ENGINEERING YEAST TO EVALUATE HUMAN PROTEINS INVOLVED IN SELECTIVE RNA PACKAGING DURING HIV PARTICLE PRODUCTION

A Thesis
presented to
the Faculty of California Polytechnic State University,
San Luis Obispo

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Biological Sciences

by Ryan Matthew Bitter
October 2018

COMMITTEE MEMBERSHIP

TITLE: Engineering yeast to evaluate human proteins involved in selective RNA packaging during HIV particle production

AUTHOR: Ryan Matthew Bitter

DATE SUBMITTED: October 2018

COMMITTEE CHAIR: Michael Black, Ph.D.
Professor, Biological Sciences

COMMITTEE MEMBER: Nathaniel Martinez, M.D., Ph.D.
Assistant Professor, Biological Sciences

COMMITTEE MEMBER: Ken Hillers, Ph.D.
Professor and Chair, Biological Sciences

ABSTRACT

Engineering yeast to evaluate human proteins involved in selective RNA packaging during HIV particle production

Ryan Matthew Bitter

Despite recent advances in antiretroviral therapy, nearly 37 million people continue to live with human immunodeficiency virus (HIV). Basic and applied research on the assembly of HIV could be enhanced by using a genetically tractable organism, such as yeast, rather than mammalian cells. While previous studies showed that expression of the HIV Gag polyprotein in *Saccharomyces cerevisiae* spheroplasts resulted in the production of virus-like particles (VLPs), many questions regarding the utility of yeast in HIV assembly remain uninvestigated. Here, we report use of *S. cerevisiae* for both the production of VLPs with selectively packaged RNA and to evaluate the human Y-box-binding protein 1 (YB-1) in selective RNA packaging into VLPs. Our data reveal: (1) When co-expressed alongside HIV-1 Gag, an RNA mammalian expression cassette is selectively encapsidated and released in VLPs produced from spheroplasts; (2) Inclusion of the 5'UTR-5'Gag RNA upstream of the mammalian expression cassette greatly increased the selectivity to which non-viral RNA was packaged into VLPs; and (3) heterologous expression of the human YB-1 protein in *S. cerevisiae* did not facilitate the selective packaging of viral RNA into VLPs, likely due to inability to bind upstream elements in the HIV-1 viral RNA. Overall, this research provides a key first step in the use of yeast for the production of viral vectors used in gene therapy, and lays a foundation for further experiments investigating the role of YB-1 and other host proteins in selective RNA packaging.

ACKNOWLEDGMENTS

I have been very fortunate to have the positive influence of many people whom I acknowledge for the completion of this thesis. First and foremost, I would like to thank my committee chair and mentor, Dr. Michael Black. I will always look up to him as an example of a great experimentalist, who is kind and considerate with his students, and has a passion for science that is contagious. He has taught me *how* to be a scientist, and has equipped me with the tools I now take as I continue on my journey in scientific research. I would also like to extend gratitude do my committee members, Dr. Nathaniel Martinez and Dr. Ken Hillers. Dr. Martinez provided expertise and guidance in the area of HIV assembly, and always met my questions or requests with enthusiasm and generosity. Dr. Hillers always showed a palpable interest in my success, and offered me great advice throughout the program. I would also like to thank my fellow M.S. graduate students in the Black lab, Kara Jew and Jared Hollows, for meaningful conversations related (and, at times, not related) to our projects, and for sharing a very small lab space with me. I would like to thank Maria Tafoya and Jorge Del Pozo for technical assistance as part of the NIH Bridges to Baccalaureate program. Lastly, I am forever indebted to my parents, Grant and Lynn, who always show unrelenting support and love for me.

Funding that supported this research came from the Center for Applications in Biotechnology and the Biological Sciences Department at Cal Poly, San Luis Obispo. Thank you to all whose tireless efforts made this funding possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Background

Human immunodeficiency virus (HIV) is estimated to infect nearly 37 million people worldwide (UNAIDS, 2016). In addition to an increased risk of cardiovascular disease, bone disease, and cancer, individuals with untreated HIV-infections suffer from CD4 T cell loss, leading to acquired immune deficiency syndrome (AIDS) (Deeks et al., 2015). Victims of AIDS are more prone to opportunistic infections and various malignancies, and over 1 million AIDS-related deaths occur annually (UNAIDS, 2016).

There have been great advances in the efficacy of antiretroviral drugs in recent decades, resulting in treatments that are effective at inhibiting HIV replication and transmission (Deeks et al., 2015). However, 19 million HIV-positive people continue to live without treatment due to associated costs and lack of availability (UNAIDS, 2016). Furthermore, because no drugs eradicate the virus, those who receive treatment are administered it for their entire lives; This prolonged use of antiretroviral treatment can result in drug toxicity and the potential for viral resistance (Freed, 2015).

1.2 Classification

HIV-1 is an enveloped, spherical (80-120 nm diameter), single-stranded positive sense RNA virus with a double-stranded DNA intermediate. The 9.2 kb ssRNA linear genome is packaged as a dimer into virions. Characterized by its long incubation period in humans, HIV-1

is a member of the Lentivirus genus, within the Retroviridae family. Based upon genome type and method of replication, it is in Group VI of the Baltimore classification.

1.3 Viral proteins and their function

The main structural protein of all retroviruses is the Gag polyprotein. HIV-1 Gag is a 55 kDa protein translated on cytoplasmic soluble polysomes (D'Agostino et al., 1992) from unspliced viral mRNA in the late phase of replication. There are six domains in the Gag precursor protein. Listed from N- to C-terminus, they are: the matrix (MA), capsid (CA), spacer peptide 1 (SP1), nucleocapsid (NC), spacer peptide 2 (SP2), and p6 domain (reviewed in Freed, E.O., 2015). MA is involved in Gag membrane binding and envelope glycoprotein incorporation; CA is involved in Gag multimerization and assembly and forms the conical core of virus particles; NC is involved in binding genomic RNA to direct selective packaging into assembling virions; p6 is required for recruitment of the host endosomal sorting complex required for transport (ESCRT) machinery, which drives membrane scission for the release of virions from infected cells. After translation, an N-terminal glycine on MA becomes myristoylated, a process essential for host membrane binding at the sites of virion assembly (Bryant and Ratner, 1990).

Heterologous expression of the Gag polyprotein is sufficient to mediate assembly and release of virus-like particles from a variety of cell types, and this phenomenon is frequently exploited in experimental systems. However, replication of HIV-1 in naturally-infected human

cells requires a greater subset of viral proteins in order to produce infectious virions. A ribosomal frameshift event during Gag translation allows bypass of the stop codon at the end of Gag and readthrough to the protease, integrase, and reverse transcriptase proteins, thereby creating a 160 kDa GagPol fusion protein. The viral protease cleaves the domains of Gag and GagPol into individual proteins, triggering virion maturation. Following infection, reverse transcriptase converts single-stranded RNA into double-stranded DNA, referred to as proviral DNA, in the cytoplasm. Integrase is required to insert proviral DNA into the host genome. The viral envelope (Env) is a single pass, type 2 integral membrane protein translated from singly spliced mRNA that undergoes glycosylation as it is trafficked through the secretory pathway, hence it is a glycoprotein (commonly referred to as gp160). A host factor cleaves gp160 near the C-terminus, resulting in gp41 and gp120. Gp120 is a highly variable surface glycoprotein which interacts with the viral receptor CD4 (Landau et al., 1988) to mediate attachment to target cells. Gp41 is a transmembrane protein that contains hydrophobic fusion peptides allowing fusion with the host cell membrane. The complex interaction between the viral structural proteins and the cellular factors they co-opt to mediate particle formation is called assembly.

1.4 HIV-1 Assembly

The late phase of HIV replication, which encompasses viral assembly, begins when viral mRNA is transcribed from proviral DNA in the nucleus. Here, unspliced and incompletely spliced viral RNA associates with the viral Rev protein, and is exported to the cytoplasm. In the

cytoplasm, Rev facilitates the localization of *gag* mRNA on polysomes, allowing efficient translation of the Gag precursor polyprotein (D'Agostino et al., 1992). After translation, the glycine residue at position two of the N-terminus of the MA domain is covalently modified by the addition of a 14-carbon fatty acid called a myristate (Gottlinger et al., 1989). Both the N-terminal myristate and a stretch of basic residues in the MA domain are required for membrane binding of Gag (Zhou et al., 1994). Interestingly, the myristoylated Gag precursor polyprotein binds with high affinity to membranes, whereas myristoylated MA domain alone does not. This observation led to the hypothesis of a "myristoyl switch", in which myristate exposure is only triggered under particular circumstances (Zhou and Resh, 1996). Tang et al. (2004) found that the myristate group is sequestered when MA is in the monomer state, and exposed when MA is in the trimer state. In other words, the exposure of the N-terminal myristate of MA is dependent upon the multimerization state of Gag. As Gag polyproteins are brought into close proximity during assembly, such as when they bind to viral RNA, MA trimerization occurs, followed by exposure of the myristate group which allows for membrane anchoring (reviewed in Resh, 2004). At the sites of assembly, Gag molecules are arranged with their N-termini in contact with the membrane and their C-termini oriented towards the center of the immature Gag shell at the plasma membrane.

Various membrane-bound subcellular compartments within the cell are candidate binding sites for the N-myristoylated MA domain of Gag. For instance, in epithelial-like 293T cells, Gag was shown to localize in late endosomes (Figure 1B) (Grigorov et al. 2006). In HeLa cells, Gag

can assemble at both the plasma membrane and at late endosomal membranes/ multivesicular bodies (LE/MVB) (Sherer et al., 2003). Disruption of AP-3, a clathrin-associated heterotetrameric adaptor protein complex, negatively affects HIV-1 release from HeLa cells, indicating that LE/MVB may play a role in HIV assembly and release (Dong et al., 2005). In human macrophages, Gag was shown to be targeted to multivesicular bodies and, despite this, subsequent release of virions from these cells occurred with high efficiency (Ono and Freed, 2004). It has been proposed that the assembly and subsequent budding of Gag into intracellular compartments in macrophages allow virions to exit the cell via microchannels (Figure 1C) (Balasubramaniam and Freed, 2011). In dendritic cells, Gag assembly occurs at points of cell-to-cell contact, called virological synapses, where virions are then accessible to filopodia of CD4+ T cells (Figure 1D) (Balasubramaniam and Freed, 2011). Lastly, some researchers contend that assembly at the plasma membrane occurs in uropod structures (Figure 1E). This localization of Gag in uropod structures has been observed in polarized T cells, where Gag co-localized with multiple uropod markers, including CD43 (Llewellyn et al., 2010).

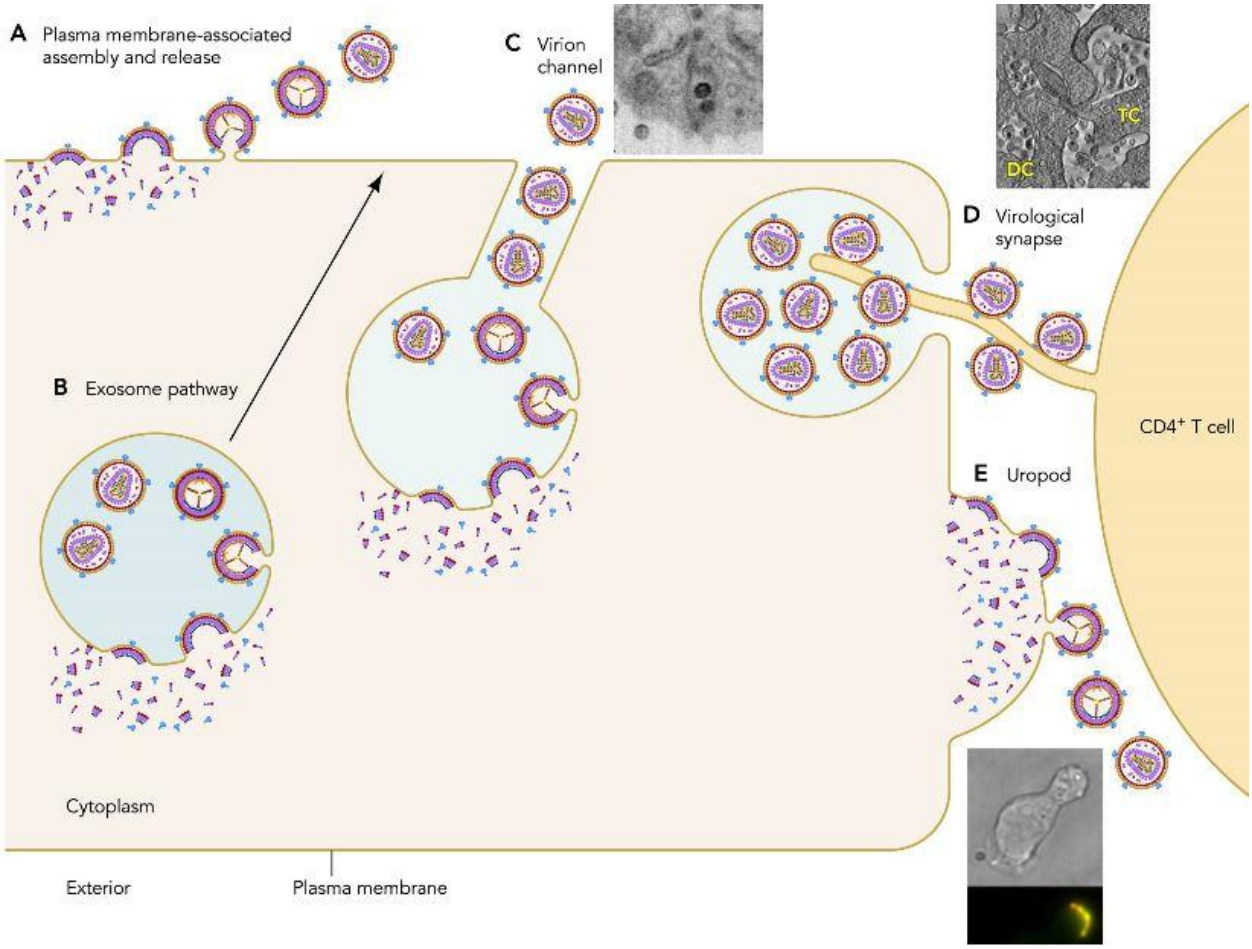


Figure 1: HIV assembly at different membrane interfaces (Balasubramaniam and Freed, 2011)

The most widely-accepted site of HIV-1 Gag assembly is the plasma membrane (Figure 2). Assembly at the plasma membrane allows efficient virion release, either extracellularly or via direct transfer to neighboring cells at points of cell-to-cell contact, called virological synapses (Balasubramaniam and Freed, 2011). By mutating endosome motility and preventing endosomal uptake, Jouvenet at al. (2006) found that Gag specifically targets the PM independent of an endosomal intermediate, and that this targeting to the PM is essential for the production of

virions. This refuted the models of assembly presented earlier, and corroborated a different study that used fluorescent imaging to determine Gag is localized at the PM shortly after synthesis (Rudner et al., 2005). Gag polyproteins are believed to arrive at the PM directly from the cytosol (Ivanchenko et al., 2009) after being trafficked along microtubule filaments via the kinesin KIF4 motor protein (Martinez et al., 2008). Once at the PM, N-myristoylated Gag anchors in the inner leaflet via interactions between MA and the phospholipid phosphatidylinositol-(4,5)-bisphosphate (Ono et al. 2004). This interaction notably occurs in lipid raft microdomains (Ono et al. 2001), which explains high levels of cholesterol and sphingolipids in the viral membrane after budding and release (Brugger et al. 2006). The negatively charged inositol headgroup of PtdIns(4,5)P_2 interacts with MA, promoting multimerization of the CA domain of Gag as the immature Gag lattice forms (Shkriabai et al., 2006). Although continuous, the immature Gag lattice contains gaps to accommodate curvature of the icosahedral virion (Briggs et al., 2009). A fully-assembled HIV-1 virion is composed of ~1,500-3000 Gag proteins with about a 20:1 ratio of Gag to GagPol, and spans a diameter of ~120 nm (Freed, E.O. 2015; Lingappa et al. 2014).

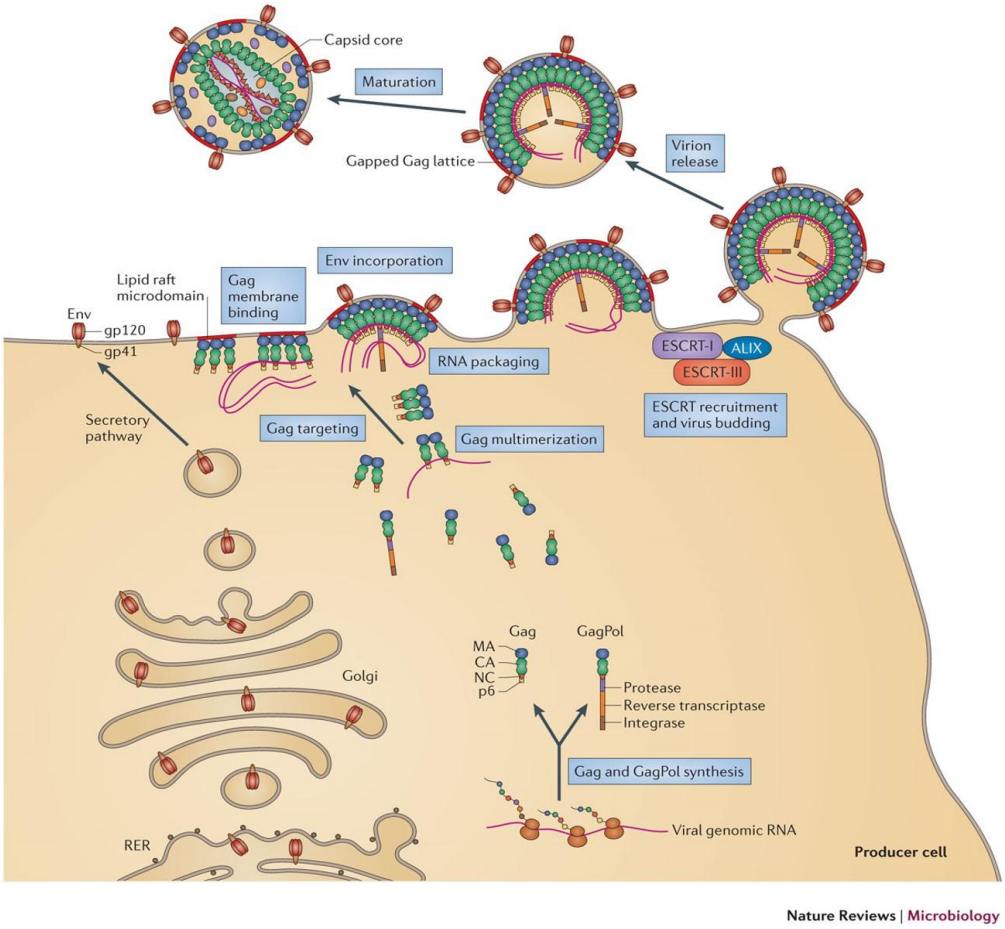


Figure 2: HIV assembly at the plasma membrane, and subsequent release of virions (Freed, 2015)

In contrast to our understanding of Gag assembly, little is known about Env incorporation into virions. Proposed mechanisms for Env incorporation include either simultaneous targeting of Gag and Env to lipid rafts in the plasma membrane, or direct recruitment of Env by Gag (Freed, E.O. 2015). A small deletion near the gp41 cytoplasmic tail disrupts Env incorporation into virions, while the process is rescued by a single amino acid change in the MA domain of Gag, indicating Gag-Env interactions mediate envelope incorporation (Murakami Freed, 2000).

