Regulation of HIV-1 Preintegration Transcription by Tat and Cellular Transcription Factors NF-kappa B and Sp1

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Beatrix Wilhelmina Meltzer
Master of Science
Johns Hopkins University, 2006

Director: Yuntao Wu, Professor
Department of Biosciences

Spring Semester 2014
George Mason University
Fairfax, VA
This work is licensed under a [Creative Commons Attribution-Nodevirs 3.0 Unported License](https://creativecommons.org/licenses/by-nd/3.0/).
DEDICATION

This dissertation is dedicated to my wonderful and loving husband, Lowell, and my two beautiful and precious children, Patries and Romy. Thank you for all your love, support, patience, and the many sacrifices you have made during my educational journey. Without your encouragement I doubt I would have ever started, nor finished this degree.

Lowell, what I have become as a result of my education in this country is for both of us. Thank you for your endless amount of support throughout the years. I am so grateful to Peter and Donna who brought us together. You mean everything to me. I appreciate and respect you more than I can say. Most of all, I love you.

This dissertation is also dedicated to my beloved father-in-law, Mel Meltzer. Mel was a captain in the US NAVY, and he and his team developed the P-3 Orion aircraft. Mel passed right before I started my Ph.D. Mel, an MIT graduate, always believed in my capabilities before he convinced me of it. He and I started my educational journey in the US. Mel drove me to my first test at the Northern Virginia Community College for an English language evaluation to enter in the “English as a Second Language” class. This first step led to my current educational status. I know that he is very proud of me. Thank you for everything Mel.
ACKNOWLEDGEMENTS

My journey through graduate school would not have been possible if it wasn’t for a core group of people who helped me get through it. I appreciate all the direct and indirect input that led to the completion of this degree. I want to start with thanking my advisor and mentor, Dr. Yuntao Wu for giving me this incredible opportunity to pursue an advanced degree under his guidance. He took me in his lab even though I had no prior research experience. With his patience and understanding, I managed to overcome this hurdle. I also would like to thank the rest of my committee members, all of whom took the time to be a part of my journey. They were always helpful and available for my questions. Furthermore, I would like to thank Dong Yang Yu who was always there with an open ear, and a willing hand to help me with my experiments. Additionally, I would like to acknowledge my fellow lab members. I would also like to thank Dr. Mudit Tyagi who helped me with the ChIP assay, and the Kashanchi lab for providing HDAC inhibitors. Thanks to our graduate academic coordinator, Diane St. Germain for helping through the process. Special thanks goes out to my dearest friend Mary Cortina, who edited my dissertation for the non-science content. For a non-science professional to take the time to read a science paper is a true test of friendship. Finally, I would like to thank Sarah Pierson, a fellow lab member who took the time to review my dissertation for the science content. Thank you all.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
<tr>
<td>List of Abbreviations or Symbols</td>
<td>xiii</td>
</tr>
<tr>
<td>Abstract</td>
<td>xv</td>
</tr>
<tr>
<td>introduction</td>
<td>17</td>
</tr>
<tr>
<td>Background of HIV-1 and AIDS</td>
<td>17</td>
</tr>
<tr>
<td>Long terminal repeats include HIV-1 basic promoter</td>
<td>20</td>
</tr>
<tr>
<td>Transcription factor Sp1</td>
<td>22</td>
</tr>
<tr>
<td>Transcription factor NF-κB</td>
<td>23</td>
</tr>
<tr>
<td>Other transcription factors</td>
<td>25</td>
</tr>
<tr>
<td>Chromatin structure</td>
<td>28</td>
</tr>
<tr>
<td>HIV-1 transcription initiation</td>
<td>31</td>
</tr>
<tr>
<td>HIV-1 transcription elongation</td>
<td>32</td>
</tr>
<tr>
<td>Viral protein Tat</td>
<td>34</td>
</tr>
<tr>
<td>Integrated versus nonintegrated HIV-1 genome</td>
<td>39</td>
</tr>
<tr>
<td>Viral latency</td>
<td>43</td>
</tr>
<tr>
<td>Neurological complications</td>
<td>47</td>
</tr>
<tr>
<td>Chapter One</td>
<td>50</td>
</tr>
<tr>
<td>Introduction</td>
<td>50</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>53</td>
</tr>
<tr>
<td>Plasmids</td>
<td>53</td>
</tr>
<tr>
<td>Cloning pCMVΔR8.2(D116N)Δtat</td>
<td>56</td>
</tr>
<tr>
<td>Cloning pCMVΔR8.2Δtat</td>
<td>62</td>
</tr>
<tr>
<td>Cloning pNL-RRE-SA-Luc</td>
<td>63</td>
</tr>
<tr>
<td>Cloning pNLΔTAR-RRE-SA-Luc</td>
<td>63</td>
</tr>
<tr>
<td>Viruses</td>
<td>67</td>
</tr>
</tbody>
</table>
Production of viruses in producer cell line .............................................................. 72
In-house p24 antigen ELISA .................................................................................. 73
Tissue Culture Media ............................................................................................... 75
Cells ........................................................................................................................ 76
Temperature gradient PCR ..................................................................................... 76
Restriction endonuclease digestion ........................................................................ 77
Ligation ..................................................................................................................... 77
Transformation ........................................................................................................ 78
Screening of clones .................................................................................................. 79
Infection of cells ....................................................................................................... 80
Luciferase assay ....................................................................................................... 80
Time-course for total, full-length HIV DNA real-time PCR assay ......................... 81
Sucrose cushion virus purification ......................................................................... 82
Reverse transcriptase assay for evaluation of the packaged viral genome ............. 83
Real-time PCR for total, full-length HIV DNA ....................................................... 84
Real-time PCR for positive control RNA ................................................................. 85
Normalization of data .............................................................................................. 86
Results ....................................................................................................................... 87
Reporter plasmid showed specific activity .............................................................. 87
Viral nonintegrated transcription is affected by a Tat negative virus ....................... 87
Providing Tat in trans in producer cells restored the defect .................................... 91
Dissecting the role of virion-associated Tat in nonintegrated transcription ............ 93
Disproportional DNA synthesis during the reverse transcription process ............ 97
Defect in transactivation of integrated viral template by Tat mutants ..................... 100
Unequal viral DNA synthesis during the reverse transcription process in an infection with a functional integrase ................................................................. 102
The lack of Tat in producer cells leads to a reduction of encapsidation of genomic RNA in virions ................................................................. 104
A TAR mutant virus infection also displayed a decrease in reporter gene expression ................................................................. 112
Low DNA synthesis during reverse transcription in a TAR mutant viral infection 115
Uneven genomic RNA packaged in a mutant TAR virus compared to a wild type TAR virus ................................................................. 117
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1-1</td>
<td>68</td>
</tr>
<tr>
<td>Table 1-2</td>
<td>70</td>
</tr>
<tr>
<td>Table 1-3</td>
<td>71</td>
</tr>
<tr>
<td>Table 3-1</td>
<td>188</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>55</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>57</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>60</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>65</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>69</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>88</td>
</tr>
<tr>
<td>Figure 1.7</td>
<td>90</td>
</tr>
<tr>
<td>Figure 1.8</td>
<td>92</td>
</tr>
<tr>
<td>Figure 1.9</td>
<td>94</td>
</tr>
<tr>
<td>Figure 1.10</td>
<td>96</td>
</tr>
<tr>
<td>Figure 1.11</td>
<td>99</td>
</tr>
<tr>
<td>Figure 1.12</td>
<td>101</td>
</tr>
<tr>
<td>Figure 1.13</td>
<td>103</td>
</tr>
<tr>
<td>Figure 1.14</td>
<td>106</td>
</tr>
<tr>
<td>Figure 1.15</td>
<td>107</td>
</tr>
<tr>
<td>Figure 1.16</td>
<td>109</td>
</tr>
<tr>
<td>Figure 1.17</td>
<td>110</td>
</tr>
<tr>
<td>Figure 1.18</td>
<td>113</td>
</tr>
<tr>
<td>Figure 1.19</td>
<td>114</td>
</tr>
<tr>
<td>Figure 1.20</td>
<td>116</td>
</tr>
<tr>
<td>Figure 1.21</td>
<td>118</td>
</tr>
<tr>
<td>Figure 1.22</td>
<td>120</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>136</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>137</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>139</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>148</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>151</td>
</tr>
<tr>
<td>Figure 2.6</td>
<td>154</td>
</tr>
<tr>
<td>Figure 2.7</td>
<td>157</td>
</tr>
<tr>
<td>Figure 2.8</td>
<td>159</td>
</tr>
<tr>
<td>Figure 2.9</td>
<td>161</td>
</tr>
<tr>
<td>Figure 2.10</td>
<td>163</td>
</tr>
<tr>
<td>Figure 2.11</td>
<td>164</td>
</tr>
<tr>
<td>Figure 2.12</td>
<td>167</td>
</tr>
<tr>
<td>Figure 2.13</td>
<td>169</td>
</tr>
</tbody>
</table>
Figure 2.14 ...................................................................................................................... 171
Figure 3.1 ........................................................................................................................ 187
Figure 3.2 ........................................................................................................................ 205
Figure 3.3 ........................................................................................................................ 207
Figure 3.4 ........................................................................................................................ 209
Figure 3.5 ........................................................................................................................ 210
Figure 3.6 ........................................................................................................................ 211
Figure 3.7 ........................................................................................................................ 212
Figure 3.8 ........................................................................................................................ 214
Figure 3.9 ........................................................................................................................ 215
Figure 3.10 ..................................................................................................................... 218
Figure 3.11 ..................................................................................................................... 220
Figure 3.12 ..................................................................................................................... 222
Figure 3.13 ..................................................................................................................... 225
Figure 3.14 ..................................................................................................................... 227
Figure 3.15 ..................................................................................................................... 229
Figure 3.16 ..................................................................................................................... 230
Figure 3.17 ..................................................................................................................... 232
Figure 3.18 ..................................................................................................................... 234
Figure 3.19 ..................................................................................................................... 235
Figure 3.20 ..................................................................................................................... 237
Figure 3.21 ..................................................................................................................... 240
## LIST OF ABBREVIATIONS OR SYMBOLS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apicidin</td>
<td>Ap</td>
</tr>
<tr>
<td>Azidothymidine</td>
<td>AZT</td>
</tr>
<tr>
<td>Carboxyl terminal domain</td>
<td>CTD</td>
</tr>
<tr>
<td>Chromatin immunoprecipitation</td>
<td>ChIP</td>
</tr>
<tr>
<td>Cyclin-dependent kinase 9</td>
<td>CDK9</td>
</tr>
<tr>
<td>Dulbecco’s modified Eagle’s medium</td>
<td>DMEM</td>
</tr>
<tr>
<td>Enzyme-linked immunosorbent assay</td>
<td>ELISA</td>
</tr>
<tr>
<td>Etravirine</td>
<td>ETV</td>
</tr>
<tr>
<td>Fetal bovine serum</td>
<td>FBS</td>
</tr>
<tr>
<td>Functional integrase</td>
<td>IN+</td>
</tr>
<tr>
<td>Green Fluorescent Protein</td>
<td>GFP</td>
</tr>
<tr>
<td>HDAC inhibitor(s)</td>
<td>HDACi(s)</td>
</tr>
<tr>
<td>Highly active anti-retroviral therapy</td>
<td>HAART</td>
</tr>
<tr>
<td>Histone</td>
<td>H</td>
</tr>
<tr>
<td>Histone acetyltransferase</td>
<td>HAT</td>
</tr>
<tr>
<td>Histone deacetylase</td>
<td>HDAC</td>
</tr>
<tr>
<td>HIV-associated dementia</td>
<td>HAD</td>
</tr>
<tr>
<td>HIV-associated neurocognitive disorders</td>
<td>HAND</td>
</tr>
<tr>
<td>Integrase</td>
<td>IN</td>
</tr>
<tr>
<td>Luciferase</td>
<td>Luc</td>
</tr>
<tr>
<td>Macrophage colony-stimulating factor</td>
<td>M-CSF</td>
</tr>
<tr>
<td>Monocyte-derived macrophages</td>
<td>MDMs</td>
</tr>
<tr>
<td>Nonfunctional integrase</td>
<td>IN-</td>
</tr>
<tr>
<td>Nuclear factor of activated T cells</td>
<td>NFAT</td>
</tr>
<tr>
<td>Nucleosome</td>
<td>Nuc</td>
</tr>
<tr>
<td>P300/CBP-associated factor</td>
<td>P/CAF</td>
</tr>
<tr>
<td>Peripheral blood mononuclear cells</td>
<td>PBMCs</td>
</tr>
<tr>
<td>Phosphate buffered saline</td>
<td>PBS</td>
</tr>
<tr>
<td>Positive transcription elongation factor b</td>
<td>P-TEFb</td>
</tr>
<tr>
<td>Pre-integration complex</td>
<td>PIC</td>
</tr>
<tr>
<td>Propidium iodide</td>
<td>PI</td>
</tr>
<tr>
<td>RNA polymerase II</td>
<td>RNA Pol. II</td>
</tr>
<tr>
<td>Roswell Park Memorial Institute medium</td>
<td>RPMI</td>
</tr>
<tr>
<td>Sodium Butyrate</td>
<td>NaBut</td>
</tr>
<tr>
<td>Specificity protein 1</td>
<td>Sp1</td>
</tr>
<tr>
<td>TATA binding protein</td>
<td>TBP</td>
</tr>
<tr>
<td>Term</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>TBP-associated factor</td>
<td>TAF</td>
</tr>
<tr>
<td>Tetracycline response element</td>
<td>TRE</td>
</tr>
<tr>
<td>Tetracycline-controlled transcriptional activator</td>
<td>rtTA</td>
</tr>
<tr>
<td>Transactivation response element</td>
<td>TAR</td>
</tr>
<tr>
<td>Transactivator of transcription</td>
<td>Tat</td>
</tr>
<tr>
<td>Valproic acid</td>
<td>VPA</td>
</tr>
</tbody>
</table>
For a productive life cycle, HIV-1 needs to integrate its genome into the human cellular chromatin. Studies have shown, however, that prior to integration, the unintegrated viral genome is transcriptional active. The role of cellular transcription factors, such as Sp1 and NF-κB, and the viral protein Tat with respect to regulation of preintegration transcription is not yet understood. This early transcriptional activity also generates three regulatory proteins: Tat, Nef and Rev. Nef from unintegrated viral DNA has been shown to downregulate the cell surface expression of CD4, CXCR4, CCR5, and MHC class I. However, the functional importance of Tat produced from preintegration transcription in the viral life cycle is unclear. In this dissertation, I demonstrate that either deleting Tat or TAR from the HIV-1 viral genome causes a decrease in the packaging of the virion genomes, resulting in a decrease in the subsequent viral reverse transcription and preintegration transcription in target cells. The Tat defect can be complemented by
providing the Tat in trans in the virion producer cells. I also demonstrate that the basal transcription from nonintegrated viral DNA is regulated by Sp1 and NF-κB; mutations in the Sp1 and NF-κB binding sites decrease preintegration transcription, and mutations in all three Sp1 binding sites cause the greatest impairment in preintegration transcription. I further demonstrate that Tat produced from unintegrated viral DNA is functional in stimulating HIV-1 LTR. Nevertheless, the nonintegrated viral DNA is less receptive to Tat transactivation than integrated provirus. I also confirm that nonintegrated viral DNA is assembled into a minichromatin which may inhibit preintegration transcription in the absence of Tat. Thus, the early Tat produced from nonintegrated viral DNA is functionally important in maintaining an open chromatin for persistent low-level transcription, which can modulate cellular activity to promote infection.
INTRODUCTION

Background of HIV-1 and AIDS

The current human immunodeficiency virus (HIV) pandemic most likely started in the mid- to late 1970s (1). The first sign of an epidemic among the population was in the beginning of the 1980s when an unusual increase in cases of Kaposi’s sarcoma (KS) (2) and Pneumocystis carinii pneumonia (PCP) (3) were observed. In 1982, CDC officially named the disease, that was caused by a still unknown pathogen, acquired immune deficiency syndrome (AIDS) (4). In that same year, it was noticed that children who either received blood transfusions, blood products, or were born to mothers with AIDS related infections, died from infections associated with AIDS (5,6). These symptoms indicated an infectious agent carried by human blood. The Institut Pasteur in France isolated the AIDS causing, blood-borne virus in May of 1983 (7), and it was named human immunodeficiency virus in 1986 (8). In that same year, a drug called azidothymidine (AZT) or also known as zidovudine, which is a nucleoside reverse transcriptase inhibitor (NRTi), was shown to slow down the progression of AIDS (9,10). AZT was approved in 1987 by the FDA as the first antiretroviral drug to treat AIDS (11).

It was thought that the HIV pandemic peaked in 1997 (12) with an estimate of 3.2 million new HIV infections globally (13). In 2010, the number of newly infected people went down by about 21% to 2.7 million incidences. This trend extended into 2012 with 2.3 million newly infected adults and children (14). Countries in sub-Saharan Africa with
the highest infection rate, such as Ethiopia, Nigeria, South Africa, Zambia and Zimbabwe, showed either stabilization of the epidemic or signs of decline in the HIV rate. On the other hand, some countries in Eastern Europe and Central Asia showed an increase in the HIV rate by more than 25% between 2001 and 2009 (13). In 2012, the estimated number of people living with HIV in the United States was between 920,000 and 1,800,000, and estimated deaths due to AIDS in that same year was between 16,000 and 27,000 (14). The numbers are still shockingly high even though so much progress has been made in the study of HIV, but it is also a testimony that more needs to be done in this field.

A primary HIV-1 infection causes in most instances fever, myalgia, rash, gastrointestinal symptoms, and in some cases neurologic complications (15). After the primary HIV-1 infection, a surge of virus replication takes place that causes high levels of viremia. This will lead to the spread of the virus to many organs, including the brain and lymphoid tissues (16). Hereafter, a prolonged period of clinical latency occurs, which is characterized by a very low number of infected blood cells. However, HIV accumulates in the lymphoid organs and is actively replicating (17). The latency state is followed by the progression to AIDS. At this point the immune system is no longer capable of containing opportunistic infections. Blood viral load increases, while the CD4+ T cells decrease to alarmingly low levels: < 200 cells/mm² of blood (18), and death is unavoidable.

Since the discovery of HIV in 1983, the virus has been abundantly studied. The virus targets mainly CD4+ T lymphocytes (19), macrophages (20), and dendritic cells
of the immune system, and microglial cells in the central nervous system (CNS) (22). Soon after the first isolation of HIV, CD4 receptors were found to be essential for the virus to enter the target cells (23). However, it was not until 1996 that Feng et al. discovered the requirement of a co-receptor for invading target cells (24). The chemokine receptors CXCR4, CCR5, and CCR3 were found to be the HIV co-receptors, and they belong to the family of G-protein-coupled receptors with seven membrane-spanning domains. CXCR4 is used as a co-receptor by viruses that are T cell-tropic (T-tropic), and CCR5 and/or CCR3 are used by macrophage-tropic (M-tropic) viruses (25–27). CCR3 is primarily used as a partner co-receptor for CCR5 for entry into macrophages and brain microglia (28). The first step of the HIV virus to enter a target cell is the interaction of the viral surface envelope glycoprotein (Env) with a CD4 receptor and the co-receptor on the cells (19). The heterodimer Env consists of a trans-membrane glycoprotein, gp41, and a surface glycoprotein, gp120. CD4 and the co-receptor of the target cell recognize gp120 (29), and upon recognition, a conformational change takes place in gp120 (30,31). As a result, gp41 activates fusion of viral and cellular membranes (29). Consequently, the viral core is released in the cytoplasm of the cell. The core consists of the viral genome (two copies of the single stranded, positive sense RNA), cellular tRNA<sub>Lys</sub> primers, and viral proteins [nucleocapsid, protease, reverse transcriptase, integrase, Nef (32,33), Vif (34), and Vpr (35)] (36). This new formed structure is now called the reverse transcription complex (RTC). However, the precise mechanism of the RTC is not well understood (37). About 2 hours post entry, the RNA is reverse transcribed in the cytoplasm into double stranded cDNA (36). A new structure is formed that is called pre-integration
complex (PIC) (38), and it is transported to the nucleus. HIV, being a lentivirus, is capable of entering the nucleus of dividing and non-dividing cells in a mitosis independent manner due to the viral nuclear-import machinery (39). After entering the nucleus, the viral genome can either be integrated into the host cell’s chromatin, which is mediated by the viral enzyme integrase, or the genome remains in the nucleus as unintegrated cDNA (40–44).

**Long terminal repeats include HIV-1 basic promoter**

The HIV-1 provirus consists of two long terminal repeats (LTRs), which are located on both extremities of the viral DNA. The complete LTRs (5’ and 3’) are produced during the reverse transcription process. The two HIV-1 RNA copies in the virion consist of the 5’ R and U5 region while the 3’ end includes the U3 and R regions. The virion-associated tRNA\text{3\textsubscript{Lys}} primer binds to the primer binding site (PBS) of the viral RNA, which is located just downstream of the 5’ U5 region, and the negative strand DNA initiation and elongation starts through the 5’ U5 and R region. At the end of the 5’ RNA, it either jumps to the 3’ end of the same RNA molecule or to the 3’ end of the second RNA copy where the R regions hybridize. The negative DNA strand is elongated through the U3 region and towards the 5’ end. At the PBS site, the second jump will occur to the same molecule where the PBS of the newly made positive DNA strand hybridizes with the PBS in the negative strand DNA to finish the production of the second LTR site (45). Each complete LTR consists of about 640 base pairs (bp) and is partitioned into 3 sections; the unique 5’ (U5), the unique 3’ (U3), and the repeat region (R) (46). The 5’ LTR functions as the HIV-1 promoter, while the 3’ LTR acts in
transcription termination and polyadenylation. The core or basal promoter region is located in the 5’ U3 region at nucleotide (nt) -78 to -1 relative to the transcription start side (47). This region includes three tandem Sp1 binding sites (47–49), which facilitates Tat activity (50), and this region also includes the TATA element (51–53), which is essential for initiation of transcription (54). Located upstream of the core promoter is the core enhancer (nt -105 to -78). The enhancer region consists of two NF-κB binding motifs (55), which are important for basal (Tat-independent) transcription from the viral LTR in cells such as the unstimulated lymphoid cells (55,56). Located upstream of the enhancer region is the modulatory region (nt -454 to -105). This region consists of many binding sites for cellular DNA-binding transcription factors such as AP-1, COUP, NFAT, and TCF-1a. These factors regulate HIV-1 expression in either a positively or negatively manner and are cell-type specific (48,55,57–63). Located downstream of the transcriptional start site in the R region, is the trans activation response (TAR) sequence, which transcribes the 59 base long functional TAR RNA (64).

Comprehensive mutagenesis studies of the HIV-1 5’ LTR revealed that this viral LTR is considered to be a “defective” promoter due to the fact that it does not contain any regulatory elements that are needed to start transcription by a completely processive polymerase. However, initiation of transcription occurs by a poorly processive polymerase (65). It appears that this is caused by a lack of a strong enhancer element in the HIV 5’ LTR (66).
**Transcription factor Sp1**

The specificity protein (Sp) 1 is a 785-amino acid, 100- to 110-kDa nuclear transcription factor, and is ubiquitously expressed in mammalian cells (67). Sp1 regulates gene expression by binding to GC-box sequences that are located in the core region of the promoter (68–70) using three C-terminal C2H2-type zinc finger motifs (70,71). This protein is a member of the Sp/Krüppel-like factor (KLF) transcription factor family. The highly conserved DNA-binding domain is shared between the family members (72). The Sp/KLF family is subdivided into the Sp and KLF subfamilies with the Sp family members preferring the GC-box DNA binding sequence, and the KLF family members favoring the GT-boxes (73). The Sp subfamily consists of Sp1–9, with Sp1-4 possessing N-terminal glutamine-rich transactivation domains, while Sp5-9 are lacking these transactivation domains (74). Sp1, Sp3, and Sp4 hold two glutamine-rich transactivation domains [domains A and B (72)] compared to Sp2, which includes only one of these domains (75). While Sp1 and Sp3 are ubiquitously expressed, Sp4 is expressed in specific tissue, with the brain as its preeminent tissue (76). In promoters with two or more Sp binding sites, Sp3 represses the Sp1-mediated transactivation when Sp3 replaces DNA bound Sp1. Sp3 is not capable of mediating transcriptional transactivation synergistically through two or more Sp sites like Sp1 can. For this reason, if Sp3 replaces Sp1 on a multiple Sp binding site promoter, it will result in a repression (77). However, the bifunctional transcription regulator, Sp3, is able to activate expression from a promoter with a single Sp binding site (78).

In the HIV-1 promoter, three tandem Sp1 binding sites are located upstream of the TATA box sequence in the core promoter of the viral 5’ LTR. It has been established that
Sp1 associates with itself (79) to activate transcription synergistically, but it also interacts with transcription factor Yin Yang (YY) 1 (80), and at least three components of the TFIID complex; the TBP (81), TAF4 [also known as TAFII130 and TAFII135 (82)] (72), and TAF7 (83) [also known as TAFII55 (84)] to functionally cooperate. In earlier investigations it has been established that the Sp1 binding sites and the TATA box sequences are required for both Tat-independent activation and for Tat-dependent activation of the LTR (49,85–87). It was revealed that Sp1 is capable of interacting with the viral transactivator Tat to form a tight protein-protein complex (50,88). The Tat region responsible for this interaction is amino acids 30 through 62 (50). The interaction between Sp1 and Tat alters the phosphorylation state of Sp1 via a double-stranded DNA-dependent protein kinase (DNA-PK), which phosphorylates Sp1 at serine 131. The DNA-PK is bound to Tat while phosphorylating Sp1. The phosphorylated Sp1 activates gene expression, and this phosphorylated Sp1 might be important only in the Tat-dependent transcription activation, but not in the Tat-independent transcription step (88).

Transcription factor NF-κB

The complex protein nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) is a potent inducible transcription factor that regulates transcription of genes that are important for the immune and inflammatory response (89) in activated T cells. The members of the NF-κB family consist of RelA, or also known as p65 (referred to as RelA/p65), c-Rel, RelB, p50, and p52. They all include an intensely conserved 300-amino acid long Rel homology domain (RHD) which is important for dimerization and DNA binding. The transcriptional activation domain is only included in RelA/p65, RelB,
and c-Rel (90). For this reason, we find RelA/p65 as the majority subunit in the nucleus of activated T cells while we find p50 as the most abundant subunit in the nucleus of resting T cells (91).

The inhibitor protein IκB-α is the most bountiful NF-κB inhibitor in cells and it keeps NF-κB sequestered in the cytoplasm. As soon as IκB-kinase (IKK) is activated by inducers, such as tumor necrosis factor (TNF)-α, the kinase will phosphorylate IκB at serine residues 32 and 36. This phosphorylation causes the inhibitory protein to be ubiquitinated at lysine residues 21 and 22, and will accordingly be degraded by the 26S proteasome. Consequently, the functional NF-κB will be translocated to the nucleus (92–95) and regulate gene expression from the promoter. However, in addition to stimuli such as mitogens, stress factors, and inflammatory cytokines, an HIV-1 gene product is also known to be skilled to cause NF-κB to be activated and shifted to the nucleus. The viral envelope glycoprotein gp120 was identified as a viral protein that activates NF-κB. Due to the binding of the gp120 to the cell surface-expressed CD4 receptors on target cells, phosphorylation and degradation of IκB takes place, and consequently p50-RelA/p65 heterodimers were able to translocate to the nucleus (96).

NF-κB also targets the enhancer of the HIV-1 promoter for activation. The viral enhancer includes two tandem NF-κB binding sites that are at least recognized by four different homo- and heterodimer proteins, which are made up of members of the NF-κB family (97). These transcription factors exercise both an activating and inhibitory impact on the HIV-1 5’ LTR. Homodimers of p50 recruit histone deacetylase (HDAC)-1 which causes tightening of the chromatin and consequently inhibition of transcription (98). On
the other hand, the replacement of p50 homodimers with p50-RelA/p65 heterodimers instigates activation of transcription. The p50-RelA/p65 heterodimer recruits the histone acetyl transferases (HATs) p300/CBP (99), which causes a more open and accessible DNA template.

Phosphorylation and acetylation at various residues of the transcriptional activating subunit RelA/p65 causes regulation of NF-κB activity (100–106). The RelA/p65 subunit is acetylated by p300/CBP at lysines 218, 221, and 310 to regulate the NF-κB function, such as DNA binding, IkB-α assembly and NF-κB transcriptional activity (105,106). Acetylation of lysine 218 hinders the assembly of RelA/p65 with the newly synthesized IkB-α. Acetylation of lysine 221 increases the DNA binding affinity of RelA/p65 to the κB enhancer, and acetylation of lysine 310 enhances the transcriptional activity of NF-κB (106,107). HDAC-3 (105) and SIRT1 (107) are capable of reversing the lysine acetylation by p300/CBP. The deacetylation of RelA/p65 at lysine 310 is prevented by viral protein Tat (108). The deacetylation of lysine 221 by HDAC-3 recovers the capacity of newly made IkB-α to bind to RelA/p65 (109). Deacetylation of RelA/p65 by HDAC-3 or SIRT1, or additional acetylation of the protein at lysine residues 122 and 123 down-regulates the NF-κB activity (105,107,110).

Other transcription factors
Another regulatory host factor important to HIV-1 transcription and replication is the nuclear factor of activated T cells (NFAT). The NFAT protein family consists of NFAT1 through NFAT5. The NFAT family members play an important role in inducible gene transcription critical for the immune response (111). Just like the NF-κB members,
the inactive NFAT, which is the phosphorylated form, is located in the cytoplasm. This protein is activated through cell-surface receptors that activate phospholipase C, which consequently causes the cytosolic Ca^{2+} concentration to increase by releasing stored Ca^{2+} located in the endoplasmic reticulum (ER). The exhaust of the ER stored Ca^{2+} supply triggers extracellular Ca^{2+} entry. Thereupon, the calcium/calmodulin-dependent serine phosphatase calcineurin is activated, and the protein dephosphorylates the N-terminal regulatory domain of NFAT to activate the protein. As a consequence, it is translocated to the nucleus. This N-terminal regulatory domain is highly conserved among the NFAT family members (111,112).

It was found that the HIV-1 envelope protein gp120 is also capable of inducing the activation of NFAT proteins and their translocation to the nucleus in nonstimulated CD4^+ primary T cells (113). This gp120-mediated NFAT activation resulted from CD4 and either CCR5 or CXCR4 signaling (113). Not only was gp120 found to be capable of inducing NFAT, but also the viral protein Nef. On the other hand, viral protein Tat was found to activate the 5’ LTR by an NFAT-independent pathway (114).

Besides the N-terminal regulatory domain, NFAT consists of a C-terminal transactivation domain and holds the same DNA binding domain as NF-κB, the rel homology region (RHR). NFAT1 and NFAT2 proteins are expressed in T cells and macrophages, which are the primary target cells for HIV-1 (113). In the HIV-1 5’ LTR, an NFAT binding site was identified upstream of the κB sites (nt -254 to -216 relative to the transcriptional start site), but no study was able to link this site to viral transcription activity (115). Nonetheless, it was observed that NFAT1 binds, as a protein dimer, to the
κB sites that are located in the enhancer of the viral promoter (116,117). In another study it has been shown that both NFAT1 and NFAT2 bind to the HIV-1 NF-κB binding site and enhance HIV-1 LTR transcription (118).

Another host cellular transcription factor is the CCAAT-enhancer-binding protein (C/EBP). The C/EBP family includes at least seven family members. This protein holds a basic leucine-zipper domain at the C-terminus. This domain is involved in DNA-binding and dimerization. Additionally, the C/EBP consists of an N-terminal transactivation region. C/EBP is expressed in adipocytes, hepatocytes and monocytes/macrophages (119). The C/EBP family member NF-IL6 (nuclear factor of interleukin 6) or also known as C/EBPβ was found to bind to three sites on the HIV-1 5’ LTR (120). The three binding sites are located upstream of the NF-κB sites. The C/EBPβ binding site that is located adjacent to the NF-κB site is the most important site (located at nt -120 to -109 relative to the transcriptional start site). It was reported that binding of C/EBPβ to the HIV-1 LTR in monocytes/macrophages causes stimulation of transcription both before and after activation of the cells (121).

The cellular factor USF (upstream stimulating factor) is a member of the conserved basic helix-loop-helix-leucine zipper transcription factor family. These factors bind to the E box regulatory sequence CANNTG (122). A conserved E box that is located at nt -162 to -167 upstream of the transcription start site in the HIV-1 5’ LTR is the binding site for USF. The binding of USF to the E box has a positive effect on LTR promoter driven HIV-1 transcription (123).
Located just upstream of the NF-κB sites [at nt -140 to -120 relative to the transcriptional start site (124)] is a well conserved cis element, RBEIII, that functions as the binding site for the nuclear factor Ras response element binding factor 2 (RBF-2) (125,126). RBF-2 consists of USF1 and USF2, and this heterodimer interacts with RBEIII with the help of the cofactor TFII-I (125). USF1 and USF2 are basic helix-loop-helix-leucine zipper DNA binding factors that bind, mostly as heterodimers, E box cis elements. However, RBEIII does not consist of an E box, and therefore the USF heterodimer also consists of an E box-independent function. It was found that the USF1/USF2/TFII-I complex is important in the transcriptional reactivation of latent viruses after T cell activation (125).

Regulation of HIV-1 transcription prior to Tat production is dependent on several host cell regulatory factors, of which only a few were briefly discussed above. There are other cellular transcription factors that interact with the HIV-1 5’ LTR. Some examples are YY1, Ets-1, AP-1, c-Myb, and COUP-TF. The different transcription factors are able to either activate or repress viral transcription. A review of cellular transcription factors that interact with the HIV-1 LTR promoter is published by Pereira et al. (127). Also, the different binding sites for the transcription factors (cis elements) are not all conserved equally. The three highest conserved cis elements in the HIV-1 5’ LTR are the binding sites for the factors NF-κB, RBF-2, and TBP (125).

**Chromatin structure**

Actively transcribed cellular genes are located in a less condensed chromatin structure called euchromatin. Non-expressed genes are located in a more condensed
structure that is called heterochromatin (128). In a chromatin structure, DNA consisting of about 146 base pairs is wrapped around an octamer of a pair of each H2A, H2B, H3 and H4 histone proteins, called core histones. The DNA is coiled around the octamer of core histones in a 1.67 left-handed super helical turn, and this complex is called a nucleosome structure. Between the nucleosomes are about 80 nucleotides that contain the H1 histone protein (129).

After the HIV-1 genome is integrated into the host genome, the viral DNA is also packaged into chromatin. Independent of the integration site, two nucleosomes, nucleosome-0 (nuc-0) and nucleosome-1 (nuc-1), are deposited onto the HIV-1 5’ LTR (130,131). In a viral latent state, nuc-0 is located at nucleotide -405 to -245, and nuc-1 at nt +20 to +165 relative to the transcription start site (46). Located between nuc-0 and nuc-2 (located just downstream of the 5’ LTR) are two important open, nucleosome-free regions that are positioned at nucleotide -244 to +19, and at nt +166 to +256 relative to the transcriptional start site (46). The first nucleosome-free DNA region contains the LTR modulatory region, and the enhancer region, which includes two NF-κB sites, and the core basal promoter, containing the three Sp1 binding sites (132) and the TATA box.

The packaging of the viral DNA controls the regulation of gene expression by transcription factors. To give transcription factors access to the viral DNA, nuc-1 needs to be remodeled (130). This action can be accomplished in two manners: 1) ATP-dependent chromatin remodeling and/or 2) histone acetylation. The ATP-dependent chromatin remodeling complexes use the energy of ATP-hydrolysis to modify the interactions between the viral DNA and the histones (133). These remodeling complexes
consist of a helicase/ATPase domain. The human SWI/SNF complex includes either the BRG-1 or hBRM as ATPase subunit (134), and both are critical in the activation of the HIV-1 LTR. Studies have determined that the SWI/SNF complex is needed in Tat transactivation of the promoter (135–137). Tat recruits hBRM to the HIV-1 LTR and the hBRM/Tat interaction is regulated by Tat acetylation at lysine 50 (136). It was observed by Mahmoudi and co-authors that Tat also interacts with BRG-1. However, this interaction requires Tat acetylation on lysines 50 and 51 by acetyltransferase p300/CBP. The Tat/BRG-1 complex localizes to the HIV promoter site that is obstructed by nuc-1 (137). Interestingly, acetylation of Tat at lysine 50 hinders Tat from interacting with hBRM (136).

Remodeling of nuc-1 can additionally be established by histone acetylation. Histone acetyltransferases (HATs) transfer an acetyl group from acetyl coenzyme A (Acetyl-CoA) onto specific lysine residues located in the N-terminus of the protruding tails from the core histones (138,139). This acetylation of the core histones results in loosening up the chromatin (138,139), and therefore the DNA is more accessible to transcription factors and the RNA Polymerase II (RNA Pol.II). Tat is known to interact with HATs, such as p300/CBP (140), P/CAF (141), GCN5 (142), Tip60 (143), and TAFII250 (144). The counterpart of HATs are Histone Deacetylases (HDACs) which remove the acetyl groups and cause the chromatin to tighten back up and eventually gene silencing will occur. The repression of the HIV-1 5’ LTR by hypo-acetylation is induced by two transcription factors, YY1 and LSF. These two host factors collectively recruit HDAC-1 to the viral LTR (145,146). Sp1 is also known to repress LTR activation by
recruiting c-Myc to the LTR, and c-Myc sequentially recruits HDAC-1 (147). The thyroid hormone receptor and the subunit of NF-κB, p50, are also known to recruit HDAC-1 to the viral LTR region (98,148).

**HIV-1 transcription initiation**

Initiation of HIV-1 gene expression is influenced by extracellular stimuli such as cytokines, which cause activation of host transcription factors (149). The HIV-1 5’ LTR includes all the required elements for initiation of transcription: two NF-κB binding sites, three Sp1 binding sites, a TATA element, and an initiator sequence (INR) (51,56,61,150). The general or basal transcription factors include TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH. The transcription initiation factor TFIID is the key basal transcription factor and is involved in the recognition and binding of the core promoter. TFIID is a multi-subunit complex and includes the TATA-binding protein (TBP) and many co-factors called TBP-associated factors (TAFs) (151–156). Interestingly, the TBP is required for basal transcription, but not the TAFs. TAFs are needed for the activated transcription (157). TFIID initiates the organized cluster of various other general transcription factors and RNA Pol. II into a transcription initiation complex or also known as the pre-initiation complex (158–160). First the TBP subunit of TFIID binds to the TATA box and subunits TAF1 and TAF2 recognize the INR. TFIIA followed by TFIIIB are recruited to stabilize the promoter bound TFIID. TFIIA promotes binding of TBP to the TATA box, and TFIIB interacts with TBP and assists in the polymerase recruitment to the core promoter. TFIIB also recruits TFIIF to the promoter. TFIIE, TFIIF, and TFIIH mediate the unwinding of the DNA (155,156). The initiation factor TFIIH consists of a kinase, cyclin-
dependent kinase (CDK) 7, that assists in promoter clearance by phosphorylating the serine 5 residues in the carboxyl terminal domain (CTD) of the RNA Pol. II (161–163). The CTD includes repeats of the consensus sequence tyrosine\(^1\)-serine\(^2\)-proline\(^3\)-threonine\(^4\)-serine\(^5\)-proline\(^6\)-serine\(^7\) (164). Transcription factor TFIIF was found not to be essential for initiation by RNA Pol. II, but the study showed that TFIIF was critical in stabilizing TFIIB within the transcription initiation complex. TFIIF was demonstrated to be important in early elongation complexes to stimulate the rate of elongation (165–170).

