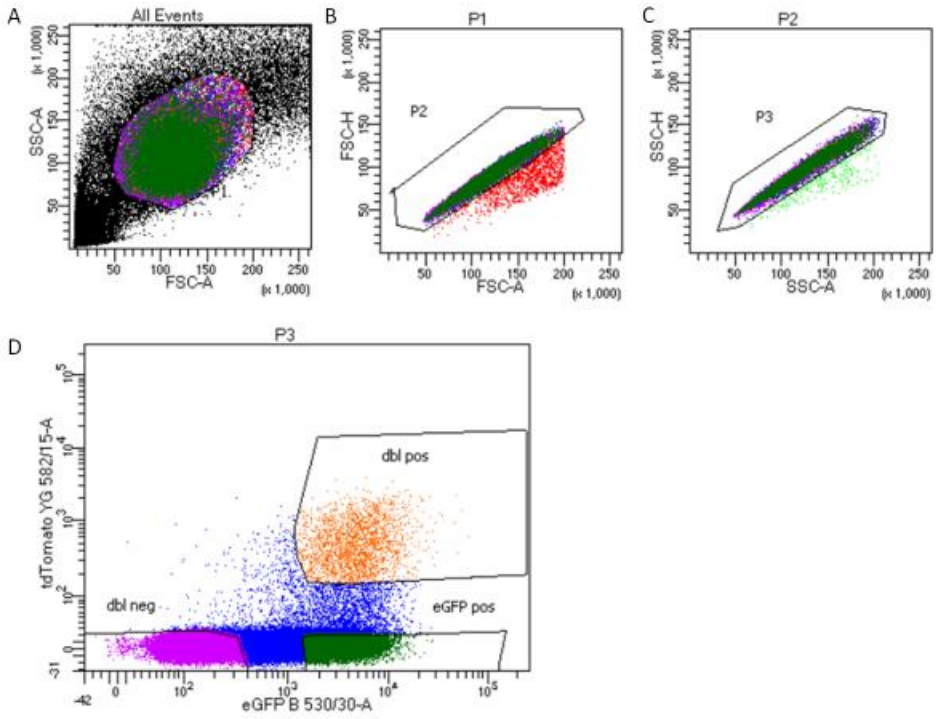
80. Holmes, M., Zhang, F. & Bieniasz, P. D. Single-Cell and Single-Cycle Analysis of HIV-1 Replication. *PLoS Pathog.* **11**, 1–23 (2015).

81. Kim, Y., Cameron, P. U., Lewin, S. R. & Anderson, J. L. Limitations of dual-fluorescent HIV reporter viruses in a model of pre-activation latency. *J. Int. AIDS Soc.* **22**, (2019).