After assembly at the plasma membrane, budding and release of the virion is mediated by the endosomal sorting complex required for transport (ESCRT). The ESCRT machinery is co-opted by the p6 domain of Gag; Specifically, the PTAP motif on the p6 domain of Gag binds the tumor susceptibility gene 101 (TSG101) subunit of ESCRT-I, and the YPXL motif of p6 binds an ESCRT-associated factor called Alix. As budding progresses, membrane scission is driven by ESCRT-III and Vps4 (Freed, E.O. 2015). This results in the release of a virion, allowing subsequent cleavage by the viral protease to trigger maturation.

While most of our understanding of HIV-1 assembly has been achieved using mammalian cells, Gag assembly can be reconstituted *in vitro* using cell-free systems. Campbell and Rein (1999) showed that Gag assembly *in vitro* produces 25-30 nm spherical particles, which is about 5-fold smaller than Gag virus-like particles (VLPs) assembled *in vivo*. Using non-viral RNAs and short oligonucleotide DNAs, the authors also showed that nucleic acids were required for the assembly of Gag with a deleted p6 domain. Most recently, a cell-free system has been used to recapitulate selective packaging of HIV-1 RNA by Gag using giant unilamellar vesicles, purified N-myristoylated Gag protein, and viral RNA (Carlson et al., 2016). These authors found that the 5' untranslated region (5'UTR) RNA of HIV-1 could be selectively packaged into Gag VLPs in the context of excess competitor RNA, while mutations in the capsid C-terminal domain abrogated RNA selectivity. Furthermore, Carlson et al. (2016) showed that excess tRNA prevented Gag membrane binding unless Gag had already bound viral RNA, indicating that selective RNA packaging facilitates Gag assembly. As the field progresses, the

ability to reconstitute HIV-1 assembly in cell-free systems will likely continue to be used to verify or refute whether particular components are required at different stages in assembly.

1.5 Host factors involved in HIV assembly

While many descriptions of HIV-1 assembly focus solely on Gag components alone, host factors co-opted by the virus remain integral in assembly. The first host protein involved in the late phase of HIV replication is Crm1, a cellular nuclear export receptor that, together with Rev, exports viral RNA from the nucleus to the cytoplasm (Yi et al., 2002). Multiple cellular factors have been implicated in trafficking of viral proteins to the assembly site at the plasma membrane. Kinesin KIF4 is a motor protein involved in Gag trafficking on microtubule filaments towards the plasma membrane (Martinez et al., 2008). Interestingly, the HIV-1 MA domain of Gag binds to KIF4, which induces KIF4 to interact with the microtubule end-binding protein EB1, thereby increasing microtubule stability during the early phase of HIV-1 infection (Sabo et al., 2013). Additionally, ABCE1, an adenosine triphosphatase, and DDX6, a DEAD box RNA helicase, are two proteins normally localized in processing bodies (P-bodies) that have been shown to be co-opted by HIV. ABCE1 binds Gag to promote multimerization at the plasma membrane, where it then dissociates upon assembly and release of virions (Dooher et al, 2007). DDX6 was shown to facilitate immature capsid assembly and is present at numerous assembly intermediate stages (Lingappa et al., 2014). In terms of facilitating Gag interactions with viral genomic RNA, the RNA binding protein Staufen 1 is believed to provide a necessary link. A previous study has led

to the idea that Staufen 1 traffics HIV-1 genomic RNA into nascent virions during assembly, largely due to the findings that Staufen 1 co-precipitates with unspliced HIV-1 RNA and is encapsidated into virions (Cochrane et al. 2006). Incorporation of the envelope glycoprotein at assembly sites is believed to be mediated by TIP47 (tail-interacting protein of 47 kDa), because TIP47 has been shown to provide the link between Gag and Env (Lopez-Verges et al., 2006).

1.6 Y-box binding protein 1

The Y-box binding protein 1 (YB-1; also known as nuclease-sensitive element-binding protein 1) is a 324 amino acid protein with a molecular mass of 36 kDa. In humans, the YBX-1 gene is located on chromosome 1. Like other Y-box binding proteins, YB-1 contains an N-terminus rich in alanine and proline (i.e. the A/P domain), a cold shock domain (CSD), and alternating positively and negatively charged amino acid residues in an elongated C-terminal domain (Figure 3). The CSD contains Ribonucleoprotein-1 (RNP-1) and Ribonucleoprotein-2 (RNP-2) domains involved in interactions with nucleic acids. The functions of YB-1, which occur in both the cytoplasm and cell nucleus, are diverse and involve interactions with both nucleic acids and proteins. YB-1 has been implicated in transcription, DNA repair, mRNA splicing, and translation (reviewed in Eliseeva et al., 2011).

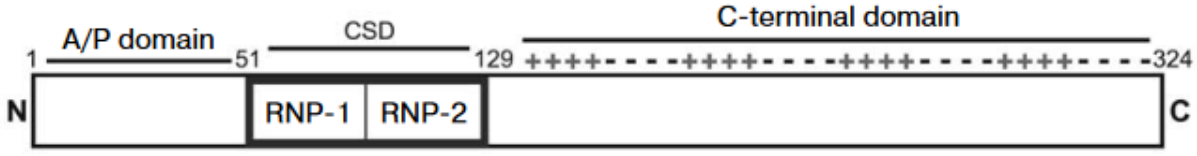


Figure 3: Y-box binding protein 1 domain organization (Eliseeva et al., 2011).

YB-1 has been implicated as a host factor in many viral processes. For instance, by activating adenoviral genes under control of promoter E2, YB-1 is believed to mediate

replication of an adenovirus in the late phase of infection (Holm et al., 2002). Furthermore, transcription of polyomaviruses is thought to be regulated by YB-1 through interactions with polyomavirus JCV promoters (Chen and Khalili, 1995; Kerr et al., 1994). Serving an important role in RNA export, YB-1 was shown to be required for the interaction of viral ribonucleoprotein complexes to properly export the nucleus of cells infected with influenza virus (Kawaguchi et al., 2012). Translation inhibition of the dengue virus was shown to occur when YB-1 binds the dengue virus 3'UTR, suggesting that, in some cases, YB-1 may even have an antiviral effect (Paranjape et al., 2007). Lastly, mouse mammary tumor virus (MMTV) was shown to have particle production decrease due to knockdown of YB-1, suggesting YB-1 facilitates assembly, and, interestingly, the authors found the interaction between Gag and YB-1 to be RNA-dependent (Bann et al., 2014). Overall, these studies suggest YB-1 is an important host factor involved in different stages of replication for a diverse set of viruses.

Previous studies indicate YB-1 may also participate in the replication of HIV-1. For instance, YB-1 was shown to bind both Tat, the HIV-1 transcriptional transactivator, and TAR, an RNA sequence at the 5' end of HIV-1 that forms a unique stem-loop structure and serves as the binding site for Tat (Ansari et al., 1999). The protein complex formed from YB-1 and Tat enhanced activation of the HIV-1 promoter, and the interaction between YB-1 and TAR RNA increases viral gene transcription (Ansari et al., 1999). Furthermore, YB-1 was shown to interact directly with stem loop 2 (SL2) in the 5' untranslated region (UTR) of HIV-1 and increased virus

production by stabilizing viral RNA (Mu et al., 2013). Overexpression of YB-1 in human embryonic kidney cells has also been linked to an increased titer in retroviral vectors, also due to stabilizing interactions of YB-1 with viral RNA (Li et al., 2012). A recent study found YB-1 interacts with HIV-1 integrase, playing a crucial role between reverse transcription of the viral RNA into proviral DNA and nuclear import of proviral DNA (Weydert et al., 2018). In a proteomic study, YB-1 was shown to interact with the matrix (MA) domain of Gag, although the assay did not determine what RNA, if any, bridged the interaction (Li et al., 2016). A subsequent proteomic study used a novel RNA hybridization and mass spectrometry technique to identify host factors that bind HIV-1 RNA, and YB-1 was among the 189 proteins identified (Knoener et al., 2017). At this time, the precise role of YB-1 in HIV-1 assembly remains unknown.

1.7 Yeast: a model system for HIV-1 basic and applied research

1.7.1 Production of recombinant retroviral vectors used in gene therapy

Gene therapy is a technique to treat or prevent genetic diseases by inserting a properly functioning gene in place of one that is malfunctioning or absent. The potential of gene therapy to ameliorate genetic disorders was first proposed forty-six years ago (Friedmann and Roblin, 1972). Since then, the effectiveness of gene therapy has been demonstrated in trials against various diseases including X-linked severe combined immunodeficiency (Hacein-Bey-Abina et

al., 2014), Wiskott-Aldrich syndrome (Aiuti et al., 2013), Hemophilia B (Nathwani et al., 2014), Leber's Congenital Amaurosis (Simonelli et al., 2010), beta-thalassemia (Persons, 2010), and metachromatic leukodystrophy (Biffi et al., 2013). These studies highlight the potential of using gene therapy to treat human disease.

The majority of gene therapy trials utilize retroviral vectors (Merten, 2004). Lentiviruses, a genus of retroviruses which includes HIV, frequently serve as vectors because they can infect both dividing and non-dividing cells, and they contain efficient mechanisms of reverse transcribing ssRNA into dsDNA for integration into the cell genome, which is crucial for long-term expression (Amado and Chen, 1999). Ironically, it is the exact molecular mechanisms that render retroviruses so damaging to human health that hold the potential to alleviate human suffering through gene therapy. In order to turn gene therapy into a practical application, an efficient means to produce retroviral vectors must be established.

The production of recombinant retroviruses for gene therapy could be enhanced by using a genetically amenable eukaryotic organism, such as yeast, rather than mammalian cells. Yeast offer benefits over mammalian cells because they are inexpensive to cultivate, technically simpler to work with, and easier to perform genetic manipulations (Andreola and Lityak, 2012). Many important genes, proteins, and biochemical pathways are conserved from yeast to humans. The baker's yeast *Saccharomyces cerevisiae* has been indispensable in the area of biotechnology, particularly for the large-scale production of insulin to regulate the blood sugar levels of people

with diabetes. Plasmid-based expression of HIV Gag in yeast led to the production of VLPs without infectious viral contamination that often plagues preparations from mammalian cell cultures (Sakuragi et al., 2002). Development of a yeast packaging cell line for large-scale production of non-pathogenic and high-titer viral retroviral vectors in gene therapy has been largely unexplored.

1.7.2 Investigation of HIV-1 assembly in yeast

The experimental tractability of yeast, which make it especially amenable to genetic manipulations and biochemical techniques, was noted in the late 1980's as the principle reason this single-celled eukaryote would contribute greatly to basic research in cell biology (Botstein and Fink, 1988). Since then, yeast have continued to be a model organism that allowed significant contributions to our understanding of functional genomics, gene expression, protein trafficking, and cell cycle regulation.

Yeast can serve as a host for replication of plant viruses such as brome mosaic virus, tomato bushy stunt virus, and cymbidium ringspot virus, as well as for replication of animal viruses such as flock house virus, nodamura virus, human papillomavirus, and bovine papillomavirus (reviewed in Zhao, 2017). Although HIV-1 is unable to replicate in yeast, plasmid-based expression of the HIV-1 Gag protein in *S. cerevisiae* results in the release of virus-like particles (VLPs) upon generation of spheroplasts, in which the cell wall is removed (Sakuragi et al., 2002). Furthermore, these HIV-1 Gag VLPs released from *S. cerevisiae*

spheroplasts were resistant to protease digestion by trypsin unless a delipidizing detergent was present, indicating the VLPs were completely enveloped by the yeast lipid membrane (Sakuragi et al., 2002). The Gag protein expressed in yeast is N myristoylated on glycine (Bathurst et al., 1989), a modification essential in anchoring Gag to the plasma membrane (Bryant and Ratner, 1990) and required for release of HIV-1 virions in both mammalian cells (Göttlinger et al., 1989) and yeast (Sakuragi et al., 2002). Gag-GFP without N-terminal myristoylation displayed cytosolic distribution in yeast cells when compared to the punctate plasma membrane distribution observed in wild type Gag-GFP (Norgan et al., 2012). Furthermore, the myristoylation mutant Gag-GFP was not detected in extracellular fractions of yeast spheroplasts putatively containing VLPs (Norgan et al., 2012). Taken together, these studies suggests that highly conserved genes, proteins, and biochemical pathways shared between yeast and humans may be exploited to use yeast as a model system for studying HIV assembly.

Despite the notable similarities outlined above, virus-like particle (VLP) assembly and release in yeast expressing HIV-1 Gag differs from that in human cells, hindering the use of yeast as a model organism for studying HIV assembly. Notably, budding of HIV-1 Gag VLPs in yeast was shown to be ESCRT-independent (Norgan et al., 2012), despite its requirement in mammalian cells. The same study also corroborated previous evidence provided by Sakuragi et al. (2002) that "late domains" (short peptide motifs such as PTAP, PPXY, and YPXL that recruit ESCRT machinery) in p6 are not required for incorporation of Gag-GFP into VLPs in yeast, and

deletion of late domains had no defect on either subcellular localization of Gag-GFP or VLP budding. Whether *S. cerevisiae* utilizes an unconventional pathway to release HIV-1 Gag VLPs, or if assembly occurs at an intracellular compartment like those described above, is currently unknown.

Another important observations that hinders yeast as a model organism for studying HIV assembly is that many of the VLPs in yeast appear to be retained in early endosomes. A similar observation in mammalian cells was reported when artificially high concentrations of microRNAs are present (Chen et al., 2014). These microRNAs are thought to be interfering with the normal assembly by outcompeting the viral gRNA for some host protein that is critical for normal virion production. This barrier to normal HIV assembly in yeast may actually allow yeast to be a suitable organism to screen various human proteins for their ability to reinstate proper HIV assembly.

A candidate human protein that may participate in HIV assembly is YB-1, which has no homolog in the yeast genome. YB-1 has been recently identified as a key player in recruiting microRNAs and loading them into exosomes via the ESCRT pathway (Shurtleff et al., 2016). This suggests that YB-1 may divert competitor RNA away from Gag in the cytoplasm, thereby allowing Gag to more efficiently multimerize and assemble around viral RNA, facilitating proper assembly. Furthermore, as described earlier in this introduction, the ability of YB-1 to bind

HIV-1 RNA led us to hypothesize that YB-1 may provide a necessary link between Gag and HIV-1 RNA throughout the assembly process.

The goals of this thesis research were to: (1) engineer yeast to produce HIV-1 Gag virus-like particles (VLPs); (2) use minimal HIV-1 RNA sequences to direct selective packaging of a mammalian RNA expression cassette into Gag VLPs; (3) investigate the role of YB-1 in selective RNA packaging.

Our experimental approach first involved expressing HIV-1 Gag in yeast, and purifying VLPs released from spheroplasts using ultracentrifugation through a sucrose gradient. An RNase protection assay was used to determine if a VLP RNA expression cassette, when co-expressed alongside Gag, could be selectively encapsidated and released in VLPs. Lastly, we used RT-qPCR with RNA extracted from purified Gag VLPs to evaluate the role of YB-1 and the 5'UTR-5'Gag RNA in selective RNA packaging. This research validates the use of a yeast system that produces recombinant retroviral vectors with selectively packaged RNA, and lays a foundation for further experiments investigating the role of YB-1 (and other host proteins) in HIV assembly.

2. MATERIALS AND METHODS

2.1 Strains and growth conditions

The *Saccharomyces cerevisiae* strain SEY6210 (MATa *ura3-52 leu2-3,-112 his3-Δ200 trp1-Δ901 lys2-801 suc2-Δ9*) was used in this study. SEY6210 was cultured at 30°C in YPD or the appropriate synthetic drop-out media to select for plasmids carrying yeast genes that complement deficiencies in the auxotrophic strain. Cloning, selection, and plasmid isolation of yeast shuttle vectors was performed using *Escherichia coli* strain MC1061 cultured at 37°C on lysogeny media with 100 µg/mL ampicillin prior to transformation into yeast.