**HIV-1 transcription elongation**

Normally, cellular gene activation is regulated by DNA-binding regulatory proteins. However, in an HIV-1 infected cell, viral gene expression is regulated by cellular proteins and the viral transactivator protein Tat. It was found that without Tat, the RNA polymerase comes to a halt close to the promoter after transcription initiation (171). The progress of HIV-1 transcription from initiation to elongation has to overcome a specific “block”. A kinase is needed to surpass this hurdle before the production of long transcripts can take place. The kinase that regulates the viral gene expression is the positive transcription elongation factor b (P-TEFb) (172).

The role of viral protein Tat in transactivation of transcription involves the RNA regulatory element, the trans-activation response element (TAR), whose sequence is located downstream of the transcription start site (nucleotides +1 through +59 relative to the transcriptional start site). For TAR to be operative, its sequence needs to be located 3’ to the viral promoter and in the proper orientation (173). The TAR RNA sequence forms a stable, nuclease-resistant stem-loop structure (174,175). Tat binds directly to the U-rich
bulge that is located close to the apex of this complex (176–178). TAR undergoes a conformational change induced by Tat during attachment (179–182) for a proper binding.

A stable ternary complex can be formed with TAR RNA, Tat, and the regulator of CDK kinases, cyclin-T1 (183). It was shown that Tat is not active in rodent cells due to an amino acid change at position 261 of cyclin-T1. The human cyclin-T1 consists of a cysteine residue at position 261, while the murine cycline-T1 includes a tyrosine residue in this amino acid position (184–190). Mouse cyclin-T1 is not capable of binding Tat to form the TAR-Tat-cyclin-T1 structure, and therefore rodent cells are deficient in HIV-1 transcription (191). Cyclin-T1 and the cyclin-dependent kinase (CDK) 9 together form the Tat-associated kinase (TAK) (192–194). This complex has been renamed after the human homolog of *Drosophila*: the positive transcription elongation factor b (P-TEFb). Viral protein Tat hijacks P-TEFb from the 7SK small nuclear RNA (snRNA) complex, where P-TEFb is inactive (195). The human 7SK snRNA is located in the nucleus and along with affiliated cellular proteins this complex controls the activity of P-TEFb (195,196), which regulates the elongation of RNA Pol. II (197,198). Binding of Tat to P-TEFb generates a conformational change in CDK9 that causes activation of the enzyme (161,183). The Tat-pTEFb complex binds to TAR from where they are recruited to the transcription complex that is formed on the template (183). The role of the TAR RNA is solely to recruit Tat and cellular co-factors. At the transcription site a large protein complex is formed that in addition to Tat and P-TEFb, also includes the human transcription factors/coactivators AFF4, ENL, AF9, and the elongation factor ELL2 (199,200). Here they stimulate processivity of the RNA Pol. II (162,163). CDK 9 hyper-
phosphorylates serines 2 of the CTD of the RNA Pol. II and assists in termination of promoter proximal pausing (161,197,201–204). In the presence of Tat, CDK 9 phosphorylates both serines 2 and 5 (205). In the absence of Tat, the negative elongation factor (NELF) and the DRB sensitivity-inducing factor (DSIF) will bind to the viral DNA, and hinder transcript elongation by the RNA Pol. II. This results in short viral transcripts of about 60 nucleotides long (206,207). In the presence of Tat, CDK 9 also phosphorylates the transcription elongation factor Spt5 (a component of the DSIF complex). This phosphorylation will prevent premature RNA Pol. II release, and transforms the DSIF to a positive elongation factor. NELF is also phosphorylated by CDK 9 and causes NELF to detach from the complex (196,208–210). The elongation factor ELL2 prevents backtracking of the RNA Pol. II (211,212). The early, Tat-independent HIV-1 transcription phase contributes to low-level basal gene expression. The later, Tat-dependent transcription phase of HIV-1 supports high levels of transcription of viral genes (213).

**Viral protein Tat**

Viral protein Tat is a multifunctional protein that is required for productive and processive viral gene expression, and viral replication. The viral *tat* gene consists of two exons, with the second exon in part overlapping the *envelope* gene (214). The first exon encodes amino acids 1 – 72 and the second exon encodes amino acids 73 – 101. The Tat protein is a small, nuclear, regulatory protein that, depending on the viral strain, consists of 86 to 101 amino acids, and 14 – 16 kDa. The laboratory strain HIV-1 vNL4-3 consists of the truncated, 86 amino acids long Tat, while most clinical isolates contain the 101
amino acid long protein (46). One of Tat’s most important functions in the viral life cycle is the activation of transcriptional elongation of the integrated provirus. Sodroski et al. published the first evidence that the protein Tat was involved in the regulation of viral gene expression in 1985. They used a reporter gene under the control of a viral LTR and observed a 200 to 300 fold increase in reporter expression in cells that were first infected with HIV-1 (215). It was found that without Tat, the RNA polymerase comes to a halt close to the promoter after transcription initiation (171). This function of Tat is described in more detail in the previous section “HIV-1 transcription elongation”.

The protein Tat consists of six domains. The first domain, the N-terminus, is the acidic domain and includes amino acids 1 through 20 (216,217) and consists of a proline-rich region and holds a conserved tryptophan residue (46). It was demonstrated that this domain is involved in the reverse transcription process of the virus (218). The second domain, including amino acids 22 through 37, consists of seven cysteines at positions 22, 25, 27, 30, 31, 34, and 37 (46) that are highly conserved among different HIV-1 isolates. This domain is called the cysteine-rich domain (217). Four of the seven cysteines are accountable for the formation of two disulfide bridges which were found to be important for efficient transactivation of HIV-1 (219). Lysine 28 in this domain is acetylated by pCAF and this modification enhances Tat binding to P-TEFb and stabilizes the TAR-Tat-cyclin-T1 complex (220,221). To reverse this reaction, HDAC-6 deacetylates Tat at lysine 28 (222). Amino acids 40 through 48 make up the third domain or core domain and consists of the hydrophobic core sequence LGISYG (46,223). It was found that lysine 41 is important in binding Tat to TFIID (224). The minimal activation domain of
HIV-1 Tat consists of the first 3 domains (225,226). Several single residue mutations in the first domain are tolerated without any serious changes in protein function, however, mutations in the second and third domain are damaging to transactivation activity (46). The fourth domain is termed the basic domain, also known as the TAR-binding domain, or ARM (arginine-rich motif). This domain consists of amino acids 49 through 57 (46). This fourth domain includes the basic, well-conserved RKKRRQRRR motif (218). This domain possesses the TAR RNA-binding capability of Tat (176,227–232). Lysine 50 is acetylated by p300/CBP and by hGCN5 (human general control of amino acid synthesis 5). This acetylation action results in the disruption of the TAR-Tat-cyclin-T1 complex. Tat detaches from TAR and binds instead to the bromodomain [acetyl-lysine binding domain (233)] of pCAF (234) and this results in the transfer of Tat and pCAF to the elongating polymerase (235). The fifth domain, residues 58 through 72, is a region rich in glutamines and shows the most genetic variability (46). Together the fourth and fifth domains are called the basic region. This basic region is important for Tat’s nuclear localization (236,237), and the residues 48 – 60 are utilized for the translocation through the plasma membrane of cells (236–239). The sixth domain is encoded by the second exon of Tat, amino acids 73 – 101, and is not required for transactivation activity. This domain includes a RGD sequence and functions as a cell adhesion signal for cellular integrin binding (240), but HIV-2 and SIV Tat do not possess this sequence (217). This terminal domain also includes a highly conserved ESKKKVE motif, but the significance of this motif is not completely identified yet (217). Some studies suggest that this motif is associated to the optimization of in vivo HIV-1 replication (241,242). In a study by
Mahlknecht et al., it was found that the second coding exon of Tat plays an important function in the NF-κB-dependent regulation of the HIV-1 transcription in T cells (243).

Protein Tat can be released from non-ruptured infected cells (244) and can be taken up by latently infected cells (245) to reactivate HIV-1 gene expression. Tat can also be taken up by uninfected cells and induce apoptosis (245–251). Tat was also found to behave as a secreted growth factor in stimulating the growth of the Kaposi-like cells (244,246,252) and in angiogenesis (253). The protein is also capable of transforming the normal organization of neurons and astrocytes (254), and at low concentrations, the protein is a neurotoxin (255) by altering neuronal calcium flux (256). Tat is additionally able to modulate cytokine expression such as TNF (257,258), II-2 (259), and II-6 (260–262).

It has also been established that HIV-1 Tat binds directly to basal transcription factors. Kashanchi et al. reported that Tat binds to general transcription factor TFIID. Tat binds the TBP subunit of TFIID directly through the amino acids 36 through 50. The protein-protein interaction is inhibited by an amino acid substitution at lysine 41 of the Tat protein. Lysine 41 is important for Tat transactivation, but does not intervene with the binding of Tat to TAR RNA, which may indicate that the TFIID and Tat interaction is important in transcriptional activation (263). It was also found that Tat interacts with TFIIB. It was proposed that the interaction of Tat with the basal transcription factors TFIIB and TBP causes the stabilization of the initiation complex that forms on the TATA box, and as a result, increases the transcriptional initiation rate (264). In a study by Zhou and Rana, they concluded that in the early transcription elongation process, after the
promoter clearance, and before the synthesis of full-length TAR RNA transcript, a Tat molecule associates with the RNA Pol. II complex. This interaction was found to be TAR and P-TEFb independent. Their study indicates a Tat binding site on the RNA Pol. II, close to the exit site for the nascent mRNA transcripts. The proposed functions of the early binding of Tat to the elongation complex were: 1) Tat prevents premature termination 2) Tat modulates the function of negative elongation factors such as DSIF and NELF, and 3) Tat plays a role in the mRNA capping reaction (265).

It is also shown that Tat has a function in the reverse transcription process of the virus (266). The viral protein reverse transcriptase and the host cell derived primer tRNA_{3}^{Lys} dominate this process. The tRNA primer binds to the viral primer binding site (PBS) (267) and the reverse transcriptase initiates DNA synthesis starting from the 5’ U5 and R elements to the strong stop DNA. Here the enzyme jumps to either the 3’ site of the same RNA genome or to the 3’ site of the second RNA genome where the R elements hybridize (36,268,269). It was reported by Harrich et al. that the Tat protein participates in initiating reverse transcription. HIV-1 virions lacking protein Tat were found to be incompetent to undergo effective reverse transcription; however, the levels of reverse transcriptase and genomic RNA were unaltered (266). The amino terminal domain of Tat was identified to be responsible for the task in reverse transcription (218). In a follow-up study by the Harrich group, they found that Tat needs to form a complex with the reverse transcriptase to be able to enhance reverse transcription. The reverse transcriptase is a heterodimer consisting of subunits p51 and p66, and it was shown that Tat interacts with subunit p51. This concept requires Tat to be present in the virion and according to this
study, the ratio of Tat to reverse transcriptase must be at least 2:1 for increased reverse transcription activity (270).

Integrated versus nonintegrated HIV-1 genome

An essential early step in a productive HIV-1 lifecycle is the integration of the viral DNA into the host genome. However, high levels of unintegrated DNA were found in both activated and quiescent CD4+ T cells (44,271). As early as 1984, Shaw et al. detected unintegrated, linear and circular double stranded HIV DNA in the H9 human T cell line and in biopsied lymph node tissue from AIDS patients (272). Pang et al. detected the three different species of HIV-1 DNA in blood and brain tissue of AIDS patients as well. The three DNA forms were the unintegrated linear, unintegrated circular (both one and two LTR containing circular form), and the integrated viral DNA (36,41). The same group detected sizable higher amounts of unintegrated viral DNA in brain tissue samples from patients with HIV encephalitis compared to samples from patients with no encephalitis (41).

The primary destination for HIV-1 is CD4+ T cells, and for a productive viral life cycle, these cells need to be in an active state. In vivo, circulating activated CD4+ T cells are rare. Stevenson et al. found that HIV-1 infects the quiescent, not replicating, T cells, but the virus was not competent to integrate its DNA into the cellular chromosomal DNA (273). However, the viral DNA stayed in the cell as extrachromosomal DNA and was found to be transcriptionally active. This unintegrated viral DNA resided in the cells for several weeks after infection of the quiescent T cell, and was capable of integrating into the host genome after cell activation and maintained an effective viral life cycle (273).
Before the viral template is able to integrate into the chromosomal DNA of an infected cell, the incoming viral genomic RNA needs to be reverse transcribed by the viral protein reverse transcriptase. The produced double-stranded viral DNA is covalently inserted in the host genome and this process is catalyzed by the viral protein integrase (274). The first step of this integration method is processing the 3’ ends of the linear viral DNA by removing 2 nucleotides that are located adjacent to the conserved dinucleotide CA. This action is called the 3’ processing step. After the transport of the viral DNA into the nucleus, the 3’ CA-OH ends are covalently joined to the 5’ phosphoryl ends of the host DNA, and this process is known as the DNA strand transfer (275,276). The joining ends are most likely repaired by cellular enzymes (277).

The viral protein integrase is encoded by the pol gene and is synthesized as part of the Gag-Pol polyprotein. During maturation, Pol is cleaved into three proteins; reverse transcriptase (RT), protease (PR), and integrase (IN). IN is a 288 amino acid, 32 kDa (278) viral protein with about 40-100 IN molecules contained in the viral particles. Besides IN catalyzing the integration reaction, it also aids in the reverse transcription process (277,279), and is involved in the nuclear import of the preintegration complex (PIC) (280). This HIV-1 virion-associated protein consists of three distinct structural and functional domains. The first domain is the N-terminal zinc-binding domain (NTD) consisting of residues 1 through 46 (281,282). The second domain is the central catalytic core domain (CCD), consisting of residues 56 through 186. The third domain is the C-terminal domain (CTD) that consists of residues 195 through 288 (276,282). All three domains are necessary for the 3’ processing and the strand transfer steps in the integration
process (283–285). However, an active site was found within the central catalytic core domain that consists of three extremely conserved amino acid residues, D-64, D-116, and E-152. These conserved residues are involved in the 3’ processing and DNA strand transfer activities. When mutated, these actions are abolished (286,287). In a mutagenesis study by Engelman et al., it was found that a mutation at amino acid 116 in the catalytic region of IN, caused the negatively charged aspartic acid (D) to change to the hydrophobic asparagine (N), and instigated the inhibition of the 3’ processing activity of IN. The ends of the viral cDNA maintained blunt ends and were unable to integrate in the host DNA (277). It was observed that cells infected with the IN mutant D116N generated a compilation of nonintegrated viral DNA species that functioned as a template for the expression of Tat (277). This mutation in the catalytic site residue did not alter any other steps upstream of the 3’ processing action of the reverse transcribed viral DNA, nor did it restrict nuclear transport (277). A virus constructed with a D116N mutation is a nonintegrating virus and provides a good system to study nonintegrated HIV-1 DNA species.

Wu and Marsh were the first pioneers to demonstrate that in primary resting T cells, transcription occurred from nonintegrated viral DNA (288). They infected quiescent CD4⁺ T cells and activated the cells with CD3-CD28 immobilized antibody beads. Cell activation was measured by IL-2 expression. They found an increase in IL-2 expression in cells that were exposed to HIV-1 compared to non HIV-1 exposed cells. By using a reverse transcription inhibitor, AZT (3’-azido-3’-deoxythymidine), they determined that the increase in IL-2 was not caused by the binding of the viral envelope
to the host receptors or by virion associated proteins, but rather caused by actions
downstream of the reverse transcription process. They also demonstrated the presence of
tat transcripts in non-active cells, with the majority being the nef transcripts. The
nef and tat transcripts are two of the three products produced from multiple splicing of
the genome (288). This result confirmed earlier work by Spina et al. (289). Wu and
Marsh also looked for the Nef and Tat proteins in infected non-active cells. Five days
after infection, the Nef protein was detected by Western analysis, but they were not able
to detect the Tat protein. This could be due to a reduced sensitivity in the Western
analysis. Looking for the possible function of Nef in non-active cells, they generated a
nef negative virus which was used to infect resting T cells. The cells were activated 5
days post infection and IL-2 was measured. They found a small increase in IL-2 excretion
in the nef negative primary resting T cell infection, which could have resulted from
potential Tat expression. To evaluate the Nef function in cell activation, they looked at
the viral replication. Only an increase in viral replication was noted in an infection with a
functional nef gene. This increase was not notable in a pre-activated cell infection (288).

In a follow-up study by Wu and Marsh, they observed that the nonintegrated
DNA was able to generate all forms of transcripts in an activated T cell. However, this
transcription was in a transient fashion, and peaked at 12 hours postinfection. This trend
is most likely due to the active replicating nature of the cell. They studied nonintegration
transcription from a mutant integrase infection (IN-) compared to a wild-type integrase
viral infection (IN+). The two researchers recognized that the early transcription pattern
of an IN- infection was the same as in an IN+ infection. This result indicated that early
transcription is produced from nonintegrated DNA in both the IN- and IN+ infection. With this result, Wu and Marsh concluded that the early transcription from nonintegrated HIV DNA is a normal progression in the life cycle of an HIV-1 infection (290).

It was demonstrated that the three early viral proteins, Nef (288,291), Rev (288,292,293), and Tat (277,294), are expressed from nonintegrated viral DNA. Nef produced from preintegration transcription was found to be important in the enhancement of the activation of T cells (288). Early Nef expression was further linked to the down regulation of cell surface CD4, CXCR4, and to a lesser extent, CCR5 (291,295). It was also found that Nef was proficient in the modulation of the cell surface expression of major histocompatibility complex-1 (MHC-1). As a consequence, the cell is able to evade the immune system more efficiently (296). In the insufficiency of integration, Rev does not support the protein expression of singly and non-spliced transcripts (290). The function of produced Tat from nonintegrated DNA is not yet fully understood.

**Viral latency**

Herpes simplex virus type 1 (HSV-1) is one of the members of the herpes virus family that establishes latency in ganglionic sensory neurons to escape host immune responses and persist in the host. HSV-1 gene expression is very low in the latent phase of the infection, but reactivation from latency by external stimuli result in high viral gene expression (297). The main action of HIV-1 to bypass host immune responses is an accelerated mutation rate (298). However, HIV-1 can also establish a latent infection state in individual cells. An efficient HIV-1 life cycle requires integration of the viral DNA in the host genome. Activated CD4+ T cells are one of the primary target cells of
HIV-1, but in most cases these cells die after about one day post infection due to viral cytopathic effects, or due to the responses of the host immune system (149,299,300). Only a few of the surviving infected cells are able to return back to a resting position. If a cell reverts back to the long-lived resting state, so will the virus. The viral genome is stably integrated in the host genome during the active infection, but while the cell is in its resting G₀ state, the provirus is transcriptionally silenced. In this form, the virus is protected from antiretroviral drugs and the host immune system (44,301–308). For reactivation, HIV-1 relies on inducible host transcription factors that are only set in motion after antigen exposure by the resting CD4⁺ T cells.

Nevertheless, there are many more possibilities that can lead to post-integrated latency. The availability and functionality of viral protein Tat is probably the most important cause to the occurrence of latent viral reservoirs (209). HIV-1 transcription can be broken down in two specific phases: the early, Tat-independent phase (basal level transcription), and the late, Tat-dependent transcription phase. The later phase is characterized with high-level viral transcription. It was shown that low levels or the absence of Tat and/or its cofactor P-TEFb correlates to HIV-1 latency (209,309).

Besides viral protein Tat, HIV-1 gene expression also relies on host cell transcription machinery. This includes host transcriptional activators and repressors. If the cellular activators are low or absent, or there is an abundance of cellular repressors, viral latency is unavoidable. The interplay between inducible transcription factors that transactivate the viral LTR, such as NF-κB(p65) (55,310), AP-1 (311–313) and NFAT (118,314,315), and transcriptional repressors, such as NF-κB(p50) (98,310), LSF and
YY1 (316,317) are crucial for the virus to either enter the latency state or becoming transcriptional active.

Post translational modifications of chromatin are associated with the preservation of viral latency. Recruiting HATs (133,141) and the SWI/SNF complex (136) will decompress the chromatin while the recruitment of HDACs (318,319), and methyl transferases (320,321) compress the chromatin, which consequently could lead to viral latency or maintenance of latent viral reservoirs. Earlier studies suggested that the integration of HIV-1 into the host genome in or near heterochromatin led to latency (322,323). However, later research located latent proviruses mainly in sites with actively transcribed genes (324–326). When HIV-1 integrates into highly active gene transcription regions (327), the integrated provirus risks transcriptional interference between the host and viral promoter which could lead to latency. For example, there will be a competition for the RNA Pol. II and the cellular transcription factors needed for transcription initiation (324). This could prohibit the formation of the pre-initiation complex. On the other hand, in a quiescent cell, it is not the competition, but the low availability of the cellular components of the pre-initiation complex that cause the lack of pre-initiation complex formation, which subsequently leads to latency (328).

Another danger of integrating into a region with actively transcribed genes is the displacement or collision of the RNA Pol. II and transcription factors. If the host gene and the provirus are aligned in opposite directions, the two polymerase complexes will collide and transcription is prevented. This can potentially lead to latency. Also, the polymerase complex of the host gene can displace the polymerase complex of the viral
LTR by read through transcription from an upstream promoter (329,330). However, if Tat is produced, the integrated provirus becomes transcriptionally active. The provirus can either turn off host gene transcription or a cooperation of expressing both genes can take place (331). Nevertheless, this approach does require the early production of Tat (324,332). [For reviews see (207,328,333)].

Two major latent viral reservoirs are identified in an HIV-1 infection. One is the latently infected resting CD4+ T cells and the other are macrophages. Compared to T cells, macrophages are better equipped to resist the cytopathic effects of the infection, and are also better in evading the defensive actions of the host’s immune system. With this capability, the macrophages can transport the virus to tissues and organs, and cross the blood-tissue barrier to the brain (334). Macrophages are among the first to be targeted by HIV-1 due to the fact that most viruses that are involved in the first round of infection use the co-receptor CCR5 to enter the host cell (335).

Highly active antiretroviral therapy (HAART) is a combination of three or more effective anti-HIV-1 drugs that each targets a different step in the viral life cycle. HAART was introduced in 1996 and considerably lengthened the life of HIV-seropositive individuals (336). HAART administered to HIV-1 infected individuals is capable of lowering viremia levels at levels below limits of detection (< 20 copies/mL) (337,338). Nevertheless, when HAART is interrupted, the level of virus in the blood plasma is rapidly increased (339). This rebound of viremia indicates that HAART is not able to eliminate the virus from the body; it keeps the level of viruses in the blood in check. The virus remains latent in cellular reservoirs that will flare up after discontinuing
of HAART (340). We should also mention that the compounds that comprise the HAART regimen do not pass through the blood-brain or blood-nerve barriers, and HIV-1 is proficient in replicating in the CNS. This could generate a reservoir of HAART-resistant viruses (341).

Unintegrated DNA can also contribute to latency. It was observed in an early study that in infected quiescent memory CD4+ T cells the predominantly species of viral DNA is the latent nonintegrated HIV-1 (44). This step of accumulating nonintegrated DNA was found to be a normal step in the HIV-1 life cycle (290) and remained stable for at least 30 days in non-dividing, metabolically active macrophages (342). In a study by the Levy group, they found that the latent nonintegrated DNA lingered for several weeks and that the viral DNA could be activated to produce virions (343).

**Neurological complications**

Multinucleated giant cells in the brain reflect a productive HIV-1 infection in this part of the patient’s body (344). Multinucleated giant cells, unintegrated HIV-1 DNA, viral proteins, HIV-1 associated dementia, and cerebral atrophy were found to be interconnected (42). Toggas et al. revealed that the extracellular protein envelope glycopolypeptide gp120 in the brain of transgenic mice was capable of leading to the same neurological complications as in patients with an HIV-1 infection (345).

One of the main side effects of an HIV-1 infection is neurological diseases. Post mortem studies measured that 79 to 94% of AIDS patient’s brains are neuropathological abnormal (346,347). HIV-1 infects microglia, macrophages, and astrocytes in the central nervous system. Microglia are a type of glial cells that are called the resident
macrophages of the brain and spinal cord. They are distributed in all brain regions at different consistencies. In some parts of the brain, the microglia make up 16% of the total cell population (348). Microglia express the receptor CD4 and co-receptors CCR5 and CCR3 and these are utilized by the virus to invade the cells (349). Astrocytes constitute over 50% of the total brain cells and play an important role in supporting the blood-brain barrier, nourish nervous tissue, maintain extra cellular ion balance, and have a role in repairing brain and spinal cord after injury (350,351). Astrocytes are a natural host for HIV-1, particularly in progressive brain disorders. However, they lack all the conventional cell receptors used by HIV-1 to enter a host cell: CD4, chemokine coreceptors CCR5, and CXCR4 (349,352). It was found that HIV’s gp120 binds to the human mannose receptor (hMR) and that this hMR-mediated HIV infection was dependent on viral entry by endocytosis (349).

Nowadays HAART has increased the lifespan of HIV-1 infected individuals, however, HIV-1 associated neurocognitive disorders (HANDs) are still a common concern among the HIV-1 seropositive community. HANDs cover a range of afflictions that are involved with the nervous system (353). These disorders include asymptomatic neurocognitive impairment (ANI), mild neurocognitive disorder (MND), and HIV-associated dementia (HAD) (354). Over 30% of AIDS patients experience some form of the neurological deterioration that is caused by HIV-1 (46). With the use of HAART, HAD incidences decreased, while the instances of ANI and MND have remained unchanged. However, the milder forms of HANDs are now occurring in individuals with less progressed immunosuppression (353). Viral protein Tat was found to be involved in
the development of the neurological disorders. Elevated tat mRNA was detected in brain tissue from patients that suffered HAD (355). Tat is known to be a neurotoxin and is secreted by infected cells as a soluble neurotoxin and is taken up by uninfected cells by endocytosis (356). It was established that Tat is a powerful inducer of monocyte chemotactic protein-1 (MCP-1), II-8, and Interferon gamma-induced protein 10 (IP-10) (357). II-8 is a chemokine that attracts neutrophils, basophils and T cells (358). IP-10 (also known as CXCL10) is a neurotoxic chemokine and attracts monocytes and T cells (359,360).

Upon HIV-1 infection of macrophages, microglia, and astrocytes in the brain, the cells become activated and secrete neurotoxins that damage the neurons which are not directly a target of an HIV-1 infection. The activated cells also cause an extended inflammation in the CNS. Viral protein Tat secreted from infected cells also contributes to the inflammation by activating non-infected cells, including microglia, astrocytes, and neurons. The activated and infected microglia and astrocytes generate cytokines such as TNF-α and II-1β, which in turn also promotes activation of bordering cells. The activated and infected microglia and astrocytes also produce chemokines such as MCP-1. The chemokines attract more cells which consequentially results in more inflammation. This results in neuronal damage and cognitive impairment (361–364).
CHAPTER ONE

The Role of Virion-Associated Protein Tat in Preintegration Transcription

Introduction

After integration of the HIV-1 provirus into the genome of the host cell, transcription is regulated by components of the host cell as well as by factors produced by the virus itself. The integrated viral genome consists of two LTRs from which the 5’ LTR functions as an RNA Pol. II promoter and the 3’ LTR assists in transcription termination and polyadenylation. Transcription initiation from the viral promoter complies with the eukaryotic transcriptional pathway. This basal transcription process relies on the kinase CDK7, present in the initiation factor TFIIH, to phosphorylate the CTD of the RNA Pol II which results in transcription initiation (65,162,163). After the early regulatory viral protein Tat is produced in sufficient quantities, the protein is involved in a positive feedback mechanism to increase the production of viral transcripts through its transactivation capability. Tat binds directly to the U-rich bulge of the RNA stem-loop structure of the TAR (177,178,365) and recruits the cellular cofactor P-TEFb. Following the phosphorylation of the CTD by TFIIH (CDK7) (161), the clearance of the promoter, and the transcription of a functional TAR, P-TEFb additionally phosphorylates the CTD of RNA Pol II (183,197,366). Tat stabilizes the RNA polymerase, and without
Tat, the polymerase will either pause or fall off the template just past TAR transcription (65,171).

Nonintegrated HIV-1 DNA is not recognized as an important factor in a productive viral life cycle. Nevertheless, it was found that the nonintegrated viral DNA is transcriptionally active in cultured cell lines (367). Wu and Marsh detected an increase in II-2, a cell activation marker, in stimulated, HIV-1 infected resting T cells compared to uninfected cells (288). They determined that this increase in II-2 production was caused by components downstream of the reverse transcription process of the infection. Full-length, nonintegrated viral DNA and transcripts of the early HIV gene \textit{nef}, and (in lower amounts) of the viral gene \textit{tat}, were detected in HIV-1 infected resting T cells. It was found that the Nef protein produced from nonintegrated viral template was the cause of the II-2 increase and that this early protein had a positive impact on viral production after stimulation. Even though \textit{tat} transcripts were observed in a resting T cell infection, the viral Tat protein could not be detected by Western analysis (288). However, this does not mean that protein Tat is not produced from nonintegrated viral DNA. In a later study by Wu and Marsh, they demonstrated that the double stranded, nonintegrated HIV DNA was fully capable of functioning as a template for the expression of early and late gene transcripts. They concluded that early transcription from nonintegrated HIV DNA is a natural part of the HIV-1 life cycle (290).

Preliminary data by Dr. Wu’s laboratory showed that an IN- HIV-1 infection projects a homologous transcription activity compared to a heterogeneous transcription activity in an IN+ virus infection. From this preliminary study we hypothesized that
transcription from the nonintegrated viral DNA template is regulated by different factors than the transcription from the integrated provirus. In this presented study, we investigated whether virion-associated Tat protein has a role in preintegration transcription. Wu and Marsh demonstrated that the transcription from nonintegrated DNA is a normal phenomenon which takes place during the early step of the HIV-1 life cycle (290). This was the rationale behind further investigating the early steps of the viral preintegrated transcription. Chertova et al. showed that Tat is incorporated in the virion in a proteomic and biochemical analysis of purified viruses from infected monocyte-derived macrophages (368). It has also been demonstrated that in a viral infection with a Tat-mutant variant, efficient reverse transcription of the genomic RNA did not occur (266). For Tat to be able to aid in a very early step of the infection, before transcription takes place, the protein Tat needs to be assembled into the virion. Tat being present early in the HIV-1 infection implies that the protein assists in the early transcription of nonintegrated viral DNA.
Materials and Methods

Plasmids

The packaging construct (also known as helper plasmid) pCMVΔR8.2(D116N)Δtat was derived from the plasmid pCMVΔR8.2(D116N). The plasmid pCMVΔR8.2(D116N) consists of the human cytomegalovirus (hCMV) immediate early promoter instead of the viral 5’ LTR promoter. The packaging signal was deleted, but the 5’ splice donor site was kept intact. The viral 3’ LTR was replaced with a poly A site and the envelope and accessory protein Vpu were made defective (369,370). A mutation in the integrase catalytic domain (Asp 116 to Asn) inhibits the 3’ processing step of the integration process (277).

The added tat mutation in plasmid pCMVΔR8.2(D116N)Δtat was made in the first exon of the gene. A polymerase chain reaction (PCR) product, which included the first exon of tat, was digested with MunI, the ends repaired with klenow fragment of DNA polymerase I and self-ligated. The blunt end ligation introduced a stop codon (TAA) after the 24th amino acid of the Tat protein. The DNA fragment with the tat mutation was re-inserted back into plasmid pCMVΔR8.2(D116N) to generate the plasmid pCMVΔR8.2(D116N)Δtat. A complete explanation of the cloning procedure is described in section “Cloning pCMVΔR8.2(D116N)Δtat.”

The plasmid pCMVΔR8.2Δtat was generated from the two plasmids pCMVΔR8.2(D116N)Δtat and pCMVΔR8.2. Plasmid pCMVΔR8.2 has been described previously (369). The plasmid pCMVΔR8.2Δtat is the counterpart of plasmid pCMVΔR8.2(D116N)Δtat, but with a wild-type integrase gene. A complete explanation of the cloning procedure is described in detail in section “Cloning pCMVΔR8.2Δtat.”
The Rev-dependent lentiviral plasmid pNL-Luc-RRE-SA has been described previously (371). The Rev-independent reporter construct pNL-RRE-SA-Luc was derived from lentiviral pNL-RRE-SA vector (372) by inserting the luciferase gene within the XhoI cloning site. This plasmid consists of a packaging signal and a retroviral LTR promoter. A complete explanation of the cloning procedure is described in section “Cloning pNL-RRE-SA-Luc.”

Plasmid pNL ΔTAR-RRE-SA-Luc was generated from the luciferase reporter plasmid pNL-RRE-SA-Luc that is described earlier. The TAR mutant sequence: 5’-TGA GCC TGG GCT CTG GCT AA-3’ consists of a 4-base deletion (AGCT) in the TAR sequence (+34 through +37 relative to the transcriptional start site) as shown in Figure 1.1. The deleted 4-base region of the TAR sequence is part of the apical loop of the bulge-loop structure that is produced by the TAR RNA (373). The result of the 4-base deletion is a disruption in the apex of the stem-loop structure and this abolishes the LTR response to the HIV-1 Tat protein (374). The TAR mutation needed to be incorporated in the 5’ LTR as well as in the 3’ LTR due to the reverse transcription process of the viral genomic RNA. The reverse transcription step of the plus-strand RNA genome starts at the 5’ end. First, the annealing of the host tRNA_{3}^{Lys} at the primer binding site (PBS) takes place. The minus-strand DNA synthesis continues through the U5 region and ends at the edge of the R region at the 5’ end. From here the first jump will occur to the 3’ end from either the second or the same RNA genome. The minus-strand DNA R region will
Figure 1.1

TAR sequence with 4-base deletion (ΔTAR) incorporated into the plasmid pNLΔTAR-RRE-SA-Luc. The plasmid pNLΔTAR-RRE-SA-Luc was constructed from a plasmid expressing the luciferase-encoding gene under the control of the HIV LTR promoter (pNL-RRE-SA-Luc). The generated ΔTAR plasmid includes a 4-base deletion (+34 through +37 relative to the transcription start site) in both the 3’ and 5’ TAR sequence.
hybridize with the 3’ R region of the plus-strand RNA genome (45). Figure 1.2 shows the reverse transcription process. Due to the hybridization of the R regions, it is necessary to include the TAR deletion in both the 5’ LTR and the 3’ LTR to prevent a reverse mutation. A complete explanation of the cloning procedure is described in section “Cloning pNLΔTAR-RRE-SA-Luc.”

The plasmid pHCMV-G that expresses the vesicular stomatitis virus (VSV) glycoprotein has been described previously (375). This plasmid provides the envelope and VSV glycoprotein envelope pseudotyped viruses enter cells through an endocytic pathway (376).

The plasmids pUC18 (# 587), pTat72 (# 597), pTat-pro6-10 (# 510), pTat-C27S (# 454), pTat-K41A (pcDNA3.1/Tat101-K41A), pTat101 (pcDNA3.1/Tat101), and pcDNA3.1+ were provided by Dr. David Harrich of the HIV Molecular Virology Laboratory of the Queensland Institute of Medical Research in Herston, Australia.

**Cloning pCMVΔR8.2(D116N)Δtat**

The cloning diagram of pCMVΔR8.2(D116N)Δtat is shown in Figure 1.3. The cloning vector pUC19 functioned as an intermediate cloning vector. This vector was digested with restriction endonuclease AflIII and the ends were repaired with Klenow fragment of DNA polymerase I and self-ligated. This action eliminated the AflIII site in pUC19, and this plasmid is called pUC19ΔAflIII (step 1 in Figure 1.3). A single site digestion by endonuclease EcoRI of plasmid pUC19ΔAflIII generated the vector backbone. An EcoRI digestion at multiple digestion sites of plasmid
Figure 1.2

**HIV-1 Reverse Transcription process** (Adapted from: Katz RA, Skalka AM. The retroviral enzymes. Annu. Rev. Biochem. 1994, 63: 133-73). Step 1: The virion-associated tRNA_{lys} primer binds to the primer binding site (PBS) of the viral plus-strand...
RNA. Step 2: The negative-strand DNA initiation and elongation starts through the 5’ U5 and R region by the enzymatic function of the viral reverse transcriptase. The RNA is digested by RNase H, which is part of the reverse transcriptase. Step 3: The minus-strand transfer reaction, also known as the first jump, takes place to the same or the second RNA genome where the R regions of the single strands hybridize. Step 4: After the minus-strand elongation crosses the polypurine-rich region called the polypurine tract (PPT), the RNase H produces a special plus-strand RNA primer. Step 5: The PPT primer is extended by the reverse transcriptase polymerase to generate the plus-strand DNA. Step 6: Plus-strand synthesis continues through the generation of the PBS and subsequently the tRNA is removed by the RNase H. Step 7: The plus-strand transfer, also called the second jump, takes place and the PBS sites on the plus- and minus-strand anneal. Step 8: Both strands are extended and form the complete provirus with the 5’ and 3’ LTR.
pCMVR8.2(D116N) resulted in a DNA fragment that included the tat gene that functioned as the insert (2646 bp long). The vector backbone and insert were ligated by sticky-end ligation, and the generated plasmid is called pUC19ΔAflIII-tat (Step 2).

To introduce the tat mutation in the first exon of the gene, a PCR product was cloned from plasmid pCMVR8.2(D116N) that included the first exon of tat. The primers that were used are F-PCMV8.2(4996 5’ – ACT TAC GGG GAT ACT TGG GC – 3’ and R-PCMV8.2(5405) 5’ – CTA TGG ACC ACA CAA CTA TTG C – 3’ (Invitrogen, Life technologies – Carlsbad, California). This PCR step generated a 431 bp PCR product that was subsequently ligated into the plasmid vector pCR®4-TOPO® and called pCR®4-TOPO®-tat (Step 3). Next, pCR®4-TOPO®-tat was digested with endonuclease MfeI to target the MunI digestion site. Restriction enzyme MfeI is an isoschizomer of enzyme MunI. The ends were repaired with klenow fragment of DNA polymerase I and self-ligated. The blunt end ligation introduced a stop codon (TAA) after the 24th amino acid of the Tat protein. The produced plasmid is named pCR®4-TOPO®Δtat (Step 4).

