EXAMINATION OF UNSPLICED HIV-1 MESSENGER RNA TRANSLATION

by

Kimberley Anne Marsh

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy.

Graduate Department of Molecular Genetics

University of Toronto

© Copyright by Kimberley Anne Marsh 2008
Examination of unspliced HIV-1 mRNA translation

Doctor of Philosophy, 2008

Kimberley Anne Marsh

Graduate Department of Molecular Genetics

University of Toronto

ABSTRACT

Replication of HIV-1 requires nuclear export and translation of the incompletely spliced 4 and 9 kb classes of HIV-1 mRNA, which encode the structural and enzymatic proteins of the virus. HIV-1 Rev binds to the Rev-responsive element (RRE) contained in the introns of incompletely spliced HIV-1 mRNAs and mediates their nuclear export via the Crm1 pathway. Sam68ΔC, is a C-terminal deletion mutant of the endogenous human protein Sam68, and has been shown to be a potent inhibitor of Rev-dependent reporters. In this study we have performed deletion analysis of Sam68ΔC, and determined the minimal mutant required for inhibition of Rev-dependent expression is Sam68Δ14Δ(45-54)-300. Sam68ΔC inhibition is specific to RRE/Rev/Crm1 transported mRNAs: the Rev/Crm1 exported reporter construct GagRRE is inhibited while the Tap/p15 transported GagCTE reporter construct is not. Previous work from our lab showed that Sam68ΔC co-localized with the Rev-exported mRNAs in perinuclear bundles. Further investigation has shown that Sam68ΔC inhibition of incompletely spliced HIV-1 mRNAs is independent of the perinuclear bundling of the viral mRNA. We go on to show that Sam68ΔC specifically inhibits the translation of the incompletely spliced HIV-1 mRNAs. Translational inhibition by Sam68ΔC is correlated with a loss of PABP1 binding with no
attendant change in abundance, polyadenylation or polyadenosine tail length of the affected mRNAs.

The selective inhibition of Crm1 exported HIV-1 mRNAs by Sam68ΔC suggests that it is able to recognize unique characteristics of these viral mRNPs. We show that Rev and the RRE are required, but individually neither is sufficient for complete Sam68ΔC inhibition. Study of the incompletely spliced HIV-1 mRNP revealed that the nuclear cap binding complex, CBP20/80, is not exchanged for eIF4E. Additionally, in cells expressing the HIV-1 provirus, CBP80 relocalizes to the cytoplasm and co-sediments with polysomes. This supports the hypothesis that incompletely spliced HIV-1 mRNAs are translated in an eIF4E-independent, CBP20/80-dependent fashion. This property of the 9kb and 4kb HIV-1 mRNAs could be utilized to develop new therapeutic approaches to controlling HIV-1 infection.
ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Alan Cochrane, for the opportunity to work in his laboratory over the past few years. I appreciate the many learning experiences he has provided me. I would also like to thank my supervisory committee: Dr. Lea Harrington and Dr. Ben Blencowe for their continued guidance and support. Also, I would like to acknowledge the funding I received through a University of Toronto Open Scholarship and Ontario HIV Treatment Network Studentships.

A huge thank you to all my colleagues in the lab, past and present: Meredith, Tina, Rade, Jodi, Kengo, Dan, Kathryn, Nicole and all the summer, project and rotation students. Thanks for helpful scientific discussions, support and most of all friendship. Nobody could ask for a better group to work with, you always kept me laughing. I will miss our coffee breaks, and Fridays just won’t be the same without Pho Hung.

I would not have made it this far without the support of my family. Mum and Dad: thanks for always encouraging me to be curious and to follow my dreams wherever they lead. Cathy, Ken and John: thanks for teaching me to keep things in perspective and not to take anything too seriously…especially myself. To my husband Matt, without your unconditional love and support I never would have made it to the finish line…this one’s for you!
TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION ................................................................................................................1
1.1. mRNA PROCESSING .........................................................................................................................4
1.1.1. Capping .........................................................................................................................................5
1.1.2. Splicing ..........................................................................................................................................6
   Splicing Control Elements ......................................................................................................................9
   Alternative Splicing ..............................................................................................................................10
   Exon Junction Complex .......................................................................................................................11
1.1.3. 3’ End Processing .........................................................................................................................13
1.2. RNA EXPORT .....................................................................................................................................16
1.2.1. Nuclear Pore Complex ..................................................................................................................16
1.2.2. Transport Receprors .....................................................................................................................18
1.2.3. mRNA Export Adaptors ................................................................................................................20
1.2.4. Crm1 mRNA Export ......................................................................................................................24
1.3. TRANSLATION ....................................................................................................................................24
1.3.1. Translation Initiation .....................................................................................................................24
1.3.2. Translation Termination and mRNA Decay ..................................................................................27
1.3.3. Nonsense-Mediated mRNA Decay ..............................................................................................30
1.4. TRANSLATIONAL CONTROL .........................................................................................................31
1.4.1. General Mechanisms of Translational Control .............................................................................31
1.4.2. Specific Mechanisms of Translational Control ............................................................................34
1.5. HIV-1 ..................................................................................................................................................37
1.5.1. HIV-1 Life Cycle ...........................................................................................................................37
1.5.2. HIV RNA Processing ....................................................................................................................40
1.5.3. HIV RNA Export ..........................................................................................................................44
1.5.4. HIV mRNA Translation ...............................................................................................................49
   Cis-acting Elements ...............................................................................................................................49
   Internal Ribosome Entry Site ...............................................................................................................49
   -1 Frameshift .........................................................................................................................................50
   Leaky Scanning ......................................................................................................................................50
   Trans-acting Factors ............................................................................................................................51
1.5.5. HIV RNA Packaging ....................................................................................................................54
1.6. SAM68 AND SAM68ΔC .......................................................................................................................55
1.7. OBJECTIVES AND RATIONALE .....................................................................................................60

CHAPTER 2: STRUCTURE AND FUNCTION ANALYSIS OF SAM68ΔC .............................................62
2.1. ABSTRACT ...........................................................................................................................................63
2.2. INTRODUCTION ..................................................................................................................................64
2.3. MATERIALS AND METHODS ............................................................................................................65
2.3.1. Expression constructs ....................................................................................................................65
2.3.2. Cell lines and transfections ..........................................................................................................67
2.3.3. Antibodies ......................................................................................................................................68
2.3.4. CAT Assays ....................................................................................................................................68
2.3.5. Western Blots ..................................................................................................................................68
2.3.6. In Situ Hybridization and Immunofluorescence .........................................................................69
2.3.8. Statistical Analysis .......................................................................................................................69
2.4. RESULTS ............................................................................................................................................70
2.4.1. Sam68ΔC inhibition is dominant over cytoplasmic stimulatory Sam68 mutants ......................70
2.4.2. Sam68ΔC co-localizes with the stimulatory mutants and unspliced HIV-1 mRNAs to perinuclear bundles ..........................................................................................................................73
2.4.3. Sam68Δ14A(45-54)–300 is the minimal inhibitory Sam68 mutant ...............................................77
CHAPTER 3: SAM68ΔC INHIBITS TRANSLATION OF INCOMPLETELY SPLICED HIV-1 MRNAS BY BLOCKING PABP1 ASSOCIATION ..........................................................94

3.1. ABSTRACT ..................................................................................................................95
3.2. INTRODUCTION .......................................................................................................96
3.3. MATERIALS AND METHODS ....................................................................................98
  3.3.1. Expression constructs .........................................................................................98
  3.3.2. Cell lines and transfections ..............................................................................98
  3.3.3. Antibodies .........................................................................................................98
  3.3.4. In Situ Hybridization and Immunofluorescence ............................................99
  3.3.5. Digitonin Permeabilization ..............................................................................99
  3.3.6. 3S labeling .........................................................................................................99
  3.3.7. RNA Analysis ....................................................................................................100
  3.3.8. RNP Immunoprecipitation .............................................................................101
  3.3.9. Statistical Analysis ..........................................................................................101
3.4. RESULTS .................................................................................................................103
  3.4.1. Perinuclear Bundling of HIV-1 mRNA by Sam68ΔC Requires Intact Microfilaments not Microtubules .............................................................103
  3.4.2. Sam68ΔC inhibits unspliced HIV-1 mRNA when perinuclear bundles are dissolved ........108
  3.4.3. Sam68ΔC does not affect the amount of HIV-1 unspliced mRNA .................110
  3.4.4. Inhibition by Sam68ΔC is Not Associated with Alterations in Polyadenylation of Viral mRNAs .................................................................113
  3.4.5. Sam68ΔC Alters the Association of Rev-dependent Viral mRNAs with PABP1 ....116
  3.4.6. Sam68ΔC inhibits HIV-1 unspliced mRNA with a histone 3'UTR ..........120
3.5. DISCUSSION .............................................................................................................125

CHAPTER 4: INCOMPLETELY SPLICED HIV-1 MRNAS RETAIN CBP20/80 COMPLEX AT THEIR 5'CAP ..............................................................................................................133

4.1 ABSTRACT ...............................................................................................................134
4.2 INTRODUCTION .....................................................................................................137
4.3. MATERIALS AND METHODS ..............................................................................137
  4.3.1. Expression constructs .....................................................................................138
  4.3.2. Cell Lines and Transfections .........................................................................138
  4.3.3. Antibodies ......................................................................................................138
  4.3.4. CAT Assays .....................................................................................................139
  4.3.5. Western Blots ..................................................................................................139
  4.3.6. RNP Immunoprecipitation ..........................................................................139
  4.3.7. Immunofluorescence .....................................................................................140
  4.3.8. Polysome Gradients ......................................................................................141
  4.3.9. Statistical Analysis .........................................................................................142
4.4. RESULTS ...............................................................................................................143
  4.4.1. Susceptibility to Sam68ΔC Repression is Conferred by the Nuclear Export Pathway ....143
  4.4.2. Sam68ΔC does not affect Rev subcellular distribution ....................................146
  4.4.3. Rev-exported HIV-1 mRNA is CBP80, not eIF4E, Bound .........................151
  4.4.4. CBP80 subcellular distribution shifts to the cytoplasm in the presence of HIV-1 provirus ....153
  4.4.5. CBP80 is associated with polysomes in pgTat/Rev transfected cells ..........155
  4.4.6. CBP80 is associated with polysomes in HIV-1 provirus transfected cells .......159
4.5. DISCUSSION ..........................................................................................................162

CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS ..................................................167

5.1. CONCLUSIONS ......................................................................................................168
5.2. Future Directions.................................................................172

5.2.1. Confirming eIF4E-independence and CBP20/80-dependence of incompletely spliced HIV-1 translation.................................................................172

5.2.2. Identification of host mRNAs translated in a CBP20/80-dependent manner .......................174

5.2.3. Analysing anti-HIV-1 therapeutic potential of Sam68ΔC based strategies ....................176

5.2.4. A deeper look at translation in the HIV-1 transfected cell......................................................177

REFERENCES 181
TABLE OF FIGURES

Figure 1.1. Mechanism of splicing. ................................................................. 8
Figure 1.2. Types of Alternative Splicing ....................................................... 12
Figure 1.3. Messenger RNA 3’ end formation. ............................................... 15
Figure 1.4. Structure of the Nuclear Pore Complex (NPC). ......................... 17
Figure 1.5. Common karyopherin-mediated RNA export pathways. ............ 19
Figure 1.6. Messenger RNA export pathways. ............................................ 23
Figure 1.7. Cap-dependent Translation Initiation .......................................... 26
Figure 1.8. Translation Termination and Deadenylation ............................... 29
Figure 1.9. General Mechanisms of Translational Control .......................... 33
Figure 1.10. Mature HIV-1 virion. ............................................................... 38
Figure 1.11. Genetic organization of HIV-1. ............................................... 41
Figure 1.12. Rev Import/Export Cycle ......................................................... 46
Figure 1.13. Sam68 and Sam68ΔC Domain Structure ................................. 57

Figure 2.1. Sam68:97-255 and Sam68:97-280 are cytoplasmic and stimulate Rev-activity. .......... 71
Figure 2.2. Sam68:97-255 and Sam68:97-280 stimulation of Rev activity is inhibited by Sam68ΔC. ................................................................. 72
Figure 2.3. Sam68:97-255 and Sam68:97-280 colocalize with Sam68ΔC in the presence of unspliced HIV mRNA .................................................. 75
Figure 2.4. Sam68:97-255 and Sam68:97-280 colocalize with unspliced HIV mRNA in the presence of Sam68ΔC. ..................................................... 76
Figure 2.5. Schematic of Sam68ΔC mutants tested. .................................... 78
Figure 2.6. Analysis of Sam68ΔC deletion mutants and their effects on Rev-activity .................. 79
Figure 2.7. The effect of Sam68ΔC and mutants thereof on HIV-1 Gag expression. .................. 83
Figure 2.8. Non-inhibitory mutants of Sam68ΔC are able to bundle unspliced HIV mRNA. ........ 85
Figure 2.9. The effect of Sam68ΔC and mutants thereof on HIV-1 Gag expression from a Rev-independent reporter. ...................................................... 88
Figure 2.10. Sam68ΔC is able to bundle unspliced HIV mRNA which is not Rev-transported. .......... 89

Figure 3.1. Determination of the minimal drug concentrations required to depolymerize the cytoskeleton. ................................................................. 105
Figure 3.2. Effect of cytoskeleton depolymerization on Sam68ΔC perinuclear bundling of unspliced HIV-1 mRNA. ...................................................... 106
Figure 3.3. Effect of cytoskeleton depolymerization on Sam68ΔC solubility. ................................................................. 108
Figure 3.4. Effect of cytoskeleton depolymerization on Sam68ΔC inhibition of unspliced HIV-1 mRNA translation. ...................................................... 110
Figure 3.5. Sam68ΔC inhibits gp160/120 expression from the HIV-1 env reporter pgTat, but does not alter the amount of mRNA. ...................................................... 113
Figure 3.6. Sam68ΔC does not alter the polyadenylation status of the affected HIV-1 viral mRNAs. ................................. 115
Figure 3.7. Sam68ΔC inhibits the association of unspliced pgTat RNAs with PABP1. ................. 118
Figure 3.8. Sam68ΔC inhibits the expression from 9kb and 4kb HIV-1 proviral mRNAs. ........... 120
Figure 3.9. Sam68ΔC inhibits the association of 9kb HxBruR/RI mRNAs with PABP1. ............ 122
Figure 3.10. pgTat-hist-3’UTR is inhibited by Sam68ΔC, but RNA levels are not affected. ........ 125
Figure 3.11. pgTat-hist-3’UTR mRNA is polyadenylated. ................................ 126
Figure 3.12. Model of Sam68ΔC inhibition of unspliced HIV-1 mRNA translation. ................... 132

Figure 4.1. Sam68ΔC inhibits translation of RRE/Rev/Crm1 exported mRNA but not Tap/p15 exported mRNA. ................................................................. 144
Figure 4.2. Complete Sam68ΔC inhibition requires the RRE. ........................ 146
Figure 4.3. The RRE-alone is not sufficient to confer Sam68ΔC sensitivity on a heterologous reporter. ................................................................. 148
Figure 4.4. Neither Sam68 nor Sam68ΔC cause Rev relocalization ........................................ 150
Figure 4.5. Unspliced, cleaved pgTat mRNA is CBP80, not eIF4E, bound ................................. 153
Figure 4.6. CBP80 partially relocalizes to the cytoplasm in the presence of HIV-1 proviral mRNA. .................................................................................................................. 155
Figure 4.7. CBP80 and eIF4E co-sediment with polysomes in cells transfected with pgTat and Rev. .................................................................................................................. 157
Figure 4.8. CBP80 and eIF4E move to lighter fractions when polysomes are dissociated by EDTA. .................................................................................................................. 158
Figure 4.9. CBP80 and eIF4E co-sediment with intact polysomes in cells transfected with HxBruR'/RI'. .............................................................................................................. 160
Figure 4.10. CBP80 and eIF4E move to lighter fractions when polysomes are dissociated by EDTA. .................................................................................................................. 161
Figure 4.11. Model of spliced versus unspliced HIV-1 mRNA translation ................................. 166
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4E-BP</td>
<td>eIF4E binding protein</td>
</tr>
<tr>
<td>4E-HP</td>
<td>eIF4E homology protein</td>
</tr>
<tr>
<td>AIDS</td>
<td>acquired immune deficiency syndrome</td>
</tr>
<tr>
<td>CAT</td>
<td>chloramphenicol acetyl transferase</td>
</tr>
<tr>
<td>CBC</td>
<td>cap binding complex</td>
</tr>
<tr>
<td>CBP20/80</td>
<td>cap binding protein 20/80</td>
</tr>
<tr>
<td>CTE</td>
<td>constitutive transport element</td>
</tr>
<tr>
<td>eIF4E</td>
<td>eukaryotic initiation factor</td>
</tr>
<tr>
<td>EJC</td>
<td>exon junction complex</td>
</tr>
<tr>
<td>GSG domain</td>
<td>GRP33, Sam68, GLD-1 domain</td>
</tr>
<tr>
<td>HAART</td>
<td>highly active anti-retroviral therapy</td>
</tr>
<tr>
<td>HIV-1</td>
<td>human immunodeficiency virus 1</td>
</tr>
<tr>
<td>hnRNP</td>
<td>heterogenous nuclear ribonucleoprotein</td>
</tr>
<tr>
<td>hRIP</td>
<td>human Rev-interacting protein</td>
</tr>
<tr>
<td>IF</td>
<td>immunofluorescence</td>
</tr>
<tr>
<td>IP</td>
<td>immunoprecipitation</td>
</tr>
<tr>
<td>IRES</td>
<td>internal ribosomal entry site</td>
</tr>
<tr>
<td>KH domain</td>
<td>hnRNP K homology domain</td>
</tr>
<tr>
<td>LTR</td>
<td>long terminal repeat</td>
</tr>
<tr>
<td>m7G</td>
<td>7-methylguanisine cap</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>NES</td>
<td>nuclear export signal</td>
</tr>
<tr>
<td>NLS</td>
<td>nuclear localization signal</td>
</tr>
<tr>
<td>NMD</td>
<td>nonsense mediated mRNA decay</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>PABP1</td>
<td>polyadenosine binding protein 1</td>
</tr>
<tr>
<td>PTC</td>
<td>premature termination codon</td>
</tr>
<tr>
<td>RIP</td>
<td>RNP immunoprecipitation</td>
</tr>
<tr>
<td>RRE</td>
<td>Rev-responsive element</td>
</tr>
<tr>
<td>snRNP</td>
<td>small nuclear ribonucleoprotein</td>
</tr>
<tr>
<td>TREX complex</td>
<td>transcription/export complex</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

Kim Marsh
Human immunodeficiency virus (HIV) is the causative agent of acquired immune deficiency syndrome (AIDS), a fatal disease for which there is no cure. HIV-1 is a member of the retrovirus family of single stranded RNA viruses. It is sexually transmitted: through exposure of the vaginal, oral or rectal mucosa; intravenously transmitted: through contaminated blood products, use of contaminated equipment in intravenous drug use or from mother to fetus; or it may be passed from mother to child during breastfeeding. HIV-1 infection leads to a drastic decrease in T-cell number, and almost total disruption of lymphoid tissue. Infected individuals become severely immunocompromised and susceptible to opportunistic infections, eventually leading to death. Although there are a number of treatments for HIV-1, there are two major obstacles to long-term treatment and survival. First, due to the high mutation rate of the virus, multi-drug resistant strains are increasingly prevalent. A single nucleotide substitution in the HIV-1 reverse transcriptase gene can provide resistance to a number of non-nucleoside reverse transcriptase inhibitors. Second, latently infected cells act as a reservoir of virus, which is not eliminated by current anti-HIV therapies. (Reviewed in (1).)

Unintegrated HIV-1 nucleic acids have a very short half-life (hours to days) in unactivated T cells (2-4). Productive infection results when resting T cells which have been infected are activated, which permits integration of the HIV-1 proviral DNA into the host DNA. The viral reservoir is composed of latently infected T cells which are productively infected, escape death and revert to a resting state (5-7). These memory T cells can not be targeted by current anti-HIV therapeutics. In patients undergoing highly active anti-retroviral therapy (HAART) researchers found that although there are very
few latently infected memory T cells, they have an estimated half-life of 44 months (8). At that clearance rate it would take 60 years to eradicate the HIV reservoir using available anti-HIV therapies (8). Other studies found that the half-life of memory T cells was closer to 6 months. This indicates that despite HAART treatment, there may be a low level of HIV replication, which is able to replenish the reservoir of latently infected cells. Lymphoid tissue, the central nervous system and genital tract may also be important HIV-1 viral reservoirs (1).

Upon activation of latently infected cells, transcription of the integrated HIV-1 provirus generates a single 9 kb transcript encoding all of the viral proteins. mRNA processing generates three classes of mRNA from this transcript: the 9 kb class encoding Gag and Gag-Pol; the 4 kb class encoding Env, Vif, Vpu and Vpr; and the 2 kb class encoding the regulatory proteins Tat, Rev and Nef. In addition, the 9 kb mRNA also serves as the HIV-1 genome and must be packaged into progeny virions. In order for HIV-1 to successfully replicate the viral proteins must be produced in stoichiometric amounts through the complex regulation of viral mRNA processing, export, and translation. However, there are currently no drug therapies which target any of these processes (9).

According to UNAIDS, in 2007, there were 2.1 million deaths due to AIDS, and 2.5 million people were newly infected with HIV bringing the total number of infected people to 33.2 million. The development of novel anti-HIV therapies is of the utmost importance to overcome emerging drug resistant strains of HIV-1 and prevent the spread of the global pandemic. The goal of this study was to examine in greater detail how the HIV-1 mRNA processing, export and translation is distinct from the metabolism of host
mRNAs. All of these processes are dependent upon host proteins, and discrepancies between host and viral mRNA metabolism may be exploited as drug targets in the future. Targeting the post-integration stage of viral replication would be advantageous in two ways. First, if a host protein is the drug target viral mutation is less likely to produce resistant strains. Second, drugs targeting the post-integration stage of infection would be efficacious in latently infected cells which have been activated. This introduction provides an overview of our existing knowledge of host and viral mRNA processing, export and translation.

1.1. mRNA Processing

Messenger RNA (mRNA) processing involves three major steps to render a mature, export competent mRNA: addition of a 5’ cap, removal of the introns, and 3’ end processing (reviewed in [10-16]). Each of these processes can be carried out in vitro independent of transcription, however, in vivo they all occur co-transcriptionally. It is important to realize that these RNA processing reactions do not occur in isolation, but rather interact with each other and the transcription machinery. Cellular mRNAs are transcribed by RNA polymerase II (pol II), which contains a unique C-terminal domain (CTD) that serves as a docking platform for many RNA processing factors. The CTD is composed of 52 heptad repeats with a consensus sequence of YSPTSPS. Ser2 and Ser5 of the repeat are phosphorylated and dephosphorylated as pol II transcribes the mRNA. At the 5’end of a nascent mRNA there is a higher ratio of Ser5 phosphorylation, and at the 3’end there is a higher ratio of Ser2 phosphorylation. RNA pol II is part of a dynamic complex, which associates with different proteins as it proceeds from the 5’ to 3’ end of the gene.
1.1.1. **Capping**

The 5’end of pol II transcribed RNAs (mRNAs, snRNAs and snoRNAs) are modified by the addition of a 7-methylguanosine (m7G) cap structure. The capping pathway is conserved between all eukaryotes, and requires 3 enzymatic reactions. First, the 5’ triphosphate is hydrolyzed to diphosphate by RNA triphosphatase (RTP). Second, the diphosphate RNA is capped with a GMP moiety by a 5’-5’ triphosphate linkage by guanylyltransferase (GT). The third and final step is the methylation of the transferred GMP at position N7, by the methyltransferase (MT). In humans the three capping reactions are carried out by two proteins: human capping enzyme (HCE1) and human cap methyltransferase (HCM1). HCE1 is a bipartite protein with an RTP domain at the N-terminus and a GT domain at the C-terminus.

The cap is added when the nascent pre-mRNA is approximately 25 bases long, soon after the 5’end emerges from the polymerase (16). The GT and MT domains are capable of binding the phosphorylated CTD (17-19). HCE1 binds to the Ser5 phosphorylated CTD, via the GT domain, this activates GT, which in turn stimulates RTP (17). HCE1 is released early in transcription by Ser5 phosphatase, while HCM1 remains associated at the 3’end of the gene. Capping must occur co-transcriptionally; it is dependant on HCE1 and HCM1 binding to the CTD of pol II and not on cis-acting elements within the RNA.

Capping serves two critical functions in mRNA metabolism (reviewed in (20)). First, the guanylate moiety prevents degradation by 5’ exonucleases (21, 22). Second, the methyl group facilitates translation initiation (23). Prokaryotes do not require the cap structure. Prokaryotes do not have 5’exonucleases and the small ribosomal subunit is
directed to the translational start site by base pairing between the 16S rRNA and the Shine-Dalgarno sequence located upstream of the translational start site.

The cap binding complex (CBC) binds to the m7G cap co-transcriptionally (24), and plays an important role in several downstream processes. The CBC is composed of two proteins, CBP20 which binds directly to the m7G cap, and CBP80 which binds to CBP20 and increases its affinity for the cap (25, 26). The CBC is able to stimulate splicing of the first, cap proximal, intron in two ways: by stimulating the interaction between the 5’ splice site and U1 snRNP (27-29), and enhancing the subsequent the interaction with U6 snRNP (30). The CBC is also required to mediate interactions with the translation machinery during the pioneer round of translation ((31, 32), discussed further in 1.3.3.).

1.1.2. Splicing

Pre-mRNAs contain protein coding regions (exons) which are interrupted by non-coding regions (introns). The introns must be accurately spliced out of the pre-mRNA, and the exons joined, so that full-length and active proteins are made. On average a human gene contains 9 introns, averaging 4130 bases in length (ten times longer than exons which average 300 bases) (33, 34). There is complicated network of cis-acting elements and trans-acting factors that work in concert to define and recognize the splice sites. However, the basic requirements for splicing consist of four cis-acting elements and the spliceosome, as summarized in Figure 1.1. From the 5’end of the intron the cis-acting elements are: the 5’ splice site (5’ss), the branchpoint site (bps), the polypyrimidine tract (ppt) and the 3’ splice site (3’ss). The spliceosome is the splicing machinery of the cell and is composed of five small nuclear ribonucleoproteins (snRNPs): U1, U2, U4, U5 and
U6, as well as other protein factors. U snRNPs are composed of a short (<300 nucleotide) U-rich RNA and tightly coupled proteins.

The assembly of the spliceosome on the pre-mRNA occurs in a stepwise manner (see Figure 1.1). The H complex is formed prior to engagement of the splicing machinery, when the nascent pre-mRNA is bound by proteins on splicing enhancer and silencer elements (see below). The interaction of these proteins with one another and with the splicing machinery will determine where (and if) splicing will occur. Subsequent E-complex formation involves recognition of the 5’ss by base pairing with U1 snRNP. At the 3’ss: splicing factor 1 (SF1) binds the bps and the heterodimer U2 auxiliary factor (U2AF) recognizes the ppt through its 65kDa subunit and the 3’ss through its 35kDa subunit. The A-complex is formed when U2 snRNP base pairs with the bps, displacing SF1 and bulging the adenosine residue within the bps. The B-complex forms when the tri-snRNP (consisting of base paired U4-U6 and U5) joins the spliceosome. Tri-snRNP binding displaces U1 snRNP and U2AF. To form the catalytically active B* complex several remodeling steps occur which bring the 5’ss and the 3’ss into close proximity. The base pairing between U4 and U6 unwinds and U4 dissociates from the complex. This permits U6 snRNP, which is base paired to the 5’ss, to simultaneously base pair with U2 snRNP in close proximity to the bps. These alterations facilitate the first of two transesterification reactions. The 5’ exon-intron junction is cleaved by nucleophilic attack of the 2’ hydroxyl of the bulged bps adenosine, generating the catalytically active C-complex. Then the second transesterification reaction occurs: the 3’ hydroxyl group on the 5’ exon attacks the 3’ exon-intron junction. This reaction releases the remaining U snRNPs, the intron lariat and the ligated exons.
Mechanism of splicing.

The consensus sequences of the important *cis*-acting elements are shown 5’ss, branchpoint site (bps), polypyrimidine tract (ppt) and 3’ss. Initially, U1 snRNP base pairs with the 5’ss, SF1 binds the bps, and U2AF-65/35 binds the ppt and 3’ss respectively. U2 snRNP binding to the bps displaces SF1, and bulges the A residue. The U4/U6-U5 tri-snRNP joins the complex, displacing U1 snRNP and U2AF. U4snRNP, dissociates, allowing U6-U2 interaction. The bulged A residue attacks the 5’ss, and the 5’ss attacks the 3’ss. Adapted from (35).
A second spliceosome was identified which is composed of U11, U12, U4atac, U5 and U6atac (reviewed in (35)). U5 snRNP is common to both spliceosomes. Splicing of this minor class of U12-type introns occurs by the same mechanism as U2-type introns. During splicing of the minor class of introns: U11 recognizes the 5’ss, U12 recognizes the bps, and U4atac and U6atac perform the same function as U4 and U6.