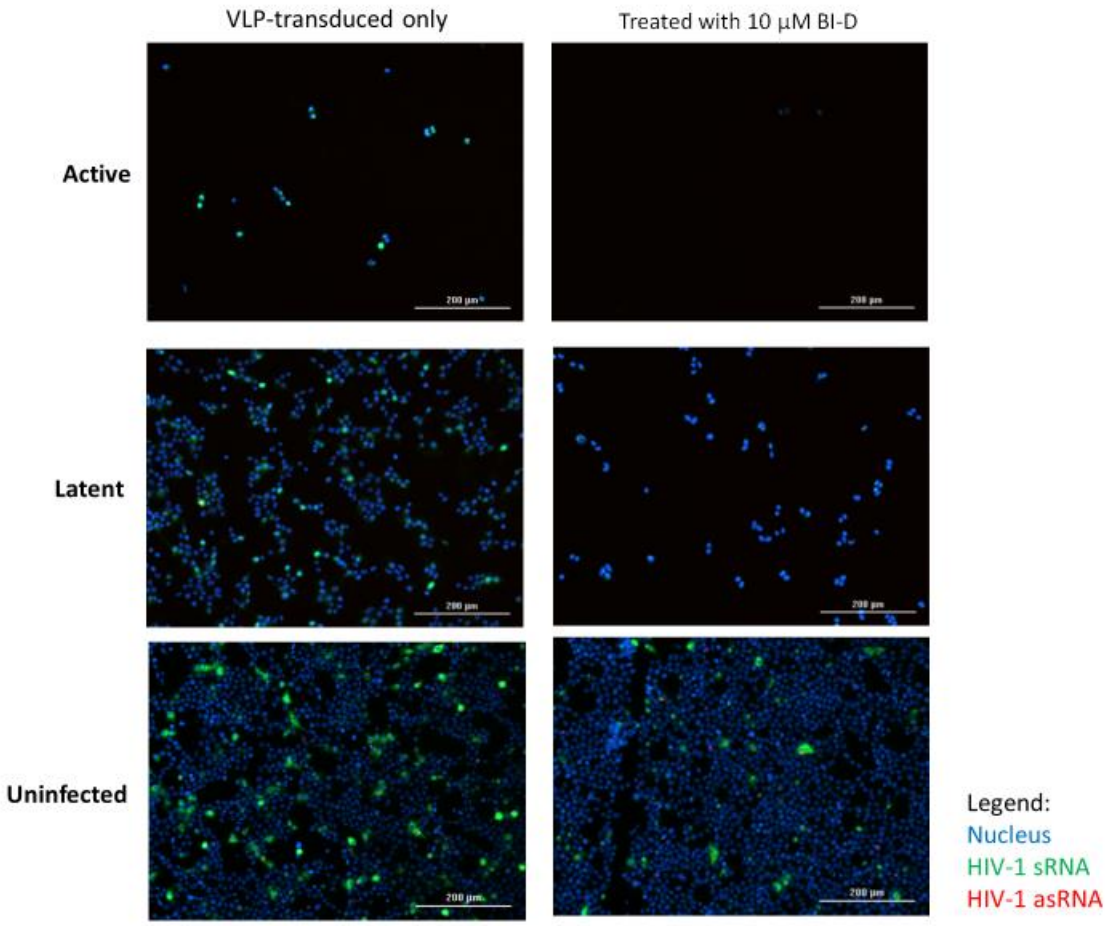
## Appendix – Supplementary Figures



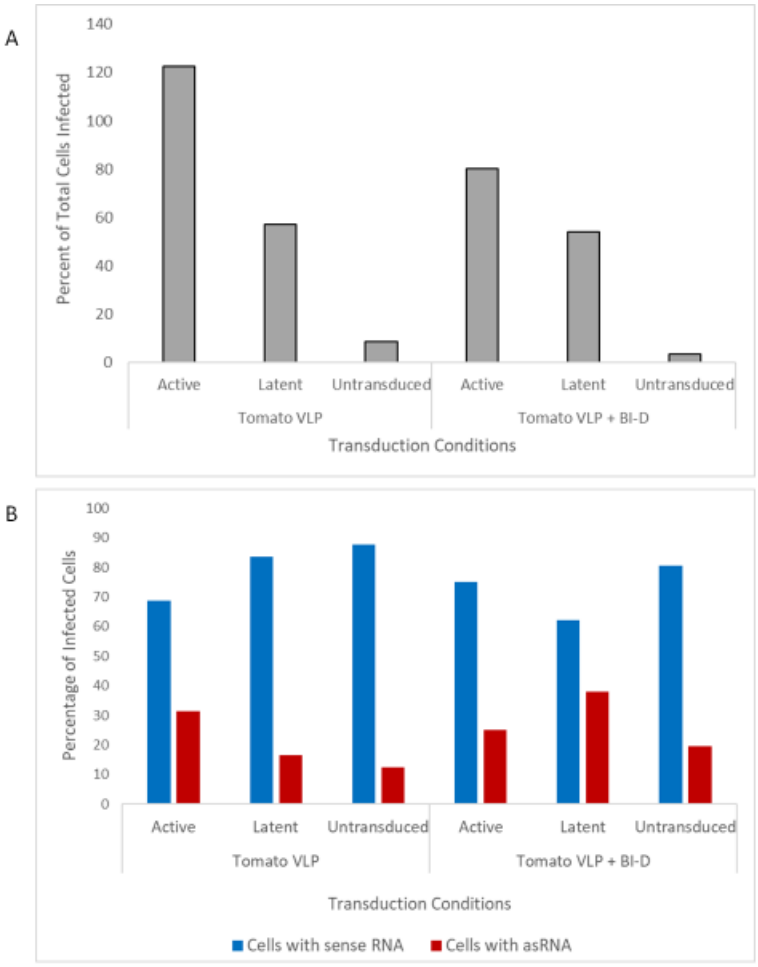

**Supplementary Figure 1.** *Quadrant gates were used to calculate the percentage of each subpopulation of cells transduced with the Hi-Fate Tomato VLP and treated with BI-D.* HEK293/17 cells were transduced with the Hi-Fate Tomato VLP (A) one sample was treated with 10 µM BI-D (B), and then analyzed for fluorescent protein expression using the Cytek Aurora spectral flow cytometer. Pseudocolor dot plots represent the sample distribution from a single representative replicate. Gates were obtained after isolating live cells, then singlets.




**Supplementary Figure 2.** *Gating strategy for fluorescence activated cell sorting experiment.* HEK293/17 cells were transduced with the Hi-Fate Tomato VLP, one sample was treated with 10 µM BI-D, and then sorted into 3 populations using the BD FACSAria II SORP cell sorter. A) Isolating live cells to generate the P1 gate. B) Cells from the P1 gate, then gated for single cells to generate the P2 gate. C) cells from the P2 gate plotted to isolate single cells and generate the P3 gate. D) Cells from the P3 gate plotted to determine the populations of cells. The gates in panel D represent the criteria that were used to identify and sort cells into each population. Dot plots represent the sample distribution from a single representative replicate. The same gates were used for both samples that were sorted. Dbl pos – double positive. Dbl neg – double negative. eGFP pos- eGFP single positive




**Supplementary Figure 3.** *Sorted populations of Hi-Fate Tomato transduced cells show sRNA and asRNA production when fixed 24 hours after seeding.* HEK293/17 cells were transduced with Hi-Fate Tomato VLPs and sorted into active, latent, and uninfected cell populations 24 hours after transduction and then fixed 24 hours after seeding on the coverslips. Cells were prepared and stained according to the modified RNAScope procedure. Sparse cells were the result of decreased cell counts in the respective populations after sorting.




**Supplementary Figure 4.** *Fixing and staining samples 24 hours after seeding results in inaccurate results.* Following VLP transduction, cell sorting, RNA staining, and imaging on the Cytation5 microscope, cells with sRNA and asRNA were counted using the BioTek software. Cell counts were obtained by identifying cell nuclei, then counting an individual cell as containing sRNA or asRNA if the sum amount of green or red fluorescence was above a 500,000- or 300,000-unit threshold, respectively, within the cell nucleus. Cells labeled infected contained sRNA or asRNA. A) Percentage of the total number of cells counted, infected and uninfected, that contained sRNA or asRNA in each sorted population when untreated or treated with BI-D. B) Percentage of cells with sRNA and asRNA out of the cells that were infected in each sorted population from two treatment groups, with and without BI-D. n =1.