To remove an extra SalI digestion site from the plasmid pUC19ΔAflIII-tat, the plasmid was digested with two endonucleases; PstI, which resulted in a 3’ overhang, and XhoI, which resulted in a 5’ overhang. The digested products were run on a 1% agarose gel and the largest DNA fragment (4499 bp) was cut out of the gel. The cut-out fragment was purified by gel extraction (Qiagen QIAquick® Gel Extraction kit #28704 – Venlo, The Netherlands). A T4 DNA polymerase reaction (New England BioLabs Inc. M0203S - Ipswich, Massachusetts) removed the 3’ overhang produced by the PstI digestion, and
Figure 1.3

Cloning diagram of pCMVAR8.2(D116N)Atat. Cloning vector pUC19 and pCR®4TOPO® were used as intermediate vectors. Step 1: Digestion of pUC19 with enzyme AflIII resulted in the elimination of the AflIII site in pUC19. Step 2: Plasmid pCMVAR8.2(D116N) was digested with EcoRI and the produced fragment with the tat gene was ligated into pUC19ΔAflIII. Step 3: A PCR product containing the tat gene from
plasmid pCMVΔR8.2(D116N) was ligated into the plasmid vector pCR®4-TOPO®. Step 4: Digestion of plasmid pCR®4-TOPO®-tat with enzyme MfeI to introduce the tat mutation. Step 5: The excess of a SalI site in plasmid pUC19ΔAflIII-tat was removed by digesting the plasmid with PstI and XhoI. Step 6: The wild-type tat gene in plasmid pUC19ΔAflIII-tatΔSal was replaced by the mutant tat gene from plasmid pCR®4-TOPO®Δtat. Step 7: The tat mutant fragment was inserted back to the original plasmid pCMVΔR8.2(D116N) by SalI and BamHI digestion to generate plasmid pCMVΔR8.2(D116N)Δtat.
the 5’ overhang, produced by the XhoI digestion, was filled in. A ligation reaction (New England BioLabs Inc. M0202 - Ipswich, Massachusetts) ligated the blunt ends that produced the plasmid pUC19ΔAflII-tatΔSal (Step 5).

Both plasmids, pCR®4-TOPO®Δtat and pUC19ΔAflIII-tatΔSal, were digested with endonucleases SalI and AflIII to generate a 273 bp insert and a 4226 bp vector backbone respectively. After separation of the DNA fragments by gel electrophoresis and gel extraction, the correct fragments were sticky-end ligated to generate plasmid pUC19ΔAflIIΔtat (Step 6).

The last cloning step is to insert the mutant tat fragment back in the original plasmid pCMVΔR8.2(D116N). Both plasmids, pCMVΔR8.2(D116N) and pUC19ΔAflIIΔtat, were digested with SalI and BamHI, and run on a 1% agarose gel to separate the different DNA fragments. The vector backbone from plasmid pCMVΔR8.2(D116N) and the insert from pUC19ΔAflIIΔtat were sticky-end ligated and this resulted in the generation of the plasmid pCMVΔR8.2(D116N)Δtat (Step 7). Figure 1.3 shows a schematic overview of the cloning procedure.

**Cloning pCMVΔR8.2Δtat**

To construct the plasmid pCMVΔR8.2Δtat, plasmid pCMVΔR8.2(D116N)Δtat was digested with endonucleases SalI and BamHI to generate the insert (1392 bp) that included the tat mutation. The plasmid pCMVΔR8.2 was also digested with endonucleases SalI and BamHI to produce the vector backbone (12075 bp). The backbone and insert were sticky-end ligated to generate the plasmid pCMVΔR8.2Δtat.
**Cloning pNL-RRE-SA-Luc**
For the cloning of plasmid pNL-RRE-SA-Luc, plasmid pNL-Luc-RRE-SA was digested with endonuclease XhoI and the DNA fragment of 1658 bp, which included the luciferase gene, functioned as the insert. The vector backbone was created from digesting the plasmid pNL-RRE-SA with the endonuclease XhoI as well to produce an 8047 bp DNA fragment. The backbone and insert were sticky-end ligated. To check for clones with the correct orientation of the insert, the plasmid DNA was digested with AflIII. Digestion of the AflIII enzyme produced two DNA fragments. The correct orientation of the insert produced the sizes of 5975 and 3732 bp compared to 6637 and 3070 bp for the incorrect orientation. Clones with correct and incorrect orientation of the insert were observed.

**Cloning pNLΔTAR-RRE-SA-Luc**
To construct the plasmid pNLΔTAR-RRE-SA-Luc, a 4-base deletion (AGCT) was made in the TAR sequence to disrupt the apex of the stem-loop structure and to terminate the LTR responsiveness to the HIV-1 Tat protein (374). First, two DNA fragments were synthesized that included either the 5’ TAR or the 3’ TAR mutant sequence by a biotechnology company (Blueheron Biotechnology - Bothell, Washington). The company inserted the synthesized DNA fragments in the vector pUCminusMCS. For the 5’ TAR mutant, a 1422 bp long DNA fragment was synthesized [MutΔTAR(5’LTR1)]. This fragment included in addition to the 4-base TAR deletion, the digestion sites BssHI and PstI. For the 3’ TAR mutant sequence, a 247 bp long DNA fragment [MutΔTAR(3’LTR2)] was synthesized and this sequence included, in addition to the 4-base TAR deletion, the digestion sites PmlI and AflIII.
The next step in this cloning process was designing primers for the pNL-RRE-SA-Luc plasmid that generated a PCR product that included the 3’ TAR region, and the digestion sites KpnI and NcoI. Additionally, the reverse primer was designed with an extra PstI digestion site. The designed primers are: KpnI-U3: 5’ – TGG GTT TTC CAG TCA CAC CTC - 3’ and NcoI-PstI-U3: 5’ - (ATGACTGCAG)TCC TGG GGT AAC ATC ACT ATT GTA - 3 ’ (Invitrogen, Life Technologies – Carlsbad, California) (Step 1 of Figure 1.4).

For the next step, the cloning vector pUC19 and the PCR product were digested with KpnI and PstI endonucleases. The digested PCR product functioned as the insert and the pUC19 served as the vector backbone. Insert and vector backbone were ligated together and the produced plasmid was called pUC19-LTR(KpnI-NcoI) (Step 2).

Next, with the use of endonucleases PmlI and AflII, the plasmid pUC19-LTR(KpnI-NcoI) and the plasmid containing the 3’ TAR sequence with the 4-base deletion [pMutΔTAR(3’LTR2)], were both digested and sticky-end ligated. The ligation caused the replacement of the wild-type 3’ TAR sequence with the mutated 3’ TAR sequence (Step 3). The obtained plasmid, pUC19LTRmutant, was sequenced to confirm the incorporated 4-base deletion in the TAR sequence (Northwoods DNA, Inc. – Solway, Minnesota). After a positive confirmation, plasmid pUC19LTRmutant was then digested with the enzymes KpnI and NcoI to insert the 3’ TAR mutant sequence back into the LTR-driven reporter plasmid pNL-RRE-SA-Luc. The produced plasmid was called pNL-RRE-SA-Luc-3’ΔTAR (Step 4).
Figure 1.4

Cloning diagram of pNLΔTAR-RRE-SA-Luc. Starting plasmid pNL-RRE-SA-Luc consists of a 5’ and 3’ wild-type HIV-1 Long Terminal Repeat (LTR). Both LTRs need to be mutated to prevent the mutation to be converted back to wild-type when the sequences
hybridize during the reverse transcription process. Step 1: Generation of the PCR product that include the 3’ TAR sequence. Step 2: Insertion of the PCR product in the pUC19 cloning vector with endonucleases PstI and KpnI. Step 3: Replacement of the wild-type 3’ TAR sequence in plasmid pUC19-LTR(KpnI-NcoI) by the mutant 3’ TAR sequence of plasmid pMutΔTAR(3’LTR2) with the enzymes PmlI and AflII. Step 4: 3’ TAR mutant insertion in the LTR-driven reporter plasmid pNL-RRE-SA-Luc. Step 5: Insertion of DNA fragment derived from plasmid pNL-RRE-SA-Luc-3’ΔTAR, which includes the 5’ TAR sequence, in the cloning vector pUC19. Step 6: Replacement of the wild-type 5’ TAR sequence in plasmid pUC19-NL-RRE-SA-Luc-3’ΔTAR by the mutant 3’ TAR sequence of plasmid pMutΔTAR(5’LTR1) and the generation of the final plasmid pNLΔTAR-RRE-SA-Luc.
Afterwards, plasmid pNL-RRE-SA-Luc-3’ΔTAR was digested with the restriction enzyme AvrII, and pUC19 was digested with enzyme Xbal. These two endonucleases generate the same overhang (CTAG). The largest DNA fragment produced from the plasmid pNL-RRE-SA-Luc-3’ΔTAR digestion was sticky-end ligated with the single cut plasmid pUC19, generating plasmid pUC19-NL-RRE-SA-Luc-3’ΔTAR (Step 5).

Next, the plasmids pMutΔTAR(5’LTR1) and pUC19-NL-RRE-SA-Luc-3’ΔTAR were digested with the enzymes BssHII and PstI. The digested plasmid pMutΔTAR(5’LTR1) generated the insert that included the 5’ TAR mutant sequence and the digested pUC19-NL-RRE-SA-Luc-3’ΔTAR generated the backbone that included the 3’ TAR mutant sequence. The two fragments were sticky-end ligated to produce the end product; plasmid pNL-RRE-SA-Luc-Δ3’Δ5’TAR, renamed to pNLΔTAR-RRE-SA-Luc (Step 6). The incorporated mutations in the 5’ and 3’ LTR regions were confirmed by sequencing the regions (Northwoods DNA, Inc. – Solway, Minnesota). Figure 1.4 shows a schematic diagram from the cloning process of plasmid pNLΔTAR-RRE-SA-Luc.

**Viruses**

The VSV-G pseudotyped, Tat negative reporter virus vHIVΔtat-Luc(D116N) was constructed by co-transfection of the packaging plasmid pCMVΔR8.2(D116N)Δtat for the necessary viral components, the envelope plasmid pHCMV-G, and the packaged reporter vector plasmid, pNL-RRE-SA-Luc. The pseudotyped Tat positive reporter virus vHIV-Luc(D116N) was generated by co-transfection of the packaging plasmid pCMVΔR8.2(D116N), the envelope plasmid pHCMV-G, and the reporter vector plasmid...
pNL-RRE-SA-Luc. The viruses were produced by co-transfection of HEK293T cells with 10 µg pNL-RRE-SA-Luc (reporter vector plasmid), 7.5 µg of either pCMVΔR8.2(D116N)Δtat or pCMVΔR8.2(D116N) (packaging plasmid), and 2.5 µg pHCMV-G (envelope plasmid) per 10 centimeter diameter petri dish (Table 1-1) using Lipofectamine™ 2000 (Invitrogen, Life Technology - Carlsbad, California). In Figure 1.5, the schematic diagrams of the constructs are presented.

**Table 1-1**
Co-transfection per 10 centimeter diameter petri dish.

<table>
<thead>
<tr>
<th>Name of virus:</th>
<th>pCMVΔR8.2(D116N)Δtat</th>
<th>pCMVΔR8.2(D116N)</th>
<th>pNL-RRE-SA-Luc</th>
<th>pHCMV-G</th>
</tr>
</thead>
<tbody>
<tr>
<td>vHIVΔtat-Luc(D116N)</td>
<td>7.5</td>
<td>10</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>vHIV-Luc(D116N)</td>
<td>7.5</td>
<td>10</td>
<td>2.5</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1.5

Diagram of viral constructs. A schematic diagram of the plasmids that were used to produce the viruses vHIVΔtat-Luc(D116N) and vHIV-Luc(D116N). The packaging plasmids pCMVΔR8.2(D116N)Δtat and pCMVΔR8.2(D116N) lack a packaging signal (ψ), consist of a mutated \textit{envelope} and \textit{integrate} gene, and the Δtat plasmid also includes a mutated \textit{tat} gene. The LTR-driven reporter plasmid pNL-RRE-SA-Luc carries the packaging signal and expresses luciferase. The viruses are pseudotyped with the VSV glycoprotein envelope from the plasmid pHCMV-G.
vHIVΔtat-Luc-pcDNA3.1+ were generated by co-transfection of HEK293T cells with 9 µg pNL-RRE-SA-Luc (reporter vector plasmid), 7.5 µg of either pCMVΔR8.2(D116N)Δtat or pCMVΔR8.2Δtat (packaging plasmid), 2.5 µg pHCMV-G (envelope), and 1 µg pTat72 or pTat-pro6-10, pTat-C27S, pUC18, pTat101, pTatK41A or pcDNA3.1+ plasmids per 10 centimeter diameter petri dish using Lipofectamine™ 2000 (Invitrogen, Life Technology - Carlsbad, California). See Table 1-2.

Table 1-2
Co-transfection per 10 centimeter diameter petri dish.

<table>
<thead>
<tr>
<th>Name of virus:</th>
<th>pCMVΔR8.2(D116N)Δtat</th>
<th>pCMVΔR8.2Δtat</th>
<th>pNL-RRE-SA-Luc</th>
<th>pTat72</th>
<th>pTat-pro6-10</th>
<th>pTat-C27S</th>
<th>pHCMV-G</th>
</tr>
</thead>
<tbody>
<tr>
<td>vHIVΔtat-Luc[D116N]-pTat72</td>
<td>7.5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vHIVΔtat-Luc[D116N]-pTat-pro6-10</td>
<td>7.5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vHIVΔtat-Luc[D116N]-pTat-C27S</td>
<td>7.5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vHIVΔtat-Luc[D116N]-pUC18</td>
<td>7.5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vHIVΔtat-Luc[D116N]-pTat101</td>
<td>7.5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vHIVΔtat-Luc[D116N]-pTat-K41A</td>
<td>7.5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vHIVΔtat-Luc[D116N]-pcDNA3.1+</td>
<td>7.5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vHIVΔtat-Luc-pTat101</td>
<td>7.5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vHIVΔtat-Luc-pTat-pro6-10</td>
<td>7.5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vHIVΔtat-Luc-pTat-C27S</td>
<td>7.5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vHIVΔtat-Luc-pTat-K41A</td>
<td>7.5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vHIVΔtat-Luc-pcDNA3.1+</td>
<td>7.5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The VSV-G pseudotyped virus HIV-Luc was generated by co-transfection of HEK293T cells with 10 µg pNL-RRE-SA-Luc (reporter vector plasmid), 7.5 µg pCMVΔR8.2 (packaging plasmid), and 2.5 µg pHCMV-G (envelope plasmid) per 10 centimeter diameter petri dish using Lipofectamine™ 2000 (Invitrogen, Life Technology - Carlsbad, California).

The VSV-G pseudotyped viruses HIV-LucΔTAR and HIV-LucΔTAR(D116N) were constructed by co-transfection of HEK293T cells with 10 µg pNLΔTAR-RRE-SA-Luc (reporter vector plasmid), 7.5 µg pCMVΔR8.2 or pCMVΔR8.2(D116N) (packaging plasmid), and 2.5 µg pHCMV-G (envelope plasmid) per 10 centimeter diameter petri dish using Lipofectamine™ 2000 (Invitrogen, Life Technology - Carlsbad, California).

See Table 1-3.

Table 1-3
Co-transfection per 10 centimeter diameter petri dish

<table>
<thead>
<tr>
<th>Name of virus:</th>
<th>pCMVΔR8.2(D116N)</th>
<th>pCMVΔR8.2</th>
<th>pNLΔTAR-RRE-SA-Luc</th>
<th>pHCMV-G</th>
</tr>
</thead>
<tbody>
<tr>
<td>vHIV-LucΔTAR</td>
<td></td>
<td></td>
<td>7.5</td>
<td>10</td>
</tr>
<tr>
<td>vHIV-LucΔTAR(D116N)</td>
<td>7.5</td>
<td></td>
<td>10</td>
<td>2.5</td>
</tr>
</tbody>
</table>
**Production of viruses in producer cell line**

Producer cells HEK293T were seeded in a 10 cm petri dish at the amount of 4 X 10^6 cells per dish. The dishes containing the cells were incubated at 37° C and 5% CO₂ overnight until approximately 70 - 80% confluency was reached. The cells were then transfected with a total of 20 µg plasmid DNA per dish. The transfection reagent used was Lipofectamine™ 2000 (Invitrogen, Life technologies – Carlsbad, California) and DMEM without serum. In a sterile tube, 60 µL of Lipofectamine™ 2000 was added to 1.5 mL of serum-free medium per dish and incubated at room temperature for 5 minutes. Plasmid DNA was mixed into 1.5 mL serum-free medium. The mixture of Lipofectamine was slowly added to the DNA mixture and incubated for 20 minutes at room temperature. During the incubation time, the medium from the petri dish containing the HEK293T cells was removed and carefully replaced with 5 mL fresh serum-free DMEM medium. After the incubation time, the Lipofectamine / DNA mixture was carefully added to the petri dish and mixed by swirling the medium around. The cells were incubated at 37° C and 5% CO₂ for 5 - 6 hours. After incubation, the medium was removed and carefully replaced with freshly made DMEM medium supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 mg/mL) followed by an incubation at 37° C and 5% CO₂ for about 48 hours. Post incubation, supernatant of the petri dish containing viral particles was collected in a 15 mL conical and any cells and cell debris were pelleted by centrifugation at 1500 rpm for 5 minutes. The viral supernatant was carefully decanted in a new conical and treated with 2.5 µL of nuclease (100 U/mL Benzonase® Nuclease, Sigma-Aldrich Co. LLC - St. Louis, Missouri) for 15 minutes at 37° C to remove free or virus-associated plasmid DNA contamination. Thereafter, the
viral supernatant was filtered through a 0.45 µm nitrocellulose membrane (Millex-HA, Millipore Corporation – Billerica, Massachusetts) to be sterilized. From the sterile viral solution, 90 µL was added to 10 µL of p24 enzyme-linked immunosorbent assay (ELISA) lysis buffer for measuring the p24 HIV antigen. The viral supernatant was aliquoted in 500 and 1000 µL aliquots and stored at -80°C. Levels of p24 in viral supernatant were measured using the Perkin Elmer Alliance p24 antigen ELISA kit (Perkin Elmer - Waltham, Massachusetts) or an in-house p24 antigen ELISA kit. Plates were kinetically read using an ELx808 automatic microplate reader (Bio-Tek Instruments, Inc. - Winooski, Vermont) at 630 nm.

**In-house p24 antigen ELISA**

The right amount of strips (containing 8 wells each) in a plate were primed with 100 µL of priming buffer that included HIV-1 p24 monoclonal antibody (183-H12-5C – obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH, from Dr. Bruce Chesebro and Kathy Wehrly). This affinity purified antibody had a stock concentration of 5 mg/mL and 8 µL was diluted into 10 mL sterile 1X PBS. The plate was covered with an ELISA plate cover and incubated overnight at 37°C. The next day, the plate was washed 3 X with 1X PBS in a plate washer (Bio-tek EL50 – Winooski, Vermont). The remaining fluid was dumped out with high force and the plate was patted dry on paper towels. Next, 250 µL blocking buffer (1X PBS + 5% FBS) was added to each well and the plate was covered with an ELISA plate cover and incubated for an hour at 37°C. During this incubation time, the samples were thawed and diluted 1:10, 1:10², 1:10³, 1:10⁴ in medium. A standard was prepared from a p24 positive control with a stock
concentration of 200 ng/mL. The standard was diluted to: 1000, 500, 250, 125, 62.5, 31.25, 15.625, and 7.8 pg/mL. After the incubation, the plate was washed 3 X in the plate washer with 1X PBS and 0.05% Tween added. After the wash, the remaining fluid was dumped out and the plate was patted dry on paper towels. Afterward, 100 µL of standard samples were added each to a well of strip #1. Thereafter, 100 µL of the $1 \times 10^2$, $1 \times 10^3$, and $1 \times 10^4$ diluted samples were added to the remaining wells. The location and dilution of each sample was tracked on a plate set-up template. After all the samples were added, the plate was covered with an ELISA plate cover and incubated for two hours at 37° C. After the incubation, the plate was washed 3 X with 1X PBS + 0.05% Tween and the wells semi-dried (like mentioned before). Next, 100 µL of diluted HIV-IG was added to each well. HIV-IG consists of 10 mL dilution buffer (1X PBS + 10% FBS + 0.5% Triton X-100) and added 0.067% stock biotin-antibodies. The plate was covered with plate sealer and incubated for 1 hour at 37° C and afterwards, the plate was washed 3 X with 1X PBS + 0.05% Tween and the wells semi-dried. Streptavidin-Horseradish peroxidase (HRP)-antibodies, consisting of 1X PBS + 0.1% BSA + 0.67% stock antibodies, was added in a volume of 100 µL to each well. The plate was covered with a plate sealer and incubated for 1 hour at 37° C. After the incubation period, the plate was washed 3 X with 1X PBS + 0.05% Tween and semi-dried. Next, 100 µL of peroxidase substrate [10 mL 0.1M Sodium Acetate + 3% of 4 mg/mL tetramethylbenzidine (TMB) in DMSO solution + 0.05% of 30% H₂O₂] was added to each well and immediately placed in the Bio-tek ELx808iu ultra microplate reader (Bio-Tek Instruments, Inc. - Winooski, Vermont) and plate kinetics were read at 630 nm. The result of the standard was used to produce a
standard curve in Excel. This standard curve was used to extrapolate the p24 values of the samples.

**Tissue Culture Media**

The medium RPMI 1640 supplemented with 10% heat-inactivated Fetal Bovine Serum (FBS) was prepared from 500 mL RPMI 1640 (Cellgro, Media Tech Inc. – Manassas, Virginia), 50 mL FBS HI (Gibco, Life Technology - Carlsbad, California), 5 mL of 100X non-essential amino acids (NEAA) (Gibco, Life Technology - Carlsbad, California), 5 mL of 100X Sodium Pyruvate (Gibco, Life Technology - Carlsbad, California), 250 µL Gentamycin (100 mg/mL) (Cellgro, Media Tech Inc. – Manassas, Virginia), and 2 µL of 1000X β-mercaptoethanol (Gibco, Life Technology - Carlsbad, California). All the components, except the PRMI 1640 were mixed and filter sterilized through a 0.22 µM membrane before they were added to the sterile RPMI 1640. The sterile medium was stored at 4ºC.

The medium DMEM, supplemented with 10% heat-inactivated Fetal Bovine Serum (FBS), was prepared from 500 mL DMEM (Cellgro, Media Tech Inc. – Manassas, Virginia), 50 mL FBS HI (Gibco, Life Technology - Carlsbad, California), 5 mL of 100X NEAA (Gibco, Life Technology - Carlsbad, California), 5 mL of 100X Sodium Pyruvate (Gibco, Life Technology - Carlsbad, California), and 5 mL of the combination of Penicillin (50 U/mL) and streptomycin (50 mg/mL). The components were combined and mixed and the sterile medium was stored at 4ºC.
**Cells**

HEK293T cells, a human embryonic kidney cell line, were cultured in DMEM supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 µg/mL) at 37° C and 5% CO₂. CEM-SS cells, a human T lymphoblastoid cell line, were cultured in RPMI 1640 medium supplemented with 10% heat-inactivated FBS at 37° C and 5% CO₂.

**Temperature gradient PCR**

To determine the most efficient annealing temperature for newly designed primers, we used a gradient PCR protocol, utilizing the thermal cycler PTC-200 Peltier (MJ Research - Hercules, California). The final volume of the PCR reaction mixture per reaction was 50 µL and consisted of 1X PCR buffer with 1.5 mM of Mg²⁺, 2.5 µL of dNTP mix (with a stock concentration of 2.5 mM for each dNTP), 1 unit of Supertaq plus polymerase (Ambion - Austin, Texas), 50 pmol of each of the primers used, and 10 ng of plasmid DNA template per reaction. The PCR conditions were set up as following: 94° C for 3 minutes, 30 cycles of: 1) 94° C for 10 seconds, 2) gradient temperature from 48° C – 68° C for 20 seconds, 3) 68° C for 1 minute per 1 kb of template. The final extension step: 68° C for 10 minutes and cooled to 4° C indefinitely. The twelve temperatures that are included in the gradient temperature are: 48, 48.5, 49.5, 51.2, 53.5, 56.4, 59.8, 62.6, 64.8, 66.4, 67.6, and 68° C. The PCR products were run on a 1% agarose gel and stained with ethidium bromide. The gel was imaged with Alpha Innotech Imager (Alpha Innotech - San Leandro, California). From the gel image, the most efficient annealing temperature was selected.
Restriction endonuclease digestion
Digestions were performed in a total volume of 20 µL that consisted of PCR-grade H2O, 1X appropriate NEBuffer, 1X Bovine Serum Albumin (BSA) (100 µg/mL) (if needed, because some restriction endonucleases require BSA), 1 µL of a chosen restriction endonucleases (variable units/µL) (New England BioLabs Inc.- Ipswich, Massachusetts), and 200 – 1000 ng of DNA. The digestion was incubated for about one and a half hours at the appropriate temperature for the used restriction endonucleases. If the digest was used as vector backbone, 1 µL (1 U/µL) calf intestinal alkaline phosphatase (CIP) was used to catalyze the removal of the 5’ phosphate groups from the DNA fragment to prevent re-circularization of the vector. Afterwards, 4 µL of 6 X MassRuler DNA loading dye (Thermo Fisher Scientific Inc. - Waltham, Massachusetts) was added to all the digests to stop the reaction. Digested material was run on a 1% agarose gel with a DNA ladder as a standard. The gel was run at approximately 125 volts for about 45 - 60 minutes and stained with ethidium bromide. The gel was imaged with Alpha Innotech Imager (Alpha Innotech - San Leandro, California).

Ligation
Most DNA fragments that were used for cloning were separated by size in a 1% agarose gel and stained with ethidium bromide. In a very rapid fashion, the correct band size was identified under UV light and cut out with a razor blade. The DNA fragments were isolated from the agarose with a gel extraction kit (QIAquick gel extraction kit – Qiagen 28704 – Venlo, The Netherlands) according to the manufacturer’s protocol. For the ligation, the vector backbone and insert were added in a 1:5 ratio (vector backbone to insert) to the ligation mix in a total volume of 20 µL. The ligation mix consisted of 1 µL
ligase, 2 µL ligase buffer (T4 DNA ligase M0202 – New England Biolabs, Inc.- Ipswich, Massachusetts), and the vector backbone and insert DNA in a 1:5 ratio. If needed, nuclease free water was added up to a total final volume of 20 µL. The microcentrifuge tube containing the ligation mixture was wrapped in parafilm and placed in a 16º C water bath for about 16 hours before ligation reaction was ready to be transformed into bacterial cells.

**Transformation**
For cell transformation, we used either DH5α or Stbl3 competent cells that were generated in our laboratory. The cells were stored at -80º C. For the transformation, aliquots of about 100 µL of competent cells were thawed on ice for 15 minutes. Per transformation, one sterile 15 mL, round-bottom culture tube was also put on ice for 15 minutes. The thawed bacterial cells were carefully added to the sterile 15 mL cooled culture tube and 5 µL of ligation mixture was added and gently mixed. The mixture was rested on ice for 30 minutes. The cells were heat-shocked in a 42º C water bath for 90 seconds and right away placed on ice for 1 – 2 minutes. Next, 800 µL of SOC media (super optimal broth with catabolite repression: 2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose) was added to the tube and placed in a shaking incubator (Analytical Instruments, LLC - Golden Valley, Minnesota) for an hour. For the DH5α cells, the temperature of the incubator was 37º C and a shaking speed of 200 - 250 rpm; and for the Stbl3 cells, the temperature was 30º C with a shaking speed of 180 rpm. Following the incubation, 50, 100, and 250 µL of the cell mixture was evenly spread with sterile glass plating beads over three 1.5% agar
plates containing Lennox Luria-Bertani (LB) media (1% tryptone, 0.5% yeast extract, 0.5% NaCl, pH 7.2-7.4) with ampicillin added (100 µg/mL, Fisher Scientific, Thermo Fisher Scientific – Waltham, Massachusetts). Plates with DH5α cells were incubated overnight (16 hours) at 37°C while plates with Stbl3 cells were incubated overnight (19 hours) at 30°C.

**Screening of clones**

From the agar plates, between 6 and 12 clones were picked up for screening. A 10 µl pipet tip was used for “stabbing” the individual colonies. Right after the “stabbing” the pipet tip was dipped into 5 mL of liquid LB media containing ampicillin (100 µg/mL) contained in a 15 mL, round-bottom culture tube. Depending on the cell type used for transformation, the inoculated colonies were placed in a shaking incubator at either 37°C and 200 - 250 rpm shaking speed for 16 hours (DH5α) or 30°C and 180 rpm shaking speed for 19 hours (Stbl3). After the incubation period, 850 µL of each culture was added to 150 µL of sterile glycerol and frozen at -80°C for storage. Plasmid DNA was extracted from the remaining culture according to the manufacturer’s protocol using the QIAprep Spin Miniprep Kit® (Qiagen - Venlo, The Netherlands). The DNA was eluted in nuclease free water, and 5 µL of the DNA solution was diluted 10 times with nuclease free water to measure DNA concentration at optical density at 260 nm (OD$_{260}$) using Eppendorf BioPhotometer (Eppendorf - Westbury, New York). Next, the clones were analyzed for correctness by digestion of the DNA with carefully chosen endonucleases and checked for the expected sizes of the DNA fragments.
**Infection of cells**

Equal amounts (p24 antigen) of input viruses were used for the infection of cells. The viral particles, stored at -80°C, were quickly thawed and the right amount added to the cells and incubated for 2 - 4 hours at 37°C and 5% CO₂. After the incubation, the cells were washed 1 or 2 times with medium and resuspended in freshly made culture medium and incubated at 37°C and 5% CO₂ for about 48 hours.

**Luciferase assay**

Equal numbers of cells were infected with equivalent amount of p24 input virus containing a luciferase reporter genome. The cells were infected for a total of 2 - 4 hours before the virus was washed away. At 48 hours post infection, either equal numbers of cells or equal volume of cell mixture was collected, spun down at 12,000 rpm for 1 minute, and the pellet resuspended in 1 mL 1X PBS. The cells were again spun down at 12,000 rpm for 1 minute and supernatant was removed without disturbing the pellet. The cells were resuspended in 100 µL of either reporter lysis buffer 1X (Promega, E397A – Madison, Wisconsin) or cell culture lysis reagent 1X (Promega, E153A - Madison, Wisconsin). The luciferase substrate used in the luciferase protocol was the Luciferase Assay Substrate (Promega, E151A - Madison, Wisconsin). To hydrate the lyophilized substrate, 10 mL of Luciferase Assay Buffer (E152A) was added. The hydrated substrate was either used right away or stored in aliquots of 1 mL at -80°C. The substrate is light sensitive and therefore the product was covered at all times in aluminum foil. Before the samples were run on the machine (Promega Glomax Multi Detection System), the PMT was activated and the machine was primed no longer than 10 minutes before the luciferase reading. The priming needs around 900 µL of Luciferase Assay Buffer with
added substrate. For each reading, a sample of “cells only” and a sample of “lysis buffer” were run for background control. The samples were pipetted on a 96-well plate. According to the selected protocol “Luciferase Assay Protocol,” 100 µL of the Luciferase Assay Buffer with added substrate was added to each sample. After reading the plate for luciferase activity, a reverse purge was performed to preserve substrate. The machine was flushed after each use with H₂O, 70% Ethanol, H₂O, and air before shut down. The data was transferred to a USB memory stick and analyzed in Excel.

**Time-course for total, full-length HIV DNA real-time PCR assay**

A total of 1.5 X 10⁶ CEM-SS cells were infected with a total of 135 ng p24 input virus. At 2 hours post infection, cells were washed twice in RPMI 1640 medium and, after last wash, the pellet was resuspended in 1.5 mL RPMI 1640 medium supplemented with 10% heat-inactivated FBS. Additionally, 500 µL of cell mixture was collected, spun down and supernatant removed. The pellet was resuspended in 175 µL lysis buffer (SV Total RNA Isolation Kit, Promega - Madison, Wisconsin) and lysate was stored at -20ºC. The remaining cells were incubated for another 4 hours, for a total of 6 hours post infection. At that point, another 500 µL of cell mixture was collected, spun down, supernatant removed, the pellet resuspended in 175 µL lysis buffer (SV Total RNA Isolation Kit, Promega - Madison, Wisconsin), and the lysate also stored at -20ºC. Genomic DNA was extracted using a SV Total RNA Isolation Kit (Promega - Madison, Wisconsin) as recommended by the manufacturer. Equal cellular DNA (10 ng/µL) was used for quantitative real-time PCR analyses of total, full-length viral DNA carried out by using the Bio-Rad iQ5 real-time PCR detection system. The TaqMan Gene Expression
Master Mix (Applied Biosystems, Life Technology – 4369016 - Carlsbad, California) was utilized in the quantitative real-time PCR with the forward primer, 5’ LTR - U5: 5’- AGA TCC CTC AGA CCC TTT TAG TCA -3’; the reverse primer, 3’ gag: 5’- TTC GCT TTC AAG TCC CTG TTC -3’; and the probe, FAM-U5/gag: 5’- (FAM) – TGT GGA AAA TCT CTA GCA GTG GCG CC – (BHQ) – 3’ (Integrated DNA Technologies, Inc. – Coralville, Iowa) at the final concentration of 300 nM each. The DNA standard used for late viral DNA quantification was constructed by using plasmid pNL4-3. DNA was extracted and linearized by XhoI digestion. For the standard, 10 to 1 X 10^6 copies of pNL4-3 mixed with DNA (10 ng/µL) from uninfected CEM-SS cells were applied.

**Sucrose cushion virus purification**

Viral supernatant was harvested at 48 hours post transfection, centrifuged for 5 minutes at 1200 rpm to remove cellular debris, treated with nuclease (100 U/mL of Benzonase® Nuclease, Sigma-Aldrich Co. LLC - St. Louis, Missouri) for 15 minutes to remove free or virus-associated plasmid DNA contamination, and filtered through a 0.45 µm nitrocellulose membrane (Millex-HA, Millapore Corporation – Billerica, Massachusetts). A 20% sucrose cushion was prepared by mixing 8 gr of sucrose to 40 mL of 1 XPBS. In an open pollyallomer, 12 mL tube (Kendro Laboratory Products, Sorvall, cat # 03699 – Newtown, Connecticut), 5 mL virus supernatant was layered carefully over a 4 mL 20% sucrose cushion (without disturbing the interface) and placed in the tube cases. The tube cases were positioned on the rotor (Sorvall TH-641 swinging bucket titanium) in a balanced fashion. The virions were pelleted through the sucrose cushion at
125,000 g (27,000 rpm) for 3 hours at 4°C. The supernatant and sucrose cushion were removed as much as possible without disturbing the invisible virion pellet. The pellet was resuspended in the remaining liquid and added in a 1:1 ratio to RNA extraction buffer [mixture of β-mercaptoethanol (BME) and 8M Guanidinium thiocyanate (GTC) at a ratio of 1:25].

**Reverse transcriptase assay for evaluation of the packaged viral genome**

The RNA from the viral pellet was extracted by using the SV Total RNA Isolation Kit (Promega - Madison, Wisconsin) as recommended by the manufacturer. For RNA reverse transcription, the M-MLV Reverse Transcriptase (Applied Biosystems Ambion, Life technologies - Grand Island, New York) was used as recommended by the manufacturer. However, either the random decamers (50 µM - # AM5722G, Applied Biosystems Ambion, Life technologies - Grand Island, New York) or the luciferase backward primer (50 µM) was utilized in this step. The sequence of the luciferase backward primer is: 5’-ATT TGC GGC CGC CTC GAG TTA CAC GGC GAT CTT TCC GCC C-3’ (Invitrogen, Life technologies – Carlsbad, California). The used concentration of the RNA template varied per experiment; this was dependent on the concentration of the recovered RNA from the sample. Combined together were 2 µL primers (50 µM) or random decamers and 10 µL of RNA and mixed and added to a thermal cycler at 80°C for 3 minutes and cooled down to 4°C for 5 minutes. Additionally, to each sample was added 2 µL of 10X complete PCR buffer (# 8707G), 4 µL dNTPs (2.5 mM - # 8228G5), 10 U RNase inhibitor (# AM2682), 100 U M-MLV Reverse Transcriptase, and if needed, nuclease-free water up to a total final volume of 20
µL. The mixture was gently mixed, fluid spun down, and incubated at 42° C in the thermal cycler for one hour. Afterwards, the thermal cycler was heated to 92° C for 10 minutes to inactivate the reverse transcriptase. The thermal cycler was cooled thereupon to 4° C and samples (cDNA) were ready for further use.

**Real-time PCR for total, full-length HIV DNA**

Ten µL cDNA was subjected to quantitative (real-time) PCR analyses for total, full-length viral template carried out by using the Bio-Rad iQ5 real-time PCR detection system. The TaqMan Gene Expression Master Mix (Applied Biosystems, Life Technology – 4369016 - Carlsbad, California) was utilized with the forward primer, 5’ LTR - U5: 5’-AGA TCC CTC AGA CCC TTT TAG TCA -3’; the reverse primer, 3’ gag: 5’- TTC GCT TTC AAG TCC CTG TTC -3’; and the probe, FAM-U5/gag: 5’-(FAM) – TGT GGA AAA TCT CTA GCA GTG GCG CC – (BHQ) – 3’ (Integrated DNA Technologies, Inc. – Coralville, Iowa). The reaction mixture per sample consisted of 15 µL 2X Gene Expression Master Mix, 1.5 µL of 20X probe (6 µM), 1.5 µL of 20X forward primers (6 µM), 1.5 µL 20X reverse primers (6 µM), 10 µL of cDNA, and 0.5 µL of PCR grade H2O for a total volume of 30 µL. The reaction mixture was added to BioRad optical tubes (8-tube strips), and inserted into the thermal cycler. The used cycling conditions were for the uracil-DNA glycosylase (UDG) incubation: 2 minutes at 50° C, AmpliTaq Gold ultra pure (UP) enzyme activation: 10 minutes at 95° C, 40 PCR cycles: denature 15 seconds at 95° C and annealing/extending 1 minute at 60° C. After all cycles were completed, the thermal cycler cooled down to 4° C. The DNA standard used for late viral DNA quantification was constructed by using plasmid pNL4-3. DNA was
extracted and linearized by XhoI digestion. For the standard, 10 to 1 X 10^6 copies of pNL4-3 mixed with DNA (10 ng/µL) from uninfected CEM-SS cells were applied. The threshold cycles (C_t) from the thermal cycler were transferred to an Excel spread sheet. A standard curve was generated from the C_t values and the log of the concentration of the DNA standard samples. A trend line was made and the equation of the trend line was used to calculate the copy number of unknown samples by taking the anti-log of the result of the trend line.

**Real-time PCR for positive control RNA**

The positive control RNA, kanamycin, was added to the samples (100 ng total) after lysis of the viral particles in the RNA extraction procedure. The RNA was extracted by using the SV Total RNA Isolation Kit (Promega - Madison, Wisconsin) as recommended by the manufacturer. Random decamers were used in the reverse transcriptase reaction and the recovery of the kanamycin was evaluated in a real-time PCR using SYBR green as a non-specific fluorescent DNA binding dye (SYBR Green ER qPCR SuperMix for iCycler – Invitrogen, Life Technologies cat. #11761-100 – Carlsbad, California). The primers used were the forward primer: KANA-F 5’- GGC TCG CGA TAA TGT CGG G -3’ and the reverse primer: KANA-R 5’- GAT GGT CGG AAG AGG C -3’ (Invitrogen, Life technologies – Carlsbad, California). The reaction mixture per sample consisted of 25 µL of SYBR GreenER qPCR SuperMix, 1 µL of forward primers (10 µM), 1 µL reverse primers (10 µM), 2.5 µL of cDNA, and 20.5 µL of PCR grade H_2O for a total volume of 50 µL. The reaction mixture was added to BioRad optical tubes (8-tube strips), and inserted into the thermal cycler. The used
cycling conditions were for the UDG incubation: 2 minutes at 50º C, UDG inactivation and DNA polymerase activation: 8 minutes and 30 seconds at 95º C, 40 PCR cycles: denature 15 seconds at 95º C and annealing/extending 1 minute at 60º C. For the melting curve analysis: 1 minute at 95º C, 1 minute at 55º C, 80 cycles of 10 seconds at 55º C + 0.5º C / cycle. After all cycles were completed, the thermal cycler cooled down to 4º C. The used DNA standard for relative copy numbers of kanamycin was constructed from 10 ng of kanamycin mRNA per reverse transcriptase reaction using random decamers (50 µM). The obtained cDNA was diluted in a 10-fold serial dilution and by experimental examination the 1/10 through 1/(1 X 10^5) dilutions were used for the generation of a standard curve.

