PROTEOMIC ANALYSES OF HIV-1 INFECTED T-CELLS AND THE FUNCTIONAL CHARACTERIZATION OF CYCLOPHILIN B DURING HIV-1 INFECTION

BY

JASON TODD DEBOER

A DISSERTATION

Submitted to the Faculty of the Graduate School of Creighton University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the Department of Medical Microbiology and Immunology

Omaha, Nebraska

May 1, 2015
ABSTRACT

The human immunodeficiency virus type 1 (HIV-1) lifecycle is complex and the virus interacts with many host cellular proteins for productive replication. Mass spectrometry analysis is a powerful tool to identify cellular proteins important for HIV replication and pathogenesis. Prior proteomic studies demonstrate that HIV-1 infection perturbs the host cell proteome.

The nuclear compartment of cells is critical for HIV replication: integration of viral DNA into the host genome is essential to virus replication and pathogenesis; numerous host transcription factors are required for HIV-1 gene expression; and subtle changes to the nuclear proteome either through translocation of cytosolic proteins or induction of regulatory factors can contribute to evasion of the host immune response. However, to date no study has looked specifically at the changes in the host nuclear proteome during HIV infection. My initial studies examined the changes in the nuclear proteome of HIV-1 infected T-cells.

On a global level, the cellular changes in response to virus infection are profound. Cells must attempt to inhibit infection as well as signal the surrounding environment to the presence of the pathogen. The results of both efforts cause changes in the expression and/or localization of host cell factors including signaling molecules, RNAs, and proteins. In the second study we examined the changes in protein expression and localization in HIV-1 infected T-cells by analyzing the cytosolic, membrane/organelle, and nuclear fractions using Sequential Window Acquisition of all Theoretical fragment-ion spectra Mass
Spectrometry (SWATH-MS) proteomic approach. In addition, factors identified as potential HIV-1 interacting proteins were evaluated for biological relevance in a cell culture model.

Cyclophilin B (CypB) is a member of the peptidyl prolyl isomerase family that is an intracellular chaperon for prolactin and IRF3. It predominantly localizes to the ER, but also contains a nuclear localization signal and can be secreted from cells. It has previously been shown to interact with the Gag protein of human immunodeficiency type 1 (HIV-1) and several proteomic and genetic studies identified it as a potential factor involved in HIV replication. In the nuclear proteome study described in Chapter II, we found that nuclear levels of CypB increased during HIV replication. In Chapter IV we show that ectopic overexpression of CypB enhances HIV infection by increasing nuclear import of viral DNA. The hydrophobic N-terminus of the protein is required for the effect. The N-terminus of CypB contains an ER leader sequence necessary for secretion as well as a nuclear localization signal. Passive transfer experiments showed that secreted CypB did not impact HIV infection. However the deletion of the N-terminus resulted in a pronounced mislocalization of CypB from the ER. Combined, these experiments indicate that intracellular CypB regulates a pathway of HIV nuclear import.

Viral proteins pose a tempting therapeutic target; however the fast viral turnover and inherent genetic variability drive the rapid evolution of resistance mechanisms to therapies directed at only viral proteins. Host cellular factors represent a unique target for drug development. These studies have identified
potential factors essential to HIV-1 replication and provide new insights into protein changes at the cellular level following infection.
ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor Dr. Michael Belshan for the continuous support of my study and research, motivation, enthusiasm, and immense knowledge. His mentorship has provided a model for leadership within the scientific community which I will carry with me throughout my career.

In addition to my advisor, I would like to thank the rest of my thesis committee: Dr. Jason Bartz, Dr. Pawel Ciborowski, Dr. Kristen Drescher and Dr. Patrick Swanson, for their encouragement, insightful comments, and hard questions.

I would like to thank my fellow labmates in Dr Belshan’s laboratory: Cameron Schweitzer, Bridget Sanford, Yan Li, Christian Madson, Zach Ploeger, and Patrick Bruck for the stimulating discussions and all the fun we have had in the last four years. In addition, I would like to thank my friends in the US Army: Dr. Kelly Wilhelms and Dr. Tony Pierson for their constant support and encouragement. I would also like to thank the US Army Medical Service Corps for providing the opportunity and funding for my research.

Last, but not the least, I would like to thank my family: my wife, Nikki and my kids, Conner and Brenna, and my parents Cindy Linder, Raymond DeBoer, James and Donna Allen. Without your support, I would not have made it through this program.
DISSEYATTON ORGANIZATION

This dissertation represents the research performed throughout my graduate education. Chapter I is an introduction to HIV-1 and our current knowledge of virus replication. It also provides discussion on current therapeutic strategies and states the goals for the research presented.

Chapter II describes a proteomic study of alterations in the nuclear proteome of HIV-1 infected C8166 T-cells. This work was published in Virology (DeBoer et al., 2014). I performed the majority of the research myself with Christian Madson providing initial training and assistance with the infections and validation blots. Teena Jagadish performed the sample preparation for mass spectrometry and Nicole Haverland provided the initial bioinformatic analysis of the mass spectrometry data.

Chapter III presents another proteomic study of the changes in the cellular proteome of Jurkat E6-1 cells following HIV-1 infection. This work is currently in preparation for publication in the Journal of Proteome Research. I performed all of the preliminary infection and cell fractionation. Nicole Haverland performed the sample preparation for mass spectrometry analysis and the z-transformed statistical analysis. Emily Rigden performed the overexpression assays. Christian Madson performed the virus production and infectivity assays. Yan Li performed the siRNA experiments. The mass spectrometry for the studies in chapters I and II was performed at the University of Nebraska Medical Center Mass Spectrometry Core Facility with Dr. Pawel Ciborowski.

Chapter IV presents research into the role of Cyclophilin B during HIV-1 infection. This work is currently in preparation for publication in Journal of vii
Virology. The data presented was performed by me except for the cell cycle analysis was performed by Dr. Greg Perry, the production of exogenous Cyclophilin B was performed by Yan Li, and the Reverse Transcription assays were performed by Christian Madson.
LIST OF TABLES

Table 1.1: HIV-1 Viral Protein Summary .......................................................... 7
Table 2.1: GI numbers unique to infected samples ............................................. 36
Table 2.2: Top unique proteins identified in control cell nuclei .......................... 38
Table 2.3: GI numbers two fold or greater changed between groups ............... 39
Table 3.1: Summary of expression changes ...................................................... 72
Table 3.2: Proteins identified in previous proteomic studies ............................ 74
Table 3.3: DDX39A interactors identified in Jurkat SWATH-MS study.......... 83
Table 3.4: Candidate factors selected for further analysis ............................... 85
Table 3.5: Summary of HIV-1 infection studies .............................................. 93

LIST OF FIGURES

Figure 1.1: The HIV-1 Genome and Virion ...................................................... 6
Figure 2.1: Schematic of Experimental Approach ............................................. 30
Figure 2.2: Infection of C8166 cells used in proteomic screen .......................... 31
Figure 2.3: Summary of MS data ..................................................................... 33
Figure 2.4: Bioinformatic analysis of candidate proteins ................................. 42
Figure 2.5: Validation of candidate factors identified using MS ...................... 43
Figure 2.6: Independent validation of candidate proteins ............................... 48
Figure 3.1: Statistical analysis using z-transformation .................................... 65
Figure 3.2: Experimental approach used for SWATH-MS analysis .............. 69
Figure 3.3: Confirmation of infection and fractionation control ........................ 70
Figure 3.4a: STRING-db analysis of cytosolic proteins increased following HIV-1 infection................................................................. 75
Figure 3.4b: STRING-db analysis of cytosolic proteins decreased following HIV-1 infection................................................................. 76
Figure 3.4c: STRING-db analysis of membrane/organelle proteins increased following HIV-1 infection ........................................ 77

Figure 3.4d: STRING-db analysis of membrane/organelle proteins decreased following HIV-1 infection ........................................ 78

Figure 3.4e: STRING-db analysis of nuclear proteins increased following HIV-1 infection ....................................................... 79

Figure 3.4f: STRING-db analysis of nuclear proteins decreased following HIV-1 infection ....................................................... 80

Figure 3.5: STRING-db PPI analysis of proteins with increased nuclear expression following HIV-1 infection ........................................ 81

Figure 3.6: DDX39A Interactome ................................................. 82

Figure 3.7: Effect of candidate factor overexpression on HIV-1 infection ........ 87

Figure 3.8: Effect of candidate factor siRNA knockdown on HIV-1 infection ......................................................................................... 88

Figure 3.9: Impact of candidate factor overexpression on HIV-1 virus production ......................................................................................... 91

Figure 4.1: Domain structure of CypB and the deletion mutants used in this study ......................................................................................... 100

Figure 4.2: Over-expression of CypB enhances HIV-1 infection ............ 110

Figure 4.3: CsA does not abrogate the ability of CypB to enhance infection ......................................................................................... 112

Figure 4.4: The N-terminus of CypB is required for the enhancement of HIV infection ......................................................................................... 113

Figure 4.5: The intracellular localization of the CypB deletion mutants ......... 115

Figure 4.6: The subcellular distribution of the CypB deletion mutants .......... 117

Figure 4.7: Secreted CypB does not enhance HIV infection in-trans .......... 120

Figure 4.8: CypB increases nuclear import of HIV ................................... 122

Figure 4.9: CypB enhancement does not require Stat5 activity ............... 124
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAR</td>
<td>adenosine deaminase, RNA-specific</td>
</tr>
<tr>
<td>AIDS</td>
<td>Acquired Immune deficiency syndrome</td>
</tr>
<tr>
<td>CA</td>
<td>Capsid</td>
</tr>
<tr>
<td>cART</td>
<td>Combined antiretroviral therapy</td>
</tr>
<tr>
<td>CypA</td>
<td>Cyclophilin A</td>
</tr>
<tr>
<td>CypB</td>
<td>Cyclophilin B</td>
</tr>
<tr>
<td>EEF1A1</td>
<td>Eukaryotic translation elongation factor 1 alpha 1</td>
</tr>
<tr>
<td>Env</td>
<td>Envelope</td>
</tr>
<tr>
<td>DEK</td>
<td>DEK oncogene</td>
</tr>
<tr>
<td>gag</td>
<td>Group specific antigen</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>HIV-1</td>
<td>Human immunodeficiency virus type 1</td>
</tr>
<tr>
<td>HSP90AB1</td>
<td>Heat Shock Protein 90 AB1</td>
</tr>
<tr>
<td>IN</td>
<td>Integrase</td>
</tr>
<tr>
<td>InSTI</td>
<td>Integrase strand transfer inhibitor</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>Liquid chromatography tandem mass spectrometry</td>
</tr>
<tr>
<td>LTR</td>
<td>Long terminal repeat</td>
</tr>
<tr>
<td>LYAR</td>
<td>Ly1 Antibody reactive</td>
</tr>
<tr>
<td>MA</td>
<td>Matrix</td>
</tr>
<tr>
<td>NAT10</td>
<td>N-acetyltransferase 10</td>
</tr>
<tr>
<td>NC</td>
<td>Nucleocapsid</td>
</tr>
<tr>
<td>Nef</td>
<td>Negative factor</td>
</tr>
<tr>
<td>NCL</td>
<td>Nucleolin</td>
</tr>
<tr>
<td>NLS</td>
<td>Nuclear localization signal</td>
</tr>
</tbody>
</table>
**NNRTI** – Non-nucleoside reverse transcriptase inhibitor

**NPM1** – Nucleophosmin

**NRTI** – Nucleoside-analog reverse transcriptase inhibitor

**PHB** – Prohibitin

**PIC** – Preintegration complex

**Pol** – Polymerase

**PR** – Protease

**Rev** – Anti-repression transactivator protein

**RT** – Reverse transcriptase

**RTC** – Reverse transcription complex

**RuvBL1** - RuvB-like AAA ATPase 1

**SERBP1** – Serpin peptidase inhibitor member 1 mRNA binding protein 1

**SU** – Surface protein

**SWATH** – Sequential Window Acquisition of all Theoretical fragment-ion spectra

**Tat** – Transactivating regulatory protein

**TM** – Transmembrane protein

**TOPOIIa** – Topoisomerase IIa

**Vif** – Viral infectivity factor

**Vpr** – Viral protein R

**Vpu** – Viral Protein U
# TABLE OF CONTENTS

List of Tables ........................................................................................................... ix

List of Figures ............................................................................................................ ix

List of Abbreviations ................................................................................................ xi

## Chapter I: Introduction ........................................................................................... 1

1.1 The Emergence of HIV and the AIDS pandemic ........................................ 1

1.2 The Virus ........................................................................................................... 4

1.3 The Infection Cycle ............................................................................................ 5

1.4 Treating the HIV-1 Infection .......................................................................... 13

1.5 *In-vitro* Biological Studies of HIV-1 Infection .......................................... 18

1.6 Research Goals ................................................................................................ 19

## Chapter II: Alterations in the Nuclear Proteome of HIV-1 Infected T-cells ...................................................................................................................... 21

2.1 Background ....................................................................................................... 21

2.2 Materials and Methods .................................................................................... 23

2.2.1 Cell culture, infections, and nuclear isolations ........................................ 23

2.2.2 In-gel tryptic digests .................................................................................. 24

2.2.3 Nano-LC-MS/MS ...................................................................................... 25

2.2.4 Bioinformatic analyses .............................................................................. 27

2.2.5 Immunoblotting ......................................................................................... 27

2.3 Results ............................................................................................................... 28

2.3.1 Experimental approach .............................................................................. 28

2.3.2 Database search and protein identification .............................................. 32

2.3.3 Identification of uniquely expressed proteins .......................................... 34

2.3.4 Identification of Other Candidate Factors .............................................. 37
Chapter III: SWATH-MS Analysis of the Host Cell Protein Response during HIV-1 Infection of T-cells ........................................................................ 58

3.1 Background .................................................................................. 58

3.2 Materials and Methods ................................................................. 60

  3.2.1 Cell culture and virus production .............................................. 60

  3.2.2 SWATH-MS ........................................................................... 61

  3.2.2.1 Infection and Sample Preparation ....................................... 61

  3.2.2.2 Mass Spectrometry ............................................................... 62

  3.2.2.3 Protein Identification and z-transformation ......................... 63

  3.2.3 Infection Assays ..................................................................... 64

  3.2.4 Immunoblotting ..................................................................... 66

  3.2.5 Endogenous RT Activity ......................................................... 67

3.3 Results .......................................................................................... 68

  3.3.1 Experimental Approach .......................................................... 68

  3.3.2 Bioinformatic Analysis ............................................................. 72

  3.3.3 Biological Evaluation of Candidate Factors .............................. 84

  3.3.3.1 Does overexpression of candidate factor alter HIV-1 infection? ........................................................................... 86

  3.3.3.2 Does siRNA knockdown of candidate factor alter HIV-1 infection? ........................................................................... 86

  3.3.3.3 Virus production after overexpression of candidate factor ...... 89
Chapter IV: Cyclophilin B Enhances HIV-1 Infection ........................................... 98

4.1 Introduction ........................................................................................................... 98

4.2 Materials and Methods .......................................................................................... 102

4.2.1 Plasmids ............................................................................................................. 102

4.2.2 Cell Culture and Virus Preparation ................................................................. 103

4.2.3 Infection Assays ................................................................................................. 104

4.2.4 Immunoblotting ................................................................................................. 105

4.2.5 Indirect Immunofluorescence ........................................................................... 105

4.2.6 Quantification of DNA for LRT, 2-LTR, and Integrated viral DNA ............... 106

4.2.7 STAT5 Inhibition Assay ....................................................................................... 107

4.3 Results ..................................................................................................................... 108

4.3.1 CypB enhances HIV-1 infection ......................................................................... 108

4.3.2 The mechanism of enhancement of infection by CypB is distinct from CypA ........................................................................................................................................ 109

4.3.3 The N-terminus of CypB is necessary to enhance HIV-1 infection ...................... 111

4.3.4 Deletion of the N-terminus alters CypB localization ........................................ 114

4.3.5 Secreted CypB does not enhance HIV-1 infection ........................................... 118

4.3.6 CypB increases nuclear import of HIV-1 DNA ............................................... 119

4.3.7 CypB potentiation does not involve the STAT5 pathway .............................. 123

4.4 Discussion ............................................................................................................. 125

4.5 Future Studies ......................................................................................................... 127
Chapter V: Project Summary and References ................................................. 129

5.1 Project Summary .................................................................................. 129

5.2 References ......................................................................................... 132
CHAPTER I: Introduction

1.1 The Emergence of HIV and the AIDS pandemic

In the late 1970’s and early 1980’s, increasing numbers of young homosexual men became ill with generalized lymphadenopathy, opportunistic infections and rare malignancies. In 1981, the CDC Morbidity and Mortality Weekly Report formally described five cases of severe immune deficiency shining light on this previously undescribed syndrome. Additional cases followed with findings of significantly reduced CD4$^+$ T-lymphocytes and the syndrome was given a name, Acquired Immune Deficiency Syndrome or AIDS (Greene, 2007). As increasing numbers of cases were examined, it became clear that a similar syndrome was affecting other groups, including blood transfusion recipients and sexual partners.

In 1983, Montagnier and colleagues at the Pasteur Institute isolated an infectious agent from the lymph nodes of a patient who presented with generalized lymphadenopathy of unknown origin (Barre-Sinoussi et al., 1983). Electron microscopy revealed characteristic retroviral features and the group named the agent the lymphadenopathy-associated virus (LAV). Subsequent characterization revealed the virus displayed magnesium-dependent reverse transcriptase (RT) activity and was found to specifically kill CD4$^+$ T-lymphocytes in cell culture (Chermann, Barre-Sinoussi, and Montagnier, 1985). One year later, in 1984, Gallo and coworkers at the National Institutes of Health isolated a retrovirus from a patient with AIDS, which they named human T-cell leukemia virus type III (HTLV-III) (Gallo et al., 1984). Additional studies provided the first
convincing serological evidence linking exposure to HTLV-III/LAV retroviruses and immunodeficient individuals (Sarngadharan et al., 1984). Concurrently, Levy et al. found another, comparable retrovirus from patients with AIDS as well healthy individuals from the highest risk groups which was named AIDS-associated retrovirus (ARV) (Levy et al., 1984). The work of both Montagnier and Levy suggested that these viruses could induce both asymptomatic and symptomatic infections. A comparison of all three new retroviruses associated with AIDS in the United States, Europe, and central Africa exhibited morphologic and genetic characteristics typical of the lentivirus genus with no significant differences between them. The common name applied to all three was human immunodeficiency virus, or HIV (Coffin et al., 1986; Rabson and Martin, 1985). In 1986, a second, immunologically distinct HIV virus was discovered in individuals residing in several west African countries such as Senegal, Ivory Coast, and Guinea-Bissau (Clavel et al., 1986). Discovery of this new virus led to the renaming of the viruses to HIV-1 and HIV-2. In comparison with HIV-1, HIV-2 has lower mortality rates, likely due to slower declines in CD4+ T-lymphocytes and much longer periods of asymptomatic infection (Gottlieb et al., 2002).

When scientists identified HIV in 1983, they believed that HIV was a new virus, one that had just emerged from the jungles of Africa. However, using advanced statistical modeling of sequence data, the emergence of HIV-1 can be traced back to colonial west central Africa in the 1920’s (Faria et al., 2014; Korber et al., 2000). It is believed that HIV-1 started by zoonotic transmission of the chimpanzee strain of the Simian Immunodeficiency Virus (SIVCPZ) to forest
hunters early in the 20th century (Taylor et al., 2008). The virus was passed, unnoticed, along the Congo River corridor for years until reaching the urban population center of Kinshasa, Zaire. Here, HIV-1 found the population density and mobility required to expand replication within the population. In fact, the oldest known infection with HIV-1 was found in samples obtained from Kinshasa in 1959 (Worobey et al., 2008; Zhu et al., 1998).

Exactly how and when HIV spread to the United States is not known, however, using phylogenetic analysis of env and gag genes, Gilbert et al. provide the most compelling evidence of the pandemic virus emergence to date (Gilbert et al., 2007). It is thought that HIV-1 migrated from Africa to Haiti around 1966 and subsequently propagated there undiscovered for years. In the late 1960’s, the virus made the initial jump to the United States, likely taking hold in the Haitian immigrant population. The virus quietly spread throughout the United States and Europe throughout the 1970’s until reaching pandemic status in the early 1980’s.

According to the World Health Organization, over 78 million people have been infected with HIV since the beginning of the pandemic, with over 39 million dying of HIV infection (http://www.who.int/gho/hiv/en/). In 1984, an optimistic Margaret Heckler, the US Surgeon General, proudly proclaimed “We hope to have such a vaccine ready for testing in approximately two years.” (Watkins, 2008). Thirty years later, the virus continues to be a significant world health problem, subverting all attempts at a cure or vaccine, and even developing
increased pathogenesis (Kouri et al.). Continued research into the mechanisms behind successful replication is essential for sustaining the fight against HIV.

1.2 The Virus

HIV-1 has been extensively studied over the past 30 years, with a great deal known about how the virus infects cells, replicates, and causes disease. However, significant gaps in knowledge remain with virus-host protein interactions, the key focus of my research. In this section I will describe the known functions and interactions of the virus and associated viral proteins.