2.2 Molecular techniques

All primer sequences are listed in Table 1. All enzymes were used in accordance with the manufacturer's instructions, unless otherwise indicated. Phusion Flash High-Fidelity PCR Master Mix (Thermo Scientific) or Platinum SuperFi DNA polymerase (Thermo Scientific) were used in PCR reactions for plasmid construction. Thermocycling parameters with Phusion were as follows: Pre-denaturation at 98°C for 10 seconds; 25 cycles of a 2-step protocol with 98°C for 1 second and a combined annealing/extension temperature of 72°C for 15 sec/kb; concluding with a final elongation at 72°C for 1 minute. Thermocycling parameters with Platinum SuperFi DNA polymerase were as follows: Pre-denaturation at 98°C for 30 seconds; 30 cycles of 98°C for 5 seconds, 68°C for 10 seconds, and 72°C for 15 sec/kb; a final elongation at 72°C for 5 minutes. Overlap-extension PCR used a 1:1 molar ratio of each fragment, not exceeding a total of 10 ng in each reaction. Digestion-independent cloning (D.I.C.) used Phusion Flash High-fidelity PCR Master Mix with 100 ng of vector and a 1:3 molar ratio of vector to insert. Thermocycling parameters for D.I.C. were as follows: Pre-denaturation at 98°C 10 seconds,

followed by 15 cycles of 98°C for 1 second and a combined annealing and extension temperature of 72°C for 2.5 minutes, concluding with a final 72°C extension for 1.5 minutes and 4°C hold. All constructs cloned in *E. coli* were isolated using the GeneJET Plasmid Miniprep Kit (Thermo Fisher). GeneJET Gel Extraction and DNA Cleanup Micro Kit (Thermo Fisher) was used for PCR cleanup and gel extraction.

Taq DNA polymerase (Promega) was used for both RT-PCR using cDNA template and colony PCR consisting of cells from a potential clone. Thermocycling parameters for Taq were as follows: Pre-denaturation at 95°C for 2 minutes; 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C extension for 1 minute/kb; followed by a final extension of 72°C extension for 5 minutes.

Restriction enzyme digests used FastDigest restriction enzymes (Thermo Scientific), and occurred at 37°C for 30 minutes for digestion followed by enzyme deactivation at the suggested temperature and duration. Ligation reactions used T4 DNA ligase (Thermo Scientific), and involved a 1:3 molar ratio of vector to insert incubated at 22°C for 1 hour for ligation followed by reaction termination at 70°C for 5 minutes.

2.3 Plasmid Construction

2.3.1 HIV-1 Gag polyprotein (p55)

The source of the HIV-1 gag coding sequence was pGag-EGFP acquired from the NIH AIDS Reagent Program. The yeast vector for Rev-independent expression of HIV-1 Gag was constructed as follows: The HIV-1 gag coding sequence was PCR amplified from pGag-EGFP using Gag-*Eco*RI-F and Gag-*Sal*I-R. The gag stop codon was omitted from the Gag-*Sal*I-R to allow read-through to a C-terminal polyhistidine tag upstream of the yeast alcohol

dehydrogenase (ADH) terminator in the pRS416 vector, thereby creating a Gag-6xHis-tag fusion protein. The 1.5 kb amplicon was digested with *Eco*RI and *Sal*I to create cohesive ends for ligation, as well as *Dpn*I to remove the pGag-EGFP template. The insert was ligated into the 5' *Eco*RI-*Sal*I 3' cloning site in the MCS of pRS416 downstream of the yeast triose-phosphate isomerase (TPI) promoter for constitutive expression in yeast. After a clone containing the insert was identified by colony PCR, the final construct was verified by sequencing at a contract lab (GENEWIZ).

2.3.2 VLP RNA expression cassette for packaging into VLPs

A GeneArt Strings DNA Fragment (Thermo Scientific; see Appendix) was synthesized with the following components: the HIV-1 5'-untranslated region (5'UTR), a CMV promoter and SV40 Poly(A) signal flanking a *Sal*I site for cloning reporter genes that will eventually be monitored in mammalian cells, and the HIV-1 Rev Response Element (RRE). The DNA string was PCR amplified with 5UTR-*Spe*I-F and RRE-*Xho*I-R. The 1.5 kb amplicon was digested with *Spe*I and *Xho*I, and ligated into the 5' *Spe*I-*Xho*I 3' cloning site of pRS414 downstream of a yeast translational elongation factor EF-1 alpha (TEF1) promoter for strong constitutive expression in yeast.

The 5' half of the gag coding region was cloned into the above construct as follows: The first 726 bp of gag was PCR amplified from pRS416 Gag-6xHis-tag with DIC-5UTR-Gag-F and DIC-CMV-Gag-R. These primers contain sequences on the 5' end that overlap the vector, allowing the 5' 726 bp of gag to be stitched between the 5'UTR and CMV promoter. DIC-5UTR-Gag-F also contained an AAG instead of AUG translation start codon at the beginning of gag to prevent expression of additional gag protein. Fragments of the vector were amplified as

follows: (1) Primers TS315-R and DIC-5UTR-R were used to PCR amplify a 954 bp fragment consisting of the upstream sequences in the VLP RNA expression cassette; that is, the yeast TEF1 promoter and HIV-1 5'UTR. (2) Primers DIC-CMV-F and TS315-F were used to PCR amplify a 1630 bp fragment consisting of the downstream sequences in the VLP RNA expression cassette; that is, the CMV promoter, SV40 Poly(A), HIV-1 RRE, and CYC1 terminator. All reactions were digested with *Dpn*I to remove contaminating plasmid template, followed by gel extraction of the expected-sized amplicon, henceforth referred to by their respective sizes.

The 5' half of gag was separately stitched to the upstream and downstream vector sequences of the VLP RNA expression cassette in the following two reactions (1) The 726 bp and 954 bp fragments were stitched and amplified using TS315-R and DIC-CMV-Gag-R to generate a 1.68 kb amplicon (2) The 726 bp and 1630 bp fragments were stitched and amplified using TS315-F and DIC-5UTR-Gag-F to generate a 2.35 kb amplicon. In summary, the 1.68 kb and 2.35 kb amplicons both contained the 5' half of gag (726 bp) overlapping. Finally, overlap-extension PCR using TS315-F and TS315-R with the 1.68 kb and 2.35 kb fragments generated a 3.3 kb amplicon, which consisted of the entire VLP RNA expression cassette with the 5' half of gag inserted between the 5'UTR and CMV promoter.

To clone the modified VLP RNA expression cassette with the additional 5' gag sequence into pRS414, digestion-independent cloning (D.I.C) was used. The original pRS414 construct containing the TEF1-5UTR-CMVp-*Sal*I-SV40Poly(A)-RRE-CYC1 expression cassette was digested with *Spe*I and *Sal*I, the vector without the 5'UTR and CMV promoter was gel extracted, and D.I.C. was used to clone the insert into the vector, as described earlier.

2.3.3 Y-box-binding protein 1

The source of the human Y-box-binding protein 1 (YBX-1) coding sequence in frame with an N-terminal c-myc tag was pDESTmycYBX1, provided by Thomas Tuschl (Addgene plasmid #19878). Myc-YBX1 was PCR amplified from pDESTmycYBX1 using either the primer pair *Asc*I-mycYBX1-F and *Bam*HI-YBX1-NoStop-R, or the primer pair *Asc*I-mycYBX1-F and *Bam*HI-YBX1-R. *Asc*I-mycYBX1-F and *Bam*HI-YBX1-NoStop-R amplified myc-YBX1 without a stop codon to allow read-through to the C-terminal polyhistidine tag upstream upstream of a yeast alcohol dehydrogenase (ADH) terminator in the pRS416 vector, thereby creating a myc-YBX1-6xHis-tag fusion protein. *Asc*I-mycYBX1-F and *Bam*HI-YBX1-R amplified myc-YBX1 with a stop codon. The PCR products were digested with *Sgs*I and *Bam*HI, followed by gel extraction of the 1.1kb amplicons. *Sgs*I is an isoschizomer of *Asc*I with the same recognition sequence. Both inserts were separately ligated into the 5' *Sgs*I-*Bam*HI 3' cloning site of pRS416 under the yeast TPI promoter for constitutive expression. After transformation and cloning in *E. coli* MC1061, the TPI-mycYBX1-ADH expression cassette was digested from pRS416 using *Xho*I and *Not*I, and the 1.7 kb cassette was gel extracted. The expression cassette was then ligated into the 5' *Xho*I-*Not*I 3' cloning site in the MCS of pRS315 to generate plasmid pRS315 TPI-mycYBX1-ADH.

2.4 Transformations

Electroporation was used to transform *E. coli* MC1061. Two microliters of each ligation were added to 40 µl electrocompetent MC1061 and transferred to a 2mm gap cuvette (Fisher Scientific). After a single pulse of 2.5 kV, transformants recovered in 500 µl S.O.C. at 37°C for 30 minutes before being plated on lysogeny media with 100 µg/mL ampicillin.

A standard lithium acetate (LiAc) transformation procedure was used to transform yeast vectors into SEY6210. A 10 mL culture of yeast grown to early log phase ($OD_{600} \approx 0.5$) was pelleted at 2,400 x g for 3 minutes, washed once with sterile water, and resuspended in 10 mL of

LiT (100mM LiAc and 10 mM Tris-Cl, pH 7.4) with 10 mM DTT. The cells were incubated with gentle shaking for 45 minutes, pelleted, and resuspended in LiT by adding 145µl per transformation. 125µl of cells were gently mixed with 0.5-1 µg plasmid DNA and 5 µl of herring sperm (10 µg/µl) and incubated for 10 minutes at room temperature. 300 µl of 50% PEG-LiT was added and gently mixed, and cells were incubated at room temperature with end-over-end rotation for another 10 minutes. Afterwards, 15µl of DMSO was added and transformation tubes were pulse vortexed and immediately placed at 42°C for 5 minutes. Cells were gently pelleted in a microcentrifuge at 4,000 rpm for 4 minutes, resuspended in 1 mL YPD, and incubated at 30°C with shaking for 45 minutes. Finally, cells were pelleted in a microcentrifuge at 5,000 rpm for 1 minute, resuspended in residual YPD, and plated onto appropriate synthetic drop-out media complemented by the plasmid(s).

2.5 Western Blotting

Whole cell lysates of *S. cerevisiae* were prepared by adding 1X SDS Sample Buffer with 1.25% Beta-mercaptoethanol to yeast pellets using 40 µl per 2.5 OD units. Next, ¾ volume of acid-washed glass beads were added, and for a total of two times samples were incubated at 99°C for 3 minutes with shaking, briefly placed on ice, then vortexed for 30 seconds. Beads were pelleted at maximum speed in a microcentrifuge for two minutes, the supernatant was collected, and samples were either cooled to room temperature and loaded onto an SDS-PAGE gel, or stored at -20°C until use. Pellets of virus-like particles (VLPs) isolated via ultracentrifugation were resuspended in 100 µl of 1X SDS Sample Buffer with 1.25% Beta-mercaptoethanol and incubated at 99°C with shaking for 4 minutes. Samples were either cooled to room temperature and loaded into an SDS-PAGE gel, or stored at -20°C until use.

Proteins were resolved on 12% SDS-PAGE for ~90 minutes at 100 Volts, and transferred to a nitrocellulose membrane using a submersible tank transfer system. Blocking occurred in 3% non-fat dry milk in TBS-T (25 mM Tris, pH 7.4; 150 mM NaCl; 0.5% Tween-20) overnight at 4°C with gentle shaking. To detect Gag-6xHis-tag, 1:1000 HisProbe™-HRP Conjugate (Thermo Scientific) diluted in TBS-T was used. To detect GFP, 1:2000 anti-GFP conjugated to HRP (Thermo Scientific) diluted in TBS-T was used. To detect myc-YB1, one of the following probing methods were used: (1) 1:5000 anti-c-Myc conjugated to HRP (Invitrogen) diluted in TBS-T with 3% non-fat milk or (2) A combination of 1:2000 anti-c-Myc (Sigma) and 1:1000 YB-1 (Cell Signaling Technology) primary antibodies diluted in PBS-T with 5% non-fat dry milk, followed by 1:2000 anti-rabbit IgG HRP-linked secondary antibody (Cell Signaling Technology) diluted in TBS-T with 5% BSA. Detection was performed using chemiluminescence (Pierce ECL Western Blotting Substrate; Thermo Scientific) on a ChemiDoc XRS+ (Bio-Rad). Band intensity was quantitated on a ChemiDoc XRS+ with Image Lab software (Bio-Rad), and a t-test was used to compare band intensity among treatments.

2.6 End-point RT-PCR

First-strand cDNA synthesis was conducted in a 20 µl reaction containing 100 pmol OdTCDS3 (Table 1), 1mM dNTP mix, and 200 U RevertAid Reverse Transcriptase (Thermo Scientific) with the appropriate amount of reaction buffer. Because regions of the viral RNA are GC rich with extensive secondary structure, total RNA and OdTCDS3 were first incubated at 65 °C for 5 minutes followed by chilling on ice prior to being added to the cDNA synthesis reaction. Reactions were incubated at 45°C for 1 hour for transcription followed by 70°C for 10 minutes to terminate the reaction. PCR amplification was conducted with GoTaq DNA

Polymerase (Promega) as described in the Molecular Techniques section, unless otherwise indicated (Figure 4).

RT-PCR targeting the VLP RNA expression cassette was conducted with either total RNA from yeast extracted using the the GeneJET RNA Purification Kit (Thermo Scientific), or from total nucleic acids extracted from VLPs using the PureLink Viral RNA/DNA Mini Kit (Invitrogen). To PCR amplify the VLP RNA expression cassette converted to cDNA, GoTaq DNA polymerase was used as described in the Molecular Techniques section with *Spe*I-CMV-F and CDShort and a 52°C annealing temperature. For the ribonucleoprotein (RNP) pull down assay, PCR amplification to detect the VLP RNA expression cassette converted to cDNA was performed with GoTaq DNA polymerase using primers RRE-qPCR-F and CDShort, a 53°C annealing temperature, and 40 cycles.

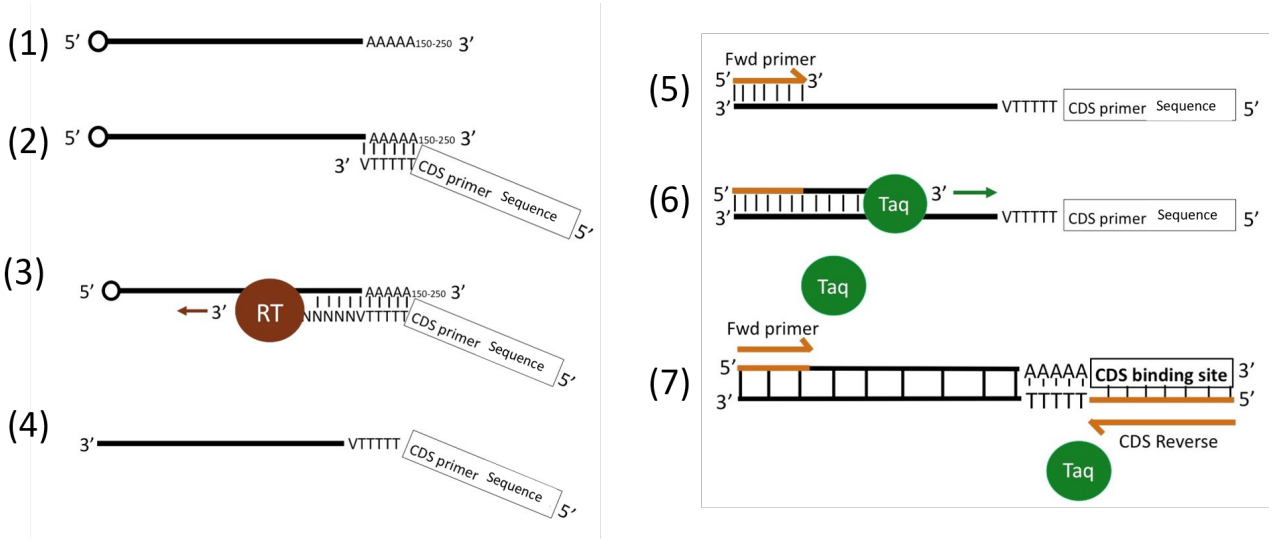


Figure 4: Schematic representation of RT-PCR used in this study to amplify the VLP RNA expression cassette. (1) Total RNA from yeast or VLPs was extracted. For simplicity, only mRNA is shown here. (2-3) First-strand complementary DNA synthesis was performed with OdTCDS3 and reverse transcriptase. (4) RNase-H activity of reverse transcriptase hydrolyzes DNA in DNA-RNA duplex. (5-6) In the beginning cycles of PCR, Taq extends the 3' end of a primer that anneals to the the CMV promoter in the VLP RNA expression cassette, thereby creating double-stranded DNA; (7) PCR amplification of the VLP RNA expression cassette.

2.7 VLP Budding Assay

Yeast transformed with the indicated constructs were grown to $OD_{600} \approx 0.5$. Twenty OD units were pelleted, resuspended in 3 mL softening buffer (0.1M Tris pH9.4, 10 mM DTT), and incubated at room temperature with gentle shaking for 10 minutes. Cells were pelleted and resuspended in 2 mL spheroplasting buffer (1M sorbitol, 3.5 mM DTT, 1X synthetic complete yeast media with 2% glucose) to which a final concentration of 1 mg/mL Zymolyase 100T (US Biological) had been added. Digestion of yeast cell walls occurred for 30 minutes at 30°C with gentle shaking. Spheroplasts were pelleted gently at 500 x g for 3 minutes, resuspended in 2 mL YPD with 1M sorbitol, and incubated overnight at 30°C with gentle shaking for VLP budding and release. After overnight (~16 hours) incubation, spheroplasts were pelleted and the clarified culture supernatant was filtered through a 0.45 micron filter to further exclude remaining spheroplasts. After filtration, the filtrate was then treated with 50 µg/mL RNase A (Thermo Scientific) either in the presence or absence of 0.5% Triton X-100 and incubated at 37°C for 30 minutes. Next, 1.8 mL of filtered supernatant were loaded onto 350 µl of 30% sucrose cushion, and VLPs were pelleted at 45,000 RPM for 2 hours at 4°C in an ultracentrifuge. After ultracentrifugation, the supernatant was gently removed, being careful to not disturb the VLP pellet. The VLP pellet was resuspended in 75 µl of PBS and subject to downstream analysis.