Normalization of data

From the real-time PCR result of the positive control (kanamycin), the lowest value was set to 1 and the other values were given a fraction compared to the lowest value. The obtained real-time PCR data for the HIV copy numbers was first subjected to the subtraction of the DNA contamination of the RNA samples used for the reverse transcriptase reaction and thereafter multiplied by the fractions according to the relative numbers of positive control DNA obtained from the real-time PCR.
Results

Reporter plasmid showed specific activity
To evaluate the newly cloned rev-independent luciferase reporter plasmid, pNL-RRE-SA-Luc, 4 µg of the plasmid was transfected into 5 X 10^5 HEK293T cells and the luciferase activity was read 48 hours after transfection. The average was taken from two read samples. For the negative control, 4 µg of the empty cloning vector pUC19 was transfected into HEK293T cells, and for comparison purposes, 4 µg of the rev-dependent luciferase reporter plasmid, pNL-LUC-RRE-SA, was transfected into cells as well. The HEK293T cells alone were evaluated for “background” by a luciferase reading. Figure 1.6 shows that the newly cloned plasmid, pNL-RRE-SA-Luc, had about 4.5-fold more specific activity compared to the rev-dependent plasmid, pNL-LUC-RRE-SA, and about 4222-fold more activity compared to the negative control, pUC19. We concluded that the newly made, rev-independent luciferase reporter plasmid consisted of specific activity.

Viral nonintegrated transcription is affected by a Tat negative virus
To be able to study transcription from a nonintegrated viral template, a virus was constructed with a nonfunctional integrase protein. Effective replication of retroviruses depends on the integration of the viral double-stranded DNA into the host genome. This integration step is catalyzed by the viral protein integrase (IN) that is encapsidated into the HIV-1 virion as part of the Gag-Pol polyprotein precursor (377). In the virion, the immature integrase protein is processed into a mature protein by the viral protein protease. The mutant integrase D116N consists of a single point mutation causing an amino acid (aa) substitution at aa 116 from Asp to Asn in the integrase catalytic domain.
**Figure 1.6**

**Specific activity of rev-independent luciferase reporter plasmid.** For each transfection into HEK293T cells, 4 µg of plasmid DNA was used or cells were left untreated (293T cells). Luciferase activity was read 48 hours post transfection. The average was taken from two read samples.
It has been shown that this mutation effectively inhibits viral DNA integration, but not disturb functions such as reverse transcription and nuclear targeting (277,378). For this reason, a virus with a D116N mutation is a nonintegrating virus and provides a good approach to study nonintegrated HIV-1 DNA species. From here on, a D116N virus is a nonintegrated virus and consists of a nonfunctional integrase (IN-) compared to an integrated virus, which includes a functional integrase (IN+).

To study the virion-associated protein Tat in the regulation of nonintegrated viral DNA transcription, a total of 5 X 10^5 CEM-SS cells were either infected with 45, 18, or 7 ng of p24 values of the IN- Tat negative virus vHIVΔtat-Luc(D116N) or the IN- Tat positive virus vHIV-Luc(D116N). Luciferase activity of a total of 1 X 10^5 and 5 X 10^5 lysed cells per sample was read at 48 hours post infection. Virus vHIVΔtat-Luc(D116N) was deficient in a virion-associated Tat protein and the reporter vector plasmid pNL-RRE-SA-Luc was not equipped to produce de novo protein Tat either. On the other hand, the virus vHIV-Luc(D116N) contained a virion-associated Tat protein, but de novo Tat was not produced. Displayed in Figure 1.7 is the result of the read luciferase activity, which presents the expression of the reporter gene (luciferase) from a nonintegrated, LTR-driven, template in the absence or presence of virion-associated protein Tat. There is a noticeable decrease in luciferase expression; an 8.1-fold decrease in luciferase activity in 5 X 10^5 cells infected with a Tat negative virus with a p24 value of 45 ng compared to an infection with a Tat positive virus. An 8.4- and 9.0-fold decrease was detected for the p24 values of 18 and 7 ng respectively. These findings suggest that the virion-associated
Figure 1.7

Transcription from nonintegrated template was impacted by the lack of Tat. CEM-SS cells were infected with 45, 18, or 7 ng input virus. After 48 hours of infection, 1 X 10^5 and 5 X 10^5 cells of each sample were lysed and the luciferase activity was read.
protein Tat has an essential function in preintegration transcription. Next we explored if the Tat function was restored in a virus infection where Tat was provided in trans in the virion producer cell line.

**Providing Tat in trans in producer cells restored the defect**

To verify that the difference in reporter gene expression was in fact caused by the absence of Tat, we co-transfected a plasmid carrying a wild-type tat gene, plasmid pTat72, with the packaging plasmid pCMVΔR8.2(D116N)Δtat, the envelope plasmid pHCMV-G, and the reporter vector plasmid, pNL-RRE-SA-Luc to construct virus vHIVΔtat-Luc(D116N)-pTat72. The pTat72 plasmid consists of the 1st exon of the tat gene, approximately 220 bp. As a negative control, we also co-transfected an empty cloning vector, pUC18, with the above mentioned plasmids to construct virus vHIVΔtat-Luc(D116N)-pUC18. For a positive control, the virus vHIV-Luc(D116N) was used, which is constructed from a packaging plasmid with a wild-type tat gene. A total of 5 X 10^5 CEM-SS were either infected with 45, 18, or 7 ng of p24 input virus. Luciferase activity of a total of 1 X 10^5 and 5 X 10^5 lysed cells per sample was read 48 hours post infection.

Figure 1.8-A shows that by providing Tat in trans in producer cells resulted in a 17-fold increase in luciferase activity of 5 X 10^5 lysed cells with a p24 virus input of 45 ng. A 10- and 17-fold increase was detected for the p24 input virus of 18 and 7 ng respectively. In Figure 1.8-B it is demonstrated that the function of Tat was fully restored when adding Tat in trans in producer cells compared to a virus constructed from a
Figure 1.8

Providing Tat in trans in producer cells restored the Tat function. A: CEM-SS cells were infected with 45, 18, or 7 ng p24 of either the IN- Tat-positive virus vHIVΔtat-Luc(D116N)-pTat72, where Tat was delivered in trans, or IN- Tat-negative virus vHIVΔtat-Luc(D116N)-pUC18. Luciferase activity of a total of 1 X 10^5 and 5 X 10^5 lysed cells per sample was read 48 hours post infection. B: CEM-SS cells were infected with 45 ng p24 of either the IN- Tat-positive virus vHIVΔtat-Luc(D116N)-pTat72, vHIV-Luc(D116N), or the IN- Tat-negative virus vHIVΔtat-Luc(D116N)-pUC18. Luciferase activity of a total of 1 X 10^5 and 5 X 10^5 lysed cells per sample was read 48 hours post infection.
packaging plasmid that included a wild-type tat, vHIV-Luc(D116N). However, the vHIVΔtat-Luc(D116N)-pTat72 infected samples showed a higher luciferase expression when compared to HIV-Luc(D116N) infected samples. Most likely, by co-transfection of a wild-type tat plasmid, protein Tat is overproduced and then causes an increase in virion-associated protein Tat which consequently leads to a higher LTR response of the reporter construct compared to virion-associated Tat in the virus vHIV-Luc(D116N) infection. Nonetheless, we concluded from this data that the virion-associated protein Tat plays a crucial role in either the production of the virus or in the transcription of the nonintegrated, LTR-driven, reporter gene. In the following experiment we explored what part of Tat is responsible for the nonintegrated transcription.

**Dissecting the role of virion-associated Tat in nonintegrated transcription**

It was discovered that protein Tat is not only involved in transactivation transcription, but also in the reverse transcription step of the HIV-1 life-cycle. Harrich et al. demonstrated that there was a significant decrease in reverse transcription in infected peripheral blood mononuclear cells (PBMCs) when protein Tat was not included in the HIV-1 virions. It was found that this phenomenon was not caused by a dysfunction of genomic RNA packaging (266). It was observed by Ulich et al. that the Tat mutants C27S and K41A (Figure 1.9) were able to complement HIV-1 reverse transcription in an HIV-1 Δtat infection and restored the defect even while there was a significant decrease in HIV-1 gene expression. The mutant pro6-10 (Figure 1.9) was not able to restore the defect in reverse transcription or in gene expression. Tat’s ability to aid in reverse transcription is
Figure 1.9

Schematic overview of the first exon of the HIV-1 Tat protein (Adapted from: Ulich C, Dunne A, Parry E, Hooker CW, Gaynor RB, Harrich D. Functional Domains of Tat Required for Efficient Human Immunodeficiency Virus Type 1 Reverse Transcription. J Virol. 1999 Mar 1;73(3):2499–508). For the Tat mutants, the changes in the amino acids are shown boxed below the native amino acid (aa) sequence. The solid line indicates multiple mutations in one plasmid (218). Also the five domains of a 72-aa Tat are shown: aa 1-20 represents the N-terminus, aa 21-40 the cysteine-rich domain, aa 41-48 the core domain, aa 49-57 is the basic domain, and aa 58-72 is the C-terminus.
mostly dependent on a complete N-terminus. The cysteine and core domains of the Tat protein are required for its ability to mediate HIV-1 transcription, but not for the activation of the reverse transcription (218). Figure 1.9 shows a schematic overview of the first exon of the HIV-1 Tat protein with the various mutation sites indicated (218).

To explore the function of virion-associated Tat in nonintegrated viral DNA transcription, we utilized the three different tat mutant plasmids: pTat-C27S, pTat-K41A, and pTat-pro6-10. HEK293T cells were co-transfected with a mutant or wild-type tat plasmid or an empty cloning vector together with the Δtat packaging plasmid, luciferase reporter plasmid and envelope plasmid as previously described. The mutants C27S and pro6-10 and the wild-type tat gene (tat72) were cloned in the vector pUC18, and therefore the empty cloning vector pUC18 was used as negative control. The mutant K41A and the wild-type tat gene (tat101) were cloned in the vector pcDNA3.1+, and for that reason the empty cloning vector pcDNA3.1+ was also used as negative control.

For this experiment, 45, 18, or 7 ng of input virus was utilized to infect 1 X 10⁵ CEM-SS cells. After 48 hours of infection, 1 X 10⁵ and 5 X 10⁵ cells per sample were lysed and analyzed for luciferase activity. The result is displayed in Figure 1.10. Figure 1.10 A shows that the IN- Tat mutant virus HIVΔtat-Luc(D116N)-pTat-pro6-10 and IN-Tat mutant virus HIVΔtat-Luc(D116N)-pTat-C27S were not capable of restoring the wild-type Tat function in nonintegrated reporter gene expression as seen by virus HIVΔtat-Luc(D116N)-pTat. Figure 1.10 B displays that the IN- Tat mutant virus HIVΔtat-Luc(D116N)-pTat-K41A was also deficient in restoring the wild-type Tat
Figure 1.10

Dissecting Tat function. A: CEM-SS cells were infected with 45, 18, or 7 ng p24 of IN- Tat mutant virus vHIVtat-Luc(D116N)-pTat72, vHIVtat-Luc(D116N)-pTat-pro6-10, vHIVΔtat-Luc(D116N)-pTat-C27S or IN- wild type Tat virus vHIVΔtat-Luc(D116N)-pTat72 or IN- Tat-negative virus vHIVΔtat-Luc(D116N)-pUC18. Cells were lysed 48 hours post infection and lysate of 1 X 10^5 and 5 X 10^5 cells per sample were analyzed for luciferase activity. B: CEM-SS cells were infected with 45, 18, or 7 ng p24 of IN- Tat mutant virus vHIVΔtat-Luc(D116N)-pTat-K41A or IN- wild type Tat virus vHIVΔtat-Luc(D116N)-pTat101 or IN- Tat-negative virus vHIVΔtat-Luc(D116N)-pcDNA3.1+. Cells were lysed 48 hours post infection and lysate of 1 X 10^5 and 5 X 10^5 cells per sample were analyzed for luciferase activity. An average was taken from two separately ran samples.
function in nonintegrated transcription. Here the result is shown from the average of
samples run in duplicate on the plate for luciferase reading.

None of the Tat mutants were capable of restoring Tat’s authentic function in
nonintegrated transcription (Figure 1.10). This result indicates that the cysteine-rich and
core domains of Tat have a major function in transcription from the nonintegrated
reporter gene template. An interesting observation is that Tat mutant pro6-10 is capable
of producing 3.5 times more luciferase activity compared to mutant C27S. Mutant pro6-
10 consists of a defect in its reverse transcription function and as a result, less double-
stranded viral DNA is generated, which consequently produced a decrease in luciferase
activity. This mutant is not capable of restoring the defect in reverse transcription and
gene transcription according to Ulich et al.(218). Mutant C27S (in cysteine-rich domain)
and K41A (core domain) both have a functional reverse transcription function, but have a
defect in its transactivation function. We expected Tat mutant pro6-10 to produce lower
reporter expression due to its defect in its reverse transcription function. The plasmids
pro6-10, C27S, K41A, pTat72 and pTat101 were sequenced to confirm that the plasmids
contained the correct sequences (Northwoods DNA, Inc. Solway, MN). To confirm that
indeed Tat has a major function in transactivating the HIV-1 promoter of the
nonintegrated DNA, we decided to investigate the reverse transcription competency of
the different mutants.

**Disproportional DNA synthesis during the reverse transcription process**

To determine if the lack of restoration in reporter gene transcription by the Tat
mutants was caused by a defect in the reverse transcription process, we investigated if
there was equal production of double-stranded viral DNA by this process. In this experiment, a total of $1.5 \times 10^6$ CEM-SS cells were infected with $3 \times 45$ ng p24 IN- Tat-positive viruses $\text{vHIV}^{\Delta\text{tat- Luc(D116N)-pTat72}}$ (Figure 1.11 A) and $\text{vHIV}^{\Delta\text{tat- Luc(D116N)-pTat101}}$ (Figure 1.11 B) and the Tat IN- mutants viruses $\text{vHIV}^{\Delta\text{tat- Luc(D116N)-pTat-pro6-10}}$, $\text{vHIV}^{\Delta\text{tat- Luc(D116N)-pTat-C27S}}$ (A), and $\text{vHIV}^{\Delta\text{tat- Luc(D116N)-pTat-K41A}}$ (B), and Tat-negative viruses $\text{vHIV}^{\Delta\text{tat- Luc(D116N)-pUC18}}$ (A) and $\text{vHIV}^{\Delta\text{tat- Luc(D116N)-pcDNA3.1+}}$ (B). Two hours after infection, cells were washed to remove the input virus and thereafter the 2-hours post infection samples were collected. Cells were lysed in 175 µL of lysis buffer for DNA/RNA extraction and stored in the -20°C freezer. The same was accomplished after 6 hours of infection to collect the 6-hours post infection samples. From the 2- and 6-hours post infection samples, DNA was extracted and 100 ng of total cellular DNA was used of each sample for real-time PCR. Total full-length HIV DNA was targeted by this amplification method.

The result shows that the occurrence of DNA synthesis during reverse transcription was not equal (Figure 1.11), and that this result is parallel with the reporter gene expression (Figure 1.10). The Tat mutant samples (pro6-10, C27S, and K41A) were impaired in producing the same amount of full-length HIV DNA as compared to the wild-type Tat samples. This result did not confirm the transactivating function of Tat in nonintegrated gene expression, but indicates that the Tat protein has an activity upstream of the reverse transcription process that caused the defect in the nonintegrated DNA gene expression in the $\Delta\text{tat}$ infection. The next step was to test if we observed the same occurrence in an infection with a functional integrase (IN+).
Figure 1.11

Unequal DNA synthesis during reverse transcription in Tat mutant viral infections.

CEM-SS cells were infected with 135 ng p24 input virus and samples were taken 2 and 6 hours post infection. DNA was extracted and 100 ng of total cellular DNA was subjected to real-time PCR targeting total, full-length viral DNA.
Defect in transactivation of integrated viral template by Tat mutants

Next, we wished to examine if protein Tat and the Tat mutants, pro6-10, C27S, and K41A, had the same impact on an infection with a functional integrase (IN+) compared to an infection with a non-functional integrase (IN-) which we investigated previously. It is known from numerous studies that the cysteine and core domain are fundamental for transactivation to occur at the HIV-1 LTR promoter of the integrated provirus (218,379–381). It should be no surprise if the mutants are not able to restore the activity of wild-type Tat. The different IN+ viruses that were constructed were vHIVΔtat-Luc-pTat101, vHIVΔtat-Luc-pTat-pro6-10, vHIVΔtat-Luc-pTat-K41A, vHIVΔtat-Luc-pTat-C27S, and vHIVΔtat-Luc-pcDNA3.1+. For producing the necessary viral components, including the integrase, the packaging plasmid pCMVΔR8.2 Δtat was used instead of the packaging plasmid pCMVΔR8.2(D116N)Δtat used in prior experiments. For this experiment, 5 X 10^5 CEM-SS cells were infected with 7, 18, or 45 ng of p24 input virus. The virus was washed away two hours post infection. Two days post infection, the cells were counted and 1 X 10^5 and 5 X 10^5 CEM-SS cells per sample were lysed and analyzed for measuring the luciferase activity.

As expected, all three Tat mutants were unable to rescue the function of wild-type Tat (Figure 1.12). Here we also noticed the higher luciferase activity from the pro6-10 mutant Tat infection compared to the C27S and K41A mutants. Although the mutants were not able to restore the level of reporter gene expression, we expected to observe comparable viral DNA synthesis during reverse transcription in the Tat mutants C27S and K41A compared to wild-type Tat virus infection. The Tat mutant pro6-10 infection was expected to produce the lowest amount of double stranded viral DNA due to its
**Figure 1.12**

Tat mutants are not able to restore wild-type Tat function in transactivation of transcription from an integrated viral template. CEM-SS cells were infected with 45, 18, or 7 ng p24 of IN+ Tat mutant viruses vHIVΔtat-Luc-pTat-pro6-10, vHIVΔtat-Luc-pTat-C27S or vHIVΔtat-Luc-pTat-K41A and IN+ wild-type Tat virus vHIVΔtat-Luc-pTat101 or IN+ Tat-negative virus vHIVΔtat-Luc-pcDNA3.1+. Cells were lysed 48 hours post infection and lysate of 1 X 10⁵ and 5 X 10⁵ cells per sample were analyzed for luciferase activity.
mutation in the N-terminal domain. The next task was to check for similar synthesis of viral DNA by the reverse transcription process in an infection with an intact integrase.

**Unequal viral DNA synthesis during the reverse transcription process in an infection with a functional integrase.**

Even though a decrease in reporter gene expression was expected in the IN+ Tat mutant C27S and K41A infections, we did expect to see equal DNA synthesis by the reverse transcription process. In this experiment, a total of $1.5 \times 10^6$ CEM-SS cells were infected with $3 \times 45$ ng p24 of IN+ Tat mutant viruses vHIVtat-Luc-pTat-pro6-10, vHIVtat-Luc-pTat-C27S or vHIVtat-Luc-pTat-K41A and IN+ wild-type Tat virus vHIVtat-Luc-pTat101 or IN+ Tat-negative virus vHIVtat-Luc-pcDNA3.1+. Two hours after infection, cells were washed to remove the input virus and thereafter the 2-hours post infection samples were collected. Cells were lysed in 175 µL of lysis buffer for DNA/RNA extraction and stored in the -20°C freezer. The same was accomplished after 6 hours of infection to collect the 6-hours post infection samples. From the 2- and 6-hours post infection samples, DNA was extracted and 100 ng of total cellular DNA from each sample was used for real-time PCR, which targeted the total, full-length viral template.

The result displayed in Figure 1.13 shows that there is unequal production of viral DNA observed in the 6-hours post infection samples. The reduction in production of double-stranded viral DNA appeared to be the same pattern as in the reporter gene expression presented in Figure 1.12. This data and previous results of the IN- infections, indicate that the wild-type Tat protein (Tat72 and Tat101) has a major function upstream of the reverse transcription process, which is independent of the presence of a functional
Amount of DNA synthesis during reverse transcription differ between wild-type Tat and Tat mutants at 6 hours post infection. CEM-SS cells were infected with 135 ng p25 input virus and samples were taken 2 and 6 hours post infection. DNA was extracted and 100 ng of total cellular DNA was subjected to real-time PCR targeting total, full-length viral DNA.
integrase protein. Subsequently, we wished to determine if the defect in reverse transcription is due to a reduction in packaging of the viral genome (luciferase reporter genome) in virion particles in the absence of Tat in the virion producer cells.

**The lack of Tat in producer cells leads to a reduction of encapsidation of genomic RNA in virions**

An event that occurs upstream of the reverse transcription process is the production of virions in the HEK293T producer cells. We next investigated the amount of viral RNA that was packaged in virions produced by the producer cell line with or without the presence of protein Tat. It is possible that wild-type Tat must be present in the producer cells for efficient encapsidation of genomic RNA. Either viral protein Tat is necessary in the process of dimerization of the viral genome or the protein is essential in the production of the viral RNA genome in the producer cell. There is no evidence in the literature that viral protein Tat is needed in the dimerization process of the genomic RNA. On the other hand, it is known that protein Tat is effective in transactivation of transcription from a viral LTR promoter. If the production of the viral genomic RNA (LTR driven) and viral components (CMV promoter driven) in the producer cells are not in a specific ratio, perhaps this could compromise the packaging of viral RNA in the produced virions.

To evaluate viral RNA packaged in virions, two different viruses were assembled in the HEK293T producer cells: an IN-positive-Tat virus, vHIV-Luc(D116N), and an IN-negative-Tat virus, vHIVΔtat-Luc(D116N). Forty-eight hours post transfection, viral supernatant was collected and viruses were pelleted by the use of a sucrose solution cushion (3 tubes per virus). After the centrifugation step, the virus pellet was resuspended
in approximately 200 µL 1X PBS. The viral mixtures (3 tubes per virus) of the two viruses were pooled together and a total of 182 µL was added to 182 µL RNA extraction buffer and viral RNA extraction was performed. For a more detailed description of the procedure, consult the section “Materials and Methods.”

For the reverse transcription step of the viral RNA, two different methods were used. In the first method random decamers were utilized as primers. Random decamers are oligodeoxyribonucleotides with random sequences that anneal to random complimentary sites on the RNA and create a pool of diverse cDNAs (Figure 1.14 A). In the second method, a luciferase primer was used to create a more homogeneous pool of cDNA (Figure 1.14 B). Afterwards, both cDNA samples were subjected to real-time PCR targeting the total, full-length HIV-1 genome (Figure 1.14). To normalize the two RNA samples, equal total RNA (96 ng) was used in the reverse transcription assay. For the real-time PCR, 10 uL cDNA was applied. The result shown in Figure 1.15 displays a noticeable difference in the HIV DNA copies that represent the packaged viral genomic RNA. The result exhibits a decrease of packaged viral genome during the virion production in the absence of Tat. There is a 21-fold decrease in total HIV DNA copies in the Tat negative sample when the random decamers were used and a 12-fold decrease when the luciferase backward primer was utilized in the reverse transcription assay. This result indicates that there was no comparable encapsidation of HIV-1 genomic RNA species when protein Tat was omitted.
Figure 1.14

**Schematic diagram of the two reverse transcription methods.** A: Random decamers were used as primers in the reverse transcription that created a pool of various cDNAs that were subjected to real-time PCR for analyzing total, full-length viral DNA. B: The backward luciferase primer was used for reverse transcription that created a homogeneous species of cDNA that was subjected to real-time PCR targeting the total, full-length viral DNA for amplification.
Figure 1.15

Packaging of viral RNAs in virions was disproportionate in the presence or absence of Tat in producer cells. Viral supernatant was pelleted through a 20% sucrose cushion and total viral RNA was extracted. Equal quantity of total viral RNA (96 ng) per virus was used for the reverse transcriptase assay; either with random decamers or with luciferase backward primer. This procedure was followed by real-time PCR targeting total, full-length viral DNA for amplification.
This experiment was repeated with the use of kanamycin RNA as a control for the efficiency of the RNA extraction. Also, after the ultracentrifugation step, the pellet was resuspended in the remaining 1X PBS and a small amount of sample (18 µL into 2 µL of p24 lysis buffer) was collected for p24 ELISA. In this experiment we used equal amount of p24 value of each virus (~190 ng) for RNA extraction. After lysis and dilution buffer was added to the samples, 100 ng of kanamycin RNA was added and the total RNA was extracted. A total of 10 µL of the extracted RNA was used in the reverse transcription process and either random decamers or luciferase backward primer was utilized in this step. In the real-time PCR targeting the kanamycin, 2.5 µL of cDNA was used that originated from RNA reverse transcribed with random decamers. Ten µL of cDNA from both the random decamers and the luciferase primer was used in the analysis of total, full-length viral DNA copies by real-time PCR. The RNA samples (with no added reverse transcriptase) were subjected to the real-time PCR for total, full-length viral DNA to check for DNA contamination. The outcome of the kanamycin real-time PCR was used to perform normalization of the data. Figure 1.16 shows the result for the RNA packaged in the virion. A decrease of 10-fold was observed in the IN- Tat negative samples in both experiments where either the random decamers or the luciferase backward primer was used. Due to the complexity of the procedure, we wished to replicate the previous data.

The experiment was repeated with a total of 16.5 ng p24 value for each virus and an average and standard deviation was taken from 6 samples. Figure 1.17 shows the result with a 17-fold decrease in total HIV DNA copies in the IN- Tat negative, random
Figure 1.16

Unequal RNA genome encapsidation during the production of virions in the presence and absence of Tat. Viral supernatant was pelleted through a 20% sucrose cushion and from resuspended viral pellet, comparable p24 values (190 ng) per virus were subjected to RNA extraction. cDNA was produced in the reverse transcriptase assay with either the utilization of random decamers or a luciferase backward primer. Total, full-length HIV genome was targeted and amplified in the real-time PCR. Kanamycin control RNA was used for normalization.
During production of virions in the presence or absence of Tat, encapsidation of RNA genome was disproportional. Viral supernatant was pelleted through a 20% sucrose cushion and from resuspended viral pellet, comparable p24 values (16.5 ng) per virus were subjected to RNA extraction. cDNA was produced in the reverse transcriptase assay with either the utilization of random decamers or a luciferase backward primer. Total, full-length HIV genome was targeted and amplified in the real-time PCR. Kanamycin control RNA was used for normalization. An average and standard deviation was taken from 6 samples.
decamers employing samples and an 8.5-fold decrease in total HIV DNA copies in the IN- Tat negative, luciferase primer employing samples.

From the data above we concluded that the viral protein Tat was needed for efficient production of virions in producer cells in the system we used. It is known that for genomic RNA to be incorporated into a budding virion, the RNA must contain an intact 5’ untranslated region, the first 200 nucleotides of the viral gene gag, and the Rev Response Element (RRE). Also, HIV-1 RNA genome encapsidation is regulated by packaging two copies of RNA regardless of its size (382).

In a study by Rulli et al., it was found that in the absence of packageable viral RNA, cellular mRNA is packaged nonselectively. In this study, they assumed that viral protein Gag is responsible for either recognizing the packaging signal (ψ) in the viral RNA or interacting with cellular mRNA in a nonspecific manner for incorporation into the nascent virion. Rulli et al. found that the species of cellular mRNA that were encapsulated was proportional to their level in the cell (383). From these and other studies we know that packaging the viral RNA genome is not a matter of the size of the genomic RNA or if the RNA includes a packaging signal. In our study, it could be possible that in the absence of Tat in the producer cells, less reporter genome mRNA was produced compared to viral proteins. The reporter plasmid was the only plasmid with an existing packaging signal and the only plasmid that was driven by an LTR promoter in this system. This promoter is known to be more efficient with protein Tat present. With less production of viral genomic RNA compared to the other viral components, it is likely
that non-specific, cellular mRNA was packaged into the produced virions instead of two identical copies of the viral genomic RNA.

To determine if the transactivation function of Tat was involved in efficient production of complete virions, we investigated the reporter gene expression, viral DNA synthesis during reverse transcription, and RNA packaging in a TAR mutant infection with a wild-type Tat and with or without the presence of a functional integrase.

**A TAR mutant virus infection also displayed a decrease in reporter gene expression**

First, the synthesized mutant TAR sequence, which included a 4 base deletion (Figure 1.18), was cloned into both the 5’ LTR and 3’ LTR of the reporter gene plasmid, pNL-RRE-SA-Luc. Even though transcription starts at the 5’ LTR, it was important to include the mutant TAR sequence also in the 3’ LTR due to the fact that in the reverse transcription process the two LTRs align and this could cause the mutant form to be corrected to the aligned LTR. Afterwards, four different viruses were constructed. A IN+ Tat-positive TAR mutant virus vHIV-LucΔTAR, a IN- Tat-positive TAR mutant vHIV-LucΔTAR(D116N) (Figure 1.18), a IN+ Tat-positive virus vHIV-Luc, and a IN- Tat-positive virus vHIV-Luc(D116N). A total of 5 X 10^5 CEM-SS cells were infected with 45, 18, or 7 ng of p24 input virus, and both 1 X 10^5 and 5 X 10^5 cells were lysed per sample and analyzed for luciferase activity at 48 hours post infection.

Figure 1.19 shows a noticeable decrease in the measured luciferase expression from the TAR mutant virus infection compared to the wild-type TAR virus infection for both the IN+ and IN- samples. There is an 18-fold decrease in luciferase activity in the IN+/TAR mutant and a 17-fold decrease in the IN-/TAR mutant observed in the 45 ng
Figure 1.18

Construction of the viruses HIV-LucΔTAR and HIV-LucΔTAR(D116N). Reporter plasmid pNLΔTAR-RRE-SA-Luc, envelope plasmid pHCMV-G and either packaging plasmid pCMVΔR8.2 (with intact integrase) or pCMVΔR8.2(D116N) (with mutant integrase) were co-transfected into HEK293T producer cells to produce IN+ Tat-positive TAR mutant virions vHIV-LucΔTAR and IN- Tat-positive TAR mutant virion vHIV-LucΔTAR(D116N).
Figure 1.19

An HIV-1 mutant TAR viral infection also resulted in a decrease in reporter gene expression. CEM-SS cells were infected with 45, 18, and 7 ng of p24 input virus. Luciferase activity of a total of 1 X 10^5 and 5 X 10^5 cells per sample was read 48 hours post infection.
input virus and $5 \times 10^5$ CEM-SS cells samples. This result indicates that, just like a $\Delta$Tat infection, a $\Delta$TAR viral infection also causes a defect in reporter expression from the viral template and that this defect is independent of the viral DNA being integrated into the host genome. Subsequently, we checked if equal HIV DNA was synthesized by the viral reverse transcription process.

**Low DNA synthesis during reverse transcription in a TAR mutant viral infection**

To evaluate the DNA synthesis in a TAR mutant viral infection, a total of $3 \times 10^6$ CEM-SS cells were infected with 150 ng input virus. The viruses used for infection were the IN+ Tat-positive virus vHIV-Luc, the IN- Tat-positive virus vHIV-Luc(D116N), IN+ Tat-positive TAR mutant virus vHIV-Luc$\Delta$TAR, and the IN- Tat-positive TAR mutant vHIV-Luc$\Delta$TAR(D116N). Cells were harvested and lysed 2- and 6-hours post infection. DNA was extracted and 100 ng of total cellular DNA was used from each sample for real-time PCR targeting total, full-length viral DNA for amplification. Clearly noticeable in the result shown in Figure 1.20 is the reduction in viral DNA synthesized at 6 hours post infection in the viral infection that included the mutant TAR sequence for both the IN+ and IN- samples. For the IN+ and IN- samples harvested 6 hours post infection, we observed a 9.4- and 7.6-fold decrease, respectively, in synthesis of total HIV DNA when the TAR sequence included a 4-base deletion. Overall, the result demonstrates that the decrease in reporter gene expression in the mutant TAR infections (Figure 1.19) was mainly caused by the lack of comparable DNA synthesis during the reverse transcription process. The same pattern was noticed in the mutant Tat infections compared to wild-type
At six hours post infection, a distinct reduction in DNA synthesis was observed in the mutant TAR infection. CEM-SS cells (3 X 10^6) were infected with 150 ng input virus and samples were taken 2 and 6 hours post infection. DNA was extracted and 100 ng of total cellular DNA was subjected to real-time PCR targeting full-length viral DNA.
Tat infections. Unequal DNA synthesis was caused by a defect in RNA genome packaging in the production of virions in the producer cells. Subsequently, we next investigated if there was less viral genomic RNA encapsidation during the assembly of virions in the presence of a ΔTAR sequence.

**Uneven genomic RNA packaged in a mutant TAR virus compared to a wild type TAR virus**

We examined the genomic RNA that was packaged in the IN+ Tat-positive virus vHIV-Luc, the IN- Tat-positive virus vHIV-Luc(D116N), IN+ Tat-positive TAR mutant virus vHIV-LucΔTAR, and the IN- Tat-positive TAR mutant vHIV-LucΔTAR(D116N). The viruses were produced by transfection (Attractene Transfection Reagent was used in this experiment, Qiagen – Venlo, The Netherlands) in HEK293T producer cells. The virus was harvested and Benzonase treated to remove residual DNA. Six mL of the virus supernatant was layered over a 3 mL - 20% sucrose cushion and the virus vHIV-LucΔTAR and vHIV-LucΔTAR(D116N) were run in duplicate. After the ultracentrifugation step, the pellet was resuspended in the remaining 1X PBS and a small amount of sample (18 µL into 2 µL of p24 lysis buffer) was collected for p24 ELISA. Equal p24 value for each virus (~5400 pg) and kanamycin control RNA were subjected to RNA extraction and reverse transcribed with either random decamers or backward luciferase primer. cDNA was subjected to real-time PCR targeting total, full-length HIV DNA for amplification. The data was normalized with subtraction of DNA contamination, and with kanamycin control RNA incases when cDNA was used that originated from RNA reverse transcribed with random decamers.
Figure 1.21

Amount of packaged viral genomic RNA was different in a mutant TAR virus compared to a wild-type TAR virus. Viral supernatant was pelleted through a 20% sucrose cushion and from resuspended viral pellet, comparable p24 values (5,400 pg) per virus were subjected to RNA extraction. cDNA was produced in the reverse transcriptase assay with either the utilization of random decamers or a luciferase backward primer. Total, full-length HIV genome was targeted and amplified in the real-time PCR. Kanamycin control RNA was used for normalization. Average was taken from two samples of the vHIV-LucΔTAR and vHIV-LucΔTAR(D116N) viral infections.
In the samples where random decamers were utilized, we observed a 1.8-fold decrease in the packaged RNA in the mutant TAR/IN+ sample and a 4-fold decrease in the IN- samples. In the samples subjected to the luciferase backward primer in the reverse transcription process, observed was a 7-fold decrease in RNA packaged in the mutant TAR/IN+ samples and a 4.3-fold decrease in the mutant TAR/IN- samples (Figure 1.21). Overall, a remarkable decrease in the RNA packaged in the produced virions with a mutant TAR sequence for both the IN+ and the IN- samples was noticed.

This experiment was repeated, but with the viruses vHIV-Luc(D116N) and vHIV-LucΔTAR(D116N) run in duplicate and a total of ~ 20,000 pg p24 value of each virus was used for RNA extraction. The result is presented in Figure 1.22 and observed was an overall decrease in the viral RNA packaged in the produced virions with a mutant TAR sequence for both the IN+ and the IN- samples. In the samples that were reverse transcribed with the random decamers, we observed a 2.7-fold decrease in the packaged viral RNA in the mutant TAR/IN+ sample and a 3.5-fold decrease in the IN- samples. In the samples subjected to the luciferase backward primer in the reverse transcription process, we observed a 1.7-fold decrease in viral RNA packaged in the mutant TAR/IN+ samples and a 3-fold decrease in the mutant TAR/IN- sample.

The results in the mutant TAR viral infection showed the same pattern as the results in the mutant Tat viral infection. From this observation we concluded that both protein Tat and the TAR sequence were essential for the effectual packaging of the viral genome in assembled virions. Tat and an intact TAR element are required when
Figure 1.22

Unequal viral genomic RNA packaged in a mutant TAR virus compared to a wild-type TAR virus. Viral supernatant was pelleted through a 20% sucrose cushion and from resuspended viral pellet, comparable p24 values (20,000 pg) per virus were subjected to RNA extraction. cDNA was produced in the reverse transcriptase assay with either the utilization of random decamers or a luciferase backward primer. Total, full-length HIV genome was targeted and amplified in the real-time PCR. Kanamycin control RNA was used for normalization. Average was taken from two samples of the vHIV-Luc(D116N) and vHIV-LucΔTAR(D116N) viral infections.
co-transfecting different plasmids in HEK293T virion producer cells when the plasmids include different promoters and at least one of the plasmids consists of an LTR promoter. Also, from this study we were unable to conclude if viral protein Tat is incorporated into the virion and if the protein participates in early transcription from the nonintegrated template.
Discussion
In this study we were not able to demonstrate that protein Tat is incorporated into virions during assembly or determine its potential function in preintegration transcription. However, we came to the conclusion that protein Tat is required in the producer cell line HEK293T when constructing viruses from plasmids with different promoters when at least one is an LTR promoter. In our case, the reporter plasmid, which provides the genomic viral RNA, was driven by the HIV LTR promoter. In our system, a Δtat virus was constructed from a packaging construct that provided the necessary viral proteins for virion assembly, but included a mutation in the Tat gene and was unable to produce a functional Tat protein. The viral 5’ LTR of the packaging construct was replaced by the human cytomegalovirus (hCMV) immediate early promoter for Tat-independent transcription (384) and the packaging signal was deleted. The VSV-G envelope plasmid pHCMV-G was also driven by a CMV promoter. The reporter plasmid used in our system consisted of a packaging signal, and was LTR-driven. These plasmids, the CMV-driven packaging and envelope plasmids and the LTR-driven reporter plasmid, were co-transfected into the virion producer cell line HEK293T to produce the Δtat virus. Most likely, less viral genomic RNA was produced due to the slow kinetics of basal transcription (tat-independent transcription) compared to the viral proteins produced from the CMV-driven plasmids. In a study by Rulli et al., cellular mRNA was found to be incorporated into the nascent virions due to the absence of viral genomic RNA with a proper packaging signal (383). With this knowledge we can hypothesize that due to the low ratio of viral genome to viral proteins, cellular mRNA was additionally included in the assembled virions in the absence of Tat. This defect has major downstream effects on
viral infections, such as a decrease in the production of viral DNA in the reverse transcription process and a subsequent loss in viral gene expression.