HIV-1 is a single stranded, non-segmented, positive sense RNA virus belonging to the family Retroviridae and genera Lentivirus (Fields, Knipe, and Howley, 2007). The HIV-1 genome is approximately 9-kb consisting of nine open reading frames encoding the characteristic retroviral Gag, Pro, Pol, and Env genes as well as additional small regulatory and accessory proteins. The virion is comprised of two copies of single-stranded viral RNA tightly bound by the protective nucleocapsid (NC) protein. Also attached to the viral RNA are the required viral enzymes Reverse Transcriptase (RT) and Integrase (IN). The viral Capsid (CA) protein forms a fullerene cone shaped core encapsulating the viral RNA complexes along with viral accessory and host derived proteins. Surrounding the viral core in a small sub-virion space are excess CA molecules and a limited amount of host proteins followed by a thin layer of viral Matrix (MA) protein (Lanman et al., 2003). The outer envelope consists of host derived
membrane with embedded viral Env glycoproteins (Env). A schematic of the HIV-1 genome and virion structure can be found in Figure 1.

The individual viral proteins can be broken down into four main groups: structural, enzymatic, regulatory, and accessory. The structural components of the virion are produced from the Gag and Env polyproteins. The Gag proteins CA and NC form the basic components of the viral core while MA and the Env surface (SU) and transmembrane (TM) proteins are found in the outer membrane envelope. The virion also contains essential enzymatic components produced from the Pol polyprotein. RT and IN are enzymes required for retroviral replication and not found within the host cell. The viral protease (PR) is essential for maturation of the virus by cleavage of the Gag and Pol polypeptide. Tat and Rev provide gene regulation with the host cell and are required for in-vivo replication. The accessory proteins are not necessary for in-vitro replication and can be further separated into intraparticle and intracellular components. The proteins Vif, Vpr, and Nef are found within the viral particle. Finally, Vpu assists in the assembly of the virions. A summary of the viral proteins can be found in Table 1.

1.3 The HIV-1 Infection Cycle

Transmission of HIV-1 to a new host occurs by introduction of the virus into the body through breaks in the mucosal surfaces, maternal transfer, or
Figure 1.1 The HIV-1 Genome and Virion. Schematic representation of the HIV-1 genome and virion structure.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Intermediate</th>
<th>Protein</th>
<th>Protein Function</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structural Proteins</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gag</td>
<td>Gag Precursor p55</td>
<td>MA</td>
<td>Plasma membrane targeting of gag and gag/pol polyprotein</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA</td>
<td>Viral core protein</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NC</td>
<td>Viral RNA packaging and transport to budding virion</td>
</tr>
<tr>
<td></td>
<td>p6</td>
<td></td>
<td>Vpr incorporation and particle release</td>
</tr>
<tr>
<td>env</td>
<td>Env Precursor gp160</td>
<td>SU</td>
<td>CD4/CCR4/CXCR5 receptor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TM</td>
<td>Membrane fusion protein</td>
</tr>
<tr>
<td><strong>Enzymatic Proteins</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pol</td>
<td>Gag/Pol Precursor p160</td>
<td>IN</td>
<td>Integration of viral DNA into host chromosome</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RT</td>
<td>Reverse transcription of viral RNA into DNA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PR</td>
<td>Proteolytic processing of gag and gag/pol polyproteins</td>
</tr>
<tr>
<td><strong>Regulatory Proteins</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tat</td>
<td></td>
<td>Tat</td>
<td>Positive regulator of transcription</td>
</tr>
<tr>
<td>Rev</td>
<td></td>
<td>Rev</td>
<td>RNA export/splicing factor</td>
</tr>
<tr>
<td><strong>Accessory Proteins</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vif</td>
<td></td>
<td>Vif</td>
<td>Viral DNA synthesis and stability</td>
</tr>
<tr>
<td>Vpu</td>
<td></td>
<td>Vpu</td>
<td>Intracellular degradation of CD4, virus release</td>
</tr>
<tr>
<td>Vpr</td>
<td></td>
<td>Vpr</td>
<td>Viral DNA nuclear transport, cell cycle arrest</td>
</tr>
<tr>
<td>Nef</td>
<td></td>
<td>Nef</td>
<td>Endocytosis and degradation of surface CD4 and MHC Class II</td>
</tr>
</tbody>
</table>
percutaneous inoculation. The most common route of infection is through mucosal abrasions caused by sexual contact or lack of mucosal integrity due to sexually transmitted diseases or genital ulcers (Abrahams et al., 2009; Haaland et al., 2009). Rather than direct blood-to-blood transmission of virus through these abrasions, it is thought that HIV-1 enters the submucosal layers via Langerhans and dendritic cells which can be infected, but more important, present HIV-1 to CD4+ T cells in the surrounding tissue and associated lymphatics (Boggiano, Manel, and Littman, 2007; Hladik et al., 2007; Lindback et al., 2000). From the peripheral lymphatic system, the virus rapidly spreads to the organized lymphoid tissues such as the gut associated lymphatic tissue (GALT) where massive destruction of CD4+ T cells follows (Brenchley et al., 2004). From this point, the virus disseminates and latent reservoirs are established.

The HIV-1 infection cycle begins with the binding and entry of the virus, one of the most studied virus/host interactions to date. A well-defined sequence of events occurs ultimately leading to release of the viral core into the cytoplasm. The primary viral protein involved in binding and entry of the virus is the HIV-1 Env. HIV-1 Env consists of two main subunits, the gp120 surface glycoprotein and the gp41 transmembrane glycoprotein. Upon arrival of the viral particle at the cell surface, gp120 selectively binds CD4, a cellular protein receptor, located primarily on T cells, but also macrophages and dendritic cells. Binding of CD4 initiates a conformational change allowing the gp120/gp41 complex to bind a co-receptor, CXCR4 or CCR5. Additional conformational changes expose the fusion domain of gp41. Concurrent to the exposure of the fusion domain, gp120 and
gp41 dissociate leaving the six-helix bundle configuration at which point the fusion protein is inserted into the host cell membrane (Gallo et al., 2003). The progressive oligomerization of gp41 fusion proteins expands the pore size leading to the fusion of the membranes and release of the viral core into the host cell.

The events following viral core release into the cytoplasm and leading up to nuclear import are some of the most important, but also the least understood. Upon release of viral cores into the cytoplasm, the CA begins to dissociate (Bukrinsky, 2004). The extent of dissociation is just one of the controversial elements of HIV-1 infection. It has generally been accepted that CA completely dissociates from the viral RNA completely during reverse transcription. Recent studies suggest that CA plays a more significant role, being associated with the pre-integration complexes (PICs) during nuclear docking and even the appearance of CA within the nucleus (Arhel, 2010; Schaller et al., 2011). In either case, it is known that reverse transcription can be initiated within the core and requires some level of CA dissociation prior to completion.

Reverse transcription is the conversion of viral single stranded RNA to a double stranded DNA molecule. HIV-1 has evolved an elaborate means of accomplishing this task, primarily through the viral enzyme, reverse transcriptase (RT). RT has two primary domains, one demonstrating DNA polymerase activity and one RNaseH activity. HIV-1 reverse transcription begins with the formation of the reverse transcription complex (RTC) which includes host cellular tRNA^{Lys,3}, host cellular factors, and the viral proteins MA, RT, NC and IN. Upon binding of tRNA^{Lys,3} to the viral RNA primer binding site, reverse transcription is initiated.
Reverse transcription progresses through an intricate mechanism of multiple strand transfers and concurrent viral RNA destruction as cDNA is transcribed (Jonckheere, Anné, and De Clercq, 2000). Complete viral cDNA is then processed by IN and forms the integration competent Pre-Integration Complex (PIC (Miller, Farnet, and Bushman, 1997)). Reverse transcription and PIC formation occurs during movement of the complex from the plasma membrane to the nucleus through the cytoplasm. Due to the relatively large size of the RTC and viscosity of the cytoplasm, active transport via the cellular cytoskeleton is required. Early movements of the virus within the cell have been associated with the actin cytoskeleton (Bukrinskaya et al., 1998). Later, directed movement of the virus towards the nucleus occurs along the microtubule network (McDonald et al., 2002). The exact mechanism of interaction between the RTC and microtubules has yet to be determined.

Nuclear import of the HIV-1 PIC is yet another poorly understood event of HIV-1 infection. Typically, retroviruses rely on breakdown of the nuclear envelope during mitosis for access to host DNA. HIV-1 and other lentiviruses possesses the unique ability to infect non-dividing cells (Weinberg et al., 1991). The central channel in the nuclear pore complex has a diameter of 25 nm, limiting passive diffusion. Through active transport, molecules up to 39 nm can be transported across the nuclear membrane (Pante and Kann, 2002). The approximate size of a HIV-1 PIC is 28 nm leaving active transport the most likely method of PIC translocation (Miller, Farnet, and Bushman, 1997). PIC entry through small ruptures in the nuclear membrane caused by Vpr has been proposed as an
alternate mechanism, although this theory has not been confirmed (de Noronha et al., 2001). Current research models favor active transport of the PIC using nuclear localization signals; either directly on PIC associated viral proteins (MA, IN, Vpr) or indirectly through binding host karyophilic proteins.

Integration of the viral cDNA is the final step in early HIV infection. Upon entry into the nucleus, the HIV-1 cDNA targets an active transcription site for integration. During formation of the PIC, the linear viral DNA ends are processed by IN prior to final attachment. Upon selection of integration site, the corresponding host DNA is processed in a similar fashion and a strand transfer event occurs (Coffin, Hughes, and Varmus, 1997). Host DNA repair machinery ligate the 5’ ends of the vDNA to the target DNA and fill in the single-strand gaps completing the integration of viral DNA into the host genome (Brown et al., 1987). Not all viral DNA that makes it into the nucleus is integrated into host DNA. Depending on which DNA joining pathway is utilized, the viral DNA can make circular DNA with either 1 or 2 long terminal repeats (LTR’s). While these circles are not fully transcriptionally competent, they can produce small amounts of mRNA coding for viral proteins (Sloan and Wainberg, 2011). More often, they are used as a biomarker for viral DNA nuclear import.

Following integration of the viral DNA, the virus can take one of two paths, active transcription or latency. What decides the fate of the provirus is unclear, but likely relies on the current health and activation status of the cell (Brooks et al., 2003). Provirus transcription is complex and relies upon many viral factors, primarily HIV-1 Tat. Initial mRNA products are limited primarily to the viral proteins
Tat, Rev and Nef. As stated earlier, Tat is primarily responsible for activating transcription of the viral DNA. Normal cellular pathways restrict unspliced mRNA export; the Rev protein provides HIV-1 a mechanism to circumvent this restriction. Rev Response Elements (RRE’s) encoded on HIV-1 mRNA is essential in this alternate pathway. Rev acts as an intermediary between mRNA and Ran/GTP allowing for nuclear export of the unspliced mRNA (Malim et al., 1989). Many functions of nef have been reported. These include downregulation of surface markers (CD4, MHC, CD3, and CD28), enhancement of virus infectivity, and regulation of various cellular pathways (Das and Jameel, 2005).

The final stages of infection include translation of viral proteins, assembly of the virion, budding, and maturation. The primary viral polyprotein produced is the p55 gag. Approximately 5% of the time, an alternate form of gag, p160 gag-pol is produced by a -1 frameshift during translation (Shehu-Xhilaga, Crowe, and Mak, 2001). The p160 gag-pol polyprotein includes the gag proteins (MA, CA, NC, p6) as well as additional enzymes (PR, RT, IN) required for HIV-1 infection. The Gag (and Gag-Pol-Pro) is subsequently transported to lipid rich "raft"-like areas of the cellular membrane where it multimerizes forming Gag-rich areas on the inner cell membrane. Through an independent mechanism, the viral Env protein is inserted cotranslationally into the ER membranes. Env travels through the secretory pathway and processed into TM and SU by the cellular protease furin. The intracellular tail of TM targets the protein to the same lipid rich domains and also mediates MA interactions promoting Env incorporation into virions (Cosson, 1996; Wyma, Kotov, and Aiken, 2000). NC captures the
genomic vRNA dimer and the complex is targeted to p6 region of Gag for incorporation into the budding virion. In addition, p6 contains binding sites for Vpr and members of the ESCRT pathway required for budding, TSG101 and ALIX (Garrus et al., 2001; Strack et al., 2003).

As the Gag proteins assemble, a lattice is formed resulting in a rounding of the cell membrane around the vRNA marking the beginning of virion budding. HIV uses the p6 bound TSG101 and ALIX to recruit additional host ESCRT pathway components, primarily ESCRT-III and VPS4, to mediate membrane fission and complete the budding process (Morita and Sundquist, 2004). Upon budding, the viral PR begins the maturation process by sequential cleavage of the gag proteins. The RT and IN enzymes as well as the NC associate with the vRNA and the CA forms a stable core around the vRNA. MA remains in close proximity to the viral membrane, while removing the gag structure allows the Env to move freely around the viral surface indicating a mature viral particle.

1.4 Treating the HIV-1 Infection

The pharmacological treatment of patients with HIV-1 presents a challenge due to the highly error-prone reverse transcription process. The HIV-1 genome is approximately 10,000 nucleotides in length, and it is estimated that one mutation is induced every 1000-10,000 nucleotides synthesized. This may result in 1 to 10 mutations per viral genome every replication cycle (Abram et al., 2010). Due to this enormous genetic diversity, early monotherapy produced rapid
drops in viral load followed by significant drug resistance. In the late 1990’s the current model of combined antiretroviral therapy (cART) was implemented. The combination of three different antiretroviral drugs proved effective at limiting the emergence of drug resistance and actually dropped the circulating viral load to below detectable levels (Lederman et al., 1998; Staszewski et al., 1996).

The current antiviral options available for HIV-1 are diverse with over 30 drugs currently approved for treatment (www.fda.gov). The primary focus of drug development thus far has been towards viral proteins, as the risk of off target effects is reduced. Current drugs fall into one of six distinct classes: (1) nucleoside-analog reverse transcriptase inhibitors (NRTIs), (2) non-nucleoside reverse transcriptase inhibitors (NNRTIs), (3) integrase inhibitors, (4) protease inhibitors (PIs), (5) fusion inhibitors, and (6) coreceptor antagonists.

The first two classes, NRTIs and NNRTIs can be grouped together as they both inhibit the reverse transcription of viral RNA into viral DNA. However, they differ significantly in their mechanism of action. NRTIs lack a 3’ hydroxyl group at the 2’-desoxyribosyl moiety which prevents the formation of a 3’-5’ phosphodiester bond between the NRTI and incoming 5’-nucleoside triphosphate terminating the growing DNA chain. Termination can occur on either RNA-dependent DNA or DNA-dependent DNA synthesis maximizing the effectiveness of the drug (Cheng et al., 1987). Resistance to NRTIs is mediated by three mechanisms: pyrophosphorolysis or removal of the NRTI, reversal of chain termination, and prevention of NRTI incorporation through mutations in the RT enzyme (Arion et al., 1998; Meyer et al., 1999; Schinazi et al., 1993). NNRTIs
inhibit enzyme activity by binding RT and changing the spatial conformation of
the substrate-binding site reducing the polymerase activity (Tantillo et al., 1994).
Resistance to NNRTIs can be easily obtained through simple amino acid
mutations with little impact on replicative fitness leading to overall stability of
NNRTI resistance (Dykes et al., 2001).

A relatively new class of drugs targets the viral IN enzyme, specifically the
strand transfer reaction (Espeseth et al., 2000). All IN inhibitors (InSTIs) currently
in development target strand transfer in a similar manner, binding a specific
complex between IN and the viral DNA. Briefly, a metal-binding pharmacophore
sequesters the active site magnesium and a hydrophobic group that interacts
with the viral DNA (Grobler et al., 2002). The common mechanism of action and
conserved binding actions present a dichotomy for development of resistance to
InSTIs. Mutations that inactivate the inhibitor often have damaging effects on
essential enzymatic functions and viral replication (Hare et al., 2010).
Nonetheless, specific mutations have been found that confer resistance to
InSTIs.

The next class of antiretroviral drugs is the protease inhibitors (PI). The
HIV-1 protease is an enzyme essential for the maturation of viral particles,
responsible for the viral gag and gag-pol cleavage (Park and Morrow, 1993).
Initial expectations were that resistance to PI would be rare due to the essential
role of the enzyme in the HIV-1 lifecycle and small size. Subsequent analysis
revealed great plasticity within the protease gene with polymorphisms in 49 of the
99 codons. Resistance to PIs occurs with mutations clustered near the active site
of the enzyme and these changes usually have a negative impact on replicative fitness (Nijhuis, Deeks, and Boucher, 2001). In order to circumvent PIs, HIV appears to follow a stepped development of resistance. First, there is selection of primary mutations conferring resistance to PIs, but losing function/viral fitness in the process. Second, there is selection of compensatory mutations that result in the rescue of the enzymatic function. Finally, the selection of mutations in the gag/gag-pol cleavage sites further enhances enzyme function fully restores viral fitness.

The final two classes of drugs are both entry inhibitors, although they form distinct subclasses based upon the target and mechanism of action. Entry inhibitors present a unique opportunity for viral control, as prevention of the virus into the cell is preferred to controlling the virus once it enters into the cell. The highly variable nature of the viral Env proteins and makeup of the lipid envelope present significant challenges to preventing viral entry. Wild et al. discovered that two homologous domains in the viral gp41 must interact to promote fusion. Disruption of this interaction by a heterologous protein can disrupt the intramolecular interactions and reduce infection (Wild et al., 1994). Subsequent development of these proteins led to a new class of drugs, the fusion inhibitors (FI). As with the InSTIs, mutations that alter the FI binding also reduce the replicative capacity of the virus (Labrosse et al., 2003).

Small molecule CCR5 antagonists present the newest class of compound available to combat HIV-1. These molecules bind to pockets within the transmembrane helices of CCR5 (Dragic et al., 2000). These molecules do not
block the interaction with HIV directly, but they make the CCR5 receptor unrecognizable. Due to the compound interacting with a host protein, development of resistance is significantly altered. Potential mechanisms include switching tropism to CXCR4, increasing affinity for the coreceptor or adaptation to the inhibitor bound receptor, and increasing the rate of entry. The most concerning is the CXCR4 tropism switching as this switch typically leads to faster disease progression (Clapham and McKnight, 2001). Due to this and difficulties determining resistance, use of CCR5 antagonists is limited.

As described above, the rapid mutation of HIV-1 requires a multi-faceted approach to treatment. The rapid development of resistance to single drugs presents a problem not uncommon in conventional treatment of infectious disease. The development of cART had a profound impact on the lives of HIV-1 infected patients. This success is not without limitations. Constant monitoring and rotation of drugs is required to remain ahead of the mutations. In addition, the development of multi-drug resistance remains a growing concern for the long-term efficacy of the existing antiviral drugs available (Megens and Laethem, 2013). Traditional drug design has been focused primarily on a viral protein which allows rapid mutation of the viral genome to alter therapeutic targets. Successful replication of HIV-1 requires a number of host cellular factors, many of which are yet to be discovered. Targeting host cellular factors presents a novel opportunity for viral control, as it is more difficult for the virus to evolve alternate replication mechanisms. Identifying alterations in host protein quantity and localization is an essential first step in the development of cell based treatments.
1.5 In-vitro Biological Studies of HIV-1 Infection

HIV-1 possesses a relatively small genome consisting of 9 open reading frames coding for 15 viral proteins (Fields, Knipe, and Howley, 2007). The overall simplicity of the viral genome requires the virus to be dependent upon host cell factors at virtually every step of replication. An extensive number of proteomic, metabolic and genetic analyses of cellular changes during HIV-1 infection have been undertaken to identify host factors critical for virus replication. Whole-genome association has been used to discover unique polymorphisms in both the HLA and RNA Polymerase I subunit expression in response to infection (Fellay et al., 2007). Extensive biochemical analysis by protein identification using mass spectroscopy has been used to evaluate a wide variety of cellular and viral proteomes such as whole virus particle (Chertova et al., 2006), viral sub-particle RTC (Schweitzer et al., 2012) and PIC (Raghavendra et al., 2010; Schweitzer et al., 2012), and viral proteins Nef (Mukerji et al., 2012) and Tat (Coiras et al., 2006) host cellular proteomes. Multiple whole genome small interfering RNA (siRNA) screens were also performed to identify critical factors for replication although very few factors were validated (Brass et al., 2008; König et al., 2008; Zhou et al., 2008). Finally, Zhang conducted full cell proteomic studies to examine differences between HIV-1 infected and non-infected whole cell lysates (Zhang et al., 2010).

At the conclusion of these varied studies, an independent meta-analysis was performed comparing results of 9 different genome/proteome studies and
the NCBI HIV-1 interactions table. This analysis produced some very interesting, albeit perplexing results. In combination, these studies identified over 1200 genes affected by HIV infection. When compared to the published NCBI HIV-1 interactions, there was only overlap of 257 genes. In fact, the highest overlap of any two studies was only 6%, likely due to significant variation between methods such as cell type used, timing of cell harvest, and differences in screening libraries. One significant finding of this study was not necessarily new individual factors, but the systemic analysis led to clusters of cellular subsystems. Through these studies, numerous host transcription and regulatory factors have been identified as required for HIV-1 replication (Bushman et al., 2009).

Further studies using newly evolving methods and focusing on cellular subsystems, such as the nuclear environment during infection, allows for more directed study and a more complete understanding of the HIV-1 lifecycle.

1.6 Research Goals

HIV-1 requires host cellular factors and processes to complete the infection cycle. At the same time, the virus must also counteract inherent protective measures employed by the host cell. Manipulation of the host cellular environment by the virus is essential for productive replication of HIV-1. To date, HIV-1 has been shown to interact with over 4225 host cellular proteins listed in the National Center for Biotechnology Information (NCBI) HIV-1 Human Interaction database (Fu et al., 2009). As large as this database is, a significant
number of interactions remain to be discovered, as our own proteomic studies revealed over 600 proteins with no known HIV-1 interaction.