Initially, U2-type introns were defined by GT and AG dinucleotides at their termini, while U12-type introns were defined by AT and AC dinucleotides at their termini. Further genetic analysis revealed that most U12-type introns contain GT-AG termini, however they are distinct from U2-type introns as they lack a ppt.

Splicing occurs both co- and post-transcriptionally: in general the 5’ introns are removed co-transcriptionally while the 3’ introns are often removed post-transcriptionally (36, 37). Direct links between the transcription and splicing machinery have remained elusive, although there are indirect links between the two processes (reviewed in (14, 15, 38)). Although the RNA pol II CTD is able to stimulate splicing in vitro and it interacts with splicing related factors in vivo, it does not interact directly with any splicing factors (39-41). The kinetics of transcription affect alternative splicing: slower rates of transcription increase exon-inclusion possibly by promoting the recognition of weak splice sites (42, 43). Constitutively spliced introns, however, are removed even at high rates of transcription. As described in section 1.1.1, splicing of the first intron is stimulated by the CBC, but subsequent splicing reactions are not.

**Splicing Control Elements**

In addition to the basic *cis*-acting elements mentioned above, there are additional sequences found in the exons and introns of a gene that promote splice site usage (ESEs
and ISEs respectively). In general, these elements are bound by members of the SR protein family. SR proteins have one or two RNA recognition motifs (RRMs) at their N-terminus, and an arginine-serine-rich (RS) domain at their C-terminus. SR proteins bound to ESEs may work in several ways to enhance splicing from adjacent 3’ or 5’ss. Through their RS domain they can interact directly with splicing factors such as U2AF to promote exon recognition, or they may interact indirectly with the splicing machinery via splicing co-activators (SRm160 and SRm300) (44-47). RS domains have also been shown to activate splicing by interacting directly with the RNA at the bps in the A complex (48, 49).

There are also a number of inhibitory elements, termed splicing silencers, found in exons and introns (ESS and ISS respectively). In general, these silencers are bound by members of the heterogenous nuclear ribonucleoprotein (hnRNP) protein family. HnRNPs are a diverse family of proteins that contain RNA- and protein-interacting domains and interact with high molecular weight nuclear RNA. HnRNPs may directly repress splicing by inhibiting splicing factor or SR protein binding. For example, the ppt binding protein (PTB) binds to the ppt, and competes with U2AF binding to inhibit splicing (50-52). Repression often requires cooperation between hnRNPs bound at multiple silencer elements and looping out of the intervening RNA (53, 54).

**Alternative Splicing**

Alternative splicing generates different mRNAs from a single gene, by joining different combinations of splice sites (reviewed in (55-57)). The resulting alternately spliced mRNAs encode different isoforms of a protein, each with potentially very different activities. Alternately, the mRNA may encode the same protein isoform but
contain different 5’ or 3’ untranslated regions. This may result in different mRNA localization, stability or translation. Studies have shown that anywhere from 40-80% of human genes are alternatively spliced, which magnifies the complexity, and coding potential, of the genome (58, 59). The coupling of a 5’ss and 3’ss does not occur until C complex formation, so E-, A- and B-complex formation are all potential targets of alternative splicing regulation. Alternative splicing is often regulated in a tissue-specific and temporal manner. Figure 1.2 illustrates some of the different types of alternative splicing that may occur.

**Exon Junction Complex**

Splicing alters the pre-mRNA by removing introns and joining exons, but it also deposits a protein complex, termed the exon junction complex on the spliced mRNA (60). The exon junction complex (EJC) is deposited on the spliced mRNA approximately 20-24 nucleotides upstream of exon-exon junctions (61, 62). Immunoprecipitation from *in vitro* splicing and crosslinking experiments, in conjunction with immunoprecipitation from *in vivo* nuclear/cytoplasmic fractionation experiments have revealed that the EJC is a dynamic structure. The core components of the EJC are eIF4AIII, Y14 and Magoh (62-66). It is postulated that the RNA helicase, eIF4AIII, is the EJC anchoring factor and that the Y14-Magoh heterodimer contacts the mRNA through it (66-72). In the nucleus other proteins associated with the EJC are the splicing co-activators SRm160, Pinin and RNPS1, the mRNA export adaptor proteins UAP56 and Ref/Aly, and the nonsense-mediated decay factor Upf3 (62-65, 73-75). UAP56, Ref/Aly, SRm160 and RNPS1 are removed from the mRNA during, or immediately after, transport to the cytoplasm. Upon export to the cytoplasm other factors join the EJC: PYM and Btz bind Y14-Magoh, and
Figure 1.2. Types of Alternative Splicing. The boxes indicate exons, the solid lines represent introns, and the dotted lines represent splicing. Some forms that alternative splicing can take are illustrated: (a) cassette exons, (b) intron retention, (c) alternative 5′ss, (d) alternative 3′ss, (e) mutually exclusive exons, (f) alternative promoters, (g) alternative polyadenylation sites. Adapted from (56).
Upf2 binds Upf3 (61, 66, 76-78). The cytoplasmic components of the EJC are removed through a pioneer round of translation. EJC deposition marks the mRNA for downstream processing, and is a major mechanism through which splicing affects subsequent mRNA export, surveillance and translation (reviewed in (79), and discussed in sections 1.2 and 1.3).

1.1.3. 3’ End Processing

Protein encoding mRNAs, except for replication dependent histone mRNAs, are modified at their 3’ end by addition of a polyadenosine, poly(A), tail. Prior to addition of the poly(A) tail the mRNA must be cleaved. Cleavage and polyadenylation are dependent on two RNA elements: an AAUAAA polyadenylation signal and a G/U-rich downstream sequence element (see Figure 1.3.a). Cleavage occurs predominantly after a CA dinucleotide between these two elements. The AAUAAA sequence is recognized and bound by the 160 kDa subunit of the cleavage and polyadenylation specificity factor (CPSF). The other CPSF subunits (-100, -73, -30 and hFip) increase the strength and specificity of the binding to AAUAAA, and the 73 kDa subunit was recently identified as the endonuclease that cleaves the RNA (80). The cleavage stimulatory factor (CstF) binds the downstream G/U-rich element via its 64 kDa subunit. CPSF and CstF exhibit cooperative binding to substrate mRNA (reviewed in (81, 82)). Cleavage factor (CF) I can bind RNA, and stabilizes the interaction of CPSF with the AAUAAA site (83). Symplekin, is a more recently identified member of the polyadenylation and cleavage machinery. It probably acts as a scaffold protein, which recruits CPSF and CstF into the same complex (84, 85). CF II and poly(A) polymerase (PAP) are also required for cleavage. After the mRNA substrate is cleaved, CPSF and PAP mediate the addition of
the poly(A) tail. As the poly(A) tail is elongated it is bound by poly(A) binding protein (PABP)-2, which enhances the processivity of PAP (86). After approximately 250 adenosine residues are added the reaction is complete.

There are many interactions between 3’end formation and transcription: RNA pol II stimulates 3’end processing and 3’end processing is required for transcription termination. 3’end processing is dependent on RNA pol II CTD both in vitro (87) and in vivo (18). CPSF and the CstF 50 kDa subunit both interact with the CTD (18, 88, 89). Evidence from immunofluorescence, in vitro binding assays and in vivo cross-linking followed by chromatin immunoprecipitations show that CPSF interacts with the CTD from the promoter all the way to the AAUAAA site (90, 91). Transcription termination requires a functional poly(A) site but not cleavage of the RNA (92, 93). The proximity of the CTD to the jaws with which RNA pol II clamps the DNA (94) suggests that dissociation of CPSF from the CTD, upon recognition of the poly(A) site, triggers a switch in pol II from processive to non-processive transcription.

The 3’end processing machinery also interacts with the 5’cap and the splicing machinery. The CBC stabilizes cleavage factor binding to the mRNA and stimulates cleavage (95). Splicing and polyadenylation of the terminal exon are strongly enhanced by each other (96). This may be mediated by U2AF-65 which has been shown to interact with the C-terminal domain of PAP (97). On the other hand, polyadenylation is inhibited by the proximity of a 5’ss. Inhibition is mediated by U1snRNP binding to the 5’ss (98). The 70kDa subunit of U1 snRNP inhibits PAP activity. This interaction inhibits polyadenylation at alternate poly(A) signals in the body of a gene, but does not affect polyadenylation at the terminal exon.
Figure 1.3. Messenger RNA 3’ end formation.
a. The polyadenylation signal (AAUAAA) is 10-30 nucleotides upstream of the cleavage site (usually after a CA dinucleotide). There is also a GU- or U-rich downstream element. The cleavage and polyadenylation machinery are illustrated: CPSF (light blue), CstF (red), CFI, CFII, symplekin and PAP. CPSF-73 is the endonuclease. Adapted from (86).
b. Histone 3’ end formation requires SLBP binding to the histone mRNA stem-loop. SLBP recruits U7snRNP (green=RNA, red=protein) which base pairs with the histone downstream element (HDE). The U7snRNP proteins ZFP100, Lsm11 and Lsm10 stabilize the stem-loop/SLBP/U7snRNP complex. This complex recruits an unknown cleavage factor which cleaves the mRNA between the stem-loop and the HDE. Adapted from (99).
Replication-dependent histone mRNAs end in a conserved stem-loop; they are the only known eukaryotic mRNAs that are not polyadenylated (reviewed in (99)). Histone 3’end formation is dependant on two RNA elements: the stem-loop and a purine-rich downstream element called the histone downstream element (HDE). The stem-loop is flanked by AC-rich nucleotides and together with the stem-loop this 25-26 nucleotide sequence is the binding site for stem-loop binding protein (SLBP). The HDE is complementary to the 5’end of U7 snRNP, and base pairing with U7 snRNP facilitates 3’end formation. U7 snRNP contains a zinc-finger protein, ZFP100, and 2 unique Sm proteins not found in other snRNPs: Lsm10 and Lsm11. ZFP100 interacts with SLBP, and this interaction has been postulated to stabilize the U7 snRNP-HDE interaction. U7 snRNP binding facilitates cleavage of the mRNA between the stem loop and the HDE, by as yet unidentified endonuclease. After cleavage, U7 snRNP dissociates while SLBP remains bound to the stem-loop of histone mRNAs as they are exported to the cytoplasm. Summarized in Figure 1.3.b.

1.2. RNA Export

1.2.1. Nuclear Pore Complex

In eukaryotes the nuclear membrane forms a physical barrier between the nucleus, in which pre-mRNAs are transcribed and processed into mature mRNAs, and the cytoplasm, in which mRNAs are translated into proteins. Transport of proteins and RNAs from one compartment to the other occurs through specialized channels, called nuclear pore complexes (NPCs) (reviewed in (100)). Embedded in the nuclear membrane, the NPC is composed of a nuclear basket, nuclear ring, spoke complex, cytoplasmic ring, and 8 short cytoplasmic fibrils (see Figure 1.4). The NPC has an eight-fold symmetry and is
Figure 1.4. Structure of the Nuclear Pore Complex (NPC). Embedded in the nuclear membrane, the NPC is composed of a nuclear basket, nuclear ring, spoke complex, cytoplasmic ring, and 8 short cytoplasmic fibrils. The NPC has an eight-fold symmetry and is composed of approximately 30 different proteins, called nucleoporins (present in 8, 16 or 32 copies per NPC). There are three classes of nucleoporins: FG nucleoporins contain phenylalanine-glycine (FG) rich domains, and line the active transport channel; nucleoporins lacking FG-rich domains are the main structural proteins of the NPC; and Nups are membrane proteins that anchor the NPC. Adapted from (128).
composed of approximately 30 different proteins, called nucleoporins, present in 8, 16 or 32 copies per NPC (101, 102). There are three classes of nucleoporins: FG nucleoporins contain phenylalanine-glycine (FG) rich domains, and line the active transport channel; nucleoporins lacking FG-rich domains are the main structural proteins of the NPC; and Nups are membrane proteins that anchor the NPC. Most nucleoporins are located symmetrically on both sides of the NPC, but some are localized asymmetrically and are responsible for conferring directionality to transport.

1.2.2. Transport Receptors

Transport through NPCs is mediated by soluble transport receptors. Transport receptors contain hydrophobic patches on their surface which can directly interact with phenylalanine residues of the FG nucleoporins (103-105). These transient interactions mediate passage of the receptor and its cargo through the NPC. By far the most abundant transport receptors in the cell belong to the karyopherin (or importin-β) family (reviewed in (106-108)). Karyopherins can recognize protein and RNA cargoes directly. Karyopherin transport is regulated by the GTPase Ran. The two regulators of Ran are differentially localized: RanGEF (Ran-GDP-exchange factor) is localized in the nucleus and RanGAP (Ran-GTPase-activating protein) is localized in the cytoplasm. The differential localizations of these regulators create a Ran gradient in which Ran is GTP-bound in the nucleus and GDP-bound in the cytoplasm. This gradient drives the movement of karyopherins from the nucleus to the cytoplasm and vice versa. Karyopherins which import cargoes to the nucleus (importins) release them when Ran-bound GDP is exchanged for GTP. Karyopherins which export cargoes from the nucleus (exportins) release them when Ran-bound GTP is hydrolyzed to GDP.
Figure 1.5. Common karyopherin-mediated RNA export pathways. tRNA, miRNA, snRNA and rRNA are exported from the nucleus by members of the karyopherin family of transport receptors. Exportins are Ran-GTP bound in the nucleus. In this state they can bind their RNA cargoes and mediate transport through the NPC, by interactions with the FG-rich domains of the FG-nucleoporins. Once in the cytoplasm, the Ran-GTPase Activating Protein (Ran-GAP) hydrolyses the GTP-moiety, causing the exportins to dissociate from their cargo. Adapted from (100).
Karyopherins are the primary nuclear export receptors for RNA (reviewed in (100)). Nuclear export of transfer RNA (tRNA) and microRNA (miRNA) is mediated by direct binding of the karyopherins exportin-t and exportin-5 respectively. SnRNA nuclear export is mediated by the karyopherin Crm1. In this instance, Crm1 interacts indirectly with the RNA through the CBC and PHAX. Ribosomal RNA (rRNA) is exported in immature ribosomal subunits by Crm1. In yeast the Tap/p15 homologues and another export factor, Arx1, also contribute to rRNA export, although it is unknown whether the same holds true in higher eukaryotes. Summarized in Figure 1.5.

Nuclear export of mRNA is mediated by the Tap/p15 transport receptor complex (109, 110). Tap/p15 is not a member of the karyopherin family, and therefore it is not regulated by Ran, but it is still able to interact with the FG-rich repeats of the FG nucleoporins. The Tap/p15-mRNA complex moves through the NPC by simple diffusion (111). Once in the cytoplasm, Tap/p15 dissociates from the mRNA prior to interaction of the mRNA with the translational machinery (112). Tap/p15 can bind directly to a subset of viral mRNAs which contain a constitutive transport element (CTE) (110), however Tap/p15 generally contacts cellular mRNA through export adaptors. There are two major classes of mRNA export adaptors: the TREX complex and SR proteins.

1.2.3. mRNA Export Adaptors

The human TREX (transcription/export) complex is composed of the stable THO complex (hTho2, hHpr1, fSAPR79, -35, and -24), the transport adaptor proteins Ref/Aly and UAP56, and a protein of unknown function, hTex1 (reviewed in (113, 114)). Ref/Aly can interact directly with Tap/p15 (115, 116), while UAP56, an RNA helicase, interacts with Ref/Aly (117, 118). The role of the TREX complex in general (and Ref/Aly and
UAP56 in particular) has often been hard to distinguish from the role of the EJC. For instance, Ref/Aly and UAP56 co-localize with splicing factors in nuclear speckles, are associated with the spliceosome, and have been identified as components of the EJC on spliced mRNAs in the nucleus (119, 120). Ref/Aly, UAP56 and EJC proteins also co-localize with spliceosome components at the site of endogenous β-globin transcription, but are not detected when a non-splicing mRNA was used (121). However, a recent paper revealed that the TREX complex is only deposited at the 5’end of an mRNA, while the EJC is deposited at each exon-exon junction (122). The authors went on to show that TREX recruitment required not only splicing, but also interactions between the CBC protein CBP80 and Ref/Aly. The TREX complex, and not the EJC, facilitates the interaction of spliced mRNAs with the export machinery. The presence of the TREX complex at the 5’end of an mRNA, may explain why export occurs in a 5’ to 3’ direction (24). Summarized in Figure 1.6a.

As the human TREX complex is recruited to an mRNA during splicing, there must be a TREX-independent manner for Tap/p15 to export endogenous intronless mRNAs. In Drosophila, RNAi was used to inhibit expression of TREX proteins, and TREX-independent mRNA export was seen (123). The additional Tap/p15 export adaptor was proposed to be SR proteins, as some SR proteins are bound to spliced mRNAs and exported to the cytoplasm as part of the mRNP (124). Subsequently, the shuttling SR proteins 9G8 and SRp20 were shown to support mRNA export (125, 126). These SR proteins interact directly with Tap/p15 in a phosphorylation-dependent manner: binding to Tap/p15 only in their hypophosphorylated form (127). Summarized in Figure 1.6b.
Dissociation of Tap/p15 complex from the mRNA soon after delivery to the cytoplasm serves as a molecular “ratchet” to prevent re-import of the mRNA (reviewed in (128, 129)). Two mechanisms have been proposed for Tap/p15 dissociation from the mRNA in the cytoplasm. The first involves hDbp5, a DEAD-box RNA helicase, that is recruited to mRNAs early in transcription. hDbp5 can shuttle between the nucleus and the cytoplasm but has a steady state localization at the cytoplasmic face of the NPC, bound to Nup214 (128, 130, 131). hDbp5 is activated by hGle1 and IP$_6$. hGle1 is tethered to hCG1 and Nup155 at the cytoplasmic face of the NPC (132) and evidence suggests that hDbp5 is activated at the cytoplasmic face of the NPC, as it enters the cytoplasm (133, 134). This may result in a localized conformational change in the RNA, and is postulated to remove Tap/p15 from the mRNA. Summarized in Figure 1.6.a.

The second mechanism of Tap/p15 dissociation has been described in yeast. In yeast the SR-like protein Nlp3 interacts with the yeast homologues of Tap/p15 in its hypophosphorylated form, and dissociates upon phosphorylation (135). An Nlp3 phosphatase, localized to the nucleus, generates a gradient: dephosphorylated-Nlp3 in the nucleus and phosphorylated-Nlp3 in the cytoplasm. As mentioned previously, the human SR proteins 9G8 and SRp20 also bind Tap/p15 in their hypophosphorylated forms (127). If, as in yeast, human SR proteins are phosphorylated in the cytoplasm it may cause dissociation of Tap/p15. Activation of hDbp5 and phosphorylation of SR proteins in the cytoplasm may control Tap/p15 export of mRNA in an analogous manner to the RanGTP-RanGDP gradient that regulates karyopherin binding to its cargo. Summarized in Figure 1.6.b.
Figure 1.6. Messenger RNA export pathways.

a) TREX-mediated mRNA export. The TREX complex, containing the export adaptor Ref/Aly, is deposited at the 5’end of the mRNA after splicing. Ref/Aly subsequently interacts with the transport receptor Tap/p15. Once exported, Tap/p15 and Ref/Aly are dissociated by activated hDbp5 (hDbp5*). hDbp5 shuttles on the mRNA from the nucleus to the cytoplasm, but is only activated in the cytoplasm, by hGle1 and IP6. b) The shuttling SR proteins 9G8 and/or SRp20 mediate export of unspliced mRNAs. They interact with Tap/p15 in their hypophosphorylated state. In the cytoplasm Tap/p15 probably dissociates when 9G8 and SRp20 are phosphorylated. Adapted from (100).
1.2.4. **Crm1 mRNA Export**

There are conflicting reports on the role of Crm1 in the export of cellular mRNAs. Initial studies in yeast and *Xenopus* concluded that mRNA export was not Crm1- or Ran-dependent (136, 137). These experiments looked at the effect of Crm1 transport inhibition on the localization of total poly(A)+ mRNA. However, experiments examining the transport of individual mRNAs in human cell lines have elucidated a role for Crm1 export of a subset of mRNAs. Crm1 export was found to play a limited role in the export of mRNAs containing AU-rich elements in their 3’ untranslated region (UTR) (138). Pp32 and APRIL were the adaptors mediating the interaction of Crm1 with HuR, which in turn binds the AU-rich elements. Crm1 has also been shown to mediate export of a distinct subset of mRNAs involved in cell cycle progression (139). In the nucleus these mRNAs are bound by eIF4E, not CBC, at the m7G cap, are exported to the cytoplasm in an eIF4E-dependent manner, and contain an approximately 50 nucleotide eIF4E-sensitivity element in their 3’UTR (139-141). Another recent article has shown that, in yeast, the Crm1 and Tap homologues both contribute to the export of an unspliced mRNA (142). There is a growing amount of evidence to suggest that subsets of mRNAs involved in similar biological processes are co-regulated, and that this is mediated by the proteins bound to elements in their UTRs (reviewed in (143-145)). Crm1-dependent mRNA transport may be a feature of one such mRNA regulon, or several different regulons, but further research is required to determine how frequently Crm1-dependent mRNA export occurs.

1.3. **Translation**

1.3.1. **Translation Initiation**
Translation initiation is a tightly controlled process and depends on the independent formation of two complexes: the 43S pre-initiation complex and eIF4F which binds the 5’cap (reviewed in (146-148), Figure 1.7). A ternary complex composed of GTP-bound eIF2 and the initiating methionyl-tRNA (Met-tRNAi) is formed. eIF2-GTP binds Met-tRNAi and recruits it to the 40S small ribosomal subunit to form the 43S pre-initiation complex. Additional components of the 43S pre-initiation complex are eIF3, eIF1, eIF1A and eIF5. The eIF4F complex contains eIF4E which interacts directly with the m7G cap. In addition to eIF4E, eIF4F contains the RNA helicase eIF4A and the scaffold protein eIF4G. eIF4F recruits the 43S pre-initiation complex to the capped mRNA through interactions between eIF4G and eIF3.

Once recruited to the 5’end of the mRNA, the 43S complex scans down the mRNA until it interacts with a canonical AUG start codon (Figure 1.7). Scanning of unstructured 5’UTRs can occur in the absence of ATP, but scanning through secondary structures requires ATP and eIF4A helicase activity (149). Base pairing between the start codon and Met-tRNAi stabilizes the interaction of the 43S subunit with the mRNA, and eIF2-GTP is hydrolyzed to eIF2-GDP by eIF5. This reaction releases eIF2-GDP, leaving Met-tRNAi in the ribosomal P-site, and facilitates recruitment of the 60S ribosomal subunit to form the 80S initiation complex. Formation of the 80S initiation complex is thought to release most of the remaining initiation factors. The 80S complex catalyzes the formation of peptide bonds, as eEF1 delivers aminoacyl-tRNAs to the A-site. The ribosome moves 5’ to 3’ along the mRNA, elongating the nascent peptide until an in frame termination codon is encountered.
Figure 1.7. Cap-dependent Translation Initiation.
The ternary complex is composed of the methionine (Met) initiator tRNA, and eIF2-GTP. The ternary complex interacts with the 40S ribosome and other initiation factors to form the 43S pre-initiation complex. The m7G cap is bound by eIF4E, which recruits eIF4G and eIF4A to form the eIF4F complex. Then the 43S pre-initiation complex binds to eIF4F via interactions between eIF4G and eIF3. The 43S pre-initiation complex scans along the mRNA until it interacts with an AUG start codon. This interaction causes the 60S ribosomal subunit to join the 40S subunit, and release of the remaining initiation factors. Adapted from (146).
Two important proteins involved in translation initiation are eIF4G and PABP1. eIF4G occurs in two isoforms, eIF4GI and eIF4GII, both of which can support translation initiation (150). eIF4G is a large scaffolding protein that contains a number of protein binding sites, including sites for: the translation initiation factors eIF4E, eIF4A, eIF3 and PABP1; the protein kinase Mnk1; and the decapping enzyme Dcp1 (151-156). PABP1 is the cytoplasmic poly(A) binding protein. PABP1 contains three domains: an N-terminus containing four RRM(s), a linker domain, and a C-terminal domain (PABC). The N-terminus of PABP1 binds to poly(A) sequences with high specificity, and mediates the interaction with eIF4G (154, 157, 158). According to the “closed loop” model of translation initiation, simultaneous eIF4G interaction with eIF4E and PABP1 circularizes the mRNA, and enhances translation initiation. Circularization would bring the dissociating ribosomal subunits into close proximity with the initiation site. The “closed loop” model proposes that the ribosomes are recycled from the 3’ end to the 5’ end. Translation usually involves a string of 80S ribosomes associating with an mRNA, and recycling ribosomes would enhance translation as initiation is a rate limiting step.

1.3.2. Translation Termination and mRNA Decay

Translation termination has been less extensively studied than initiation, but offers interesting insights into mRNA decay. The stop codon is recognized by both the small ribosomal subunit, through interaction with the 18S rRNA, and the eukaryotic release factor 1 (eRF1) (159, 160). eRF1 associates with the ribosomal A-site: it mimics the shape of a tRNA and recognizes all three stop codons through two loops located in its N-terminus (161-163). The C-terminus of eRF1 is in a stable complex with eRF3-GTP and eRF1 interaction with the termination codon and ribosome stimulates the GTPase activity
of eRF3 (164-167). eRF3-GTP hydrolysis is required for the cleavage of the ester bond by the large ribosomal subunit, and subsequent release of the nascent protein (168). The C-terminal domain of eRF3, which binds GTP and eRF1, is sufficient for the translation termination reaction (169). Summarized in Figure 1.8a.

A model of translation termination-coupled mRNA decay has emerged (170). In yeast strains expressing N-terminal deletions of eRF3, deadenylation is inhibited; and overexpression of the N-terminal domain prolongs the half-life of mRNA (170). The N-terminal domain of eRF3 interacts with the PABP1, through PABC (171). While the PABC domain has no effect on translation of an mRNA, it is required for mRNA deadenylation and decay (172, 173). Therefore, interactions between eRF3 and PABP1, which do not affect translation, provide a link between the recognition of the stop codon by eRF1 and deadenylation.

In mammalian cells, deadenylation precedes decapping in normal mRNA decay (reviewed in (174, 175), summarized in Figure 1.8.b). Deadenylation is carried out by three exonucleases with strong preferences for poly(A) sequences: PARN, Pan2-Pan3 and Ccr4-Caf1. PARN is localized to the nucleoplasm at steady state, and is not essential for mRNA decay (176, 177). By using RNAi and dominant-negative mutants of mRNA decay enzymes, decay was revealed to occur in two phases (177). The first phase is observed as the synchronous, stepwise deadenylation of the poly(A) tail from ~250 nucleotides to ~110 nucleotides, by Pan2-Pan3 (177). The second phase of deadenylation is mediated by Ccr4-Caf1, and it is accompanied by decapping by Dep1-Dep2 (177). Once the cap and poly(A) tail are removed, the body of an mRNA is rapidly degraded by Xrn1, the 5’-3’ exonuclease, and/or the exosome, the 3’-5’ exonuclease complex. The
Figure 1.8. Translation Termination and Deadenylation

a. All three stop codons are recognized by the N-terminus of eRF1. The C-terminus of eRF1 interacts with eRF3-GTP. Interaction with the stop codon and ribosome activates the GTP-ase activity of eRF3, causing release of the nascent peptide. The N-terminus of eRF3 interacts with Upf1 and PABP-1. These interactions facilitate deadenylation and decay. Adapted from (192).

b. Deadenylation occurs in two discrete stages. This is due to repression/stimulation by bound PABP-1. Adapted from (177).
length of the poly(A) tail probably affects deadenylase activity through the activity of bound PABP1. PABP1 stimulates the Pan2-Pan3 deadenylase (178, 179) and Pan3 interacts with the PABC-domain of PABP1 (180). However, PABP1 inhibits the activity of both Ccr4-Caf1 (181) and Dcp1-Dcp2 (182). As the poly(A) tail is shortened the amount of PABP1 decreases, therefore Pan2-Pan3 activity decreases and Ccr4-Caf1 and Dcp1-Dcp2 activities increase. The PABC-domain of PABP1 provides the link between translation termination and deadenylation as it interacts with both eRF3 and Pan3.

1.3.3. Nonsense-Mediated mRNA Decay

Surveillance of mRNAs is essential, as it ensures that mRNAs containing a premature termination codon (PTC) are not translated into truncated proteins which may have deleterious effects on the cell. One of the best understood and most well studied mechanisms of mRNA surveillance is nonsense-mediated mRNA decay (NMD) (reviewed in (183-185)). As the CBC-bound m7G cap emerges from the NPC, it is engaged by the translation initiation machinery in a “pioneer” round of translation (32, 186-188). During this pioneer round of translation, the translating ribosome removes EJCs as it progresses along the mRNA (32, 186). NMD occurs when the translating ribosome encounters a termination codon upstream of an EJC. With few exceptions, the PTC must be located more than 50-55 nucleotides upstream of an exon-exon junction to induce NMD (189, 190). If an mRNA is not targeted for degradation by the NMD machinery, the CBC is exchanged for eIF4E and translation proceeds as outlined in 1.3.1.