One of the researchers who has studied the viral protein Tat extensively is Dr. David Harrich from the Queensland Institute of Medical Research, Herston, Queensland, Australia. Harrich and his colleagues discovered that Tat (218,266) and the TAR element (385) were both involved in the reverse transcription process of HIV-1. In the two studies that determined Tat was engaged in the reverse transcription process, the researchers examined the RNA content in the Δtat virus and compared it to the content of a wild-type Tat virus (218,266). Their findings established that there were comparable levels of viral RNA encapsidated into the virions in either the absence or presence of Tat. However, the Harrich group used a different approach for constructing the Δtat viruses. They constructed a stably-transfected cell line with the HIV-1 Δtat provirus. To overcome the requirement of Tat’s transactivation function, they used the cell line HEK293T, which expresses the adenovirus transactivator proteins E1A and E1B (266). These proteins have been shown to greatly boost HIV-1 gene expression in a manner that is Tat and TAR independent (386,387). The Harrich group also replaced the nef gene with an antibiotic resistant gene to select for clones that contained the provirus (388). They were successful in producing Δtat and TAR mutant viruses for their studies.

Even though we used the producer cell line HEK293T for constructing the viruses in the presented research, there were major alterations in HIV-1 RNA encapsidation in the Δtat and the TAR mutant viruses (Figure 1.15, Figure 1.16, Figure 1.17, Figure 1.21, and Figure 1.22). The data presented in Figure 1.7, where an IN- wild-type Tat infection
is compared to an IN- Tat negative infection, reflects the defect of unequal RNA genome incorporated into the virions. From the three Tat mutants used in this study (pro6-10, C27S, and K41A), Tat mutant pro6-10 was deficient in the reverse transcription activity, while K41A and C27S were deficient in transactivation of the viral LTR promoter activity. Although deficient in reverse transcription, the Tat mutant pro6-10 consists of an intact domain responsible for the transactivation activity of Tat, this explains the highest luciferase activity among the mutants (Figure 1.10) and the highest total HIV DNA produced during reverse transcription (Figure 1.11). This indicates that Tat mutant pro6-10 was complementing the defect of the Δtat in the packaging plasmid in the virion producer cells and presumably did produce an equivalent amount of viral RNA genome that was encapsidated into the virions (this was not examined in this study). However, due to its defect in reverse transcription activity, both the luciferase expression and the DNA synthesis were not comparable to the wild-type Tat virus infection.
**Future Studies**

A functional assay to evaluate the presence of Tat in the HIV-1 virion could be very valuable. We could use the cell line 1G5 that is stably transfected with an HIV-LTR-luciferase construct. This cell line is characterized by a low basal level luciferase activity, HIV infectability, and high responsiveness to Tat expression (389). We would need to infect these cells with a Δtat virus constructed with Tat provided in trans in HEK293T producer cells. If the protein Tat is incorporated into the virion, it should be able to transactivate the LTR promoter of the stable transfected reporter plasmid in the 1G5 cell line. A Δtat virus constructed with an empty vector in trans could be used as a negative control. We know from our research that this virus includes less viral genomic RNA when compared to the Δtat virus constructed with Tat provided in trans, but we are not evaluating the genomic RNA in this system. To investigate the functionality of the virion-associated Tat, if incorporated into virions during assembly, we may construct Δtat viruses with different amounts of the wild-type tat plasmid that is provided in trans in the producer cell line. If Tat is incorporated into the virions, we assume that more Tat will be incorporated when increasing the provided Tat in trans. We hypothesize that this will also increase the reporter activity in the 1G5 cell line.

To be able to study the function of virion-associated Tat on transcription of nonintegrated DNA with the use of a LTR-driven reporter plasmid, we could slow down the production of viral particles so the ratio of viral genome to viral particles will provide a more realistic representation. To accomplish this, both the packaging construct, and the VSV-G plasmid need to be under control of a viral LTR promoter and co-transfected in the producer cell line HEK293T, which expresses the transactivators E1A and E1B from
adenovirus. In a Tat negative viral production cycle, using all LTR-driven plasmids for the virus assembly, it is expected that there will be a decrease in the amount of produced viruses. However, the produced viruses are expected to have two copies of the viral genomic RNA incorporated. The lower production of virus can be compensated for by using equal p24 virus values in a mutant Tat versus a wild-type Tat infection.

Another system that could be used to look at the impact of Tat on nonintegrated viral DNA transcription is the Tet-On system. This system is described in detail in Chapter Three. In this system we could use a Δtat virus constructed from the packaging plasmid pCMVΔR8.2(D116N)Δtat for the necessary viral components, the envelope plasmid pHCMV-G, and the packaged reporter vector plasmid, pNL-RRE-SA-Luc. The viral particles will be constructed in HEK293T producer cells by co-transfection of the plasmids. Next, HeLa Tet-On cells will be transfected with the doxycycline-induced Tat response plasmid. Thereafter, the cells will be infected with the Tat negative (IN-) virus vHIVΔtat-Luc(D116N) and after infection, doxycycline will be added to selected samples. The doxycycline will induce the production of Tat, which represents the virion-associated Tat. To compare infected cells with and without added doxycycline, we are then able to study the impact of virion-associated Tat on nonintegrated viral DNA transcription. Even though the packaged viral genome in the Tat negative virus is lower compared to a Tat positive virus, the amount will be equal in the compared samples.
CHAPTER TWO

Cellular Transcription Factors Sp1 and NF-κB are involved in Preintegrated Transcription

Introduction
The HIV-1 genome is bordered by two LTRs (long terminal repeats) and each LTR is organized into three domains, the U3, R, and the U5 domains (36,58). Even though the 5’ and 3’ LTR are nearly the same, they have different functions in the life cycle of HIV-1. The 5’ LTR functions as the viral promoter for the complete genome while the 3’ LTR acts in transcription termination, provides polyadenylation to nascent viral RNA, and encodes the viral accessory protein Nef (390). The U3 region of the 5’ LTR is subdivided into three different regions; the modulatory region, the enhancer or core enhancer, and the core region which is also known as the basal promoter (127). The long modulatory region, which spans nucleotide -454 to -104 (where nucleotide +1 is the transcription start site) (58), is completely overlapped by the nef gene (391), and the probability that this region is important in HIV-1 gene expression and AIDS pathogenesis is small (392). However, in earlier studies it was proposed that the modulatory region contains a negative regulatory element (NRE) between nucleotides -340 and -184 as deletions within this region increased viral transcription and replication (393,394). It has also been found that various cellular proteins such as c-Myb, NFAT, USF and COUP (58,395–398) interact with the modulatory region. The enhancer region (nt -105 to -79)
includes two κB sites, which are binding sites for the inducible transcription factor NF-κB. The core region (nt -78 to -1) contains the TATA box, a highly conserved AT-rich domain located 25 to 30 bp upstream from the transcription start site (399). The core region also includes three GC-rich DNA regions, which are the binding sites for transcription factor Sp1. Sp1 is a transcription factor that is ubiquitously expressed in the mammalian cells (58). The core or basal promoter region also includes the transcription start site (400). The 59 bp long trans activation response (TAR) element is located in the R region (nt +1 to +60) (64).

Basal gene transcription involves general transcription machinery. After the viral genome is integrated in the active chromatin of the host genome, the pre-initiation complex is assembled. First, the TATA binding protein (TBP), which is part of the transcription factor IID (TFIID), binds to the TATA box and functions as a scaffold for the pre-initiation complex assembly. TFIIA followed by TFIIB are recruited to stabilize the promoter bound TFIID. TFIIB recruits the complex RNA polymerase II (RNA Pol. II) and TFIIF. Together the TFIID, TFIIA, TFIIIB, and RNA Pol. II/TFIIF form a stable promoter complex. Next, TFIIE and subsequently, TFIIH are recruited to form the pre-initiation complex [for review (156)].

Besides the TATA box where the pre-initiation complex assembles, upstream of this box are two important enhancer elements that are binding sites for the proteins NF-κB and Sp1. The complex protein NF-κB is a phorbol ester- and tumor necrosis factor alpha (TNF-α)-inducible transcription factor. The protein binds to the enhancer of the κ light chain, HIV-1, and the upstream sequence of the Il-2 receptor α-chain gene
The NF-κB family of proteins includes two sub families: the NF-κB and the Rel proteins. All the members of the sub families have a highly conserved DNA-binding/dimerization domain, the Rel homology domain (RHD) (403). The subfamily Rel contains c-Rel, RelB, and RelA (also known as p65) and this group is characterized by their C-terminal transactivation domain. The NF-κB subfamily includes proteins that do not comprise the transcription activation capability, such as p50 and p52. Homodimers of p50 or p52 repress transcription and need a nuclear cofactor, such as Bcl-3, to form a ternary complex to be able to activate transcription or these family members have to form dimers with the members of the Rel subfamily to be capable of activating transcription. We do have to mention that the transcription factor nuclear factor of activated T-cell (NFAT) also holds the RHD. NFAT is also capable of binding to the same DNA sequences as the Rel/NF-κB dimers [for review: (404)]. From here on the term NF-κB refers to the p50-RelA/p65 heterodimer, the most commonly formed dimer and major complex in most cells (405,406).

In un stimulated cells, the dimerized nuclear transcription factor NF-κB is held hostage in the cytoplasm by a specific inhibitory protein, the IκB (407) and forms a NF-κB/IκB complex. The inhibitory protein covers the nuclear localization signal of NF-κB and obstructs the DNA binding sequences. There are two well-known pathways that lead to the activation of NF-κB in the cytoplasm. In the classical pathway or also known as the canonical pathway, a ligand binding to a receptor on the cell surface, such as a Tumor Necrosis Factor (TNF) receptor or a Toll-like receptor, subsequently recruits an adaptor protein to the cytoplasmic domain of the receptor. In turn, the adaptor protein recruits and
activates the IκB kinase (IKK) complex. The IKK complex includes two catalytic kinase subunits, IKKα and IKKβ (408,409), and a regulatory subunit, the NF-κB essential modulator (NEMO), also called IKKλ or IKKAP1 (410,411). The activated IKK complex is now capable of phosphorylating IκB (412), and this phosphorylated protein will be accordingly degraded by a proteasome. NF-κB is now capable of entering the nucleus and activating its target genes. This pathway consists of an auto-regulatory action where NF-κB activates also the IκB gene and this leads to the reappearance of the inhibitory complex in the cytoplasm. This classical pathway is used for the activation of the heterodimer p50-RelA/p65 (413). The second pathway, the alternative or non-canonical pathway, activates mostly the heterodimer p100-RelB during B- and T-cell organ development. This pathway is activated by only certain receptors, such as the B-cell activating factor (BAFF) receptor and CD40. In this case, the IKK complex includes two IKKα subunits, but not NEMO. The kinase NIK (NF-κB-inducing kinase) (414) is activated as a consequence of receptor binding and consecutively phosphorylates and activates the IKK complex. The IKK complex in turn phosphorylates the C-terminal IκB domain of p100 and this leads to the partial proteolysis and consequently to the release of the p52-RelB complex. This IKKα-dependent p52-RelB pathway is crucial to maintaining the production of CXCL12/SDF-1 for cell migration towards HMGB1 (High Mobility Group Box1) (415). Other, not well understood, pathways do exist. For example, the pathways of the activation of homodimers that enter the nucleus and are able, with the help of the IκB-like co-activator Bcl-3 (or IκBζ) which contains a transactivation domain, to become transcriptional activators (416). The proto-oncoprotein
Bcl-3 is an IκB family member with six to seven ankyrin repeat motifs and is located predominantly in the nucleus (417). Studies indicate that Bcl-3 is capable of activating the p50 and p52 transcriptionally inactive homodimers (416), or disconnecting the binding homodimers from the NF-κB DNA binding sites to activate transcription (418,419). In an HIV-1 infection, several viral elements are known to stimulate the NF-κB pathway such as the viral components gp120, Tat, Nef and Vpr (114,420–422).

Sp1 (specificity protein 1) is a mammalian, ubiquitously expressed, nuclear transcription factor that can activate or repress transcription in response to physiological and pathological stimuli. It regulates various cellular functions including: cell growth, differentiation, apoptosis, angiogenesis, and immune response (70,71,423,424). Sp1 holds three C2H2-type zinc fingers that function as DNA binding domains (425). This transcription factor binds to GC-boxes (68,426) and, in lesser affinities, to CT-boxes and GT-boxes (427). Sp1 is a transactivator and interacts with basal transcription machinery. The protein consists of two glutamine-rich transactivation domains, A and B (79,423,428), and each directly interacts with both the TBP (81,429) and TAF4 (430,431), which are units of the basal transcription factor TFIID. Sp1 was also found to be involved in the assembly of TFIIB and TFIIE into the pre-initiation complex (432). This protein also directly or indirectly interacts with TFIIA to heighten transcription in vitro (433). Jeang et al. reported that viral protein Tat and cellular protein Sp1 form a protein-protein complex (50). However, it was found that this is not a direct interaction, but is mediated most likely by a viral or cellular protein (434,435). The double-stranded DNA-dependent protein kinase (DNA-PK) was identified as an Sp1 kinase (436) which
can regulate Sp1 DNA binding or transcriptional activity. This Sp1 kinase is also competent at phosphorylating the CTD of RNA Pol. II (437) and this action is enhanced by the presence of adjacent transcriptional activator domains (438). Another feature of Sp1 is its ability to interact with a subset of NF-κB binding sites and this action may substitute for the absence of active, nuclear NF-κB molecules and increase the basal NF-κB-dependent gene expression (439).

HIV-1 transcriptional regulation depends on host nuclear factors such as NF-κB and Sp1 in the absence of viral protein Tat. The binding sites for NF-κB and Sp1 in the HIV-1 LTR are highly conserved among differing isolates, which indicates an important function in regulating gene expression (49,55,440,441). In a study employing a transient transfection assay using plasmids with different LTR mutants in HeLa cells, it was found that the enhancer NF-κB was more important to basal transcription (Tat-independent) from the LTR than Sp1, with site NF-κB (I) the most essential of the two sites (54). Mutations in the individual Sp1 sites had little consequence on the basal transcription rate in this study (54). However, basal expression was significantly reduced when the mutations were combined in two or more Sp1 sites. This study also demonstrated that if the two intact NF-κB sites are located closer to the TATA box (by removing the three Sp1 binding sites), an increase in basal transcription is observed (54). When the system used in this study was co-transfected with a Tat-expressing plasmid, NF-κB had a reduced impact on Tat-dependent transcription, while Sp1 had a more important role in Tat-induced expression compared to the basal transcription. Surprisingly, modifications
in the TATA sequence had an increased negative impact on Tat-dependent transcription, compared to a minor impact on basal transcription (54).

In a study by Perkins et al., they demonstrated in a transfection assay that the NF-κB binding to their binding sites in the viral LTR is by itself insufficient to initiate HIV-1 basal gene expression. Rather, an interaction must take place between the adjacent bound proteins of NF-κB and Sp1 (56,442). However, a study in 2001 found that NF-κB was capable of recruiting P-TEFb to the bound RNA Pol. II and stimulate transcriptional elongation. The RelA/p65 subunit of NF-κB associates with cyclinT1, a subunit of P-TEFb. This study suggests that the kinase of P-TEFb, CDK 9, phosphorylates the serine 2 and 5 of the CTD in the RNA Pol. II which causes elongation of transcription in the absence of Tat (443).

The HIV-1 virion associated viral protein R (Vpr) is known to increase the viral replication rate via transactivation of the HIV-1 LTR (444). Vpr is able to bind to TFIIB of the basal transcriptional machinery (445). It has also been demonstrated that this viral protein is proficient at interacting with cellular protein Sp1 when bound to the HIV-1 LTR. The transactivation of Vpr using Sp1 could be important in the early transcription of HIV-1 when transcription factors such as NF-κB are restricted or inactive (446). In a study by Sawaya et al. it was found that Vpr is also capable of interacting with viral protein Tat and cellular protein cyclinT1. Tat’s ability to interact with cyclinT1 is well known, but the binding sites of Vpr and Tat to cyclinT1 are different and the binding to cyclinT1 is independent from each of the viral proteins (447).
In the next presented study we investigated the regulation of gene transcription from nonintegrated viral DNA. In a preliminary study by Dr. Wu, it was demonstrated that an HIV-1 infection with a nonfunctional integrase (IN-) showed a homologous transcription activity compared to a heterogeneous transcription activity in an HIV-1 infection with a functional integrase (IN+) incorporated into the virion. From this preliminary data, we assume that the transcriptional regulation of the integrated and nonintegrated viral DNA is not the same. This study focuses on the regulation of basal transcription by transcription factors NF-κB and Sp1 and whether the regulation occurs in the same manner in integrated viral DNA compared to nonintegrated viral DNA.
Materials and Methods

Plasmids

The packaging construct pCMVΔR8.2 consists of the human cytomegalovirus (hCMV) immediate early promoter which replaced the U3 region of the viral 5’ LTR, resulting in a Tat-independent transcription plasmid. The packaging signal was deleted, but the 5’ splice donor site was kept intact. The viral 3’ LTR was replaced with a poly A site and the envelope and accessory protein Vpu were made defective (369,370). The plasmid pCMVΔR8.2(D116N) is the integrase negative counterpart of plasmid pCMVΔR8.2 that carries a mutation in the integrase catalytic domain (Asp 116 to Asn) which makes this protein nonfunctional (D116N) (277). The plasmid pHCMV-G expresses the vesicular stomatitis virus (VSV) glycoprotein and has been described previously (375). This plasmid provides the envelope and viruses pseudotyped with the VSV glycoprotein envelope, enter cells through an endocytic pathway (376).

A set of five Sp1 target MISSION® TRC shRNA lentiviral plasmids were purchased from Sigma-Aldrich (St. Louis, Missouri) to knockdown the transcription factor Sp1. The clones arrived as frozen bacterial glycerol stocks (Luria Broth, carbenicillin at 100 µg/mL and 10% glycerol) in Escherichia coli cells. The shRNA clones are lentiviral plasmids, pLKO.1-puro, which contains the puromycin selection marker for transient or stable transfection. The Sp1 shRNA target set is visualized in Figure 2.1. The vector map of the plasmid pLKO.1-puro is shown in Figure 2.2.
<table>
<thead>
<tr>
<th>TRC No.</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRCN0000020444</td>
<td>CCGGCCCAAGTTATTTCTCTCTTAATCTCGAGTAAGAGAGAAATAAAACTTTGGGT</td>
</tr>
<tr>
<td></td>
<td>Clone ID : NM_138473.2-7276s1c1</td>
</tr>
<tr>
<td></td>
<td>Accession Number(s) : NM_138473.2</td>
</tr>
<tr>
<td></td>
<td>Region : 3UTR</td>
</tr>
<tr>
<td>TRCN0000020445</td>
<td>CCGGCCAGCAAATTCTCGAGAATTCTCCTGAGAAATCTGCTGGTTGCTGCTTGTTTTTT</td>
</tr>
<tr>
<td></td>
<td>Clone ID : NM_138473.2-290s1c1</td>
</tr>
<tr>
<td></td>
<td>Accession Number(s) : NM_138473.2</td>
</tr>
<tr>
<td></td>
<td>Region : CDS</td>
</tr>
<tr>
<td>TRCN0000020447</td>
<td>CCGGCCAGTGTGAAAACCAACAGATTTCTCTGAGAGAATCTCCTGAGATGTATTCCATCAACCACCAGCTTTTT</td>
</tr>
<tr>
<td></td>
<td>Clone ID : NM_138473.2-719s1c1</td>
</tr>
<tr>
<td></td>
<td>Accession Number(s) : NM_138473.2</td>
</tr>
<tr>
<td></td>
<td>Region : CDS</td>
</tr>
<tr>
<td>TRCN0000020448</td>
<td>CCGGGCTGTGCTGGATGAGATAACATCTCGAGAGAATCTCCTGAGATGTATTCCATCAACCACCAGCTTTTT</td>
</tr>
<tr>
<td></td>
<td>Clone ID : NM_138473.2-1805s1c1</td>
</tr>
<tr>
<td></td>
<td>Accession Number(s) : NM_138473.2</td>
</tr>
<tr>
<td></td>
<td>Region : CDS</td>
</tr>
<tr>
<td>TRCN0000020446</td>
<td>CCGGCCACTCTCTCAGGCCCTATTACTCGAGTAATAAGGGCTGAGGATGGTTTTTT</td>
</tr>
<tr>
<td></td>
<td>Clone ID : NM_138473.2-2310s1c1</td>
</tr>
<tr>
<td></td>
<td>Accession Number(s) : NM_138473.2</td>
</tr>
<tr>
<td></td>
<td>Region : CDS</td>
</tr>
</tbody>
</table>

**Figure 2.1**

**Mission™ TRC shRNA.** shRNA oligonucleotide sequences for the Sp1 target site (Sigma-Aldrich - St. Louis, Missouri).
Figure 2.2

**pLKO.1-puro vector description and features** (Sigma-Aldrich - St.Louis, Missouri).

The HIV-1 LTR-driven, Rev independent, reporter plasmid pNL-RRE-SA-Luc was derived from plasmid pNL-RRE-SA by inserting the *luciferase* gene within the XhoI cloning site. The cloning procedure is described in more detail in the previous chapter. Different mutations and deletions were inserted in the Sp1 and NF-κB binding sites located in the U3 region in the 3’ LTR of plasmid pNL-RRE-SA-Luc. The mutations and deletions were transferred to the 5’ LTR of the reporter plasmid after reverse transcription. The plasmid pUC19-LTR(KpnI-Ncol), described in Chapter One, was utilized to clone the following mutant plasmids: pNL-Luc-Mut-Sp1(I) consists of a two base mutation in the Sp1(I) binding site (GG → TT); pNL-Luc-Mut-Sp1(II) consists of a two base mutation in the Sp1(II) binding site (GG → TT); pNL-Luc-Mut-Sp1(III) consists of a two base mutation in the Sp1(III) binding site (GG → TT); pNL-Luc-Mut-Sp1(all) consists of a two base mutation in the Sp1(I), Sp1(II), Sp1(III) binding site (GG → TT); pNL-LucΔSp1 contains a 38-base deletion that includes all three Sp1 binding sites; pNL-Luc-Mut-NF-κB(I) contains a three base mutation in the NF-κB(I) binding site (GGG → TCT); pNL-Luc-Mut-NF-κB(II) contains a three base mutation in the NF-κB(II) binding site (GGG → TCT); pNL-Luc-Mut-NF-κB(all) contains a three base mutation in the NF-κB(I) and the NF-κB(II) binding site (GGG → TCT); pNL-LucΔNF-κB, consists of a 24-base deletion that includes both NF-κB binding sites; pNL-Luc-Mut-NF-κB(I)-Sp1(III) consists of a three base mutation in the NF-κB(I) binding site (GGG → TCT) and with a two base mutation in the Sp1(III) binding site (GG → TT). Figure 2.3 illustrates the distinctive mutations in each plasmid.
Figure 2.3

Mutations in Sp1 and NF-κB binding sites (Adapted from Burnett JC, Miller-Jensen K, Shah PS, Arkin AP, Schaffer DV. Control of Stochastic Gene Expression by Host Factors at the HIV Promoter. PLoS Pathog. 2009 Jan 9;5(1):e1000260). An overview of the mutations made in the two NF-κB (red and underlined) and the three Sp1 (green and italic) binding sites in the 5’ viral LTR. The boxed bases indicate the mutation and the dashed line indicates the deleted sequence.
Cells
PBMCs (peripheral blood mononuclear cells) were obtained from healthy, HIV negative donors at the Student Health Center, George Mason University (GMU) and purified by Ficoll-Hypaque density gradient centrifugation. The purified cells were diluted at a concentration of 1 X 10^6 T cells/mL in RPMI 1640 medium supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 µg/mL). The cells were pre-stimulated with anti-CD3/CD28 antibody-conjugated beads (2 beads/cell) and incubated overnight at 37° C and 5% CO₂ before any treatment.

HeLa and HEK293T cells were cultured in DMEM supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 µg/mL) at 37° C and 5% CO₂. CEM-SS were cultured in RPMI 1640 medium supplemented with 10% heat-inactivated FBS at 37° C and 5% CO₂.

Propagation and downstream purification of the shRNA clones
In a sterile 500 mL Erlenmeyer flask, 150 mL of LB broth medium and 150 µL carbenicillin (100 µg/mL) were mixed. To this nutritionally rich medium, 50 µL of the shRNA bacterial glycerol stock was added. The bacterial cultures were incubated in a shaking incubator (Analytical Instruments, LLC - Golden Valley, Minnesota) for 16 - 18 hours at 37° C at a shaking speed of 200 rpm. After the incubation period, 850 µL was added to 150 µL of sterile glycerol for long-term storage at -80° C. Plasmid DNA was extracted from the remaining culture according to the instructions of the Promega Pure Yield™ Plasmid Midiprep System (Promega – Madison, Wisconsin) protocol. The DNA was eluted in nuclease-free water, and 5 µL of the DNA solution was diluted 10 times
with nuclease free water to measure DNA concentration at optical density at 260 nm (OD$_{260}$) using Eppendorf BioPhotometer (Eppendorf - Westbury, New York).

**Generation of Sp1 shRNA lentiviral particles**

The Sp1 shRNA lentiviruses were generated by co-transfection of HEK293T cells with 10 µg shRNA lentiviral plasmid (transfer vector plasmid), 7.5 µg pCMVΔR8.2 (packaging construct), and 2.5 µg pHCMV-G (envelope plasmid) per 10 cm (diameter) dish. For the transfection, Lipofectamine™ 2000 was used according to manufacturer’s instructions. Dishes were incubated for about 48 hours at 37° C and 5% CO$_2$.

**Concentration of shRNA lentiviral particles**

After transfection and the 48 hours incubation period, the viral supernatant from the dishes with the transfected HEK293T cells was collected and stored in a 4° C refrigerator. Freshly prepared DMEM medium supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 mg/mL) was added back to the cells and incubated for another 24 hours at 37° C and 5% CO$_2$. The next day, the supernatant was collected and combined with the supernatant collected the day before. Any cells or cell debris in the virus supernatant was pelleted by centrifugation at 1500 rpm for 5 minutes. The supernatant was carefully decanted into a new conical and stored on ice. For the concentration procedure, the Vivaspin 20 centrifugal concentrator (Sartorius Stedim North America Inc. – Bohemia, New York) with a 100,000 MWCO (Molecular Weight Cut Off) PES (Polyethersulfone) membrane was used. The centrifuge was cooled down to 4° C and up to 20 mL of the viral supernatant was added to the concentrator and centrifuged for 20 minutes at 5,000 g at 4° C. Depending on the volume of the viral
mixture, this step was repeated multiple times until all the collected supernatant was used and the desired volume was reached in the upper chamber of the concentrator. The flow-through was discarded. The concentrated viral mixture was collected from the upper chamber by pipetting up and down over the membrane and subsequently added to a new, labeled tube that was placed on ice. Afterwards, the viral mixture was filtered by adding the solution to a sterile syringe (10 mL luer-lok™ tip, Becton, Dickenson and Company – Franklin Lakes, New Jersey) that was attached to a 0.45 µm pore size nitrocellulose membrane filter unit (Millex-HA, Millapore Corporation – Billerica, Massachusetts). The syringe plunger was squeezed carefully until all the supernatant was filtered. From the sterile viral supernatant, 5 µL was added to 45 µL of p24 ELISA lysis buffer (a 10-fold dilution) used for measuring the p24 value. The remaining solution was aliquoted on ice in 500 and 250 µL aliquots and stored at -80°C.

**Transfection of shRNA in HeLa cells for Western blot analysis**

Added were 5 X 10⁵ HeLa cells to each well of a 6-well plate (with a growth area of 9.5 cm²/well) in a total volume of 2 mL of DMEM supplemented with 10% heat-inactivated FBS medium and incubated overnight at 37°C and 5% CO₂. The cells were microscopically examined to see if they had reached approximately 80% confluency, which is an optimal state for the transfection of the shRNA plasmids. For the transfection, 4 µg of plasmid DNA was used. The used plasmids were the Sp1 target MISSION® TRC shRNA lentiviral plasmids, TRCN20444, TRCN20445, TRCN20446, TRCN20447, TRCN20448, and a non-targeting control (NTC) plasmid. This control plasmid contains an shRNA insert that does not target any known genes from any species. The plasmids
were added individually to 250 µL of serum-free DMEM medium. At the same time, 10 µL of Lipofectamine™-2000 (Invitrogen, Life technologies – Carlsbad, California) was added to 250 µL of serum-free DMEM medium and incubated for 5 minutes at room temperature. After the incubation period, the medium and lipofectamine mixture was slowly added to the DNA sample and incubated for 20 minutes at room temperature. During incubation, the medium in the 6-well plates was replaced with 1.5 mL fresh serum-free DMEM medium. Following incubation, DNA/lipofectamine mixture (about 500 µL) was added to the correct well of the labeled 6-well plate with the HeLa cells. The cells were incubated at 37° C and 5% CO₂ for about 5 hours. After the 5 hour incubation, the medium in the wells was replaced with 2 mL freshly made DMEM supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 µg/mL). Subsequently, the 6-well plates were incubated at 37° C and 5% CO₂ for approximately 48 hours.

Collecting cells for Western blot
To dislodge the cells from the wells, the medium was removed, and a total of 500 µL of 0.25% Trypsin-EDTA (1X) Phenol Red (Gibco, Life Technologies - Grand Island, New York) was added to each well. The plates were incubated for about 5 minutes at 37° C and 5% CO₂. After confirmation by microscopy that the cells were dislodged from the wells, 3 mL of DMEM supplemented with 10% heat-inactivated FBS medium was added to stop the trypsin reaction. The cell mixture was pipetted up and down to de-clump the cells and thereafter the mixture was added to a 15 mL conical. A small amount (10 µL) of cell mixture was diluted 1:5 in 0.4% Trypan Blue Solution (Gibco, Life Technologies -
Grand Island, New York) and approximately 10 µL was added to a hemocytometer. Only life cells (bright cells) were counted in the four corner squares (1mm²) and averaged. The average was multiplied by the dilution factor, 5. This cell number represents the amount of cells per 10⁻⁴ cm³ (1 mm X 1 mm X 0.1 mm). By multiplying by 10⁴, the number of cells is per cm³, which is approximately equivalent to the number of cells per 1 mL.

Afterwards, the total amount of cells was centrifuged for 5 minutes at 1200 rpm and the supernatant was discarded. The cells were washed in 1 mL of 1X PBS and centrifuged again at 1200 rpm for 5 minutes. Carefully, the supernatant was removed without disturbing the pellet and cells were lysed in 100 µL per 1 X 10⁶ cells of NuPAGE® Lithium Dodecyl Sulfate (LDS) Sample Buffer (Invitrogen, Life technologies – Carlsbad, California). Samples were heated to about 95° C in a heat block for 10 minutes and sonicated for preparing protein lysate.

**Western blot**

The samples were centrifuged at 13,000 rpm for 10 minutes before loading 20 µL per sample on a NuPage® Novex 4-12% Bis-Tris 1.0 mm gel (Invitrogen, Life technologies – Carlsbad, California). The gel was run at 150 volts for approximately 45 minutes. The separated samples were transferred to a nitrocellulose membrane (Invitrogen, Life technologies – Carlsbad, California) at 30 volts for 2 hours on ice. Afterwards, the membrane was washed with Tris-Buffered Saline and Tween 20 (TBS-T) (Tris 50 mM, NaCl 150 mM, and Tween20 0.05%) for 5 minutes, 3X on a shaker to remove the residual gel components and transfer buffer. Next, the gel was blocked for 30 minutes at room temperature in either blocking buffer [StartingBlock™ (TBS) Blocking
Buffer – Thermo Fisher Scientific Inc. - Waltham, Massachusetts] or 2.5% skim milk powder (Difco™ Skim milk – Becton, Dickenson and Company – Franklin Lakes, New Jersey) dissolved in TBS-T and gently rocked. The membrane was washed again with TBS-T for 5 minutes, 3X. Following the wash, the membrane was incubated with the primary antibody at a recommended dilution in 2.5% skim milk powder dissolved in TBS-T and gently rocked overnight in the cold-room at 4°C. The next day, the membrane was washed 15 minutes, 3X with TBS-T. Afterwards, the membrane was incubated with a secondary antibody at a recommended dilution in 2.5% skim milk dissolved in TBS-T and gently rocked for an hour at room temperature. Hereafter, the membrane was washed again 15 minutes, 3X with TBS-T on a shaker and subsequently the blot was developed with Super Signal West Femto Maximum Sensitivity Substrate (Pierce – Thermo Scientific - Rockford, Illinois). Images were taken at various time points and captured with a CCD camera (FluorChem 9900 Imaging Systems, Alpha Innotech Corp. – San Leandro, California).

**Antibodies for Western blotting**

Rabbit polyclonal antibodies specific for transcription factor Sp1 (PEP 2 sc-59) were purchased from Santa Cruz Biotechnology, Inc. (Dallas, Texas) and used at a 1:500 dilution. A goat anti-rabbit horseradish peroxidase-conjugated antibody (Kirkegaard and Perry Laboratories, Inc. – Gaithersburg, Maryland) was used as secondary antibody at a 1:5000 dilution. For loading control, the blots were stripped and reprobed with antibodies against GAPDH (Abcam – Cambridge, Massachusetts) at a 1:1000 dilution.
Generation of the LTR mutant plasmids

Mutations and deletions were generated in the Sp1 and NF-κB binding sites that are located in the 5’ LTR of the LTR-driven plasmid pNL-RRE-SA-Luc. The HIV-1 promoter consists of three tandem Sp1 binding sites, namely I, II, and III. To be able to investigate each site independently, a mutation was designed in each binding site. A mutant version was also designed where the individual mutations were combined or where all three binding sites were deleted (85,448). Also located in the HIV-1 promoter are two tandem binding sites for NF-κB, namely I and II. Mutations were also designed in each NF-κB binding site independently, but also combined or both NF-κB binding sites were deleted (448). One designed mutant was a combination of a mutation in the Sp1 site III and a mutation in the NF-κB site I (448). An overview of the distinctive mutations is shown in Figure 2.3.

The synthesized DNA fragments containing the different mutations and deletions (Blue Heron Biotechnology - Bothell, Washington) were cloned into the cloning vector pUCminusMCS, transformed into GC10 E. coli competent cells, and shipped as a bacterial agar stab (1% tryptone, 0.5% yeast extract, 1% sodium chloride, 0.7% agarose) by the company. To obtain the recombinant DNA, a sterile 10 µL pipet tip was gently inserted into the agar stab and streaked out on a 1.5% agar plate containing LB broth and ampicillin (100µg/mL). The plates were inverted and incubated overnight at 37°C. The next day, by using sterile techniques, single colonies were inoculated from the plates into 5 mL LB growth medium containing 5 µL of ampicillin (100µg/mL). The cultures were incubated in a shaking incubator at 37°C and 220 rpm shaking speed overnight (16 - 18 hours). After the incubation period, 850 µL of the culture was added to 150 µL of sterile
glycerol and frozen at -80°C for long-term storage. The plasmid DNA was extracted from the remaining culture according to the plasmid extraction kit protocol (QIAprep Spin Miniprep Kit, Qiagen – Venlo, The Netherlands). The DNA was eluted in nuclease free H2O.

To be able to retrieve the synthesized de novo DNA fragments containing the mutations or deletions from the cloning vector DNA, two digestion sites were included in the designed DNA sequence. The added digestion sites were PmlI (CACGTG) and AflII (CTTAAG), shown in Figure 2.4 as the colors cyan and magenta, respectively. Figure 2.4 illustrates the complete synthesized DNA fragments with mutations or deletions. The next step was to replace the wild-type 3’ LTR sequence with the mutated sequence. Due to the reverse transcription process of HIV-1 (Figure 1.2), the mutations were cloned in the 3’ LTR of the plasmid DNA. The plasmid pUC19-LTR(KpnI-NcoI) was generated previously and the cloning steps that led to the production of this plasmid are described in Chapter One (Figure 1.4). For the generation of the LTR mutant plasmids, the pUC19-LTR(KpnI-NcoI) plasmid functioned as the vector backbone. This plasmid contained the required part of the 3’ LTR. Both the pUC19-LTR(KpnI-NcoI) plasmid and the plasmid containing the synthesized DNA fragment, p3’LTRmutant (named after the included mutation), were digested with the endonucleases PmlI and AflII. The digestion produced a 4069 bp vector backbone and a 227 bp insert and the two DNA fragments were ligated together. This ligation generated the plasmid pUC19LTRmutant (named after the included mutation). Plasmids pUC19LTRmutant and the LTR-driven reporter plasmid pNL-RRE-SA-Luc were subsequently digested with the enzymes KpnI and NcoI. The
Figure 2.4

Sequences of DNA synthesis. Complete sequences of the DNA synthesis for generating the mutations of the NF-κB and Sp1 binding sites in the HIV-1 promoter. The added
digestion site PmlII is shown in cyan and the added AflII site is shown in magenta. The two NF-κ binding sites are displayed in red and underlined. The three Sp1 binding sites are represented in green and italic. The boxed bases depict the mutation.
smaller DNA fragment from the digested plasmid pUC19LTRmutant was sticky-end ligated into the larger digest product of plasmid pNL-RRE-SA-Luc. This cloning step replaced the wild-type sequence with the individual LTR mutant into the 3’ LTR of plasmid pNL-RRE-SA-Luc. A complete cloning schema is shown in Figure 2.5. The incorporated mutations and deletions in the 3’ LTR region were confirmed by sequencing the region with the primer: 3’ – TGC ATG GAA TGG ATG ACC CTG AGA – 5’ (Northwoods DNA, Inc. – Solway, Minnesota).

**Generation of reporter viral particles**

The VSV-G pseudotyped LTR-driven luciferase virus particles were generated by co-transfection of HEK293T cells with 10 µg pNL-RRE-SA-Luc (reporter vector plasmid), 7.5 µg pCMVΔR8.2(D116N) or pCMVΔR8.2 (packaging construct), and 2.5 µg pHCMV-G (envelope plasmid) plasmids per 10 cm dish. For the transfection, Lipofectamine™ 2000 was used.

**Generation of LTR mutant reporter viral particles**

The VSV-G pseudotyped LTR mutant viruses were generated by co-transfection of HEK293T cells with 10 µg pNL-RRE-SA-Luc-LTR mutant or wild type pNL-RRE-SA-Luc (reporter vector plasmid), 7.5 µg pCMVΔR8.2(D116N) or pCMVΔR8.2 (packaging construct), and 2.5 µg pHCMV-G (envelope plasmid) plasmids per 10 cm dish. For the transfection Lipofectamine™ 2000 was used.
Figure 2.5
Cloning design of the 3’ LTR mutants. Mutations in the NF-κB and Sp1 binding sites were incorporated in the 3’ LTR of the LTR-driven reporter plasmid: pNL-RRE-SA-Luc.

Step 1: PmlI and AflII digestion of the vector backbone plasmid pUC19-LTR(KpnI-NcoI) and the plasmid containing the individual LTR mutants, p3’LTRmutant. The DNA fragments were ligated and this generated the different plasmids pUC19LTRmutant.