2.8 Quantitative RT-PCR

VLP RNA was converted to cDNA as described in the End-point RT-PCR section with the following modification: 10 µl of 5 ng/µl total RNA extracted from SEY6210 transformed with pRS416 GFP-6xHis-tag was added to 15 µl of VLP RNA, and 19.5 µl of this combined RNA was added to a 30 µl reaction with 300 U RevertAid Reverse Transcriptase (Thermo

Fisher). The addition of total yeast RNA was intended to be an internal control, allowing us to normalize data by accounting for variability in cDNA synthesis and RT-qPCR between treatments; However, low qPCR efficiency using the primers intended for this purpose led to us seeking alternative methods for normalization (e.g., conducting the experiment in three independent replicates, and determining amount of VLP recovery via quantitative western blot). RT-qPCR reactions were performed in either 10µl or 20µl reaction volumes with Maxima SYBR Green/ROX qPCR Master Mix (Thermo Fisher), 0.3 µM RRE-qPCR-F, 0.3 µM CDShort, and 1 µl cDNA. Real-time thermocycling was performed on a CFX96™ Real-Time PCR Detection System (BioRad Laboratories, Inc.) with the following thermal profile: Initial denaturation 95°C 10 minutes; 40 cycles of 95°C for 15 seconds followed by a combined annealing/extension of 60°C for 1 minute. To create a template for the standard curve, total RNA from yeast expressing the VLP RNA expression cassette was converted to cDNA, and RT-PCR amplification using RRE-qPCR-F and CDShort was used to generate the target-sized product. The standard curve consisted of five consecutive 2-fold serial dilutions of this product.

To determine the relative fold-change in RNA packaged into VLPs due to either the YB-1 protein or the 5'UTR-5'Gag RNA, the following equation was used:

$$\text{Relative fold-change} = 2^{(\Delta Ct)}, \text{ where } \Delta Ct = (Ct \text{ control} - Ct \text{ treatment})$$

"Ct treatment" refers to VLP RNA from yeast expressing the molecule of interest; that is, either YB-1 protein or the 5'UTR-5'Gag RNA. "Ct control" refers to VLP RNA from yeast not expressing the molecule of interest.

2.9 Ribonucleoprotein (RNP) Pull Down

SEY6210 containing the indicated constructs were grown overnight in synthetic drop-out media and diluted to OD_{600} of 0.1 in YPD the following morning. Once early log phase growth ($OD_{600} \approx 0.35$ to 0.5) had been achieved, 30 OD units were pelleted, resuspended in 3 mL softening buffer (0.1M Tris pH9.4, 10 mM DTT), and incubated at room temperature with gentle shaking for 10 minutes. Aliquots of 10 OD units were made, pelleted in a microcentrifuge at 5,000 RPM for 1 minute, and cells were then spheroplasted using the procedure described earlier. Spheroplasts were pelleted in a microcentrifuge at 5,000 RPM for 1 minute, the supernatant was discarded, and spheroplast pellets were snap-frozen in liquid nitrogen before being stored at -70°C.

After thawing samples on ice, 600 µl of ice-cold lysis buffer (100 mM HEPES, pH 7.5; 400 mM NaOAc, pH 7.5; 4 mM EDTA; 10 mM MgOAc; 0.25% NP-40; 1X Thermo Scientific Halt Protease Inhibitor Cocktail) was used to gently resuspend 10 OD units of spheroplasts. For lysate incubated with HisPur™ Ni-NTA Resin (Thermo Scientific), EDTA was omitted from the lysis buffer following the manufacturer's instructions. Samples were sonicated briefly until noticeable clearing was observed (Setting 5, with a two short pulses of 1-2 seconds) and cellular debris was removed by centrifugation. Lysate was added to 30 µl Pierce™ Anti-c-Myc Agarose (Thermo Scientific) for constructs expressing myc-YB1 without a 6xHis-tag, or 60 µl of HisPur™ Ni-NTA Resin (Thermo Scientific) for constructs expressing myc-YB1-6xHis-tag. Samples were incubated on an end-over-end rotator for 1 hour at 4°C. The resin containing bound RNP complexes was gently pelleted in a microcentrifuge at 7,000 RPM for 1 minute at 4°C, and 400 µl of the supernatant containing unbound RNP complexes were stored at -70°C for later RNA extraction representing the "unbound" fraction. Resin was then washed three times in 700 µl TBS-T and stored at -70°C until downstream analysis.

To extract RNA bound to the resin in RNP complexes, 400 µl of RNP disassociation solution (1.25 mg/mL Proteinase K; 40 mM DTT) was added to each resin, and samples were

incubated at 56°C for 15 minutes. To ensure equal treatment of unbound RNP complexes, Proteinase K and DTT were added to obtain the same final concentrations to unbound samples. Total RNA was purified from all samples using GeneJET RNA Purification Kit (Thermo Scientific) and eluted with 100 µl of elution buffer, following the manufacturer's instructions. Samples were converted to cDNA and used in RT-PCR as described above.

Table 1: Oligonucleotide primer sequences used in this study. All sequences listed are in the 5' to 3' polarity.

Primer Name	**Sequence**	**Tm (°C)**
Gag-*Eco*RI-F	GCCGAATTCATGGGTGCGAGAGCGTCAG	71
Gag-*Sal*I-R	ACGCGTCGACTTGTGACGAGGGGTCGTTGC	73.6
PRS4-Ext-F	TCGACCTTGAGCATGTGG	56
PRS4-Ext-R	CCGGAACGGTATTTGTGC	56
GagEGFP-*Sal*I-R	AGGTCGACTTACTTGTACAGCTCGTCCATG	68.1
5UTR-*Spe*I-F	GCCACTAGTGGTCTCTCTGGTTAGACCAG	69.5
RRE-*Xho*I-R	TTTTCTCGAGACTAGCATTCCAAGGCACAGCAG	69.5
DIC-5UTR-Gag-F	GCGGAGGCTAGAAGGAGAGAGAaGGGTGCGAGA GCGTCAGTATTAAGC	78.8
DIC-CMV-Gag-R	TGACCCCGTAATTGATTACTATTAATAACTAGGTA CTAGTAGTTCCTGCTATGTCACTTC	73.5
TS315-R	CGACTGGAAAGCGGGCAGTG	63.4
DIC-5UTR-R	TCTCTCTCCTTCTAGCCTCCGC	64
TS315-F	CCATTCAGGCTGCGCAACTG	61.4
DIC-CMV-F	TAGTTATTAATAGTAATCAATTACGGGGTC	59.9
OdTCDS3	GGAACAAAAGCTGGAGCTCTTTTTTTTTTTTTTV	63.2
CDShort	GGAACAAAAGCTGGAGCTC	56.7

ENV-*Eco*RI-F	GCCGAATTCATGAGAGTGAAGGAGAAATATCAG	67
ENV-*Sal*I-R	AGGTCGACTAGCAAAATCCTTTCCAAGCCCTG	69.5
ENV-*Xma*I-F	CTGCCCGGGACAGAAAAATTGTGGGTCACAG	70.8
ENV-myc-XmaI-R	TGTCCCGGGCAGATCTTCTTCAGAAATAAGTTTTTGTTCAGCACTACAGATCATCAACATC	75.4
*Spe*I-CMV-F	GCCACTAGTTAGTTATTAATAGTAATCAATTACGGGGTCAT	67.4
Act-1 F	CGCTTACTGCTTTTTTCTTCC	55.9
*Asc*I-mycYBX1-F	TTTGGCGCGCCATGGAACAAAAACTTATTTCTGAAG	68.3
*Bam*HI-YBX1-NoStop-R	CCCGGATCCCTCAGCCCCGCCCTGCTCAG	78
*Bam*HI-YBX1-R	CCCGGATCCTTACTCAGCCCCGCCCTGCTCAG	77.2
RRE-qPCR-F	CAGGCAAGAATCCTGGCTG	58.8

3. RESULTS

3.1 Expression of HIV-1 Gag in yeast

As a first step in creating a yeast system to produce virus-like particles (VLPs), we sought to express HIV-1 Gag polyprotein in yeast. The HIV-1 Gag coding sequence in-frame with a C-terminal polyhistidine tag was cloned into pRS416 downstream of a yeast triose-phosphate isomerase (TPI) promoter for constitutive expression. After plasmid construction, the Gag-6xHisTag coding sequence was verified by sequencing (see Appendix). After transformation into *Saccharomyces cerevisiae* SEY6210, a western blot of whole-cell lysate using Ni-HRP to detect the C-terminal polyhistidine tag showed expression of Gag 6xHis-tag as a protein of the expected size of 55 kDa (Figure 5). To ensure the Ni-HRP, which has high affinity for histidine residues, was not simply binding an endogenous yeast protein rich in histidine, SEY6210 transformed with a construct that expressed GFP with a 6xHis-tag instead of Gag with a 6xHis-tag was analyzed in the adjacent lane. An expected sized band of 28 kDa in the GFP 6xHis-tag lane was detected, while the 55 kDa protein representing Gag 6xHis-tag was not detected, indicating the 55 kDa protein in the Gag 6xHis-tag lane was not the result of Ni-HRP binding an endogenous yeast protein rich in histidine residues.

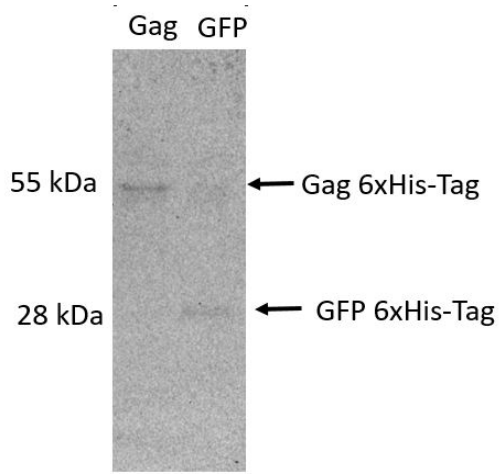


Figure 5: Expression of HIV-1 Gag polyprotein in *Saccharomyces cerevisiae*. Western blot analysis using Ni-HRP to detect the C-terminal HisTag fused to the HIV-1 Gag polyprotein. As a control, lysate from yeast that expressed GFP with a HisTag was analyzed in the lane on the right. Expected sizes: HIV-1 Gag = 57 kDa; GFP = 28 kDa.

3.2 Production of HIV-1 Gag VLPs from yeast spheroplasts

A system for recombinant retroviral vector production is contingent upon efficient production and isolation of virus-like particles (VLPs). Retroviral VLPs are icosahedral assemblages of the viral Gag polyprotein that assemble, bud, and are released from cells. Although VLP budding resembles that of infectious virions, VLPs are non-infectious because they do not contain necessary viral genetic information or envelope glycoproteins and do not undergo proteolytic cleavage by the viral protease to trigger maturation into a mature virion.

Once the HIV-1 Gag protein was successfully expressed in yeast, a VLP budding assay was used to determine if Gag would assemble, bud from the plasma membrane, and be released from spheroplasts as VLPs. Spheroplasts of SEY6210 transformed with pRS416 Gag-6xHis-tag were incubated overnight in isotonic medium for VLP budding and release. Next, clarified culture supernatant was subject to ultracentrifugation through a 30% sucrose cushion to pellet the VLPs. Lastly, the resuspended ultracentrifuge pellet was analyzed by western blot using Ni-

HRP to detect the C-terminal polyhistidine tag on Gag. In addition to analyzing the ultracentrifuge pellet, cell lysates were also analyzed for the presence of Gag. Two necessary negative controls were included: (1) SEY6210 transformed with pRS416 Gag-6xHis-tag that had not been converted to spheroplasts (i.e., had an intact cell wall) (2) spheroplasts of SEY6210 that did not contain pRS416 Gag-6xHis-tag, and therefore were not expressing Gag.

HIV-1 Gag VLPs were detected in the ultracentrifuge pellet of spheroplasts expressing Gag 6xHis-tag, as evidenced by a 55 kDa protein detected by western blot (Figure 6). Similarly, Gag was also detected in the lysate of the same culture. Gag was not detected in the ultracentrifuge pellet of non-spheroplasted yeast, despite its presence in the cell lysate from the same culture. This provides evidence that VLPs are only efficiently released from yeast spheroplasts. Spheroplasts that did not contain the Gag construct were void of any similar sized band in both cell lysate and ultracentrifuge pellet, as expected.

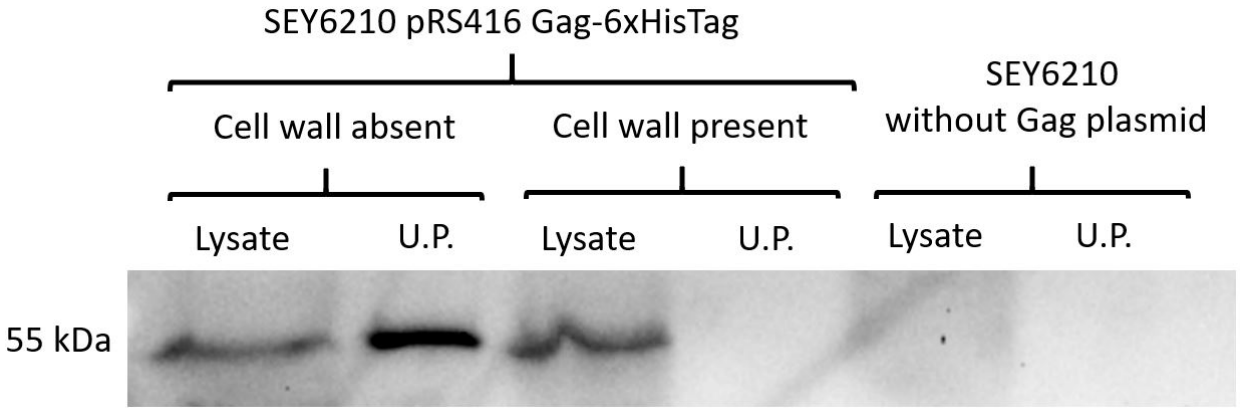


Figure 6: Purification of HIV-1 Gag virus-like particles from *Saccharomyces cerevisiae* after cell wall degradation. Western blot analysis using Ni-HRP to detect HIV-1 Gag-6xHisTag in either the whole-cell lysate or ultracentrifuge pellet (U.P.) of clarified culture media. HIV-1 Gag = 55 kDa.

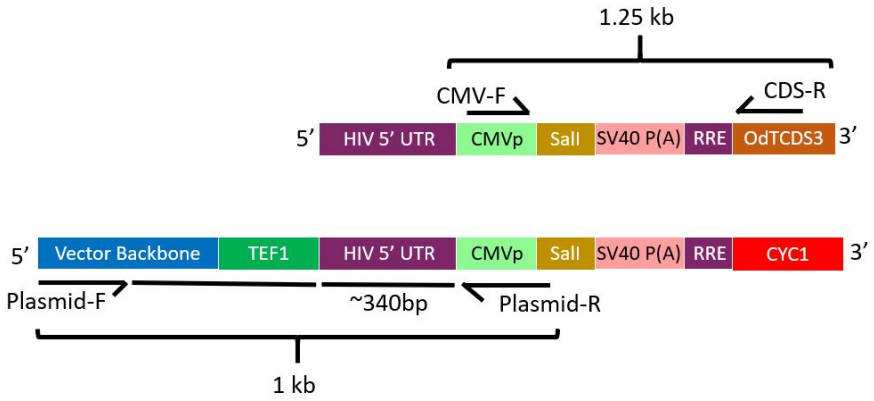
3.3 Transcription of VLP RNA expression cassette in yeast

Having observed expression of HIV-1 Gag and production of VLPs in yeast, we next sought to express a RNA expression cassette that could participate in Gag-mediated selective packaging into VLPs. A yeast plasmid was constructed that contained the HIV-1 5'-untranslated region (5'UTR), a cytomegalovirus (CMV) promoter and simian vacuolating virus 40 (SV40) Poly(A) signal flanking a *Sal*I site, and the HIV-1 Rev Response Element (RRE) (sequence provided in Appendix). The 5'UTR contains the HIV-1 packaging signal. The purpose of the CMV promoter and SV40 Poly(A) signal flanking a *Sal*I site is for cloning reporter genes that will eventually be monitored in mammalian cells after transfection by VLPs. The RRE was included because it is implicated in HIV-1 genome packaging, as well as export of unspliced and partially spliced mRNA from the nucleus in the presence of Rev (Kim et al., 1989). This VLP RNA expression cassette cloned downstream of a yeast translational elongation factor EF-1 alpha (TEF1) promoter for strong, constitutive expression in yeast.

Yeast were transformed with the VLP RNA expression cassette construct containing the 5'UTR, or an identical construct that lacked the 5'UTR; The construct that lacked the 5'UTR was created in order to evaluate the role of the 5'UTR in later experiments. Total RNA was extracted and converted to cDNA with OdTCDS3, an oligo-dT with a unique primer sequence overhanging the 5' end (referred to as CDS3). First-strand cDNA synthesis with OdTCDS3 results in cDNA linked to the 5' sequence on OdTCDS3. When paired with a forward primer complementary to the target cDNA, CDS3 allows PCR amplification of the target cDNA while avoiding amplification from contaminating DNA, since only a single primer (i.e., the forward primer) is complementary to contaminating DNA. RT-PCR using CDS3 paired with a forward

primer targeting the CMV promoter amplified the expected sized band of 1.25 kb from both constructs (Figure 7). A series of controls were used to verify that the 1.25 kb amplicon was the result of transcription of the VLP RNA expression cassette. RNA that had not been reverse transcribed into cDNA (i.e., No-RT control) did not lead to amplification of any detectable product, indicating the 1.25 kb amplicon was a product of cDNA. Primers that annealed to the plasmid backbone and CMV promoter were used to determine that plasmid DNA was present in cDNA, indicating the total RNA extraction had contaminating DNA. However, the primers used in RT-PCR did not amplify the 1.25 kb product from purified plasmid template, and instead amplified a ~300 bp fragment (likely the product of non-specific binding), once again showing CDS3 specificity for cDNA only. Taken together, these data suggest: (1) the primers used in RT-PCR only amplify a 1.25 kb product from the VLP RNA expression cassette converted to cDNA and (2) the VLP RNA expression cassette is being transcribed in yeast.