The key components of the human NMD pathway are the Upf proteins: Upf1, Upf2 and Upf3 (-a or -b). In the nucleus, Upf3 associates with a spliced mRNA as a member of the EJC. Upon transport to the cytoplasm Upf3 recruits Upf2 to the EJC. On
the other hand, Upf1 interacts with the N-terminus of eRF3 (191), as part of the SURF complex (Smg1, Upf1, eRF1 and eRF3) (192). Upf1 contains a number of N- and C-terminal serines which are phosphorylated by Smg1 (193-196). Upf1 can interact with Upf2 and Upf3. Smg1 phosphorylation of Upf1 occurs when the SURF complex interacts with the EJC (192). Phosphorylated Upf1 recruits Smg5, Smg6 and/or Smg7, which in turn recruit the mRNA degradation machinery (197-200). Exactly how the mRNA degradation machinery is recruited is not clear. In the case of Smg7, Smg7 interaction with the mRNA may localize the mRNA to P-bodies where degradation enzymes are concentrated (183). In yeast the first step of NMD is decapping, but in mammalian cells deadenylation precedes decapping (174, 175). In fact, in mammalian cells NMD occurs with the same biphasic dynamics as wild type mRNA decay, outlined in Section 1.3.2 (177).

A PTC may be generated as a result of improper RNA processing or naturally occurring alternative splicing events (201). As much as 10-20% of the transcriptome may be subject to regulation by NMD (202-204). For example, PTB mRNA is regulated by NMD (205). At high concentrations, the splicing factor PTB induces skipping of exon 11 of its own pre-mRNA. As a consequence of exon skipping, a PTC occurs in exon 12, inducing NMD. Therefore, NMD appears to be an important feedback mechanism which links translation, alternative splicing and mRNA decay. Using bioinformatics analysis of cassette exons, it has been predicted that more than 30% of mRNAs that are alternatively spliced are targets for NMD (206).

1.4. Translational Control

1.4.1. General Mechanisms of Translational Control
General cellular translation is controlled through the modulation of eIF4F binding to the m7G cap and eIF2 phosphorylation (summarized in Figure 1.9). A major target of translational control is the recognition of the m7G cap by eIF4E, which is present in limiting amounts. eIF4E activity is controlled by phosphorylation on serine 209, which increases eIF4E affinity for the m7G cap and stimulates translation (207, 208). Mnk1 protein kinase, which interacts with eIF4G, can phosphorylate eIF4E on serine 209 both *in vitro* and *in vivo* (209, 210). Phosphorylation of eIF4E is particularly important for translation initiation of mRNAs with long 5’UTRs, as increased eIF4E affinity results in increased eIF4A helicase association with the mRNA. Increased eIF4A association facilitates unwinding of long 5’UTRs and 43S scanning from the 5’ cap.

Formation of the eIF4F complex is regulated by a family of eIF4E binding proteins (4E-BP) 1-3. 4E-BP1 binds eIF4E, blocking interaction with eIF4G and therefore inhibiting translation (211-213). Hypophosphorylated 4E-BP1 binds very tightly to eIF4E but dissociates upon phosphorylation (212, 214, 215). Interestingly, in cells undergoing apoptosis, caspase cleavage of 4E-BP1 generates a variant which can not be phosphorylated and therefore shuts off cap-dependent translation (216). Two additional isoforms of eIF4E have been identified (eIF4E-2 and eIF4E-3) (217, 218). Both isoforms can bind the 5’ cap, although with less affinity than eIF4E itself, but neither can support translation (217-219). In *Drosophila* eIF4E-2 has been shown to inhibit translation of *caudal* mRNA, but the role of these two isoforms in human translation has yet to be determined (220). Control of translation initiation through eIF4E can occur by three general mechanisms: phosphorylation of eIF4E; disruption of eIF4F by 4E-BPs; and possibly competition for cap-binding by eIF4E-2 and -3.
**Figure 1.9.** General Mechanisms of Translational Control.

a. Control of eIF4E.

i. eIF4E phosphorylation on serine 209 by Mnk1 kinase increases affinity for the m7G cap. This is particularly important for translation of mRNAs with long 5’UTRs.

ii. Hypophosphorylated 4E-BPs bind tightly to the eIF4G binding site of eIF4E, blocking eIF4F formation and translation.

iii. eIF4E-2 and eIF4E-3 can bind the cap but cannot support translation.

b. Ternary complex formation is controlled through phosphorylation of eIF2α. Phosphorylated eIF2α-GDP is sequestered by eIF2B, not recycled into eIF2α-GDP. Adapted from (146).
A second major mechanism of translational control is the formation of the ternary complex (eIF2-GTP-Met-tRNAi). eIF2 is composed of three subunits (α, β, γ), and formation of the ternary complex is inhibited by phosphorylation of eIF2α on serine 51. Phosphorylation of eIF2α increases the affinity of eIF2-GDP for the guanine nucleotide exchange factor, eIF2B. This increase in affinity blocks the recycling of GDP for GTP, thereby blocking ternary complex formation, and translation initiation. The eIF2α protein kinase family phosphorylates the α-subunit of eIF2 in response to a variety of cell stresses (such as amino acid starvation and ER stress). PKR is a member of the eIF2α protein kinase family which is activated by binding to double stranded RNA (221, 222). PKR is ubiquitously expressed at low levels in most mammalian tissues and is also transcriptionally induced by the interferon anti-viral response (223). The cell has two major mechanisms to control translation: modulation of m7G cap recognition by eIF4E and phosphorylation of eIF2. These control mechanisms have global affects on mRNA translation.

1.4.2. Specific Mechanisms of Translational Control

Translation of a specific mRNA, or a subset of mRNAs, is controlled by trans-acting factors binding to cis-acting elements. These cis-acting elements are often located in the 3’UTR of an mRNA. A number of important message specific translational control mechanisms have been discovered in *Xenopus* and *Drosophila* (reviewed in (146)). In *Xenopus*, the cytoplasmic polyadenylation element (CPE) regulates translation of maternal mRNAs. The CPE is bound by CPE binding protein (CPEB), CPEB binds Maskin, and Maskin binds to eIF4E in competition with eIF4G (224). This prevents formation of eIF4F, and inhibits translation. In *Drosophila*, two analogous types of
inhibition occur. Smaug binds to the 3’UTR of unlocalized nanos mRNA, and recruits Cup which binds eIF4E and inhibits translation (225). Bicoid binds to unlocalized caudal mRNA and interacts directly with eIF4E to inhibit translation (226).

There are also some human examples of specific mRNA translational control mechanisms. AU-rich elements (AREs) confer mRNA instability and translational repression (reviewed in (227)). AREs are found in the 3’UTRs of cytokines and proto-oncogenes, tight regulation of which is required to avoid disease. Two examples of ARE-binding proteins which affect translation are T-cell-restricted intracellular antigen 1 (TIA-1) and the ELAV/Hu-type protein, HuR. TIA-1 binds to AREs located in TNFα and COX-2, and inhibits translation but not mRNA abundance. Translational repression may be due to inhibition of translation initiation and sequestration of the inhibited mRNAs in stress granules. HuR, on the other hand, stimulates the translation of p53 and p21 through binding to their AREs and increasing mRNA association with polysomes. hnRNP K and hnRNP E1 inhibit translation of 15-lipoxygenase mRNA by binding to the differentiation-control element (DICE) in its 3’UTR (228). In this instance, the 43S pre-initiation complex is recruited to the mRNA and scans to the start AUG, however recruitment of the 60S ribosomal subunit is blocked (229).

Another mechanism to control translation of specific mRNAs is through miRNAs (reviewed in (230)). Endogenous miRNAs are produced from transcripts that form stem loops. In the nucleus, the RNase III endonuclease Drosha and DGCR8 process the stem-loop into 65-75 nucleotide pre-miRNAs. In the cytoplasm, final processing into ~21-26 nucleotide miRNAs is completed by another RNase III endonuclease, Dicer. Dicer also facilitates incorporation of the miRNA and Argonaute (Ago) proteins into the RNA-
induced silencing complex (RISC). Ago proteins contain an miRNA binding PAZ domain (231, 232) and an RNase H-related PIWI domain (233). Endogenous miRNAs can mediate either the endonucleolytic cleavage (234, 235) or translational silencing of a target RNA (236, 237). The miRNA and Ago components of the RISC complex determine whether a target is cleaved or translationally silenced. Cleavage is generally favoured when there is perfect base-pairing between an miRNA and the target mRNA. Also, although there are 4 human Ago proteins, only Ago2 can cleave RNA (238, 239).

Translation inhibition by miRNAs does not affect the amount of target mRNA (240-242). In contrast to cleavage activity, inhibition can tolerate substantial mismatch between the miRNA and target mRNA. Also, all 4 human Ago proteins are believed to be capable of inhibiting translation. Multiple miRNAs bound to a single target may act in an additive manner to efficiently inhibit translation (243). The mechanism of translation inhibition is largely unknown. In mammalian cells, RISC may target mRNAs to P-bodies (244-247). RISC may inhibit translation initiation, as tethering eIF4E or eIF4G to the effected mRNA rescues translation (244). However, in C. elegans the lin-4 miRNA inhibited translation of the lin-14 mRNA, but did not affect lin-14 association with polysomes (248). In this case, miRNA appears to be inhibiting translation at a step post-initiation. Translation inhibition by miRNAs is important to understand, as up to 30% of human genes may be regulated by miRNAs (249).
1.5. HIV-1

1.5.1. HIV-1 Life Cycle

The mature HIV-1 virion is a spherical particle of approximately 100-120nm diameter (Figure 1.10). The outer envelope of the virion consists of a lipid bilayer of cellular origin and contains approximately 8-10 spikes made up of heterodimers of viral envelope (Env) proteins (250, 251). The Env heterodimers consist of two glycoproteins: gp41, a transmembrane domain anchored in the lipid bilayer which non-covalently binds gp120; and gp120 which contains the receptor binding domains of HIV-1. The exposed gp120 first binds to the CD4-receptor, inducing a conformational change, and exposing the gp120 co-receptor binding domain. The HIV-1 co-receptors are the chemokine receptors CCR5 and CXCR4. Concomitant binding to CD4 and CCR5 or CXCR4 by gp120 induces a conformational change in gp41, fusing the viral and plasma membranes. Gp41 forms pores in the plasma membrane, permitting entry of the HIV-1 core into the target cell (252-255).

The core of HIV-1 consists of three structural proteins processed from the precursor HIV-1 Gag polyprotein: matrix (MA), capsid (CA) and nucleocapsid (NC) (reviewed in (9)). MA forms a shell inside the lipid bilayer by interaction of its N-terminus with the inside of the membrane (256). CA assembles in hexameric rings to form the capsid surrounding the nucleocapsid (257). Approximately 1500-2000 copies of NC coat two RNA copies of the viral genome, forming a dense cone-shaped core (9). In addition, the core contains viral protease, reverse transcriptase, integrase, Vpu, Vif, Vpr, Nef and some cellular proteins. Upon entry of the core into the cytoplasm of a target cell, uncoating of the viral genome is facilitated by cellular factors, MA, Nef and Vif. The
Figure 1.10. Mature HIV-1 virion.
The outer envelope of an HIV-1 virion consists of a lipid bilayer of cellular origin and contains approximately 8-10 spikes made up of gp120/gp41 heterodimers. The matrix protein associates with the inside of the lipid bilayer. The core is surrounded by capsid protein arranged in hexameric rings. Approximately 1500-2000 copies of nucleocapsid coat two RNA copies of the viral genome, forming a dense cone-shaped core. In addition, the core contains viral protease, reverse transcriptase, integrase, Vpu, Vif, Vpr, Nef and some cellular proteins. Adapted from (9).
genome is reverse transcribed by HIV-1 reverse transcriptase using the host tRNA (Lys3-tRNA) as a primer. Reverse transcription generates a full-length dsDNA copy of the genome, which assembles into a pre-integration complex (PIC), and is subsequently directed to the nuclear membrane by Vpr (258). The PIC enters the nucleus through the nuclear pore (259, 260). The proviral dsDNA is integrated into the host chromosomal DNA by HIV-1 integrase (261-263). Integrated proviral DNA may persist as a latent infection, and act as a reservoir of virus that can not be eradicated by drug treatment.

Upon activation by pro-inflammatory cytokines, integrated proviral DNA is transcribed by cellular RNA pol II. Transcription generates a single 9 kb transcript which is processed into over 30 different mRNAs through alternative splicing. Three classes of HIV-1 mRNA are generated by alternative splicing: the unspliced 9 kb class encoding Gag and Gag-Pol; the incompletely spliced 4 kb class encoding Env, Vif, Vpu and Vpr; and the completely spliced 2 kb class encoding the regulatory proteins Tat, Rev and Nef. Initially, transcription occurs very inefficiently and only the 2 kb class of HIV-1 mRNA is exported and translated. Once sufficient amounts of Tat have accumulated, Tat goes to the nucleus where it binds to the TAR element in the proviral LTR and in conjunction with other cellular factors stimulates transcription (264-267). When enough Rev protein has accumulated, Rev facilitates export of the 9 and 4 kb HIV-1 mRNAs to the cytoplasm for translation (discussed in detail in section 1.5.3). The Env mRNA is translated in association with the endoplasmic reticulum, the precursor gp160 glycosylated and processed into gp120 and gp41. Gp120 and gp41 migrate and insert into the plasma membrane. The Gag and Gag-Pol polyproteins are translated on membrane free
ribosomes, and migrate to the plasma membrane. At the plasma membrane Gag, Gag-Pol, two copies of the 9 kb genomic mRNA, various cellular and viral accessory proteins assemble into an immature core. This complex buds through the plasma membrane generating an immature virion. Budding activates the HIV-1 protease, which cleaves the Gag and Gag-Pol polyproteins into the mature structural core and enzymatic proteins of the virus.

Current anti-HIV-1 drug targets fall into one of four major categories: reverse transcriptase inhibitors, protease inhibitors, integrase inhibitors and gp41-fusion inhibitors (reviewed in (9)). In order for HIV-1 to successfully replicate, stoichiometric amounts of each viral protein must be made (see section 1.5.4). Stoichiometry is maintained through the complex regulation of viral mRNA processing, export, and translation. Therefore, study of HIV-1 mRNA metabolism may lead to novel anti-retroviral treatments.

1.5.2. HIV RNA Processing

HIV-1 mRNAs are exported and translated by the host machinery, as a result they must be capped, spliced and polyadenylated. As mentioned above, the HIV-1 integrated provirus generates a single 9kb primary transcript which is processed into three major classes of mRNA: 9, 4 and 2 kb classes. To produce stoichiometric amounts of structural proteins, and replication competent viral particles, over 40 different mRNAs are generated through suboptimal and alternative splicing (reviewed in (268, 269)). The primary 9kb transcript contains four 5’ss (SD1-4) and eight 3’ss (SA1-3, SA4c,a,b, SA5, SA7) (Figure 1.11) (270, 271). While the four 5’ss have high constitutive activity, alternative splicing is regulated through the inefficient use of the eight 3’ss.
**Figure 1.11.** Genetic organization of HIV-1. 

a. The coding regions, and the reading frame of the HIV-1 proteins. b. The 9kb primary HIV-1 transcript is shown with the 5' cap and poly(A) tail. The 4 splice donors (SD) are indicated by solid arrows, the eight splice acceptors (SA) are indicated with hollow arrows. The enhancers and silencers are colour coded to show which splice sites they affect (see text for details). Adapted from (272).
Although the 3’ss contain suboptimal branchpoint or polypyrimidine tract sequences, they are primarily regulated through ESEs and ESSs located in adjacent exons. Kammler et al analysed the relative strengths of the 3’ss by in vivo splicing assays. These splicing assays were performed with and without the exonic regulatory sequences flanking the 3’ss (272). In the absence of flanking sequences, only SA2 and SA3 had any significant activity. However, in the presence of exonic regulatory sequences the activity of SA2 and SA3 are decreased and the strongest 3’ss is SA1. Under these conditions, SA7 and SA5 have significant activity, while SA4c,a,b only has significant activity if the GAR ESE downstream of SA5 is included.

The position of the regulatory splicing elements and the splice sites they affect are summarized in Figure 1.11. As with host ESSs, the HIV-1 ESSs interact with members of the hnRNP family of proteins which block assembly of the spliceosome and repress use of the upstream 3’ss. The HIV-1 ESSs interact with hnRNP A/B proteins (273-276), except ESS2p which interacts with hnRNP H (277). SA3 is regulated by ESS2p, ESS2 and ESE2. HnRNP A1 binds ESS2, and oligomerizes along the RNA (278, 279). Oligomerization of hnRNP A1 along the RNA blocks ESE2, the binding site of SC35 and SRp40 (280, 281). Overexpression of SC35 and SRp40 increases use of SA3, indicating that ESE2 enhances splicing by inhibiting hnRNP A1 oligomerization along the RNA (282). A similar mechanism regulates the use of SA7. At SA7, co-operative binding of hnRNP A1 at ESS3a,b and the ISS may loop out the splice site preventing recognition (276, 283). In the exon the primary binding site of hnRNP A1 is ESS3b, and it oligomerizes up and down the RNA (278, 279). Oligomerization is inhibited by SF2/ASF binding to ESE3 (279, 284, 285). So in general, HIV-1 ESSs bind hnRNP A/B proteins to
inhibit formation of the spliceosome; and HIV-1 ESEs bind SR proteins, block oligomerization of hnRNP A1, and thereby inhibit function of the adjacent ESS. Unlike ESE2 and ESE3, GAR does not appear to function by inhibition of an adjacent ESS. Instead, the GAR ESE binds SF2/ASF and SRp40 enhancing both SA4c,a,b and SD4 recognition (272).

The HIV-1 polyadenylation signal and GU-rich element are located in the long terminal repeat (LTR) and therefore are present at both the 5’ and 3’ end of the viral transcript. Use of the 5’ proximal polyadenylation signal would prevent viral replication and is inhibited by U1snRNP binding to SD1 (97, 286, 287). A protein component of U1snRNP, U1-70K, interacts with PAP and inhibits its activity (98, 287). As there is no 5’ss downstream of the 3’ proximal polyadenylation signal, polyadenylation and cleavage at this site is not inhibited. However, 3’ end formation of the 9kb and 4kb HIV-1 mRNAs does appear to have some unique features. hnRNP U has been implicated in the positive regulation of 3’ end formation of these HIV-1 mRNAs (288). Overexpression of the N-terminal fragment of hnRNP U inhibits export of the incompletely spliced HIV-1 mRNAs, in a polyadenylation signal-dependent manner. In HIV-1 reporter constructs, deletion of ESE3 inhibits splicing (276, 289, 290). These constructs are also retained at the site of transcription due to inhibition of 3’ end formation (276, 289, 290). Results indicate that ESE3 counteracts the activity of ESS3, indicating that ESS3 is an inhibitor of splicing and 3’ end formation. The exact mechanism of hnRNP U and role of ESS3 on proviral 3’ end formation have yet to be determined.

HIV-1 proteins are themselves able to affect the various stages of RNA processing. Tat binds guanylyltransferase and enhances capping \textit{in vitro} (291). Tat also
increases expression of CPSF, which in turn enhances polyadenylation and cleavage (292). Vpr dephosphorylates PAP, stimulating polymerase activity (293). Also, HIV-1 infection generally inhibits 3’ss recognition by inhibiting SR protein activity (294). Infection decreases the amount of SC35 and SF2/ASF, alters SR protein subcellular localization and decreases SR protein phosphorylation. The exact mechanism of HIV-1 inhibition of SR proteins is unclear, although it may be mediated through effects on SR protein kinases (294).

1.5.3. **HIV RNA Export**

The intron containing 9kb and 4kb classes of HIV-1 mRNA require the HIV-1 regulatory protein Rev to mediate their export (summarized in Figure 1.12, (295-297)). In contrast, completely spliced 2kb mRNAs are exported by the host export machinery, Tap/p15, as outlined in section 1.2. Retention of the 9kb and 4kb classes of mRNA may be attributed to partial assembly of the spliceosome (298) or nuclear retention sequences in the RNA (299-306). The Rev-responsive element (RRE) is an approximately 350 nucleotide stem-loop structure in the env coding region, and is present in all incompletely spliced HIV-1 mRNAs (297, 307). Rev binds to the RRE and facilitates export of the HIV-1 intron containing mRNAs via the Crm1 export pathway (308-310). Rev is an 116 amino acid phosphoprotein, which contains three important functional domains: the arginine-rich domain, the nuclear export signal (NES) and the multimerization domain (reviewed in (311)).

The arginine-rich domain, amino acids 38-49, acts as a nuclear localization signal (NLS) and an RNA binding domain. This domain interacts directly with importin-β (312, 313). Importin-β/Ran-GDP facilitates the import of Rev through the nuclear pore.
In the nucleus RanGEF converts Ran-GDP to Ran-GTP dissociating the Rev/importin-β complex. Released from importin-β, the arginine-rich domain is free to interact with the RRE (314). Rev binds with high affinity to stem IIB of the RRE (315), which contains a non-Watson-Crick G-G base-pair (316, 317). This forms an open major groove which is bound by the α-helical arginine-rich domain of Rev (316, 317). Rev binding to this high affinity site facilitates cooperative binding of multiple Rev molecules along the RRE (318-322).

In addition to its role as an NLS and in RNA binding, the arginine-rich motif and the surrounding sequences of Rev play a role in protein multimerization (323, 324). Importantly, Rev multimerization is required for nuclear import (325). Also, it has recently been shown that the arginine residues of the arginine-rich domain of Rev can be asymmetrically methylated by PRMT6 (326). PRMT6 methylation of Rev reduces its affinity for the RRE, and reduces expression from a Rev-dependent reporter (326).

The Rev NES, amino acids 71-82, is a leucine-rich domain which facilitates nuclear export (315, 327) through direct interactions with the exportin Crm1 (308-310). Rev only exports nascent transcripts, indicating that Rev is probably recruited to RNAs co-transcriptionally (328). Rev binding to the RRE exposes the Rev NES, which binds Crm1/Ran-GTP. Crm1/Ran-GTP mediates interactions with FG nucleoporins, specifically Nup214 and Nup98, to mediate transport to the cytoplasm (309, 329-331). In the cytoplasm interaction with RanGAP hydrolyzes Ran-GTP, dissociating the RRE/Rev/Crm1 complex. The arginine-rich domain of Rev is exposed and interacts with importin-β to continue the Rev shuttling cycle. Summarized in Figure 1.12.
Figure 1.12. Rev Import/Export Cycle.
Early in the HIV-1 replication cycle only the completely spliced 2kb viral mRNAs are exported from the nucleus via the Tap/p15 export pathway. The 2kb mRNAs are translated into Tat, Rev and Nef. Rev multimerizes, interacts with importin-b-Ran-GDP and enters the nucleus. Rev multimers are released from importin-b by the action of Ran-GEF. The Rev-responsive element (RRE) is an approximately 350 nucleotide stem-loop structure in the env coding region, and is present in all incompletely spliced HIV-1 mRNAs. Rev binds to the RRE and facilitates export of the HIV-1 intron containing mRNAs via the Crm1 export pathway. In the cytoplasm the RRE-Rev-Crm1 complex is dissociated by the action of Ran-GAP. The intron-containing HIV-1 mRNAs are translated and Rev is available to be re-imported to the nucleus. Adapted from (311).
Proper subcellular localization of Rev is important for its function, as two proteins which interfere with its steady state nuclear/nucleolar localization also impede its activity. DDX1 is a DEAD-box RNA helicase, which binds to Rev and the RRE (332). SiRNA knockdown of endogenous DDX1 caused cytoplasmic accumulation of Rev and abrogated its activity (332). The authors went on to show that in astrocytes, which are impaired for HIV-1 replication (333-336), both DDX1 and Rev show aberrant, cytoplasmic localization (337). Overexpression of DDX1 in astrocytes restored Rev localization to the nucleus and increased expression of HIV-1 p24 from a proviral reporter (337). Another protein which affects Rev subcellular distribution is nuclear factor 90 C-terminal variant (NF90ctv) (338). The authors found that overexpression of NF90ctv partially relocalized Rev to the cytoplasm, and moderately inhibited expression from a Rev-dependent reporter construct. NF90ctv can interact with the RRE and with Rev. Therefore the authors postulate that NF90ctv can compete with Rev for binding to the RRE and inhibit Rev multimerization, thereby inhibiting Rev function (338).

Once transit through the nuclear pore has occurred, multiple fates remain for HIV-1 RNA. However, several pieces of data indicate that movement away from the perinuclear space does not occur by simple diffusion, but depends upon two cis-acting elements and several trans-acting factors. Two trans-acting factors facilitate movement from and/or through the nuclear membrane: hRIP and DDX3. Depletion of the host factor hRIP (human Rev interacting protein) results in accumulation of unspliced viral RNA in the perinuclear space and loss of structural protein expression (339). A second host factor, DDX3, was identified by differential display as a Tat-induced cellular factor (340). The authors went on to show that DDX3 is an ATP-dependent RNA helicase that
shuttles between the nucleus and the cytoplasm, but it is localized to the cytoplasmic face of the nuclear pore at steady state. DDX3 interacts with Crm1 and Rev, is required for Rev mediated RNA export, and probably acts in an analogous manner to hDbp5 (340). Together these findings suggest the viral RNP must undergo some remodeling following export to allow it to continue on the path of translation or packaging into assembling virions.

HIV-1 genomic RNA contains two hnRNP A2 response elements (A2REs): A2RE-1 is located in the gag gene while A2RE-2 is in the overlap between the vpr and tat genes (341). In oligodendrocytes, hnRNP A2 was shown to bind the A2RE of myelin basic protein (MBP) mRNA, facilitate transport of the mRNA along microtubules and enhance translation (342-344). In oligodendrocytes, both HIV-1 A2REs are capable of transporting viral mRNA and a heterologous reporter construct (341). The authors went on to show that activity of the A2REs is context dependent: in the context of a Tat encoding mRNA the 5'UTR may interfere with hnRNP A2 recognition of A2RE-2 (341). Subsequent studies in COS7 cells confirmed the interaction of hnRNP A2 with HIV-1 A2REs, and mutation of A2RE induced a marked accumulation of the RNA in the nucleus, but did not have gross effects on p55 Gag expression (345). More recently, studies in HeLa cells show that hnRNP A2 knockdown induced perinuclear accumulation of HIV-1 9kb RNA around the microtubule-organizing centers (MTOC), decreased viral production, but had little effect on p55 Gag production (346). Overexpression of another host factor, Rab7-interacting lysosomal protein (RILP), recruits dynein motor complexes to the MTOC and also induces accumulation of genomic HIV-1 RNA at the MTOC (346). Together, these results indicate that hnRNP A2 mediates transport of HIV-1
A2RE-containing RNAs from the nuclear periphery through the cytoplasm along microtubules, probably in a dynein-dependent fashion.

1.5.4. HIV mRNA Translation

HIV-1 mRNA translation is entirely dependent on the translation machinery of the host cell. For HIV-1 to replicate successfully, it must usurp the host translation machinery such that the viral mRNA is translated in large quantity. However, this must be balanced with the requirement for host cell viability, which means that host cellular mRNAs must continue to be translated. The virus must avoid activating the host antiviral response, which would shutoff translation, induce apoptosis, and prevent the virus from completing its lifecycle. In addition, the virus must produce stoichiometric amounts of its various structural, enzymatic and regulatory proteins to replicate successfully (347). In order to achieve all of these goals HIV-1 has evolved various cis-acting elements and trans-acting factors.

Cis-acting Elements

Internal Ribosome Entry Site

HIV-1 uses the scanning mechanism in the translation of many of its mRNAs, however recent studies have highlighted additional strategies. In the case of HIV-1 genomic RNA, scanning from the 5’ cap is particularly problematic given the extensive secondary structure within the 5’UTR that includes the signals for primer binding (PBS), Tat transactivation (TAR), control of polyadenylation, dimerization (DIS) and packaging (psi sequence). Normally, such extensive secondary structure would negatively impact on the translational efficiency of any mRNA. HIV-1 has circumvented this problem by encoding an internal ribosome entry site (IRES) between +104 to + 336 nt of the RNA.
The presence of the IRES allows HIV-1 to bypass cap-dependent translation initiation. Although initial reports suggested that the HIV-1 5’UTR had no IRES activity \textit{in vitro} or \textit{in vivo} (348), subsequent experiments have determined that the HIV-1 IRES functions in a cell cycle dependent fashion, being most active during the G2/M phase of the cell cycle (349). Such a property is advantageous given that HIV-1 Vpr is known to arrest cells in this state (259). It may also permit viral protein synthesis under conditions of cell stress. A second IRES has also been mapped within the coding region of HIV-1 Gag (350). This IRES is unique as it is located within the Gag ORF. Also, it is able to stimulate translation from the upstream full-length Gag AUG start site and from a downstream AUG start site to induce expression of an N-terminal truncated Gag isoform (p40).

\textit{-1 Frameshift.}

The -1 frameshift is a mechanism used by HIV-1 to both maximize its coding potential and to coordinate the amount of Gag to Gag-Pol that is translated (reviewed in (351)). The ratio of Gag to Gag-Pol translated is approximately 20:1, and it is crucial for viral infectivity that this ratio be maintained (347). In order to translate \texttt{gag-pol} the ribosome must slip backwards 1 nucleotide just before encountering the \texttt{gag} stop codon. This -1 frameshift event is dependent on two RNA elements: the heptanucleotide slippery site and a pseudoknot. The pseudoknot is located 2 nucleotides downstream of the heptanucleotide slippery site, and causes ribosome pausing (352). The two ribosome bound tRNAs translocate -1 nucleotide, leaving the \texttt{gag} reading frame, and continue translating in the \texttt{pol} reading frame.