Step 2: Replacement of wild-type 3’ LTR region in the plasmid pNL-RRE-SA-Luc by the individual mutations.
Results

Our goal in the next set of experiments was to establish a functional system where we stably transduced cells with an shRNA (short hairpin RNA) targeting the Sp1 gene. Afterwards, these transduced cells were infected with an LTR-driven reporter particle that either contained a functional integrase (IN+) or did not contain a functional integrase (IN-). This allowed us to study the transcriptional regulation by transcription factor Sp1 on nonintegrated viral DNA versus integrated viral DNA. We purchased an shRNA target set with 5 different oligonucleotide sequences and each sequence targeted a different site of the Sp1 sequence. Our first goal was to determine which sequence of the target set was most efficient in knocking down the Sp1 gene (targeted gene silencing). Next, a lentivirus was generated with the most efficient shRNA as vector plasmid. Subsequently, we transduced cells with this lentivirus containing the shRNA targeting Sp1, and as a negative control we used a lentivirus containing shRNA that did not target any known mammalian genes, which is the NTC (non-targeting control). After the cells were transduced and clonally selected with puromycin, the cells were infected with a pseudotyped HIV virus. Two different viral particles were used, a wild-type (IN+) and a D116N (IN-) reporter virus. The wild-type virus was generated from the packaging construct pCMVΔR8.2, for the envelope we used the vesicular stomatitis virus (VSV) glycoprotein expressed by plasmid pHCMV-G, and for the reporter gene we used a luciferase expressing plasmid that was driven by an HIV-1 LTR, the plasmid pNL-RRE-SA-Luc. The HIV-1 5’ LTR region consists of binding sites for, and is regulated by the transcription factors Sp1 and NF-κB (277,449). The D116N virus was generated with the same envelope and vector plasmid, but with a mutation in the packaging construct,
pCMVΔR8.2(D116N). D116N contains a mutation of the integrase in its active site and this affects the virus only at the step of integration (277). This mutant D116N virus served as a good system to study nonintegrated DNA forms, so that luciferase expression from the mutant population represented nonintegrated viral DNA forms. The luciferase expression from the mutant and wild-type viral DNA template made it possible to study the regulation of transcription by a specific transcription factor, Sp1, which binds to the binding sites in the HIV-1 promoter.

**Knockdown of the Sp1 gene with shRNA accomplished most efficiently by set # TRCN0000020448**

A Mission™ TRC shRNA Target Set of plasmids (pLKO.1-puro, see Figure 2.2) was purchased from Sigma-Aldrich carrying 5 different sequences targeting Sp1 (Figure 2.1), and a non-targeting control (NTC). To determine which shRNA oligonucleotide sequence of the target set most efficiently silenced or reduced gene expression, we first transfected the lentiviral vectors containing the shRNA targeting Sp1 and the NTC into HeLa cells. Forty-eight hours post transfection, the cells were harvested and counted for western blot. Equal amount of cells were lysed and loaded on a gel and western blot assay performed according protocol (see Materials and Methods). Primary antibodies targeting Sp1 and GAPDH (loading control) were used. As shown in Figure 2.6, the shRNA sequence #48 (TRCN0000020448) is the most efficient in silencing the Sp1 gene in HeLa cells. The shRNA #48 lentiviral vector will be used to produce shRNA lentivirus particles.
**Figure 2.6**

**Immunoblotting for Sp1 in transfected HeLa cells.** Transfected HeLa cells with plasmid DNA of the shRNA target set and NTC analyzed by western blotting using an anti-Sp1 antibody. The same blot was stripped and reprobed with anti-GAPDH for loading control. The set numbers of the shRNA are the last two digits of the original set numbers (Figure 2.1).
Transduction of primary T cells using shRNA lentivirus resulted in low numbers of live cells

To stably transduce cells with the shRNA lentiviral construct targeting Sp1, we used primary T cells to represent more accurately an in vivo system. Primary T cells were transduced with shRNA transduction particles that target Sp1 or used were the NTC control particles. For each transduction, 1 X 10^6 primary cells were used and transduced with 200, 300, and 400 µL of lentiviral particles with a p24 concentration of 2558 ng/mL. The cells were incubated for 24 hours followed by removing of the anti-CD3/CD28 antibody-conjugated beads that were used for pre-stimulation. Next was added 1 ug/mL puromycin for selection of stable transduced clones. The cultures were incubated for another 24 hours at 37°C and 5% CO2. After 48 hours post transduction, the cultures were observed under the microscope. A low number of cells were noticed and after culturing the cells for a couple more days, it was decided by multiple lab members and Dr. Wu that the cells were not of harvesting quality for this experiment. Considering that the negative control (NTC) also resulted in low numbers of live cells, this could indicate that the infectivity of the virus particles was not high enough to establish a population after selection of stable clones that express the shRNA. Unfortunately, no control of only primary T cells was used. In the next experiment we will use a cell line instead of primary cells.

Transduction of CEM-SS cells using shRNA lentivirus particles resulted in cell death

To be more successful in stably transducing cells with a lentivirus, we used a higher concentration of lentiviral particles and added the puromycin 48 hours post transduction for selection. Instead of primary cells, we used the cell line CEM-SS. The
CEM-SS cell line is a human T4-lymphoblastoid cell line initially derived by G.E Foley et al. and biologically cloned by P.L. Nara et al. (450). We used 2 X 10^5 CEM-SS cells for shRNA transduction and as a control, the NTC shRNA transduction particles were used. In this experiment, 100 µL of lentiviral particles were used with a p24 concentration of 13,850 ng/mL for the Sp1 shRNA lentivirus and 10,911 ng/mL for the NTC lentivirus. The high p24 values were accomplished by centrifugal concentration. For selection of stably transduced cells that carry the integrated shRNA constructs, puromycin was added 48 hours post transduction at a final concentration of 1 µg/mL. Cultures were evaluated 5 days post transduction with flow cytometry for analyzing cell viability using propidium iodide (PI). PI stains extracellular DNA and DNA in dead cells. We found low viability in the transduced CEM-SS cells as is shown in Figure 2.7. A high concentration of DNA was stained by PI staining in the shSp1 transduced sample compared to the shNTC transduced CEM-SS cells. We concluded that knocking down Sp1 in CEM-SS cells led to cell death as Sp1 regulates various cellular functions some of which are cell growth, differentiation, apoptosis, angiogenesis, and immune response (70,71,423,424). Since in this experiment, the shNTC expressing clones were viable, we used primary T cells again in the next experiment with the same concentration of lentiviral particles.

**Most optimum time for infection of transduced primary T cells was found to be 3 days post transduction**

The aim of the next experiment was to look for an optimal time to infect the transduced primary T cells. We were looking for a small window where transduced cells
Figure 2.7

Cell death in CEM-SS cells transduced with shSp1. Cell cultures were stained with Propidium Iodide (PI) and analyzed for cell viability using flow cytometry.
expressing the shRNA were still viable and suitable for infection of the viral reporter particles. In this experiment we utilized primary T cells, but the concentration of the shRNA lentiviral p24 value was increased and addition of the puromycin for selection was delayed by one day. The freshly purified primary T cells (~1 X 10^6 / transduction) were transduced with 200 µL of shSp1 lentiviral particles with a p24 concentration of 13,850 ng/mL. As control, 200 µL of the shNTC viral particles were used for transduction of the cells with a p24 concentration of 10,911 ng/mL. Also, a control sample with unmodified primary cells was established. After 24 hours incubation at 37º C and 5% CO₂, the anti-CD3/CD28 antibody-conjugated beads for pre-stimulation were removed and at this time point, the sample with unchanged primary cells were split into two samples with one sample functioning as an “added puromycin” control and the other sample as a “cells only” control. The cells were incubated for 24 hours at 37º C and 5% CO₂, and 48 hours post transduction, 1 µg/mL final concentration of puromycin was added to the appropriate cell cultures for selection. The next day, 3 days post transduction, cells were counted and 250,000 cells were collected for western blotting and lysed in 50 µL NuPAGE® LDS lysis buffer. In the “cells only” sample, very few live cells were observed by microscopy and for this reason, no cells were harvested for western blotting at this time point. The next day, 4 days post transduction, the Sp1 knockdown culture was observed under the microscope as not healthy, however, the shNTC culture was observed to be very viable. The culture “added puromycin” did not seem to contain any live cells and the culture “cells only” was observed to be not healthy. Most likely this was due to the previous splitting of the culture. Cells were collected for
Figure 2.8

**Knockdown of Sp1 in primary T cells.** Primary T cells were transduced with an shSp1 or shNTC (negative control) lentiviral particle and puromycin was added 48 hours post transduction for selection of puromycin-resistant clones. Cells were harvested 3 and 4 days post transduction for immunoblotting for Sp1 and GAPDH.
western blotting; for the Sp1 knockdown sample, about 235,000 cells (all remaining cells) and for the shNTC samples, 250,000 cells. The cells were lysed in 50 µL of NuPAGE® LDS lysis buffer. As shown in Figure 2.8, both 3 and 4 days post transduction samples demonstrated a knockdown in Sp1 gene expression compared to the control samples (shNTC). Due to the short life span of the shSp1 transduced primary cells observed in this experiment, we concluded that 3 days post transduction is the most optimal time for infection.

**Knockdown of Sp1 showed no transcriptional regulation on reporter gene expression**

Primary T cells were transduced the same way as in the previous experiment. Two days post transduction, puromycin (1 µg/mL) was added for selection of puromycin-resistant clones and 3 days post transduction, cells were infected with the IN+ reporter virus particles vHIV-Luc and IN- virus vHIV-Luc(D116N). This showed us the result of the Sp1 knockdown on the transcription regulation of the reporter gene in integrated and non-integrated species. For infection, cells were counted with the cellometer auto T4 cell counter (Nexcelom bioscience) and 5 X 10^5 cells were used per infection. The cells were subsequently infected with 45 ng of p24 input virus. The cells were washed after about 2 hours post infection to remove residual viral particles and incubated for 24 hours at 37ºC and 5% CO₂. After the incubation period, cells were counted with the cellometer auto T4 cell counter (Nexcelom bioscience), and 2 times 2 X 10^5 cells were collected from each sample and lysed in 100 µL of luciferase cell culture lysis reagent (Promega E1500). Figure 2.9 shows the result of the average of two analyzed samples. We concluded
Figure 2.9

**Knockdown of Sp1 in IN- and IN+ infected T cells.** Primary T cells (1 X 10^6) were transduced with shSp1 and shNTC lentiviral particles. Two days post transduction, puromycin was added for clonal selection. Three days post transduction; cells were infected with either an IN- or IN+ reporter virus. Luciferase activity was read from 2 X 10^5 lysed cells 1 day post infection. An average was taken from two samples. Primary cells were used from three different donors.
from this result that there is no significant difference between the reporter activities of the Sp1 knockdown cells compared to the cells transduced with the NTC.

This experiment was conducted in primary T cells of three different donors. The same infection was repeated with transduced primary T cells 4 days post transduction. The cells were incubated for 24 hours post infection and thereafter, counted and 2 samples of 2 X 10^5 cells were collected from each sample and lysed in 100 µL of luciferase cell culture lysis reagent (Promega E1500). The average was taken of the reporter activity from both samples. Primary T cells were used from 2 different donors. Figure 2.10 shows the result and here we can also conclude that the Sp1 knockdown cells compared to wild-type Sp1 expression does not show a significant difference in luciferase activity either in an IN- or IN+ infection. Figure 2.11 shows the reporter activity from induced cells that were infected 4 days post induction and the cells were harvested 2 days post infection. Primary cells of one donor were used in this experiment. The result is the average of two read samples. Due to the viability of the cells after longer shRNA induction periods, which was observed in earlier experiments, this result is less accurate. In a study it was found that preventing the binding of Sp1 to its specific binding site, altered many genes involved in the p53 signaling pathway. The overexpression of p21 caused growth arrest and the overexpression of Bak induced apoptosis (451,452).

Considering the fact that in the knockdown assay system, all the genuine cellular functions of the Sp1 protein are prevented, a new method was chosen to study the impact of transcription factors Sp1 and NF-κB on the transcription of nonintegrated viral DNA. A more analytical way of studying the impact of a transcription factor on transcription is
Figure 2.10

Knockdown of Sp1 in IN- and IN+ infected T-cells. Primary T cells (1 X 10^6) were transduced with shSp1 and shNTC lentiviral particles. Two days post transduction, puromycin was added for clonal selection. Four days post transduction, cells were infected with either an IN- or IN+ reporter virus. Luciferase activity was read from 2 X 10^5 lysed cells 1 day post infection. An average was taken from two samples. Primary cells were used from two different donors.
Figure 2.11

**Knockdown of Sp1 in IN- and IN+ infected T-cells.** Primary T cells (1 X 10^6) were transduced with shSp1 and shNTC lentiviral particles. Two days post transduction, puromycin was added for clonal selection. Four days post transduction, cells were infected with either an IN- or IN+ reporter virus. Luciferase activity was read from 2 X 10^5 lysed cells 2 day post infection. An average was taken from two samples.
to mutate the binding sites of the proteins. The HIV-1 5’ LTR consists of three binding sites for Sp1; namely I, II, and III. To be able to investigate each site separately, a mutation was designed in each binding site. A mutant version was also designed where the individual mutations were combined or where all three binding sites were deleted completely (85,448). The HIV-1 5’ LTR consists also of two NF-κB binding site; namely I and II. Here the mutations were also designed in each binding site separately, and either combined or both NF-κB binding sites were deleted (448). One designed mutation was a combination of a mutation in the Sp1 site III and the NF-κB site I (448). An overview of the mutations is shown in Figure 2.3. The different mutations were synthesized and cloned into the LTR-driven reporter plasmid, pNL-RRE-SA-Luc, and mutant viral reporter particles were constructed. To indicate that the transcription after infection originated from newly synthesized viral DNA, we treated cells with a reverse transcriptase inhibitor, etravirine (Etravirine-Intenence, TMC125, obtained from NIH AIDS Research & Reference Reagent Program).

The reverse transcriptase inhibitor etravirine has an inhibitory effect on reporter gene transcription

A pilot study was performed to obtain information on the inhibitory effect of the non-nucleoside reverse transcriptase inhibitor etravirine (ETV) on the transcription of a reporter gene. Infected were 5 X 10^5 CEM-SS cells with 45 and 18 ng of p24 input virus. The viruses used for this infection were an IN+ reporter virus: vHIV-Luc and an IN-reporter virus: vHIV-Luc(D116N). Before infection, the cells were incubated with or without 250 nM ETV for an hour at 37°C and 5% CO₂. After incubation, the cells were infected and incubated at 37°C and 5% CO₂ for 4 hours. Following the incubation period,
the cells were washed and resuspended in 2mL RPMI 1640 medium supplemented with 10% heat-inactivated FBS with or without 250 nM (final concentration) ETV and added to a 6-well plate and incubated for 48 hours at 37° C and 5% CO₂. After the incubation period, 3 different volumes, 100, 500, and 1000 µL, were taken from the culture after mixing very well by pipetting up and down. The cells were added to a centrifuge tube and washed in 1X PBS, followed by addition of 250 µL luciferase cell culture lysis reagent (Promega E1500). For reading of luciferase activity, 125 µL of lysate was added to a 96-well plate. The result is shown in Figure 2.12. The data shows a reduction up to 10.8 fold for the IN+ samples and up to 3.6 fold reduction in the IN- samples. The luciferase activity in the ETV-treated samples was most likely due to the cells containing some plasmid DNA carried over by virion particles. From this experiment we may conclude that ETV is efficient in inhibiting the function of the HIV-1 reverse transcriptase.

**No significant difference in the regulation of Sp1 and NF-κB on gene transcription between integrated and nonintegrated viral DNA**

By infecting CEM-SS cells with different mutant Sp1 and NF-κB binding site virus particles, we investigated the Sp1 and NF-κB regulation on either an IN+ or IN-LTR driven reporter expression. ETV was used to show that gene expression originated from newly synthesized DNA and a wild-type reporter virus was used as positive control. For infection, equal p24 values were used, and for luciferase activity reading, equal amounts of cells were lysed. A detailed description of the used viruses is discussed in the Materials and Method section of this chapter. The LTR mutant reporter viruses did include or did not include (D116N) a functional integrase in the virion. For the positive
Inhibitory effect of etravirine on reporter virus infection in CEM-SS cells. CEM-SS cells were pre-treated with ETV (250 nM) for one hour, and then infected with either IN+ reporter viral particle vHIV-Luc or IN- reporter virus vHIV-Luc(D116N). For infection, a p24 value of 45 and 18 ng was used and for the luciferase reading, cells in 1000, 500, and 100 µL of culture was lysed in 250 µL of lysis buffer. Of the lysate, 125 µL was analyzed for luciferase activity.
control, we used a wild-type LTR reporter virus that also included or excluded (D116N) functional integrase in the virion.

For this experiment, 5 X 10⁵ CEM-SS cells per infection were used. Two sample sets were utilized; either the cells were pre-treated with ETV or cells were not pre-treated with ETV before infection. The ETV-treated infection set demonstrated that the reporter gene expression occurred from newly synthesized DNA. The ETV-treated cells and the non ETV-treated cells were incubated for one hour at 37°C and 5% CO₂ before infection. From here on, the ETV-treated cells were never in the absence of the drug. For infection, equal p24 values were used and samples were incubated 4 - 5 hours after infection at 37°C and 5% CO₂. After the incubation period, the cells were washed to eliminate residual viral particles. Following the wash and the cells being resuspended in freshly made RPMI 1640 medium supplemented with 10% heat-inactivated FBS with or without 250 nM ETV, the samples were incubated at 37°C and 5% CO₂ for roughly 48 hours before samples were harvested for luciferase activity reading. The cells were counted, and equal numbers of cells were collected and washed in 1X PBS before adding 250 µL of luciferase cell culture lysis reagent (Promega E1500). From each sample, two different amounts of cells were lysed. From the lysate, 100 µL was added to 96-well plate according to a plate design and luciferase activity was detected. This experiment was repeated 3 different times. In Figure 2.13, the results of 3 individual experiments are presented. In the presented data, the wild-type LTR was set at 1.0. Overall, the data shows that there is no significant difference notable in the regulation of reporter gene transcription between
Figure 2.13

Relative activity of reporter gene regulated by modified NF-κB and Sp1 elements in the HIV-1 LTR. The contributions of Sp1 and NF-κB elements to basal transcription of integrated (IN+) and nonintegrated (IN-) viral DNA are shown. The basal expression from the wild-type LTR was set at 1.0. CEM-SS cells were infected with LTR-driven reporter particles with modified enhancers or with wild-type enhancers that either included a functional integrase or did not include a functional integrase. Open symbols represent the mutation and a line represents a deletion.
the integrated and non-integrated samples by Sp1 and NF-κB. The combination of all three mutations in the Sp1 binding sites, mutant Mut.Sp1(all), results in the lowest basal gene expression. The data also suggests that the combination of the binding sites NF-κB(I) and Sp1(III) are important in basal transcription from both the integrated and nonintegrated template (Figure 2.14). Due to the fact that viral protein Tat was not present, this data represents basal transcription. Figure 2.14 represents the mean and standard deviations from the 3 independent experiments.
Figure 2.14

The mean and standard deviation of the relative reporter activity regulated by modified NF-κB and Sp1 elements in the HIV-1 LTR. The mean and standard deviations from the 3 independent experiments depicted in Figure 2.13 are represented here.
Discussion

Viral gene expression is directed by a complex interplay between various DNA-binding factors. Protein factors bind DNA enhancers that are usually located upstream of the TATA motif (54). These interactions likely affect either the transcription initiation complexes or weaken the complexes to advance to elongation. Two of the upstream enhancers in the 5’ HIV-1 LTR are the Sp1 and NF-κB. In this study we analyzed the individual impact of both enhancers on nonintegrated transcription by mutating the 5’ HIV-1 LTR at different positions in the enhancers. The mutated LTR drove a reporter gene and was delivered by a pseudotyped lentiviral particle. However, most of the previous studies concerning enhancer impact on transcription were performed as a transfection and reporter gene assay.

In an HIV-1 infection, basal transcription is thought of as the early state of transcription when protein Tat is not yet expressed. HIV-1 displays a low basal LTR activity and in the presence of Tat, the expression of the LTR increases 100 – 500 fold (54,215). The transcription studies that were performed in this section are considered basal transcription.

It is by now well established that transcription occurs from nonintegrated viral DNA (277,288–293,378). However, this process is not understood completely. In this study we investigated the nonintegrated basal transcription by comparing the relative activity of the basal transcription from nonintegrated viral DNA to the integrated form and to compare our data to published data. In a study by Berkhout and Jeang, it was determined that basal transcription depends more on NF-κB enhancers than on the Sp1 elements and especially the NF-κB (I) site (54). This is only partially shown in our data.
Individual mutations that display the least impact on transcription are the Sp1(I) and Sp1(II) for both integrated (IN+) and nonintegrated (IN-) form. However, mutant Sp1(III) demonstrates a much lower relative activity in both forms compared to the controls, and the lowest reporter gene activity was observed in the mutant Sp1(all) where all 3 mutant Sp1 binding sites are combined. Besides, mutant NF-κB(II) does not seem to have a large effect on basal transcription in our data, for both integrated and nonintegrated form (Figure 2.13) and mutant NF-κB(I) shows a 55% gene expression compared to the control. In a 2009 study it was found that κB site I recruits mostly the active heterodimer of NF-κB, the p50-RelA/p65, while κB II is partial towards recruiting the homodimer p50-p50, which represses gene expression (448). Also demonstrated in published data is that Sp1 binding to enhancer site III and NF-κB binding (p65) (56,442) to its enhancer site I interact with each other and activate gene expression from the HIV enhancer (56).

In our data shown in Figure 2.13, we can see that indeed the combined mutant NF-κB(I)-Sp1(III) shows a much lower relative activity compared to the control with a wild-type LTR. For the integrated form, this is averaged around 37%, while the nonintegrated form averaged about 27% of the positive control (Figure 2.14). In a study by Burnett et all., it is demonstrated that the mutants Sp1(all) and NF-κB(I)-Sp1(III) did not show active gene expression in an integrated reporter vector (448). This data is comparable to our data where these two mutants (both IN+ and IN-) show the least reporter gene expression. It was also observed and confirmed in our data that NF-κB by itself is not adequate for gene expression (56). Mutants Sp1(all) and ΔSp1 both show low relative transcription activity in both the integrated form and nonintegrated form. However, the ΔSp1 sample, where
the Sp1 elements were deleted, is less effected by the absence of Sp1, which could be caused by the closer positioning of NF-κB elements to the TATA box (54). Nevertheless, there are some studies that demonstrate that NF-κB is capable of activating a reporter gene by itself (54,443). In one study, the two NF-κB elements by themselves produced a higher activation compared to both NF-κB elements combined with the three Sp1 motifs (54). In another study, they concluded that P-TEFb is an important co-activator of NF-κB and that the active NF-κB subunit, RelA/p65, recruits P-TEFb to the transcriptional apparatus and stimulates elongation of reporter transcription (Tat-independent) (443). It is well established that viral protein Tat recruits the P-TEFb to the paused RNA Polymerase II to assist in transcription elongation by hyper-phosphorylating the CTD of the RNA Polymerase II (Tat-dependent).

Overall, the presented data in this chapter shows that basal transcription from the integrated and nonintegrated viral template is regulated by Sp1 and NF-κB in a similar fashion. The combination of all three mutant Sp1 elements, mutant Mut.Sp1(all), resulted in the lowest basal gene expression. The data also suggests that the combination of the binding sites NF-κB(I) and Sp1(III) are important in basal transcription from both the integrated and nonintegrated template.
**Future Studies**

The studies discussed in this chapter involved analysis of regulation of nonintegrated basal transcription by the upstream enhancers Sp1 and NF-κB by modifying the LTR. The next step is to examine the regulation of Tat-induced transcription by the two enhancers. It is possible that Tat-induced transcription of the nonintegrated DNA is not regulated in the same manner as the integrated form. In published data, the Sp1 upstream elements in integrated Tat-induced transcription were found to be more important compared to NF-κB elements (54). Interestingly, it was also observed that an HIV-1 LTR missing the upstream enhancers was adequate for basal transcription, but not for Tat-induced transcription (54). The follow-up on the study described in this chapter, involves analyzing the nonintegrated DNA transcription in the presence of Tat. The Tet-On system, described in Chapter Three, will be a potential good strategy to express protein Tat in an IN+ and IN- infection with an LTR-driven reporter viral particle, with or without modified enhancers.
CHAPTER THREE

The Role of Tat in Preintegration Transcription.

Introduction
In the early phase of an HIV-1 infection, the reverse transcribed viral DNA is translocated to the nucleus and there the viral genome has two destinies, either to be integrated into the host genome or stay nonintegrated in the nucleus. For an effective HIV life cycle, integration of the HIV genome in the host chromatin is an essential step, while nonintegrated viral DNA is still seen as a dead end. However, in more recent studies it is established that transcription does occur from nonintegrated viral DNA (288,290,293,453). In proliferating T cells, the nonintegrated viral DNA is observed transiently (290) and preintegration transcription peaks at about 12 hours post infection and is abolished within 3 days due to T cell division (454,455). Nevertheless, nonintegrated viral DNA was the predominant form found in the asymptomatic phase of the infection (44).

Wu and Marsh were the first pioneers to believe that preintegrated viral DNA may have an exclusive role in the early phase of the HIV-1 life cycle (290). They studied quiescent T cells, which do not sustain viral replication, but are the most abundant circulating T cells. It was found that in infected quiescent cells, the most detectable transcript was nef (288,456) and to a lesser degree, tat. The same transcription pattern
was found in an HIV-1 infection with a non-functional integrase. Also, five days post infection, protein Nef was detected by western analysis in an infection both with and without a functional integrase in a similar quantity and it was established that this early-produced Nef aids in the activation state of infected quiescent cells. Protein Tat could not be detected by the western blotting method (288). However, that does not necessarily mean that Tat is not produced. Subsequently, Wu and Marsh established that an infection with a non-functional integrase produced all forms of transcripts, and the pattern of the early transcription, up to 12 hours post infection, is comparable to the early transcription in a wild type HIV infection. Integration of the HIV genome was found at 14 hours post infection and, with this in mind, the early transcription from both infections must have come from nonintegrated viral DNA (290). Proposed was that the transcription from nonintegrated viral DNA not only occurs in an infection where integration is blocked, but is a normal event in the early phase of the infection. In a novel, interesting study by the Levy group, it was presented that nonintegrated viral DNA is competent to produce viral particles (343). In this study, it was also found that virion associated protein Vpr was fundamental for nonintegrated HIV-1 DNA gene expression and for virus production. The group demonstrated that the virion associated Vpr was adequate for the early gene expression from the nonintegrated viral DNA (343). This gives us another reason to believe that nonintegrated HIV-1 DNA has a practical use in the viral life cycle and for these reasons we wished to further study nonintegrated viral DNA.

The mutant integrase D116N consists of a single point mutation causing an amino acid (aa) substitution at aa 116 from aspartic acid to asparagine in the integrase catalytic
domain. It has been shown that this mutation effectively inhibits viral DNA integration, but does not disturb functions such as reverse transcription and nuclear targeting (277, 378). In an earlier study by our lab, it was confirmed by Alu-PCR amplification that no integration in a D116N infection occurred (342). It has been established that viruses produced from an HIV-1 genome with the integrase mutation D116N, are proficient in transactivating the expression of an indicator gene that is under the control of the HIV-1 LTR (277, 378). This demonstration suggests that protein Tat is expressed from the nonintegrated viral DNA.

Studying transcription from nonintegrated DNA is complicated in proliferating cells, such as cell lines or activated primary T cells. The nonintegrated HIV DNA is diluted out of the cell population and transcription from these nonintegrated templates is perceived as a transient process (290). Nevertheless, macrophages are non-dividing cells, but metabolically active, and are seen as an earlier target of HIV-1 (334). Macrophages are more accommodating to the HIV-1 virus by possessing more protection to the cytopathic effects of the infection and being adept at circumventing the immune system as compared to infected T cells (334). It was also found that macrophages contain large amounts of nonintegrated viral DNA. In some cases, nonintegrated DNA exceeded integrated DNA 10 times (41).

In cells, the 2-meter long human genome is firmly arranged and organized in a chromatin structure during mitosis. Chromatin consists of repeating units of nucleosomes. A nucleosome can be dissected into 8 core histones, two of each H2A, H2B, H3 and H4. Draped around the histone octamer are 145 - 147 base pairs of DNA to form a
An H1 linker histone connects two adjacent nucleosomes with the help of linker DNA and allows the 10-nm nucleosomal arrangement to form a 30-nm fiber that is capable of folding into higher-order structures (457–459). The core histone tails refer to the N-terminus of the core histones that protrude from the nucleosome (460). The tails are post-translationally modified by phosphorylation, methylation and acetylation which impacts the interactions of the tails and therefore affects the state of the chromatin fiber (461–464). For example, acetylation of the lysine residues within the N-terminal tails neutralizes the positively charged lysine residues, and as a consequence, their affinity for the negatively charged backbone of the DNA decreases. This causes the nucleosome to unfold and change the condition of the chromatin (464). Acetylation of histones by histone acetyltransferases (HATs) are well studied and their counterpart histone deacetylases (HDACs) as well. The collaboration between HATs and HDACs play a regulatory role in gene expression. H3 and H4 acetylation by HATs translates into loosening of the highly packed chromatin structure and restoring accessibility of proteins and enzymes to the DNA. HDACs counteract HATs and cause the chromatin to tighten back up by removing acetyl groups [reviewed in (465)]. The ATP-dependent chromatin-remodeling complexes consist of a helicase/ATPase domain. The human SWI/SNF complex includes either the BRG1 or hBRM as ATPase subunit and is involved in nucleosome remodeling by modifying the interaction between histones and DNA in an ATP-dependent process (134).

Following infection, if possible, proviral HIV-1 DNA gets integrated into the host genome. Consequently, at specific locations in the 5’ LTR, nucleosomes are positioned
and this event is independent of the location of the integration site (130,466,467). In basal conditions, the viral 5’ LTR encompasses two exactly situated nucleosomes that orchestrate the HIV-1 gene expression. Nucleosome 0 (nuc-0) is positioned at nt 40 - 200 (with the transcription start site at nt 455), nucleosome-1 (nuc-1) at nt 465 – 610, and a nucleosome-free region in between (nt 200 - 452/465). This nucleosome-free region consists of the core promoter which includes the DNA sequence of the TATA box and the Sp1 binding sites. This nucleosome free region also includes the enhancer region where the two tandem NF-κB elements are located. A second nucleosome-free region was found downstream of the initiation site at nt 610 - 720 and comprises the primer binding site, but no indication of a regulatory function was observed. Due to its location, nuc-1 is the most essential in transcriptional regulation. In a suppressive chromatin, nuc-1 needs to be remodeled before transcription activation can take place (130,467).

After activation of an HIV-1 infected CD4+ cell, cellular acetyl-transferases are recruited to the viral promoter, such as p300/CBP, p300/CBP-associated factor (P/CAF) and GCN5, to acetylate histones H3 and H4 of the nucleosomes associated to the 5’ LTR (132). The adaptors p300 and CBP are very similar in DNA sequence and function and therefore referred to together as p300/CBP. They are coactivators for many different transcription factors (468,469). It has been found that these adaptors are also independent histone acetyltransferases (470,471). Viral protein Tat was shown to form an intracellular multiprotein complex with p300 and P/CAF, two HATs. However, Tat most likely recruited P/CAF as the acetyl transferase, while p300 functioned as an adapter protein to aid in the activation of integrated LTRs. Surprisingly, P/CAF and p300 were not involved
in activating expression of nonintegrated templates (141). In that particular study they both used a stably integrated LTR driven CAT reporter construct in HeLa cells or HeLa cells transiently transfected with a LTR-driven CAT reporter plasmid (141). In a study a couple of years later, it was found that the presence of two HATs, CBP and P/CAF, at the promoter was related to Tat induction. Interestingly, P/CAF was found to be responsive to Tat activation and not to phorbol ester TPA (PMA) activation (132). TPA activates protein kinase C (PKC), which leads to NF-κB activation (472). It was also shown that Tat recruits p300 and CBP to the LTR and increases the Tat transactivation level of the integrated LTR (140). Tat was also found to recruit and interact with the ATP-dependent chromatin-remodeling complex, SWI/SNF. Tat interacts directly with the DNA-dependent ATPase subunit hBRM of SWI/SNF. This interaction was found to be necessary for a Tat-mediated transactivation of the HIV-1 promoter of an integrated provirus (136,137).

HDACs are a group of enzymes that remove acetyl groups from the tails of the core histones that make up the nucleosomes. This modification to the N-terminal of the tails causes the chromatin to tighten back up. However, over 50 non-histone proteins are currently classified as substrates for HDACs. A more suitable label for the enzymes would be “lysine deacetylases” (473). HDACs are grouped according to their homology to yeast protein into three main classes: Class I, II (IIA and IIB), and IV. Class I consists of HDAC-1, 2, 3, and 8. Class II comprises of HDAC-4, 5, 7, 9 (IIA), 6, and 10 (IIB), and class IV constitutes HDAC-11 (474). It was revealed that HDAC-1, 2 and 3 play a role in HIV-1 transcriptional regulation (98,319). The transcription factors YY1 (146),

181
NF-κB p50 homodimer (98), AP-4 (475), CRIP2 (476), Sp1, c-Myc (147), and CBF-1 (477) are known to recruit class I HDACs to the HIV-1 LTR and repress viral transcription. HDAC inhibitors (HDACis) disrupt the recruitment of HDACs to the LTR and this results in LTR activation. HDACis consist of structurally different compounds used as anticancer drugs. Some of the structural divisions are hydroxamates, cyclic peptides, aliphatic acids, and benzamides [for review (474)]. In transformed cells, HDACis can cause various phenotypes, such as growth arrest, differentiation and/or apoptosis (478), activation of the extrinsic and intrinsic apoptotic pathway that lead to cell death (473,479), autophagic cell death, reactive oxygen species (ROS)-induced cell death, mitotic cell death and senescence (473). On the other hand, non-transformed cells are more resistant to HDACis-induced abnormalities (473,474).

Following integration of the HIV DNA in the host genome, a small portion of the HIV-1 infected CD4+ cells turn into a latent state where transcription is blocked in replication competent viruses. Highly active anti-retroviral therapy (HAART) is very effective in suppressing viral replication, but is not capable of entirely eradicating the disease due to this latency (303,305,308). The half-life of the latent reservoir was found to be about 44 months. With a latent reservoir of 1 \times 10^5 cells, it will take about 60 years for the virus to be eliminated from the body while the patient is on HAART (306). To have a better chance to eradicate the disease, reactivating of the latent reservoir might be a potential solution. One strategy that is pursued in earlier studies is the activation of T cells. For example HAART in combination with interleukin (IL)-2 or anti-CD3 antibodies (OKT3) therapy was used in an effort to purge the latent viral reservoir (480–
However, these approaches were not successful and might have increased the number of cells that were susceptible to infection. The result of the increase in infected cells could not be contained with HAART (486). Another approach to reactivating latent reservoirs is the use of HDACis. The drugs valproic acid (VPA) and vorinostat (suberoylanilide hydroxamic acid) are two FDA-approved HDACis. Vorinostat is used to treat cutaneous T cell lymphoma and binds to the catalytic domain of HDACs. This drug activates the phosphatidylinositol-3-kinase (PI3K)/Akt pathway. The PI3K/Akt pathway leads to the phosphorylation of HEXIM1 and as a result, the active P-TEFb is released (487) and increases transcription from the HIV-1 LTR (488). VPA is used to treat seizures (epilepsy) and is already administered to HIV patients for treatment of other conditions (489). Both drugs have been tested in primary cell systems and have shown to be successful in reactivating the latent viral reservoir in these systems (488). VPA is the only HDACi that is actually tested in several clinical trials in combination with HAART in HIV-1 infected subjects (490–494), but no significant reduction of the latent virus pool was observed. This outcome might be due to the fact that VPA is a weak, non-specific inhibitor and does not particularly target HIV-1 associated HDACs with HDAC-1 and 2 the most important (495). Recently, a study was performed with various HDACis in primary cell systems whereas most drugs were studied earlier in cell lines. The study concluded that successful use of HDACis depends on extended treatment and that the result, the reactivation of the latent virus, is reversible. The virus can easily return back to latency after ending the treatment (496). Overall, this study sees potential in HDACis to reduce or eliminate the reservoir of latent virus.
In a 2009 study it was observed that not only integrated proviral DNA is assembled into chromatin, but researchers demonstrated that nonintegrated HIV-1 DNA was also organized in a chromatin structure that regulates gene expression. Remodeling of the structure by histone modification was required for gene expression shown by a HDACi study (497). In the following study we explored a potential function of early transcribed viral protein Tat from the nonintegrated viral DNA. The pre-integration Tat may be involved in antagonizing repressive chromatin to maintain persistent transcription.
Materials and Methods

Plasmids

The HIV-1 envelope mutant plasmid pNL4-3(KFS) was kindly donated by Eric Freed (498). The double mutant plasmid pNL4-3(KFS)(D116N), was constructed by introducing a point mutation (Asp 116 to Asn) into the integrase catalytic domain of pNL4-3(KFS) as previously described (277).

Plasmids pCMVΔR8.2, pCMVΔR8.2(D116N), and pHCMV-G are described in detail in Chapter One. Plasmids pNL-eGFP(pEV658) and pNL-TAT-IRES-eGFP(pEV731) were kindly donated by Dr. E. Verdin from the Gladstone Institute of Virology and Immunology in San Francisco, California. pNL-TAT-IRES-eGFP(pEV731) is an LTR-driven expression plasmid that contains the two exon Tat (Flag) and an internal ribosome entry site (IRES)-driven enhanced green fluorescent protein (GFP). Plasmid pNL-GFP(pEV658) lacks the tat gene and is a chimeric 5’ LTR-driven plasmid. To make the plasmid Tat independent, the HIV U3 region was replaced with the Rous Sarcoma Virus (RSV) U3 sequence (499).

Plasmid pNL-RRE-SA-Luc is described in detail in Chapter One. Plasmid pNL-Tat-IRES-Luc was generated from the plasmids pNL-TAT-IRES-eGFP(pEV731) and pNL-RRE-SA-Luc. A PCR product was produced from plasmid pNL-TAT-IRES-eGFP(pEV731) that included the tat gene and IRES and two external EcoRI digestion sites. The EcoRI target sequence was added to both the forward and reverse primers. The name and sequence of the forward primer is: Tat-IRES-EcoRI-F – ATT AGA ATT CAT GGA GCC AGT AGA TC, and the name and sequence of the reverse primer is: Tat-IRES-EcoRI-R – CTA CGA ATT CTG TGG CCA TAT TAT CAT (Invitrogen, Life
technologies – Carlsbad, California). The PCR product and plasmid pNL-RRE-SA-Luc were both digested with EcoRI and the insert generated from the PCR product, which included the Tat-IRES sequence, was ligated just upstream of the luciferase gene of plasmid pNL-RRE-SA-Luc.