A



B

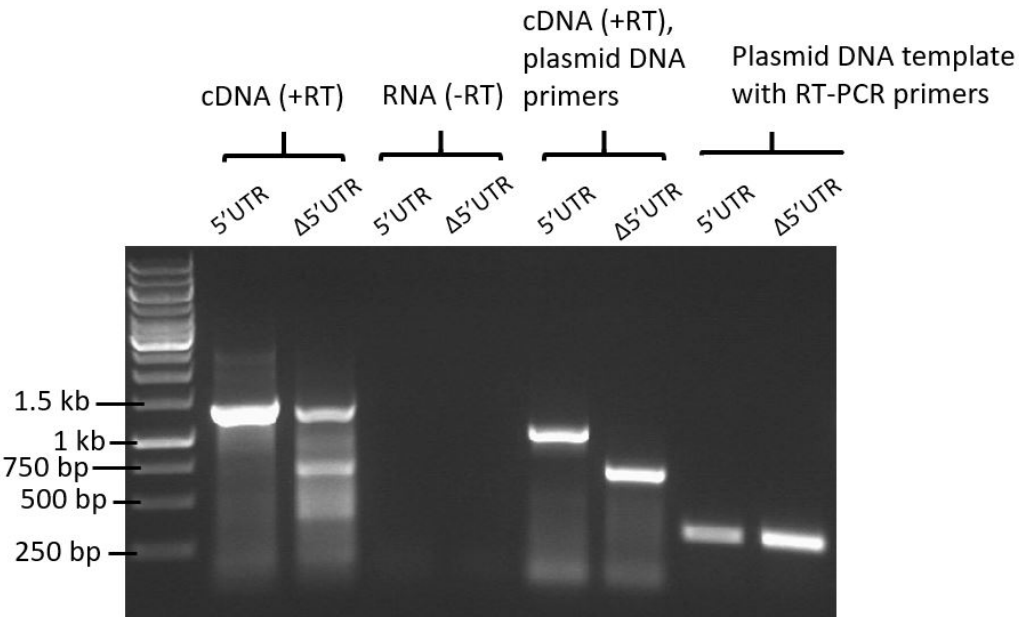


Figure 7: RT-PCR to verify transcription of the VLP RNA expression cassette in yeast. (A) Diagram showing the template and expected sizes of amplicons from the VLP RNA expression cassette (top) and contaminating plasmid DNA (bottom). To accommodate text clarity, the diagram is not drawn to scale. (B) RT-PCR targeting the VLP RNA expression cassette transcribed in yeast amplified the expected sized band of 1.25 kb (Lane 1 and 2). This 1.25 kb band was not present in the No-RT control (Lane 3 and 4). Despite the presence of contaminating plasmid DNA (Lane 5 and 6), the primers used in RT-PCR did not amplify the 1.25 kb band from purified plasmid DNA template, and instead amplified a 300 bp fragment, likely the product of non-specific priming (Lane 7 and 8).

While it was long believed the HIV-1 packaging signal in the 5'UTR was the only requisite for selective packaging of viral RNA into assembling virions, recent evidence suggests that the efficiency of selective packaging is heavily influenced by downstream viral RNA sequences (Liu et al. 2017). When expressed downstream of the 5'UTR, the 5' half of gag coding sequence was shown to increase selective packaging of non-viral RNA into VLPs from background levels (1-3%) up to 25% (Liu et al. 2017). Given these recent findings, we used digestion-independent cloning to insert the 5' half of gag downstream of the 5'UTR in the construct expressing the VLP RNA expression cassette, and transcription of this new construct in yeast was verified by RT-PCR as described above (data not shown). Since the *Spe*I-CMF-F and CDS3 primer pair amplified a product of identical size to the amplicon when the 5' half of gag was not present, we verified the insertion of this additional sequence in the plasmid by performing a diagnostic digest with an added restriction site near the end of the gag coding sequence (data not shown).

3.4 Selective packaging of the VLP RNA expression cassette into HIV-1 Gag VLPs

The nucleocapsid domain of Gag binds the HIV-1 packaging signal, a sequence in the 5'UTR, to direct selective packaging of viral genomic RNA into assembling virions (Freed, 2015). Having produced Gag VLPs from yeast and verified transcription of the VLP RNA expression cassette, we sought to determine if the VLP RNA expression cassette could be selectively packaged into Gag VLPs.

A limitation of pelleting VLPs through a sucrose cushion as opposed to sucrose gradient centrifugation is that it is not possible to separate VLPs from exosomes or other extracellular, membrane-bound vesicles that contain cellular mRNA. We reasoned that these vesicles, likely

contaminating our ultracentrifuge pellet, could incorporate the VLP RNA expression cassette, providing a false positive with RT-PCR if the VLP RNA expression cassette is not packaged into Gag VLPs. To test this, we determined if the presence of the VLP RNA expression cassette in the VLP fraction depended on Gag presence, reasoning that if Gag were absent the ultracentrifuge pellet would not contain the VLP RNA expression cassette. Using spheroplasts that expressed the VLP RNA expression cassette in the absence of Gag (i.e., Gag VLPs were not produced from this strain), RT-PCR produced a very faint, barely visible band corresponding to the VLP RNA expression cassette in the ultracentrifuge pellet (Figure 8). The abundance of the VLP RNA expression cassette seemed to be near the minimum for detection, even for a sensitive technique like RT-PCR, and particularly in comparison to when Gag was present (Figure 8). This suggests the VLP RNA expression cassette is present at extremely low abundance in exosomes or extracellular vesicles that are pelleted with VLPs, while the predominant source of the VLP RNA expression cassette in the ultracentrifuge pellet comes from its association with Gag.

Having determined that the VLP RNA expression cassette in the ultracentrifuge pellet is associated with Gag rather than contaminating extracellular vesicles, we conducted an RNase protection assay to determine if the VLP RNA expression cassette was encapsidated in enveloped Gag VLPs, rather than associated with Gag outside of this context. The rationale behind the RNase protection assay was that if an RNA is only sensitive to RNase in the presence of a detergent, then a membrane (i.e., the plasma membrane enveloping a VLP) is protecting the RNA. Spheroplasts expressing both Gag and the VLP RNA expression cassette were incubated overnight in isotonic medium for VLP budding and release. Clarified culture supernatant was passed through a 0.45 micron filter, and equal aliquots were treated with RNase either in the

presence or absence of a detergent. Samples were then ultracentrifuged through a 30% sucrose cushion, and the VLP pellet was subject to both RT-PCR and western blot analysis. We found that the VLP RNA expression cassette could be detected in purified Gag VLPs when a detergent was not added prior to treatment with RNase (Figure 8). However, when VLPs were treated with a detergent before treatment with RNase, the VLP RNA expression cassette was no longer detected in the ultracentrifuge pellet (Figure 8). Western blot analysis to detect Gag in detergent-treated VLPs showed only trace amounts of Gag, indicating the detergent had indeed abrogated Gag interactions. These data suggest that a membrane is protecting the VLP RNA expression cassette found in the ultracentrifuge pellet from yeast co-expressing Gag and the VLP RNA expression cassette, supporting that the RNA is packaged into bonafide Gag VLPs.

Having shown that the VLP RNA expression cassette is packaged into Gag VLPs, we sought to determine the selectivity to which the RNA was packaged. In addition to RT-PCR targeting the VLP RNA expression cassette, we also conducted RT-PCR targeting actin mRNA with the same samples. Actin mRNA was not detected in the ultracentrifuge pellet regardless of whether Gag was expressed, and despite the detection of actin mRNA from total yeast RNA. The exclusion of actin mRNA from VLPs supports that packaging of the VLP RNA expression cassette is selective, to the extent that one particular, highly-abundant cytoplasmic mRNA is not incorporated into Gag VLPs.

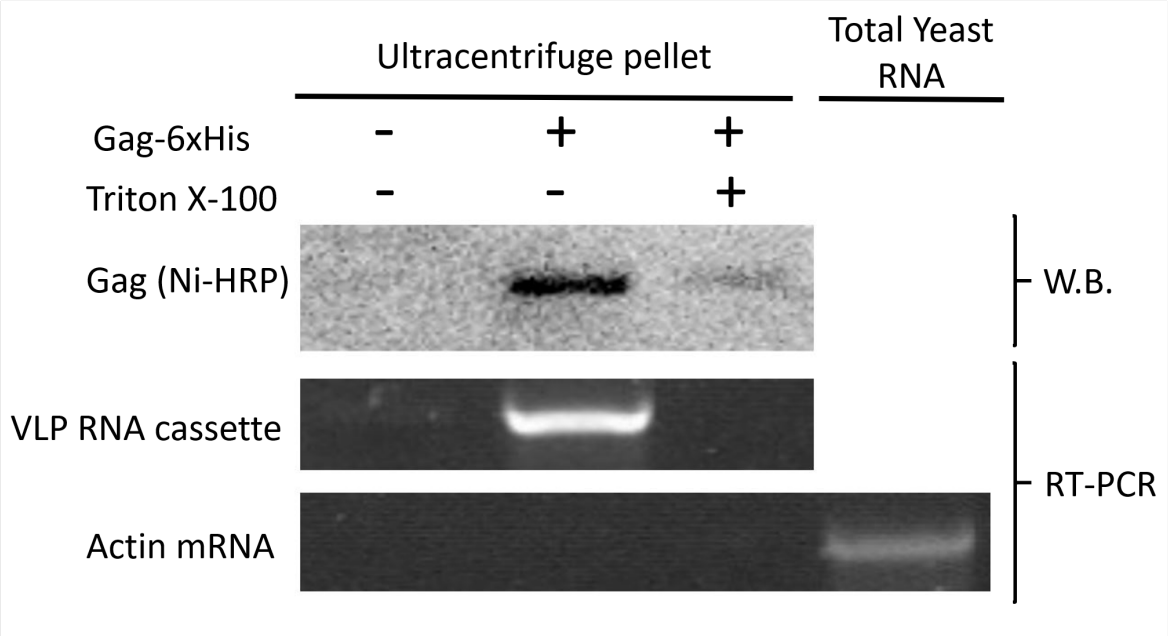


Figure 8: RNase protection assay showing VLPs produced in yeast contain a selectively packaged VLP RNA expression cassette. All yeast used in the experiment expressed the VLP RNA expression cassette, and all VLPs were treated with RNase before ultracentrifugation. In the indicated lane, a detergent was used to dissolve membranes protecting RNA. All VLP pellets were also subject to RT-PCR using primers that targeted actin mRNA; As a positive control to detect actin mRNA, total yeast RNA was used. RT-PCR representative of two independent replicates. W.B.= Western Blot.

3.5 Expression of YB-1 in yeast

The creation of a yeast system that supports HIV-1 assembly and selective RNA packaging provided us an opportunity to evaluate host factors that may participate in these processes. We chose to focus on the human Y-box binding protein 1 (YB-1), which has been implicated in facilitating retroviral assembly via interactions with viral genomic RNA (Bann et al., 2014). YBX-1 was cloned into two yeast centromeric vectors and constitutively expressed under a TPI promoter. One construct expressed YB-1 with an N-terminal c-Myc tag, and the other construct expressed YB-1 with both an N-terminal c-myc tag and a C-terminal

polyhistidine tag. Both constructs were verified as being expressed in yeast via western blot using anti-c-myc conjugated to HRP (Figure 9).

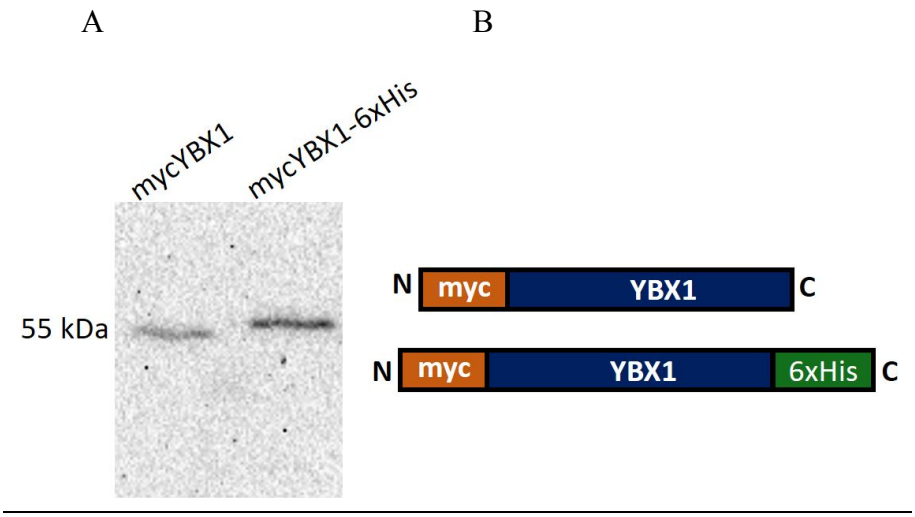


Figure 9: Expression of YB-1 in *Saccharomyces cerevisiae*. (A) Western blot analysis using anti-c-myc conjugated to HRP to detect expression of myc-YB-1 (left) and myc-YB1-6xHisTag (right) in yeast. Expected size of YB-1 = 55 kDa. (B) Linear depiction of YB-1 fusion proteins expressed in yeast

3.6 Effects of YB-1 and 5'UTR-5'Gag RNA on selective RNA packaging into VLPs

Having shown Gag-mediated selective packaging of the VLP RNA expression cassette into VLPs, the effects of the human YB-1 protein and the viral 5'UTR-5'Gag RNA was evaluated on selective RNA packaging. In addition to HIV-1 Gag, yeast were transformed with additional plasmids to set up four trials presented in Figure 10. After incubating spheroplasts in isotonic medium for VLP budding and release, clarified culture supernatant was treated with RNase to remove RNA not packaged into VLPs. VLPs were then purified and equal aliquots of the resuspended VLP pellet were analyzed by western blot to detect Gag and quantitative RT-PCR to determine the amount of VLP RNA expression cassette packaged into VLPs. The entire experiment was repeated in triplicate. No major or consistent differences in Gag VLP yield were

visually observed by western blot when YB-1 or the 5'UTR-5'Gag RNA was expressed alongside Gag compared to when those factors were absent (Figure 11A). Using band intensity data collected with an imaging software, a t-test revealed no significant differences (p-value > 0.19) between treatments (Figure 11B). These results suggest that any differences in RNA packaging are due to either the 5'UTR-5'Gag RNA or YB-1, since there were no apparent differences in Gag VLP recovery.

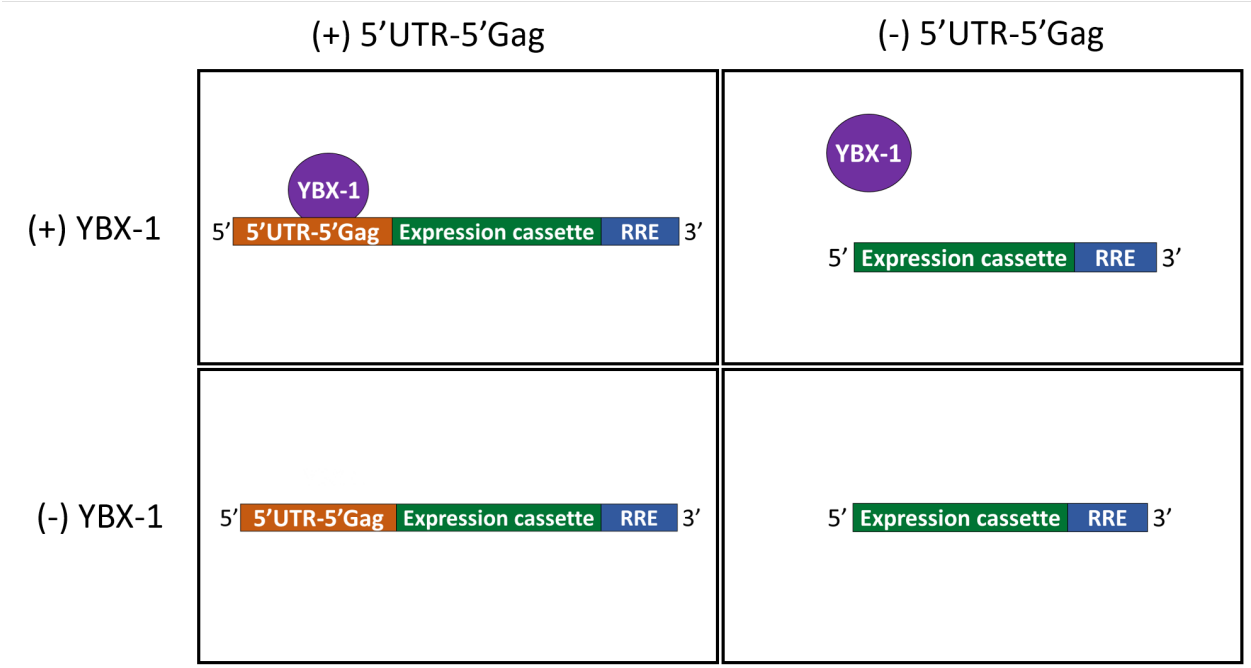
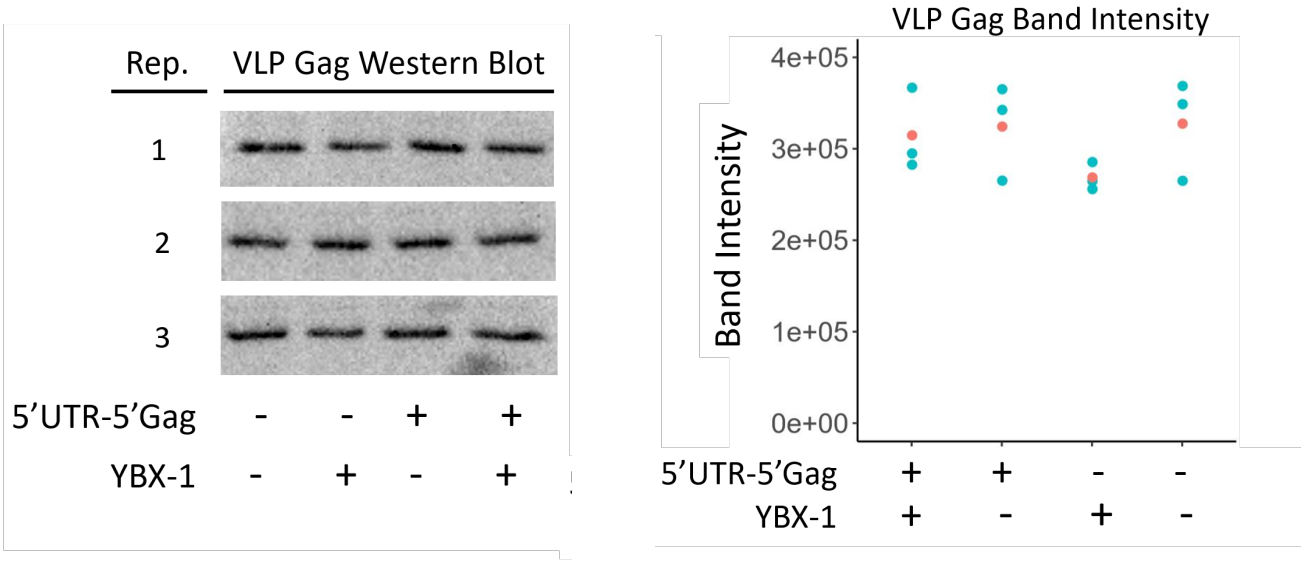


Figure 10: Trials used to determine the role of YB-1 and the 5'UTR-5'Gag RNA in selective RNA packaging into VLPs. In addition to expressing Gag, yeast expressed one of four presented combinations of YB-1 and the VLP RNA expression cassette. These strains were used to evaluate the role of YB-1 and the 5'UTR-5'Gag RNA in selective RNA packaging.