\textit{Leaky Scanning.}
While use of an IRES overcomes problems of secondary structure within the 5’UTR of HIV-1 genomic RNA, additional strategies have been suggested to allow high level synthesis of Env (gp160). In all mRNAs encoding Env, the AUG for the env reading frame is not the first AUG after the 5’ cap. Rather, it is preceded by the AUG for Rev and/or Vpu depending on the 3’ss used in generating the singly spliced RNA (SA4c,a,b versus SA5) (268, 269). Whereas one might anticipate that the presence of a 5’ AUG would negatively impact use of 3’ reading frames, recent evidence indicates that this is not the case for any of the env mRNAs. Using constructs representing individual forms of env mRNA, it was found that the level of gp160 synthesis was constant regardless of the number of upstream AUG codons (353). One mechanism to explain this observation is that initiation proceeds by a leaky scanning mechanism (269, 354). In support of this hypothesis is the finding that mutation of the upstream weak Vpu AUG to a canonical AUG suppresses env translation, while deletion of the weak Vpu AUG enhanced Env expression, indicating that the Vpu AUG is inefficiently recognized (355). Evidence also supports the possibility that env translation initiation occurs via discontinuous ribosome scanning from the 5’end or a ribosome shunt (353). A third model implicates translation termination and re-initiation. Directly upstream of the Vpu AUG is a highly conserved AUG codon, followed immediately by a stop codon, which overlaps with the weak Vpu AUG (355). Mutation of the highly conserved upstream AUG was found to negatively affect Env synthesis.

**Trans-acting Factors**

Along with its crucial role in transcription, Tat plays an important role in translational regulation. First and foremost, Tat acts as a pseudosubstrate for PKR, one of
the components involved in the cellular antiviral response (222, 356-359). HIV-1 TAR RNA has been shown to activate PKR, resulting in eIF2α phosphorylation and translational shutoff (360). Tat competes with eIF2α for binding to the C-terminus of PKR and can reverse the inhibition of translation mediated by active PKR. The importance of this activity cannot be overstated as PKR activation by HIV-1 TAR RNA, if unchecked, would block HIV-1 replication (360). The second role postulated for Tat in modulating translation is mediated through binding to the elongation factor (EF)-1δ subunit. EF-1 is required for delivery of aminoacyl tRNAs to the ribosome. The interaction of Tat with EF-1δ subunit was shown to inhibit cellular but not viral RNA translation (361).

The TAR RNA binding protein (TRBP) binds to the TAR element and synergistically activates transcription with Tat. TRBP has also been found to bind to PKR and inhibit its kinase activity (363). It has been reported that TRBP binding to TAR RNA restores translation inhibited by the TAR structure (362) and may help to sequester this dsRNA element away from PKR, thereby inhibiting the antiviral response induced by PKR binding to viral RNA.

In addition to Rev’s role in the export of viral RNA to the cytoplasm, it is also believed to enhance the translation of these mRNAs. It has been shown to promote the loading of gag-pol and env mRNA on to polysomes (365). Rev elevates the amount of gag mRNA in the cytoplasm by only 8-16 fold while the amount of Gag protein increases by over 800 fold (365). In addition, experiments using cytoplasmically transcribed env mRNA show that Rev is still able to enhance translation (366). mRNA must contain the
RRE for Rev-induced stimulation, but the RRE alone is not sufficient, indicating the requirement for additional cis-acting elements or trans-acting factors.

The HIV-1 protease is encoded by the pol gene and cleaves the Gag and Gag-Pol polyproteins into their mature forms. In addition, it cleaves host proteins chief among which are the translation factors eIF4GI and PABP1 (367-370). It cleaves eIF4GI in 3 places: twice between the eIF4E and first eIF4A binding site, and once in the second eIF4A binding site (370). HIV-1 protease cleaves PABP1 twice: once in RRM3 and once in the C-terminal domain (368). This activity is similar to that of the poliovirus 2A protease which cleaves eIF4GI, eIF4GII, and PABP1 (150, 371-375), resulting in rapid host translational shutoff and enhancement of IRES mediated translation. It is possible that HIV-1 protease cleavage of host factors similarly influences IRES-mediated translation, but further investigation is required. However, while HIV-1 protease cleavage of these translation factors has been shown to decrease global cellular translation, there is not a complete shutoff of cap-dependent translation (369).

Recent work by Huang et al. has suggested that resting CD4+ T cells may actively suppress HIV-1 gene expression at the level of translation (376). The group observed that resting CD4+ T cells had reduced ability to support expression of constructs containing the terminal exon of HIV-1. This activity was associated with increased expression of several miRNAs (miR-28, miR-125b, miR-150, miR-223, miR-382) that could bind to regions in the HIV-1 terminal exon. Selective inhibition of these miRNAs was associated with increased HIV-1 expression. Given that no change in viral RNA abundance or processing was observed, it is probable that these miRNAs are acting to inhibit translation (see section 1.4.2). TRBP is a component of RISC, and TRBP binding to TAR
RNA reduces RISC activity (364, 377). By disrupting RISC activity, HIV-1 may have evolved a strategy to minimize the effect of innate miRNA antiviral defenses.

1.5.5. HIV RNA Packaging

In the cytoplasm there are three major fates for HIV-1 genomic RNA: translation, packaging or degradation (378). In contrast, the singly spliced and multiply spliced viral RNAs are committed to translation which, in the case of env mRNAs, is done in association with the endoplasmic reticulum (378). Unspliced HIV-1 mRNA serves as both a template for synthesis of Gag and Gag-Pol, as well as a substrate for packaging. Packaging occurs through the interaction of the NC portion of Gag with the packaging signal (psi) within the mRNA. Therefore, it is anticipated that after activation of proviral DNA transcription, the bulk of the unspliced viral RNA would be committed to translation but as Gag protein levels increase there would be a shift to packaging. Consistent with this expectation, Gag has a bimodal effect on viral RNA translation; at low levels of Gag, translation of viral RNA is enhanced while at higher concentrations, it is inhibited (379). Stimulation maps to regions within MA while inhibition requires the NC sequence (379). The inhibitory effect is dependent on the integrity of the packaging signal within the RNA, supporting the hypothesis that it is the result of direct interaction between Gag and the RNA. There is research to support the model that in the case of HIV-1, there are not distinct pools of unspliced viral RNA committed to either encapsidation or translation (380). Rather, Gag is able to select from RNAs either actively engaged in translation or within the free RNP pool. Alternately, there is evidence that the Gag-genomic RNA interaction first occurs within the MTOC and perinuclear space to initiate the packaging process (381).
HnRNP E1 and Staufen are two endogenous proteins which are incorporated into HIV-1 virions and affect the encapsidation versus translation of HIV-1 genomic RNA (382-385). HnRNP E1 overexpression inhibits both Gag and Env synthesis, and conversely knockdown of hnRNP E1 enhances Gag and Env synthesis (382). This ability to suppress viral protein synthesis does not involve any change in viral RNA abundance, splicing or cytoplasmic accumulation. Given the effects on translation and presence in viral particles, the authors propose that hnRNP E1 may inhibit HIV-1 RNA translation initiation and promote encapsidation. Staufen is also present in the viral particle and was shown to selectively bind the 9kb HIV-1 genomic RNA and Gag (383, 384). While overexpression of Staufen increases HIV-1 RNA incorporation into particles, it leads to reduced infectivity (383). Additional investigations have shown that Staufen affects the oligomerization of Gag (385). Both overexpression and knockdown of Staufen increase Gag multimerization, but have deleterious affects on viral infectivity; indicating once again that appropriate stoichiometry is required for successful HIV-1 replication (383-385).

1.6. Sam68 and Sam68ΔC

Sam68 is a member of the STAR (signal transduction and activation of RNA) family of proteins (reviewed in (386)). Proteins belonging to this family of proteins contain four major domains: the GSG, proline-rich, arginine-rich, and tyrosine-rich domains (Figure 1.13). The GSG domain was named after the first three proteins identified that contain this conserved domain (GRP33, Sam68, GLD-1). The GSG domain is an RNA binding domain, approximately 200 amino acids long, which contains an hnRNP K homology (KH) RNA binding motif and conserved N- and C-terminal
flanking sequences (387, 388). Both the KH-domain and the flanking sequences are required for RNA binding activity (389, 390). Bacterially produced Sam68 used in SELEX (systematic evolution of ligands by exponential enrichment) experiments identified A/U-rich sequences, particularly UAAA and UUUA, as high affinity RNA targets (390). Endogenous Sam68 has been shown to bind with high affinity to RNA homopolymers poly(A) and poly(U) (389, 391). In addition, in vivo experiments have shown Sam68 binding to elements in the 3’UTR of a number of endogenous mRNAs, including β-actin (392). The GSG domain has been shown to mediate oligomerization, and Sam68 is believed to bind RNA as an oligomer (389).

The proline- and tyrosine-rich domains of Sam68 mediate a number of protein-protein interactions. There are six proline-rich motifs in Sam68 (P0-P5), and these mediate interactions with SH3 and WW domains. A subset of the proteins that bind to the proline-rich motifs of Sam68 through their SH3 domains are the tyrosine kinases: p60src (391, 393), p59fyn (394) and Sik/BRK (395), and the adaptor protein Grb-2 (396). Sam68 proline-rich motifs P3 and P4 can also bind the WW domains of FBP21 and FBP30 (397). Sam68 contains 16 tyrosine residues in a 50 amino acid region near the C-terminus. These tyrosine residues are phosphorylated by the tyrosine kinases mentioned above (p60src, p59fyn and Sik/BRK). Tyrosine-phosphorylated Sam68 can interact with SH2 domain-containing proteins, including the aforementioned tyrosine kinases. Oligomerization is disrupted by phosphorylation of the tyrosine-rich domain (389). In the nucleus Sam68 probably oligomerizes, binds RNA and WW motifs; and in the cytoplasm Sam68 binds SH3 domain-containing proteins, is phosphorylated and binds SH2 domain-containing proteins (reviewed in (386)).
Figure 1.13. Sam68 and Sam68ΔC Domain Structure. Sam68 contains a GSG domain, six proline-rich motifs (P0-P5), a tyrosine-rich domain (YY), a nuclear localization signal (NLS), and two RGG-boxes (hatched boxes). The GSG domain contains an embedded hnRNP K (KH) RNA binding domain. The entire GSG domain is required for RNA-binding activity of Sam68. a. Sam68. b. Sam68ΔC. Adapted from (386).
Sam68 also contains two RGG boxes which can be methylated in vivo by PRMT1 (398). Hypomethylated Sam68 is mislocalized to the cytoplasm, but methylation does not affect Sam68-protein or -RNA interactions (398). As arginine methylation is most likely irreversible, arginine methylation may occur as part of the constitutive maturation of Sam68 (399). The RGG boxes also contribute to homopolymeric RNA binding (389). Sam68 is predominantly localized to the nucleus, due to an NLS located in its C-terminal 24 amino acids (400).

Sam68 was first identified as a 62 kDa tyrosine-phosphorylated protein, in cells transformed with tyrosine kinase oncogenes (401, 402). It was subsequently identified as a substrate of p60src during mitosis, and migrated as 68 kDa (thus Src-associated during mitosis, 68 kDa gives Sam68) (391, 393). As a substrate of Src during mitosis, Sam68, was proposed to function in cell cycle and signaling. Additional roles for Sam68 have been elucidated in cell growth, tumorigenesis and alternative splicing (reviewed in (386)).

Importantly, Sam68 and a deletion mutant that removes the last 100 amino acids, Sam68ΔC, have potent affects on HIV-1 Rev activity. Initially, overexpression of Sam68 was found to stimulate expression of Rev-transported HIV-1 mRNAs, and the authors asserted that Sam68 could functionally substitute for Rev (403). Sam68 interactions with two endogenous proteins, hnRNP K and hsp22, inhibits its enhancement of Rev-mediated RNA transport (404, 405). Sam68 acting as a substitute for Rev is controversial, and has been disputed in subsequent studies (406-408). These discrepancies may be due to the use of different cell types and HIV-1 reporter constructs. Sam68 has been shown to enhance 3’ end processing of unspliced HIV-1 mRNAs (290). This provides an explanation for Sam68 stimulation of Rev-mediated export other than as a functional substitute for Rev.
By enhancing 3’ end formation, Sam68 increases the amount of RNA available for Rev-transport to the cytoplasm, and subsequent translation.

Many studies have indicated that endogenous Sam68 is required for expression of HIV-1 Rev-transported mRNAs. Astrocytes are impaired for HIV-1 replication (333-336). Astrocytes have reduced expression of Sam68, and overexpression of Sam68 is able to partially rescue the block to HIV-1 replication (409). In HIV-1 replication competent cell lines (293, 293T and Jurkat) Sam68 knock down by antisense RNA impedes Rev-Crm1 mediated transport (410). The use of Sam68 antisense RNA for knockdown was controversial. Sam68-like mammalian protein-1 and -2 (Slm-1 and -2) are also able to stimulate Rev-mediated export (403, 406), and due to nucleotide sequence homology with Sam68 are likely to be knocked down by an antisense RNA. However, another group confirmed that Sam68 was required for Rev-mediated export using Sam68 specific siRNA (411).

The Rev-inhibitory Sam68 deletion mutant, Sam68ΔC, is localized to the cytoplasm (403). Previous work in our lab showed that Sam68ΔC bundled the incompletely spliced HIV-1 mRNAs on the outer periphery of the nucleus (406). Our lab went on to show that when a heterologous nuclear localization signal (NLS) is fused to Sam68ΔC it relocalized to the nucleus and stimulated Rev activity to a similar extent as full-length Sam68. This indicates that the localization of the protein determines its activity, and that Sam68ΔC inhibits an RNA processing step downstream from the step influenced by full-length Sam68. Zhang et al show that Sam68ΔC inhibits HIV-1 Rev-dependent mRNA export to the cytoplasm (412). They go on to map the domain of Sam68ΔC required for HIV-1 inhibition to amino acids 269-321 (412). They show that
Sam68ΔC can partially relocalize exogenous GFP-Sam68 to the cytoplasm, while mutants lacking amino acids 269-321 can not (412). They postulate that the relocalization of Sam68 to the cytoplasm depletes nuclear Sam68, inhibiting Rev activity to levels seen in Sam68 deficient cell lines such as astrocytes (409, 412).

1.7. **Objectives and Rationale**

Current anti-HIV-1 drugs fall into one of four major categories: reverse transcriptase inhibitors, protease inhibitors, integrase inhibitors and fusion inhibitors (reviewed in (9)). Patients are generally treated with a combination of three or more of these drugs in a treatment regime termed HAART (highly active anti-retroviral therapy). There are two problems currently facing anti-HIV-1 treatment. First, although HAART is successful in reducing the circulating viral load, it does not eradicate the viral reservoir of latently infected cells in the host. Second, due to the high error rate of the HIV-1 reverse transcriptase, HIV-1 mutates very quickly. As a result, there are many multi-drug resistant strains of HIV-1. Sam68ΔC, potently inhibits the metabolism of the Rev-transported HIV-1 mRNAs. This is a stage of the viral life-cycle not targeted by existing anti-retroviral therapeutics, but shows potential for inhibition. Anti-retroviral strategies that target the metabolism of the 9 and 4 kb classes of HIV-1 mRNA may offer two advantages over existing drug treatments. First, if the metabolism of these HIV-1 mRNAs uses unique combinations of host proteins, not used by cellular mRNAs, then host proteins can be targeted. This is advantageous as it would probably be very difficult for HIV-1 to mutate around this treatment. Second, drugs targeting the post-integration stage of infection would be efficacious in latently infected cells which have been activated.
The objectives of my research have been three-fold. The first objective was to determine whether or not Sam68ΔC inhibition of HIV-1 Rev-transported mRNAs was specific, and to map the regions of Sam68ΔC required for this inhibitory activity. The second objective was to determine what step of HIV-1 mRNA metabolism Sam68ΔC inhibited, and how this inhibition was manifested. The third objective was to determine how HIV-1 Rev-transported mRNAs are distinguished from host mRNAs at a molecular level. We hoped to discover novel aspects of HIV-1 mRNA metabolism, by using Sam68ΔC as a tool to inhibit this process. This would add to our fundamental understanding of HIV-1 mRNA metabolism, and could possibly lead to novel anti-HIV-1 strategies.
Chapter 2: Structure and Function Analysis of Sam68ΔC

Kim Marsh, Meredith McLaren, Vanessa Soros and Alan Cochrane

All of the experiments in this chapter were performed by Kim Marsh except for the experiments depicted in Figure 2.1, which were performed by Meredith McLaren.
2.1. Abstract

Replication of HIV-1 requires nuclear export and translation of the incompletely spliced 4 and 9 kb classes of HIV-1 mRNA, which encode the structural and enzymatic proteins of the virus. HIV-1 Rev binds to the Rev-responsive element (RRE) contained in the introns of incompletely spliced HIV-1 mRNAs and mediates their nuclear export via the Crm1 pathway. Sam68ΔC is a C-terminal deletion mutant of the endogenous human protein Sam68, and has been shown to be a potent inhibitor of Rev-dependent reporters. Previous work from our lab showed that Sam68ΔC colocalized with the Rev-exported mRNAs in perinuclear bundles. In this study we have performed deletion analysis of Sam68ΔC, and determined the minimal mutant required for inhibition of Rev-dependent expression is Sam68Δ14Δ(45-54)-300. We have also determined that the amino acids 15-28 and 263-299 are required for inhibition: neither Sam68Δ28ΔC nor Sam68:5-262 are able to inhibit Rev-dependent RNA expression. Sam68ΔC inhibition is specific to RRE/Rev/Crm1 transported RNAs. The Rev/Crm1 exported reporter construct GagRRE is inhibited while the Tap/p15 transported GagCTE is not. We have determined that sequestration of mRNA in perinuclear bundles is not sufficient for inhibition as Sam68Δ28ΔC and Sam68:5-262 both bundle the Rev-transported RNA while enhancing its expression. In addition, Sam68ΔC bundles the GagCTE reporter but does not inhibit its expression. This indicates that Sam68ΔC requires properties in addition to perinuclear bundling of the target mRNA to inhibit expression of the mRNA.
2.2. Introduction

Successful HIV-1 replication requires that stoichiometric amounts of its structural and enzymatic proteins be produced. To achieve this goal, the 9kb primary transcript is suboptimally spliced into over 40 different mRNAs (reviewed in (268, 269)). These mRNAs fall into 3 classes: the unspliced, 9 kb class, encoding Gag and Gag-Pol; the singly spliced, 4 kb class, encoding Vif, Vpr, Vpu and Env; and the multiply spliced, 2 kb class, encoding Tat, Rev and Nef. Initially, only the multiply spliced 2 kb class of mRNA is transported from the nucleus to the cytoplasm, using the host Tap/p15 export pathway. In the cytoplasm, these mRNAs are translated into Tat, Rev and Nef. Rev multimerizes, is imported into the nucleus, and binds to the Rev-responsive element (RRE) of the intron containing 4 and 9 kb mRNAs. Rev multimerizes along the RRE, and interacts with Crm1 which facilitates transport of the Rev-mRNA complex into the cytoplasm for transport (reviewed in (311)).

Sam68 is a member of the STAR (signal transduction and activation of RNA) family of proteins (reviewed in (386)). Sam68 was found to potently stimulate expression from Rev-dependent reporters, and in some instances to functionally substitute for Rev (403). Sam68 is localized to the nucleus, but a C-terminal truncation mutant (Sam68ΔC) is localized to the cytoplasm and was found to be an inhibitor of Rev function (403). Previous work in our lab showed that Sam68ΔC bundled the incompletely spliced HIV-1 mRNAs on the outer periphery of the nucleus (406). Addition of a heterologous nuclear localization signal (NLS) to Sam68ΔC relocalized it to the nucleus and stimulated Rev activity to a similar extent as full-length Sam68. This observation indicates that the
localization of the protein determines its activity, and that Sam68ΔC inhibits an RNA processing step downstream from the step influenced by full-length Sam68.

In this chapter we have performed deletion analysis of Sam68ΔC, to further define the regions required for its inhibition of HIV-1. Overexpression of the Sam68 variants comprising only the GSG-domain stimulates expression from a Rev-responsive reporter. Interestingly, these mutants are whole cell/cytoplasmic, and are able to bundle the Rev-responsive mRNAs in the perinuclear region. Sam68ΔC is a dominant inhibitor of the stimulatory activity. Immunofluorescence and in situ hybridization analysis reveals that Sam68ΔC is able to bundle the RRE-containing mRNAs and the cytoplasmic stimulatory mutants in the perinuclear region. Further deletion analysis revealed that the minimal mutant for inhibition is Sam68Δ14Δ(45-54)-300. Sam68ΔC and the minimal mutant inhibit expression from a Rev-responsive reporter, but not a Tap/p15 exported reporter construct, indicating that inhibition is specific to the RRE/Rev/Crm1 export pathway.

2.3. Materials and Methods

2.3.1. Expression constructs

The following constructs have been previously described: SV-Hygro, SV-H6Rev, CMVmyc3xterm (289), pDM128 (413), pgTat (297), Bl-env-HindIII (305), Sam68 and Sam68ΔC (406). Gag-CTE and Gag-RRE plasmids were provided by Dr. K. Boris-Lawrie, Ohio State University. Sam68:97-255 and Sam68:97-280 were generated by PCR using the forward primer: 5’-AAA CTG CAG ATG GAG CCG GAG AAT-3’ and the reverse primer: 5’-CGG GAT CCG GCA GCT CCT CGT CCT CAC-3’. The PCR product was digested with PstI and BamHI, and ligated into PstI/BamHI digested
CMVmyc3xterm. Due to PCR errors, the product was truncated at the C-terminus generating two clones Sam68:97-280 and Sam68:97-255. To make HA-tagged Sam68ΔC, the following oligos were used: 5’-CGA TAT GGC TTA CCC ATA CGA TGT TCC AGA TTA CGC TAA GCT TCC TGC A-3’ and 5’-GGA AGC TTA GCG TAA TCT GGA ACA TCG TAT GGG TAA GCC ATA T-3’. The oligos were annealed and ligated into Clal/BamHI digested Sam68ΔC.

The C-terminal deletion mutants of Sam68ΔC were generated by PCR using the forward primer: 5’-CCA TTA ACG CAA ATG GGC GGT A-3’, and the reverse primers: 5’-CCG CTC GAG AAC AGG TGG AGG-3’ (Sam68:5-300) and 5’-CCG CTC GAG AAC CAA AGC TCC-3’ (Sam68:5-314). The amplicons were digested with EcoRI and XhoI and ligated into EcoRI/XhoI digested Sam68ΔC. Sam68:5-262 was made by EcoRI and EcoRV digest of Sam68ΔC and ligated into the EcoRI and SmaI sites of CMVmyc3xterm. To generate Sam68Δ14ΔC Quickchange mutagenesis (Stratagene) was carried out using Sam68ΔC as a template and the following primers: 5’-GGG AAT TCG AGA TCG GGC CGC AGC TG-3’ and 5’-GCA GCT GCG GCG CTA CCT TCG TGA ATT CCC-3’. The other N-terminal deletion mutants were generated by PCR using the reverse primer: 5’-CGG GAT CAT GGA TGT GCA CCT CT-3’ and the forward primers: 5’-CGG AAT TCC CCT CGG TGC GTC TGA C-3’ (Sam68Δ28ΔC), 5’-CGG AAT TCC CCA GAG GAG GCG CTC GTG-3’ (Sam68Δ50ΔC), and 5’-CGG AAT TCC ACG CGA CGG TGG GTG-3’ (Sam68Δ75ΔC). The amplicons were digested with EcoRI and BamHI and ligated into CMVmyc3xterm. Sam68Δ96ΔC was generated by the same reverse primer and the forward primer 5’-AAA CTG CAG ATG
GAG CCG GAG AAT-3’. The PCR product was digested with PstI and BamHI and ligated into PstI/BamHI digested CMVmyc3xterm.

To generate internal deletions, Quickchange mutagenesis (Stratagene) was carried out using Sam68ΔC as a template and the following primer pairs: Δ(15-24): 5’-CGC CTC ACC CGG TCC GGT GCC CAC CCC TCG-3’ and 5’-CGA GGG GTG GCC ACC GGA CCG GGT GAG GCGR3’; Δ(45-54): 5’-GCT TCC TCA CCG GCC CGC TCG GGC TCT GCC C-3’ and 5’-GGG CGA GGC CCG AGC CCG GTG AGG AAG C-3’.

To generate point mutants, Quickchange mutagenesis (Stratagene) was carried out using Sam68ΔC as a template with the following primer pairs: 5’-ACC CGG TCC TCG GGC GCC AGC TGC TCC-3’ and 5’-GTC CTT GGA GCA GCT GGC GCC CGA GGA-3’ (Sam68ΔC-R17A), 5’-TCC TCG GGC CGC AGC GCC TCC AAG GAC CCG TCA-3’ and 5’-TGA CGG GTC CTT GGA GGC GCC CGA GGA-3’ (Sam68ΔC-C19A), 5’-GGC CGC AGC TGC TCC GCC GAC CCG TCA GGT GCC-3’ and 5’-GGC ACC TGA CGG GTC GGC GGA GCA GCT GCG GCC-3’ (Sam68ΔC-K21A), and 5’-CGC AGC TGC TCC AAG GCC CCG TCA GGT GCC CAC-3’ and 5’-GTG GCC ACC TGA CGG GCC CTT GGA GCA GCT GCG-3’ (Sam68ΔC-D22A).

**2.3.2. Cell lines and transfections**

HeLa, 293T and COS7 cells were maintained in Iscove’s modified Dulbecco’s medium (IMDM) supplemented with 10% fetal bovine serum (FBS) (v/v), 50µg/mL gentamycin sulfate and 2.5µg/mL amphotericin B. Vectors were introduced to 293T cells by calcium phosphate transfection (414). Vectors were introduced to HeLa cells by
Fugene 6 transfection reagent (Roche) following manufacturer’s protocol. Cells were harvested two days post-transfection.

2.3.3. Antibodies

The following antibodies were used: mouse anti-myc (Invitrogen), mouse anti-tubulin (Sigma), rabbit anti-HA (Zymed) and mouse anti-p24 (clone 183-1112-5C). The following secondary antibodies were used: donkey anti-mouse IgG conjugated HRP, donkey anti-rabbit IgG conjugated Texas red, donkey anti-mouse IgG conjugated Texas red (Jackson Immuno Research) and sheep anti-digoxigenin conjugated FITC (Boehringer Mannheim).

2.3.4. CAT Assays

2x10^5 293T cells were transfected in triplicate, per condition. For Rev-dependent expression cells were transfected with: 0.125 µg pDM128, 0.025 µg SV-Hygro or SV-H6Rev, and 0.5 µg pcDNA3.1, Sam68, Sam68:97-255 or Sam68:97-280. For analysis of Sam68ΔC dominance 2x10^5 COS7 cells were transfected in triplicate, per condition. Cells were transfected with: 0.125 µg pDM128, 0.025 µg SV-Hygro or SV-H6Rev, 0.5 µg pcDNA3.1, Sam68, Sam68:97-255 or Sam68:97-280, and 2.0 µg pcDNA3.1 or Sam68ΔC. CAT assays were performed as previously described (415). Conversions were normalized by protein concentration as determined by Bradford assay (BioRad).

2.3.5. Western Blots

6x10^5 293T cells were transfected as follows for Rev-dependent expression: 0.5 µg GagRRE, 0.1 µg SV-Hygro or SV-H6Rev, and 2.0 µg pcDNA3.1, Sam68ΔC or mutants thereof. For Rev-independent expression cells were transfected with: 2.0 µg GagCTE, and 2.0 µg pcDNA3.1, Sam68ΔC or mutants thereof. Cells were harvested in
whole cell lysis buffer (150mM NaCl, 10mM Na$_2$HPO$_4$, 1% Triton X-100 v/v, 0.1% SDS w/v, 0.2% sodium azide w/v, 0.5% sodium deoxycholate w/v, 1mM sodium orthovanadate), fractionated on SDS-PAGE gels and transferred to PVDF membrane (Pall Life Sciences). Bound antibodies were detected using Western Lightning (Perkin Elmer).

### 2.3.6. In Situ Hybridization and Immunofluorescence.

3x10$^5$ HeLa cells were transfected. For examination of Sam68ΔC dominance, cells were transfected with: 0.25 µg SV H6Rev, 1.25 µg pgTat, 5.0 µg myc-tagged Sam68, Sam68ΔC, Sam68:97-255 or Sam68:97-280, and 5.0 µg HA-Sam68ΔC. For examination of Sam68ΔC mutants ability to bundle unspliced HIV-1 mRNA, cells were transfected with: 0.25 µg SV H6Rev, 1.25 µg pgTat and 5.0 µg Sam68ΔC, Sam68Δ14Δ(45-54)-300, Sam68Δ28ΔC or Sam68:5-262. For examination of Sam68ΔC ability to bundle GagCTE mRNA, cells were transfected with: 1.25 µg GagCTE and 5.0 µg pcDNA3.1 or Sam68ΔC. 48 hours post-transfection, cells were fixed for immunofluorescence or in situ hybridization as previously described (305). Digoxigenin-labeled Env-HindIII probe (antisense to HIV-1 env mRNA) and digoxigenin-labeled Gag probe (antisense to HIV-1 gag mRNA), were made with 10 x digoxigenin RNA labeling kit (Roche). RNA probe was detected with anti-digoxigenin conjugated FITC (Boehringer Mannheim). Myc-tagged proteins were detected with monoclonal anti-myc antibody (Invitrogen). HA-tagged Sam68ΔC was detected with polyclonal anti-HA antibody (Zymed). Immunofluorescence was detected using a Leica DMR microscope at either 400x or 630x magnification.