The goal of my studies was to identify HIV-1 interacting proteins using innovative proteomic approaches to identify potential druggable targets for development of novel therapeutics. Previous proteomic approaches utilized either global screens of infected cells or interactome studies of individual viral proteins. However, subtle proteome changes may not be apparent in broad studies, or too severely limited in interactome proteomics. My initial studies focused on identifying changes to nuclear protein expression in T cells following HIV-1 infection and are described in Chapter II. Additional studies presented in Chapter III focus on whole cell analyses of protein expression and localization in HIV-1 infected T cells. Initial bioinformatic and functional screens were performed on a number of factors with Cyclophilin B (CypB) identified as a lead candidate protein. My experiments found that CypB enhanced HIV replication by increasing nuclear import of viral DNA and studies of the CypB/HIV interaction can be found in Chapter IV.

These comprehensive studies of the changes in the nuclear and subcellular proteomes during HIV-1 infection elucidate intracellular alterations which contribute to virus replication and define the host cell response. That knowledge may identify new cellular targets for therapeutic intervention as well as improve the fundamental knowledge of HIV-1 biology.
Chapter II: Alterations in the Nuclear Proteome of HIV-1 Infected T-cells

2.1 BACKGROUND

AIDS is a devastating disease that has claimed the lives of an estimated 39 million people. The etiologic agent of AIDS is HIV-1, a member of the \textit{lentiviridae} subfamily of retroviruses. HIV-1 infects cells of the immune system, specifically CD4\(^+\) T-cells, macrophages, and dendritic cells. Left untreated, HIV-1 infection results in chronic activation and eventual destruction of the immune system followed by an onslaught of opportunistic infections. HIV-1 has a small genome and expresses only 15 proteins but exhibits a complex life-cycle. Like all viruses it is dependent on host cell proteins and factors for productive replication and spread. These interactions with host factors promote replication, neutralize host defenses, and elicit pathogenesis.

Uncovering the network of host cell responses to viral invasion may reveal novel targets for cell-based anti-viral therapeutics. To this end, an extensive number of genetic and proteomic analyses of the cellular changes during HIV-1 infection have been completed (reviewed in (Giri et al., 2006)). Multiple whole genome small interfering RNA (siRNA) screens were completed to identify host factors critical for replication ((Brass et al., 2008; Konig et al., 2008; Liu et al., 2011; Zhou et al., 2008); summarized in (Bushman et al., 2009) and others). There have also been whole-genome association screens to discover host polymorphisms associated with virus acquisition, set point, and control (Fellay et al., 2007; Lingappa et al., 2011; Petrovski et al., 2011). Proteomic studies have investigated a wide variety of cellular and viral proteomes including T cell
(Ringrose et al., 2008; Sheng and Wang, 2009), macrophage (Haverland, Fox, and Ciborowski, 2014; Kraft-Terry et al., 2010), intact HIV particles (Bregnard et al., 2013; Chertova et al., 2006; Saphire, Gallay, and Bark, 2006), purified HIV cores (Fuchigami et al., 2002; Santos et al., 2012), and HIV reverse transcription and preintegration complexes (Raghavendra et al., 2010; Schweitzer et al., 2013). Individual viral protein interactomes have also been mapped, including the viral proteins Gag (Engeland et al., 2014), Nef (Mukerji et al., 2012), Tat (Coiras et al., 2006), and Rev (Naji et al., 2012), as well as a comprehensive study of all HIV-1 proteins (Jager et al., 2011). Subcellular studies of the nucleolus (Jarboui et al., 2012) and the nuclear membrane (Monette, Pante, and Mouland, 2011) have also been reported. Comprehensive summaries of the current interactome of HIV-1 can be found on the National Institutes of Health NCBI website (http://www.ncbi.nlm.nih.gov/projects/RefSeq/HIVInteractions/).

The goal of this study was to investigate the changes in the nuclear proteome of T cells during HIV-1 infection. Subtle changes to the nuclear proteome either through translocation of cytosolic proteins or induction of regulatory factors can contribute to the successful integration of the viral genome, changes in viral gene expression, as well as modulate the host cell defense against viral infection. Numerous host transcription and regulatory factors have been identified as required for HIV-1 integration, transcription, splicing, and RNA export (summarized in (Bushman et al., 2009)). To identify and characterize novel nuclear proteins associated with HIV-1 infection we used mass spectrometry to investigate the changes in nuclear proteins at 20 hours post infection (hpi) in
C8166-45 T cells. Using a high threshold of identification, a total of 51 proteins were identified as unique in the nuclei of infected or control cells in at least two biological replicates. An additional 15 proteins were predicted to be differentially expressed by having a greater than two-fold change in total Proteome Discoverer score between the infected and control samples. We validated the MS analysis by examining the subcellular expression of several candidate proteins by immunoblot. Differences in the nuclear expression of eight proteins between control and infected cells were confirmed in an independent T-cell line across a time-course infection. Combined these data map changes in the nuclei proteome of HIV infected T-cells as well as confirm the altered expression of several candidate proteins during infection.

2.2 MATERIALS AND METHODS

2.2.1 Cell culture, infections and nuclear isolations

C8166-45 and Jurkat E6-1 T cell lines (NIH AIDS Research and Reference Reagent Program, Germantown, MD) were cultured in RPMI 1640 media. 293T (NIH AIDS Research and Reference Reagent Program, Germantown, MD) cells were cultured in DMEM media. All media was supplemented with 10% Fetalclone III (Hyclone, Logan, UT USA), 8 mM L-glutamine, 100 U/mL penicillin, and 100 U/ml streptomycin. Cells were cultured in humidified incubators at 37°C and 5% CO₂. HIV-1 NLX virus stocks were produced by transient transfection of 293T
cells using Polyethylenimine (PEI) as previously described (Schweitzer et al., 2013). Viral supernatants were collected over 72 hours, filter concentrated and stored at \(-80^\circ C\). Virus was treated with 100 U/ml Turbo DNase (Ambion, Austin, TX USA) for 60 min at 37° C prior to infecting cells. For nuclear isolation experiments 1x10^8 C8166-45 cells were infected by the spinoculation technique as described previously (O'Doherty, Swiggard, and Malim, 2000; Schweitzer et al., 2013). At 20 hours post infection the subcellular compartments (cytosol, membrane/organelle, nuclear, and cytoskeleton/insoluble) were isolated using the QProteome Cell Compartment kit as described by the manufacturer (Qiagen, Valencia, CA USA). Protein concentrations were determined by BCA protein assay (Thermo Scientific, Rockford, IL USA). For time-course experiments, 4x10^7 Jurkat E6-1 cells were spinoculated with VSVg pseudotyped HIV-1 NLX virus. At each time-point (24, 48, 72 hpi), 1x10^7 cells were removed and subcellular compartments isolated and processed as described above.

2.2.2 In-gel tryptic digests

Nuclear fractions (20 µg/lane) were separated by standard SDS-PAGE and protein bands were visualized by staining with Coomassie brilliant blue. Each lane was cut into 28 individual slices that were processed individually in glass tubes. In-gel tryptic digests were performed essentially as described before (Schweitzer et al., 2012) using a modified protocol adapted from (Shevchenko et al., 2001). Briefly, the Coomassie blue stain was removed by washing with 50%
H₂O/50% acetonitrile (ACN) followed by 50% ACN/50 mM \( \text{NH}_4\text{HCO}_3 \), and 50% ACN/10 mM \( \text{NH}_4\text{HCO}_3 \). Each slice was dried by centrifugation under vacuum and incubated with 10 \( \mu \text{L} \) (0.1 \( \mu \text{g/\mu L} \)) modified trypsin (Promega, Madison, WI USA) at room temperature for 10 minutes. Then 50 \( \mu \text{L} \) \( \text{NH}_4\text{HCO}_3 \) was added and the sample incubated overnight at 37° C. Digested peptides were extracted with three washes of 0.1% Trifluoroacetic acid in 60% ACN with shaking for 60 minutes at room temperature. The washes were combined and peptides purified using C18 Zip Tips\textsuperscript{®} according to manufacturer's procedure (Millipore, Billerica, MA USA). Samples were dried by vacuum centrifugation and resuspended in 0.1% formic acid in HPLC-grade water for liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis.

### 2.2.3 Nano-LC-MS/MS

LC-MS/MS was performed as described in (Schweitzer et al., 2013). Briefly, samples were separated using an Eksigent nano-LC system equipped with two alternating peptide traps and a PicoFrit C18 column-emitter from New Objectives (Woburn, MA USA) and analyzed with a LTQ Orbitrap XL (Thermo Scientific, Inc., San Jose, CA USA). Samples were loaded onto the peptide trap with 2% acetonitrile (ACN) + 1% formic acid and eluted using a 60 minute linear gradient of 2-60% ACN + 1% formic acid. The data-dependent mode was used for the acquisition method with one precursor scan in the Orbitrap, followed by fragmentation of the 5 most abundant peaks in the CID, detected in the LTQ.
Resolution of the precursor scan was set to 60,000, scanning from 300-2000 m/z. Precursor peaks with a minimum signal count of 50,000 were dynamically excluded after two selections for 60 seconds within a range +/-10 ppm with Monoisotopic Precursor selection (MIPS) enabled. Previously found background peaks were included in a mass rejection list. Collision energy was set to 35 using an isolation width of 2 and an activation Q of 0.250.

Proteome Discoverer™ 1.2 software (Thermo Fisher Scientific) was used to generate peak lists. The method outline used in the spectrum selector node of the Proteome Discoverer™ software included the following scan event filters: fragmentation method, ionization source, and Unrecognized Activation Type Replacements parameters. All filters listed above and the precursor charge state (high/low), retention time, minimum peak count, total intensity threshold value were all set to default settings. The Max and Min precursor mass settings were 5000 Da and 350 Da, respectively. The NCBI.fasta database from http://ftp.ncbi.nih.gov was created on November 2011 with a file size of 118783 kB and was restricted to Homo sapiens. The search parameters were: Fully tryptic digested peptides only, two missed cleavages allowed, variable modifications of oxidized methionine (+15.995Da), and fixed modification of carboxymethyl cysteine (+58.00Da). The spectra were searched using Sequest™ algorithm in Proteome Discoverer 1.2 software (Thermo Scientific Inc.) using the following parameters: threshold for Data generation = 10000, peptide tolerance for ion trap ms/ms was 1.80Da, and precursor ion mass tolerance = 10 ppm. Database NCBI.fasta from http://ftp.ncbi.nih.gov was used.
for protein identification with at least two unique peptides and two missed cleavage sites allowed for identification. The criteria used for acceptance of peptide assignments are as follows: The minimal X-corr value for charge state > 7 was 1.0, minimal X-corr value for each charged state ranging from 1 to 7 was 1.50, 2.0, 2.25, 1.0, 1.0, 1.0 and 1.0 respectively. Two target values, a strict FDR of 0.01 and a relaxed FDR of 0.05 were applied for a decoy database search.

**2.2.4 Bioinformatic analyses**

Bioinformatic analysis was performed only on the 163 proteins identified in the study. Protein-protein interactions were investigated using STRING 9.1 (http://string-db.org) and contain experimental, co-expression, database, and text mining evidence using a confidence setting of greater than 0.4 (medium confidence). Centrality analysis was performed using Cytoscape v. 3.0.2 (www.cytoscape.org) and the CentiScaPe 2.0 plug-in. Node eccentricity, radiality and closeness within the network were individually ranked and the summed rank was used for the total centrality measure.

**2.2.5 Immunoblotting**

Protein concentrations of all samples were measured by BCA protein assay (Pierce Biotechnology). Samples were normalized by protein concentration with PBS, mixed 1:1 with 2x SDS-PAGE loading buffer, boiled for 10 min. and
separated by SDS-PAGE. Proteins were transferred to PVDF and detected by Western blot using the following primary antibodies: anti-ADAR (D-8), anti-CypA (C-14), anti-CypB (C-15), anti-DEK (H-300), anti-EEF1A1 (CBP-KK1), anti-GAPDH (6C5), anti-SERBP1/PAI-RBP1 (1B9), anti-HSP90AB (N-17), anti-PHB (H-80), anti-NAT10 (B-4), anti-RuvBL1 (N-15), anti-B23/NPM1 (H-106), and anti-C23/NCL (MS-3) were all obtained from Santa Cruz Biotechnology (Santa Cruz, CA USA); anti-TOPO IIa (31) from BD Biosciences (San Jose, CA USA); and the anti-LYAR antibody (P01) from Abnova (Taipei City Taiwan). HRP conjugated anti-rabbit, anti-mouse IgG secondary (GE Healthcare, Piscataway, NJ USA) or anti-goat IgG secondary (Sigma, St. Louis, MO USA) antibody were used to detect primary antibody and visualized by chemiluminescent staining (Pierce Biotechnology). Images were captured using radiographic film, scanned to computer, adjusted for brightness and contrast if necessary, and cropped for size.

2.3 RESULTS

2.3.1 Experimental approach

The objective of this study was to identify changes in the expression of nuclear proteins during the early steps of HIV-1 infection. We previously defined a method to produce large quantities of preintegration complexes by infecting C8166-45 cells with HIV-1 pseudotyped with the glycoprotein of vesicular
stomatitis virus (VSVg) using the spinoculation technique (Schweitzer et al., 2013). Because this method employs replication competent virus, a burst of co-infection occurs and a high level of reverse transcription and preintegration complexes are recovered at 20 hpi. We therefore hypothesized that this time-point may be enriched for factors that promote integration and/or activate virus transcription. The overall experimental approaches (both biological and bioinformatic) are summarized in Fig. 2.1. Nuclear fractions were isolated using a subcellular fraction kit which separates cells into: 1) the cytosol; 2) the membranes and membrane-bound organelles; 3) the nuclear; and 4) cytoskeletal and insoluble proteins using sequential lysis with differing buffer/detergent concentrations (Fig. 2.1A). In this protocol the nuclear fractions are treated with benzonase and therefore contain both soluble and nucleic acid bound proteins. Proteins were separated by one dimensional SDS-PAGE, parsed into 28 slices, digested in-gel with trypsin, and analyzed by LC-MS/MS. Proteins assignments were made using Proteome Discoverer software. Each biological replicate in the study represented an independent infection, nuclear isolation, and MS analysis.

Since the yield and purity of nuclear isolations is dependent on the method and cell type used (Ori, Andres-Pons, and Beck, 2014), the infection and quality of nuclear isolation of each biological replicate was confirmed by western blot analysis (Fig. 2.2). The integrity of the nuclear fractions was assessed by immunoblot of Glyceraldehyde-3-phosphate Dehydrogenase (GAPDH) and topoisomerase IIa (TOPO IIa; Fig 2.2. top two panels). GAPDH was consistently detected in the cytosolic, membrane, and insoluble fractions, but not the nuclei of
Figure 2.1 Schematic of experimental approach used to investigate nuclei of HIV-1 infected cells. Work-flow representative of one complete biological replicate is shown in (a). Bioinformatic analyses scheme is shown in (b).
Figure 2.2 Infection of C8166-45 cells used in proteomic screen. Western blot analysis of subcellular fractions of infected (left panels) and uninfected (right panels) C8166-45 cells used for proteomic analyses. Cells were fractionated into cytosolic (Cy), membrane/organelle (Me), nuclear (Nu), and cytoskeletal/insoluble (In) fractions. The proteins detected are listed on the left. Blots were performed for all biological replicates and representative blots are shown.
C8166-45 cells, indicating an absence of cytosolic or membrane/organelle contamination in the nuclear samples. TOPO IIa was detected predominantly in the nuclear fraction, with a small amount in the cytoskeletal/insoluble fraction, confirming that the nuclei remained intact during the extraction procedure. The infection of cells was confirmed by checking for the presence of the viral Integrase and Matrix proteins (Fig 2.2, bottom two panels). At 20 hpi, both viral proteins were readily detected in the infected samples. Integrase was detected in all 4 fractions, but predominantly found in the membrane fraction. Matrix was readily detected in the cytosolic, membrane, and nuclear fractions, but only slightly in the insoluble fraction. Interestingly we consistently observed different migrating forms of Matrix in each fraction. Three forms of Matrix were observed: The expected size (17 kDa), a slightly smaller form of Matrix (15-16 kDa) in the membrane/organelle and nuclear fractions, and a second much smaller MW form (~7-8 kDa) in nuclei.

2.3.2 Database search and protein identification

The overall summary of the data work-flow is shown in Fig. 2.1B. Peptides were identified from the MS/MS scans with Proteome Discoverer 1.2 software utilizing the NCBI protein database. The total number of peptides identified across all experimental replicates was 8093 and 5339 (control and infected, respectively; Fig. 2.3A). Protein assignments were initially designated by NCBI GenInfo Identifier (GI) number for each peptide. Based on individual GI
Figure 2.3 Summary of MS data. Graphical representation of the overall number of peptides identified (a), and proteins identified (b) from all biological replicates. (C) Chart of the number of cellular proteins identified in two or more biological replicates across the study.
numbers, a total of 3868 protein assignments were made. Collating the data to account for duplicate identifications across biological replicates resulted in the identification of 1738 proteins in the control set and 1054 in the infected set (Fig. 2.3B). The ratio of identified proteins (control/infected) was 1.65, suggesting that the final results would contain an overrepresentation of factors in the control samples and possibly mask some low abundant proteins in the infected samples.

Next, the data sets were purged of all non-human records, any records marked as “removed” in the NCBI database, and all keratin records, which were considered background contamination.

The data was further transformed for additional bioinformatic analyses. The remaining NCBI GI numbers were converted to Uniprot IDs using the SwissProt knowledgebase. This step removed protein redundancies and formatted the data set for Centrality analysis (see below). Finally, the Uniprot IDs were converted to NCBI Gene IDs. This data set was used for the subtractive analysis to identify unique proteins, the protein-protein interaction analysis, as well as the knowledge-based analyses shown on the right of Tables 2.1-2.3.

2.3.3 Identification of uniquely expressed proteins

The main objective of our study was to identify proteins which were enriched or depleted in HIV-infected nuclei. A simple subtractive analysis was used to measure relative protein abundance and identify differentially expressed factors. The control cell data set was used as reference to identify unique proteins in the HIV-1-infected nuclear samples and the vice versa performed to discover unique
proteins in the uninfected samples. The thresholds of identification were set as at least two peptides per protein identified with not less than 99% confidence, which favored the identification of moderately to highly abundant proteins, and the criterion that any given protein must be identified in at least two or more of the five biological replicates. Based on these criteria, our subtractive analysis narrowed list of proteins to 163 proteins: 13 and 38 proteins were unique to the infected and control data sets, respectively; and 112 proteins were common to both (Fig. 2.3C) making the list of candidates for follow-up validation manageable. Forty three of these proteins are not currently listed in the HIV interaction database. Table 2.1 shows the 13 proteins found only in infected cell nuclei data set. The list is ranked by the total Proteome Discoverer™ protein score across all biological replicates, which is the sum of the scores of the individual peptides from any given protein identified in the analyzed samples. The protein score provided by Proteome Discoverer™ correlates with a spectral count which is relative to the abundance of a peptide (Bridges et al., 2007; Stevenson et al., 2009; Tu et al., 2014). Therefore a higher protein score reflects higher peptide abundance thus higher relative quantity of a protein across the samples.

A knowledge-based analysis with several previous RNAi (Brass et al., 2008; König et al., 2008) and proteomic studies (Chertova et al., 2006; Haverland, Fox, and Ciborowski, 2014; Jager et al., 2011; Kraft-Terry et al., 2010; Raghavendra et al., 2010), as well as the NCBI HIV interaction database (Fu et al., 2009) was performed to identify factors with known HIV interactions. Factors identified previously in those studies are denoted in the columns to the right of the table.
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene Name</th>
<th>PD Score</th>
<th>Previous Studies[^4]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1915</td>
<td>EEF1A1, eukaryotic translation elongation factor 1 alpha 1</td>
<td>99.48</td>
<td>□</td>
</tr>
<tr>
<td>7913</td>
<td>DEK oncogene</td>
<td>66.79</td>
<td>□</td>
</tr>
<tr>
<td>55646</td>
<td>LYAR, Ly1 antibody reactive</td>
<td>63.01</td>
<td>□</td>
</tr>
<tr>
<td>1973</td>
<td>EIF4A1, eukaryotic translation initiation factor 4A1</td>
<td>60.76</td>
<td>□</td>
</tr>
<tr>
<td>1786</td>
<td>DNMT1, DNA (cytosine-5-)-methyltransferase 1</td>
<td>60.69</td>
<td>□</td>
</tr>
<tr>
<td>10212</td>
<td>DDX39A, DEAD (Asp-Glu-Ala-Asp) box polypeptide 39A</td>
<td>54.05</td>
<td>□</td>
</tr>
<tr>
<td>55749</td>
<td>CCAR1, cell division cycle and apoptosis regulator 1</td>
<td>48.15</td>
<td>□ □</td>
</tr>
<tr>
<td>3315</td>
<td>HSPB1, heat shock 27kDa protein 1</td>
<td>31.06</td>
<td>□</td>
</tr>
<tr>
<td>26135</td>
<td>SERBP1, SERPINE1 mRNA binding protein 1</td>
<td>23.38</td>
<td>□</td>
</tr>
<tr>
<td>9276</td>
<td>COPB2, coatamer protein complex, subunit beta 2 (beta prime)</td>
<td>22.13</td>
<td>□ □ □</td>
</tr>
<tr>
<td>6625</td>
<td>SNRNP70, small nuclear ribonucleoprotein 70kDa (U1)</td>
<td>21.81</td>
<td>□</td>
</tr>
<tr>
<td>26986</td>
<td>PABPC1, poly(A) binding protein, cytoplasmic 1</td>
<td>20.29</td>
<td>□ □ ^4</td>
</tr>
<tr>
<td>3418</td>
<td>IDH2, isocitrate dehydrogenase 2 (NADP+), mitochondrial</td>
<td>18.27</td>
<td>□ □ □</td>
</tr>
</tbody>
</table>

[^1]: Ranked by Proteome Discoverer Score
[^2]: NCBI Gene name linked to G1 number identified in study. Unknown protein hits were identified by BLAST homology search.
[^3]: The sum total of Proteome Discoverer (PD) scores for all identifications across the study.
[^4]: The summary of previous HIV-1 protein interaction studies cited in text.
Eight of the proteins in Table 2.1 have previously described interactions with HIV, five of those were interactions defined on the NCBI website.