The plasmids pTRE-Tat86 and pTet-On were kindly donated by Dr. Chauhan from the University of South Carolina, School of Medicine. The plasmid pTRE-Tat86 is the response plasmid in the Tet-On system. The HeLa-Tet-On cell line is stably transfected with the regulatory plasmid that expresses the tetracycline-controlled transcriptional activator (rtTA). This transactivator only binds to the tetracycline response element of the response plasmid (pTRE-Tat86) in the presence of doxycycline (tetracycline was replaced by doxycycline in this system) and only then the tat gene will be expressed (Figure 3.1). The expressed viral protein Tat is the 86 amino acids long protein version.

Viruses

The virus vNL4-3(KFS) was constructed by co-transfection of the envelope mutant packaging construct pNL4-3(KFS) and the envelope construct pHCMV-G in a 1:1 ratio into HEK293T cells. Virus vNL4-3(KFS)(D116N) was produced by co-transfecting envelope and integrase mutant packaging plasmid pNL4-3(KFS)(D116N) and envelope plasmid pHCMV-G in the ratio of 1:1 into HEK293T cells.

The viruses vNL-Tat-IRES-eGFP(D116N), vNL-Tat-IRES-eGFP, vNL-eGFP(D116N), and vNL-eGFP were generated by co-transfection of either the packaging plasmid pCMVΔR8.2(D116N) or pCMVΔR8.2, reporter vector plasmids
Figure 3.1

**Tet-On System.** (Adapted from: Clontech Laboratories, Inc. – Mountain View, California) In the Tet-On system, the rtTA (tetracycline-controlled transcriptional activator) binds to the TRE (tetracycline response element). In this system the tetracycline was replaced by doxycycline. In the presence of doxycycline, transcription activity takes place and protein Tat is produced.
pNL-TAT-IRES-eGFP(pEV731) or pNL-eGFP(pEV658), and the envelope plasmid pHCMV-G at a ratio of 3:4:1 respectively (see Table 3-1).

Table 3-1
Co-transfection per 10 centimeter diameter petri dish

<table>
<thead>
<tr>
<th>Name of virus:</th>
<th>pCMVΔR8.2</th>
<th>pCMVΔR8.2(D116N)</th>
<th>pNL-TAT-IRES-eGFP(pEV731)</th>
<th>pNL-eGFP(pEV658)</th>
<th>pHCMV-G</th>
</tr>
</thead>
<tbody>
<tr>
<td>vNL-Tat-IRES-eGFP</td>
<td>7.5</td>
<td>10</td>
<td>2.5</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>vNL-Tat-IRES-eGFP(D116N)</td>
<td>7.5</td>
<td>10</td>
<td>2.5</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>vNL-eGFP</td>
<td>7.5</td>
<td>10</td>
<td>2.5</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>vNL-eGFP(D116N)</td>
<td>7.5</td>
<td>10</td>
<td>2.5</td>
<td>2.5</td>
<td></td>
</tr>
</tbody>
</table>

Virus vNL-Luc and virus vNL-Luc(D116N) were constructed by co-transfection of HEK293T cells with either the packaging plasmid pCMVΔR8.2 or the integrase mutant packaging plasmid pCMVΔR8.2(D116N), the reporter vector plasmid pNL-RRE-SA-Luc, and the envelope plasmid pHCMV-G in a 3:4:1 ratio respectively.

The viruses vNL-Tat-IRES-Luc and vNL-Tat-IRES-Luc(D116N) were generated from either the packaging plasmid pCMVΔR8.2 or the integrase mutant packaging plasmid pCMVΔR8.2(D116N), the reporter vector plasmid pNL-Tat-IRES-Luc, and the
envelope plasmid pHCMV-G in a 3:4:1 ratio respectively by co-transfection of HEK293T cells.

**Cells**

The cell line 1G5 (NIH AIDS Reference and Reagent Program, catalog # 1819) was derived from a human T cell line (Jurkat cells) that is stably transfected with an HIV-LTR-luciferase construct. This cell line was selected for a low basal level luciferase activity, HIV infectability, and high responsiveness to Tat expression (389). This suspension cell line was cultured in RPMI 1640 medium supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 mg/mL) at 37°C and 5% CO₂.

The CEM-SS cell line was cultured in RPMI 1640 medium supplemented with 10% heat-inactivated FBS and kept in a growth stage between 2 X 10⁵ and 8 X 10⁵ cells per mL at 37°C and 5% CO₂. The cell lines U2-OS-Luc Tet-On and HeLa Tet-On Advanced were purchased from Clontech Laboratories, Inc. (Mountain View, California), and were cultured in DMEM medium supplemented with 10% Tet System Approved Fetal Bovine Serum (Clontech Laboratories, Inc. - Mountain View, California). To maintain the stable transfected cell line HeLa Tet-On Advanced, G418 (100 µg/mL medium) was added to the DMEM medium. The cells were incubated at 37°C and 5% CO₂.

**Luciferase assay**

After infection, the cells were incubated for 48 hours at 37°C and 5% CO₂. Either equal numbers of cells or equal cell culture volume was used for the luciferase assay. The
cells were collected and spun down at 12,000 rpm for 1 minute and the pellet was
resuspended in 1 mL 1X PBS. The cells were again spun down at 12,000 rpm for 1
minute and supernatant was removed without disturbing the pellet. The cells were lysed
in 100 – 500 µL 1X reporter lysis buffer (Promega, E397A – Madison, Wisconsin) or 1X
cell culture lysis reagent (Promega, E153A - Madison, Wisconsin). Either 100 or 150 µL
of sample was pipetted to each well of a 96-well plate. After luciferase activity was read
(Promega Glomax Multi Detection System), the data was transferred to a USB memory
stick and analyzed in Excel. This assay is described in more detail in Chapter One.

**Concentration of virus**

After the petri dish containing the transfected HEK293T cells were incubated at
37° C and 5% CO₂ for about 48 hours, the supernatant was collected and stored at 4° C.
Right away, new, freshly made DMEM medium supplemented with 10% heat-inactivated
FBS, penicillin (50 U/mL), and streptomycin (50 mg/mL) was added back to the dish.
The petri dish was incubated for another 24 hours at 37° C and 5% CO₂. After the
incubation period, the supernatant was collected and was combined with the viral
supernatant collected the day before. Any cells and cell debris were pelleted by
centrifugation at 1500 rpm for 5 minutes. The supernatant was carefully decanted into a
new conical and stored on ice. For the concentration procedure, the Vivaspin 20
centrifugal concentrator (Sartorius Stedim North America Inc. - Bohemia, New York)
with a 100,000 MWCO (molecular weight cut off) PES (polyethersulfone) membrane
was used. The centrifuge was cooled down to 4° C and up to 20 mL of the viral
supernatant was added to the concentrator and centrifuged for 20 minutes at 5,000 g at 4°
C. Depending on the volume of the viral mixture, this step might have been repeated multiple times till all the collected supernatant was used and the desired volume was reached in the upper chamber of the concentrator. The flow-through was discarded. The concentrated viral mixture was collected from the upper chamber by pipetting up and down over the membrane and subsequently added to a new, labeled tube that was placed on ice. Afterwards, the viral mixture was filtered by adding the solution to a sterile syringe (10 mL luer-lok™ tip, Becton, Dickenson and Company – Franklin Lakes, New Jersey) that was attached to a 0.45 µm pore size nitrocellulose membrane filter unit (Millex-HA, Millapore Corporation – Billerica, Massachusetts). The syringe plunger was squeezed carefully until all the supernatant was filtered. From the sterile viral supernatant, 5 µL was added to 45 µL of p24 ELISA lysis buffer (a 10-fold dilution) used for measuring the p24 value. The remaining solution was aliquoted on ice in 500 and 250 µL aliquots and stored at -80°C.

**Luciferase induction of U2-OS-Luc Tet-On control cell line**

In a 6-well plate, 5 X 10⁴ U2-OS-Luc Tet-On cells (Clontech Laboratories, Inc. 631123 - Mountain View, California) were seeded in 5 mL of DMEM medium with addition of 10% Tet-On Approved FBS. The U2-OS-Luc Tet-On is a double-stable Tet-On/pTRE-Luc premade cell line. The Tet-On plasmid expresses the transcriptional activator that will activate expression from the pTRE-Luc plasmid only in the presence of doxycycline. Tetracycline is replaced by doxycycline in this system. Doxycycline was found to have a longer half-life and lower concentrations can be used for optimal performance compared to tetracycline (500). In the presence of doxycycline, the reporter
luciferase is induced. Different concentrations of doxycycline (100, 300, 500, 800, and 1000 ng/mL) were added to each well with seeded cells. The cells were incubated at 37\(^\circ\) C and 5\% CO\(_2\) for about 48 hours. Thereafter, cells were trypsinized, harvested, washed and spun down at 12,000 rpm for 1 minute. After supernatant was carefully removed without disturbing the pellet, the cells were lysed in 100 \(\mu\)L 1X cell culture lysis reagent (Promega, E153A - Madison, Wisconsin) and pipetted on a 96-well plate and luciferase activity was measured (Promega Glomax Multi Detection System). The data was transferred to a USB memory stick and analyzed in Excel.

**Check for doxycycline inducible expression of Tat**

In a 6-well plate, 5 \(\times\) 10\(^5\) HeLa Tet-On cells were seeded in 2 mL of DMEM medium with addition of 10\% Tet-On Approved FBS. For the selection of cells that were stably transfected with the regulatory plasmid, pTet-On, G418 (100 \(\mu\)g/mL) was added to the seeded HeLa Tet-On cells. The cells were incubated for about 24 hours at 37\(^\circ\) C and 5\% CO\(_2\). After the incubation, the cells were transfected with 2, 1, and 0.5 \(\mu\)g of the plasmid pTRE-Tat86 or 2 \(\mu\)g of plasmid pcDNA3.1+ (empty vector). The transfected cells were incubated for 5 - 6 hours. Thereafter, the medium was replaced with DMEM with added 10\% Tet-On Approved FBS, penicillin (50 U/mL), and streptomycin (50 mg/mL). Twenty-four hours post transfection, 1000 ng/mL of doxycycline was added to selected wells and cells incubated for another 24 hours at 37\(^\circ\) C and 5\% CO\(_2\). Afterwards, the cells were trypsinized with 500 \(\mu\)L 0.25\% Trypsin-EDTA (1X) Phenol Red (Gibco, Life Technologies - Grand Island, New York) and incubated for 5 minutes at 37\(^\circ\) C. Next, 3 mL of DMEM supplemented with 10\% Tet-On Approved FBS was added to the
trypsinized cells and by pipetting up and down the cells were de-clumped. The total amount of cells was added to a 15 mL conical and was pelleted at 1200 rpm for 5 minutes. The supernatant was decanted and 1 mL of 1X PBS was added to wash the cells. The cells were pelleted at 1200 rpm for 5 minutes, supernatant decanted, and 350 μL lysis buffer was added for extracting DNA/RNA (Promega, SV Total RNA Isolation Kit - Madison, Wisconsin). RNA was extracted as recommended by the manufacturer. For RNA reverse transcription, the M-MLV reverse transcriptase (Applied Biosystems, Ambion - Grand Island, New York) was used as recommended by the manufacturer with equal RNA (200 ng) and random decamers (50 μM). The cDNA was used in the quantitative real-time PCR. The Bio-Rad iQ5 real-time PCR detection system was used for the quantitative real-time PCR and the TaqMan Gene Expression Master Mix (Applied Biosystems, Life Technology, 4369016 - Carlsbad, California) was utilized as reagent mix. The used primers and probe were: the forward primer, Tat129F 5’-AGG CAT CTC CTA TGG CAG GAA GAA -3’; the reverse primer, Tat227R 5’- TGG GAG GTG GGT TGT CTT GAT AGA -3’; and the probe, FAM/Tat171P 5’- (56-FAM) – ACC TCC TCA AGG CAG TCA GAC TCA A – (3BHQ_1) – 3’. The DNA standard used for relative Tat copy numbers was constructed by using the pTRE-Tat86 plasmid containing the 86 amino acid long Tat from 10 to 10⁷ copies.

**Time-course Tet-On assay**

In a 6-well plate, 5 X 10⁵ HeLa Tet-On cells were seeded in 2 mL of DMEM medium with addition of 10% Tet-On Approved FBS. For the selection of cells that were stably transfected with the regulatory plasmid, pTet-On, G418 (100 μg/mL) was added to
the seeded HeLa Tet-On cells. The cells were incubated for about 24 hours at 37° C and 5% CO₂. After the incubation period, the cells were either transfected with 2 µg of the plasmid pTRE-Tat86 or 2 µg of plasmid pcDNA3.1+ (empty vector) or cells were untreated. The transfected cells were incubated for 5 - 6 hours. Thereafter, the medium was replaced with DMEM with added 10% Tet-On Approved FBS, penicillin (50 U/mL), and streptomycin (50 mg/mL). Twenty-four hours post transfection, the cells were infected with 45 ng (p24 value) of the IN- virus HIVΔtat-Luc(D116N)-pTat72 or the IN+ virus HIVΔtat-Luc-pTat72 (see Table 1-2) or cells were not infected. Two hours post infection, the medium was removed and cells were washed with 3 mL of DMEM with added 10% Tet-On Approved FBS. The medium was removed and 2 mL of freshly made DMEM with added 10% Tet-On Approved FBS was added to each well and to selected wells was additionally 1000 ng/mL of doxycycline added. The cells were incubated for 24 hours at 37° C and 5% CO₂. After the incubation period, the cells were trypsinized with 500 µL 0.25% Trypsin-EDTA (1X) Phenol Red (Gibco, Life Technologies - Grand Island, New York) and incubated for 5 minutes at 37° C. Next, 3 mL of DMEM supplemented with 10% Tet-On Approved FBS and with or without doxycycline was added to the trypsinized cells and by pipetting up and down the cells were de-clumped. Of the cell mixture, a total volume of 1.75 mL was added to each well of a new 6-well plate with the addition of 250 µL of DMEM with added 10% Tet-On Approved FBS and with or without doxycycline. The remaining 1.75 mL of the cell mixture was added to a clean micro centrifuge tube and pelleted at 12,000 rpm for 2 minutes. The supernatant was removed and cells resuspended in remaining liquid. Cells were washed in 2 mL of
1X PBS and pelleted at 12,000 rpm for 2 minutes. Thereafter, supernatant was removed without disturbing the pellet and cells were lysed in 330 µL 1X cell culture lysis reagent (Promega, E153A - Madison, Wisconsin). The lysate was stored at -80°C. The 6-well plates containing the other half of the cells were placed back in the incubator at 37°C and 5% CO2 for 24 hours. After the incubation period, half of the cells were harvested and lysed as described above and lysate stored at -80°C while the other half was added back to a new 6-well plate and incubated for 24 hours at 37°C and 5% CO2. After the incubation period, again half of the cells were harvested and lysed as described above and lysate stored at -80°C and the other half of the cells was discarded. The next day, samples were thawed and 3 times 100 µL per sample was pipetted in individual wells of a 96-well plate and luciferase activity was read.

**Spin-inoculation**
Viral particles were added to the cells in a 5 mL round bottom falcon tube. Next, the sample tubes were arranged into swing buckets and subjected to centrifugation for 2 hours at 600 g (1800 rpm). After centrifugation, the medium was replaced with 1 mL of freshly made RPMI 1640 medium supplemented with 10% heat-inactivated FBS and incubated at 37°C and 5% CO2 for about 48 hours.

**Flow cytometry assay**
For this assay, we started with 2 X 10^5 CEM-SS cells that were in their growth stage, between ~ 2 X 10^5 and 8 X 10^5 cells/mL. The cells were pelleted and resuspended in 100 µL RPMI 1640 supplemented with 10% heat-inactivated FBS in a 5 mL round bottom Falcon tube. The cells were infected with equal p24 input virus (5000 pg). Used
for the infection were the viruses vNL-Tat-IRES-eGFP(D116N), vNL-Tat-IRES-eGFP, vNL-eGFP(D116N), and vNL-eGFP. Cells and viral particles were incubated for 4 hours at 37ºC and 5% CO₂. Thereafter, the cells were washed once with medium and resuspended in 1 mL freshly made RPMI 1640 supplemented with 10% heat-inactivated FBS. Two days post infection, either 400 or 200 µL of cell mixture with an addition of 200 µL of RPMI 1640 was used for propidium iodide (PI) cell staining. PI fluorescence occurs when it intercalates into double-stranded nucleic acids. However, PI only penetrates the cell membrane of dying or dead cells. For PI staining, 1 µL per 400 µL of cell mixture was used and the samples were run on a FACSCalibur (Becton, Dickenson and Company, Biosciences - San Jose, California) and GFP was examined in the viable cell population and density plots were generated.

**Culturing of PBMC-derived macrophages**

PBMCs (peripheral blood mononuclear cells) were collected from healthy donors at the Student Health Center of George Mason University (GMU), at either the Fairfax location or Manassas, Virginia. All protocols involving human subjects were reviewed and approved by the GMU Institutional Review Board (IRB). Fresh blood samples were diluted 1:1 in sterile 1X PBS. Diluted blood was carefully layered onto 15 mL lymphocyte separation medium up to a total volume of 50 mL. At room temperature and with the brakes off, the blood samples were centrifuged at 160 g (830 rpm) for 20 minutes. After centrifugation, about 20 mL of the supernatant (yellow top layer) was removed to reduce the amount of platelets in the sample. The samples were centrifuged again at room temperature with the brakes off at 350 g (1200 rpm) for another 20
minutes. Next, the PBMCs were collected from the interface and added to a 50 mL conical. The cells were washed in cold 1X PBS - 0.1% BSA (bovine serum albumin) that was added to the cells up to a total volume of 50 mL. The cells were pelleted at 400 g (1300 rpm) for 5 minutes at 4°C. Subsequently, the supernatant was poured off and the cells were resuspended in the remaining liquid. The cells were washed again in cold 1X PBS - 0.1% BSA in a total volume of 50 mL and centrifuged at 225 g (1000 rpm) for 5 minutes at 4°C. Carefully the supernatant was poured off and the pellet resuspended in the remaining fluid. RPMI 1640 with no added serum or other supplements was added to the cells up to a total volume of 50 mL. The cells were counted and between 2 X 10⁶ and 2.5 X 10⁶ PBMCs resuspended in serum-free RPMI 1640 were added to each well of a 6-well plate. The plates were incubated for an hour at 37°C and 5% CO₂ till most cells had adhered. After the incubation period and cell adhesion was confirmed by microscopy, the medium was replaced with 2 mL RPMI 1640 medium supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 µg/mL). In addition to FBS and antibacterials, recombinant macrophage colony-stimulating factor (M-CSF) (Research and Diagnostic Systems Inc., #216-MC – Minneapolis, Minnesota) was added to the purified PBMCs to a final concentration of 10 ng/mL. The cells were incubated at 37°C and 5% CO₂. Every other day, for about two weeks, 1 mL of the medium was replaced with 1 mL of freshly made RPMI 1640 medium supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), streptomycin (50 µg/mL), and M-CSF (10 ng/mL). After the two weeks of culturing, the monocyte-derived macrophages (MDMs) were ready for further use.
Infection of monocyte-derived macrophages

From the 6-well plates containing the MDMs, about 1 - 1.5 mL of the medium was removed and consequently about 500 µL of the medium remained in the well to cover the cells. The frozen virus supernatant was thawed quickly in either a 37°C heat block or directly in the hand. Immediately, appropriate amount of virus (equal p24 input virus) was added to the cells, and if needed, additional medium was added to compensate for the volume differences between viral supernatant volumes. The cells were incubated at 37°C and 5% CO₂ for about 4 hours. Afterwards, the cells were washed twice with RPMI 1640 and finally, 2 mL of freshly prepared RPMI 1640 medium supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 µg/mL) was added to the cells. Every other day, 1mL of the medium was replaced with freshly made RPMI 1640 medium supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 µg/mL).

Induction of monocyte-derived macrophages

The different HDACis (histone deacetylase inhibitors) were diluted at different concentrations in RPMI 1640 medium supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 µg/mL). The total amount of medium from the well was removed and immediately replaced with 2 mL medium with the HDACi. For the wells with the control cells, the medium was replaced with RPMI 1640 medium supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 µg/mL). The cells were incubated for approximately 24 hours before 1 mL of the medium was replaced with 1 mL of freshly prepared RPMI 1640 medium supplemented with 10% heat-inactivated FBS, penicillin (50 U/mL), and streptomycin (50 µg/mL) that
either contained the HDACi or did not contain the HDACi (control cells). The cells were incubated for another 24 hours before they were ready for harvesting.

**Harvesting macrophages for luciferase assay**

The medium was removed from the wells and the cells were rinsed twice with cold 1X PBS. After the second rinse, the PBS was removed and replaced with 500 µL of 1X reporter lysis buffer (Promega, E397A – Madison, Wisconsin) or 1X cell culture lysis reagent (Promega, E153A - Madison, Wisconsin). The 6-well plates were placed on a shaker or rocker for about 10 minutes or, if needed, longer. After conformation by microscopy that most cells were lysed, the lysate was added to a labeled centrifuge tube and either used right away or stored at -80°C for later analysis for luciferase activity.

**Total RNA and DNA Isolation from macrophages**

The macrophages were lysed in the well of the 6-well plate in 500 µL of lysis buffer (SV Total RNA Isolation Kit, Promega - Madison, Wisconsin). For the total RNA and DNA isolation, 175 µL of the lysate was transferred to a new labeled 1.5 mL centrifuge tube and 350 µL of dilution buffer (SV Total RNA Isolation Kit, Promega - Madison, Wisconsin) was added. The samples were heated at 70°C in a heat block for 2 minutes and centrifuged at maximal speed for 10 minutes. The supernatant was transferred to a new labeled centrifuge tube and 150 µL of 95% ethanol was added and mixed. The samples were transferred to a spin column assembly and spun for 1 minute. The flow-through contained the RNA, while the used column contained the bound DNA. First the RNA was isolated by adding 300 µL of 95% ethanol to the flow-through and mixed. This mixture was transferred to a new spin column assembly and spun for 1
minute at maximal speed. The flow-through was discarded and 600 µL of SV RNA wash solution was added to the column and spun at maximal speed for 1 minute. For DNaseI digestion, a mixture was made of 0.09 M MnCl₂, yellow buffer, and DNaseI in a ratio of 1:8:1 respectively. Of the mixture, 50 µL was added to each column and incubated at room temperature for 15 minutes. After incubation, 200 µL of stopping buffer was added to the columns and spun for 1 minute. Afterwards, 600 µL of SV washing buffer was added to each column and spun for 1 minute and flow-through was discarded. An additional 250 µL of SV washing buffer was added to each column and spun for 1 minute, and the flow-through was discarded. In the final step, 50 µL of nuclease-free water was added to the filter without touching the filter to elute RNA. The columns were incubated for 5 minutes at room temperature and spun for 1 minute. Another 50 µL of nuclease-free water was added to the filter for eluting the remaining RNA. Again, the columns were incubated for 5 minutes at room temperature and spun for 1 minute. A small amount of RNA was diluted 10 times into nuclease free water to measure RNA concentration by the optical density at 260 nm (OD₂₆₀) using Eppendorf BioPhotometer (Eppendorf - Westbury, New York).

The columns containing the bound DNA were washed with 600 µL of 70% ethanol and spun for 1 minute. The flow-through was discarded before the addition of 400 µL of 70% ethanol for a second wash and spun for 1 minute. The flow-through was discarded. The columns were spun for an additional 1 minute to dry the columns. The DNA was eluted in 100 µL of nuclease-free water. A small sample of DNA was diluted 10 times into nuclease free water to measure the dsDNA concentration by the optical
density at 260 nm (OD$_{260}$) using Eppendorf BioPhotometer (Eppendorf - Westbury, New York).

**Reverse transcriptase reaction**

For RNA reverse transcription, the M-MLV reverse transcriptase (Applied Biosystems Ambion, Life technologies, AM2044 - Grand Island, New York) was used and as primers, random decamers (AM5722G) were utilized. Per sample, 2 µL of random decamers (50 µM) and 10 µL of RNA were mixed and added to a thermal cycler at 80°C for 3 minutes and cooled down to 4°C for 5 minutes. Thereafter, added to each sample was 2 µL of 10X complete PCR buffer (# 8707G), 4 µL dNTPs (2.5 mM - # 8228G5), 10 U RNase inhibitor (# AM2682), 100 U M-MLV reverse transcriptase, and if needed, nuclease-free water up to a total volume of 20 µL. The reaction mixture was gently mixed and fluid spun down. The mixture was incubated at 42°C in the thermal cycler for one hour. Afterwards, the thermal cycler was heated up till 92°C for 10 minutes to inactivate the reverse transcriptase. Thereafter, the thermal cycler was cooled down to 4°C and samples containing cDNA were ready for further use.

**Real-time PCR reaction**

For the real-time PCR assay, the Taqman® Gene Expression Master Mix (Applied Biosystems, Life Technology – 4369016 - Carlsbad, California) was used. The reaction mixture per sample consisted of 25 µL 2X Gene Expression Master Mix, 5 µL 10X probe, 5 µL 10X forward primers, 5 µL 10X reverse primers, and 10 µL DNA. The reaction mixture was added to BioRad optical tubes (8-tube strips), and inserted into the thermal cycler. The cycling conditions used were for the uracil-DNA glycosylase (UDG)
incubation: 2 minutes at 50°C, AmpliTaq Gold ultra pure (UP) enzyme activation: 10 minutes at 95°C, 40 PCR cycles: denature 15 seconds at 95°C and annealing/extending 1 minute at 60°C. After all cycles were completed, the thermal cycler cooled down to 4°C. A DNA standard was used and the threshold cycles (C_t) from the thermal cycler were transferred to an Excel spread sheet. A standard curve was generated from the C_t values and the log of the concentration of the DNA standard samples. A trend line was made and the equation of the trend line was used to calculate the copy numbers of the unknown samples by taking the anti-log of the result of the trend line.
Results

**Tat is produced by nonintegrated viral DNA and is functional**

It has been demonstrated in earlier studies that an integrase mutant (IN-) HIV-1 viral infection is capable of expressing an indicator protein via the transactivation of the HIV-1 LTR linked to the indicator gene (277,378). In these studies, Tat was assumed to be produced by the nonintegrated viral DNA. However, so far, a method has not been determined to physically detect Tat in IN- infected cells. In this study we would like to show that an IN- HIV-1 viral particle, vNL4-3(KFS)(D116N), is able to transactivate the expression of a reporter gene under the control of the HIV-1 LTR. The reverse transcriptase inhibitor, etravirine (ETV), was used to demonstrate that the *tat* gene expression occurred from newly synthesized DNA.

Utilized in this experiment was the 1G5 cell line, which is stably transfected with an LTR-driven luciferase construct. For each infection, 5 X 10^5 1G5 cells were used and pre-treated with ETV (250 nM final concentration) for one hour at 37° C and 5% CO₂ and the untreated cells were incubated as well for an hour. From this point on, the cells that were pre-treated with the drug were always cultured in medium with the drug present, including washes. The cells were infected with p24 values of 300, and 500 ng of the IN-virus vNL4-3(KFS)(D116N). In the preparation of this virus, no nuclease was used. The cells were incubated for 4 hours at 37° C and 5% CO₂ after the infection. After the incubation time, the cells were washed 2 times and resuspended in freshly made RPMI 1640 medium supplemented with 10% heat-inactivated FBS and either with ETV or without ETV and transferred to a 6-well plate. The plates were incubated at 37° C and 5% CO₂ for 48 hours. For evaluation of the luciferase activity, the cells were counted and
from each sample equal numbers of cells (120,000 cells) were lysed in 250 µL lysis buffer. The result was a 16-fold increase in reporter gene expression in the 500 ng p24 input virus infection without the drug ETV compared to the infection with the drug (Figure 3.2). In the infection with 300 ng input virus, observed was a 19-fold increase in the samples without the treatment of the reverse transcription inhibitor compared to the samples with the treatment (Figure 3.2). From this result, we can conclude that viral protein Tat is produced from the IN- infection with the virus vNL4-3(KFS)(D116N) and that the protein was functional in transactivating the HIV-1 LTR promoter of the reporter construct.

**Tat is able to transactivate the nonintegrated HIV-1 DNA genome**

To show that newly made (de novo) Tat is also capable of transactivating a nonintegrated HIV-1 DNA genome, we infected cells with an IN- virus particle that packaged an LTR-driven reporter genome including the tat gene through the use of an internal ribosome entry site (IRES). The control counterpart did not carry the tat gene. The other set of viruses used in this experiment were the IN+ viral particles with and without the tat gene. In this system we not only indicate the production of Tat, but also if it was capable of transactivating the nonintegrated DNA genome.

In the next experiment, 2 X 10⁵ CEM-SS cells that were in their growth stage were infected with the GFP and Tat expressing IN- virus vNL-Tat-IRES-eGFP(D116N), the GFP and Tat expressing IN+ virus vNL-Tat-IRES-eGFP, the GFP expressing IN- virus vNL-eGFP(D116N), and the GFP expressing IN+ virus vNL-eGFP at equal p24
Figure 3.2

Increase in reporter expression by an integrase mutant viral infection. Cells (1G5) that were stably transfected with an LTR-driven luciferase reporter plasmid were infected with the IN- virus vNL4-3(KFS)(D116N). Equal amounts of cells (120,000) were lysed and used for luciferase activity reading. The control group was treated with the reverse transcriptase inhibitor, etravirine, to demonstrate that protein Tat originated from newly synthesized DNA.
values (1200 pg). To increase the infection efficiency, we performed a spin-inoculation after the infection. GFP expression was evaluated from 200 µL per sample, 2, 3, 5, and 7 days post infection by flow cytometry. At 3 days post infection, 500 µL of fresh RPMI 1640 medium supplemented with 10% heat-inactivated FBS was added to each sample. The result is shown in Figure 3.3. From this experiment we can conclude that Tat is produced by an infection with an IN- viral particle and that Tat is functional. Tat was capable of transactivating the LTR promoter of the nonintegrated HIV-1 DNA genome. The cells infected with the IN- virus vNL-eGFP(D116N) and with the IN+ virus vNL-eGFP represent the basal reporter transcription, while the cells infected with the Tat expressing IN- virus vNL-Tat-IRES-eGFP(D116N) and with the Tat expressing IN+ virus vNL-Tat-IRES-eGFP show the Tat-induced transcription of the reporter genome. Observed in Figure 3.3, is up to a 49-fold increase (day 5 post infection) in GFP expression in the IN- sample with Tat expression compared to the basal expression. This result indicates that Tat is capable of transactivating a nonintegrated DNA template.

**De novo Tat involved in nonintegration transcription**

To confirm the result from the previous experiment, we investigated the involvement of the newly made Tat on nonintegrated transcription by using the Tet-On system. This system is illustrated in Figure 3.1. First a pilot experiment was performed to establish the correct amount of doxycycline to be used. The U2-OS-Luc-Tet-On cell line was cultured in medium with addition of 10% Tet system approved FBS, an FBS that is tetracycline-free (Clontech). The control cell line U2-OS-Luc-Tet-On expresses the regulatory plasmid (reverse tet-responsive transcriptional activator – rtTA) and the
Figure 3.3

Transactivation of nonintegrated and integrated HIV-1 DNA by Tat. CEM-SS cells were infected with the Tat positive IN- virus vNL-Tat-ires-eGFP(D116N) (Tat+/IN-), the Tat positive IN+ virus vNL-Tat-ires-eGFP (Tat+/IN+), Tat negative IN- virus vNL-eGFP(D116N) (IN-), and the Tat negative IN+ virus vNL-eGFP (IN+) at equal p24 values. The GFP expression was read by flow cytometry at 2, 3, 5, and 7 days post infection.
response plasmid that holds the inducible luciferase gene. Figure 3.4 displays a flow chart of the experiment. For this experiment, 5 X 10^4 of U2-OS-Luc-Tet-On cells in 5 mL of medium were used. Doxycycline was added to the medium at the concentrations of 100, 300, 500, 800, and 1000 ng/mL or cells were cultured without the presence of doxycycline. The cells were harvested for luciferase activity reading after 48 hours of incubation at 37°C and 5% CO₂. The result is shown in Figure 3.5, and from the bar graph we concluded that 1000 ng/mL of doxycycline is the appropriate dose. This dose increased the luciferase activity through induction by doxycycline over 1000-fold compared to cells without induction.

In the following experiment, we examined the response plasmid, pTRE-Tat86, which expresses Tat in the presence of doxycycline. The Tet-On HeLa cell line is stably transfected with the regulatory plasmid that expresses the transcriptional activator (rtTA). For selection of transfected HeLa cells, the selective antibiotic G418 was used for the neomycin resistance gene on the stably transfected plasmid. The HeLa cells were transfected with different amounts of response plasmid (0.5, 1, and 2 µg) and induced with 1000 ng doxycycline or transfected cells were not induced at all. As negative control, HeLa cells were transfected with 2 µg of an empty cloning vector (pcDNA3.1+). The produced Tat protein was measured by tat mRNA in equal amounts of cellular RNA. A schematic flow of this experiment is displayed in Figure 3.6. The result is shown in Figure 3.7. The relative quantity of tat transcripts are presented in the bar graph. Observed is a 6.7-fold increase in tat transcripts when 2 µg of response plasmid was used and 1000
Figure 3.4

Control cell line U2-OS-Luc-Tet-On. The U2-OS-Luc-Tet-On cell line expresses both the regulatory (rtTA) and the response plasmid (luciferase). The response plasmid expresses luciferase upon induction by doxycycline.
Figure 3.5

Pilot study for optimal doxycycline concentration. U2-OS-Luc-Tet-On cells were cultured with different concentrations of doxycycline added to the growth medium or cells were cultured without the addition of the doxycycline to the medium. After 48 hours incubation, all cells were lysed and the luciferase activity was analyzed.
**Figure 3.6**

**Induction of Tat by doxycycline.** The Tet-On HeLa cell line expresses the transcriptional activator (rtTA). The cells are transfected with the response plasmid that expresses Tat, but only in the presence of doxycycline. Doxycycline binds to the transcriptional activator and the activator undergoes a conformational change. The tetracycline response element (TRE)-bound transcriptional activator is thereupon capable of activating transcription of the *tat* gene.
Figure 3.7

**Pilot study for optimal response plasmid concentration.** A total of $5 \times 10^5$ Tet-On HeLa cells were seeded in a well of a 6-well plate and selected for the presence of the regulatory plasmid with selective antibiotic G418. After 24 hours incubation, the cells were transfected with 0.5, 1, or 2 µg of the response plasmid (pTRE-Tat86) or with 2 µg of empty cloning vector pcDNA3.1+ (pEmpty). Twenty-four hours post transfection, 1000 ng/mL doxycycline was added to appropriate samples (bar 1, 2, and 3). After 24 hours incubation, the cells were harvested and RNA extracted. Equal cellular RNA was reverse transcribed and equal volume of cDNA amplified by real-time PCR targeting the tat transcripts.
ng/mL of doxycycline for induction compared to the sample transfected with 2µg of response plasmid, but not induced by doxycycline. The system is also observed to be “leaky”, because over 22,000 relative tat transcripts are produced by non-induced cells (bar # 4 in Figure 3.7) compared to non-induced cells, transfected with an empty vector (bar # 7). It was decided to use 2 µg of the pTRE-Tat86 response plasmid in our following experiment.

Next, we examined the effect of the de novo Tat on the transcription of the reporter gene, luciferase. Tet-On HeLa cells were transfected with either the response plasmid pTRE-Tat86 or plasmid pcDNA3.1+, which is an empty vector. After transfection, the cells were infected with 45 ng of either IN+ or IN- reporter viral particles: vNL-Luc and vNL-Luc(D116N) respectively. The reporter vector is LTR-driven and contains the HIV genomic packaging signal, but no functional HIV proteins are expressed from this vector. Its counterpart, the IN- virus: vNL-Luc(D116N) was constructed with an integrase mutant packaging plasmid. Induction was performed 2 hours post infection. For a schematic work flow, see Figure 3.8. The result is shown in Figure 3.9. After collecting cells from equal volumes of cell mixtures 1, 2, and 3 days post infection, the cells were lysed and luciferase activity was measured. The average was taken from luciferase activity readings of 3 samples. In the IN+ infection, the integrated template shows a steady increase over the 3 days in luciferase expression when Tat is induced compared to the non-induced samples. Day 1, 2, and 3 show a 5.4-, 8.2-, and 12.5-fold increase in transcription respectively (Figure 3.9 – B). It is well established and confirmed that Tat-induced transcription is more productive compared to basal...
Figure 3.8

Illustration of the work flow for Tet-On Tat induction in an infection. Tet-On HeLa cells, expressing transcriptional activator (rtTA) were transfected with the response plasmid, pTRE-Tat86. Subsequently, the cells were infected with either an IN+ or IN- viral reporter particle and induced with doxycycline 2 hours post infection. The induced Tat transactivates the viral LTR promoter and luciferase activity was analyzed.
Figure 3.9

Involvement of de novo Tat in transcription from nonintegrated viral template.

Presented is a diagram of the experiment (A). Tet-On HeLa cells were transfected with 2 µg of the response plasmid, pTRE-Tat86 (B and D), or an empty vector (pcDNA3.1+) (C and E). Twenty-four hours post transfection, the cells were infected with 45 ng p24 value of either an IN+ reporter (luciferase) viral particle (vNL-Luc) (B and C) or with an IN-reporter viral particle [vNL-Luc(D116N)] (D and E). The infected cells were either induced or not induced with doxycycline (1000 ng/mL) to produce de novo Tat (B - E).
transcription of the integrated viral template. In the IN- virus infection (Figure 3.9 D and E) we observe a steady decline in induced transcription. Day 1, 2, and 3 show a 2.8-, 1.9-, and 1.6-fold increase in the Tat-induced transcription of the nonintegrated reporter template compared to basal transcription (D). The decline in transcription is most likely due to the dilution of the nonintegrated reporter template in the proliferating cells (290). This pattern is similar to our previously presented data (Figure 3.3). Still, the induced Tat has a positive effect on transcription of the nonintegrated reporter template. In Figure 3.9 C and E we presented a control experiment that examined if doxycycline by itself was able to induce an increase in transcription. We can see from the presented data that the addition of doxycycline is not stimulating transcription from the reporter gene, but has rather an inhibitory effect. Next we wanted to explore the function of de novo Tat in chromatin remodeling.

**Low level induction by HDAC inhibitors in IN- infected macrophages**

CD4 positive macrophages and microglia are natural target cells for HIV-1 (501–503). In these non-cycling, but metabolically active primary cells, the non-integrated HIV-1 genome is persistent compared to dividing cells such as activated T cells (342,504). It is demonstrated by Kantor et al. that in SupT1 cells gene expression from a nonintegrated HIV-1 genome can be induced greatly by 5 mM Sodium Butyrate (NaBut) or valproic acid (VPA). However, in a primary monocyte-derived macrophage system, they used 1 mM VPA. (497). We wished to test 4 different HDACis to see if they were able to induce a nonintegrated LTR-driven reporter gene from infected macrophages. In the first experiment, human monocyte-derived macrophages (MDMs) were used and
infected with about 108 ng p24 value of the IN- virus vNL-Luc(D116N). For the induction by HDACis the following drugs were used: Apicidin (kindly provided by Dr. Kashanchi’s lab), Oxamflatin (kindly provided by Dr. Kashanchi’s lab), Sodium Butyrate (Merck-Darmstadt, Germany), and Valproic acid (Merck-Darmstadt, Germany) at the final concentrations of 150 nM, 150 nM, 1 mM, and 1 mM respectively. The induction was performed 2 days post infection, for a total of 2 days before the cells were lysed and harvested. The cells were lysed with lysis buffer on the plate. The result is shown in Figure 3.10. The induction of the transcription of the reporter gene in the samples treated with Apicidin (Ap) and Sodium Butyrate (NaBut) resulted in a ~1.2 fold increase compared to the infected macrophages without induction. The sample treated with Oxamflatin (Ox) shows a ~1.1 fold increase, and the induction with Valproic acid (VPA) did not result in any increase in reporter expression. In the next experiment we used various concentrations of HDACis to investigate the induced reporter expression.