### 2.3.8. Statistical Analysis
The data are presented as mean +/- one standard deviation. Data were compared using Student’s t-test. * indicates a p-value < 0.05. ** indicates a p-value < 0.01.

2.4. Results

2.4.1. Sam68ΔC inhibition is dominant over cytoplasmic stimulatory Sam68 mutants

Previous work by our lab showed that fusion of a heterologous NLS to the N-terminus of Sam68ΔC, relocalized the protein to the nucleus and converted it from an inhibitor to a stimulator of Rev-dependent reporter expression (406). This observation indicated that the subcellular distribution determined protein function. We made further deletions of Sam68ΔC, to see what domains were required for inhibition. Initially, we made a Sam68 deletion composed solely of the GSG-domain of Sam68, Sam68:97-280. A second clone, Sam68:97-255, was generated by PCR error and was also tested. 293T cells were transfected with the Rev-dependent reporter pDM128, which contains the chloramphenicol acetyl transferase (CAT) open reading frame within the HIV-1 4kb intronic sequences (Figure 2.1.a). Cells were also co-transfected with Sam68:97-255 or Sam68:97-280, in the presence or absence of Rev. 48 hours post-transfection the cells were harvested and the lysates were tested for CAT activity. Rev-activity was measured as the amount of CAT activity in the lysates, and normalized to the total amount of protein as measured by Bradford assay. In the absence of Rev, there is a basal amount of CAT expression. Sam68:97-255 and Sam68:97-280 were not inhibitors of Rev-dependent reporter expression. Surprisingly, they stimulated Rev activity to a similar extent as Sam68 (see Figure 2.1.b, p<0.01). We also looked at their localization by immunofluorescence. We found that both Sam68:97-255 and Sam68:97-280 were localized throughout the cell, but predominantly in the cytoplasm (Figure 2.1.c).
Figure 2.1. Sam68:97-255 and Sam68:97-280 are cytoplasmic and stimulate Rev-activity.

a. pDM128, a Rev-responsive reporter construct. CAT open reading frame is contained within the HIV-1 env intron. CAT expression is a measure of Rev activity. b. 293T cells were transfected with pDM128, Rev and Sam68 or mutants thereof. 48 hours post-transfection cells were harvested and analysed for CAT expression. Sam68 and mutants stimulate Rev activity. c. HeLa cells were transfected with Sam68 or mutants thereof and fixed for immunofluorescence 48 hours post-transfection.
Figure 2.2. Sam68:97-255 and Sam68:97-280 stimulation of Rev activity is inhibited by Sam68ΔC.

COS7 cells were transfected with pDM128, Rev, Sam68:97-255 or Sam68:97-280 in the presence or absence of Sam68ΔC. Total amount of transfected DNA was equalized with empty vector. Cells were harvested 48 hours post-transfection, and analysed for CAT expression as a measure of Rev-activity. Overexpression of the GSG-domain of Sam68, Sam68:97-255 and Sam68:97-280, stimulates the Rev-activity. Co-transfection of Sam68ΔC and Sam68:97-255 or Sam68:97-280 inhibits Rev activity.
Previous results by our lab and others showed that Sam68ΔC is a dominant inhibitor of Sam68 Rev-stimulation: when both proteins are transfected into cells, Rev activity is inhibited (403, 406). We postulated that as Sam68ΔC bundles the Rev-dependent mRNAs on the nuclear periphery, it was acting at a step downstream of Sam68, which is localized to the nucleus. Unlike Sam68, Sam68:97-255 and Sam68:97-280 are localized in the cytoplasm and therefore may interfere with Sam68ΔC inhibition of Rev activity. We wanted to determine whether or not Sam68ΔC was a dominant inhibitor of Sam68:97-255 and Sam68:97-280 stimulatory activity. COS7 cells were used for these experiments, as the amount of plasmid DNA transfected was toxic to 293T cells. COS7 cells were transfected with the same Rev-dependent reporter used previously, pDM128. Cells were co-transfected with Sam68:97-255 or Sam68:97-280, in the presence or absence of Rev and Sam68ΔC. 48 hours post-transfection the cells were harvested and the lysates tested for CAT activity. Rev-activity was measured as the amount of CAT activity in the lysates, and normalized to the total amount of protein as measured by Bradford assay. In COS7 cells, there was no CAT expression in the absence of Rev. Rev-activity was stimulated by Sam68:97-255 and Sam68:97-280, and Sam68ΔC was a dominant inhibitor of this activity (Figure 2.2, p<0.01).

**2.4.2. Sam68ΔC co-localizes with the stimulatory mutants and unspliced HIV-1 mRNAs to perinuclear bundles**

Next we investigated the mechanism of Sam68ΔC dominance over the cytoplasmic, stimulatory Sam68 mutants. We wanted to look at the effects of Sam68ΔC on the localization of the unspliced HIV-1 mRNA in the presence of the stimulatory mutants. We transfected HeLa cells with Rev and pgTat, a Rev-dependent reporter
construct encompassing the HIV-1 \textit{env} gene. Cells were co-transfected with HA-tagged Sam68\textDelta C and either myc-tagged Sam68\textDelta C, Sam68, Sam68:97-255 or Sam68:97-280. 48 hours post-transfection cells were fixed with paraformaldehyde for immunofluorescence or \textit{in situ} hybridization.

Figure 2.3 shows the immunofluorescence results with anti-HA and anti-myc antibodies. The top three panels confirm that both myc-Sam68\textDelta C and HA-Sam68\textDelta C co-localize in perinuclear bundles. The second set of panels show that myc-Sam68 remains nuclear when HA-Sam68\textDelta C bundles are formed. However, comparison of the localization of Sam68 in the presence of HIV-1/Sam68\textDelta C bundles with its localization in the absence of Sam68\textDelta C reveals a difference. In the absence of HIV-1/Sam68\textDelta C perinuclear bundles, Sam68 displays a disperse nuclear, non-nucleolar stain (Figure 2.1.c). In the presence of HIV-1/Sam68\textDelta C perinuclear bundles Sam68 appears more concentrated at the inner nuclear membrane (Figure 2.3). Sam68:97-255 and Sam68:97-280 both co-localize to the HIV-1/Sam68\textDelta C perinuclear bundles (Figure 2.3).

To confirm that the perinuclear bundles shown by anti-HA immunofluorescence in Figure 2.3 do indeed contain unspliced HIV-1 mRNA, \textit{in situ} hybridizations were performed. Cells were transfected with Rev, pgTat, HA-Sam68\textDelta C, and either myc-tagged Sam68\textDelta C, Sam68, Sam68:97-255 or Sam68:97-280. \textit{In situ} hybridization was performed using a DIG-labelled probe complementary to the intron of pgTat. In addition, we performed immunofluorescence using anti-myc antibodies. The \textit{in situ} results shown in Figure 2.4 recapitulate the results shown in Figure 2.3. Myc-tagged Sam68\textDelta C, Sam68:97-255 and Sam68:97-280 all co-localize in perinuclear bundles with the
Figure 2.3. Sam68:97-255 and Sam68:97-280 colocalize with Sam68ΔC in the presence of unspliced HIV mRNA.

HeLa cells were transfected with pgTat, Rev, HA-Sam68ΔC, and myc-tagged Sam68 mutants. 48 hours post-transfection, the cells were fixed for immunofluorescence, stained with anti-myc antibodies (green) and anti-HA antibodies (red). Nuclei were stained with DAPI (blue). Cells which showed perinuclear bundling of HA-Sam68ΔC were scored as positive for pgTat, Rev and HA-Sam68ΔC. HA-Sam68ΔC co-localizes with the cytoplasmic, stimulatory myc-tagged mutants of Sam68 in the presence of pgTat and Rev.
Figure 2.4. Sam68:97-255 and Sam68:97-280 colocalize with unspliced HIV mRNA in the presence of Sam68ΔC.

HeLa cells were transfected with pgTat, Rev, HA-Sam68ΔC, and myc-tagged Sam68 mutants. 48 hours post-transfection, the cells were fixed for RNA in situ hybridization (green) and co-stained with anti-myc antibodies (red). Nuclei were stained with DAPI (blue). Cells which showed perinuclear bundling of the unspliced pgTat RNA were scored as positive for pgTat, Rev and HA-Sam68ΔC. The cytoplasmic, stimulatory myc-tagged mutants of Sam68 colocalize in perinuclear bundles with the unspliced HIV mRNA in the presence of HA-Sam68ΔC.
unspliced HIV-1 mRNA. We also see that myc-Sam68 is localized to the inner nuclear membrane, and is not diffuse throughout the nucleus, in the presence of HIV-1/Sam68ΔC perinuclear bundles (Figure 2.4).

2.4.3. Sam68Δ14Δ(45-54)−300 is the minimal inhibitory Sam68 mutant

To map the regions of Sam68ΔC required for inhibition, we made a series of N- and C-terminal deletion mutants (summarized in Figure 2.5). Initially, we tested these deletion mutants with the Rev-responsive reporter pDM128 (summarized in Figure 2.6). The results show that deletions at the N-terminus are not well tolerated. Sam68Δ14ΔC is an inhibitor of Rev activity, as measured by CAT expression. Rev activity in the presence of Sam68Δ14ΔC is significantly less than in the presence of the empty vector pcDNA (p<0.01). CAT expression in the presence of Sam68Δ14ΔC is not significantly different from the basal level of CAT expression in the absence of Rev, or in the presence of Rev and Sam68ΔC (p<0.01). Overexpression of Sam68ΔC mutants with further N-terminal deletions (ie. Sam68Δ28ΔC) stimulated Rev-activity. CAT expression is significantly stimulated by overexpression of these N-terminal deletion mutants compared to Rev alone (p<0.01), and is not significantly different from stimulation by Sam68:97-255 (p<0.01). Deletions at the C-terminus are better tolerated (Figure 2.6). Sam68:5-314 retains inhibitory activity similar to that seen with Sam68ΔC (p<0.01). Although overexpression of Sam68:5-300 is also inhibitory to CAT expression (p<0.05), CAT expression is stimulated significantly over levels seen in the absence of Rev or in the presence of Rev and Sam68ΔC (p<0.01). Sam68:5-262 is a stimulator of Rev activity (p<0.01).
Figure 2.5. Schematic of Sam68ΔC mutants tested. Sam68 contains a GSG domain, six proline-rich motifs (P0-P5) and two RGG-boxes (hatched boxes). The GSG domain contains an embedded hnRNP K (KH) RNA binding domain. The entire GSG domain is required for RNA-binding activity of Sam68.
Figure 2.6. Analysis of Sam68ΔC deletion mutants and their effects on Rev-activity. 293T cells were transfected with pDM128, Rev and Sam68ΔC or mutants thereof. 48 hours post-transfection cells were harvested and analysed for CAT expression as a measure of Rev-activity. Graph summarizing the effects of Sam68ΔC, and deletion mutants thereof, on Rev-activity using pDM128 reporter construct.
The two RGG-boxes of Sam68 have been implicated in non-specific, homopolymeric RNA-binding (389). The C-terminal RGG-box is not present in Sam68:5-300 and only a portion is retained in Sam68:5-314, however the N-terminal RGG-box is retained in all of the Sam68ΔC mutants which have retained inhibitory activity. We wanted to test whether the N-terminal RGG-box was required for the inhibitory activity of Sam68ΔC (Figure 2.6). Sam68Δ(45-54)ΔC contains a deletion of the N-terminal RGG-box, and significantly inhibits Rev-activity to the same level as Sam68ΔC (p<0.01). This RGG-box deletion was combined with inhibitory N- and C-terminal deletions. Sam68Δ14Δ(45-54)ΔC and Sam68Δ14Δ(45-54)-314 are both significant inhibitors of Rev activity, comparable to Sam68ΔC itself (p<0.01). As with Sam68:5-300, Sam68Δ14Δ(45-54)-300 consistently inhibits Rev activity (p<0.05). However, CAT expression is still above the levels seen in the absence of Rev or in the presence of Rev and Sam68ΔC (p<0.01). Further analysis is required to decipher the ambiguous activities of Sam68:5-300 and Sam68Δ14Δ(45-54)-300.

To better assess the activity of these Sam68ΔC deletion mutants, we repeated the assay using a more biologically relevant reporter. GagRRE encodes the entire HIV-1 gag-pol gene as well as the HIV-1 RRE (Figure 2.7.a). It expresses p55 which is processed to p24 by the viral protease. Figure 2.7 summarizes the effects of Sam68ΔC, and mutants thereof, on expression from the GagRRE reporter construct. Western blots were initially probed with anti-p24 antibodies, which detect both p24 and the precursor p55. Blots were stripped and re-probed with anti-myc antibodies to confirm expression of each of the mutants, then stripped again and re-probed with anti-tubulin antibodies to confirm equal loading between wells. It should be noted that stripping of the western blots preferentially
stripped smaller proteins from the membrane. As a result, the expression of lower molecular weight proteins such as Sam68:97-255 appears to be less than the expression of higher molecular weight proteins such as Sam68ΔC. However, when anti-myc blots were performed first, expression of all of the myc-tagged mutants appears approximately equal (data not shown). Therefore differences between Sam68ΔC, or mutants thereof, on GagRRE expression can not be attributed to dosage effects.

Basal p55 expression is seen in the absence of Rev, but expression is stimulated by Rev and inhibited below basal levels by Sam68ΔC (Figure 2.7.b-d). p24 is only detected in the presence of Rev, and Sam68ΔC completely inhibits this Rev-dependent expression (Figure 2.7.b-d). Figure 2.7.b, summarizes the effects of the C- and N-terminal truncation mutants of Sam68ΔC. Complementary to the pDM128 results summarized in Figure 2.6, deletions at the N-terminus are not well tolerated. While Sam68Δ14ΔC retains inhibitory activity, Sam68Δ28ΔC has lost inhibitory activity, and Sam68Δ50ΔC, Sam68Δ75ΔC, and Sam68Δ96ΔC stimulate Rev-function. Deletions at the C-terminus are better tolerated, Sam68:5-314 and Sam68:5-300 both retain inhibitory activity, while Sam68:5-262 is a stimulator of Rev-activity.

The internal deletion mutant Δ(45-54) is inhibitory in the context of Sam68ΔC, and when combined with the other inhibitory Sam68ΔC mutants (Figure 2.8.c). Combination of internal, C- and N-terminal deletion shows that Sam68Δ14Δ(45-54)-300 is the minimal Sam68ΔC mutant which retains inhibitory activity in the context of GagRRE (Figure 2.8.c). As GagRRE is a more biologically relevant reporter than pDM128, in future experiments Sam68Δ14Δ(45-54)-300 will be used as a positive control to assess Sam68ΔC properties required for Rev inhibition.
Sam68Δ(15-24)ΔC was made to determine if this region is required for Rev activity. The internal mutant Sam68Δ(15-24)ΔC shows no effect on Rev-activity, however the myc-blot reveals that this mutant is not expressed (Figure 2.8.c). This internal deletion most likely causes misfolding and subsequent destabilization of the protein. To determine whether specific residues in this region confer inhibitory activity on Sam68ΔC, we mutated candidate residues to alanine (Figure 2.8.d). Unlike Sam68Δ(15-24)ΔC, all of these point mutants were expressed to a similar level as Sam68ΔC. None of the mutated residues resulted in a complete loss of inhibitory activity. In the case of Sam68ΔC:K21A, some p24 is expressed, however this level of p24 expression is comparable to that seen in the absence of Rev. Therefore, the inhibitory activity of Sam68ΔC can not be attributed to any one of the mutated residues.
Figure 2.7. The effect of Sam68ΔC and mutants thereof on HIV-1 Gag expression.
a. GagRRE reporter construct. GagRRE is a Rev-dependent reporter which expresses HIV-1 Gag/pol (p24). b-d. 293T cells were transfected with GagRRE, Rev and Sam68ΔC or mutants thereof. 48 hours post-transfection cells were harvested and effects on p24 expression was measured by western blot. N- and C-terminal deletion mutants are summarized in b, internal deletion mutants and combinatorial mutants are summarized in c, and point mutants in the region 15-25 are summarized in d.
2.4.4. Inhibitory and Stimulatory Sam68ΔC mutants retain perinuclear bundling of unspliced HIV-1 RNA

Through our deletion analysis, we defined the minimal inhibitory mutant Sam68Δ14Δ(45-54)-300, and the maximal stimulatory mutants Sam68Δ28ΔC and Sam68:5-262. To assess the ability of these mutants to bundle unspliced HIV-1 RNA, HeLa cells were transfected with pgTat, Rev and Sam68ΔC or deletion mutants thereof. The cells were fixed for in situ hybridization, and probed with a DIG-labelled probe complementary to the intron of pgTat, and anti-myc antibodies. When transfected with an empty vector, the unspliced pgTat mRNA is found in both the nucleus and the cytoplasm (Figure 2.8, top 3 panels). In the presence of Sam68ΔC, the unspliced HIV-1 RNA is localized to the nucleus and co-localizes with Sam68ΔC in perinuclear bundles. Surprisingly, all of the Sam68ΔC deletion mutants tested were able to bundle the unspliced HIV-1 RNA in the perinuclear region, irrespective of their effect on its subsequent translation. This observation indicates that perinuclear bundling of the HIV-1 RNA is not correlated with inhibition of viral protein synthesis.
**Figure 2.8.** Non-inhibitory mutants of Sam68ΔC are able to bundle unspliced HIV mRNA.

HeLa cells were transfected with pgTat, Rev and pcDNA, Sam68ΔC or mutants thereof. The localization of unspliced HIV-1 mRNA (green) and myc-tagged Sam68ΔC, and deletion mutants thereof (red), are shown by *in situ* hybridization and immunofluorescence. The nuclei were stained with DAPI (blue).
2.4.5. Sam68ΔC inhibits Rev-, but not Tap/p15-, transported HIV-1 mRNAs

It was important to confirm that Sam68ΔC inhibition of Rev-dependent reporter constructs is specific, and not a general effect on RNA transport and/or translation. Therefore, we used a Rev-independent reporter construct GagCTE, which encodes the HIV-1 gag-pol gene but replaces the HIV-1 RRE with the constitutive transport element (CTE) from Mason-Pfizer Monkey virus (Figure 2.9.a). The CTE interacts directly with Tap/p15, and is exported in a Rev/Crm1 independent manner. Figure 2.9.b-d summarizes the effects of Sam68ΔC, and mutants thereof, on expression from the GagCTE reporter construct. As with the analysis of GagRRE, the membranes were initially probed with anti-p24 antibodies, re-probed with anti-myc antibodies to confirm expression of each of the mutants, and then re-probed with anti-tubulin antibodies to confirm equal loading between wells. There is a very low basal expression of p24 from GagCTE in the presence of empty vector, pcDNA (Figure 2.9.b-d). In contrast to its effects on GagRRE, Sam68ΔC stimulates p24 expression from GagCTE. All of the Sam68ΔC deletion mutants had a stimulatory effect on this reporter construct, irrespective of their effects on GagRRE. These results show that Sam68ΔC and Sam68Δ14Δ(45-54)-300 inhibition of Rev-activity is specific to an element of the RRE/Rev/Crm1 export complex.

To determine if the GagCTE RNA was still bundled by Sam68ΔC, HeLa cells were transfected with GagCTE and empty vector or Sam68ΔC. 48 hours post-transfection the cells were analyzed for the localization of the RNA and myc-tagged Sam68ΔC (Figure 2.10). The results show that in the absence of Sam68ΔC the Gag-CTE RNA is localized throughout the cytoplasm. In the presence of Sam68ΔC Gag-CTE RNA is
localized to the nucleus and perinuclear bundles. Sam68ΔC co-localizes to the perinuclear bundles with the unspliced GagCTE RNA. Although GagCTE is not inhibited by Sam68ΔC it is still bundled in the perinuclear region. This observation confirms that perinuclear bundling does not correlate with inhibition of protein expression.
Figure 2.9. The effect of Sam68ΔC and mutants thereof on HIV-1 Gag expression from a Rev-independent reporter.

a. GagCTE reporter construct. GagCTE is a Rev-independent reporter which expresses HIV-1 Gag/pol (p24). b-d. 293T cells were transfected with GagRRE, Rev and Sam68ΔC or mutants thereof. 48 hours post-transfection cells were harvested and effects on p24 expression was measured by western blot. N- and C-terminal deletion mutants are summarized in b, internal deletion mutants and combinatorial mutants are summarized in c, and point mutants in the region 15-25 are summarized in d.
**Figure 2.10.** Sam68ΔC is able to bundle unspliced HIV mRNA which is not Rev-transported.
HeLa cells were transfected with GagCTE with or without myc-Sam68ΔC. Cells were fixed for *in situ* hybridization/immunofluorescence, and probed with antisense GagCTE RNA (green) and anti-myc antibodies (red). The nuclei were stained with DAPI (blue).
2.5. Discussion

Sam68ΔC was identified as a potent inhibitor of HIV-1 Rev activity (403). Subsequent work from our lab revealed that Sam68ΔC bundled Rev-exported HIV-1 RNA in the perinuclear region, and we postulated that this inhibited these RNAs from accessing the translational machinery of the cell (406). We also showed that inhibition of Rev-activity was dependent upon the cytoplasmic localization of Sam68ΔC; relocalization to the nucleus by fusion of a heterologous NLS converted Sam68ΔC to a stimulator of Rev-activity (406). In this study we have further explored the mechanism of inhibition and its relationship with the subcellular localization of Rev, Sam68ΔC and the unspliced HIV-1 RNAs.

Deletion mutants of Sam68ΔC comprised of the GSG-domain alone display a whole cell/cytoplasmic localization but are stimulators of Rev-activity (Figure 2.1). This indicates that in addition to its cytoplasmic localization, Sam68ΔC requires domains beyond the GSG-domain for inhibition. The GSG-domain of Sam68 is responsible for RNA binding and multimerization (389, 390). Therefore, it is possible that these stimulatory, cytoplasmic mutants would be able to form heteromultimers with Sam68ΔC and inhibit its activity, or compete for binding to susceptible mRNAs. However, our results reveal that Sam68ΔC inhibition is dominant over stimulation by Sam68:97-255 and Sam68:97-280 (Figure 2.2). Immunofluorescence and in situ hybridization experiments reveal that in the presence of pgTat and Rev, Sam68ΔC relocalized Sam68:97-255 and Sam68:97-280 in perinuclear bundles (Figure 2.3, 2.4). Under the same conditions, Sam68 remains nuclear but its localization is subtly altered. Sam68 has
a disperse nuclear, non-nucleolar distribution (Figure 2.1), but in the presence of pgTat and Sam68ΔC perinuclear bundles, Sam68 is more concentrated at the inner nuclear membrane (Figure 2.3, 2.4). This contrasts to results seen by Zhang et al, who show that Sam68 partially relocates to the cytoplasm in the presence of Sam68ΔC (412). These discrepancies may be due to the fact that Zhang et al were using 293T cells for immunofluorescence or that they were using GFP-Sam68 (412). The large GFP moiety at the N-terminus may alter how Sam68 and Sam68ΔC interact. In our lab we have observed that fusion of a large moiety, such as a tandem affinity purification tag, at the N-terminus of Sam68ΔC abrogates its function (data not shown).

We have also defined the minimal mutant of Sam68ΔC required for inhibition, Sam68Δ14Δ(45-54)-300 (Figure 2.6, 2.7). Sam68Δ14Δ(45-54)-300 lacks both RGG boxes, but retained significant inhibitory activity (Figure 2.6, 2.7). Recent analyses in another lab determined that the region spanning amino acids 269-321 are essential for the inhibitory property of Sam68ΔC (412). Our results agree with this result, Sam68:5-262 is no longer inhibitory. However, we have found that Sam68:5-300 was inhibitory. This indicates that the region required for inhibition can be further narrowed down to the amino acids 263-299. In addition, we found that at the N-terminus amino acids 15-28 were also required for inhibition: Sam68Δ28ΔC lost inhibitory activity while Sam68Δ14ΔC is still inhibitory (Figure 2.6, 2.7).

We have also found that the effects of Sam68ΔC are specific to Rev-dependent RNAs. The GagCTE reporter is not inhibited by Sam68ΔC, or mutants thereof (Figure 2.9). In fact several of the Sam68ΔC mutants appear to stimulate expression from GagCTE. In addition, when we looked at the localization of the GagCTE RNA in the
presence of Sam68ΔC, we found that the RNA and Sam68ΔC were localized in
perinuclear bundles (Figure 2.10). This was surprising because translation of this RNA is
not inhibited. However it confirms the results seen in Figure 2.8, where the non-
inhibitory Sam68ΔC mutants, Sam68Δ28ΔC and Sam68:5-262, were able to bundle the
GagRRE RNA. Although we do not see an instance where an inhibited RNA is not
bundled in the perinuclear region, we see several examples where the expression of a
bundled RNA is not inhibited. Therefore, there must be some additional means through
which Sam68ΔC and Sam68Δ14Δ(45-54)-300 are able to inhibit the expression from
these RNAs. Perinuclear bundling alone is not sufficient for inhibition, and the
translational machinery must be able to access these RNAs in certain circumstances.

Recent studies have highlighted the changes in RNP composition that occur
during mRNA processing, export and surveillance: as the mRNA moves from its site of
synthesis through the nuclear pore to the translational apparatus. In particular, a number
of hnRNP proteins are removed, the exon junction complex (EJC) is added following
splicing then subsequently removed during translation, and there is an exchange of
factors at the extreme 5’ and 3’ ends of the RNA (CBP20/80 and PABP2 are exchanged
for eIF4E and PABP1, respectively) (13, 416, 417). What remains unknown is the extent
to which the protein composition of a specific RNP is determined by its sequence or its
processing pathway. In the case of the EJC, the act of splicing deposits the complex ~20-
25 nt 5’ of the splice site in a sequence independent fashion (13). In other instances,
binding of factors to specific sequence elements (Zipcode localization elements, AU-rich
elements) affects localization, stability or translation of an RNA (417). It has generally
been assumed that, despite an alternative export pathway, remodeling of incompletely
spliced HIV-1 mRNPs would occur in the same manner as host mRNPs. However, in this study we provide evidence that Sam68ΔC is able to discriminate between RNAs based on the export pathway used.
Chapter 3: Sam68ΔC Inhibits Translation of Incompletely Spliced HIV-1 mRNAs by Blocking PABP1 Association

Kim Marsh and Alan Cochrane

All of the experiments in this chapter were performed by Kim Marsh.

Figures 3.2, 3.4-3.9 and 3.12 have been submitted to Retrovirology as part of “Selective Translational Repression of HIV-1 RNA by Sam68ΔC Occurs by Altering PABP1 Association.” Marsh, K., Soros, V., and Cochrane, A. (2008).
3.1. Abstract

HIV-1 structural proteins are translated from incompletely spliced 9 kb and 4 kb mRNAs, which are transported to the cytoplasm by Crm1. It has been assumed that once in the cytoplasm, translation of incompletely spliced HIV-1 mRNAs occurs in the same manner as host mRNAs. While investigating the translational inhibition of incompletely spliced HIV-1 mRNAs by Sam68ΔC, we determined that the effect was independent of the perinuclear bundling of the viral mRNA. We demonstrate that inhibition by Sam68ΔC is correlated with a loss of PABP1 binding with no attendant change in abundance, polyadenylation or polyadenosine tail length of the affected mRNAs. The capacity of Sam68ΔC to selectively inhibit HIV-1 RNAs exported by Crm1 suggests that it is able to recognize unique characteristics of these viral mRNPs, a property that could lead to new therapeutic approaches to controlling HIV-1 replication.
3.2. Introduction

Expression of the full coding potential of the HIV-1 genome is dependent upon a number of post-transcriptional processes. The primary 9kb transcript from the integrated provirus can be spliced into over 40 mRNAs through suboptimal splicing events (268, 269, 378, 418). Resulting HIV-1 mRNAs can be grouped into three classes: the unspliced, 9 kb class, encoding Gag and GagPol; the singly spliced, 4 kb class, encoding Vif, Vpr, Vpu and Env; and the multiply spliced, 2 kb class, encoding Tat, Rev and Nef. Incompletely spliced mRNAs are normally retained in the nucleus but the virus has evolved a mechanism for the transport of the 9 kb and 4 kb viral mRNAs to the cytoplasm. The Rev protein is translated in the cytoplasm, then shuttles back into the nucleus where it multimerizes on the 

HIV-1 gene expression may be controlled at several steps including transcription, splicing, polyadenylation, nuclear export and translation (378, 418, 419). All of these processes depend upon host cell factors (420). Recent work in our laboratory has focused on Sam68, a member of the STAR/GSG family of proteins (386). These proteins share an RNA binding motif, the KH domain, embedded within a larger conserved GSG domain, which mediates multimerization. Sam68 is a nuclear, non-shuttling protein, and contains both proline- and tyrosine-rich domains mediating binding to multiple kinases (i.e. Src, Lck, Sik/BRK, ZAP-70) through SH3 and SH2 domains, respectively (386, 421). Given its interaction with kinases involved in signal transduction, Sam68 has been
suggested to serve as a signal mediator, having effects on multiple cellular processes including cell cycle regulation, tumour suppression, alternative splicing, and RNA 3’ end formation (290, 386, 421-427). More pertinent to HIV-1, overexpression of Sam68 and other members of the GSG family have been shown to significantly enhance HIV-1 gene expression (403, 406, 428, 429). Furthermore, two groups have reported that depletion of Sam68 results in the loss of HIV-1 structural protein expression in several cell lines (409, 411).