Table 2.2 lists the 15 candidate proteins from the control samples with the highest total Proteome Discoverer™ score. These were factors found only in the control samples and may represent proteins that are down-regulated or removed from the nucleus during HIV infection of the cell. As seen in Table 2.2, a large percentage of the top 15 are previously described HIV factors. Overall, 30 of the 38 proteins in the control dataset have HIV-1 interactions defined in the referenced studies and/or the NCBI database.

2.3.4 Identification of Other Candidate factors

Focusing solely on unique proteins using stringent thresholds may miss important factors with less moderate changes in expression. To expand the number of proteins of interest, we sought to identify proteins differentially regulated between the control and infected nuclei. To do this we utilized an approach we used previously in a comparative proteomic study of HIV preintegration complexes (Schweitzer et al., 2013) in which we compared the total Proteome Discoverer™ score of each protein assignment across the infected and control samples to identify up- and down-regulated proteins. We distinguished those proteins with a greater than two-fold change in total score between the infected and control nuclei. These are listed in Table 2.3 and the
### TABLE 2.2: Top unique proteins identified in control cell nuclei

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene Name</th>
<th>PD Score</th>
<th>Previous Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>58</td>
<td>ACTA1, actin, alpha 1, skeletal muscle</td>
<td>302.04</td>
<td></td>
</tr>
<tr>
<td>3008</td>
<td>HIST1H1E, histone cluster 1, H1e</td>
<td>109.82</td>
<td></td>
</tr>
<tr>
<td>55226</td>
<td>NAT10, N-acetyltransferase 10 (GCN5-related)</td>
<td>53.54</td>
<td></td>
</tr>
<tr>
<td>3010</td>
<td>HIST1H1T, histona cluster 1, H1t</td>
<td>47.24</td>
<td></td>
</tr>
<tr>
<td>117159</td>
<td>DCD, darmacidin</td>
<td>41.03</td>
<td></td>
</tr>
<tr>
<td>56660</td>
<td>PRPF40A, PRPF40 pre-mRNA processing factor 40 homolog A</td>
<td>36.27</td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>ADAR, adenosine deaminase, RNA-specific</td>
<td>65.18</td>
<td></td>
</tr>
<tr>
<td>3106</td>
<td>HLA-B, major histocompatibility complex, class I, B</td>
<td>31.81</td>
<td></td>
</tr>
<tr>
<td>4869</td>
<td>NPM1, nucleosomin</td>
<td>27.52</td>
<td></td>
</tr>
<tr>
<td>4926</td>
<td>NUMA1 (nuclear mitotic apparatus protein 1)</td>
<td>25.28</td>
<td></td>
</tr>
<tr>
<td>27044</td>
<td>SND1 staphylococcal nuclease and tudor domain containing 1</td>
<td>25.1</td>
<td></td>
</tr>
<tr>
<td>292</td>
<td>SLC25A5, solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 5</td>
<td>24.03</td>
<td></td>
</tr>
<tr>
<td>10291</td>
<td>SF3A1, splicing factor 3a, subunit 1, 120kDa</td>
<td>23.24</td>
<td></td>
</tr>
<tr>
<td>6742</td>
<td>SSEP1, single-stranded DNA binding protein 1</td>
<td>22.75</td>
<td></td>
</tr>
<tr>
<td>255626</td>
<td>HIST1H2BA, histone cluster 1, H2ba</td>
<td>21.45</td>
<td></td>
</tr>
</tbody>
</table>

1. Ranked by Proteome Discoverer Score
2. NCBI Gene name is linked to GI number identified in study. Unknown protein hits were identified by BLAST homology search.
3. The sum total of Proteome Discoverer (PD) scores for all identifications across the study.
4. The summary of previous HM protein interaction studies cited in text.
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene Name</th>
<th>PD Ratio</th>
<th>Previous Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>7153</td>
<td>TOP2A, DNA topoisomerase II, alpha</td>
<td>6.558</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>ACTG1, actin, gamma 1</td>
<td>3.898</td>
<td></td>
</tr>
<tr>
<td>4691</td>
<td>NCL, nucleolin</td>
<td>3.398</td>
<td></td>
</tr>
<tr>
<td>10382</td>
<td>TUBB, tubulin, beta class I</td>
<td>3.022</td>
<td></td>
</tr>
<tr>
<td>3148</td>
<td>HMGE2, high mobility group box 2</td>
<td>2.736</td>
<td></td>
</tr>
<tr>
<td>3326</td>
<td>HSP90AB1, heat shock protein 90kDa alpha (cytosolic), class B member 1</td>
<td>2.709</td>
<td></td>
</tr>
<tr>
<td>8243</td>
<td>SMC1A, structural maintenance of chromosomes 1A</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>8607</td>
<td>RUVBL1, RuvB-like 1</td>
<td>2.479</td>
<td></td>
</tr>
<tr>
<td>10992</td>
<td>SF3B2, splicing factor 3b, subunit 2</td>
<td>2.369</td>
<td></td>
</tr>
<tr>
<td>6426</td>
<td>SRSF1, serine/arginine-rich splicing factor 1</td>
<td>2.364</td>
<td></td>
</tr>
<tr>
<td>3017</td>
<td>HIST1H2BC, histone cluster 1, H2bc</td>
<td>2.314</td>
<td></td>
</tr>
<tr>
<td>7112</td>
<td>TMPO, thymopoeitin</td>
<td>2.302</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>ACTN4, alpha actinin 4</td>
<td>2.268</td>
<td></td>
</tr>
<tr>
<td>5479</td>
<td>PPIB, peptidylprolyl isomerase B (cyclophilin B)</td>
<td>2.235</td>
<td></td>
</tr>
<tr>
<td>5245</td>
<td>PHB, prohibitin</td>
<td>0.414</td>
<td></td>
</tr>
</tbody>
</table>

1. Ranked by Proteome Discoverer Score
2. NCB1 Gene name linked to GI number identified in study. Unknown protein hits were identified by BLAST homology search.
3. The ratio of Proteome Discoverer (PD) scores (Inf/Cont) for all identifications across the study.
4. The summary of previous HIV-protein interaction studies cited in text.
ratio of the infected/control score is provided. Fourteen proteins were found with a >2-fold higher score in infected nuclei, whereas only one protein, Prohibitin, was found to be enriched in control cell nuclei.

### 2.3.5 Bioinformatic analysis

Next, protein-protein interaction and centrality analysis were performed on the candidate proteins to identify important pathways and protein networks within the data set that might be missed by looking solely at individual proteins. First, protein-protein interaction analysis was performed with StringDB (v9.1, searched Jan. 6, 2014 (Franceschini et al., 2013); Fig. 2.4A). The algorithm produced four distinct clusters of interacting proteins (denoted with roman numerals in Fig. 2.4A). A small cluster (I) at the center of the map connected to the other three clusters. The three central proteins in this cluster were Nucleolin (NCL), Nucleophosmin (NPM1), and RuvBL1. NPM1 alone connected to all clusters. Cluster II included several factors associated with cell homeostasis, including Heat Shock Protein 90 (HSP90AB1), Actin (ACTG1), and PolyA Binding Protein (PABPC). The largest cluster (III) contained eight proteins, and was comprised of numerous splicing factors including SRSF1, SF3A1, and SF3B2. The fourth cluster contained four proteins, TOP2A, TMP0, HMGB2, and DEK, associated with chromosomal and nuclear architecture. To complement the protein-protein interaction analysis we also measured the summed ranked for Centrality for the candidate proteins ((Smoot et al., 2011); Fig. 2.4B). This analysis scored proteins
based on their position relative to the center of cellular pathways. Identifying these factors may point toward pathways important for HIV infection. Consistent with the protein-protein interaction analysis, NPM1 was the factor with the highest rank and three of the top four proteins were from the central cluster (I) in Fig. 2.4A. Combined the bioinformatic results indicated that HIV infection alters the nuclear expression of proteins involved in splicing, chromosomal and nuclear architecture, as well as nuclear homeostasis.

2.3.6 Validation of MS data

The proteomic approach we used made protein identifications and rankings based on probability algorithms, necessitating the biochemical validation of the phenotypes of candidate proteins. To confirm protein assignments we examined the expression profiles of selected candidate proteins in the subcellular fractions of C8166-45 cells by immunoblot. For these studies all four subcellular compartments were analyzed since not all the candidate proteins are reported in the literature to be exclusively nuclear. Equal quantities of fraction were analyzed by SDS-PAGE and immunoblot for candidate proteins. The factors that were detected as predicted by MS in two or more biological replicates are shown in Fig. 2.5.

The factors from Table 1 that showed distinct phenotypes are presented in Fig. 2.5A. Eukaryotic translation elongation factor 1 alpha 1 (EEF1A1) was the highest ranked protein based on Proteome Discoverer score. This was not
Figure 2.4 Bioinformatic analysis of candidate proteins. (A) Cluster analysis of unique and differentially regulated proteins. Analysis was performed using STRING db (Jan 17, 2014) to generate a protein-protein interaction network containing experimental, co-expression, database, and textmining evidence. Lines between factors denote known or predicted interactions and stronger associations are represented by thicker lines. (B) Rank of centrality of the candidate factors. Centrality was analyzed for all factors identified in unique and differentially regulated data sets (see text). Bars are coded to data set as indicated by legend.
Figure 2.5 Validation of candidate factors identified using MS.
Western blot confirmation of candidate factors identified in MS studies.
Proteins detected are listed on the left and the subcellular fractions are
the same as described in Fig 1. (A) Proteins identified as unique to HIV-1
infected cells (Table 1). (B) Factors from unique control data set (Table
2). (C) Expression of protein predicted to be changed in infected nuclei (Table
3). (D) Nucleolar factors from cluster I of Fig. 4A bioinformatic analysis.
Blots are representative of two or more biological replicates.
surprising as it interacts with several HIV proteins including Matrix (Cimarelli and Luban, 1999), Rev (Naji et al., 2012), Tat (Wu-Baer, Lane, and Gaynor, 1996), Integrase (Allouch and Cereseto, 2011; Warren et al., 2012), and Reverse Transcriptase (Warren et al., 2012). The level of EEF1A1 was confirmed to be higher in the nuclei of infected versus uninfected C8166-45 cells (Fig. 5A, top panels, compare lane 3 to 7). In addition to its predicted size, we observed the presence smaller MW bands in the nuclei of infected cells. These smaller bands may represent be products of cleavage as EEF1A1 is cleaved by HIV-1 protease (Cimarelli and Luban, 1999). Several other proteins from Table 2.1, including DEK oncogene (DEK), Ly1 Antibody Reactive (LYAR), and Serpin Peptidase Inhibitor member 1 (SERPINE1) mRNA binding protein 1 (SERBP1) were also validated by immunoblot to be present at substantially higher levels in infected cell nuclei (Fig. 2.5A). SERBP1 was originally identified as a SERPINE1 mRNA interacting protein (Heaton et al., 2001). SERPINE1 is a regulatory protein of the fibrinolytic system that acts on tissue plasminogen activator and urokinase to limit activation of the fibrinolytic pathway (Binder et al., 2002). In addition to the increase in nuclear SERBP1 in infected cells, we also discovered that levels of SERBP1 in the membrane/organelle fraction were substantially reduced in infected cells compared to uninfected cells (4th panel of Figure 2.5A, compare lane 2 to 6), suggesting deregulation of its intracellular localization.

A number of the candidate proteins found at the threshold cutoff to be unique to control data set or predicted to be down regulated in infected cells were evaluated by immunoblot. Prohibitin (PHB) and N-acetyltransferase 10 (NAT10)
were the only proteins confirmed to have lower expression in the nuclei of HIV infected C8166-45 cells by immunoblot (Fig. 2.5B). PHB levels were also decreased in the cytoskeletal/insoluble fraction of infected cells, but showed increased expression in the membrane/organelle fraction. PHB is a multifunctional protein involved in cell signaling, growth and tumor suppression. It can be a negative regulator of the cell cycle as well as regulate mitochondrial metabolism and cellular aging (Coates et al., 2001). Previous reports of PHB expression during HIV infection are mixed. PHB appears to be over-expressed when cells were exposed to HIV envelope protein (Molina et al., 2007), but Vpr expression reduces PHB levels (He et al., 2009). Recently PHB was found to interact directly with HIV-1 envelope protein (Emerson et al., 2010; Jager et al., 2011). Overall NAT10 levels in the nuclei of C8166-45 cells were low compared to the insoluble fraction, and decrease further upon HIV infection (Fig. 5B, lower panel lanes 3 vs. 7). NAT10 is a lysine acetyltransferase involved in histone acetylation, telomerase activity, and microtubular organization and nuclear architecture (Larrieu et al., 2014; Lv et al., 2003; Shen et al., 2009). Inhibition of NAT10 reduces DNA damage signaling and improves cell health, suggesting that down-regulation during HIV infection may support T-cell survival and enhanced virus replication.

Fig. 2.5C shows the immunoblot analysis of two candidate proteins from Table 2.3, CypB and HSP90AB. Consistent with MS data, peptidyl prolyl isomerase B, or cyclophilin B (CypB) showed higher nuclear levels in HIV infected cells (top panels). Two additional higher MW bands were also observed
in the nuclear fraction of infected cells (lane 3). CypB levels were also higher in the cytosolic fractions of infected cells (lane 1 vs. 5). CypB is a member of the immunomodulatory cyclophilin family. It shares significant homology in its core domain with Cyclophilin A (CypA), a well-described HIV-1 co-factor that interacts with HIV-1 Capsid and is required for efficient HIV replication. However, the expression of CypA was distinct from CypB (middle panels). The overall levels of CypA appeared lower in HIV infected cells, but it was absent in the nuclear fractions of both infected and control cells. Combined this demonstrated that despite their homology, the expression of CypB and CypA are differentially regulated in HIV infected cells. The overall expression of HSP90AB, including in the nucleus, was higher in HIV infected cells, consistent with infection-induced cellular stress (lower panels). Notably, HSP90 can promote HIV-1 replication by increasing HIV transcription (Roesch et al., 2012).

Finally we focused on the proteins in cluster I of Fig. 2.4A- NCL, NPM1, and RuvBL1. Only low levels of NCL were found in uninfected C8166-45 cells, but upon HIV infection there was a distinct increase in its nuclear levels. (Fig. 2.5D, top panels). Two distinct bands were detected in the NCL blots, possibly consistent with auto-cleavage that induces apoptosis (Li et al., 2009). In multiple replicates we saw a decrease of NPM1 in the cytosolic and membrane/organelle fractions, but only a slight increase in nuclei (middle panels). RuvBL1 exhibited slightly higher levels in infected C8166-45 nuclei, as well as an increase in the cytoskeletal/insoluble fraction (lower panels). It was less abundant in the membrane/organelle fraction of infected cells and predominantly expressed in
the insoluble fraction where its levels appeared to increase during HIV infection. Overall, these data demonstrate that HIV infection modulates nucleolar protein expression.

2.3.7 Independent confirmation of localization phenotypes

The above experiments investigated the nuclear expression of the candidate proteins as predicted by MS in the same cells used for the MS analysis. To extend these observations and confirm the altered nuclear expression of the candidate proteins during HIV-1 infection we investigated each protein in an independent infection model using Jurkat E6-1 T-cells. For these experiments we performed a series of time-course infections and harvested cells at 24, 48, and 72 hpi to provide more data points of analysis. Cells were fractioned into subcellular compartments as before. Infections were monitored by anti-Gag immunoblot (Fig. 2.6A, top panel) and we consistently detected p55 Gag and p24 capsid beginning at 24 hpi in the cytosolic and membrane fractions. The integrity of nuclear fractions was again confirmed by the absence of GAPDH and the presence of TOPO IIa only in nuclear fractions (Fig. 2.6A, bottom two panels).

All of the proteins mentioned above were investigated in the Jurkat experiments. Those proteins that showed a notable phenotype are shown in Fig. 2.6B-D. CypB and SERBP1 both showed an increase in the nuclei of HIV infected cells verses uninfected cells, confirming the phenotype we observed in C8166-45 T-cells. In addition to an increased nuclear accumulation, CypB also
Figure 2.6 Independent validation of candidate proteins. Immunoblot evaluation of candidate factors during time course infection of Jurkat E6-1 cells. Infected Jurkat cells were harvested and fractionated at hours post infection (hpi) indicated at top. Proteins listed at left were evaluated as described in Figs 1 and 5. Control blots (a), factors upregulated in infected (b) and control (c) cells, and cluster l factors (d) are shown. Data is representative of at least two independent infection experiments.
showed a slight increase in cytoplasmic levels in infected Jurkat cells. Notably, we did not observe the presence of the large MW bands in Jurkat lysates, suggesting that those forms were unique to C8166-45 cells. As with the C8166-45 experiments, we also looked at the distribution of CypA. In the Jurkat cells CypA localized distinctly to the cytoplasmic fraction, and its levels did not appear to change substantially. SERBP1 showed a very similar response to HIV infection in the Jurkat cells as with the C8166-45 cells. In uninfected cells it was present in both the cytosolic and membrane/organelle fractions. As HIV infection progressed there was a loss of membrane/organelle-associated SERBP1 and a gain of SERBP1 in the nuclear fraction, suggesting that its intracellular localization changed, but the overall levels of SERBP1 did not.

In contrast to the C8166-45 cells, PHB was not detected in the nuclei of Jurkat cells. Low levels were detected in the membrane/organelle fraction of the Jurkat cells, and it increased during HIV-1 infection (Fig. 2.6C, top panels). The phenotype of NAT10 during HIV infection of Jurkat cells was similar to that seen with the C8166-45 cells- expression in the nuclear and insoluble fractions, both of which were reduced during HIV replication (middle panels). Of note, ADAR, a candidate found unique in control nuclei that we were unable to validate in the C8166 samples, showed a reduction in nuclei over the course of Jurkat cell infection (bottom panels).

The Cluster I factors- NCL, NPM1, and RuvBL1 are shown in Fig. 2.6D. In contrast to the C8166-45 cells, NCL showed strong cytosolic localization in uninfected Jurkat cells. This is consistent with previous observations of leukemic
cells (Otake et al., 2007). In agreement with the C8166-45 infection data, the nuclear level of NCL also increased in infected Jurkat E6-1 cells (Fig 2.6D, top panels). Notably however, only a single band was detected in the Jurkat cells compared to the C8166-45 cells. NPM1 showed a decrease in the cytosolic and membrane/organelle fractions similar to the C8166-45 cells, but we did not see any remarkable change in nuclei (middle panels). In contrast to the C8166-45 infection, RuvBL1 levels appeared to decrease in the nuclei of HIV-1 infected Jurkat cells and the higher molecular weight isoform decreased in the insoluble fraction (bottom panels).

### 2.4 DISCUSSION

Combined antiretroviral therapy effectively controls HIV infection, but it does not lead to virus eradication. Moreover, resistance to available antivirals continues to rise. Targeting of the host cell factors or pathways is a promising approach to address these issues and augment existing therapies. Numerous MS studies have sought to define the role of host factors during HIV-1 infection (Birse et al., 2013; Chen, Wang, and Chan, 2012; Engeland et al., 2014; Haverland, Fox, and Ciborowsk, 2014; Naji et al., 2012; Schweitzer et al., 2013). The approach used in this study, examining the changes in the proteome of T-cell nuclei upon HIV-1 infection, aimed to identify nuclear proteins important for HIV pathogenesis either as factors that promote HIV replication or ones that combat virus infection. Such factors may reflect the translocation of a protein in
or out of the nucleus, or *de novo* synthesis or degradation of a resident nuclear protein. Overall, our data analysis strategy identified fifty one unique factors in either infected or control nuclei and an additional 15 differentially regulated proteins between the two conditions. Forty three of these factors are not currently listed on the NCBI HIV interaction database. The validation experiments distinguished 13 candidate proteins that had altered expression patterns in either the C8166-45 or Jurkat E6-1 cells by immunoblot.

Our goal was to map the global changes in protein expression in HIV infected nuclei. Two other proteomic studies have focused on the nuclear compartment, specifically changes to the nucleolus upon Tat expression (Jarboui et al., 2012), and alterations in the nuclear membrane of HIV-infected T-cells (Monette, Pante, and Mouland, 2011). Our data had little overlap with the results of the Tat over-expression study. This might be because of our use of replicating virus, but may also suggest that our study captured more factors related to nuclear import and/or integration. Consistent with that, we saw a greater overlap with the nuclear membrane study which observed changes in NCL, NPM1, and CypB (among other factors). Notably, our approach likely removed the nuclear membrane as we did not detect nucleoporins or other nuclear membrane proteins. Interestingly, both ADAR and HSP90AB were common factors to all three nuclear studies, suggesting they have important roles in the nuclear compartment during HIV infection.