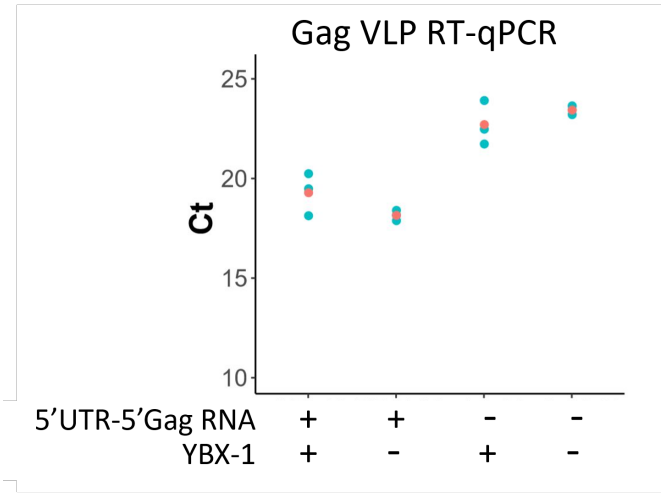
A B

Figure 11: Gag VLP yield from quantitative RNA packaging experiment. Equal aliquots of resuspended VLP pellets were subject to both western blot analysis (presented here) or RT-qPCR (presented later). The purpose of western blot analysis was to ensure that differences in RT-qPCR data were due to either the viral 5'UTR-5'Gag RNA or human YB-1 protein rather than differences in Gag VLP recovery (i.e., experimental error). (A) Western blot detecting the Gag polyprotein released in VLPs. Ni-HRP was used to detect the C-terminal His-tag on Gag. All bands correspond to ~55 kDa, the expected size of Gag. (B) Band intensity from the adjacent blot was quantified using Image Lab Software (Bio-Rad). A t-test showed no significant differences between treatments ($p>0.19$). Mean values are depicted in red.

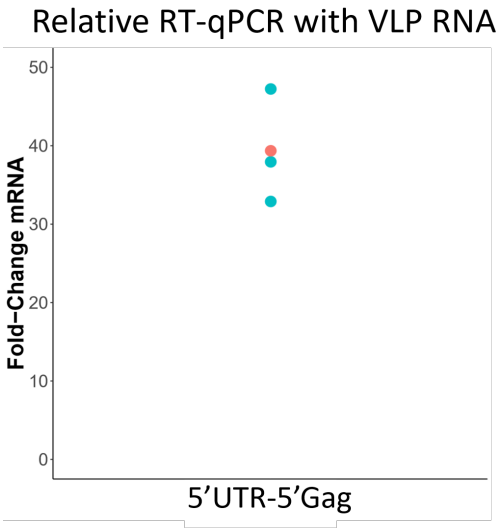
The role of viral RNA elements in selective packaging is widely studied, while no studies have examined the role of YB-1 in selective packaging, despite the ability of YB-1 to bind the HIV-1 5'UTR (Mu et al., 2013). We conducted RT-qPCR to quantitate the effects of the 5'UTR-5'Gag RNA and YB-1 in selective packaging of the VLP RNA expression cassette into VLPs. Primers were designed to amplify the last 300 bp on the 3' end of cDNA transcribed from the VLP RNA expression cassette. The forward primer targeted the HIV-1 RRE, and the

CDS-R reverse primer was shown in Figure 7 to specifically target cDNA. The primer pair was evaluated for specificity and efficient amplification with agarose gel electrophoresis. No-template controls and No-RT controls were evaluated by RT-qPCR to ensure false amplification or contamination would not confound data from experimental replicates. The entire experiment was conducted in triplicate. Based upon standard curve analysis, RT-qPCR reaction efficiency ranged from 93% to 94.7% (mean efficiency=93.7%), and r^2 values ranged from 0.99 to 0.911 (mean r^2 > 0.95). Ct values corresponding to the VLP RNA expression cassette packaged into Gag VLPs were recorded (Figure 12A). Using these Ct values, we found roughly 38-fold enrichment of the VLP RNA expression cassette in VLPs when the 5'UTR-5'Gag RNA was present compared to when it was absent (Figure 12B). This result demonstrates the importance of the 5'UTR-5'Gag RNA sequences in packaging non-viral RNA into VLPs. YB-1 had a negligible impact on selective packaging; The VLP RNA expression cassette was packaged into VLPs at a mean fold-increase of 0.5 when YB-1 was expressed compared to when it was absent (Figure 12C). YB-1 expression in the particular yeast strain used in the experiment was verified by western blot analysis (Figure 12D). Since our experiment relied on heterologous expression of YB-1 in yeast, these observations may be due to a non-functional YB-1 rather than a negligible role of YB-1 in RNA packaging (see Discussion section).

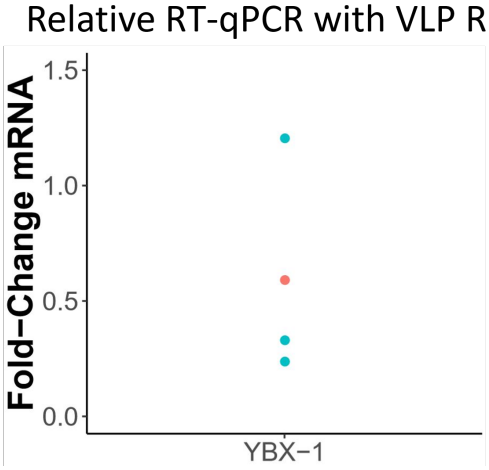
A



B



C



D

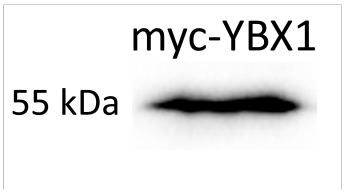


Figure 12: Impact of YB-1 and 5'UTR-5'Gag RNA on selective packaging of a mammalian RNA expression cassette into VLPs produced in *Saccharomyces cerevisiae*. RT-qPCR was used to quantitate the VLP RNA expression cassette packaged into purified VLPs. Both the VLP budding assay and the RT-qPCR were conducted with three independent replicates. Means depicted in red. (A) Ct values determined by RT-qPCR targeting the VLP RNA expression cassette packaged into Gag VLPs. (B) Relative RT-qPCR revealed a 38-fold enrichment of the VLP RNA expression cassette in Gag VLPs when the 5'UTR-5'Gag RNA was upstream the mammalian RNA expression cassette compared to when it was absent. (C) Relative RT-qPCR showed an average of 0.5-fold enrichment of the VLP RNA expression cassette in Gag VLPs when YB-1 was expressed compared to when it was absent. (D) Western blot verifying YB-1 expression in the yeast strain used in the experiment investigating the role of YB-1 in selective packaging.

Previous research has suggested that YB-1 interacts with Gag in an RNA-dependent manner (Bann et al. 2014). The YB-1 responsive element in HIV-1 RNA has been shown to be the stem loop 2 of the 5'UTR (Mu et al., 2013). Having not observed a difference in selective packaging in the presence of YB-1, we conducted a ribonucleoprotein (RNP) pull-down assay to determine if any interaction between YB-1 and the viral sequences of the VLP RNA expression cassette could be detected. Lysate from yeast expressing YB-1 and the VLP RNA expression cassette was incubated with anti-c-myc agarose beads to allow binding of mycYB-1, and RT-PCR was used to detect the VLP RNA expression cassette in both the unbound and bound fraction. The VLP RNA expression cassette was detected in the unbound fraction, with only background levels bound to the anti-c-myc resin (Figure 13). We conducted the same experiment using the mycYB1-6xHis-tag construct with Ni-NTA and found similar results (data not shown). It should be noted that western blot analysis was not used to verify YB-1 had successfully been pulled down. These results indicate that heterologous YB-1 expressed in yeast does not bind HIV-1 RNA sequences in the 5'UTR, despite what has been reported in mammalian cells.

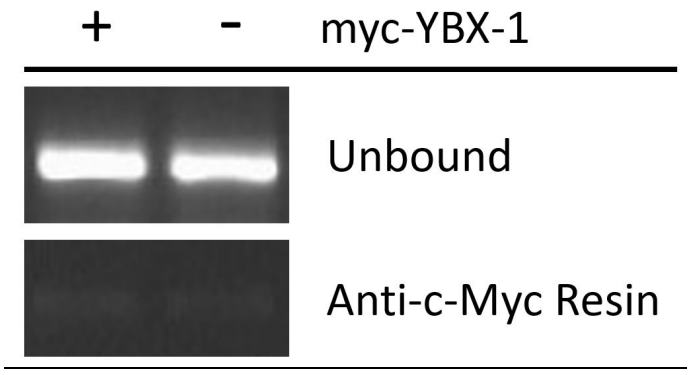


Figure 13: Ribonucleoprotein (RNP) pull-down assay to detect interactions between YB-1 and the 5'UTR-5'Gag RNA sequences of HIV-1. The N-terminal Myc tag on YB-1 was used to bind YB-1 to an anti-c-Myc resin, and RT-PCR was used to detect the target RNA in both bound and unbound fractions. Data representative of two independent replicates.

4. DISCUSSION

Our study identified viral RNA elements important in Gag-mediated selective packaging of a mammalian RNA expression cassette into membrane-bound VLPs produced in yeast. VLP production was shown to be dependent on HIV-1 Gag polyprotein expression in yeast spheroplasts. Packaging was determined to be Gag-mediated because it was dependent upon co-expression of the VLP RNA expression cassette alongside Gag . Packaging of the VLP RNA expression cassette was selective to the extent that the highly abundant, cytosolic actin mRNA was not detected in VLPs. This VLP RNA expression cassette was determined to be enclosed within a membrane because it was only sensitive to treatment with RNase after a detergent was added. Comparison of packaging efficiency of the VLP RNA expression cassette both with and without the 5'UTR-5'Gag RNA sequence revealed that inclusion of this viral RNA sequence upstream of the VLP RNA expression cassette greatly increased its abundance in VLPs. Our study did not find evidence that heterologous expression of the human YB-1 protein in yeast facilitates selective packaging of a VLP RNA expression cassette, and this was likely due to the inability of YB-1 to bind viral RNA sequences present in the expression cassette.

A previous report has demonstrated selective packaging of viral genomic RNA into Gag VLPs produced in yeast (Tomo et al., 2013). These authors showed that selective packaging was greatly dependent on the R-U5-SL region of HIV-1 RNA, and that HIV-1 genomic RNA was present in VLPs at a 500-fold increase compared to its presence in cells. Because our research is aimed at producing recombinant retroviral vectors that could be used in gene therapy, we were interested in only including the minimal HIV-1 RNA sequence necessary to facilitate selective packaging. A recent study utilizing mammalian cells showed that in addition to the 5'UTR, which contains the HIV-1 packaging signal, the 5' half the Gag is important in packaging reporter genes into VLPs (Liu et al. 2017). Given the indirect impact of downstream sequences on RNA packaging, it would be interesting to know the effect of our mammalian RNA

expression cassette on selective RNA packaging when compared to the full-length HIV-1 genomic RNA.

The cytomegalovirus (CMV) promoter included in our VLP RNA expression cassette, while intended to drive transcription of reporter genes in mammalian cells, has actually been shown to be active in yeast (Becskei et al. 2001). Due to the location of the CMV promoter in our VLP RNA expression cassette, it is possible that mRNA transcripts that lacked the 5'UTR-5'Gag RNA were generated from the plasmid containing, and intended to express, the 5'UTR-5'Gag RNA. We have not found any literature on the activity of the simian vacuolating virus 40 (SV40) polyadenylation site in yeast. In addition to amplifying the expected 1.25 kb fragment with RT-PCR targeting the VLP RNA expression cassette, we consistently observed amplification of a smaller band which could have corresponded to one of the above scenarios. It is unknown how these possibilities could have affected our RT-qPCR results that aimed to quantitate the impact of the 5'UTR-5'Gag RNA on selective packaging, since the forward primer targeted the downstream Rev Response Element (RRE) in the VLP RNA expression cassette. While the HIV-1 RRE has been shown to bind Gag (Carlson et al. 2016), it was surprising to see the VLP RNA expression cassette that lacked the 5'UTR-5'Gag RNA yet contained the RRE was packaged into VLPs, albeit with much less affinity than when it was present.

The HIV-1 genomic RNA has been shown to support the assembly of Gag intermediates at the plasma membrane (Yang et al. 2018). Additionally, evidence suggests that the localization of gag-pol mRNAs regulate the site of virion assembly (Becker and Sherer, 2017). It would be interesting to know if the viral RNA sequence expressed alongside Gag in this study facilitated VLP assembly and release, perhaps by nucleating Gag assembly and facilitating Gag-Gag multimerization either in the cytosol or at the plasma membrane.

A candidate human protein we examined that may participate in selective packaging of HIV-1 RNA into VLPs was the human Y-box-binding protein 1 (YB-1). Our attention was first

drawn to YB-1 because its involvement in recruiting microRNAs into exosomes (Shurtleff et al. 2016). Retention of Gag in endosomes has been reported in mammalian cells when artificially high levels of microRNAs are present (Chen et al. 2014). Similarly, using previously published fluorescence microscopy images of Gag-GFP localization during assembly (Norgan et al. 2012), we observed that Gag VLPs in yeast appeared to be retained in endosomes. We reasoned that yeast are missing a homolog of a critical host protein required for normal HIV-1 assembly, particularly by facilitating Gag-RNA interactions. Interestingly, there is no YB-1 homolog in the yeast genome. Our excitement grew as we read reports of direct involvement of YB-1 in HIV-1 replication via RNA interactions. For instance, YB-1 was shown to interact directly with the HIV-1 genomic RNA (Knoener et al. 2017), supporting previous studies that mapped the YB-1-responsive elements to be both the the SL2 (Mu et al. 2013) and TAR (Ansari et al., 1999) located in the 5'UTR of unspliced HIV-1 genomic RNA. Furthermore, YB-1 was shown to increase VLP production of retroviral vector systems via RNA interactions (Bann et al. 2014; Mu et al. 2013).

We did not observe a functional role of YB-1 in selective packaging of RNA into Gag VLPs. Despite the observed 0.5-fold increase in selective RNA packaging in the presence of YB-1, this small increase lies within the range of experimental error. Furthermore, the ribonucleoprotein (RNP) pull-down assay did not detect an interaction between viral sequences on the VLP RNA expression cassette and YB-1, despite how YB-1 has been previously shown to bind the 5'UTR in mammalian cells (Mu et al., 2013). The most likely explanation for the observed differences in YB-1 protein function in yeast compared to mammalian cells is that post-translational modifications to YB-1 required for functionality in HIV-1 assembly are not occurring. In particular, these modifications may affect protein folding that are required for function. There are numerous residues in YB-1 that are phosphorylated in mammalian cells that may affect protein folding required for function (reviewed in Prabhu et al. 2015b). The cold-shock domain (CSD; amino acid residues 51-129) of YB-1 is the region responsible for RNA-

binding, which may impact binding of YB-1 to HIV-1 RNA. Activated Akt phosphorylates YB-1 at Ser102 (Dunn et al, 2005). Sch9 (AGC family kinase) is a yeast homolog to Akt (Fabrizio et al. 2001). Ser165 of YB-1 is also known to be phosphorylated, and it is hypothesized to occur by casein kinase II (CKII) (Prabhu et al. 2015a). Yeast contain the casein kinase 2 catalytic subunit (CKA2). It remains to be seen if YB-1 is modified by these homologs in a way that render it functional, or if upstream activation of these kinases modulating activity of YB-1 is occurring in yeast. Current research in the Black lab is attempting to change the primary structure of YB-1 to hopefully mimic folding that occurs when various serine residues are phosphorylated.

Conversely, there could be post-translational modifications to YB-1 rendering the protein *in*active occurring in yeast. For example, poly-ubiquitination of YB-1 at the sites of Gag VLP assembly (i.e., the plasma membrane) would trigger multivesicular body (MVB) biogenesis via nearby ESCRT machinery, leading to degradation of YB-1 in the proteasomal pathway. If poly-ubiquitination of YB-1 is occurring, this may explain some of our difficulty in detecting a strong signal for YB-1 on western blots analyzing total yeast lysate, compared to other exogenous proteins expressed under the same promoter (data not shown).