In contrast to the full length protein, a truncation mutant of Sam68 lacking the C-terminal 112 amino acids, Sam68ΔC, is a potent inhibitor of HIV-1 protein expression (403, 406). Unlike Sam68, Sam68ΔC is localized predominantly in the cytoplasm and its inhibitory function requires this distribution (406). Previous experiments in our lab showed that Sam68ΔC induced accumulation of HIV-1 4 kb mRNAs into perinuclear bundles suggesting that it might act by sequestering the viral mRNA from the translational apparatus (406). However, subsequent studies show that perinuclear bundling of an mRNA does not correlate with inhibition of mRNA expression (Chapter 2). In this study we set out to define the mechanism of Sam68ΔC inhibition. We show that depolymerization of microfilaments disrupted the perinuclear bundles, dispersing the viral mRNA throughout the cytoplasm but failed to restore the synthesis of the HIV-1 structural proteins (Gag, Env). This finding demonstrates that the block to expression is a block to translation, not transport. Our studies determined that Sam68ΔC has no effect on viral mRNA abundance, polyadenylation or poly(A) tail length. Inhibition of translation by Sam68ΔC was not associated with any changes in viral mRNA localization, abundance, or processing but is correlated with changes in the composition of the mRNP.
We show that Sam68ΔC inhibits translation of these HIV-1 mRNAs by blocking association of PABP1.

3.3. Materials and Methods

3.3.1. Expression constructs.

The following constructs have been previously described: SV-Hygro, SV-H6Rev, CMVmyc3xterm (289), pDM128 (413), pgTat (297), Bl-env-HindIII (305), Sam68 and Sam68ΔC (406). Cloning of Sam68Δ14Δ(45-54)-300, Sam68Δ28ΔC, Sam68:5-262 were described in Chapter 2.3.1. Bl-actin was generated by PCR from cDNA using primers: 5’-GCT ACG AGC TGC CTG AC-3’ and 5’-TCC TTC TGC ATC CTG TC-3’. The amplicon was cloned into the EcoRV site of Bluescript. Bl-SD-Gag was amplified from HxBruR’/RI using: 5’-CGG GGA TCC GAA GTA GTG TGT GCC CGT CTR3’ and 5’-CCC AAG CTT CCC TGC TTG CCC ATA CTA TAR3’. The amplicon was digested with BamHI and HindIII and ligated into Bluescript. pgTat-hist-3’UTR was made by PCR amplification of the histone H3 3’UTR from cDNA using the primers: 5’histone: 5’-GCT CTA GA G TCC CTG CCG GGA CCT GGC GCR3’, and 3’histone: 5’-CCG GAA TTC GAA TAA GTG ACC AAG CTC TT3’. The amplicon was digested with Xbal/EcoRI and ligated into Xbal/EcoRI digested pgTat.

3.3.2. Cell lines and transfections.

HeLa and 293T cells were maintained in Iscove’s modified Dulbecco’s medium (IMDM) supplemented with 10% fetal bovine serum (FBS) (v/v), 50µg/mL gentamycin sulfate and 2.5µg/mL amphotericin B. Vectors were introduced to 293T cells by calcium phosphate transfection (414). Vectors were introduced to HeLa cells by Fugene 6 transfection reagent (Roche) following manufacturer’s protocol. Cells were harvested two
days post-transfection (except pgTat-hist-3'UTR transfections which were harvested three days post-transfection).

3.3.3. Antibodies.

The following antibodies were used: mouse anti-myc (Invitrogen), mouse anti-tubulin (Sigma), mouse anti-gp120AK (courtesy of H. Schaal), rabbit anti-Rev (2/99) and rabbit anti-human PABP1 (aa 462-633) (courtesy of N. Sonenburg, McGill University). The following secondary antibodies were used: donkey anti-rabbit IgG conjugated HRP, donkey anti-mouse IgG conjugated HRP, donkey anti-rabbit IgG conjugated Texas red, donkey anti-mouse IgG conjugated Texas red (Jackson Immuno Research), protein G conjugated HRP (Molecular Probes), and sheep anti-digoxigenin conjugated FITC (Boehringer Mannheim).

3.3.4. In Situ Hybridization and Immunofluorescence.

3x10^5 HeLa cells were transfected with 0.25 µg SV H6Rev, 1.25 µg pgTat and 5.0 µg pcDNA3.1 or Sam68ΔC. 48 hours post-transfection, cells were treated with fresh media containing no drug, colcemid (0.1µg/mL), nocodazole (1.0µg/mL), cytochalasin D (0.5µg/mL), or latrunculin B (0.5µg/mL) for 2 hours prior to fixation. In situ hybridization was performed as previously described (305). Digoxigenin-labeled Env-HindIII probe, antisense to HIV-1 env mRNA, was made with 10 x digoxigenin RNA labeling kit (Roche). Myc-tagged proteins were detected with monoclonal anti-myc antibody (Invitrogen). Immunofluorescence was detected using a Leica DMR microscope at either 400x or 630x magnification.

3.3.5. Digitonin Permeabilization
6x10^5 HeLa cells were transfected with 0.25 µg SV-H6Rev, 1.25 µg pgTat and 5.0 µg Sam68ΔC. 48 hours post-transfection, cells were treated with drugs for 2 hours, as outlined above. Cells were rinsed quickly with cytoskeletal buffer (CSK, 10mM PIPES (pH 6.9), 0.1M NaCl, 3mM MgCl₂, 0.3M sucrose) two times. To each well 2mL of 50 µg/mL digitonin (in CSK) was added and incubated on ice for 5 minutes with shaking. Cells were rinsed with CSK twice. 1mL of CSK was added and cells scraped into 1.8mL microfuge tube. Cells were pelleted by centrifugation 6,000 rpm, 5 minutes, 4ºC. Supernatant was aspirated and the cell pellet resuspended in 30 µL whole cell lysis buffer (150mM NaCl, 10mM Na₂HPO₄, 1% Triton X-100 v/v, 0.1% SDS w/v, 0.2% sodium azide w/v, 0.5% sodium deoxycholate w/v, 1mM sodium orthovanadate). 1 µL cell lysate was used for a Bradford Assay (BioRad). 25 µg of lysate was loaded per well, and fractionated on a 10% SDS-PAGE gel (v/v), and transferred to PVDF membrane (Pall Life Sciences). Membranes were blocked in 5% milk (w/v) / 1xPBS, incubated with anti-myc antibody (1/2000), and anti-mouse-HP (1/5000) and signals detected using Western Lightning (Perkin Elmer).

3.3.6. ³⁵S labeling.

6x10^5 HeLa cells were transfected with 0.2 µg SV-H6Rev, 1.0 µg Gag-RRE and 4.0 µg pcDNA3.1 or Sam68ΔC. 48 hours post-transfection, cells were treated with drugs for 2 hours, as outlined above. Cells were then labeled with 100 µCi of ³⁵S- methionine for 4 hours. Cells were harvested in whole cell lysis buffer then diluted with 3 volumes IPP150 (150mM KCl, 10mM Tris-HCl pH 7.5, 0.1% NP40 v/v, 0.1% sodium azide w/v). The precleared lysates were incubated with anti-p24 antibody for 1 hour at 4ºC, then immunoprecipitated with Gammabind plus sepharose (GE Healthcare) for 1 hour. Beads
were washed with IPP150 buffer and the immunoprecipitated protein was run out on 10% SDS-PAGE (v/v). Labeled proteins were detected following exposure to a Phosphor Imager screen.

3.3.7. RNA Analysis.

3x10^6 293T cells were transfected with 0.4 µg SV-Hygro or SV-H6Rev, 2.0 µg pgTat and 8.0 µg of pcDNA3.1, Sam68AC or mutants thereof. 48 hours post-transfection 25% of the cells were harvested in whole cell lysis buffer for protein analysis, and 75% of the cells were harvested for RNA isolation (430). RNase protection assays were performed as previously described (290). To monitor the polyadenylation status of the pgTat mRNA 10 µg total RNA was selected using oligo(dT)25 beads according to manufacturers directions (Dynal Biotech). The selected RNA was then input into the RNase protection assay using Bl-Tat X/X probe (290). RACE-PAT (random amplification of cDNA ends-polyadenylation test) cDNA was synthesized using an anchor primer (5’-CTC GCC GGA CAC GCT GAA CTT TTT TTT TTT TTT TTT TTT-3’) with MMLV-RT (Invitrogen). The cDNA generated was then used to generate both spliced and unspliced amplicons using the forward spliced (5’-AGC GGA GAC AGC GAC GAA GAG-3’) or unspliced (5’-CGA CCT GGA TGG AGT GGG ACA-3’) and the reverse primer (5’-CTC GCC GGA CAC GCT GAA C-3’). Amplification used the following cycle parameters: 94°C, 1 minute; 66°C, 1 minute; 72°C, 2 minute; for 30 cycles. Amplicons were fractionated on native PAGE gels and detected by exposure to Phosphor Imager screen. For proviral RNA analysis, 3x10^6 293T cells were transfected with 2.0 µg HxBruR/RI and 8.0 µg of Sam68AC or mutants thereof. The probes used for RNase protection assay were Bl-actin and Bl-SD-Gag.
3.3.8. RNP Immunoprecipitation.

3x10^6 293T cells were transfected with 0.4 µg SVH6Rev, 2.0 µg pgTat and 8.0 µg of pcDNA3.1, Sam68ΔC or mutants thereof. Whole cell lysates were generated by a modification of the protocol of Siomi et al (431). 48 hours post-transfection, cells were washed twice with 1xPBS and once with Buffer A (110mM potassium acetate, 2mM magnesium acetate, 2mM DTT, 10mM HEPES pH 7.5). The cell pellet was resuspended in 400 µl Buffer B (10mM potassium acetate, 2mM magnesium acetate, 2mM DTT, 5mM HEPES pH 7.5, 20µM cytochalasin D) and incubated on ice for 10 min. Cells were disrupted by passage through 25 gauge needle, 5 times. KCl was added to 100mM final concentration. Cells debris was pelleted by centrifugation at 1,500 rpm. Supernatant was adjusted to 0.5% NP40 (v/v) and incubated on ice for 20 min. Insoluble material was pelleted by centrifugation at 14,000 rpm and supernatant retained for use in subsequent immunoprecipitations. 1/42 of the supernatant was retained for analysis of the RNA input, and 1/42 was retained for analysis of protein input. For each immunoprecipitation, 10/42 of the soluble fraction was diluted with 10 volumes RIPA buffer (150mM NaCl, 50mM Tris-HCl pH 7.5, 1% NP40 v/v, 1mM EDTA pH 7.5, 0.5% sodium deoxycholate w/v, 0.05% SDS w/v) and 100u RNaseOUT (Invitrogen). The sample was precleared by incubation with Gammabind plus sepharose (GE Healthcare) for 1 h at 4°C. Supernatant was incubated overnight at 4°C with the indicated antibody. Immune complexes were collected by addition of Gammabind plus sepharose and incubation at 4°C for 1 h. The resin was washed 5x with 1 ml of RIPA buffer and then resuspended in 100 µl of RIPA. RNA was extracted from 150 µl of the resin slurry (430) and 50 µl of slurry was used for protein analysis following addition of 50 µl 2x dissociation buffer.
3.3.9. **Statistical Analysis.**

The data are presented as mean +/- one standard deviation. Data were compared using Student’s t-test. * indicates a p-value < 0.05. ** indicates a p-value < 0.01.

3.4. **Results**

3.4.1. **Perinuclear Bundling of HIV-1 mRNA by Sam68∆C Requires Intact Microfilaments not Microtubules**

Previously, we reported that Sam68∆C expression induced accumulation of both unspliced viral mRNA and Sam68∆C in bundles at the outer periphery of the nucleus (406). The formation of these perinuclear bundles by Sam68∆C might sequester the mRNA from the translational apparatus. If true, disruption of these complexes should restore translation of the viral RNAs. We examined the effect of various agents which disrupt the cytoskeleton on the integrity of the Sam68∆C/viral mRNA perinuclear bundles, and the expression of viral proteins. To minimize secondary drug effects, it was important to determine the minimum amount of each drug required to depolymerize its target in one hour (Figure 3.1). To do this, HeLa cells were incubated with decreasing amounts of each drug for 1 hour, then fixed for immunofluorescence. The top panels of Figure 3.1.a show the intact microfilaments with no drug treatment, and the dissociated microfilaments after treatment with 0.5 µg/mL cytochalasin D or 0.5 µg/mL latrunculin B. The top panels of Figure 3.1.b show the intact microtubules with no drug treatment, and the dissociated microtubules after treatment with 1.0 µg/mL nocodazole or 0.01 µg/mL colcemid. It should be noted that no concentration of colcemid was able to completely dissociate the microtubules in 1 hour, this drug was only effective after 2
hours. Therefore, in all subsequent experiments, the colcemid incubation was started 1 hour prior to the other drug treatments.

Once the minimal concentration of each cytoskeletal depolymerizing drug was determined, we wanted to determine their effects on Sam68ΔC perinuclear bundling. HeLa cells were transfected with Rev, pgTat and Sam68ΔC. Cells were treated with the cytoskeleton depolymerizing drugs and then fixed for in situ hybridization. The cells were probed with a DIG-labelled probe complementary to the intron of pgTat, and antimyc antibodies (Figure 3.2). While treatment of cells with either nocodazole or colcemid had no effect on localization of viral mRNA or Sam68ΔC, disassembly of the microfilaments by treatment with cytochalasin D or latrunculin B resulted in dispersal of both unspliced HIV-1 mRNA and Sam68ΔC throughout the cytoplasm (Figure 3.2). The distribution of the HIV-1 mRNA which has been released from the perinuclear bundles is similar to that seen in the absence of Sam68ΔC (Figure 3.2, top panels).

We wanted to confirm that depolymerization of the microfilaments, and not the microtubules, dissociated the perinuclear bundles. We transfected HeLa cells with Sam68ΔC in the absence or presence of Rev and pgTat. Cells were treated with nocodazole to depolymerize the microtubules or cytochalasin D to depolymerize the microfilaments. After drug treatment, the cell membrane was permeabilized by treatment with digitonin. The cells were washed to remove the soluble fraction, and then harvested. Bradford assays were used to determine the protein concentration of the lysates, and 25 µg of lysate loaded onto a SDS-PAGE gel. The lysates were analysed by anti-myc western blot to determine the retention of Sam68ΔC under the different conditions. The amount of Sam68ΔC retained was quantitated using ImageQuant. Statistical analysis of
Figure 3.1. Determination of the minimal drug concentrations required to depolymerize the cytoskeleton.

HeLa cells were fixed for immunofluorescence after 1 hour treatment with drugs which depolymerize the cytoskeleton. The minimal concentration required to depolymerize the cytoskeleton was determined for each drug. a. The fixed cells were stained for intact microfilaments (red) and nuclei were stained with DAPI (blue). Cells were treated with no drug (top), 0.5 µg/mL cytochalasin D (middle) or 0.5 µg/mL latrunculin B (bottom). b. The fixed cells were stained for intact microtubules (red) and nuclei were stained with DAPI (blue). Cells were treated with no drug (top), 1.0 µg/mL nocodazole (middle) or 0.01 µg/mL colcemid (bottom). The colcemid treatments required 2 hours for complete depolymerization.
<table>
<thead>
<tr>
<th>Drug Treatment</th>
<th>DAPI</th>
<th>unspliced HIV mRNA</th>
<th>Sam68ΔC</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td></td>
<td></td>
<td></td>
<td>n/a</td>
</tr>
<tr>
<td>none</td>
<td></td>
<td></td>
<td></td>
<td>n/a</td>
</tr>
<tr>
<td>nocodazole</td>
<td></td>
<td></td>
<td></td>
<td>microtubules</td>
</tr>
<tr>
<td>colcemid</td>
<td></td>
<td></td>
<td></td>
<td>microtubules</td>
</tr>
<tr>
<td>cytochalasin D</td>
<td></td>
<td></td>
<td></td>
<td>microfilaments</td>
</tr>
<tr>
<td>latrunculin B</td>
<td></td>
<td></td>
<td></td>
<td>microfilaments</td>
</tr>
</tbody>
</table>

**Figure 3.2.** Effect of cytoskeleton depolymerization on Sam68ΔC perinuclear bundling of unspliced HIV-1 mRNA. HeLa cells were transfected with Rev, pgTat and pcDNA (top panel) or Sam68ΔC (except top panel). 48 hours post-transfection, cells were treated with cytoskeleton depolymerizing drugs for 1-2 hours as required for complete depolymerization. Cells were fixed for *in situ* hybridization and immunofluorescence. The fixed cells were probed with an anti-sense unspliced HIV RNA (green), anti-myc antibodies (red) and nuclei were stained with DAPI (blue).
several experiments was not possible as the degree of digitonin permeabilization seemed to vary between experiments, however the trends seen between experiments were the same. A representative result is shown in Figure 3.3.

The amount of Sam68ΔC in untreated, unpermeabilized cells was set to 100% (Figure 3.3, lane 1). In the absence of drug treatment, Sam68ΔC was retained both in the absence (37%) and presence (63%) of Rev and pgTat (lanes 2 and 3). After nocodazole treatment slightly less Sam68ΔC was retained, both in the absence (27%) and presence (49%) of Rev and pgTat (lanes 4 and 5). After cytochalasin D treatment considerably less Sam68ΔC was retained, both in the absence (1%) and presence (8%) of Rev and pgTat (lanes 6 and 7). There are several interesting observations to be made in this experiment. First, in the unpermeabilized cells there are two bands corresponding to Sam68ΔC, but in the permeabilized cells only the larger form is retained (lane 1 versus lanes 2-7). This indicates that only the larger form is insoluble (binds microfilaments). Second, although in the absence of Rev and pgTat we do not see any perinuclear bundling of Sam68ΔC, it must still associate with the insoluble fraction of the cell as it is retained (lanes 2, 4, 6). Third, there is always more Sam68ΔC retained in the presence of Rev and pgTat, when perinuclear bundles are formed. Fourth, although microtubule depolymerization results in a slight decrease in Sam68ΔC retention, microfilament depolymerization results in a considerable decrease in Sam68ΔC retention. These results confirm the in situ hybridization experiments: perinuclear bundling of Sam68ΔC and unspliced HIV-1 mRNA is dependent on intact microfilaments.
Figure 3.3. Effect of cytoskeleton depolymerization on Sam68ΔC solubility.
HeLa cells were transfected with Sam68ΔC (all lanes) and Rev and pgTat (lanes 3, 5, 7). 48 hours post-transfection, cells were treated with no drug (lanes 1-3), nocodazole (lanes 4, 5) or cytochalasin D (lanes 6, 7). Cells were then permeabilized with digitonin (lanes 2-7) and washed to remove the soluble material. The insoluble fraction was harvested, and 25 mg lysate run out on SDS-PAGE. An anti-myc western blot showed the retention of Sam68ΔC under the different condition.
3.4.2. **Sam68ΔC inhibits unspliced HIV-1 mRNA when perinuclear bundles are dissolved**

Next we questioned whether the released viral mRNA was translated. Cells were transfected with Rev and GagRRE in the presence or absence of Sam68ΔC. Forty-eight hours post-transfection, cells were treated with the cytoskeleton depolymerizing drugs as before. Synthesis of viral protein post drug treatment was monitored by incubation in the presence of $^{35}$S-methionine. Cells were lysed, the $^{35}$S-labelled Gag immunoprecipitated, the immunoprecipitates run out on SDS-PAGE gels, and the gels exposed to a phosphorscreen overnight (Figure 3.4). The 55kDa Gag polyprotein (p55) was not detected in the immunoprecipitate from the mock transfected cells, but was only in the immunoprecipitate from the cells transfected with Gag-RRE and Rev (Figure 3.4, lanes 1-5 versus lane 6). Therefore, the immunoprecipitation was specific and the p55 signal could be used as a measure of Gag-RRE translation.

None of the drugs had any effect on the level of Gag expression in the absence of Sam68ΔC indicating that they had no significant effect on translation (Figure 3.4, lanes 1-6). None of the drugs induced expression of Gag in the presence of Sam68ΔC (Figure 3.4, lanes 7-12) despite latrunculin B/cytochalasin D shifting the subcellular distribution of both Sam68ΔC and the viral mRNA. Therefore, integrity of the Sam68ΔC/viral RNA perinuclear bundles is not essential for translational repression. This observation suggests that Sam68ΔC inhibits viral mRNA translation, not by changing its subcellular distribution, but through alterations in viral mRNA structure or composition of the mRNP.
Figure 3.4. Effect of cytoskeleton depolymerization on Sam68ΔC inhibition of unspliced HIV-1 mRNA translation. HeLa cells were mock transfected (lanes 6 and 12), transfected with GagRRE, Rev and pcDNA (lanes 1-5) or Sam68ΔC (lanes 7-11). 48 hours post-transfection, cells were treated with no drug or drugs to depolymerize the cytoskeleton. After 2 hours, the media was spiked with 100μCi $^{35}$S labelled methionine and incubated for 4 hours. The cells were harvested and Gag (p55/p24) was immunoprecipitated from the lysates. The immunoprecipitates were run out on SDS-PAGE, and $^{35}$S-labelled proteins were detected by exposure on a Phosphor Screen.
3.4.3. Sam68ΔC does not affect the amount of HIV-1 unspliced mRNA

Splicing, 3’ end cleavage, and polyadenylation are tightly coupled events within the nucleus. As HIV-1 mRNA is inefficiently spliced, numerous forms of viral mRNA are generated which have failed to complete some or all steps of mRNA processing. This is essential for the production of the incompletely spliced 4 and 9 kb HIV-1 mRNAs which encode the structural and enzymatic viral proteins (268, 269, 378, 418). The HIV-1 env reporter, pgTat, expresses gp160/120 from unspliced, cleaved mRNA ((290), see Figure 3.5.a). Previously, we showed that full-length Sam68 increases the amount of unspliced, cleaved pgTat mRNA available for polyadenylation, export and translation into gp160/120 (290). We have also shown that restoring nuclear accumulation of Sam68ΔC by addition of a heterologous nuclear localization signal (NLS) converted the protein into a stimulator of Rev function comparable to full-length Sam68 (406). In light of these results, we wanted to assess whether Sam68ΔC inhibition is due to alterations in the abundance or processing of env mRNA

Sam68ΔC contains a number of well-defined domains that mediate specific RNA binding (KH-domain), non-specific RNA binding (RGG-boxes) or protein-protein interactions (proline-rich domain) (386). In Chapter 2 we made a number of deletions of Sam68ΔC to define a minimal inhibitory mutant, Sam68Δ14Δ(45-54)-300, to serve as a positive control for any Sam68ΔC effects. For simplicity we will call this mutant Sam68Δ14-300 for the remainder of this chapter. We also defined two non-inhibitory mutants: Sam68Δ28ΔC and Sam68:5-262, to serve as negative controls for any Sam68ΔC effects. First, we confirmed that the effects of these proteins on gp160/120 expression from pgTat agreed with their effects on GagRRE (Figure 3.5.b). Indeed, Sam68ΔC and
Sam68Δ14-300 completely inhibit expression of gp160/120, while Sam68Δ28ΔC and Sam68:5-262 do not.

Next we examined the extent of splicing and cleavage of pgTat mRNA by RNase protection assay (RPA). The RPA probe used in this analysis spans both the 3’ splice site and the polyadenylation site and thus yields four RNase protection products: unspliced, uncleaved (US-UC); unspliced, cleaved (US-C); spliced, uncleaved (S-UC); and spliced, cleaved (S-C) (Figure 3.5.a). Sam68ΔC did not cause any marked change in the amount of US-C mRNA that could account for the loss of gp160/120 expression (Figure 3.5.c). In contrast, the inactive mutants, Sam68:5-262 and Sam68Δ28ΔC, increased the abundance of US-C mRNA to varying extents (Figure 3.5.c). The stimulation of cleavage of unspliced mRNA by Sam68:5-262 and Sam68Δ28ΔC is consistent with the known activity of full length Sam68 previously reported by our laboratory (290).
**Figure 3.5.** Sam68ΔC inhibits gp160/120 expression from the HIV-1 env reporter pgTat, but does not alter the amount of mRNA.

a. The pgTat reporter construct encodes HIV-1 env. gp160/120 are expressed from unspliced, cleaved pgTat. The black line indicates the position of the RNase Protection Assay (RPA) probe used. It covers the 3’ss and the polyadenylation site. 293T cells were transfected with pgTat. Rev, and pcDNA, Sam68ΔC or mutants thereof. 48 hours post-transfection, the cells were harvested for protein and RNA analysis. b. Western blots were probed for anti-gp120, anti-myc and anti-tubulin (loading control). c. RNA analysis of pgTat by RPA. There are four protections: unspliced-uncleaved (us-uc), unspliced-cleaved (us-c), spliced-uncleaved (s-uc) and spliced-cleaved (s-c). gp120 is translated from the us-c isoform of the RNA. 20 µg of total RNA was input to the assay.
3.4.4. Inhibition by Sam68ΔC is Not Associated with Alterations in Polyadenylation of Viral mRNAs

Deadenylation of mRNA is a common form of translational repression (81). To assess for changes in the polyadenylation state of viral mRNAs in the presence of Sam68ΔC, total RNA was extracted and the poly(A)+ fraction purified using oligo(dT)$_{25}$ beads. Given that 3’end cleavage and polyadenylation are tightly coupled processes, all cleaved mRNAs are expected to have a poly(A) tail and bind to the oligo(dT)$_{25}$ beads. Appearance of cleaved mRNA in the poly(A)- fraction would be indicative of a deadenylation event. Analysis of the poly(A)- and poly(A)+ fractions by RPA revealed that the fractionation was successful; uncleaved viral mRNAs (US-UC and S-UC) being predominantly in the poly(A)- fraction and cleaved versions (US-C and S-C) in the poly(A)+ fraction (Figure 3.6.a). Sam68ΔC did not shift the distribution of the unspliced RNAs between the fractions, indicating that the US-C form of pgTat mRNA retained a poly(A) tail of sufficient size to bind to the column (Figure 3.6.a).

To assay whether there were any significant changes in poly(A) tail length, we used a random amplification of cDNA ends, polyadenylation test (RACE-PAT) (432). An anchor primer consisting of an eGFP specific 15-mer followed by an oligo(dT) 20-mer was used to make cDNA. This anchor primer anneals randomly along the length of the poly(A) tail, generating cDNAs with a range of poly(A) tail lengths. Using an eGFP reverse primer and a gene specific forward primer we measured the lengths of the poly(A) tail on spliced and unspliced viral RNA (Figure 3.6.b). Without a poly(A) tail, the spliced amplicon is 340 nucleotides and the unspliced amplicon is 415 nucleotides. Therefore, the tail length of spliced pgTat mRNA is approximately 10-60 nucleotides and
**Figure 3.6.** Sam68ΔC does not alter the polyadenylation status of the affected HIV-1 viral mRNAs.

Cells were transfected with pgTat, Rev and pcDNA or Sam68ΔC. Total RNA was harvested and used in the assays below. a. 20 µg of total RNA was selected using oligo(dT) beads. Both the poly(A)- and poly(A)+ fractions were input into the RPA. b. RACE-PAT analysis. The illustration at top shows the position of the anchor primer used to make cDNA and the relative positions of the PCR primers. The anchor primer can anneal anywhere along the length of the poly(A) tail to make cDNA. The resultant PCR generates a smear representing various lengths of poly(A) tail. The amplicons were analyzed by fractionation on PAGE gels, and sizes of products determined by comparison to markers.
that of the unspliced mRNA is 200-250 nucleotides (Figure 3.6.b). No change in amplicon size was observed upon addition of Sam68ΔC. Thus, the loss of gp120 expression upon Sam68ΔC co-transfection cannot be attributed to deadenylation of the affected mRNA.