The NCBI HIV interaction database continues to grow. Overall, a majority of the proteins identified in this study to be differentially regulated in both uninfected
and infected samples have been previously reported as HIV-1 dependent factors (Tables 1-3). Analysis of protein-protein interactions identified four clusters of factors. The central cluster contained NCL, NPM1, and RuvBL1, which were three of the top four proteins in the Centrality analysis of the data set. These proteins are components of nucleoli and involved in ribosomal synthesis (Penman, Smith, and Holtzman, 1966; Perry, 1962), cell cycle progression (Liu and Yung, 1999), as well as cell signaling (Inder et al., 2009). In our studies, NCL was found to be upregulated in the nuclei of both C8166-45 and Jurkat cells.

Several studies suggest that NCL is an important HIV factor: It is present in HIV-1 virions possibly through interaction with the HIV RNA via its 4 RNA binding domains (Ghisolfi-Nieto et al., 1996); and it also interacts with HIV-1 Envelope protein (Nisole, Krust, and Hovanessian, 2002), and both the Gag polyprotein and Matrix in budding virions (Ueno et al., 2004). Both NPM1 and RuvBL1 also showed changed expression in HIV infected cells, although the changes were more notable in the cytosolic and membrane/organelle fractions. Combined, these data suggest strongly that this cluster of proteins play an important role in HIV replication or pathology and warrant further investigation. Interestingly, another factor our screen identified, LYAR, reportedly plays a critical role in the nucleolar stabilization of NCL. The nuclear levels of LYAR, like NCL, increased during infection. LYAR is a nucleolar zinc finger protein initially identified in T-cell leukemia. The exact function of LYAR is not understood, but it has been implicated in the modulation of cell growth, tumorgenesis, self-renewal of embryonic stem cells, and medulloblastoma (Li et al., 2009; Su, Hershberger,
and Weissman, 1993; Swartling et al., 2010). Thus, it is tempting to speculate that the upregulation of LYAR in nuclei may serve an anti-apoptotic mechanism in HIV infected T-cells.

CypB showed increased nuclear levels during infection of both C8166-45 and Jurkat cells. CypB is a member of the cyclophilin family, known for inducing immunosuppression when complexed with Cyclosporine A (CsA). CypA and B are the two primary members of the family and share significant homology in their core domains. CypB has longer N- and C-termini- the CypB N-terminus contains an ER leader and nuclear localization sequence (Rycyzyn et al., 2000); and the C-terminus has an ER retention signal (Arber, Krause, and Caroni, 1992). CypB has been shown to be involved multiple cellular processes such as the interferon-β response to viral infection through modulating IRF3 translocation to the nucleus (Obata et al., 2005). CypB is also essential for the nuclear retrotransport of Prolactin (PRL) following internalization of Prolactin-bound receptor (PRL/PRLr). Finally, CD147 stimulation by CypB induces intracellular Ca^{2+} flux stimulating T cell proliferation and produces a chemotactic response in primary human neutrophils (Yurchenko et al., 2001).

Both CypA and CypB were identified as HIV-1 Gag binding proteins in a yeast two-hybrid cDNA screen (Luban et al., 1993). Subsequent studies found CypB binds at a stronger affinity due to a hydrophobic N-terminal ER signal sequence rendering the Gag-CypB interaction less sensitive to disruption by CsA compared to Gag-CypA complexes (Braaten, Ansari, and Luban, 1997). CypB has not been investigated at as an independent factor, likely due to the assumption that with its
homology to CypA it has similar functions. Our results demonstrate clearly that
the expression of CypB is independently regulated during HIV infection, arguing
that it may play a distinct role in HIV infection. Studies in Chapter IV investigate
what that role may be.

SERBP1 showed a consistent phenotype in both C8166-45 and Jurkat cells,
suggesting that it plays an important role in HIV-1 infection. Consistent with that
idea it was previously identified in a proteomic screen of HIV-1 preintegration
complexes (Raghavendra et al., 2010). SERBP1 is an RNA binding protein that
binds the cyclic nucleotide-responsive sequence of the SERPINE1 mRNA and
may regulate its stability (Heaton et al., 2001), although its exact effect is not
known. SERPINE1 is a key component of the fibrinolytic pathway that regulates
blood coagulation (Binder et al., 2002). It blocks tissue plasminogen activator and
urokinase to limit activation of the fibrinolytic pathway. Blood levels of SERPINE1
are elevated in HIV Associated Lipidodystrophy Syndrome which may be the
cause of the associated thrombocytosis (He et al., 2005). The regulatory
mechanisms for the expression of SERPINE1 are not fully understood. It is
tempting to speculate that HIV-1 induced expression of SERPBP1 expression
may be involved in SERPINE1 regulation. However, we did not detect the 45 kDa
form of SERPINE1 in either the C8166-45 or Jurkat E6-1 cells (data not shown).
SERBP1 also interacts with the C-terminal region of the chromo-helicase-DNA-
binding domain protein 3 (CHD-3), which suggests it could play a role in the
regulation of other nuclear functions such as chromatin remodeling (Lemos et al.,
2003). Interestingly, both SERPINE1 and CHD-3 have been shown to interact
with HIV-1 Tat. A recent study found that SERBP1 localizes to nucleoli upon stress (Lee et al., 2014). Ongoing studies are working to determine what role the up-regulation of nuclear SERBP1 plays during HIV infection.

As these studies show, mass spectrometry is a powerful tool to uncover cellular networks and pathways associated with disease. As with all "omics" studies, there are limitations to this study, including the use of a single cell line, single virus clone, and a high MOI. In general, at least four issues heavily influence progress in discovering cellular proteins that might be used as potential targets for new therapies in infectious disease. First is the high dynamic range of abundance between constitutively expressed structural or housekeeping proteins (on the high end) versus specialized, less abundant regulatory factors (on the low end). Second is the short time span when the differential expression of cellular proteins can be measured and detected during infection. Third is the distribution of specific factors between cellular compartments where proteins perform their function(s). Fourth is the accurate representation of in vivo pathology by in vitro models. The first issue is addressed at the level of instrumentation and methods of data acquisition. New technologies, such as SWATH-MS (Haverland, Fox, and Ciborowski, 2014), show promise for high resolution mapping of a wide range of expression levels in single samples. The solution for the second issue lies in selecting an early time point, to capture early proteomic changes within the cell, while the third issue can be resolved by interrogating organelle proteomes. The solution to the fourth issue is less transparent. Cell lines are advantageous in proteomic experiments for their reproducible infections and homogenous
proteomes as compared to primary cells which can suffer from donor to donor variability. Moreover, a more consistent level of infection can be achieved across biological replicates, thus addressing issue two. However, as demonstrated by our data, factors may show disparate expression and/or subcellular localization between cell lines. Even with the small number of proteins we looked at, several proteins showed different expression patterns between the C8166-45 and Jurkat E6-1 cell lines. These differences may result from the cause of transformation-C8166-45 cells derive from an HTLV-infected leukemia (Salahuddin et al., 1983), whereas the Jurkat cells are a clonal derivative from a non-viral leukemia (Weiss, Wiskocil, and Stobo, 1984). Additional studies utilizing primary peripheral blood mononuclear cells or monocyte-derived macrophages may more closely mimic the in vivo situation in infected individuals. Again, new technology and approaches, such as SWATH-MS should be able to address the issues of donor-to-donor variation or variability in infection kinetics between donors to identify factors important for HIV infection in vivo.

2.5 FUTURE STUDIES

This study provides a template for a focused subcellular proteomics study. The subtractive analysis presented effectively reduced the total candidate proteins from over 3500 to a more manageable 163. This approach was highly effective to significantly reduce the candidate population, but may have been too simplistic, and potential factors may have been missed. Reevaluation of the data
using more advanced bioinformatic analysis may yield additional factors. STRING-db PPI provided an overarching view of the altered proteome interactions; however cluster analysis was not performed. Cluster analysis could provide evidence of strong interactions between candidate proteins. In addition, advanced pathway analysis using PANTHER or similar software, would provide additional insight into the nuclear changes during HIV infection. To confirm factors have a similar *in-vivo* effect, similar fractionation studies could be performed on primary lymphocytes. Finally, a biological evaluation of candidate proteins using overexpression and knockdown studies could be used to determine the overall impact on HIV replication.
CHAPTER III: SWATH-MS Analysis of the Host Cell Protein Response
during HIV-1 Infection of T-cells

3.1 BACKGROUND

Human immunodeficiency virus type 1 (HIV-1), like all viruses, is an obligate intracellular parasite. As such, viruses must induce cellular changes to establish infection and promote replication. At the same time, host cells initiate innate responses to counter the pathogen as well as signal nearby cells to the presence of the microbe. In response, they must also alter the cellular environment to combat these innate cellular pathways. Combined, these changes result in the overall pathology displayed by HIV-1. Investigation of these cellular responses will identify critical pathways for virus replication and may offer new targets for pharmaceutical disruption of virus replication.

Despite the availability of a variety of therapies, HIV-1 remains a significant global pathogen, as the disease remains incurable. The prospect of life-time treatment coupled with the number of infected individuals makes the development of drug resistant viruses an important concern. The identification of new targets for therapy must be continued until a cure is found. One approach to finding new targets for therapy is to identify cellular factors critical for virus replication. A large number of genetic, proteomic, and other "-omics" studies have sought to identify such factors. Proteomic studies examining the host cell response have typically relied on so-called "shot-gun" approaches. These have included studies of whole cell lysates, individual cellular compartments, or the
Sequential Window Acquisition of all Theoretical fragment-ion spectra mass spectrometry (SWATH-MS) is an alternative approach to proteome analysis initially described by Gillet et al. (Gillet et al., 2012). SWATH-MS is a quantitative, label-free MS approach which combines traditional "shotgun proteomics" discovery method of LC-MS/MS with the high specificity of a targeted data extraction strategy. First, a spectral library is constructed using traditional data-dependent acquisition (DDA) where the selected fragment ion spectra are assigned to their peptide sequences by sequence database searching. Next, samples are ran through a second LC-MS/MS analysis in the data-independent mode (DIA) in which the MS cycles through precursor acquisition windows (400-1200 m/z) covering the range of which most peptide precursors fall. During each cycle, all precursors from a given window (25 Da) will be fragmented and a complete, high accuracy fragment ion spectrum will be recorded. The cycle is repeated and the same range is fragmented for a time-resolved of spectra of fragment ions. These ions are then compared against the spectral library constructed in the first step for peptide and subsequent protein identification.

In this study we utilized the SWATH-MS method to analyze changes to the subcellular proteome of HIV infected T-cells. Specifically, we investigated the cytosolic, membrane and organelle, and nuclear fractions of cells at 48 hpi. A total of five biological replicates were performed. The spectral data was log\(_{10}\) transformed to attain normal distribution and a z-Score based statistical method
used to identify changes in protein abundance in each fractions. A total of 6828 proteins were identified and 1167 of these factors exhibited statistically significant alteration in HIV-infected cells. Using multitude of bioinformatic methods, 14 proteins were selected for biological evaluation.

3.2 MATERIALS AND METHODS

3.2.1 Cell Culture and Virus Production.

Jurkat E6-1 cells (NIH AIDS Research and Reference Reagent Program, Germantown, MD) were cultured in RPMI 1640 media supplemented with 10% Fetalclone III (Hyclone, Logan, UT USA), 8 mM L-glutamine, 100 U/mL penicillin, and 100 U/ml streptomycin. 293T (AIDS Reagent Program, Germantown, MD) cells were cultured in DMEM media with the same supplementation and all cells were cultured in humidified incubators at 37°C and 5% CO₂. HIV-1 NLX virus stocks were produced by transient transfection of 293T cells using PolyJet as described by the manufacturer (SignaGen, Gaithesburg, VA). NLXLuc-VSVg virus was produced by transfection of the viral molecular clone (pNLX-luc) DNA and pMD2.G vesicular stomatitis virus glycoprotein G (VSVg) expression vector (Addgene Plasmid Repository, Cambridge, MA). Viral supernatants were collected over 48 hours, clarified by centrifugation at 4000 xg for 5 minutes and stored at -80°C.
3.2.2 SWATH-MS.

3.2.2.1 Infection and Sample Preparation.

Jurkat cells were infected by the spinoculation technique as described previously (O’Doherty, Swiggard, and Malim, 2000; Schweitzer et al., 2013) and 1x 10^7 Jurkat cells were processed in parallel with no virus as control cells. At 48 hours post infection, the cells were split into two groups and processed for MS. To generate the Jurkat reference spectral library, 1x 10^7 cells were lysed in 4% (w/v) SDS (Fisher Scientific, Waltham, MA), 0.1M dithiothreitol (Fisher Scientific, Waltham, MA), 0.1M Tris-HCl (Sigma-Aldrich, St. Louis, MO), and 100 units/mL Benzonase Nuclease, pH 7.6 (Life Technologies, Grand Island, NY). In addition, 1x10^7 cells were subjected to subcellular fractionation using the QProteome Cell Compartment Kit (Qiagen, Valencia, CA). Total protein for all samples was determined by Pierce 660 nm Protein Assay supplemented with 50 mM ionic detergent compatible reagent (Thermo Fisher, Waltham, MA).

To build the Jurkat reference library, 25 µg of 4% SDS lysate and 20 µg (each) of QProteome fraction 1-3 were used. Experimental DIA samples were prepared using QProteome fractions 1-3 only. All samples were processed using filter-assisted sample preparation (Wisniewski et al., 2009). Library samples were fractionated by isoelectric point using OFFGEL electrophoresis using pH 3-10 OFFGEL strips (Agilent, Santa Clara, CA), cleaned using an Oasis mixed cation exchange cartridge following manufacturers protocols (Waters, Milford, MA) and
2 µg of protein prepared for mass spectrometry using C18 Zip-Tips® (Millipore, Billerica, MA). Experimental DIA samples were digested with trypsin and cleaned using C18 spin columns following manufacturer protocols (Thermo Fisher, Rockford, IL)

3.2.2.2 Mass Spectrometry.

Mass spectrometry was performed as described previously (Haverland, Fox, and Ciborowski, 2014). Peptide samples were resuspended in 0.1% formic acid in HPLC grade water and were analyzed by reverse-phase high-pressure liquid chromatography electrospray ionization tandem mass spectrometry (RP-HPLC-ESI-MS/MS) using an Eksigent NanoLC-Ultra 1D plus (Exsigen, Dublin, CA) and nanoFlex cHiPLC system coupled to a 5600 Triple-TOF (AB Sciex; Concord, Canada) mass spectrometer with a nanospray needle voltage set to 2400V. Samples were loaded using a stepwise flow rate of 10 µL/min for 8.5 min and 2 µL/min for 1 min using 0.1% (v/v) formic acid in HPLC water (solvent A). Peptides were eluted from the analytical column using a 5-35% linear gradient of solvent B (95% (v/v) acetonitrile with 0.1% (v/v) formic acid over the course of 180 minutes with a flow rate of 0.3 µL/min.

The reference library was generated by running samples in the traditional DDA mode with a 250-ms survey scan and the top 50 ions selected for subsequent MS/MS. Ion selection criteria included a charge state from +2 to +5, intensity of greater than 100 counts/s, mass tolerance of 50 mDa, and were not
found on a dynamic exclusion list. Ions that were fragmented and analyzed by MS/MS were excluded from further analysis for 15 seconds.

Experimental samples were processed using cyclic DIA of mass spectra using 25-Da swaths as described in Liu et al (Liu et al., 2013). Briefly, a 50-ms survey scan (MS1) was performed and all precursors within a given 25-Da swath were fragmented and analyzed (MS2). Each cycle was composed of 34 25-Da swaths which covered 400Da to 1200 Da. Total Cycle time was 3.314 s using an accumulation time of 96 ms per 25-Da swath.

3.2.2.3 Protein Identification and z-Transformation

The DDA spectral data was searched using ProteinPilot software v4.2 (AB Sciex) using the Paragon algorithm with the following parameters: methyl methanethiosulfonate (MMTS) cysteine alkylation, digestion by trypsin (Thermo Scientific, Rockford, IL) and no special factors. Protein identification was through the UniProt Swiss-Prot database (November 2012 release) containing both human and HIV-1 proteins. In addition, a false discovery rate analysis was performed and peptides identified from proteins with an FDR of <1% were used as the Jurkat reference spectral library.

The DIA data was subject to spectral alignment and targeted data extraction by PeakView v1.2 (AB Sciex) software using the Jurkat reference spectral library. Identification of proteins used an extraction window of 20 min and the following parameters: 8 peptides, 5 transitions, exclude shared peptides, extracted ion
chromatogram (XIC) width set at 50 ppm, and peptide confidence of >99%.

Reversed sequences and laboratory contaminants were removed from the data set.

The 5 biological replicates were normalized by z-Score transformation (Cheadle et al., 2003). In summary, the natural log (ln) of the protein intensity was assigned a z-score, where $x$ is the experimental value, $\mu$ is the mean of all experimental values and $\sigma$ is the standard deviation of all experimental values. The $\Delta z$ was calculated for each protein per replicate ($\Delta z = z_{INF} - z_{CONT}$) and the average $\Delta z$ across all replicates was calculated. To determine statistically significant changes, a paired $z$-test was used where $\Delta z_{avg}$ is the average $\Delta z$ across all replicates, $D$ is the hypothesized mean (null hypothesis) of pairwise differences, $\sigma_d$ is the standard deviation of the pairwise differences per protein, and $\sqrt{n}$ is the square root of the sample size (number of biological donors). The equations can be found in Figure 3.1. The $p$-value for the computed $z$-test statistic was assigned using the standard normal distribution.

### 3.2.3 Infection assays

293T cells were seeded in triplicate wells of a 6-well plate to obtain 60% confluency the following day. For overexpression assays, each well was transfected with 1µg plasmid using PolyJet transfection (Signagen, Gaithersburg, MD) reagent. For siRNA knockdown assays, 2 pmol siRNA was transfected using PepMute (Signagen) reagent according to manufacturer’s directions. The
A. 
\[ z = \frac{x - \mu}{\sigma} \]

B. 
\[ z = \frac{\Delta z_{\text{avg}} - D}{\sigma_d / \sqrt{n}} \]

Figure 3.1 Statistical Analysis using z-Transformation. Equation used to determine z-score of the transformed raw intensity of SWATH-MS identified proteins is shown in (A). The z-test equation used to determine statistical difference in protein expression between HIV-1 infected cells and uninfected control cells is shown in (B).
next day, wells were inoculated with HIV-luc + VSVg and incubated for 48 h at 37°C. Cells were lysed with M-PER solution (Pierce Biotechnology, Rockford, IL) and clarified by centrifugation. One-glo luciferase reagent (Promega, Madison, WI) was used to determine Luciferase activity. Total protein concentration was determined by BCA protein assay (Thermo Scientific, Rockford, IL) and luciferase activity normalized to the total protein. Data shown represents at least three independent experiments. Western blot was used to confirm overexpression and protein knockdown. MTT assays were performed on siRNA transfected cells using the CellTiter 96 non-radioactive cell proliferation assay according to the manufacturer's specifications (Promega, Madison, WI).

3.2.4 Immunoblotting

Protein concentration in each sample was normalized with PBS and mixed 1:1 with 2x SDS-PAGE loading buffer. Samples were boiled for 10 min., separated by SDS-PAGE, and transferred to PVDF. Proteins were detected by Western blot using the following primary antibodies: anti-ERGIC (H-245), Sumo2/3 (N-18), GAPDH (6C5), ALYREF (11G5), SF2/ASF (3G268), COPINE1 (E18), C23/NCL (MS-3), SRPK1 (EE13), and TXNDC12/ERP19 (C-7). All were obtained from Santa Cruz Biotechnology (Santa Cruz, CA). Anti-GSTK1 (PA5-26866) and ARGHAP15 (PA5-31530) were obtained from ThermoScientific (ThermoScientific, Rockford, IL). Anti-MTPN (13508-1AP) and LSM5 (10700-1AP) were obtained from Proteintech (Proteintech, Chicago, IL). Human HIV-1
Immunoglobulin was obtained from the NIH AIDS Reagent Program (Germantown, MD). HRP conjugated anti-Flag (M2) primary antibody (Sigma-Aldrich, St Louis, MO), HRP Conjugated cMyc (SC-40) and Actin (C-4) primary and HRP conjugated anti-rabbit, anti-mouse, and anti-human IgG secondary (Santa Cruz, CA USA) antibodies were used for detection and visualized by West Pico (Thermo Scientific, Rockford, IL) chemiluminescent staining. Images were captured using radiographic film, scanned to computer, adjusted for brightness and contrast if necessary, and cropped for size.

3.2.5 Exogenous RT Activity

The procedure used was a modification of one previously described (Quan et al., 1996). Standard reaction mixtures contained 50 mM Tris (pH 7.9), 75 mM KCl, 2 mM DTT, 0.1875 mM ATP, 5 mM MgCl2, 0.05% Nonidet P-40, 2 µM dTTP, and 2 µCi [\(^{32}\text{P}\)]-α-TTP, and 10µL supernatant in a final volume of 40 µL. All reactions were incubated for 4 hours at 37°C. 10µL RT products were added to as individual spots on Whatman filter paper and allowed to dry completely. The paper was washed three times with 2x saline-sodium citrate (SSC) buffer and allowed to dry. After drying, \(^{32}\text{P}\) activity was visualized and quantitatively analyzed using a storage phosphor plate and scanned on a Typhoon 9410 scanner (GE Healthcare Bio-Sciences Corp. Piscataway, NJ). The amount of radioactivity in reverse transcribed DNA that was generated by virus produced in
the presence of excess factor was normalized to activity produced in untransfected cells.