Since YB-1 is a known constituent of stress granules, P-bodies, exosomes and, at times, imported into the nucleus (all reviewed in Eliseeva et al. 2011), the HIV-1 RNA in the cytosol may not be accessible to YB-1 in yeast. Furthermore, YB-1 binding to endogenous yeast proteins not present in mammalian cells may hide the active site required for function in HIV-1 assembly. Lastly, the interaction between YB-1 and HIV-1 RNA may be bridged by another, unknown, protein or part of a larger protein complex that is present in mammalian cells but not in yeast.

REFERENCES

Aiuti, A., Biasco, L., Scaramuzza, S., Ferrua, F., Cicalese, M.P., Baricordi, C., Dionisio, F., Calabria, A., Giannelli, S., Castiello, M.C., et al. 2013. Lentiviral Hematopoietic Stem Cell Gene Therapy in Patients with Wiskott-Aldrich Syndrome. *Science* 341, 1233151.

Amado, R.G., and Chen, I.S.Y. 1999. Lentiviral Vectors--the Promise of Gene Therapy Within Reach? *Science* 285, 674.

Andreola, M.-L., and Lityak, S. 2012. Yeast and the AIDS Virus: The Odd Couple. *J. Biomed. Biotechnol* e549020.

Ansari, S.A., Safak, M., Gallia, G.L., Sawaya, B.E., Amini, S., and Khalili, K. 1999. Interaction of YB-1 with human immunodeficiency virus type 1 Tat and TAR RNA modulates viral promoter activity. *J. Gen. Virol.* 80(10), 2629–2638.

Balasubramaniam, M. and Freed, E.O. 2011. New Insights into HIV assembly and Trafficking. *Physiology* 26(4):236-251.

Bann DV, Beyer AR, Parent LJ. 2014. A Murine Retrovirus Co-Opts YB-1, a Translational Regulator and Stress Granule-Associated Protein, To Facilitate Virus Assembly. Ross SR, ed. *Journal of Virology* 88(8):4434-4450.

Bathurst, I.C., Chester, N., Gibson, H.L., Dennis, A.F., Steimer, K.S., and Barr, P.J. 1989. N myristylation of the human immunodeficiency virus type 1 gag polyprotein precursor in Saccharomyces cerevisiae. *J. Virol.* 63, 3176–3179.

Becker, J.T., and Sherer, N.M. 2017. Subcellular Localization of HIV-1 *gag-pol* mRNAs Regulates Sites of Virion Assembly. *Journal of Virology* 91(6) 1-22.

Becskei, A., Seraphin, B., Serrano, L. 2001. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO* 20(10): 2528-2535.

Biffi, A., Montini, E., Lorioli, L., Cesani, M., Fumagalli, F., Plati, T., Baldoli, C., Martino, S., Calabria, A., Canale, S., et al. 2013. Lentiviral Hematopoietic Stem Cell Gene Therapy Benefits Metachromatic Leukodystrophy. *Science* 341.

Botstein, D., and Fink, G.R. 1988. Yeast: an experimental organism for modern biology. *Science* 240, 1439–1443.

Briggs JAG, Riches JD, Glass B, Bartonova V, Zanetti G, Kräusslich H-G. 2009. Structure and assembly of immature HIV. *Proceedings of the National Academy of Sciences of the United States of America* 106(27):11090-11095.

Brugger, B. *et al. 2006.* The HIV lipidome: a raft with an unusual composition. *Proc. Natl Acad. Sci. USA* 103, 2641–2646.

Bryant M, Ratner L. 1990. Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proc Natl Acad Sci U S A.* 87(2):523-7

Campbell, S., and Rein, A. 1999. In Vitro Assembly Properties of Human Immunodeficiency Virus Type 1 Gag Protein Lacking the p6 Domain. *Journal of Virology* 73(3): 2270-2279.

Carlson, L.A., Bai, Y., Keane, S.C., Doudna, J.A., Hurley, J.H. 2016. Reconstitution of selective HIV-1 RNA packaging in vitro by membrane-bound Gag assemblies. *eLife* 5:e14663

Chen, N.N., and Khalili, K. 1995. Transcriptional regulation of human JC polyomavirus promoters by cellular proteins YB-1 and Pur alpha in glial cells. *J. Virol.* 69, 5843–5848.

Chen, A.K., Sengupta, P., Waki, K., Engelenburg, S.B.V., Ochiya, T., Ablan, S.D., Freed, E.O., and Lippincott-Schwartz, J. 2014. MicroRNA binding to the HIV-1 Gag protein inhibits Gag assembly and virus production. *Proc. Natl. Acad. Sci.* 111, E2676–E2683.

Chertova E, Chertov O, Coren LV, Roser JD, Trubey CM, Bess JW Jr, Sowder RC 2nd, Barsov E, Hood BL, Fisher RJ, Nagashima K, Conrads TP, Veenstra TD, Lifson JD, Ott DE. 2006. Proteomic and biochemical analysis of purified human immunodeficiency virus type 1 produced from infected monocyte-derived macrophages. *J Virol.* 80:9039–9052.

Cochrane AW, McNally MT, Mouland AJ. 2006. The retrovirus RNA trafficking granule: from birth to maturity. *Retrovirology* 3:18.

D'Agostino DM, Felber BK, Harrison JE, Pavlakis GN. 1992. The Rev protein of human immunodeficiency virus type 1 promotes polysomal association and translation of gag/pol and vpu/env mRNAs. *Molecular and Cellular Biology* 12(3):1375-1386.

Deeks, S.G., Overbaugh, J., Phillips, A., and Buchbinder, S. 2015. HIV infection. *Nat. Rev. Dis. Primer* 1, 15035.

Dong X, Li H, Derdowski A, Ding L, Burnett A, Chen X, Peters TR, Dermody TS, Woodruff E, Wang JJ, Spearman P. 2005. AP-3 directs the intracellular trafficking of HIV-1 Gag and plays a key role in particle assembly. *Cell* 120:663–674.

Dooher JE, Schneider BL, Reed JC, Lingappa JR. 2007. Host ABCE1 is at plasma membrane HIV assembly sites and its dissociation from Gag is linked to subsequent events of virus production. *Traffic* 8:195–211.

Sutherland, B.W., et al. 2005 Akt phosphorylates the Y-box binding protein 1 at Ser102 located in the cold shock domain and affects the anchorage-independent growth of breast cancer cells. *Oncogene* 24(26):4281-92.

Eliseeva, I.A., Kim, E.R., Guryanov, S.G., Ovchinnikov, L.P., and Lyabin, D.N. 2011. Y-box-binding protein 1 (YB-1) and its functions. *Biochem. Biokhimiia 76,* 1402–1433.

Fabrizio et al. 2001. Regulation of longevity and stress resistance by Sch9 in yeast. *Science* 292(5515):288-290.

Freed, E.O. 2015. HIV-1 assembly, release and maturation. *Nat. Rev. Microbiol. 13*, 484–496.

Friedmann, T., and Roblin, R. 1972. Gene Therapy for Human Genetic Disease? *Science* 175, 949–955.

Göttlinger, H.G., Sodroski, J.G., and Haseltine, W.A. 1989. Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. U. S. A.* 86, 5781–5785.

Grigorov B, Arcanger F, Roingeard P, Darlix JL, Muriaux D. 2006. Assembly of infectious HIV-1 in human epithelial and T-lymphoblastic cell lines. *J Mol Biol.* 359:848–862.

Hacein-Bey-Abina, S., Pai, S.-Y., Gaspar, H.B., Armant, M., Berry, C.C., Blanche, S., Bleesing, J., Blondeau, J., de Boer, H., Buckland, K.F., et al. 2014. A Modified γ-Retrovirus Vector for X-Linked Severe Combined Immunodeficiency. *N. Engl. J. Med. 371, 1407*–1417.

Holm, P.S., Bergmann, S., Jurchott, K., Lage, H., Brand, K., Ladhoff, A., Mantwill, K., Curiel, D.T., Dobbelstein, M., Dietel, M., et al. 2002. YB-1 relocates to the nucleus in adenovirus-infected cells and facilitates viral replication by inducing E2 gene expression through the E2 late promoter. *J. Biol. Chem.* 277, 10427–10434.

Ivanchenko S, Godinez WJ, Lampe M, Krausslich HG, Eils R, Rohr K, Brauchle C, Muller B, Lamb DC. 2009. Dynamics of HIV-1 assembly and release. *PLoS Pathog.* 5:e1000652.

Jouvenet N, Neil SJD, Bess C, et al. 2006. Plasma Membrane Is the Site of Productive HIV-1 Particle Assembly. Emerman M, ed. *PLoS Biology* 4(12):e435.

Kawaguchi A, Matsumoto K, Nagata K. 2012. YB-1 functions as a porter to lead influenza virus ribonucleoprotein complexes to microtubules. *J Virol.* 86(20):11086–95.

Kerr, D., Chang, C.F., Chen, N., Gallia, G., Raj, G., Schwartz, B., and Khalili, K. 1994. Transcription of a human neurotropic virus promoter in glial cells: effect of YB-1 on expression of the JC virus late gene. *J. Virol.* 68, 7637–7643.

Kim SY, Byrn R, Groopman J, Baltimore D. 1989. Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. *J Virol.* 63(9):3708-13

Knoener RA, Becker JT, Scalf M, Sherer NM, Smith LM. 2017. Elucidating the *in vivo* interactome of HIV-1 RNA by hybridization capture and mass spectrometry. *Sci Rep.* 7(1):16965.

Landau NR, Warton M, Littman DR. 1988. The envelope glycoprotein of the human immunodeficiency virus binds to the immunoglobulin-like domain of CD4. *Nature* 334(6178):159-62

Li, W., Wang, X., and Gao, G. 2012. Expression of YB-1 enhances production of murine leukemia virus vectors by stabilizing genomic viral RNA. *Protein Cell 3*, 943–949.

Li, Y., Frederick, K.M., Haverland, N.A., Ciborowski, P., and Belshan, M. 2016. Investigation of the HIV-1 matrix interactome during virus replication. *Proteomics Clin. Appl. 10*, 156–163.

Lingappa, J.R., Reed, J.C., Tanaka, M., Chutiraka, K., Robinson B.A. 2014. How HIV-1 Gag assembles in cells: Putting together pieces of the puzzle. *Virus Research* 193:89-107.

Liu, Y., Nikolaitchik, O.A., Rahman, S.A., Chen, J., Pathak, V.K., Hu, W.S. 2017. HIV-1 Sequence Necessary and Sufficient to Package Non-viral RNA into HIV-1 Particles. *J. Mol. Biol.* 429(16):2542-2555.

Llewellyn GN, Hogue IB, Grover JR, Ono A. 2010. Nucleocapsid promotes localization of HIV-1 gag to uropods that participate in virological synapses between T cells. *PLoS Pathog*. 6:e1001167.

Lopez-Verges, S., Camus, G., Blot, G., Beauvoir, R., Benarous, R., Berlioz-Torrent, C. 2006. Tail-interacting protein TPI47 is a connector between Gag and Env and is required for Env incorporation into HIV-1 virions. *Proc. Natl Acad Sci* 103:14947-14952.

Martinez NW, Xue X, Berro RG, Kreitzer G, Resh MD. 2008. Kinesin KIF4 regulates intracellular trafficking and stability of the human immunodeficiency virus type 1 Gag polyprotein. *J Virol.* 82:9937–9950.

Murakami, T. & Freed, E. O. 2000. Genetic evidence for an interaction between human immunodeficiency virus type 1 matrix and α-helix 2 of the gp41 cytoplasmic tail. *J. Virol.* 74, 3548–3554.

Merten, O. 2004. State-Of-The-Art of the Production of Retroviral Vectors. *J Gene Med.* 6 Suppl 1:S105-24.

Mu, X., Li, W., Wang, X., and Gao, G. 2013. YB-1 stabilizes HIV-1 genomic RNA and enhances viral production. *Protein Cell 4*, 591–597.

Nathwani, A.C., Reiss, U.M., Tuddenham, E.G.D., Rosales, C., Chowdary, P., McIntosh, J., Della Peruta, M., Lheriteau, E., Patel, N., Raj, D., et al. 2014. Long-Term Safety and

Efficacy of Factor IX Gene Therapy in Hemophilia B. *N. Engl. J. Med. 371*, 1994–2004.

Norgan, A.P., Lee, J.R.E., Oestreich, A.J., Payne, J.A., Krueger, E.W., and Katzmann, D.J. 2012. ESCRT-Independent Budding of HIV-1 Gag Virus-Like Particles from Saccharomyces cerevisiae Spheroplasts. *PLOS ONE 7*, e52603.

Ono A, Freed EO. 2004. Cell-type-dependent targeting of human immunodeficiency virus type 1 assembly to the plasma membrane and the multivesicular body. *J Virol.* 78:1552–1563.

Ono, A., Ablan, S. D., Lockett, S. J., Nagashima, K. & Freed, E. O. 2004. Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. *Proc. Natl Acad. Sci. USA* 101, 14889–14894.

Ono, A. & Freed, E. O. 2001. Plasma membrane rafts play a critical role in HIV-1 assembly and release. *Proc. Natl Acad. Sci.* USA 98, 13925–13930.

Paranjape SM, Harris E. 2007. Y box-binding protein-1 binds to the dengue virus 3' untranslated region and mediates antiviral effects. *J Biol Chem.* 282(42):30497–508.

Pelchen-Matthews A, Kramer B, Marsh M. 2003. Infectious HIV-1 assembles in late endosomes in primary macrophages. *J Cell Biol.* 162:443–455

Persons, D.A. 2010. Gene therapy: Targeting β-thalassaemia. *Nature 467*, 277–278.

Prabhu et al. 2015a. Critical role of phosphorylation of serine 165 of YBX1 on the activation of NF-κB in colon cancer. *Oncotarget* 6(30):29396-29412

Prabhu et al. 2015b. Role of post-translational modifications of the Y box binding protein 1 in human cancers. *Genes and Diseases* 2:240-246.

Rudner L, Nydegger S, Coren LV, Nagashima K, Thali M, Ott DE. 2005. Dynamic fluorescent imaging of human immunodeficiency virus type 1 gag in live cells by biarsenical labeling. *Journal of Virology* 79:4055–4065.

Sabo Y, Walsh D, Barry DS, et al. 2013. HIV-1 induces the formation of stable microtubules to enhance early infection. *Cell host & microbe* 14(5):10.1016/j.chom.2013.10.012.

Sakuragi, S., Goto, T., Sano, K., and Morikawa, Y. 2002. HIV type 1 Gag virus-like particle budding from spheroplasts of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. 99*, 7956–7961.

Sherer NM, Lehmann MJ, Jimenez-Soto LF, Ingmundson A, Horner SM, Cicchetti G, Allen PG, Pypaert M, Cunningham JM, Mothes W. 2003. Visualization of retroviral replication in living cells reveals budding into multivesicular bodies. *Traffic* 4:785–801.

Shkriabai, N. et al. 2006. Interactions of HIV‑1 Gag with assembly cofactors. *Biochemistry* 45, 4077–4083.

Shurtleff, M.J., Temoche-Diaz, M.M., Karfilis, K.V., Ri, S., and Schekman, R. 2016. Y-box protein 1 is required to sort microRNAs into exosomes in cells and in a cell-free reaction. *eLife 5*, e19276.

Simonelli, F., Maguire, A.M., Testa, F., Pierce, E.A., Mingozzi, F., Bennicelli, J.L., Rossi, S., Marshall, K., Banfi, S., Surace, E.M., et al. 2010. Gene therapy for Leber's congenital amaurosis is safe and effective through 1.5 years after vector administration. *Mol. Ther. J. Am. Soc. Gene Ther. 18*, 643–650.

Tang C, Loeliger E, Luncsford P, Kinde I, Beckett D, Summers MF. 2004. Entropic switch regulates myristate exposure in the HIV-1 matrix protein. *Proceedings of the National Academy of Sciences of the United States of America* 101(2):517-522.

Tomo, N., Goto, T., and Morikawa, Y. 2013. Trans-packaging of human immunodeficiency virus type 1 genome into Gag virus-like particles in *Saccharomyces cerevisiae*. *Microb. Cell Factories 12*, 28.

UNAIDS. 2016. AIDS by the numbers — AIDS is not over, but it can be.

Weydert C, van Heertum B, Dirix L, De Houwer S, De Wit F, Mast J, et al. 2018. Y-box-binding protein 1 supports the early and late steps of HIV replication. *PLoS ONE* 13(7): e0200080.

Yang, Y., Qu, N., Tan, J., Rushdi, M.N., Krueger, C.J., Chen, A. 2018. Roles of Gag-RNA interactions in HIV-1 virus assembly deciphered by single-molecule localization microscopy. *PNAS*, In Review

Yi R, Bogerd HP, Cullen BR. 2002. Recruitment of the Crm1 nuclear export factor is sufficient to induce cytoplasmic expression of incompletely spliced human immunodeficiency virus mRNAs. *Journal of Virology* 76:2036–2042.

Zhao RY. 2017. Yeast for virus research. *Microbial Cell* 4(10):311-330.

Zhou W, Parent LJ, Wills JW, Resh MD. 1994. Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. *Journal of Virology* 68(4):2556-2569.

Zhou W, Resh MD. 1996. Differential membrane binding of the human immunodeficiency virus type 1 matrix protein. *Journal of Virology* 70(12):8540-8548.