3.4.5. Sam68ΔC Alters the Association of Rev-dependent Viral mRNAs with PABP1

PABP1 has been shown to be an important promoter of mRNA translation through its direct interaction with eIF4G and resultant indirect interactions with the cap binding proteins (433, 434). Therefore, we examined PABP1 association with pgTat mRNA. The PABP1 mRNP immunoprecipitations (RIPs) efficiently precipitated PABP1 (Figure 3.7.a). RNA analysis determined that pgTat mRNAs were specifically precipitated with PABP1 as compared to control rabbit IgG RIPs (Figure 3.7.b). The S-C form of pgTat was the dominant isoform pulled down with anti-PABP1 antibody but US-C mRNA was also detected in the immunoprecipitates. PABP1 association was quantitated by comparing the ratio of US-C to S-C pgTat mRNA in the RIP relative to the input (Figure 3.7.c). Sam68Δ28ΔC and Sam68:5-262 had no significant effect on US-C mRNA association with PABP1. Sam68ΔC and Sam68Δ14-300 significantly reduced US-C mRNA association with PABP1 (p<0.05, Figure 3.7.b,c). Expression of Sam68ΔC had no effect on the immunoprecipitation of mRNPs containing the S/C form of pgTat mRNA, consistent with Sam68ΔC being able to discriminate between the US/C and S/C forms of the viral mRNA.
### a. WB: α-PABP

<table>
<thead>
<tr>
<th></th>
<th>pcDNA</th>
<th>Sam68ΔC</th>
<th>Δ14-300</th>
<th>Δ28ΔC</th>
<th>5-262</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
</tr>
<tr>
<td><strong>Ab:</strong></td>
<td>R-IgG</td>
<td>PABP</td>
<td>R-IgG</td>
<td>PABP</td>
<td>R-IgG</td>
</tr>
<tr>
<td><strong>IP</strong></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
</tr>
</tbody>
</table>

### b. Input and IP

<table>
<thead>
<tr>
<th></th>
<th>pcDNA</th>
<th>Sam68ΔC</th>
<th>Δ14-300</th>
<th>Δ28ΔC</th>
<th>5-262</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
</tr>
<tr>
<td><strong>IP</strong></td>
<td><img src="image16.png" alt="Image" /></td>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Legend:
- us-uc
- us-c
- s-uc
- s-c
Figure 3.7. Sam68ΔC inhibits the association of unspliced pgTat RNAs with PABP1. Cells were transfected with pgTat, Rev and pcDNA, Sam68ΔC or mutants thereof. Cell lysates were prepared 48 h post-transfection. a. PABP1 western blot showing the input and immunoprecipitated (IP) protein using either rabbit IgG (R-IgG) or anti-PABP1 (PABP). b. RPA showing input and immunoprecipitated (IP) pgTat mRNAs. RNA was extracted from either rabbit IgG (R-IgG) or anti-PABP1 (PABP) immunoprecipitates and analyzed by RPA. c. Quantitation of RPA shown in b. The ratio of us-c to s-c pgTat mRNA immunoprecipitated with anti-PABP1 standardized to the input. The amount of US-C precipitated in the presence of pcDNA was set to 1.0. Error bars represent one standard deviation. Data was quantitated from 3 independent experiments. Asterisk indicates a value significantly different (p-value<0.05) to pcDNA.
To confirm that a similar response is also seen in the context of the HIV-1 provirus, the experiment was repeated using a replication inactive form of the virus (HxBruR²/RΓ, Figure 3.8.a). Sam68ΔC strongly reduces both p24 and gp160/120 expression from this construct (Figure 3.8.b). However, the mutants have slightly different effects on the provirus than on the env reporter construct pgTat. While Sam68:5-262 enhances expression of gp160/120 and p24, Sam68Δ28ΔC and Sam68Δ14-300 have little effect. None of the mutants had any effect on Rev protein levels (top band, Figure 3.8.b). PABP1 was immunoprecipitated uniformly in all samples (Figure 3.9.a). The proviral mRNA was analysed using an RPA probe that spans the 5’ splice site within Gag, yielding 2 protection products. One protection product corresponds to the 9 kb and the other to the 2 and 4 kb classes of HIV-1 mRNAs. As translation of both the 9 kb (p55/p24) and 4 kb (gp160/120) classes of viral mRNAs are affected by Sam68ΔC, the band corresponding to the 2 and 4 kb classes could not be used for quantitation. Relative PABP1 association was quantitated by the ratio of 9 kb mRNA to actin mRNA in the RIP relative to the input. Sam68ΔC significantly decreased PABP1 association with HIV-1 9 kb mRNA relative to pcDNA, while Sam68:5-262 significantly increased PABP1 association with HIV-1 9 kb mRNA (p<0.01, Figure 3.9.b,c). Both Sam68Δ28ΔC and Sam68Δ14-300 showed slight reductions in PABP1 association with 9 kb viral mRNA, but this was not significant even at a higher p-value (p<0.05, Figure 3.9.b,c).
Figure 3.8. Sam68ΔC inhibits the expression from 9kb and 4kb HIV-1 proviral mRNAs.

a. Schematic of HxBruR'/RI', an HIV-1 proviral construct, which has point mutations in reverse transcriptase (RT) and integrase (IN). b. 293T cells were mock transfected or transfected with HxBruR'/RI' and pcDNA, Sam68ΔC or mutants thereof. 48 hours post-transfection total cell lysates were harvested and analysed by western blot. Westerns were probed with the following anti-HIV antibodies: anti-gp120 (4kb product), anti-p24 (9kb product) and anti-rev (2kb product, top band). Westerns were also probed with anti-myc antibodies (to confirm Sam68ΔC and mutant expression) and anti-tubulin (as a loading control).
a. WB: α-PABP

<table>
<thead>
<tr>
<th>Input</th>
<th>pcDNA</th>
<th>Sam68ΔC</th>
<th>Δ14-300</th>
<th>Δ28ΔC</th>
<th>5-262</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ab:</td>
<td>R-IgG</td>
<td>PABP</td>
<td>PABP</td>
<td>PABP</td>
<td>PABP</td>
</tr>
<tr>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b. IP

<table>
<thead>
<tr>
<th>Input</th>
<th>pcDNA</th>
<th>Sam68ΔC</th>
<th>Δ14-300</th>
<th>Δ28ΔC</th>
<th>5-262</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ab:</td>
<td>R-IgG</td>
<td>PABP</td>
<td>PABP</td>
<td>PABP</td>
<td>PABP</td>
</tr>
<tr>
<td>9 kb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 &amp; 4 kb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>actin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.9. Sam68ΔC inhibits the association of 9kb HxBruR'/RI' mRNAs with PABP1.

Cells were transfected with HxBruR'/RI' and pcDNA, Sam68ΔC or mutants thereof. Cell lysates were prepared 48 hours post-transfection. a. PABP1 western blot showing the input and immunoprecipitated (IP) protein using either rabbit IgG (R-IgG) or anti-PABP1 (PABP). b. RNase protection assays showing input and immunoprecipitated (IP) HxBruR'/RI' and actin mRNAs. RNA was extracted from either rabbit IgG (R-IgG) or anti-PABP1 (PABP) immunoprecipitates and analyzed by RPA. c. Quantitation of RPA shown in b. The ratio of 9kb HIV-1 to actin mRNA immunoprecipitated with anti-PABP1 standardized to the input. The amount of 9kb mRNA immunoprecipitated in the presence of pcDNA was set to 1.0. Error bars represent one standard deviation. Data was quantitated from 3 independent experiments. Asterisks indicate a value significantly different (p-value<0.01) to pcDNA.
3.4.6. *Sam68ΔC inhibits HIV-1 unspliced mRNA with a histone 3’UTR*

Histone transcripts are unique as they have a stem loop at the 3’ end instead of a poly(A) tail (reviewed in (99)). Histone 3’end formation is dependent on two RNA elements: the stem-loop and a downstream element called the histone downstream element (HDE). The stem-loop binding protein (SLBP) binds to the stem loop and stabilizes base-pairing between the HDE and the 5’end of U7snRNP. The SLBP-U7snRNP complex recruits other proteins, facilitating cleavage of the mRNA between the stem loop and the HDE. After cleavage SLBP remains bound to the stem-loop of histone mRNAs as they are exported to the cytoplasm, and facilitates translation. The stem loop-SLBP complex is analogous in function to the poly(A) tail-PABP1 complex in promoting translation of the mRNA. In light of our finding that Sam68ΔC inhibits HIV-1 unspliced mRNA translation by preventing the interaction with PABP1, we wanted to test whether an unspliced HIV-1 mRNA with a 3’ histone stem loop, not a poly(A) tail, would be susceptible to inhibition.

We removed the polyadenylation and cleavage site from pgTat and replaced it with the histone H3 3’UTR (Figure 3.10a). SLBP expression is replication dependent (high levels only in S-phase), making expression of histone mRNAs replication dependent. Therefore, to maximize the number of cell divisions, and thus SLBP and histone 3’UTR expression, we seeded our experiments with half the normal amount of cells and harvested them after 72, not 48, hours. 293T cells were transfected with Rev, pgTat-hist-3’UTR and pcDNA, Sam68ΔC or mutants thereof. The lysates show that pgTat-hist-3’UTR is Rev-responsive, there being no gp160/120 expression in the absence of Rev (Figure 3.10b). In addition, pgTat-hist-3’UTR responds to Sam68ΔC and its
mutants in a similar manner to HxBruR7/RI: Sam68ΔC and Sam68Δ28ΔC completely inhibited gp120 expression while Sam68Δ14–300 and Sam68:5-262 had no effect. We also looked at the effects of Sam68ΔC and mutants thereof on the levels of pgTat-hist-3′UTR mRNA (Figure 3.10c). The levels of the unspliced pgTat-hist-3′UTR mRNA do not change in the presence of Sam68ΔC or the mutants. It is worth noting that the amount of spliced pgTat-hist-3′UTR and Rev mRNAs was increased by Sam68ΔC. This does not correlate with any inhibitory effects as Sam68Δ28ΔC had no effect on the abundance of these two mRNAs.

The susceptibility of pgTat-hist-3′UTR to inhibition by Sam68ΔC led us to consider several possible explanations. First, it is possible that Sam68ΔC is generally able to block access to the 3′end of an mRNA; it could block SLBP access as easily as PABP1. Second, Sam68ΔC may inhibit translation through an alternate mechanism; loss of PABP1 binding is a symptom of this inhibition, not the cause. Third, it is possible that overexpression of pgTat-hist-3′UTR may make it cell cycle independent. To ensure that pgTat-hist-3′UTR was not polyadenylated, we extracted total RNA and purified the poly(A)+ fraction using oligo(dT)25 beads. The results show that a fraction of all of the mRNAs probed for: actin, rev, pgTat-hist-3′UTR spliced and unspliced were polyadenylated (Figure 3.11). Even if only a fraction of the unspliced pgTat-hist-3′UTR mRNA is polyadenylated, it would be impossible to draw any meaningful conclusions about the requirements of a poly(A) tail-PABP1 complex for Sam68ΔC inhibition.
Figure 3.10. pgTat-hist-3’UTR is inhibited by Sam68ΔC, but RNA levels are not affected.

293T cells were transfected with pgTat-hist-3’UTR, Rev and pcDNA, Sam68ΔC or mutants thereof. Cells were harvested 72 hours post-transfection. a. Schematic of pgTat-hist-3’UTR reporter construct. The pgTat polyadenylation and cleavage site was removed and the histone H3 3’UTR was inserted. The position of the histone stem loop (hist) is indicated, as is the position of the pgTat RPA probe. b. Western analysis of total proteins. c. 20µg of total RNA was input to the RPA.
Figure 3.11. pgTat-hist-3’UTR mRNA is polyadenylated. 293T cells were transfected with pgTat-hist-3’UTR, Rev and pcDNA, Sam68ΔC or mutants thereof. Cells were harvested 72 hours post-transfection and total RNA was harvested. 20 µg of total RNA was selected using oligo(dT) beads and both the poly(A)- and poly(A)+ fractions were input into the RNase protection assay.
3.5. Discussion

Previously, we reported that Sam68ΔC expression induced accumulation of both unspliced viral mRNA and Sam68ΔC in bundles at the outer periphery of the nucleus (406). Subsequent deletion analysis of Sam68ΔC revealed that although perinuclear bundling was seen with inhibitory mutants, it was also seen with some of the stimulatory mutants (see Chapter 2, Figure 2.8). This indicates that perinuclear bundling does not correlate with Sam68ΔC inhibition. In this chapter we show that integrity of the Sam68ΔC/viral mRNA perinuclear bundles is dependent on intact microfilaments (Figure 3.2, 3.3) but bundling is not essential for translational repression (Figure 3.4). This observation suggests that Sam68ΔC inhibits viral mRNA translation, not by changing its subcellular distribution, but through alterations in viral mRNA structure or composition of the mRNP.

Translational control has been studied extensively, and there are several well-defined mechanisms: changes in subcellular distribution (stress granules, P-bodies) (435, 436), changes in mRNA structure (deadenylation, cytoplasmic adenylation) (81), destabilizing mRNA elements (AREs) (174, 437), and blocking eIF4E initiation (4E-BPs) (438) to name a few. Our data show that inhibition by Sam68ΔC is not dependent upon changes in mRNA subcellular distribution (Figure 3.2, 3.4). These observations conflict with a recent report by Zhang et al that purported the mechanism of inhibition was a block to nuclear export (412). As evidence of this transport block, they performed nuclear/cytoplasmic fractionation in which Sam68ΔC sensitive mRNAs co-sedimented with the nuclear fraction. We clearly see that the inhibited mRNAs are in perinuclear
bundles in the cytoplasm, in the presence of Sam68ΔC (Figure 3.2). However, their association with the microfilaments makes these mRNPs insoluble (Figure 3.2, 3.3 and data not shown). As a result in such a fractionation assay they co-sediment with the nuclei and would be incorrectly scored as untransported. Sam68ΔC does not affect target mRNA abundance (Figure 3.5c), polyadenylation (Figure 3.6a), or polyA tail length (Figure 3.6b), only PABP1 binding (Figure 3.7, 3.9). This effect could be achieved either by Sam68ΔC blocking PABP2 exchange for PABP1, or removing PABP1 from the mRNA. We believe that inhibition of PABP1 binding to an mRNA, by Sam68ΔC, represents a novel mode of translation inhibition. Further studies will be necessary to identify the stage of cytoplasmic mRNA processing being affected by Sam68ΔC.

Sam68ΔC consistently inhibited expression of the incompletely spliced HIV-1 mRNAs, and had a corresponding effect on PABP1 binding. However, the mutants had different effects on pgTat and HxBruR7/RI reporters. Importantly, the effect of the mutants on viral protein synthesis was mirrored in the effects they had on PABP1 binding. For instance, Sam68:5-262 had no effect on gp160/120 expression from pgTat, but increased the amount of gp160/120 and p24 from HxBruR7/RI (Figure 3.5b, 3.8). Correspondingly, Sam68:5-262 had no significant effect on PABP1 binding to US-C pgTat mRNA but did significantly increase PABP1 binding to the 9 kb HxBruR7/RI mRNA (Figure 3.7, 3.9). In contrast, Sam68Δ14-300 was able to block expression of pgTat mRNA but not the 9 kb and 4 kb mRNAs of the provirus (Figure 3.5b, 3.8). Analysis of the effect of Sam68Δ14-300 revealed that it decreased binding of PABP1 to US/C pgTat mRNA but resulted in no significant alteration in PABP1 interaction with viral mRNAs (Figure 3.7, 3.9). The differences among the Sam68ΔC mutants may reflect
changes in the factors associated with these proteins that facilitate or inhibit remodeling of the viral mRNP following export from the nucleus.

The perinuclear accumulation of viral mRNA induced by Sam68ΔC, while not essential for the translational inhibition, is similar in phenotype to the effect seen upon depletion of human Rev-interacting protein (hRIP) (439). This finding suggests that both Sam68ΔC and hRIP may be influencing a similar step in viral mRNA metabolism. These observations are consistent with a model whereby the viral mRNP and hRIP interact at the nuclear periphery (directly or indirectly) to initiate remodelling of the viral mRNP and enhance its translation. A part of this remodelling is the exchange of PABP2 for PABP1. Sam68ΔC may interfere with this remodelling and prevent binding of PABP1, thereby inhibiting translation initiation (Figure 3.12). The recent identification of two RNA helicases (DDX1 and DDX3) that play essential roles in Rev function suggests that remodeling of the viral mRNP plays an important role in ensuring efficient expression of the HIV-1 structural genes (332, 340). How and where these helicases act remains to be determined.

To determine if the poly(A) tail/PABP1 complex is required for Sam68ΔC to inhibit translation of a target mRNA we assessed the effects of Sam68ΔC on pgTat-hist-3’UTR. In this reporter construct, the pgTat polyadenylation and cleavage site has been replaced with the human H3 histone 3’UTR. Sam68ΔC, and mutants thereof, affected gp120 expression from pgTat-hist-3’UTR in a similar manner to HxBruR7/RI: Sam68ΔC and Sam68Δ28ΔC completely inhibited gp120 expression while Sam68Δ14–300 and Sam68:5-262 had no effect (Figure 3.10). Subsequent analysis revealed that the pgTat-hist-3’UTR spliced and unspliced mRNAs were polyadenylated (Figure 3.11). As the
unspliced pgTat-hist-3’UTR mRNA is polyadenylated, it is impossible to draw any meaningful conclusions about the requirements of a poly(A) tail-PABP1 complex for Sam68ΔC inhibition.

There are several possible reasons for the aberrant polyadenylation of the histone 3’UTR in this context (reviewed in (99)). There are several differences between endogenous histone genes and pgTat which may account for polyadenylation of the pgTat-hist-3’UTR. First, histone transcripts are small (400-750 nucleotides) while the unspliced pgTat is approximately 3,000 nucleotides and spliced pgTat is approximately 800 nucleotides. Second, histone genes do not contain introns but unspliced pgTat mRNA does, and spliced pgTat mRNA would have an EJC deposited 5’ of the exon-exon junction. Additionally, the correct processing of histone 3’ends requires U7snRNP binding to the HDE. Unlike spliceosomal U5snRNPs which are present in 500,000-1,000,000 copies per cell, U7snRNP is present in only a few hundred molecules per cell. Histone gene clusters are localized to Cajal bodies, which also contain U7snRNPs. As pgTat-hist-3’UTR is expressed from transiently transfected plasmid DNA, it is not localized to Cajal bodies, and therefore may not be accessible to U7snRNP. Also, due to the large quantities of pgTat-hist-3’UTR mRNA there may not be sufficient cellular U7snRNP for correct 3’end processing. Overexpression of U7snRNA would be unlikely to rescue 3’end processing of pgTat-hist-3’UTR mRNA as several of the unique protein components of U7snRNP (Lsm10, Lsm11 and ZFP100) would also be present in limiting quantities.

In summary, we have discovered that Sam68ΔC is able to use the features of Crm1 mRNA export to specifically inhibit translation of Rev-dependent HIV-1 mRNAs
Inhibition occurs by a novel mechanism: inhibiting the association of PABP1 with the target mRNA (Chapter 3). Based on the data presented here, and other published data (332, 340, 439), we hypothesize that mRNAs transported via this export pathway have a unique mechanism through which they engage the translation machinery, possibly involving Rev, hRIP, DDX1 and DDX3 (Fig. 3.12). If indeed this is true, it would indicate that mRNAs using different export pathways have distinct RNP compositions and therefore separate control mechanisms. Future research into the details of translation initiation for Rev-dependent HIV-1 mRNAs may lead us to discover further discrepancies from host mRNAs, and lead to new therapeutic approaches to control HIV-1 gene expression.
Figure 3.12. Model of Sam68ΔC inhibition of unspliced HIV-1 mRNA translation.
Chapter 4: Incompletely Spliced HIV-1 mRNAs retain CBP20/80 Complex at Their 5’Cap

Kim Marsh, Alper Yilmaz¹, Kathleen Boris-Lawrie¹ and Alan Cochrane

1. Center for Retrovirus Research, Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Road, Columbus, Ohio, USA.

All of the experiments in this chapter were performed by Kim Marsh, except the polysome fractionations which were performed by Alper Yilmaz (Figure 4.7-10.a). Subsequent RNA and protein analysis of the polysomal fractions was performed by Kim Marsh.

Figures 4.1 and 4.3 have been submitted to Retrovirology as part of “Selective Translational Repression of HIV-1 RNA by Sam68ΔC Occurs by Altering PABP1 Association.” Marsh, K., Soros, V., and Cochrane, A. (2008).
4.1 Abstract

HIV-1 structural proteins are translated from incompletely spliced 9 kb and 4 kb mRNAs, which are transported to the cytoplasm by Crm1. It has been assumed that once in the cytoplasm, translation of incompletely spliced HIV-1 mRNAs occurs in the same manner as host mRNAs. We previously demonstrated that inhibition by Sam68ΔC is correlated with a loss of PABP1 binding with no attendant change in polyadenosine tail length of the affected mRNAs. The ability of Sam68ΔC to prevent PABP1 binding to Crm1 transported mRNAs led us to believe that these mRNAs engage with the translational machinery in a unique manner. In this chapter, we show that Rev and the RRE are required, but individually neither is sufficient, for complete Sam68ΔC inhibition. The selective inhibition of Crm1 exported HIV-1 mRNAs by Sam68ΔC suggests that it is able to recognize unique characteristics of these viral mRNPs. Study of the unspliced HIV-1 mRNP revealed that the nuclear cap binding complex, CBP20/80, is not exchanged for eIF4E. Additionally, in cells expressing the HIV-1 provirus, CBP80 relocates to the cytoplasm and co-sediments with polysomes. This supports the hypothesis that incompletely spliced HIV-1 mRNAs are translated in an eIF4E-independent, CBP20/80-dependent fashion. This property of the 9kb and 4kb HIV-1 mRNAs could be utilized to develop new therapeutic approaches to controlling HIV-1 infection.
4.2 Introduction

In the preceding chapters we have explored the mechanism through which Sam68ΔC is able to specifically inhibit expression of unspliced HIV-1 mRNAs. We have defined the minimal domains of Sam68ΔC required for inhibition (Chapter 2). We have also determined that Sam68ΔC perinuclear bundling of a target mRNA is not sufficient for inhibition (Chapter 2). We found that Sam68ΔC inhibits translation, not transport of target mRNAs (Chapter 3). Translational inhibition of these Sam68ΔC sensitive mRNAs is not due to deadenylation nor does Sam68ΔC affect poly(A) tail length (Chapter 3). We found that Sam68ΔC significantly inhibits PABP1 association with target mRNAs; PABP1 association with a target mRNA correlates with the effects of Sam68ΔC, or mutants thereof, on translation of the mRNA (Chapter 3). PABP1 forms oligomers on the poly(A) tail of the mRNA, inhibiting mRNA turnover and, through interaction with eIF4G, stimulates translation of the mRNA (433, 434, 440). Loss of PABP1 binding therefore renders an mRNA untranslatable.

Completely spliced HIV-1 mRNAs, and the vast majority of host mRNAs, are exported via the Tap/p15 export pathway (reviewed in (113, 114)). The TREX complex is deposited at the 5’ end of an mRNA during splicing, and requires interaction between the TREX components, Ref-Aly, and the nuclear cap binding complex, CBP20/80 (122). Tap/p15 binds to host mRNAs via the TREX complex and mediates passage from the nucleus to the cytoplasm through the nuclear pore complex. In the cytoplasm the RNA helicase, hDbp5, removes the Tap/p15 transport receptor from mRNAs. Removal of Tap/p15 acts as a molecular ratchet to prevent re-import of the mRNA, and permits
translation of the mRNAs in the cytoplasm (reviewed in (128, 129)). While the “pioneer” round of translation is mediated via CBP20/80, eIF4E replaces CBP20/80 at the cap and mediates subsequent rounds of translation (32, 186-188).

We wanted to further understand the specificity of Sam68ΔC inhibition of unspliced HIV-1 mRNA translation. As we have shown previously, mRNAs transported via the RRE/Rev/Crm1 pathway are susceptible to inhibition, but mRNAs transported via the CTE/Tap/p15 pathway are not (Chapter 2). As it is translation, not transport to the cytoplasm, that is blocked (Chapter 3) this indicates that mRNAs transported by the Crm1 pathway interact with the translational apparatus in a unique manner. Indeed, there are many indications that this may be the case. First, several labs have shown that Rev plays a role in promoting the translation of incompletely spliced HIV-1 mRNAs, independent of its role in transport (365, 366). Second, the host factor hRIP is required for Rev transported mRNAs to move away from the perinuclear region, and their subsequent translation (339). Third, the host RNA helicase, DDX3, is required for Rev-mediated Crm1 transport (340). It has been proposed that DDX3 acts in a manner analogous to hDbp5 (132-134). We wanted to explore how Rev transported HIV-1 mRNAs are distinct from Tap/p15 transported mRNAs in order to determine what makes them susceptible to inhibition by Sam68ΔC.

In this chapter we have further explored the contributions of the RRE, Rev and Crm1 to mRNAs sensitivity to Sam68ΔC inhibition. Our results confirmed that Sam68ΔC could significantly inhibit expression of an RRE/Rev/Crm1 transported mRNA, but not a Tap/p15 transported mRNA. We went on to show that the RRE contributes to Sam68ΔC recognition of a target mRNA, but can not confer sensitivity on a Tap/p15 exported
reporter. Susceptible mRNAs are exported via the Crm1 pathway and contain an RRE, indicating that these mRNAs engage in the translational machinery in a unique manner. We went on to examine the proteins bound to the 5′cap of HIV-1 mRNAs. We were surprised to find that while completely spliced HIV-1 mRNAs are associated with eIF4E, unspliced HIV-1 mRNAs are not. We went on to show that both unspliced and spliced HIV-1 mRNAs associate with CBP80. This indicates that the unspliced HIV-1 mRNAs retain the nuclear CBP20/80 cap binding complex in the cytoplasm during translation. To confirm this, we show that CBP80 relocalizes to the cytoplasm in cells expressing HIV-1 provirus. We confirm that CBP80 is bound to actively translated mRNAs, by showing its presence in the polysome fraction. This finding corroborates the growing body of knowledge that unspliced HIV-1 mRNAs interact with the translational machinery in a manner distinct from Tap/p15 exported mRNAs.

4.3. Materials and Methods

4.3.1. Expression constructs.

The following constructs have been previously described: SV-Hygro, SV-H6Rev, CMVmyc3xterm (289), pDM128 (413), pgTat (297), Bl-env-HindIII (305), Sam68 and Sam68ΔC (406). Bl-actin and Bl-SD-Gag were described in Chapter 3.3.1. pDM128-4xMS2 and pCMV-MS2-Rev were provided by Dr. B. Cullen, Duke University (441). Gag-CTE and Gag-RRE plasmids were provided by Dr. K. Boris-Lawrie, Ohio State University. FS-Gag-GFP was provided by C. Liang, McGill University. HxBruR'/RI' was provided by Eric Cohen, Universite de Montreal. Bl-CMV-CAT-RREpA was generated by restriction digest of Bl-CMV-RREpA with BamHI and XbaI. The RRE fragment was ligated into the corresponding sites of Bl-CMV-CATpA.
4.3.2. **Cell Lines and Transfections.**

HeLa and 293T cells were maintained in Iscove’s modified Dulbecco’s medium (IMDM) supplemented with 10% fetal bovine serum (FBS) (v/v), 50µg/mL gentamycin sulfate and 2.5µg/mL amphotericin B. Vectors were introduced to 293T cells by calcium phosphate transfection (414). Vectors were introduced to HeLa cells by Fugene 6 transfection reagent (Roche) following manufacturer’s protocol. Cells were harvested two days post-transfection. To generate FSRGag HeLa cells: HeLa cells were transduced with the HIV-1 provirus FSR-GagGFP, and selected by FACS.

4.3.3. **Antibodies.**

The following antibodies were used: mouse anti-myc (Invitrogen), mouse anti-tubulin (Sigma), mouse anti-p24 (clone 183-1112-5C), rabbit anti-Rev (2/99), rabbit anti-CBP80 (courtesy of I.W. Mattaj, EMBL), and mouse anti-eIF4E (P2) (Santa Cruz). The following secondary antibodies were used: donkey anti-rabbit IgG conjugated HRP, donkey anti-mouse IgG conjugated HRP, donkey anti-rabbit IgG conjugated Texas red, donkey anti-rabbit IgG conjugated FITC, donkey anti-mouse IgG conjugated Texas red (Jackson Immuno Research), and protein G conjugated HRP (Molecular Probes).

4.3.4. **CAT Assays.**

2x10^5 293T cells were transfected in triplicate, per condition. For Rev-dependent expression cells were transfected with: 0.125 µg pDM128, 0.025 µg SV-Hygro or SV-H6Rev, and 0.5 µg pcDNA3.1 or Sam68ΔC. For Rev-independent expression cells were transfected with: 0.5 µg CMV-CAT-RRE and 0.5 µg pcDNA3.1 or Sam68ΔC. CAT assays were performed as previously described (415). Conversions were normalized by protein concentration as determined by Bradford assay (BioRad).
4.3.5. Western Blots.

6x10^5 293T cells were transfected as follows: 0.5 µg GagRRE, 0.1 µg SV-Hygro or SV-H6Rev, and 2.0 µg pcDNA3.1 or Sam68∆C. For Rev-independent expression cells were transfected with 2.0 µg GagCTE, and 2.0 µg pcDNA3.1 or Sam68∆C. Cells were harvested in whole cell lysis buffer (150mM NaCl, 10mM Na₂HPO₄, 1% Triton X-100 v/v, 0.1% SDS w/v, 0.2% sodium azide w/v, 0.5% sodium deoxycholate w/v, 1mM sodium orthovanadate), fractionated on SDS-PAGE gels and transferred to PVDF membrane (Pall Life Sciences). Bound antibodies were detected using Western Lightning (Perkin Elmer). For quantitation of western blots, films were scanned and analyzed using Imagequant software. Signals for both p55 and p24 were combined for analysis and data normalized using the tubulin signals. Quantitation shown was generated from three independent experiments.