3.3 RESULTS

3.3.1 Experimental Approach

This study focused on HIV-infected T-cells using Jurkat E6-1 cells. Prior studies in our laboratory found an exponential jump in exogenous RT activity indicating an increase in viral production starting at 48 hours as well as a reduction in CD4+ T-cells at the same time point (unpublished data). We hypothesized 48h as the ideal time point to evaluate the host cellular proteome changes just prior to maximum virus production. The approach used is summarized in Fig. 3.2. Jurkat T-cells were infected for 48 h and the cytosolic, membrane/organelle and nuclear fractions isolated. For the construction of the SWATH-MS reference library, we also isolated proteins by whole cell lysis with 4% SDS. A total of 5 biological replicates (infections, subcellular fractionations, and MS runs) were performed and used for library construction as well as the SWATH-MS experimental comparison between uninfected and HIV-infected cells. The collected spectra were normalized by z-transformation within each MS replicate allowing direct comparison between replicates and statistical analysis. Viral infection was confirmed for each replicate experiment by immunoblot of HIV-1 IgG (Fig 3.3A). Cellular markers were used to test the purity of the cell
Figure 3.2 Experimental approach used for SWATH-MS analysis of HIV-1 infected T-cells. Work-flow representative of one complete biological replicate.
Figure 3.3 Confirmation of infection and fractionation control. Western blot confirmation of HIV-1 infection in Jurkat cells used for SWATH-MS analysis is shown in (A). Control samples for cell fractionation are shown in (B). Representative blots from one biological replicate shown.
fractionations and include GAPDH (cytosol), ERGIC (membrane/organelle), and Sumo2/3 (nuclear) (Fig 3.3B).

A summary of the data obtained from the infection experiments is presented in Table 3.1. In total, 1167 proteins were found to significantly change expression/localization in infected Jurkat cells. Validating our approach, the HIV proteins Gag, Rev, and Env were found unique to infected cells and CD4+ was the most down-regulated factor in the membrane fraction.

3.3.2 Bioinformatic Analysis

A multifaceted approach was used to determine which factors would be identified for further investigation. Normalization of replicates was accomplished by $z$-Transformation of each replicate where the individual protein intensity was normalized to the mean and standard deviation of the entire replicate. Initially, we sought to determine if $z$-Score sorting was a practical method to discover targets. Data was sorted by $z$-Score and examined by independent fraction. While $z$-Score is not a direct quantitative measurement, we did find correlation with approximate quantity as HIV factors Gag, Rev, and Env were found to have high $z$-Scores. By fraction, the proteins with the highest or lowest $z$-Score and no HIV interaction were selected for further study.

In conjunction with the $z$-Score analysis, a comparative analysis was used to determine which factors were previously identified in prior proteomic studies within our laboratory (DeBoer et al., 2014; Schweitzer et al., 2013). A total of 5
<table>
<thead>
<tr>
<th>Fraction</th>
<th>Total Proteins</th>
<th>Increased Expression</th>
<th>z-Score Range</th>
<th>Decreased Expression</th>
<th>z-Score Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytosol</td>
<td>2055</td>
<td>154</td>
<td>0.01 to 2.24</td>
<td>313</td>
<td>-0.01 to -1.94</td>
</tr>
<tr>
<td>Membrane/Organelle</td>
<td>2103</td>
<td>326</td>
<td>0.01 to 2.87</td>
<td>388</td>
<td>-0.01 to -2.73</td>
</tr>
<tr>
<td>Nuclear</td>
<td>1895</td>
<td>132</td>
<td>0.01 to 1.63</td>
<td>74</td>
<td>-0.01 to -1.68</td>
</tr>
</tbody>
</table>
proteins had more than two hits in prior studies. Table 3.2 provides a summary of the proteins and their historical hits.

Additional analysis using the STRING-db software focused on protein-protein interactions (PPI) of factors within the data set. The individual fractions were sorted by z-Score and PPI determined for the entire group of proteins differentially regulated during HIV-1 infection. Initial examination included all proteins dysregulated in every fraction, however no clear patterns emerged (data not shown). Next, proteins were examined separately by fraction and whether they were up or down regulated. This data is shown in Figure 3.4a-f. When the proteins were separated in this manner, distinct clusters were discovered, with Q3 displaying the most prominent clusters.

Increasing the interaction confidence in STRING-db resulted in fewer interactions and allows for easier identification of central interacting proteins. This enhanced analysis resulted in identification of two additional factors for study and can be seen in Figure 3.5. Nucleolin (NCL) posed an interesting candidate due to the central nature of PPI as well as the predominance in previous studies as seen in Table 3.2. In addition, ATP-dependent RNA helicase DDX39A (DDX39A) provided another central protein with no previously published HIV interaction. However, DDX39A is found in the Staufen/HIV-1 gag complex of proteins essential for HIV-1 RNA incorporation into viral particles (Chatel-Chaix et al., 2004). A schematic of the known interactions can be found in Figure 3.6. The entire data set was searched for the Staufen/HIV-1 interactors, and many were found to be significantly altered following HIV-1 infection as seen in Table 3.3. In
Table 3.2: Proteins Identified in Previous Proteomic Studies

<table>
<thead>
<tr>
<th>Factor</th>
<th>Function</th>
<th>Proteomic Study¹</th>
<th>PICs</th>
<th>MA-interactor</th>
<th>C8166</th>
<th>NCI HIV</th>
<th>SWATH</th>
<th>Findings²</th>
</tr>
</thead>
<tbody>
<tr>
<td>CypB</td>
<td>Immunophilin; multiple f(x)s</td>
<td>Yes</td>
<td>Upregulated in infected nuclei; over-expression enhances HIV infection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DDX39A</td>
<td>DEAD-box helicase; RNA regulation</td>
<td></td>
<td></td>
<td>Over-expression enhances HIV infection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCL/C23</td>
<td>RNA regulation; transcription</td>
<td>Yes</td>
<td>Upregulated in infected nuclei</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NF45</td>
<td>RNA regulation; transcription</td>
<td>Yes</td>
<td>Over-expression enhances HIV gene expression; binds HIV RNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRPK1</td>
<td>SR protein kinase</td>
<td></td>
<td>Upregulated in infected cells</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ PIC and C8166 studies are published (Schweitzer et al. 2013, DeBoer et al. 2014), other studies unpublished.
² From our laboratory.
Figure 3.4a STRING-db analysis of proteins altered following HIV infection. Graphic analysis of cytosolic proteins with increased expression during HIV infection.
Figure 3.4b STRING-db analysis of proteins altered following HIV infection. Graphic analysis of cytosolic proteins with decreased expression during HIV infection.
Increased Membrane/Organelle Expression

Figure 3.4c STRING-db analysis of proteins altered following HIV infection. Graphic analysis of membrane/organelle proteins with increased expression during HIV infection.
Figure 3.4d STRING-db analysis of proteins altered following HIV infection. Graphic analysis of membrane/organelle proteins with decreased expression during HIV infection.
Figure 3.4e STRING-db analysis of proteins altered following HIV infection. Graphic analysis of nuclear proteins with increased expression during HIV infection.
Figure 3.4f STRING-db analysis of proteins altered following HIV infection. Graphic analysis of nuclear proteins with decreased expression during HIV infection.
Figure 3.5 Enhanced STRING-db PPI analysis of proteins with increased expression in the nucleus following HIV infection. Increased interaction confidence of PPI interactions shows NCL and DDX39A as central interacting proteins in the nucleus of HIV infected cells.
Figure 3.6 DDX39A Interactome. Known interactome of DDX39A in the Staufen Complex.
**Table 3.3: DDX39 Interactors Identified in Jurkat SWATH-MS Study.**

<table>
<thead>
<tr>
<th>Uniprot</th>
<th>ID</th>
<th>Fraction 1</th>
<th>Fraction 2</th>
<th>Fraction 3</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>O00571</td>
<td>DDX3X</td>
<td>-0.13</td>
<td>-0.09</td>
<td></td>
<td>ATP-dependent RNA helicase DDX3X</td>
</tr>
<tr>
<td>Q00148</td>
<td>DX39A</td>
<td>0.13</td>
<td></td>
<td></td>
<td>ATP-dependent RNA helicase DDX39A</td>
</tr>
<tr>
<td>Q13838</td>
<td>DX39B</td>
<td>0.17</td>
<td></td>
<td></td>
<td>Spliceosome RNA helicase DDX39B; also UAP56 or BAT-1</td>
</tr>
<tr>
<td>P35537</td>
<td>FUS</td>
<td>0.89</td>
<td>0.45</td>
<td></td>
<td>RNA-binding protein, also TLS</td>
</tr>
<tr>
<td>P82979</td>
<td>SARNP</td>
<td>0.56</td>
<td></td>
<td></td>
<td>CIP29</td>
</tr>
<tr>
<td>Q86V81</td>
<td>THOC4</td>
<td>-0.10</td>
<td>0.34</td>
<td></td>
<td>Aly</td>
</tr>
</tbody>
</table>
addition to DDX39A, THOC4 (ALYREF) and SARNP were selected for additional investigation. Chapter IV is dedicated to the investigation of CypB as a HIV-1 factor. Table 3.4 provides a summary of the proteins selected for additional biological investigation.

### 3.3.3 Biological Evaluation of Candidate Factors.

Proteomic screens and bioinformatic analyses provide important clues as to potential factors that may influence HIV-1 infection. However, these studies do not provide a complete story. Evaluation of factors in a biological setting provides insight into whether the alterations in cellular proteins directly impact HIV-1 infection or are simply byproducts of cellular functions. In the following studies, factors identified in 3.3.2 are investigated further for their impact on HIV-1 infection. It is important to note, in many of the studies the overall fold-change difference may not appear to be relevant, even with statistically significant changes. In a seminal study to establish CD4 as the primary HIV-1 receptor, Maddon et al. only found a ~7.5 fold increase in infectivity following the addition of CD4 to a CD4− permissive cell line (Maddon et al., 1986). Therefore, relatively small changes in the proteome have the potential to significantly alter overall viral fitness and should not be ruled out completely. Altering the overall protein levels through exogenous expression vectors or siRNA mediated protein knockdown provides insight into how these proteins alter HIV-1 infection, if at all.
### Table 3.4: Candidate Factors Selected for Further Analysis

<table>
<thead>
<tr>
<th>Candidate Factor</th>
<th>Comparative$^1$</th>
<th>z-Score$^2$</th>
<th>PPI$^3$</th>
<th>NCBI HIV Database?</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALYREF</td>
<td>▼Q2, ▲Q3</td>
<td></td>
<td>■</td>
<td>Yes</td>
</tr>
<tr>
<td>ARHGAP15</td>
<td>▼Q1</td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>ASF/SF2</td>
<td>■</td>
<td>▼Q1, Q2, Q3</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>CNOT10</td>
<td>▼Q2</td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Copine1</td>
<td>▼Q3</td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>DDX39A</td>
<td>■</td>
<td>▲Q3</td>
<td>■</td>
<td>No</td>
</tr>
<tr>
<td>GSTK1</td>
<td>▲Q1</td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>LSM5</td>
<td>▲Q2, Q3</td>
<td></td>
<td>■</td>
<td>No</td>
</tr>
<tr>
<td>MTPN</td>
<td>▲Q2</td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>NCL</td>
<td>■</td>
<td>▲Q3</td>
<td>■</td>
<td>Yes</td>
</tr>
<tr>
<td>SARNP</td>
<td>▲Q3</td>
<td></td>
<td>■</td>
<td>No</td>
</tr>
<tr>
<td>SRPK1</td>
<td>■</td>
<td>▼Q1, Q3</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>TM261</td>
<td>▼Q2</td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>TXNDC12</td>
<td>▲Q2</td>
<td></td>
<td></td>
<td>No</td>
</tr>
</tbody>
</table>

$^1$ Proteins identified in prior proteomic studies within this laboratory

$^2$ Z-Score alteration by Protein 4 Fraction

$^3$ Proteins identified using STRING-db Protein-Protein Interactions
3.3.3.1 *Does overexpression of candidate factors alter HIV-1 infection?*

Gain-of-function studies provide a mechanism to evaluate if the over-expression of candidate factors effects HIV infection. To do this we pre-transfected 293T cells with expression vectors 24 h prior to infection with a VSVg-pseudotyped HIV luciferase reporter virus (HIV-Luc). At 48 hpi, luciferase activity was used as a measurement of HIV infection. Figure 3.7A summarizes the results for all 14 factors evaluated. Western blot was used to confirm expression and representative blots can be found in 3.7B. Overexpression of ARHGAP, COPINE1, DDX39A, MTPN, and TM261 proteins enhanced HIV infection, and one factor reduced infection, ALYREF.

3.3.3.2 *Does siRNA Knockdown of candidate factors alter HIV-1 infection?*

To examine the inhibitory effect a protein may have on HIV infection, siRNA was used to deplete cells of the candidate factors prior to HIV infection. The protocol is similar to the overexpression assays except 293T cells were transfected with siRNA 24 h prior to infection, and infected with the same VSVg-pseudotyped HIV-Luc reporter virus. At 48 hpi, luciferase activity was measured and compared to the activity of cells transfected with control siRNA. Figure 3.8A summarizes the results for the factors evaluated. Data is not available for all
Figure 3.7 Effect of candidate factor overexpression on HIV infection. HIVLuc activity following overexpression of candidate factors shown in (A). Activity is normalized to the control pcDNA3.1 vector. Statistically significant changes in HIV infection are shaded. Representative blots confirming overexpression are shown in (B). “F-” indicates Flag-tagged construct used and anti-Flag antibody used. “cM-” indicates cMyc-tag used for construct identification. All others are antibodies directed at the candidate factor.
Figure 3.8 Effect of candidate factor siRNA knockdown on HIV infection. HIVLuc activity following siRNA of candidate factors shown in (A). Activity is normalized to the control siRNA. Statistically significant changes in HIV infection are shaded. Representative blots confirming candidate factor knockdown are shown in (B) with control blots for Actin beneath each factor.
factors screened due to either unsuccessful knockdown of the protein or not having the ability to confirm knockdown. Confirmation of knockdown was visualized by Western blot and representative blots are shown in Figure 3.8B. Statistically significant increases in HIV-1 infection following knockdown of ASF/SF2, COPINE1, GSTK1, NCL, SRPK1, and TXNDC12 was observed.

ALYREF was the lone factor to reduce HIV-1 infection following alteration of protein expression. In addition, COPINE1 and SRPK1 knockdown resulted in increased infection correlating with lower endogenous expression following infection as seen by SWATH-MS. In contrast, the depletion of NCL and TXNDC12 resulted in increased infection following knockdown while the SWATH-MS data indicated that the levels increased following infection.

3.3.3.3 Virus Production after overexpression of candidate factor.

The HIV-Luc reporter assay is an effective tool to measure HIV-1 infection. However, the assay is not without limitations. The HIV-Luc virus is created on the HIV backbone with Env replaced by the Luciferase gene. This results in a replication deficient virus limiting infection cycle analysis to events up to mRNA translation. To investigate the later stages in HIV-1 infection, we chose to look at the quantity of virus produced by cells overexpressing candidate proteins. 293T cells were co-transfected with the NLX-Luc molecular clone and protein expression vector. At 48 hpi, supernatant was collected and RT activity measured to quantify the amount of virus released relative to cells transfected
with a pcDNA3.1 empty vector control. A summary of results is found in Figure 3.9. ARHGAP15, NCL, and SRPK1 significantly increased HIV-1 virus production while ASF/SF2, DD39A, and SARNP decreased virus production.

The overexpression of three factors, SRPK1, NCL, and ARHGAP15 enhanced the production of virus with ARHGAP15 being the only one with no published HIV interaction. The overexpression of ASF/SF2, SARNP, and DDX39A all significantly decreased virus production. Of note, these proteins are all RNA processing factors.

3.4 DISCUSSION

Proteomics as a discipline continues to evolve and mature (Hause et al., 2011; Silberring and Ciborowski, 2010; Vaudel, Sickmann, and Martens, 2012). As such, there is major emphasis on the development of high throughput technologies and methodologies to unlock the remaining mysteries of virus biology and provide potential therapeutic targets. However, the application of proteomics in addressing questions of biological importance has shifted from global, unbiased profiling studies in which cataloging and comparing proteins in relevant samples was the major goal, to much more focused (Lai, Nice, and Schilling, 2013) and “surgical” proteomic studies in which we measure more subtle yet functionally crucial changes in the biological system being experimentally manipulated (Chen, Na, and Peng, 2012; Vuckovic et al., 2013). The approach used in this study sought to profile the global changes to the T-cell
Figure 3.9 Impact of candidate factor overexpression on late replication events and HIV virus production. 
Endogenous RT activity of supernatant following co-transfection of pNLX and candidate factors is shown. Activity is normalized to the control pcDNA3.1 vector. Statistically significant changes in HIV infection are shaded.
proteome while providing a template for selection of protein targets for focused biological study.

Determining the significance of protein changes identified in proteomic studies presents a unique challenge. Many proteomic studies rely primarily on the quantitative increase or decrease of overall protein amount, while subtle protein changes in the cellular proteome may have a significant impact viral replication. In this study, we sought to identify proteins altered in HIV-1 infection through multiple normalized replicates and analyzed by statistical change, not quantitative protein changes.

Correlating biological relevance with proteomic findings is a key goal of this study. Alterations to the cellular proteome can be a result of the cellular response to viral infection, or it may be a specific action caused by the virus. In this study we used a novel proteomic strategy coupled with bioinformatic analyses encompassing evaluation of the SWATH-MS data based upon $z$-Score analysis for statistical change, prior proteomic studies, and known protein-protein interactions to determine a condensed data set consisting of 14 potential factors out of over 1000 proteins altered in response to HIV-1 infection. These proteins were then subjected to a battery of biologic studies to determine the effect on HIV-1 infection. Table 3.5 provides an overall summary of the 14 candidate factors including SWATH-MS protein expression studies and all infection studies.

Proteins found in the NCBI database for prior HIV interactions were not excluded from the infection studies. However, the results obtained in our
<table>
<thead>
<tr>
<th>Candidate Factor</th>
<th>z-Score</th>
<th>OE</th>
<th>KD</th>
<th>VP</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALYREF</td>
<td>↓Q2, ↑Q3</td>
<td>↓</td>
<td>↓</td>
<td></td>
</tr>
<tr>
<td>ARHGAP15</td>
<td>↓Q1</td>
<td>↑↑</td>
<td>NP</td>
<td>↑↑</td>
</tr>
<tr>
<td>ASF/SF2</td>
<td>↓Q1, Q2, Q3</td>
<td>↑</td>
<td></td>
<td>↓↓↓↓</td>
</tr>
<tr>
<td>CNOT10</td>
<td>↓Q2</td>
<td></td>
<td>NP</td>
<td></td>
</tr>
<tr>
<td>Copine1</td>
<td>↓Q3</td>
<td>↑</td>
<td>↑↑↑↑</td>
<td></td>
</tr>
<tr>
<td>DDX39A</td>
<td>↑Q3</td>
<td>↑↑</td>
<td>NP</td>
<td>↓↓↓↓</td>
</tr>
<tr>
<td>GSTK1</td>
<td>↑Q1</td>
<td>↑↑↑↑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSM5</td>
<td>↑Q2, Q3</td>
<td></td>
<td>NP</td>
<td></td>
</tr>
<tr>
<td>MTPN</td>
<td>↑Q2</td>
<td>↑↑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCL</td>
<td>↑Q3</td>
<td>↑↑↑↑</td>
<td>↑↑</td>
<td></td>
</tr>
<tr>
<td>SARNP</td>
<td>↑Q3</td>
<td>NP</td>
<td></td>
<td>↓↓↓↓</td>
</tr>
<tr>
<td>SRPK1</td>
<td>↓Q1, Q3</td>
<td>↑↑↑↑</td>
<td>↑↑</td>
<td></td>
</tr>
<tr>
<td>TM261</td>
<td>↓Q2</td>
<td>↑</td>
<td>NP</td>
<td></td>
</tr>
<tr>
<td>TXNDC12</td>
<td>↑Q2</td>
<td>↑↑</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OE - Overexpression, KD - siRNA Knockdown, VP - Virus Production, NP - Not Performed
experiments mirror those of the previously published studies for ALYREF (Taniguchi, Mabuchi, and Ohno, 2014), ASF/SF2 (Jacquenet et al., 2005), GSTK1 (Barrero et al., 2013), and NCL (Lee and Park, 2009). While our results do not provide new information, our similar findings serve to validate the experimental approach.

DDX39A and SARNP are both nucleic acid binding proteins with SARNP enhancing the RNA unwinding activity of DDX39A; both may be an integral part of transcription, splicing and RNA export (Sugiura, Sakurai, and Nagano, 2007). DDX39A shows a slight enhancement of infectivity following overexpression, however, both show a significant decrease in viral production following overexpression. These results suggest this complex plays a key role in the later stages of HIV-1 infection.

The overexpression of MTPN and TM261 result in a slight increase in HIV-1 expression with no change following siRNA knockdown. While both of these offer potential as HIV-1 enhancement factors, neither are required for infection. Likewise, knockdown of TXNDC12 produces a slight increase in infectivity, with no decrease following overexpression suggesting TXNDC12 is a restriction factor, but not a block to replication.

The remaining factors, ARHGAP15, COPINE1, and SRPK1 are the most likely factors to contribute to HIV-1 replication. ARHGAP15, Rho GTPase-activating protein 15, is an activator for the Rho-type GTPases which regulate a variety of regulates a variety of important cell signaling and growth control pathways (Van Aelst and D'Souza-Schorey, 1997). While no direct interaction
with HIV-1 has been published for ARHGAP15, Rho-GTPases have been found to be important for CXCR4 conformation and actin polymerization during budding (Audoly, Popoff, and Gluschankof, 2005; Zoughlami et al., 2012).