**HDAC inhibitors stimulate LTR activity from non-integrated viral DNA**

In the next experiment we tested 4 different concentrations of each HDACi. A reporter virus was used with luciferase as the reporter. The virus did not carry any functional viral genes and no de novo Tat was produced. The reporter expression only represents basal transcription. We used an IN- virus from which its genome (the reporter gene) behaved as a nonintegrated viral template and was LTR stimulated. The IN+ virus was used as a control. MDMs were infected with 20 ng p24 value of the IN- virus pNL-Luc(D116N) or the IN+ virus pNL-Luc. Five days post infection and for 2 days, reporter gene expression was induced by HDACis at 4 different concentrations; Apicidin (Ap)
Figure 3.10

Relative luciferase expression in MDMs stimulated with different HDACis.

Monocyte-derived macrophages (MDMs) were infected with an IN-LTR-driven reporter viral particle and induced with HDACis 2 days post infection for 2 days. Cells were lysed on the plate and equal amounts of lysate was evaluated. Ap: Apicidin, Ox: Oxamflatin, NaBut: Sodium Butyrate, VPA: Valproic acid
150, 750, 3750, and 18750 nM final concentrations, Sodium Butyrate (NaBut) 1, 5, 25, and 125 mM, and Valproic acid (VPA) 1, 5, 25, and 125 mM. The other half of the infected cells were induced 18 days post infection for a total of 2 days with the same drugs and concentrations. The cells were lysed and harvested on the plate and an equal amount of lysate was analyzed for luciferase activity. The highest concentration of the HDACis resulted in low data due to cell death that was observed by microscopy (data not shown).

As presented in Figure 3.11, the basal expression of the reporter at 5 days post infection is about 8.5 X 10⁵ relative luciferase activity for the integrated form (white bar in the upper left graph), while the nonintegrated form has a basal activity of 1.5 X 10⁴ (white bar in the upper right graph). It is well established that the nonintegrated viral DNA template demonstrates a lower level of transcriptional activity compared to its integrated counterpart (293). As seen in the lower half of Figure 3.11, 18 days post infection, the basal activity dropped to the levels of background luciferase expression. Our hypothesis was that expression is blocked by the repressive chromatin assembled not only on the integrated viral DNA, but also on the nonintegrated form. As demonstrated in Figure 3.11 (upper graphs), 5 days post infection, HDACis Apicidin and NaBut were capable of inducing gene expression from the integrated and nonintegrated viral DNA. However, VPA was not able to do so. Most likely due to its weak, non-specific inhibitory action and that it does not particularly target HIV-1 associated HDACs which are predominantly HDAC-1 and 2 (495). In the nonintegrated form, an increase was observed of up to 3.4-fold in the HDACi Apicidin with a concentration of 3.75 µM while NaBut
Figure 3.11

Stimulation of the LTR-driven reporter by induction. MDMs were infected with an IN+ and IN- luciferase reporter virus where the reporter gene is controlled by the LTR promoter. Reporter expression was induced by HDACis Apicidin, Sodium Butyrate (NaBut), and Valproic acid (VPA) at 3 different concentrations.
induces expression up to 4.0-fold compared to basal activity in the infected, non-induced cells. For the integrated control group, we see up to 2.7- and 2.0-fold increase respectively in the induced samples. In the induction 18 days post infection (lower half of Figure 3.11), we see a fold increase as high as 209 (Apicidin) in the nonintegrated samples compared to basal expression in the infected, non-induced cells. In the integrated group this fold increase is up to 5533 times. Even the weak HDACi VPA is able to induce expression after 18 days post infection in the integrated as well as in the nonintegrated samples. The induction of reporter expression by HDACis was most likely caused by counteracting the repressive chromatin that was assembled onto the viral DNA by the functioning of HDACs. Nevertheless, it is demonstrated that HDACis cause cell death in transformed cells (473, 479). HDACs not only target histones, but have many other protein substrates. When using HDACis, the functions of proteins involved in gene expression regulation, cell proliferation, and cell death are altered (473). Also, using HDACis at high concentrations can cause cell death in transduced primary resting CD4+ T cells (496). We were not able to count the viable cells, and therefore we normalized the data by investigating the quantity of the viral DNA template in the lysate. The total luciferase expression was divided by the amount of total viral DNA template in equal volume of lysate. About 56 ng of total cellular DNA was used for quantitative real-time PCR for the samples induced 5 days post infection and 37 ng for the samples induced 18 days post infection. Total HIV-1 genome (late viral DNA) was targeted for amplification. Used was the forward primer, 5’LTR-U5: 5’-AGA TCC CTC AGA CCC TTT TAG TCA
Figure 3.12

Normalized data of the induced reporter expression. The relative luciferase activity (Figure 3.11) was divided by the quantity of full length viral DNA present in the lysate.
-3’; the reverse primer, 3’ gag: 5’- TTC GCT TTC AAG TCC CTG TTC -3’; and the probe, FAM-U5/gag: 5’- (FAM) – TGT GGA AAA TCT CTA GCA GTG GCG CC – (BHQ) – 3’. Overall, the normalized data (Figure 3.12) shows the same pattern as seen in the non-normalized data (Figure 3.11).

Our hypothesis is that viral protein Tat opposes the assembly of a suppressive chromatin. To prove this hypothesis, we performed the same experiment as above, but with a wild-type virus that did generate de novo Tat.

**Stimulation by HDAC inhibitors is reduced in a wild-type infection**

We demonstrated in the previous experiment that the nonintegrated viral genome, which consisted of the reporter gene under control of the viral LTR, seemed to form a repressive chromatin configuration that could be induced by HDACis. The same was found to be true in the integrated form. In the next experiment we investigated if the same transpired in a VSV-G-pseudotyped HIV virus. This virus contained a packaged genome that was competent to produce all the viral proteins de novo, except for the envelope and in the case of the D116N virus; also a functional integrase protein was not produced.

MDMs were infected with about 20 ng p24 value of the IN- virus vNL(KFS)(D116N) and the IN+ virus vNL(KFS). For the induction, the drugs Apicidin (Sigma-Aldrich - St. Louis, Maryland), Sodium Butyrate (Merck - Darmstadt, Germany), and Valproic acid (Merck - Darmstadt, Germany) were used. The induction was performed 5 and 20 days post infection, for 2 days before the cells were lysed and harvested. The cells were lysed in 500 µL cell lysis buffer (SV Total RNA isolation System, Promega - Madison, Wisconsin) on the plate and RNA and DNA were extracted.
from equal amounts of lysate. For the reverse transcriptase reaction, we used 62 ng of total cellular RNA for the samples that were induced 5 days post infection and 57 ng of total cellular RNA for samples induced 20 days post infection. In the real-time PCR, we targeted the viral nef transcripts. We used the probe: FAM-Nef/Rev 5’-(56-FAM) CGG AGA CAG CGA CGA AGA GCT CAT C (3 BHQ_1)- 3’ (Integrated DNA Technologies - Coralville, Iowa) where 10X is 2 µM. We used primers: 5’ Nef – 5’ – GGC GGC GAC TGG AAG AA – 3’ and 3’ Rev – 5’ – AGG TGG GTT GCT TTG ATA GAG AAG – 3’ (Integrated DNA Technologies - Coralville, Iowa) where 10X is 3 µM. A quantitative standard was generated by a 10-fold dilution series of PCR-Nef copies diluted in nuclease-free water from 1 X 10^2 – 1 X 10^9 copies.

Interestingly, the induced samples infected with an IN+ virus do not show any enhancement of the nef transcripts when induced 5 days post infection (upper left graph of Figure 3.13) and a minor increase (1 mM NaBut) in the amount of transcripts in the samples induced 20 days post infection (lower left graph of Figure 3.13) compared to the Tat-dependent transcription in infected, non-induced cells. Due to contamination, we lost the samples induced with VPA at 5 days post infection. In the IN- infected cells (right graphs), observed is a slight increase (< 2-fold) in some samples. Overall we concluded that the stimulatory effect of HDACis to induce transcription was counteracted most likely by one or more of the de novo produced proteins.

To normalize the data, we used the extracted DNA from the samples. About 63 ng of total cellular DNA was used for quantitative real-time PCR and total HIV-1 genome (late viral DNA) was targeted for amplification. The quantity of nef transcripts were
Figure 3.13
HIV-1 LTR mediated transcription is not stimulated by HDACis induction. MDMs were infected with an IN+ and an IN- HIV-1 virus that produced de novo Tat. HDACis were added 5 and 20 days post infection for 2 days. Cells were lysed on the plate and total RNA was extracted from equal amounts of lysate. The nef transcript was targeted for real-time PCR after production of cDNA.
divided by total HIV-1 template. Only the NaBut induced samples were normalized and as we can see in Figure 3.14, the data is not as consistent as in the non-normalized data. In the results of the induction of transcription at 5 days post infection, observed is an increase of up to 1.3-fold (5 mM) of the integrated template (upper left graph) compared to the infected, non-induced cells. In the nonintegrated template, observed is an increase in nef transcripts per HIV genome of 4.6-fold (1 mM) compared to the control. However, as we can see in the lower part of Figure 3.14, in the samples induced 20 days post infection; there is only an increase of 1.9-fold (1 mM) in integrated and 1.7-fold in nonintegrated transcription.

In the next experiment, we combined the two previous experiments so we excluded donor-dependent variation in our data. We only used HDACi NaBut for induction, but included the concentration 125 mM. MDMs were infected with the four different viruses; the IN- reporter genome virus; vNL-Luc(D116N), and its counterpart IN+ virus; vNL-Luc, the IN- viral genome virus; vNL(KFS)(D116N) and its IN+ counterpart vNL(KFS). All the used MDMs came from one donor (101211) and were treated all the same. The macrophages were infected 15 days after PBMC were purified from fresh blood with about 20 ng p24 values of the viruses stated above. Five and 19 days post infection; transcription induction took place with the HDACi NaBut (Merck - Darmstadt, Germany) for 2 days before cells were lysed and harvested. The cells infected with viruses vNL-Luc(D116N), and vNL-Luc were lysed in 500 µL 1X cell culture lysis reagent (Promega - Madison, Wisconsin) and the cells infected with the viruses vNL(KFS)(D116N), and vNL(KFS) were lysed in 500 µL cell lysis buffer (SV Total
Figure 3.14

Normalized data of the stimulation of the HIV-1 LTR by induction. The relative Nef copies (Figure 3.13) were divided by the quantity of whole length viral DNA template.
RNA isolation System, Promega - Madison, Wisconsin) on the plate. From the latter, RNA and DNA were extracted from equal amounts of lysate per sample. For the luciferase assay, 100 µL of lysate was added to 3 different 96-well plates to run each sample in triplicate. For the reverse transcriptase reaction, we used 55 ng (samples 5 days post infection) and 43 ng (samples 19 days post infection) of total cellular RNA and for the real-time PCR, we targeted the viral nef mRNA. Used were the probe: FAM-Nef/Rev (Integrated DNA Technologies - Coralville, Iowa) where 10X is 2 µM and the primers: 5’ Nef and 3’ Rev (Integrated DNA Technologies - Coralville, Iowa) where 10X is 3 µM. Used for the standard was a 10-fold dilution series of PCR Nef copies diluted in nuclease-free water, 1 X 10³ – 1 X 10⁹ copies. In general, the morphology of the MDMs induced with the lowest HDACi concentration and the MDMs with and without infection, and no induction, did look the healthiest compared to the MDMs with the induction at higher concentrations. The morphology of the MDMs with the HDACis at higher concentrations were more rounded and seemed to come loose from the plate compared to the more “fried egg” shape of healthy macrophages.

Figure 3.15 shows the induction of basal gene expression 5 and 19 days post infection from the reporter gene. The result shows a steady increase of the reporter expression in 1, 5, and 25 mM NaBut induction. The 125 mM samples show low activity most likely due to cell death. We also observed that the infected, non-induced samples shown in the lower half of Figure 3.15 (“induction 19 days post infection”) do not show a high repressed luciferase activity like we saw in a previous study (Figure 3.11). This is most likely due to donor variation (505). As shown in Figure 3.16, the induction from the
Figure 3.15

Stimulation of the LTR-driven reporter by induction. MDMs were infected with an IN+ and an IN- luciferase reporter virus where the reporter gene is controlled by the HIV-1 LTR promoter. HDACi Sodium Butyrate (NaBut) was used to induce reporter gene expression.
**Figure 3.16**

**Stimulation of the LTR from HIV-1 viral template by induction.** MDMs were infected with an IN+ and an IN- HIV-1 virus. HDACi Sodium Butyrate (NaBut) was added 5 and 19 days post infection for 2 days. Total RNA was extracted from equal amounts of lysate per sample and *nef* transcripts were targeted for real-time PCR after production of cDNA.
HIV-1 template (with the production of de novo Tat) shows an inconsistent pattern. Overall, the 5 mM NaBut produced the highest relative quantity of nef transcripts. Induction 5 days after infection displays a 1.2-fold increase of nef transcripts from integrated HIV-1 template compared to infected, non-induced cells. For the nonintegrated form, the increase is 1.6-fold. In the induced samples 19 days post infection, observed is no increase from the integrated form and a 1.3-fold increase from the nonintegrated template. Due to low amounts of lysate from samples infected with the reporter virus, only the samples infected with the HIV-1 genome virus were normalized. The normalized data (Figure 3.17) shows that the highest induction came from the samples induced with 5 mM NaBut. Overall, the observed increase of transcription by induction does not exceed 2-fold. This experiment confirms the result obtained by a previous experiment that induction of an HIV-1 genome by the HDACi NaBut is obstructed by a de novo viral protein.

**Tat opposed the stimulatory effects of an HDACi**

In the following experiment we wished to explore if de novo Tat is potentially the protein that caused the obstruction of the HDACi induced transcription. We used a virus with an LTR based construct that co-expresses the tat gene and a reporter gene (luciferase) by using an internal ribosome entry site (IRES). This IN+ virus is called vNL-Tat-IRES-Luc and the IN- virus is called vNL-Tat-IRES-Luc(D116N). For comparison purpose, we also used a virus with a reporter construct lacking the tat gene. For the IN+ the virus is named vNL-Luc and for the IN-: vNL-Luc(D116N). We infected
Figure 3.17

Normalized data of the LTR stimulation by induction from an HIV-1 virus infection. The relative Nef copies (Figure 3.16) were divided by the quantity of full length viral DNA that was present in equal amounts of lysate.
MDMs with the above discussed viruses (220 ng p24 value) and as HDACi we used NaBut. Cells were induced 5 days post infection. Noticed was the phenotype of the induced cells that seemed to be more rounded compared to the more “fried egg” form in the non-induced cells. Cells were lysed and harvested after induction only at 5 days post infection due to the quality of the MDMs. From the data of the luciferase expression in Figure 3.18 in the upper half, observed is a trend that if de novo Tat is produced, induced activity from the reporter template by HDACis is lacking or low. However, if Tat is absent, there is induced transcription of the reporter gene by NaBut. In the IN+ infection lacking Tat, an increase of 2.3-fold is observed and in the IN- infection lacking Tat, a 3.6-fold increase is shown in Figure 3.18. In the normalized data (Figure 3.18, lower half), we observed the same pattern. These observations suggest that, indeed, protein Tat is the de novo protein that interferes with the induction of transcription by HDACis. We wished to duplicate this experiment with only the nonintegrated (IN-) viral genome infection to confirm our findings.

We repeated the above experiment and used 20 ng of p24 value virus pNL-Luc-Tat(D116N) and pNL-Luc(D116N) for infection. Samples were induced 5 and 14 days post infection and cells lysed and harvested. The data was normalized with total HIV genome. The columns “Raw data” in Figure 3.19 illustrates the luciferase activity before normalization of the data. The infected cells that express Tat do not exhibit any induced transcription of the reporter by NaBut. When Tat is not expressed, there is induction of transcription by NaBut ranging from 3.3-fold to 4.6-fold increase of reporter expression. This result does confirm the results of earlier experiments. Tat is able to either inhibit
Figure 3.18
Stimulation of LTR-based reporter by HDACi induction with or without Tat.

MDMs were infected with an IN+ and IN- reporter virus particle that either co-expressed a tat gene or did not. The cells were induced with the HDACi Sodium Butyrate (NaBut) 5 days post infection for 2 days. Luciferase expression was evaluated from equal volumes of lysate. The data was normalized by dividing the luciferase expression by total, full length HIV DNA (lower half).
Figure 3.19

Stimulation of the LTR-driven reporter gene by an HDAC inhibitor. MDMs were infected with an IN- reporter virus particle that either included a tat gene or did not. The cells were induced with the HDAC inhibitor Sodium Butyrate (NaBut) 5 and 18 days post infection for 2 days. Luciferase expression was evaluated in equal volumes of lysate. The data was normalized by dividing the luciferase expression by total, full length HIV DNA.
integrated and nonintegrated DNA from assembling into repressive chromatin or Tat keeps the chromatin in a nonrepressive state.

**Chromatin assembled onto nonintegrated reporter vector**

Organization of viral DNA into chromatin-like structures has been detected in DNA viruses such as SV40, human and bovine papilloma virus (HPV and BPV respectively), EBV, and HBV (506–511). It has also been demonstrated that nonintegrated HIV-1 and HIV-1 based vector genomes are arranged into chromatin structures and with a histone modification pattern that is distinctive of chromatin that is transcriptionally restricted (497). In the next experiment we wished to show that chromatin is assembled on the nonintegrated reporter vector DNA by the chromatin immunoprecipitation (ChIP) assay. U937 cells, a monocyte cell line, were infected with the IN- reporter particle vNL-Luc(D116N) at a p24 value of 312 ng and spin-inoculated for 2 hours at 600 g (1800 rpm). Following the centrifugation, the cells were incubated at 37°C and 5% CO₂ for a total time after infection of 4, 12, and 24 hours before cross-linking the cells. A standard protocol for the ChIP assay was used (477) and Dr. Mudit Tyagi performed this part of the experiment. On every run, 5% of each sample was analyzed by quantitative real-time PCR to determine the amount of sample immunoprecipitated by individual antibodies. Specific primer sets were used to amplify different regions of the LTR. The reading obtained with preimmune sera was subtracted as background counts. Observed in Figure 3.20 is the binding of the RNA polymerase II at the promoter, and the high levels of HDAC-1, a key enzyme in repressive chromatin, at nuc-1 compared to nuc-2. Detected are also the acetylated H3 and H4 and the p65 subunit of NF-κB, which indicate that
Figure 3.20

Chromatin assembled onto nonintegrated reporter genome. U937 cells were infected with the IN- reporter viral particle: vNL-Luc(D116N). The cells were spin-inoculated for 2 hours and cells harvested for chromatin immunoprecipitation (ChIP) assay 4, 12, and 24 hours post infection.
remodeling of a chromatin structure takes place at nuc-1 compared to low levels at nuc-2. This data suggests that the nonintegrated viral DNA is packaged in a chromatin structure that can be either in a restricted or open form.

**Nonintegrated DNA transcription in microglia**
The appearance of HIV-associated neurocognitive disorders (HAND) in infected HIV-1 individuals is about 50%. HAND is defined by the development of cognitive, behavioral and motor abnormalities (512,513). Nowadays, an HIV infection is a chronic, controllable disease when on HAART. Although incidences of HIV-associated dementia (HAD) has decreased with patients on HAART (514), the medication is inadequate to infiltrate the CNS (515) and this defect generates a protected viral reservoir in this part of the body. HIV-1 infects and replicates efficiently in brain macrophages, microglia, and monocytes (512,516). Microglia are a type of glial cells that are called the resident macrophages of the brain and spinal cord. They are distributed in all brain regions at different concentrations, up to 16% of the total glial cell population within the brain (348). Neuronal cell death and apoptosis is caused by release of proinflammatory factors and neurotoxins by HIV infected cells and the activated immune cells in the brain, but also by viral proteins such as Tat, Vpr, and gp120. This neuronal cell death intensifies HAND (517–519). Tat was found to have an important role in the development and maintenance of HAND due to its ability to induce the activation of microglia (361). We received a 75 cm² flask with human microglia from the Virginia Commonwealth University in Richmond, Virginia. We wanted to explore if these brain cells were capable of supporting nonintegrated viral DNA transcription. We infected the cells with a VSV-G
pseudotyped IN- HIV virus, vNL4-3(KFS)(D116N). About 500 µL of virus (~288.5 ng p24 value) was added to the cells in the flask and incubated at 37° C and 5% CO2 for 4 hours and followed by a wash. The cells were lysed after 72 hours post infection in 600 µL of lysis buffer. A total of 175 µL of the lysate was used for RNA and DNA extraction. The nef transcript was targeted in the total cellular RNA and total, full-length HIV DNA was targeted by PCR in the total cellular DNA. The data presented in Figure 3.21 shows that microglia do support transcription from nonintegrated viral DNA. Due to the fact that we only had one flask to work with, no controls could be provided.
**Figure 3.21**

**Microglial cells support transcription from nonintegrated viral DNA.** Microglial cells were infected with the VSV-G pseudotyped IN- HIV virus vNL4-3(KFS)(D116N) and cellular RNA was evaluated for the presence of *nef* transcripts and total cellular DNA was evaluated for total, full-length HIV-1 DNA by real-time PCR.
Discussion

In the initial state of an HIV-1 infection, transcription from the viral promoter is activated by cellular transcription factors such as Sp1 and NF-κB. Nevertheless, the promoter demonstrates a defect in elongation and basal transcription resulting in excessive short transcripts of about 50-60 nucleotides long. This defect could be caused by the lack of pTEF-b complex at the promoter, or the nucleosome that is located just downstream of the transcription start site, nuc-1, repressing elongation. Viral protein Tat is involved in resolving both these limitations. Pre-mature termination of the transcript is not absolute and some full-length RNAs are produced and Tat can be generated.

In one of the very early studies, researchers found that gene expression from a LTR-driven reporter template was 200 to 300 fold more in HIV-1 infected cells, compared to non-infected cells (215). In subsequent studies it was identified that Tat activates gene expression from the LTR with the cooperation of the TAR RNA located at the 5’ end of the maturing viral RNAs (174,380,381). Tat recruits PTEF-b, which consists of the kinase CDK 9 and the regulatory subunit cyclin-T1 (203,520), to the TAR RNA. RNA elongation by RNA Pol. II is mainly regulated through PTEF-b by phosphorylating the DSIF (521) and serines 2 residues in the CTD repeats of RNA Pol. II (197). DSIF induces a stable, paused RNA Pol. II when not phosphorylated, but when phosphorylated it works as a positive elongation factor (521). DSIF works together with the NELF, which is also phosphorylated by pTEF-b and upon phosphorylation, released from the RNA (206). Not only is Tat a trans-activator of transcription elongation, the protein is involved in many more functions. Tat is involved in inducing apoptosis (522,523), in changing cell surface protein expression profile such as MHC class I (144,524), mannose receptors
(525), CXCR4 (526), and CCR5 (527). Tat is a neurotoxin (255) and dysregulates cytokine expression by activating the production of IL-1 (528), IL-6 (261), and TNF-α (529), and more. Tat is also known to recruit chromatin-remodeling complexes. The ATPase-dependent chromatin remodeling complex SWI/SNF is known to be recruited by Tat and it was found that Tat interacts with a subunit of the complex needed to stimulate expression from the LTR (136). Also, Tat was found to form a ternary complex with p300 and P/CAF, two HATs. However, only P/CAF was required for Tat-mediated promoter activity and only from integrated LTRs, not from nonintegrated LTRs (141).

In our studies, we also found that protein Tat was essential in LTR activation in both integrated and nonintegrated templates. We detected a reporter expression reduction ranging from 82 to 92% when Tat was absent in an IN+ infection. In an IN- infection we observed reductions ranging from 93 up to 98% in reporter expression from a nonintegrated template (Figure 3.3). From this data, we concluded that Tat was important in the transcription from nonintegrated DNA. However, this demand for Tat does require that the protein be available in the early phase of the infection. One hypothesis is that protein Tat is, in low amounts, packaged in the virion (266,530) and therefore delivered and available in an early stage of the infection, and triggers the initial production of de novo Tat. However, we were not able to prove this hypothesis.

Another hypothesis is that the nonintegrated viral template produces Tat in a Tat-independent manner. This early produced Tat could “prepare” the host cell for a productive HIV-1 infection. In earlier studies by Wu and Marsh, it was established that nonintegrated DNA in quiescent T-cells is capable of producing both the early and late
gene transcripts, but that only the multiply spliced transcripts were translated, and that this event most likely was due to insufficient Rev activity (290). Though protein Tat was never detected by the Western blot assay, it is always assumed that it is produced from a nonintegrated viral template (277,290,294). In this presented study, we also demonstrated that Tat is most likely produced by a nonintegrated template. We used a double mutant HIV-1 virus, vNL4-3(KFS)(D116N). This virus is VSV G-pseudotyped due to a mutation in the env gene of the packaging plasmid and this plasmid also does not produce a functional integrase protein (D116N). The cell line 1G5, which includes a stably transfected LTR-driven luciferase reporter plasmid, was used for the infection. To show that Tat originated from a newly synthesized viral DNA template, cells were pre-treated with a reverse transcriptase inhibitor, etravirine. As shown in Figure 3.2, an increase in reporter activity was observed in the infected cells up to 19-fold compared to the infected cells pre-treated with the reverse transcriptase inhibitor. We assume that this increase is caused by the activity of de novo Tat produced by the nonintegrated viral template. That Tat is involved in nonintegrated transcription is presented in Figure 3.9. Here we used a Tet-On system that produced Tat after the addition of doxycycline to the system. Cells were infected with IN+ and IN- viruses that were both not capable of producing de novo Tat, but Tat could be delivered in the virion. The samples were induced or not induced with doxycycline to activate the production of de novo Tat. Observed is an increase in reporter activity from both the integrated and nonintegrated DNA in induced samples compared to non-induced samples. Here we confirm that newly made Tat is necessary for
productive transcription from the LTR in the nonintegrated as well as in the integrated system.

To determine if nonintegrated viral DNA assembles into a repressive chromatin structure, we used HDAC inhibitors to detect if reporter activity was inducible. A repressive chromatin is characterized by the presence of HDACs at the integrated HIV LTR (145). Identified are two ubiquitous transcription factors, YY1 and LSF that are responsible for the repressive nature of the LTR (146,316,531). LSF attaches to the DNA sequence located at base pairs -10 to +27 of the LTR and recruits YY1 and together thereafter they recruit HDAC1 to the LTR to exert its repression upon the promoter (146). Also the inactive p50 subunit of NF-κB is known to recruit HDACs (98). HDACs induce hypoacetylation of H3 and H4 at conserved lysine residues located in the N-terminal tail of the histones. Methylation of H3 at lysine residues 9 and 27 by the histone methyltransferase Suv39H1 recruits the adapter protein heterochromatin protein 1 (HP1) that binds to the methylated lysine 9 residue of histone H3 to contribute to the transcriptional silencing (321,476). Recently it was shown that a nonintegrated HIV-1 and HIV-1 based vector genomes are arranged in the structure of chromatin (497). As shown in Figure 3.20 we also confirmed that chromatin remodeling takes place in the nonintegrated reporter DNA. We compared the less remodeled nucleosome 2 (nuc-2) with nuc-1. Nuc-1 is located just downstream of the transcription start site and when repressive chromatin is formed on the nonintegrated DNA, nuc-1 needs to be remodeled before efficient transcription from the template can take place. Shown is a higher quantity of HDAC1 at the nuc-1 site compared to nuc-2. Also, the higher amounts of H3 and H4
acetylation and p65 subunit of NF-κB at nuc-1 compared to nuc-2 are indicative of chromatin remodeling. We also determined that HDACis are capable of remodeling the chromatin in a relaxed form that allows a boost in basal transcription in the nonintegrated DNA with the integrated DNA functioning as control group. We used primary macrophages derived from monocytes. Macrophages are non-dividing cells, but metabolically active and it was demonstrated that macrophages are able to maintain persistent viral infections and support continuous LTR activity from viral nonintegrated DNA (342). Three different HDACis were used; Apicidin, Sodium Butyrate (NaBut), and Valproic acid (VPA) for induction. It was demonstrated that HDACis cause cell death in transformed cells (473,479) and it was found in one study that a specific HDACi (LBH589) caused selective cell death in CD4+ T lymphocytes and macrophages (532). HDACs not only target histones, but have many other protein substrates. When using HDACis, the functions of proteins involved in gene expression regulation, cell proliferation, and cell death are altered (473). Also, HDACis used at high concentrations can cause cell death in transduced primary resting CD4+ T cells (496). We did notice cell death in our samples, but we were not able to count the viable cells. Therefore, we normalized the data by the observed total, full-length viral DNA. We found that not only the integrated HIV-1 DNA was transcriptionally inducible with the HDACis, but also the nonintegrated form as shown in Figure 3.11 and Figure 3.12. In the normalized data (Figure 3.12), we observed in the IN+ infected samples that were induced 5 days post infection, up to 3.7-fold increase in reporter activity. For the IN- infection the increase is over 5-fold compared to infected, non-induced samples. The fold increase is much higher
in the samples induced 18 days post infection. The infected, non-induced samples show an almost complete silencing of the basal transcription at 18 days post infection. This event is most likely due to the repressive chromatin structure formed on both the integrated and the nonintegrated template, as this latency is inducible with the HDACis. Observed is up to 7990-fold increase in transcription from integrated DNA in the 25 mM NaBut induced sample and up to a 112-fold increase of reporter expression from nonintegrated DNA in the 25 mM NaBut induced sample compared to the infected, non-induced cells. The next experiment was conducted to explore if the same pattern was observed in a wild-type HIV-1 genome. We used an IN+ and IN- virus that was assembled with a construct that expressed all functional HIV-1 proteins except the envelope and by virtue of the D116N, it also did not express a functional integrase. The construct did contain the HIV-1 genomic packaging signal. MDMs were used as the primary cell for infection. Surprisingly, the HDACis were insufficient in significantly inducing transcription from either the integrated HIV provirus or the nonintegrated viral DNA (Figure 3.13 and Figure 3.14). In most instances the increase did not exceed 2-fold. However, in the normalized data of the IN- induction with 1 mM NaBut, we see a 4.6-fold increase in reporter activity (Figure 3.14). This experiment was repeated and as shown in Figure 3.17, the increase in transcription due to induction of NaBut compared to infected cells only is up to 2-fold. Overall, in HIV-1 infected macrophages where de novo Tat was produced, the pattern of induced versus non-induced transcription does not compare to the pattern of the reporter virus infected macrophages where no Tat was produced from the template. We hypothesized that newly-made Tat in the system
antagonizes the stimulatory effect of the HDACis of both an integrated as well as a nonintegrated template. To investigate this hypothesis, we generated an LTR-based expression vector with an LTR-driven tat gene and the luciferase gene is expressed through an internal ribosome entry site (IRES). For comparison, we also used the LTR-based luciferase expressing vector without the tat gene. We showed that Tat is indeed antagonizing the stimulatory effect of the HDACi NaBut in both the integrated and nonintegrated viral DNA (Figure 3.18 and Figure 3.19).

Even though over 90% of the cDNA in infected CD4+ T cells is not successfully integrated in the host genome (40,44,272,289), the function of nonintegrated viral DNA is not fully explained as being part of the HIV-1 life cycle. A prior study by Wu and Marsh concluded that the nonintegrated transcription is a normal early course of action in an HIV infection (290). The CD4+ T cells are the primary target cells for HIV-1 multiplication (533), however, most of these cells are in a quiescent state (in Go phase) which does not maintain virus replication. When HIV-1 infects these resting T cells, integration is blocked and the reverse-transcribed genome remains as extra chromosomal viral DNA. However, integration of the genome will take place after activation of the cell by mitogens or antigens such as the phorbol ester, phorbol 12-myristic 13-acetate (PMA) (393), and submitogens such as cytokine interleukin (IL)-2, IL-4, IL-7, and IL-15 (534). Also viral proteins Nef (535–537) and Tat (538–540) are known to stimulate T cell activation. For these viral components to be responsible for the activation of the quiescent cells, they first need to be produced from the provirus. This could be accomplished by transcription from nonintegrated viral DNA (288). Nef protein is produced, and
presumably Tat too, from nonintegrated HIV-1 DNA in non-activated T cells (288). The question remains what initiates the nonintegrated transcription. In this study we determined that basal transcription from the nonintegrated DNA is generally the same compared to integrated basal transcription (Figure 2.13). Viral protein R (Vpr) is a highly conserved, virion-associated protein with various functions in viral replication. Some of the functions are nuclear transport of the PIC (541,542), induction of cell cycle G2 arrest (543,544), induction of apoptosis (545), and LTR promoter activation (546,547) [see review (548,549)]. Vpr interacts with transcription factors Sp1 (446) and NF-κB (422,550), and with the transcription factor of the basal transcription machinery, TFIIB (445), and also the co-activator p300/CBP (546,547). Effective transcription of nonintegrated viral DNA requires Vpr. In a Vpr mutant, observed was a 10- to 20-fold decrease in nonintegrated viral DNA expression (551). The ability of Vpr to regulate gene expression from the LTR of nonintegrated viral DNA could play an essential role in the immediate early phase of the HIV-1 life cycle. In an early study it was determined that the cell activation of infected resting T cells was not caused by the binding of the HIV envelope to cellular receptors or by the virion-associated Nef (288). The immediate early gene expression regulation from the nonintegrated viral DNA in resting T cells could produce Nef and Tat to stimulate cell activation. Thus, the early Tat produced from nonintegrated viral DNA is functionally important in maintaining an open chromatin for persistent low-level transcription, which can modulate cellular activity to promote infection.
Future Studies

It is possible that virion-associated Vpr is required for nonintegration transcription. It was found that in the absence of Vpr, a 10- to 20-fold decrease in preintegration transcription occurred (551). Similar finding was reported in a recent study by the Levy group. The authors also observed the need for Vpr in preintegration transcription (343). For future studies, the exploration of the involvement of Vpr in nonintegrated transcription would be valuable. It is important to know what part of the basal transcription machinery Vpr interacts with. The Sp1 and NF-κB element mutants described in this dissertation in combination with the presence or absence of Vpr in the virion could potentially answer that question. Also, another question is if this action of Vpr in LTR transactivation originates from virion-associated Vpr or de novo synthesized Vpr. This question can be addressed by comparing an infection with a virus with or without Vpr in the virion particle and with or without de novo synthesized Vpr.

The involvement of viral protein Tat in antagonizing repressive chromatin may have an application in gene therapy. Gene expression of a transgene is influenced by the chromatin environment at the site of integration in the human genome. A repressive chromatin structure is one of the factors that diminish transgene expression. Tat may be co-expressed on the vector expressing the transgene. The vector may also include a promoter that can activate genes in quiescent or non-activated cells. It has been reported that the tob gene was selectively expressed in anergic cells and needs to be down-regulated for activation of T cells (554). The tob gene promoter might be a good candidate in this application. However, due to the toxicity of protein Tat, the protein should be altered into a non-toxic mutant. It is known that in HIV-1 infection, Tat is
released from infected cells and enter uninfected cells with the use of the cellular heparan sulfate proteoglycans as receptors that interact with the basic domain of the extracellular Tat (555,556). In the uninfected cell, Tat is able to exert its activity on modifying specific gene expression (557) which may induce apoptosis. On the other hand, it is also demonstrated that Tat recruits HATs to the HIV-1 promoter region, such as CBP/p300, which leads to alleviation of the repressed chromatin (133,558). This action of Tat could be the cause of the chromatin antagonization. This important function of Tat should be preserved.

The same application can be used for purging HIV-1 from latent reservoirs. The Tat expressing vector may be made HIV-1 independent so it will express Tat in latent cells and as a consequence activate the HIV-1 provirus in these latent cells.
REFERENCES


145. He G, Margolis DM. Counterregulation of Chromatin Deacetylation and Histone Deacetylase Occupancy at the Integrated Promoter of Human Immunodeficiency Virus Type 1 (HIV-1) by the HIV-1 Repressor YY1 and HIV-1 Activator Tat. Mol Cell Biol. 2002 May;22(9):2965–73.


184. Bieniasz PD, Grdina TA, Bogerd HP, Cullen BR. Recruitment of a protein complex containing Tat and cyclin T1 to TAR governs the species specificity of HIV-1 Tat. EMBO J. 1998 Dec 1;17(23):7056–65.

185. Bieniasz PD, Grdina TA, Bogerd HP, Cullen BR. Recruitment of cyclin T1/P-TEFb to an HIV type 1 long terminal repeat promoter proximal RNA target is both necessary and sufficient for full activation of transcription. PNAS. 1999 Jul 6;96(14):7791–6.

186. Chen D, Fong Y, Zhou Q. Specific interaction of Tat with the human but not rodent P-TEFb complex mediates the species-specific Tat activation of HIV-1 transcription. PNAS. 1999 Mar 16;96(6):2728–33.


193. Herrmann CH, Rice AP. Lentivirus Tat proteins specifically associate with a cellular protein kinase, TAK, that hyperphosphorylates the carboxyl-terminal


247. Feinberg MB, Baltimore D, Frankel AD. The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation. PNAS. 1991 May 1;88(9):4045–9.


314. Fortin JF, Barbeau B, Robichaud GA, Paré ME, Lemieux AM, Tremblay MJ. Regulation of nuclear factor of activated T cells by phosphotyrosyl-specific phosphatase activity: a positive effect on HIV-1 long terminal repeat-driven


358. IL8 Gene - GeneCards | IL8 Protein | IL8 Antibody [Internet]. [cited 2013 Oct 26]. Available from: http://www.genecards.org/cgi-bin/carddisp.pl?gene=IL8


405. NF-kB Transcription Factors | Boston University [Internet]. [cited 2014 Apr 22]. Available from: http://www.bu.edu/nf-kb/


472. Lin X, O’Mahony A, Mu Y, Geleziunas R, Greene WC. Protein kinase C-theta participates in NF-kappaB activation induced by CD3-CD28 costimulation through


CURRICULUM VITAE

Beatrix Wilhelmina Meltzer was born on January 31, 1965 in Groningen, The Netherlands. She immigrated to the United States of America in 1996 with her two daughters, Patries and Romy, and married Lowell Meltzer. She received her Bachelor of Science in Biology from George Mason University in 2001. She was employed as a Molecular Lab Technologist, and as an Assistant in Research and Development at the Genetics & IVF Institute in Fairfax, Virginia for six years. During this time, Beatrix received her Master of Science in Biotechnology from Johns Hopkins University in 2006. While working on her Ph.D., she worked as a Teaching Assistant at George Mason University teaching Genetics lab. Beatrix became an American citizen in April of 2014.