APPENDIX

i. CRISPR-Cas9 genome integration of Gag-eGFP in *Saccharomyces cerevisiae*

Our first effort to express the HIV-1 Gag polyprotein in *Saccharomyces cerevisiae* utilized CRISPR-Cas9 genome integration through homology-directed repair. The lithium acetate transformation was used to transform SEY6210 with the following three plasmids: (1) pRS414-Cas9 (provided by George Church; Addgene plasmid # 43802), a construct that constitutively expressed *Streptococcus pyogenes* Cas9 under a yeast translational elongation factor EF-1 alpha (TEF1) promoter; (2) pRPR-gRNA (created by Ben Lewis, CIRM M.S. thesis), a construct created that expressed a 20-bp protospacer RNA targeting the PRS4 locus; (3) pHDR-GagEGFP (previously created by undergraduates in the Black lab), a construct that contained Gag-eGFP under a TEF1 promoter and yeast cytochrome c (CYC1) terminator flanked by 5' and 3' homology arms complementary to the regions of the double-stranded break (DSB) in the PRS4 locus. Transformants were plated on SD-LEU/-TRP media, and colony PCR using primers Gag-*Eco*RI-F and PRS4-Ext-R was used to identify CFUs with GagEGFP integration at the PRS4 locus.

The TEF1-Gag-eGFP-CYC1 expression cassette was integrated into the yeast genome at the PRS4 locus, as determined by PCR using genomic DNA of a CRISPR transformant (Figure 14A, 14B). However, the Gag-eGFP protein was not detected on a western blot using anti-GFP antibodies, despite the detection a GFP positive control (Figure 14C). Because the matrix (MA) domain of Gag is involved plasma membrane binding, separate fractions of small, water soluble proteins (i.e. the supernatant from whole-cell lysate) and proteins pelleted with cellular debris including membranes (i.e. the pellet from whole-cell lysate) were both analyzed by western blot.

As an alternative approach to determine if the Gag-eGFP fusion protein was expressed, CRISPR transformants were viewed with fluorescent microscopy. The enhanced GFP protein was not detected with fluorescent microscopy, corroborating the western blot results that the Gag-eGFP fusion protein was not being expressed (data not shown).

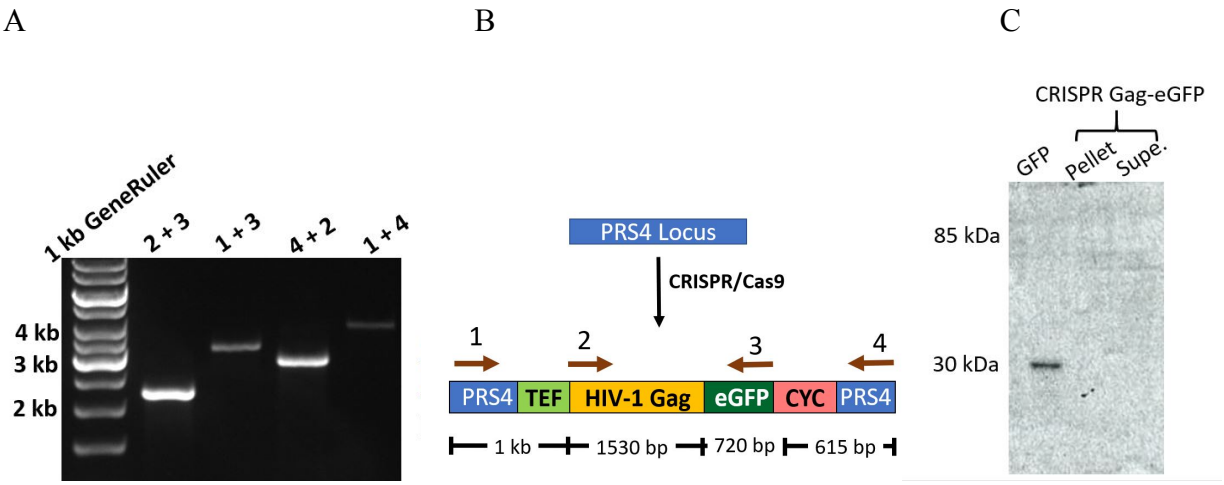


Figure 14: *CRISPR-Cas9 genome integration of GagEGFP under a TEF1 promoter and upstream of a CYC1 terminator, inserted into the PRS4 locus of the yeast genome.* (A) PCR using genomic DNA from a yeast CRISPR transformant. Primer sequences are listed in Materials and Methods section; Numerical representations: (1) PRS4-Ext-F (2) Gag-*Eco*RI-F (3) GagEGFP-*Sal*I-R (4) PRS4-Ext-R. (B) Diagram showing template and expected amplicon sizes after CRISPR/Cas9 insertion of TEF1-Gag-eGFP-CYC1 expression cassette into the PRS4 locus. (C) Western blot analysis using anti-GFP to detect expression of Gag-eGFP from a CRISPR transformant.

Despite CRISPR-Cas9 integration of the TEF1-Gag-eGFP-CYC1 expression cassette into the PRS4 locus of the yeast genome, we did not observe Gag-eGFP protein expression. Similarly, Gag-eGFP amplified from the mammalian pGag-eGFP construct was cloned into a yeast expression vector downstream of a yeast TPI promoter and, again, expression was not observed by western blot (data not shown). Gag-GFP has been expressed in yeast with no issue regarding cleavage of Gag away from GFP (Norgan et al. 2012) and, regardless, we did not observe any GFP expression. Since the same yeast vector backbone and pGag-eGFP template

was eventually used to successfully express the Gag-6xHis-tag protein used in VLP production, the most likely reason for failed expression of Gag-eGFP is an undefined error in the eGFP gene or sequence linking Gag to eGFP. For instance, a stop codon between Gag and eGFP could explain why a functional Gag coding sequence was cloned into a yeast vector, but Gag fused to a C-terminal GFP was not expressed. Additionally, a frameshift mutation after the Gag coding sequence would lead to the same result. Since antibodies raised against Gag were never used, so we cannot say with certainty that a functional Gag protein was expressed from the Gag-eGFP construct; However, we sequenced the Gag coding region in the yeast expression vector and found the sequence to be correct.

ii. Sequencing Results

The coding region of HIV-1 Gag with a C-terminal 6xHisTag, cloned into pRS416 and used extensively throughout this study, was verified by sequencing at a contract lab (GENEWIZ). The sequence in 5' to 3' polarity is as follows:

ATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGATGGGAAAAAAT
TCGGTTAAGGCCAGGGGGAAAGAAGAAGTACAAGCTAAAGCACATCGTATGGGCA
AGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGG
CTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAGGAGC
TTCGATCACTATACAACACAGTAGCAACCCTCTATTGTGTGCACCAGCGGATCGAGA
TCAAGGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAGTCCAA
GAAGAAGGCCCAGCAGGCAGCAGCTGACACAGGACACAGCAATCAGGTCAGCCAA
AATTACCCTATAGTGCAGAACATCCAGGGGCAAATGGTACATCAGGCCATATCACCT
AGAACTTTAAATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGT
GATACCCATGTTTTCAGCATTATCAGAAGGAGCCACCCCACAGGACCTGAACACGA
TGTTGAACACCGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATC
AATGAGGAAGCTGCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGC
ACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACTACTAGTACC
CTTCAGGAACAAATAGGATGGATGACAAATAATCCACCTATCCCAGTAGGAGAGAT
CTACAAGAGGTGGATAATCCTGGGATTGAACAAGATCGTGAGGATGTATAGCCCTA
CCAGCATTCTGGACATAAGACAAGGACCAAAAGAACCCTTTAGAGACTATGTAGAC
CGGTTCTATAAAACTCTAAGAGCTGAGCAAGCTTCACAGGAGGTAAAAAATTGGAT
GACAGAAACCTTGTTGGTCCAAAATGCGAACCCAGATTGTAAGACCATCCTGAAGG

CTCTCGGCCCAGCGGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTAGGA
GGACCCGGCCATAAGGCAAGAGTTTTGGCTGAAGCAATGAGCCAAGTAACAAATTC
AGCTACCATAATGATGCAGAGAGGCAATTTTAGGAACCAAAGAAAGATTGTTAAGT
GTTTCAATTGTGGCAAAGAAGGGCACACAGCCAGAAATTGCAGGGCCCCTAGGAAA
AAGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTACTGAGA
GACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCCTACAAGGGAAGGCCAGGGAAT
TTTCTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTCTGG
GGTAGAGACAACAACTCCCCCTCAGAAGCAGGAGCCGATAGACAAGGAACTGTATC
CTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCACAAGTCGACCTGC
AGCCAAGCCACCACCACCACCACTAA

iii. GeneArt Strings DNA Fragment for VLP RNA expression cassette

In creating a yeast vector that expressed the VLP RNA expression cassette, we used GeneArt Strings DNA Fragment service (Thermo Fisher Scientific) to synthesize a linear dsDNA containing the HIV-1 5'-untranslated region (5'UTR), a cytomegalovirus (CMV) promoter and simian vacuolating virus 40 (SV40) Poly(A) signal flanking a *Sal*I site, and the HIV-1 Rev Response Element (RRE). The sequence we ordered is presented in Table 2. (Note: This does not contain the 5' half of *gag* coding sequence downstream of the 5'UTR because that sequence was cloned in later.) Importantly, given the technology used to synthesize this oligonucleotide of considerable length, the manufacturer required certain changes me made in the sequence, which deviated from the "wildtype" version of these genes. Those changes are identified in Table 2, and the prior nucleotide that occupied that position is presented in Table 3.

Vector Element	Sequence (5'→3')	Comments
HIV-1 5' UTR	ggtctctctggttagaccagatctgagcctgggagctctctggctaactagggaacccactgcttaagcctcaataaagc ttgccttgagtgcttcaagtagtgtgtgcccgtctgttgtgtgactctggtaactagagatccctcagacccttttagtc agtgtggaaaatctctagcagtggcgcccgaacagggacctgaaagcgaaagggaaaccagaggagctctctcgacgcag gactcggcttgctgaagcgcgcacggcaagaggcgaggggcggcgactggtgagtacgccaaaaattttgactagcggag gctagaaggagagagatgggtgcga	Sequence from HXB2 reference strain. U3, which contains the viral promoter, has been removed and will be replaced by the TEF1 promoter for constitutive expression in yeast.
CMV promoter	tagttattaatagtaatcaattacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggAaa atggccAgcAtggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaata gggactttccGttgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtacaAGCagtgtatcatatgcc aagtacgccccctattgacgtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggactttc ctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtgga tagcggtttgactcacggggatttccaagtctccTccAcattgacgtcaatgggagtttgttttggcaccaaaatcaacg ggactttccaaaatgtcgtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataa gcagagctggtttagtgaaccgtcagatc	Sequence from Clontech pEGFP-N1 Vector (Catalog #6085-1; GenBank Accession #U55762) GeneArt Strings DNA Fragments recommended modifications in red text.
SalI	GTCGAC	Restriction site intended for cloning reporter genes
SV40 PolyA	tctagatcataatcagccataccacatttgtagaggttttacttgctttaaaaaacctcccacacctcccctgaacctg aaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttatatggttacaaataaagcaatagcatcac aaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtccaaactcatcaatgtatc	Sequence from Clontech pEGFP-N1 Vector (Catalog #6085-1; GenBank Accession #U55762)
HIV-1 Rev Response Element	agtagcacccaccaaggcaaagagaagagtggtgcagagagaaaaaagagcagtgggaataggagctttgttccttgggt tcttgggagcagcaggaagcactatgggcgcagcctcaatgacgctgacggtacaggccTgacaattattgtctggtata gtgcagcagcagaacaatttgctgagggctattgaggcgcaacagcatctgttgcaactcacagtctggggcatcaagca gctccaggcaagaatcctggctgtggaaagatacctaaaggatcaacagctcctggggatttggggttgctctggaaaac tcatttgcaccactgctgtgccttggaatgctagt	Sequence from HXB2 reference strain. GeneArt Strings DNA Fragments recommended modification in red text.

Table 2: GeneArt Strings DNA Fragment (Thermo Fisher Scientific) synthesized for the VLP RNA expression cassette. The entire fragment was synthesized as a single (1,524 bp) linear dsDNA in the following order (5' to 3'): 5'UTR-CMV-*SalI*- SV40Poly(A)-RRE.

Vector Element	Nucleotide position inside that element	Original nucleotide:	Changed to:
CMV promoter	78	T	A
CMV promoter	87	C	A
CMV promoter	90	C	A
CMV promoter	171	A	G
CMP promoter	223-225	TCA	AGC
CMV promoter	435	A	T
CMV promoter	438	C	A
HIV-1 RRE	140	A	T

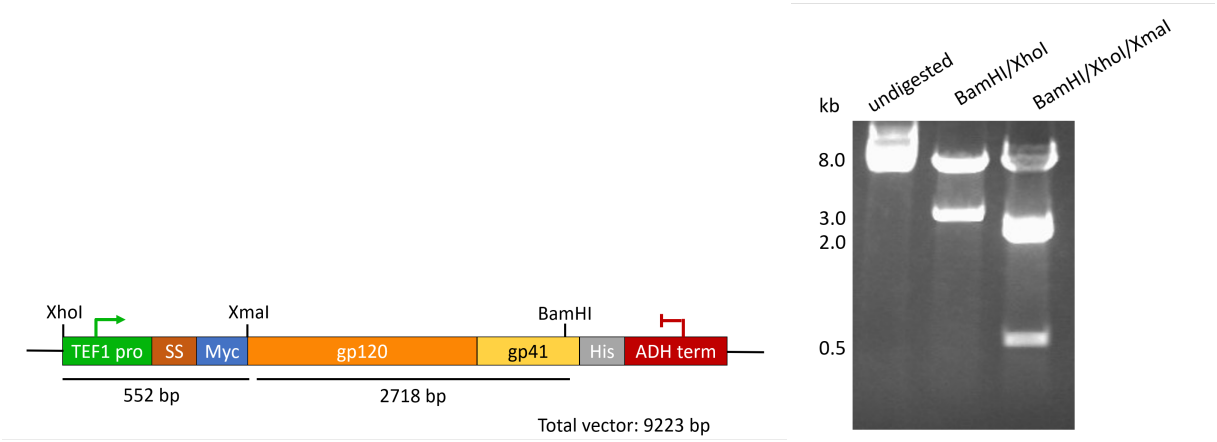
Table 3: Mutations in the VLP RNA expression cassette required by GeneArt Strings DNA Fragments (Thermo Fisher Scientific) for synthesis. Modifications to the CMV promoter occurred in the enhancer region (which spans nucleotides 59-465) and did not occur in either the TATA box (nucleotides 554-560) or Transcription Start Point (583).

iv. Yeast Expression Vector: HIV-1 Envelope (gp160)

Retroviral vectors in used gene therapy must first attach to and fuse with target cells before therapeutic genes can be delivered. In HIV-1 replication, attachment and entry is mediated by the envelope glycoprotein. After gp160 proteolytic cleavage, gp120 binds CD4 receptors on the surface of T cells, allowing coreceptor (CCR5 or CXCR4) binding and subsequent exposure of the fusion peptide of gp41. Having created a yeast system that produces Gag VLPs with selectively packaged RNA, we constructed a yeast expression vector to express the HXB2 envelope (gp160) protein.

The source of the HIV-1 envelope (gp160) coding sequence was pHXB2-Env acquired from the NIH AIDS Reagent Program. The HIV-1 envelope coding sequence was PCR amplified from pHXB2-Env using ENV-*Eco*RI-F and ENV-*Sal*I-R using the thermocycling profile described for Phusion DNA polymerase, except with a 58°C annealing temperature. The 2.6 kb amplicon was gel extracted and digested with *Eco*RI and *Sal*I to create cohesive ends for ligation, as well as *Dpn*I to remove residual pHXB2-Env template. The insert was then ligated into the 5' *Eco*RI-*Sal*I 3' cloning site in the MCS of pRS416 after dephosphorylation of the vector with alkaline phosphatase. After transformation, cloning, and plasmid isolation, the TPI-ENV-HisTag-ADH expression cassette was digested from pRS416 using *Xho*I and *Xba*I. The 3.2 kb expression cassette was gel extracted and ligated into the 5' *Xho*I-*Xba*I 3' cloning site in the MCS of pRS315. Since the envelope protein is trafficked through the secretory pathway, we reasoned that a c-Myc tag upstream of the signal peptide may interfere with ER translocation and protein orientation. Therefore, we inserted the c-Myc tag immediately downstream of the signal sequence in-frame with the N-terminus of gp120. To add an N-terminal c-Myc tag to gp120 after the signal peptide, pRS315 TPI-ENV-HisTag-ADH (described above) was used as a template in the following two PCR reactions: (1) ENV-*Xma*I-F and TS315-F were used to amplify a 2.9 kb fragment (2) ENV-myc-*Xma*I-R and TS315-R were used to amplify a 736 bp fragment. The two fragments were ligated together using a total of ~500 ng DNA in the ligation

and the 3.6 kb fragment corresponding to the TPI- SS-Myc-ENV-HisTag-ADH expression cassette was gel extracted to isolate it from other products from the ligation. PCR with TS315-F and TS315-R was then used to amplify the entire expression cassette for D.I.C. To prepare a vector for D.I.C., the TPI promoter and ENV coding sequence (without the c-Myc tag) of pRS315 TPI-ENV-HisTag-ADH was cut out with *Xho*I and *Sal*I, and the 6.2 kb vector corresponding to pRS315 was gel extracted. D.I.C. was then used to insert the TPI-SS-Myc-ENV-HisTag-ADH expression cassette into the pRS315 using the thermocycling profile outlined above for D.I.C., except with a final 72°C extension time of 5 minutes. The final pRS315 TPI-SS-Myc-ENV-HisTag-ADH construct contains a LEU2 auxotrophic marker so it can be expressed alongside pRS416 TPI-Gag-HisTag-ADH, which carries a URA3 marker. A diagnostic restriction digest revealed that the construct was indeed as expected (Figure 15).



A B

Figure 15: HIV-1 Envelope (gp120) plasmid validation. (A) Diagram of the TPI-SS-Myc-ENV-HisTag-ADH expression cassette cloned into pRS315, including restriction sites used in the diagnostic digest (B) Diagnostic digest of pRS315 TPI-SS-Myc-ENV-HisTag-ADH. The expected band sizes upon incubation with BamHI and EcoRI is 6505 bp and 2718 bp, while additionally including XmaI (which is inserted with the c-Myc tag) generates expected band sizes of 6505 bp, 2166 bp, and 552 bp.