SWATH-MS analysis showed a decrease of ARHGAP15 in the cytoplasm of infected cells. Overexpression of ARHGAP15 leads to an increase in HIV-1 infectivity. Knockdown of protein expression could not be confirmed, so we are unable to determine if ARHGAP15 is required HIV-1 replication. An increased amount of virus was produced when ARHGAP is overexpressed. Combined, these results suggest that ARHGAP is an important, but not required factor for HIV-1 replication and that the decrease in cellular ARHGAP15 may be a cellular response to HIV-1 infection.

COPINE1 is a calcium-dependent membrane-binding protein which may function in membrane trafficking and regulate molecular events at the interface of the cell membrane and cytoplasm (Tomsig, Sohma, and Creutz, 2004). Copine 1 has no published direct HIV-1 interaction, but knockdown by siRNA increases TNFα-stimulated NF-κB transcription which can result in increased HIV-1 transcription (Coiras et al., 2007; Ramsey et al., 2008).

SWATH-MS showed a decrease in COPINE1 in the nuclear fraction of HIV-1 infected cells. Overexpression shows a slight increase in infection while knockdown shows a significant increase in infection. Overexpression had no impact on overall virus production. These studies coupled with the known molecular interaction with the well-known HIV-1 factor NF-κB (Fiume et al., 2012; Roux et al., 2000) reduce the likelihood that COPINE1 is an independent HIV-1
factor. Rather, the enhancement of HIV-1 infection in the siRNA knockdown experiments was possibly NF-κB mediated.

SRPK1 is a Serine/arginine-rich protein-specific kinase which specifically phosphorylates its substrates at serine residues located in regions rich in arginine-serine dipeptides and is involved in the phosphorylation of SR splicing factors and the regulation of splicing. No HIV-1 protein interactions have been published, however, the main substrate of isoform 2 of SRPK1 is ASF/SF2, a known HIV-1 factor (Pongoski, Asai, and Cochrane, 2002).

SWATH-MS analysis shows reduced SRPK1 in the cytoplasm and nucleus of HIV-1 infected cells. Overexpression of SRPK1 produced no increase in HIV-1 infection, but did significantly increase virus production. Knockdown of SRPK1 also produced an increase in HIV-1 infection. These results are contrary to expected results as the overexpression of SRPK1 should results in an increased amount of its primary substrate, ASF/SF2, and decreased virus production as seen earlier in this study. This suggests SRPK1 is acting on a different substrate warranting further investigation.

The study as performed did not succeed in finding a factor essential for HIV-1 replication. However, information relevant to the HIV lifecycle and potential interacting partners can be gleaned from the experiments.
3.5 FUTURE STUDIES

One critical element missing from most proteomic studies is the ability to map changes in the localization of proteins, which may identify critical pathways of activation or repression within the cell. Analysis of proteins altered in 2 or more fractions will yield additional candidates based upon the movement of proteins within the cell following HIV infection.

The data analysis used in this study to determine protein levels was a purely statistical measure rather than a quantitative amount of protein. SWATH-MS analysis provides the possibility to re-evaluate the data set determining the quantitative change in protein levels. The quantitative change in protein levels may provide additional factors worth evaluating.

Finally, Additional bioinformatic analyses using more advanced pathway analysis (PANTHER, KEGG, etc) may also provide additional insight into the subcellular systems altered during HIV infection. This analysis may provide additional factors that may not be evident by the other evaluation methods.
Chapter IV: Cyclophilin B Enhances HIV-1 Infection

4.1 INTRODUCTION

Human Immunodeficiency Virus type 1 (HIV-1) is the causative agent of Acquired Immune Deficiency Syndrome (AIDS) and has remained a global health problem for over thirty years. HIV-1 is a retrovirus which primarily targets the CD4+ T-cells of the immune system. The genome consists of 9 open reading frames encoding only 15 viral proteins (Frankel and Young, 1998). This limited genome necessitates the utilization of host proteins and pathways for effective replication while evading the innate and acquired immune responses. Thus far, approximately 4000 human proteins have been found to affect HIV-1 infection in some capacity (http://www.ncbi.nlm.nih.gov/genome/viruses/retroviruses/hiv-1/interactions/; (Ako-Adjei et al., 2014)). The major challenge is identifying which of these proteins can be targeted for cell-based viral inhibitors. This necessitates both characterizing protein function as well as deciphering how a factor interacts with HIV and impacts viral replication.

Cyclophilin B (CypB) was identified in a number of proteomic studies performed by our laboratory, including: the interactome of Matrix and Integrase, a study of partially purified preintegration complexes, and a study of the nuclei of HIV-infected T-cells ((DeBoer et al., 2014; Schweitzer et al., 2013; Schweitzer et al., 2012); unpublished data). Additionally, it was found to be significantly upregulated in the nuclei of HIV-infected monocyte-derived macrophages (Haverland, Fox, and Ciborowski, 2014). CypB, also known as peptidyl prolyl isomerase B (PPIB), is a member of the cyclophilin family of immunophilins.
CypA, the initial member of this family, was identified by its ability to bind the immunosuppressant cyclosporine A (CsA). The CypA/CsA, and to a lesser extent, CypB/CsA heterodimers bind and inhibit calcineurin preventing dephosphorylation and nuclear translocation of the transcription factor NFAT. NFAT is critical for production of immunostimulating cytokines, including IL2.

Cyps also contain a peptidyl prolyl isomerase (PPIase) function which lowers the rotational energy barrier in prolyl imide bonds. The core domain of CypB (aa 48-143) has 80% homology with CypA and contains its isomerase domain and CsA binding region (Figure 4.1). However, the functions of CypA and CypB are distinct. CypB contains a N-terminal hydrophobic leader region with an endoplasmic reticulum (ER) signal sequence that are cleaved after expression, as well as a C-terminal ER retention signal (Arber, Krause, and Caroni, 1992). Whereas CypA is localized primarily in the cytosol, CypB localizes to the ER/golgi and plasma membrane (Price et al., 1994). As much as 50% of CypB is secreted from the cell (Caroni et al., 1991). It was proposed early that the secreted form of CypB may be a chaperon for proteins bound for export (Price et al., 1994). Interestingly, CsA, which reduces the isomerization activity of CypB (Spik et al., 1991), also causes an increase in the secretion of CypB (Price et al., 1994). CypB has been found to act as an intracellular chaperon for two proteins-CAML, an activator of the NF-AT pathway in T-cells (von Bülow and Bram, 1997), and interferon regulatory factor 3 (IRF-3; (Obata et al., 2005)). In the latter capacity it was demonstrated that knock-down of CypB reduced virus induced phosphorylation of IRF-3 and its subsequent induction of INF-β.
<table>
<thead>
<tr>
<th></th>
<th>ER Leader and NLS</th>
<th>CsA Binding and PPI active site</th>
<th>ER Ret Seq</th>
</tr>
</thead>
<tbody>
<tr>
<td>CypB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δ5/10/20</td>
<td>5</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Δ33</td>
<td>NLS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δ40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δ206</td>
<td>ER Leader and NLS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPlm</td>
<td>ER Leader and NLS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CypA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4.1 Domain structure of CypB and the deletion mutants used in this study.** The approximate location of previously described domains are shown and the core domain conserved between CypB and CypA is highlighted in gray.
Most of the immunophilin studies with HIV have focused on CypA. Both CypA and CypB bind HIV-1 Gag, but CypB also interacts with SIV Gag (Franke, Yuan, and Luban, 1994; Luban et al., 1993). Moreover CypB binds Gag with a stronger affinity due to its hydrophobic N-terminal ER signal sequence. This also renders Gag-CypB interaction less sensitive to disruption by CsA compared to Gag-CypA complexes. Initial immunoblot studies found that CypA, but not CypB, was incorporated into virions (Franke, Yuan, and Luban, 1994; Thali et al., 1994). A later study detected CypB in virions by mass spectrometry (Chertova et al., 2006). The interaction of CypA with viral cores is necessary for capsid stabilization, proper uncoating, and protection of viral cores from cellular restriction factors (Sayah et al., 2004), and efficient HIV replication is inhibited by CsA \textit{in vitro}. CypA interacts with Gag via proline residues and mutagenesis of proline 222 (Pro222Ala) ablates CypA binding which inhibits replication. Interestingly, this mutation does not affect the interaction of CypB and Gag, indicating that CypB interacts in an independent manner with Gag that CypA (Franke, Yuan, and Luban, 1994).

CypB is essential for successful infection of multiple viruses, including Hepatitis C Virus (HCV; (Watashi et al., 2005)), Japanese encephalitis virus (JEV; (Kambara et al., 2011)), and human papillomavirus type 16 (HPV16; (Bienkowska-Haba, Patel, and Sapp, 2009)). CypB has been shown to be required for the interferon-β response to viral infection through modulating IRF3 phosphorylation, dimerization and translocation to the nucleus (Obata et al., 2005). In addition to IRF3, CypB has been found to be essential for the nuclear
retrotransport of Prolactin (PRL) following internalization of PRL/PRLr (Rycyzyn et al., 2000). This intranuclear PRL/CypB complex acts as a transcriptional inducer enhancing Stat5-mediated gene expression (Rycyzyn and Clevenger, 2002). Finally, CD147 stimulation by CypB induces intracellular Ca2+ flux stimulating T cell proliferation and produces a chemotactic response in primary human neutrophils (Yurchenko et al., 2001).

Given the repeated identification of CypB in HIV proteomic studies, we hypothesized that it played an important role in HIV replication. We discovered that ectopic expression of CypB enhances HIV infection at the step of nuclear import of the viral DNA. This enhancement is distinct from CypA and specific for HIV. The N-terminus of CypB is necessary for the effect. Deletion of the leader sequence of CypB alters localization of the CypB and hinders its secretion from the cell. However, media transfer experiments suggested that secreted CypB does not play a role in increasing infection. Finally, inhibition of the Stat5 pathway does not interfere with the enhancement. Combined, these results suggest CypB modulates HIV nuclear import.

4.2 MATERIALS AND METHODS

4.2.1 Plasmids

The C-terminial FLAG-tagged CypB and CypA ORFs were obtained from Origene Technologies. The CypB ORF was moved by PCR into the pTarget
eukaryotic expression vector, the backbone used for the deletion mutant ORFs which were inserted by PCR mutagenesis (Promega Corp., Madison WI USA). The PPI mutant was constructed in pTarget by PCR-overlap mutagenesis and contained the amino acid substitutions R96 to A96 and F101 to A101. The primers used for mutagenesis were designed using NCBI Reference Sequence # NM000942.4.

4.2.2 Cell culture and virus preparation

293T and HeLa cells (NIH AIDS Research and Reference Reagent Program, Germantown, MD) were maintained in Dulbecco’s modified eagle medium (DMEM) supplemented with 10% fetal clone 3 (HyClone, Logan, UT), 8 mM L-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin. Viruses were produced by transient transfection of 293T cells using PolyJet reagent following the manufacturer’s instructions (SignaGen Laboratories, Gaithersburg, MD). Briefly, cells were seeded into 10 cm dishes at 70% confluency one day prior to transfection. NLXLuc-VSVg virus was produced by transfecting 3.5 µg of viral molecular clone (pNLX-luc) DNA and 1.5 µg of pMD2.G vesicular stomatitis virus glycoprotein G (VSVg) expression vector (Addgene Plasmid Repository, Cambridge, MA). Media was collected at 24 and 48 h post transduction and clarified by centrifugation at 4000 xg for 5 min. MLV-luc viruses were produced following the same protocol using 2.5 µg viral molecular clone (pFB-luc), 1.5 µg
packaging vector (pCG-gag-pol), and 1 μg pMD2.G. Viruses were titered on 293T cells and a TCID<sub>50</sub> used for all infection experiments.

4.2.3 Infection assays

293T cells were seeded in triplicate wells of a 6-well plate. The following day each well was transfected with 500 ng plasmid using PolyJet transfection reagent. The next day, wells were inoculated with 100 μl of HIV-luc or MLV-luc and incubated for 48 h at 37°C. Cells were lysed in 300 μl M-PER solution (Pierce Biotechnology, Rockford, IL) and clarified by centrifugation. Luciferase activity was determined using One-glo luciferase reagent (Promega, Madison, WI) and measured using a Spectramax L (Molecular Devices, Sunnyvale, CA). Protein concentrations were determined by BCA protein assay (Thermo Scientific, Rockford, IL) and luciferase activity was normalized to the total protein. Data shown represents at least three independent experiments. Subcellular compartments (cytosol, membrane/organelle, nuclear, and cytoskeleton/insoluble) were isolated using the Qproteome Cell Compartment kit as described by the manufacturer (Qiagen, Valencia, CA). MTT assays were performed on transfected cells using the CellTiter 96 non-radioactive cell proliferation assay according to the manufacturer’s specifications (Promega, Madison, WI).
4.2.4 Immunoblotting

Samples were normalized by protein concentration with PBS, mixed 1:1 with 2x SDS-PAGE loading buffer, boiled for 10 min., separated by SDS-PAGE, and transferred to PVDF. Proteins were detected by Western blot using the following primary antibodies: anti-CypA (C-14), anti-CypB (k2E2), anti-ERGIC (H-245), Sumo2/3 (N-18), and anti-GAPDH (6C5). All were obtained from Santa Cruz Biotechnology (Santa Cruz, CA). HRP conjugated anti-Flag (M2) primary antibody (Sigma-Aldrich, St Louis, MO) and HRP conjugated anti-rabbit, anti-mouse IgG secondary (Santa Cruz, CA USA) antibodies were used for detection and visualized by West Pico (Thermo Scientific, Rockford, IL) chemiluminescent staining. Images were captured using radiographic film, scanned to computer, adjusted for brightness and contrast if necessary, and cropped for size.

4.2.5 Indirect immunofluorescence

HeLa cells were plated on sterile coverslips in a six-well plate at 50% confluence. The following day, cells were transfected with CypB expression plasmids and incubated overnight. Cells were washed with PBS and fixed for 5 min by adding 2mL of 4% paraformaldehyde to each well. Cells were permeabilized using 2 mL of 0.2% Triton X-100 in PBS for 5 min. FITC conjugated anti-Flag (M2) antibody (Sigma-Aldrich, St Louis, MO) was diluted 1:1000 in PBS+1% FBS and 2 mL added to cells and incubated at 37° C for 60
min. Cells were washed 3x with PBS+1% FBS and coverslips mounted using ProLong® Gold antifade reagent with DAPI (Life Technologies, Eugene, OR). Slides were dried overnight and imaged using a Nikon Eclipse Ci (Nikon USA, Melville, NY) microscope and images captured using Infinity Analyze software (Lumemera, Ottawa, ON).

4.2.6 Quantification of DNA for LRT, 2-LTR and Integrated viral DNA

293T cells were seeded into 6 well plates to achieve 50% confluence the following day. Cells were transfected with control vector, CypB, or Δ40 and incubated overnight. NLX+VSVg stocks were treated with 2 U/ml Turbo DNase (Ambion, Austin TX) for 1 h at 37°C prior to inoculation. Extra-chromosomal DNA was isolated at indicated times using the modified HIRT protocol (Arad, 1998; Belshan et al., 2009). DNA was amplified using Bullseye EvaGreen qPCR Master Mix (Midsci, Valley Park MO) on CFX Connect real-time PCR detection system (Bio-Rad). Late reverse transcription (LRT) viral DNA was quantified using gag-specific primers NL919 (5’-TTCGCAGTTAATCCTGGACTT-3’) and NL1054 (5’-GCACACAATAGAGGACTGCTATTGTA). LRT was normalized to detection of mitochondrial DNA using primers MitoFor (5’-ACCCACTCCCTCTTAGCCAATATT-3’) and MitoRev (5’-GTAGGGCTAGGCCCACCG-3’). 2-LTR circles were quantified using primers NL500 (5’-AACTAGGGAACCCACTGCTTAAG-3’) and NL9126 (5’-TCCACAGATCAAGGATATCTTGTC-3’) and normalized to the level of LRT viral
DNA. Chromosomal DNA was isolated using the EZNA Tissue DNA Kit (Omega Biotek, Frederick CO) and integration was assessed by nested Alu-PCR, using primers cNL658 (5’-TTTCAGGTTCCCTGTTCGGGCGCCAC-3’), Alu1 (TCCCAGCTACTGGGGAGGCTGAGG-3’), and Alu2 (GCCTCCCAAAAGTGCTGGGATTACAG-3’) for the initial PCR and primers NL493 (5’-TCTGGCTAACTAGGGAACCCAC-3’) and cNL616 (5’-CTGACTAAAAGGGTCTGAGG-3’) for the second PCR (Brussel and Sonigo, 2003). Integrated DNA was normalized to detection of the β-globin intron using primers BGlobinFor (5’-GCAAGAAAGTGCTCGGTGC-3’) and BGlobinRev (5’-CTACTCAGTGTGGCAAAG-3’). Heat inactivated virus (30 min at 65°C) was used as a negative control for each experiment.

**4.2.7 STAT5 inhibition assay**

A STAT5 inhibitor (N’-((4-Oxo-4H-chromen-3-yl)methylene)nicotinohydrazide) was obtained from Santa Cruz Biotechnology and resuspended in DMSO to a concentration of 10mg/mL. 293T cells were seeded into 6 well plates to achieve 50-60% confluence the following day. The STAT5 inhibitor was diluted in complete DMEM in a range of concentrations from 1 µM to 200 µM. Cells were pre-treated with inhibitor for 4 hours prior to transduction with NLXLuc VSVg virus. At 24 hours, virus was removed and new media containing inhibitor added. At 48 hours post-transduction, cells were lysed with MPER and luciferase activity determined as described above. Cell cycle analysis was used to verify STAT5
inhibition with $1 \times 10^6$ 293T cells treated 24 hours with inhibitor, harvested by centrifugation, and resuspended in Vindelov’s reagent (10mM Tris (pH 7.6), 10 μg/ml RNase A, 75 μM propidium iodide, and 0.1% Igepal CA-630). Cells were analyzed at the Creighton University flow cytometry core using a FACS Aria flow cytometer (BD Biosciences, San Jose, CA) and Flowjo software (Treestar Inc., Ashland, OR). The MTT assay was used to determine cell viability in the presence of inhibitor.

### 4.3 RESULTS

#### 4.3.1 CypB enhances HIV-1 infection.

CypB has been identified in a number of proteomic studies of HIV infected cells and purified virions (Chertova et al., 2006; DeBoer et al., 2014; Haverland, Fox, and Ciborowski, 2014; Schweitzer et al., 2013). In a recent study we observed it to be increased in the nuclei of HIV infected T-cells (DeBoer et al., 2014). To study the role of CypB we initially attempted depletion studies using RNAi. Despite several attempts to knock-down CypB expression using small interfering or short hairpin RNAs we were unsuccessful suggesting that CypB is an essential factor for cell survival. Therefore we moved to gain-of-function studies to see if the over-expression of CypB affected HIV infection. To do this we pre-transfected 293T cells with increasing amounts of FLAG-tagged CypB expression vector 24 h prior to infection with a VSVg-pseudotyped HIV luciferase
reporter virus (HIV-Luc). As shown in Figure 4.2, increased expression of CypB enhanced the infection of HIV in a dose dependent manner. To assess if this effect was non-specific toward retroviruses, we performed parallel experiments with a VSVg-pseudotyped murine leukemia virus luciferase reporter virus (MLV-Luc). Increased expression of CypB did not alter MLV infection, indicating that CypB did not have a general effect on retrovirus transduction.

4.3.2 The mechanism of enhancement of infection by CypB is distinct from CypA.

Many studies have demonstrated that CypA is required for efficient HIV infection. The core domain of CypB is ~80% homologous to CypA, but CypB contains distinct N- and C-termini (Figure 4.1). Moreover, both CypA and CypB bind HIV-1 Gag (Luban et al., 1993). The interaction of CypA and Gag occurs via the PPI domain of CypA binding PRO 222 of Gag. We also investigated if CypA could also potentiate HIV infection in our assays. Despite robust over-expression of CypA, no substantial change in HIV infection was observed (Figure 4.2). This result suggested that CypB affects HIV infection in a mechanism distinct from CypA.

To further examine this issue, we tested whether cyclosporine A (CsA) would abrogate the ability of CypB to enhance HIV infection. CsA binds both CypA and CypB inhibiting the peptidyl prolyl isomerase (PPI) activity of both CypA and CypB, and blocks the interaction of CypA and calcineurin (Liu et al., 1991). CsA
Figure 4.2 Over-expression of CypB enhances HIV-1 infection. Cells were transfected with increasing amounts of CypB or CypA as indicated that was balanced with empty vector control plasmid. The following day, cells were infected with either a VSVg-pseudotyped HIV-Luc or MLV-Luc reporter virus (CypB only). At 24 hpi, cells were harvested and infection quantified by Luc assay which was normalized to protein concentration. Results are normalized to control cells. Cell lysates were blotted for CypB or CypA expression and GAPDH was used as loading control (bottom panels). Data shown is representative of three independent experiments with triplicate infections. Error bars in this and all other figures represent standard error of the mean.
treatment disrupts the interaction of CypA and Gag resulting in an inhibition of HIV infection (Franke and Luban, 1996). To test if CsA would mitigate the enhancement by CypB, the over-expression experiments were repeated with cells treated with CsA (Figure 4.3). As previously reported (Franke and Luban, 1996), CsA partially inhibited HIV infection in mock transfected cells. However, CsA had no effect on the ability of CypB to enhance HIV infection. These results showed that the core/PPI domain of CypB was not required for its ability to enhance HIV infection and confirmed that the effect of CypB is independent of CypA.

4.3.3 The N-terminus of CypB is necessary to enhance HIV-1 infection.

The previous experiments suggested that a region outside the core domain of CypB was responsible for promoting HIV infection. To map the domains of CypB that were required for the enhancement effect, we constructed a series of deletion mutants, as well as a PPI-inactive mutant (Figure 4.1). Similar to the parental construct, all the mutants were created with a C-terminal FLAG epitope to allow for detection. The ability of these mutants to affect HIV infection was assessed as in previous experiments. Deletion of the C-terminus of CypB (Δ206) or disruption of the PPI motif did not affect the ability of CypB to augment HIV infection (Figure 4.4A). In contrast, deletions in the N-terminus of CypB blocked its ability to potentiate infection. Removal of the first 10 amino acids alone
Figure 4.3 CsA does not abrogate the ability of CypB to enhance infection. Cells were transfected with CypB or empty vector and pretreated with CsA prior to infection with HIV-Luc VSVg. Cells were harvested and protein normalized Luc activity measured to quantify infection. Results are presented as normalized to control DMSO treated cells. Cell lysates were blotted for CypB and CypA expression. GAPDH was used as loading control (bottom panels).
Figure 4.4 The N-terminus of CypB is required for the enhancement of HIV infection. (A) Infection experiments were performed as described in legend of FIG 1 using indicated deletion mutants and a PPI inactive mutant. Results shown were normalized to cells transfected with empty vector control and are representative of a minimum of three replicate experiments. Mutants that failed to significantly enhance infection compared to WT CypB are indicated with an asterisks (p<0.05 by two-tailed t-test). At the end of each experiment the expression of each mutant was verified by immunoblot of cell lysates (bottom panels). (B) Overall cell viability following expression of CypB and each mutant was monitored by MTT assay.
completely reduced the ability of CypB to increase HIV infection. The expression of two larger N-terminal deletion mutants (Δ33 or Δ40) caused a slight reduction of infection compared to mock-transfected cells. Control immunoblots showed that all the mutants expressed (lower panels), although mutation of the PPI domain or deletion of the C-terminus at amino acid 206 appeared to enhance the expression of a larger isoform of the protein, which is putatively a glycosylated form (Spik et al., 1991). Interestingly, this isoform was not present in cells transfected with the Δ10, Δ33, or Δ40 mutants, suggesting the N-terminus is the site of modification. Finally, MTT assays confirmed that over-expression of CypB and the derivative mutants did not impact overall health of the cells (Figure 4.4B).

4.3.4 Deletion of the N-terminus alters CypB localization.

The N-terminus of CypB contains an ER signal sequence, a putative nuclear localization signal (NLS), and is also necessary for the secretion of CypB (Rycyzyn et al., 2000). To assess the effect the deletions had on the intracellular localization of CypB we performed immunofluorescence assays on cells transfected with the different CypB deletion mutants (Figure 4.5). Wild-type CypB showed strong perinuclear localization typically with one distinct foci of fluorescence. The Δ5 mutant exhibited less perinuclear staining than full-length CypB, but maintained distinct foci of fluorescence near the nuclei of cells. In contrast, the Δ10 mutant showed strong perinuclear staining, but less distinct
Figure 4.5 The intracellular localization of the CypB deletion mutants. HeLa were transfected with CypB or indicated mutant expression vectors. After 24 h the cells were fixed, permeabilized and protein localization detected by immunofluorescence using a FITC conjugated anti-FLAG antibody. Nuclei were stained with DAPI. Panels are representative images of the localization of each protein as observed in two independent experiments.
foci. Both the Δ33 and Δ40 mutants lost both the perinuclear localization and distinct foci and displayed diffuse cellular localization. The Δ206 C-terminal mutant show staining similar to wild-type CypB and the PPI mutant maintained generally perinuclear staining but with more cytoplasmic puncti compared to wild-type CypB. Overall, all the deletions except Δ5 appear to distinctly perturb the localization of CypB correlating with a loss of enhancement. Combined, these results show that the mutants that lost the ability to enhance HIV infection (Δ10, Δ33, and Δ40) also failed to form perinuclear foci.

We also examined the localization of the CypB mutants by cell fractionation and immunoblotting (Figure 4.6). 293T cells expressing the tagged CypB mutants were fractionated into cytoplasmic, membrane/organelle, nuclear and insoluble/cytoskeletal fractions. Fraction consistency was monitored by detection of GAPDH (cytosol), ERGIC (membrane/organelle), and Sumo2 (nuclei) (Figure 6, bottom blots). Wild-type CypB was observed in the cytoplasmic, membrane/organelle, and nuclear fractions consistent with previous reports (Price et al., 1994) as well as our immunofluorescence data. Several isoforms of CypB were detected in the membrane fractions. The mature form of CypB in cells lacks the N-terminal leader sequence with removal of the first 32 amino acids (Spik et al., 1991). The larger band likely represented either the full length “pro” form of CypB or the mature form with a modification (consistent with this idea, WT, Δ5, Δ10, and Δ33 all migrate at a similar size (Figure 4). Interestingly, only the mature form of CypB was detected in the cytoplasm of cells and conversely, only larger isoforms of WT CypB were detected in nuclei. Both the Δ5 and Δ10
Figure 4.6 The subcellular distribution of the CypB deletion mutants. 293T cells transfected with CypB or indicated mutant were fractionated into cytosolic (Cyt), membrane/organelle (Mem), nuclear (Nu), and cytoskeletal/insoluble (In) subcellular fractions. Fractions were separated by SDS-PAGE and immunoblotted with anti-FLAG antibodies. Bottom three panels show control immunoblots for cell fractionation. Results shown are representative of two independent experiments.
mutants fractionated similar to WT CypB with the notable exception that the mature isoforms were detected in nuclei instead of the larger forms. Consistent with the IFA data as well as the removal of the ER leader sequences, both Δ33 and Δ40 were detected in the cytoplasmic fractions only. The Δ206 mutant was present in the cytoplasmic, organelle, and nuclear fractions but consistent with the deletion of the C-terminal ER retention signal (Arber, Krause, and Caroni, 1992), it exhibited a higher proportion of cytoplasmic to membrane localization compared to WT CypB. Notably, three isoforms of Δ206 were observed, confirming that any modification, if present, was not altered by deletion of the C-terminus of CypB. Finally, the PPIm showed a localization distribution similar to Δ206, with a higher proportion in the cytoplasmic fractions of cells. Overall, these results show that the N-terminal truncation of CypB caused aberrant localization and expression of the CypB isoforms. Thus, the CypB mutants that lost the ability to enhance HIV infection showed changes in their localization compared to wild-type CypB.

4.3.5 Secreted CypB does not enhance HIV infection.

CypB is secreted from cells through the ER pathway and found in milk and blood in the body (Mariller et al., 1996). Since all the enhancement-null mutants displayed changes to their intracellular localization, we next tested if the secretion of CypB was necessary for the enhancement of HIV infection. First we
investigated the secretion of the mutants by immunoblot (Figure 4.7A). Only the PPI mutant was detected in supernatants at a level equivalent to wild-type CypB. Importantly, the loss of secretion of the Δ206 mutant suggested that the secretion of CypB was not critical for its ability to enhance infection as it potentiated infection as well as WT CypB. To further test if extracellular CypB was involved in the enhancement of infection, we performed media transfer experiments in which cells were infected with virus pre-mixed 1:1 with media from cells over-expressing CypB, CypA, or mock transfected cells. Both CypB and CypA could be detected in the media from transfected cells (Figure 4.7B, top blot). Trans-addition of the media from CypB over-expressing cells to target cells only slightly enhanced HIV infection, and no effect was seen with the CypA media (Figure 4.7B, graph). Notably, we did not observe any uptake of either CypB or CypA into the recipient cells (Figure 4.7B, bottom blot, 2º lanes). We concluded from these results that secreted CypB does not play a significant role in the enhancement of HIV infection.

4.3.6 CypB increases nuclear import of HIV DNA.

To determine how CypB enhanced HIV infection we analyzed the early steps of HIV infection by quantifying viral DNA species by real-time PCR. In addition to over-expressing CypB, we also included the following controls: mock transfected cells, cells infected with heat-inactivated virus, as well as cells transfected with the Δ40 mutant that did not potentiate infection. Late RT product accumulation
Figure 4.7 Secreted CypB does not enhance HIV infection in trans. (A) Secretion of CypB mutants. Cells were transfected with CypB or indicated mutants. After 24 h supernatants and cell lysates were collected and analyzed for protein expression by immunoblot. (B) Media transfer experiments. Donor 293T cells were transfected with indicated expression constructs. At 24 h post-transfection the media was changed to eliminate excess transfection complexes. After an additional 24 h supernatants and cell lysates were collected. NLX-Luc virus stock was mixed 1:1 with donor supernatant and used to infected untransfected recipient cells. At 48 hpi recipient cells were harvested and infection levels measured by Luc assay. The secretion of CypB and CypA was monitored by immunoblot of donor supernatants (top blot) and the ectopic expression of both CypB and CypA (1°) as well as uptake into recipient cells (2°) was monitored by immunoblotting with anti-FLAG antibodies (bottom blots). Graphed data represents the compilation of two independent experiments and blots are representative images.
was analyzed at 24 hpi using \textit{gag}-specific primers and normalized to the level of mitochondrial DNA as an internal RT control (Figure 4.8A). The over-expression of CypB did not alter the total amount of late products compared to mock-transfected cells or the cells over-expressing the Δ40 mutant. This indicated that CypB did not enhance steps prior to, and through reverse transcription. The import of viral DNA into the nuclei of cells was assessed by quantifying 2-LTR circle accumulation. The levels of 2-LTR circles were normalized to the levels of late RT products to account for differences in the magnitude of infection. As shown in Figure 4.8B, the level of 2-LTR circles was higher in cells over-expressing CypB compared to the mock-transfected cells. The amount of 2-LTR circles was also slightly decreased in the Δ40 transfected cells. To confirm the increase in nuclear import in CypB expressing cells, we also quantified the levels of viral DNA integrated in the genome of cells using an established \textit{alu}-based nested PCR assay (Brussel and Sonigo, 2003). Consistent with the nuclear import/2-LTR data, the amount of integrated viral DNA was higher in cells over-expressing CypB (Figure 4.8C). Moreover, cells expressing the Δ40 mutant had reduced levels of integrated viral DNA. Combined, the correlation of the 2-LTR and integrated viral DNA data with the infection data suggests that the over-expression of CypB enhances infection by increasing the nuclear import of the viral DNA.
Figure 4.8 CypB increases nuclear import of HIV. In all experiments 293T cells were transfected with either CypB or the Δ40 mutant and infected the following day with NLX+VSVG. (A) For Late RT experiments extrachromosomal DNA was isolated at 24 hpi. Late RT products quantified by qPCR using gag-specific primers. Levels were normalized for DNA isolation efficiency by quantification of mitochondrial DNA (mtDNA) levels by qPCR. (B) 2-LTR circles were quantified from samples harvested at 48hpi. To account for infection efficiency, the levels of 2-LTR circles were normalized to the level of late RT DNA in each sample (which was normalized to mtDNA level). (C) Measurement of integrated viral DNA. Chromosomal DNA was isolated at 48 hpi and integrated viral DNA measured by an alu-based nested qPCR as outlined in Methods. Mock transfected and cells infected with heat inactivated (HI) virus were used as controls in all experiments. Data presented in A-C is normalized to the mock transfected group and combined from three independent experiments. Asterisks indicate p<0.05 by two-tailed t-test.
4.3.7 CypB potentiation does not involve the STAT5 pathway.

Previous reports suggest that CypB is critical for the nuclear localization of IRF-3 and prolactin (Obata et al., 2005; Rycyzyn et al., 2000), suggesting it may interact with a nuclear import pathway. CypB interacts with prolactin and mediates its nuclear transport (Rycyzyn et al., 2000). The nuclear CypB-prolactin complex interacts and activates the Stat5 transcription factor (Rycyzyn and Clevenger, 2002). Ectopic expression of Stat5 has been shown to increase HIV-1 production in primary T-cells (Selliah et al., 2006). To investigate if Stat5 activation was involved in the CypB potentiation of HIV infection, we tested whether overexpression of CypB would enhance HIV infection in the presence of a Stat5 inhibitor. As shown in Figure 4.9A, treatment of cells with the Stat5 inhibitor alone increased HIV infection, but at all treatment levels CypB still potentiated infection. Stat5 did not appear to alter cell metabolism as measured by MTT assay (Figure 4.9B), but did cause a decrease in DNA synthesis as evidenced by few cells in S and G2 phases of the cell cycle (Figure 4.9C). These results coupled with the fact that 1) deletion of the 40 N-terminal amino acids of CypB retains prolactin binding (Rycyzyn et al., 2000), but blocks nuclear import of prolactin; and 2) mutation of the PPI domain of CypB, which still enhances HIV infection, does not activate Stat5 (Rycyzyn and Clevenger, 2002), lead us to conclude that the enhancement effect of CypB is independent of Stat5 activation.
Figure 4.9 CypB enhancement does not require Stat5 activity. (A) Infection assays. 293T cells were transfected with either pcDNA3.1 empty vector or CypB. The following day media was changed and the cells pre-treated with indicated concentrations of STAT5 inhibitor for 4 h prior to transduction with NLX-Luc. At 48hpi cells were lysed and assayed for Luc activity, which was normalized to protein total protein concentration. Results are presented relative to infected control cells treated with DMSO. (B) MTT assay of cells treated with Stat5 inhibitor. (C) Cell cycle analysis of cells treated with 5, 50, or 200 μM Stat5 inhibitor.
4.4 DISCUSSION

The results from these studies show that the over-expression of CypB enhances the infection of HIV-1, suggesting that CypB mediates an important step of infection. In the studies first describing CD4 as the receptor for HIV, the addition of CD4 to CD4− fibroblasts potentiated infection ~7.5 fold (Maddon et al., 1986). In our studies, addition of excess CypB enhanced infection ~3.5-4 fold suggesting a significant role for CypB in HIV infection. Moreover, unlike the CD4− fibroblasts, the cells in our study all express endogenous CypB. Unfortunately, we were unable to knock-down CypB without cytopathic effects. CypB had been identified in virions and in numerous proteomic studies but has not been extensively studied likely due to its homology to CypA and an assumption that it has a related function. Indeed, both CypA and CypB are inhibited by CsA and both bind Gag; albeit, CypB binds at higher affinity due to its hydrophobic N-terminus and mutation of PRO222 does not affect CypB binding (Franke, Yuan, and Luban, 1994). The enhancement was clearly distinct from CypA as overexpression of CypA did not increase infection, and CsA treatment, which blocks CypA-Gag binding, did not alter the effect.

Consistent with the effect being distinct from CypA, deletion analyses demonstrated that the N-terminus of CypB is required for the effect. In addition, deletion of this region of CypB resulted in a loss of ER localization and secretion. However, it appears that the secretion of CypB was not involved in potentiating HIV infection as media transfer experiments showed that CypB did not enhance
infection when provided in trans nor did it enter cells or bind virions (data not shown). Moreover, CsA increases the secretion of CypB (Price et al., 1994), but it did not affect (either positively or negatively) the ability of CypB to boost infection. Combined, these results suggest that disruption of the intracellular localization of CypB ablates its ability to potentiate infection. The defective mutants Δ33 and Δ40 showed a loss of detectable membrane/organelle and nuclear localization in our immunoblot experiments of fractionated cells. The Δ10 N-terminal mutant lost the ability to potentiate infection and showed less concentrated perinuclear staining in cells, although the immunoblot analyses showed it still could enter the nuclear compartment. Combined, this suggests that the perinuclear localization of CypB is necessary for increasing HIV infection.

CypB has been shown to contribute to the nuclear localization of IRF-3 and prolactin (Obata et al., 2005; Ryczyn et al., 2000), and contains a “classical” lysine rich NLS in its N-terminus. The deletion of this region ablates prolactin nuclear import, but the mechanism of CypB nuclear import and its interactions with the nuclear transport machinery have yet to be defined. Arguing against a direct role of CypB facilitating nuclear import of the PIC, deletions N-terminal to the NLS ablated the ability of CypB to increase infection. Further study of the interactions of CypB with the HIV RTC and PIC as well as the nuclear transport machinery is needed to clarify the role of CypB as a nuclear chaperon. We theorize two possible mechanisms of CypB enhancement. First, that CypB may interact with the viral capsid, RTC, or PIC and facilitate interactions with cellular pathways to increase the nuclear import of the vDNA. Alternatively, the over-
expression of CypB may activate one or more cellular pathways that boost import of the vDNA. *In vitro*, CypB did not affect reverse transcription or integration, and we did not observe the association of CypB with viral RNA or DNA. Moreover, CsA, which blocks CypA and B binding to Gag did not inhibit enhancement by CypB. Therefore, although we do not rule out the possibility that CypB binds the viral import complex, based on our findings thus far our future studies will focus on the CypB interactome in HIV infected cells to map factors that support HIV nuclear import.

4.5 FUTURE STUDIES

The studies described above clearly demonstrate that CypB plays an important role in the HIV lifecycle, nuclear import of vDNA in particular. The biologic function as well as the protein-protein interactome of CypB has not been fully described. Our studies suggest proteins acting in conjunction with CypB potentiate HIV-1 infection rather than CypB alone. A potential first step to map the CypB interacting partners would be to immunoprecipitate CypB during HIV infection and use MS to determine the interactome. Fluorescent microscopy could be used to confirm interaction through co-localization studies. In addition, time course fractionation studies as described in ChII, and re-evaluation of the SWATH-MS data could be used to study the intracellular translocation of interacting proteins. To further characterize the interaction, additional studies using the alternate CypB constructs could be used to determine interaction sites.
Initial studies are currently underway to determine the intracellular CypB interactome potentially identifying co-factors contributing to the CypB mediated enhancement of HIV-1 infection.
Chapter V: Project Summary and References

5.1 Project Summary

HIV-1 is an obligate intracellular parasite, requiring host cell proteins and systems to effectively replicate. Identification and characterization of these proteins is essential for understanding the HIV-1 replication process. In addition, host factors present a novel approach to antiretroviral drug targeting. A variety of methods have been employed to identify host cellular factors essential for HIV-1 replication including siRNA, cDNA, whole cell proteomic, fractionated cells, and virus particles. In fact, the NCBI has a dedicated database of proteins that have been identified with HIV-1.

The investigation of the nuclear proteome described in Chapter II provided unique insight into protein changes at the subcellular level. Traditional whole cell studies lack the sensitivity to identify subtle changes in the subcellular proteome due to the sheer volume of data obtained. Our approach of only looking at the nuclear proteome allowed for a more focused interrogation of the data leading to the high confidence identification of 163 proteins altered following HIV-1 infection. Of these proteins, 13 were validated using Western blot expression in an independent cell line. While this study may not be overwhelming with the total number of proteins identified, the overall quality of the identifications is much greater than previous studies due to the experimental design and data analysis techniques used. Biological evaluation of a select group of these proteins uncovered an often overlooked factor, Cyclophilin B, which enhances HIV-1 infection.
The SWATH-MS studies described in Chapter III are another unique approach to proteomic analysis of HIV-1 infected cells. The use of subcellular fractionation and new SWATH-MS methodology combine to provide information unavailable in previous studies. Subcellular fractionation provides the ability to map protein changes within the cellular environment providing details that may be missed in traditional whole cell proteomics which only looks at overall change in protein expression. In addition, the increases sensitivity of SWATH-MS provides a means to find subtle changes that may have been missed in previous studies. Using basic bioinformatic analysis tools we identified 14 proteins with a high likelihood of impacting HIV-1 infection. Biological evaluation of these factors included overexpression, knockdown, and overall virus production experiments. Three proteins, ALYREF, COPINE1, and SRPK1 were found to have a significant impact on HIV-1 replication and warrant further investigation. Additional bioinformatic analysis is ongoing and will lead to additional factors for investigation.

Determining the biological relevance of protein interactions identified in proteomic studies presents a significant challenge. Many proteins in the NCBI HIV-1 interaction database have been shown to simply interact with HIV-1 in some manner, but not necessarily influence HIV-1 replication. In addition, many of these interactions were found in cultured cell lines and have not been independently validated through alternate cell lines or primary cells. Our studies sought to address these shortcomings and identify factors with biological relevance to HIV-1 infection by investigating as many factors as possible through
validation of changes in protein expression in independent cell lines, as well as
gain- and loss-of-function studies. Our laboratory identified CypB as a potential
HIV-1 factor. The functional characterization of CypB during HIV-1 infection was
performed to determine how CypB enhanced HIV-1 infection. Initial studies
sought to establish CypB enhancement as specific for HIV-1 and importantly,
independent of the known role of CypA during HIV-1 infection. Subsequent
functional domain analysis revealed that the mechanism for enhancement lies
within the N-terminus of the protein. HIV replication cycle analysis revealed that
CypB enhances nuclear import of the PIC through an unknown mechanism.
Additional studies into this mechanism may unlock a key mystery in the HIV-1
replication cycle, what cellular pathways modulate nuclear import of the PIC.

Taken together, these studies further the basic knowledge of HIV-1 infection
and sheds light on the complex protein changes within the cell during infection.
Understanding the host cell proteome is the first step in developing novel cell
based treatments to combat the ever-changing HIV-1 virus. Modulation of the
cellular proteome may provide additional treatment options in the future.
5.2 References