4.3.6. RNP Immunoprecipitation.

3x10^6 293T cells were transfected with 0.4 µg SVH6Rev, 2.0 µg pgTat and 8.0 µg of pcDNA3.1, Sam68∆C or mutants thereof. Whole cell lysates were generated by a modification of the protocol of Siomi et al (431). 48 hours post-transfection, cells were washed twice with 1xPBS and once with Buffer A (110mM potassium acetate, 2mM magnesium acetate, 2mM DTT, 10mM HEPES pH 7.5). The cell pellet was resuspended in 400 µl Buffer B (10mM potassium acetate, 2mM magnesium acetate, 2mM DTT, 5mM HEPES pH 7.5, 20µM cytochalasin D) and incubated on ice for 10 min. Cells were disrupted by passage through 25 gauge needle, 5 times. KCl was added to 100mM final concentration. Cells debris was pelleted by centrifugation at 1,500 rpm. Supernatant was adjusted to 0.5% NP40 v/v and incubated on ice for 20 min. Insoluble material was
pellet by centrifugation at 14,000 rpm and supernatant retained for use in subsequent immunoprecipitations. 1/42 of the supernatant was retained for analysis of the RNA input, and 1/42 was retained for analysis of protein input. For each immunoprecipitation, 10/42 of the soluble fraction was diluted with 10 volumes RIPA buffer (150mM NaCl, 50mM Tris-HCl pH 7.5, 1% NP40 v/v, 1mM EDTA ph 7.5, 0.5% sodium deoxycholate w/v, 0.05% SDS w/v) and 100u RNaseOUT (Invitrogen). The sample was pre-cleared by incubation with Gammabind plus sepharose (GE Healthcare) for 1 h at 4°C. Supernatant was incubated overnight at 4°C with the indicated antibody. Immune complexes were collected by addition of Gammabind plus sepharose and incubation at 4°C for 1 h. The resin was washed 5x with 1 ml of RIPA buffer and then resuspended in 100 µl of RIPA. RNA was extracted from 150 µl of the resin slurry (430) and 50 µl of slurry was used for protein analysis following addition of 50 µl 2x dissociation buffer.

4.3.7. Immunofluorescence.

Transiently transfected proteins were detected as follows: 3x10^5 HeLa cells were transfected with 0.25 µg SV-H6Rev, 1.25 µg pgTat and 5.0 µg pcDNA3.1, Sam68 or Sam68ΔC. 48 hours post-transfection, cells were fixed for immunofluorescence as previously described (305). Rev was detected with rabbit anti-Rev antibody (2/99) and myc-tagged proteins were detected with monoclonal anti-myc antibody (Invitrogen). Immunofluorescence was detected using a Leica DMR microscope at either 400x or 630x magnification.

Stably transfected and endogenous proteins were detected as follows: 3x10^5 HeLa or FS-Gag HeLa cells were plated on coverslips in 6-well dishes. 24 hours after plating they were fixed as previously described (305). For p24 and CBP80 co-staining: cells were
probed with the primary antibodies anti-p24 and anti-CBP-80, and the secondary antibodies donkey anti-mouse-IgG conjugated FITC and donkey anti-rabbit-IgG conjugated Texas Red. For p24 and eIF4E co-staining both primary antibodies were mouse IgG. Therefore, anti-eIF4E was conjugated with AlexaFluor594 using a kit from Zenon, according to the manufacturer’s protocol. Cells were initially probed with the primary anti-p24, then the secondary donkey anti-mouse-IgG conjugated FITC, and finally with anti-eIF4E conjugated AlexaFluor594. Anti-eIF4E was conjugated with AlexaFluor594 using a kit from Zenon, according to the manufacturer’s protocol. Immunofluorescence was detected using a Leica DMR microscope at either 400x or 630x magnification.

4.3.8. Polysome Gradients.

For ribosomal profile analysis, $1 \times 10^7$ 293 cells in T150 flasks were transfected and treated 48 h later with 100 µg of cycloheximide/ml for 15 min at 37°C. The cells were harvested in phosphate-buffered saline. Cells were resuspended in hypotonic lysis buffer (10 mM HEPES [pH 7.9], 1.5 mM MgCl$_2$, 10 mM KCl, and 0.5 mM dithiothreitol) and placed on ice for 10 min, followed by gentle vortexing for 10 s. Nuclei were pelleted by centrifugation at 3,000 g for 2 min at 4°C. The cytoplasmic supernatant was then layered onto a 10-ml linear gradient of 15 to 45% sucrose (w/v) in 10 mM HEPES containing 10 mM NaCl, 3 mM CaCl$_2$, 7 mM MgCl$_2$, and 1 mM dithiothreitol and was centrifuged at 36,000g at 4°C for 3.5 h in a Beckman SW41 rotor. Gradients were fractionated, and the $A_{254}$ was monitored using an ISCO (Lincoln, Nebr.) fractionation system.
Each fraction was split in half for RNA and protein analysis. The RNA fraction was precipitated with ethanol, and 1 µg tRNA, overnight at -80°C. The RNA was pelleted by centrifugation at 12,000 g for 20 minutes at 4°C. The RNA pellet was resuspended in TRI Reagent (Sigma) and extraction followed manufacturer’s directions. The final RNA pellet was resuspended in 20 µl DEPC-H2O and analyzed by RNase protection assay (as previously described Chapter 3.3.7. and (290)). The protein fraction was precipitated on ice with an equal volume of 40% trichloroacetic acid (w/v), and 50 µg BSA as carrier. After 1 hour, the protein was pelleted by centrifugation at 12,000g for 20 minutes at 4°C. The protein pellet was washed 5 times with acetone, dried in a speed-vac (medium rate, cryo-pump on) and resuspended in 2 x dissociation buffer. The proteins were run out on 10% SDS-PAGE gels (v/v) for analysis as per section 4.3.5.

4.3.9. **Statistical Analysis.**

The data are presented as mean +/- one standard deviation. Data were compared using Student’s t-test. * indicates a p-value < 0.05. ** indicates a p-value < 0.01.

4.4. **Results**

4.4.1. **Susceptibility to Sam68ΔC Repression is Conferred by the Nuclear Export Pathway**

The ability of Sam68ΔC to selectively suppress expression of the 9 kb and 4 kb classes of HIV-1 mRNAs suggested that there is some unique feature that renders them susceptible to repression. Cellular mRNAs use the Tap/15 export pathway, while HIV-1 9 and 4 kb mRNAs are incompletely spliced and contain sequences preventing their export by Tap/p15 (299-306). These incompletely spliced HIV-1 mRNAs are exported from the nucleus via HIV-1 Rev which mediates interactions with the host protein Crm1
Two possible explanations for repression of the 9 kb and 4 kb HIV-1 mRNAs by Sam68ΔC are immediately apparent: either they contain unique RNA sequences recognized by Sam68ΔC or export via the Crm1 pathway marks the viral mRNA for inhibition.

To address this question directly, we examined the ability of Sam68ΔC to inhibit expression from HIV-1 Gag expression vectors utilizing different nuclear export elements; the constitutive transport element (CTE) from Mason-Pfizer Monkey virus that interacts directly with Tap/p15 (GagCTE), or the RRE that requires Rev and Crm1 (GagRRE) (Figure 4.1.a). Gag mRNA generates a 55 kDa polyprotein that is subsequently processed by the viral protease into matrix (p17), capsid (p24), nucleocapsid (p9) and p6. Gag expression was measured by anti-p24 western blot which recognizes the p55 precursor as well as p24. GagRRE expression is dependent upon Rev, and is reduced to baseline by Sam68ΔC (Figure 4.1.b,c). In contrast, Sam68ΔC had no significant effect on expression from GagCTE (Figure 4.1.b,c). This demonstrates that it is not the Gag sequence, but rather the RRE and/or the Crm1 export pathway that dictates inhibition by Sam68ΔC.
Sam68ΔC inhibits translation of RRE/Rev/Crm1 exported mRNA but not Tap/p15 exported mRNA.

a. GagRRE and GagCTE reporter constructs. GagRRE is a Rev-dependent reporter which expresses HIV-1 Gag-pol (p55/p24) and GagCTE is a Rev-independent reporter which expresses HIV-1 Gag-pol (p55/p24). b. 293T cells were transfected with GagRRE, Rev and pcDNA or Sam68ΔC. Or 293T cells were transfected with GagCTE and pcDNA or Sam68ΔC. 48 hours post-transfection cells were harvested and the lysates run out on SDS-PAGE gels. Western blots were probed with anti-p24, anti-myc and anti-tubulin (loading control) antibodies. c. Quantitation of Western blots from 3 independent experiments. Error bars indicate 1 standard deviation.
To determine whether or not the RRE is a determinant of Sam68ΔC sensitivity we tested the Rev-dependent reporter pDM128, and the variant pDM128-4xMS2. pDM128 contains the chloramphenicol acetyl transferase (CAT) open reading frame within the HIV-1 env intron, which also contains the RRE (Figure 4.2.a). CAT expression is a measure of Rev-activity. pDM128-4xMS2 has had the RRE removed and replaced with four MS2-high affinity RNA binding sites (Figure 4.2.a). Rev-binding is mediated by fusion of the MS2-coat protein domain to the N-terminus of Rev. Expression of pDM128-4xMS2 in the presence of MS2-Rev is dependent on transport via the Crm1 pathway (441). As shown in Chapter 2, pDM128 is significantly inhibited by Sam68ΔC, and is not significantly different from basal CAT expression in the absence of Rev (Figure 4.2.b, p<0.01). CAT expression from pDM128-4xMS2 is significantly stimulated by MS2-Rev (Figure 4.2.b, p<0.01). Sam68ΔC inhibits MS2-Rev induced CAT expression from pDM128-4xMS2, but CAT expression is still significantly more than in the absence of MS2-Rev (Figure 4.2.b, p<0.05). Therefore, Sam68ΔC has reduced capacity to inhibit the expression of the Crm1 transported mRNA which lacks the RRE. This indicates that the RRE contributes to making a reporter Sam68ΔC sensitive.
Figure 4.2. Complete Sam68ΔC inhibition requires the RRE.
a. pDM128 and pDM128-4xMS2 reporter constructs. pDM128 is a Rev-dependent reporter which contains a CAT open reading frame within the HIV-1 env intronic sequences. In pDM128-4xMS2 the RRE has been replaced with four MS2-binding sites. 
b. 293T cells were transfected with pDM128, with or without Rev, and pcDNA or Sam68ΔC. Or 293T cells were transfected with pDM128-4xMS2, MS2 or MS2-Rev, and pcDNA or Sam68ΔC. 48 hours post-transfection cells were harvested and the lysates assayed for CAT-activity. Quantitation of CAT assays from 3 independent experiments. Error bars indicate 1 standard deviation.
To test whether the RRE is sufficient to confer Sam68ΔC sensitivity on a reporter, the RRE was inserted into the Rev-independent reporter, CMV-CAT and effects on CAT expression measured (Figure 4.3.a). Rev had no significant effect on the level of CAT expression, while Sam68ΔC was marginally stimulatory, in the presence of Rev (Figure 4.3.b, p<0.05). The results from these experiments demonstrate that while the RRE may contribute to Sam68ΔC recognition of a target mRNA, it is not sufficient to confer Sam68ΔC repression. All susceptible mRNAs contain the RRE and are exported via the Rev/Crm1 pathway, indicating that there may be something unique in the structure/composition of Crm1 transported mRNPs.
Figure 4.3.  The RRE-alone is not sufficient to confer Sam68ΔC sensitivity on a heterologous reporter.
a. CMV-CAT-RRE reporter construct. CMV-CAT-RRE expresses CAT, and contains the HIV-1 RRE inserted 3’ of the CAT open reading frame. b. 293T cells were transfected with CMV-CAT-RRE, with or without Rev, and pcDNA or Sam68ΔC. 48 hours post-transfection cells were harvested and the lysates assayed for CAT-activity. Quantitation of CAT assays from 3 independent experiments. Error bars indicate 1 standard deviation.
4.4.2. Sam68ΔC does not affect Rev subcellular distribution

Our results suggest that it is the Rev/RRE complex which is recognized by Sam68ΔC (Figure 4.1-4.3). This complex may be recognized directly or indirectly by Sam68ΔC. If the Rev/RRE complex is recognized directly by Sam68ΔC, Rev may re-localize to the perinuclear bundles in the presence of pgTat and Sam68ΔC. Results from a number of labs have revealed that the steady state localization of Rev to the nucleus is required for Rev activity. SiRNA knockdown of DDX1 or overexpression of NF90ctv both result in relocalization of Rev to the cytoplasm and loss of Rev activity (332, 337, 338). Rev was shown to play a role in loading its mRNA cargo into polysomes, and this activity is distinct from its role in mRNA transport (365, 366). Therefore, we wanted to examine whether Sam68ΔC was altering the subcellular distribution of Rev, and sequestering Rev in perinuclear bundles. If it does, it could indicate that Sam68ΔC blocks Rev-mediated mRNA loading into polysomes.

HeLa cells were transfected with Rev, pgTat, and myc-tagged Sam68 or Sam68ΔC, and 48 hours post-transfection the cells were fixed for immunofluorescence. Anti-Rev and anti-myc antibodies were used to determine the localization of Rev and myc-Sam68 or myc-Sam68ΔC respectively (Figure 4.4). In the presence of Sam68 or Sam68ΔC, Rev retains its nuclear/nucleolar localization. Unfortunately, simultaneous probing for unspliced HIV-1 pgTat mRNA was not possible as we did not have 4 filters available. However, we were able to confirm the inhibitory activity of Sam68ΔC by the formation of perinuclear bundles (Figure 4.4). Based on this observation, it is unlikely that Sam68ΔC directly binds to the Rev/RRE complex. It is more likely that Sam68ΔC is
Figure 4.4. Neither Sam68 nor Sam68ΔC cause Rev relocalization. HeLa cells were transfected with pgTat, Rev and pcDNA, Sam68 or Sam68ΔC. Cells were fixed for immunofluorescence and probed with anti-myc (red) and anti-Rev (green) antibodies. Nuclei were stained with DAPI (blue).
able to indirectly recognize the Rev/RRE complex, perhaps through the unique composition of RRE/Rev/Crm1 transported mRNPs.

4.4.3. Rev-exported HIV-1 mRNA is CBP80, not eIF4E, Bound

We have previously determined that Sam68ΔC is able to modify the structure of a sensitive mRNP, by removing or preventing the binding of PABP1 to the poly(A) tail of a target mRNA (Chapter 3). This change in mRNP composition prevents the translation of these mRNAs, but we are not sure if it causes the inhibition or is a consequence of Sam68ΔC translation inhibition. Next we wanted to look at the effects of Sam68ΔC on proteins which mediate translation by binding to the 5’cap: eIF4E and CBP80. eIF4E displaces CBP20/80 at the 5’cap of the mRNA, after the pioneer round of translation, and recruits translation initiation factors (32). To monitor association of these proteins with the HIV-1 mRNAs of interest, RNP immunoprecipitation (RIP) assays (442) were performed using antibodies to either eIF4E or CBP80. Immunoprecipitated viral mRNA was detected by RPA. The perinuclear, microfilament bound, bundles created by Sam68ΔC in the presence of unspliced HIV-1 mRNA, are highly insoluble and special lysis conditions were required to solubilize the mRNP complex. We adapted a technique used by Siomi et al (431), which uses a hypotonic lysis buffer containing cytochalasin D to solubilize the proteins. As shown in Figure 3.4, the diffuse HIV-1 mRNA is still translationally repressed by Sam68ΔC, so this technique maintains functionally relevant conditions.

Initial experiments examining the association of eIF4E with HIV-1 mRNAs provided a surprising result. Although anti-eIF4E antibody was successful in precipitating eIF4E protein (Fig. 4.5.a), analysis of the associated mRNAs revealed that only the S-C
form of pgTat mRNA was pulled down (Fig. 4.5.b). In the absence of Sam68ΔC, no signal was obtained for the US-C form of pgTat mRNA despite the fact that western blots confirm that the mRNA is being translated under these conditions (as illustrated by the synthesis of gp160/120, Figure 3.5.b). The initial cap-binding complex deposited on the mRNA in the nucleus is CBP20/80 (24). To assess whether the CBP20/80 complex remains associated with the viral mRNA, RIPS were also carried out using antibody to CBP80. Analysis of the associated viral mRNA revealed the presence of all env mRNA species. This observation is consistent with previous studies that demonstrated co-transcriptional association of CBP20/80 with RNA polymerase II-generated RNAs (24). Sam68ΔC did not alter association of viral mRNA with CBP80, indicating that translational repression could not be attributed to alterations in cap function.
Figure 4.5. Unspliced, cleaved pgTat mRNA is CBP80, not eIF4E, bound. 293T cells were transfected with pgTat, Rev and pcDNA or Sam68ΔC. Cells were harvested 48 hours post-transfection and the lysates were immunoprecipitated with anti-eIF4E or anti-CBP80 antibodies. a. Analysis of the proteins pulled down in the immunoprecipitations. b. Analysis of the RNA pulled down in the immunoprecipitations.
4.4.4. CBP80 subcellular distribution shifts to the cytoplasm in the presence of HIV-1 provirus

We wanted to confirm that the unspliced HIV-1 mRNAs retain CBP20/80 at their cap during translation. To do this we examined the subcellular distribution of CBP80 and eIF4E in untransfected HeLa cells and HeLa cells stably expressing an HIV-1 proviral construct (HeLa FS-Gag-GFP). eIF4E has a similar distribution (predominantly cytoplasmic) in both cell types (Figure 4.6). The subcellular distribution of CBP80, on the other hand, shifted in the presence of HIV-1 provirus. In HeLa cells, the majority of the protein is nuclear, with a small fraction of the protein present throughout the cytoplasm (Figure 4.6). In HeLa FS-Gag-GFP cells, there is a distinct increase in the cytoplasmic fraction of CBP80 (Figure 4.6). This data is consistent with CBP20/80 complex being retained on cytoplasmic, Crm1 transported HIV-1 mRNAs.
Figure 4.6. CBP80 partially relocalizes to the cytoplasm in the presence of HIV-1 proviral mRNA. HeLa cells and HeLa cells stably expressing the proviral construct FS-Gag-GFP were plated on coverslips and harvested 24 hours later. Cells were fixed for immunofluorescence and probed with anti-p24 (green) and anti-CBP80 (red) or anti-eIF4E (red) antibodies. Nuclei were stained with DAPI (blue).
4.4.5. **CBP80 is associated with polysomes in pgTat/Rev transfected cells**

If CBP20/80 is retained at the cap of all incompletely spliced HIV-1 mRNAs during translation, CBP80 should be associated with the polysomal fraction of the cells. We transfected 293T cells with Rev and the HIV-1 *env* reporter construct pgTat. 48 hours post-transfection the cells were harvested and the lysates passed through a sucrose gradient. The absorbance at 254 λm ($A_{254}$) was measured, and the profile fits a typical polysome gradient (Figure 4.7.a). The free RNPs are near the top (fraction 1), then 40S ribosomal subunit, 60S ribosomal subunit, 80S monosomes and the polysome fraction (fractions 9-21). Fractions of the gradient were taken off starting at the top (fraction 1). To avoid any cross-contamination between samples, only every other fraction was analysed. These fractions were analysed for their protein content (Figure 4.7.b) and their RNA content (Figure 4.7.c). eIF4E is found in the free RNPs (fraction 1) and throughout the polysome fraction (fractions 9-21). CBP80 is found predominantly in fractions 5-13, but there is some in the heavier polysome fractions (fractions 15-21) as well. When we look at the distribution of the HIV-1 mRNA we see that the S-C is found in fractions 7-21 (and predominantly in fractions 13-19), and the US-C is found in fractions 11-21. As a control the localization of endogenous actin mRNA was probed. It is found in increasing amounts from fraction 9 through 21.

To confirm that the presence of CBP80, eIF4E and the mRNA in the lower fractions of the gradient is due to association with polysomes, the experiment was repeated in the presence of EDTA which dissociates polysomes (Figure 4.8). The $A_{254}$ profile over the gradient shows a shift in protein/RNA to the lighter fractions (1-11), indicating polysome dissociation (Figure 4.8.a). The eIF4E and CBP80 localizations are
CBP80 and eIF4E co-sediment with polysomes in cells transfected with pgTat and Rev. 293T cells were transfected with pgTat and Rev. After 48 hours, polysomes were purified through a sucrose gradient. Every other fraction was analysed. a. Polysome profile. b. Western blot with anti-eIF4E and anti-CBP80 antibodies. c. RPA analysis of RNA.
Figure 4.8. CBP80 and eIF4E move to lighter fractions when polysomes are dissociated by EDTA. 293T cells were transfected with pgTat and Rev. 48 hours post-transfection, the lysates treated with EDTA to dissociate the polysomes and then the lysates were purified through a sucrose gradient. Every other fraction was analysed. a. Polysome profile. b. Western blot with anti-eIF4E and anti-CBP80 antibodies. c. RPA analysis of RNA.
also shifted to the lighter fractions: eIF4E to fractions 9-15 and CBP80 to fractions 3-17, but concentrated in fractions 7 and 9 (Figure 4.8.b). The mRNA is similarly shifted to the lighter fractions: pgTat US-C to fractions 9-15, pgTat S-C to fractions 5-13, and actin to fractions 7-17 (Figure 4.8.c). These results confirm that eIF4E and CBP80 are found in the polysome fraction, as are the translated pgTat mRNAs (S-C and US-C) and actin mRNA.

4.4.6. **CBP80 is associated with polysomes in HIV-1 provirus transfected cells**

We wanted to confirm that the CBP80 is present in the polysomes in the presence of a more biologically relevant reporter construct, the HIV-1 proviral construct HxBruR⁻/R₁. We transfected 293T cells with HxBruR⁻/R₁, 48 hours post-transfection the cells were harvested and the lysates passed through a sucrose gradient as before. The A₂₅₄ profile fits a typical polysome gradient (Figure 4.9.a). The free RNPs are near the top (fraction 1), and the polysome fraction is near the bottom (fractions 11-21). Again, the gradient fractions were taken off starting at the top (fraction 1) and every other fraction was analysed, to avoid any cross-contamination. Fractions were analysed for their protein content (Figure 4.9.b) and their RNA content (Figure 4.9.c). eIF4E is found in the free RNPs (fraction 1) and throughout the polysome fraction (fractions 13-21). CBP80 is found in fractions 7-21 but predominantly in fractions 13-19. When we look at the distribution of the HIV-1 mRNA we see that the 9kb mRNA is found in fractions 13-21, and the 2 and 4kb mRNAs are found in fractions 13-21, but mostly 19 and 21. Again, the localization of endogenous actin mRNA was used as a control. It is found in increasing amounts from fraction 9 through 21.
Figure 4.9. CBP80 and eIF4E co-sediment with intact polysomes in cells transfected with HxBruR/R1.
293T cells were transfected with HxBruR/R1. 48 hours post-transfection, the polysomes were purified through a sucrose gradient. Every other fraction was analysed. a. Polysome profile, from top to bottom (left to right). b. Western blot with anti-eIF4E and anti-CBP80 antibodies. c. RPA analysis of RNA.
Figure 4.10. CBP80 and eIF4E move to lighter fractions when polysomes are dissociated by EDTA.

293T cells were transfected with HxBruR/RI. 48 hours post-transfection, the lysates treated with EDTA to dissociate the polysomes and then were purified through a sucrose gradient. Every other fraction was analysed for RNA and protein. a. Polysome profile, from top to bottom (left to right). b. Western blot with anti-eIF4E and anti-CBP80 antibodies. c. RPA analysis of RNA.
To confirm that the presence of CBP80, eIF4E and the mRNA in the lower fractions of the gradient is due to association with polysomes, the experiment was repeated in the presence of EDTA which dissociates polysomes (Figure 4.10). The $A_{254}$ profile over the gradient shows a shift in protein/RNA to the lighter fractions (1-11), indicating polysome dissociation (Figure 4.10.a). The eIF4E and CBP80 localizations are also shifted to the lighter fractions: eIF4E to fractions 7-15 and CBP80 to fractions 9-17, but concentrated in fractions 11 and 13 (Figure 4.10.b). The mRNA is similarly shifted to the lighter fractions: 9kb HIV-1 mRNA to fractions 13 and 15, 2 and 4kb HIV-1 mRNA to fractions 7-15, and actin mRNA to fractions 7-15 (Figure 4.10.c). These results confirm that eIF4E and CBP80 are found in the polysome fraction, as are the HIV-1 and actin mRNAs. The presence of CBP80 in the polysome fraction in the presence of pgTat and Rev, or HxBruR/R confirms that the results seen in Figures 4.5 and 4.6: incompletely spliced, translated HIV-1 mRNAs retain CBP20/80 at their 5’cap.

4.5. Discussion

Sam68ΔC is a specific and potent inhibitor of HIV-1 (Chapters 2, 3, Figure 4.1). We have shown here that targets marked for inhibition require the RRE, but this element alone is not sufficient to define a target (Figures 4.2 and 4.3). The initial study by Reddy et al showed that GST-Sam68 but not GST-Sam68ΔC could bind the RRE in vitro (403). There are several possible explanations for this apparent contradiction. First, in our experience fusion of a large domain to the N-terminus of Sam68ΔC renders it inactive (data not shown). These fusions may interfere with the action of amino acids 15-28 which we have shown are essential for inhibition (Chapter 2). Second, recognition may not be dependent on the RRE alone. Our results suggest that it is the Rev/RRE complex which is
recognized by Sam68ΔC (Figure 4.1-4.3). This complex may be recognized directly by Sam68ΔC, but this is unlikely as we do not see relocalization of Rev to the perinuclear bundles in the presence of Sam68ΔC (Figure 4.4). Therefore, it is more likely that Sam68ΔC is able to indirectly recognize the Rev/RRE complex and that the presence of CBP20/80 at the cap may also be required for susceptibility to Sam68ΔC inhibition (see discussion below).

It has generally been assumed that, despite an alternative export pathway, translation of incompletely spliced HIV-1 mRNAs occurs in the same manner as host mRNAs. Our data shows that eIF4E is only associated with the spliced, cleaved isoform of pgTat mRNA, while CBP80 is associated with all 4 mRNA isoforms, including the unspliced, cleaved (Figure 4.5). This indicates that translation of the unspliced, cleaved isoform of pgTat is eIF4E-independent; translation occurs either via a CBP20/80-dependent mechanism or in cap-independent mechanism (via an internal ribosome entry site (IRES)). Although two IRES have been identified in HIV-1 it is unlikely that they can account for the eIF4E-independent translation seen in these experiments. The first IRES element is localized in 5’UTR of HIV-1 and the second is in the coding region of Gag (349, 350). Neither IRES is present in the pgTat reporter construct used in Figure 4.5, so neither can account for the translation of the unspliced, cleaved pgTat mRNA. Also, activation of the 5’UTR IRES requires Vpr-mediated cell cycle arrest in G2/M phase: pgTat does not express Vpr and the cells do not arrest in G2/M phase (259).

Therefore, translation of the incompletely spliced HIV-1 mRNAs must occur in a CBP20/80-dependent fashion. The nuclear cap binding complex is able to mediate translation initiation, as it is responsible for the pioneer round of translation (32,
There are three compelling pieces of evidence which support this theory, in addition to the fact that CBP80, not eIF4E, is associated with the unspliced, cleaved isoform of pgTat (Figure 4.5). First, we see a redistribution of CBP80 from predominantly nuclear to whole cell in the presence of HIV-1 provirus (Figure 4.6). Second, we show CBP80 and unspliced, cleaved pgTat mRNA are both associated with polysomes in pgTat and Rev transfected cells (Figure 4.7). Third, we show that CBP80 and 9kb HIV-1 mRNA are both associated with polysomes in provirus transfected cells (Figure 4.9). Importantly, upon dissociation of the polysomes with EDTA both the CBP80 and the incompletely spliced HIV-1 mRNAs shift to the lighter fraction (Figure 4.8, 4.10). It should be noted that although each of these polysomal fractionations was performed twice, the EDTA polysome dissociation was not successful the second time. Although the pgTat and proviral polysomal fractionations do agree with one another, another repeat of the experiment would be advantageous. However, taken as a whole, these observations confirm that the HIV-1 mRNA and CBP80 are indeed in the polysomal fraction and not in another large mRNP structure, such as a P-body.

These observations indicate that the mRNP of actively translated, unspliced, cytoplasmic HIV-1 mRNAs is unique as it retains CBP20/80 at the 5’cap. During the pioneer round of translation the CBP20/80 complex is exchanged for eIF4E at the 5’cap (32). Avoiding the exchange of cap binding proteins at the 5’cap may provide HIV-1 with two advantages. First, by using the Crm1 export pathway, incompletely spliced HIV-1 mRNAs evade the pioneer round of translation, thereby avoiding a host mechanism of mRNA surveillance (183-185). Second, eIF4E is available in limiting amounts in the cell
Transformed cells overcome the eIF4E deficit by overexpressing eIF4E; by using CBP20/80-dependent translation HIV-1 achieves the same goal.

How can CBP20/80 and Rev/RRE complex, which are so far apart, be simultaneously recognized by Sam68ΔC? There have been some reports that indicate that Rev may also act as a translational activator (365, 366). This activation may be mediated through other host proteins such as human Rev-interacting protein (hRIP). hRIP was recently shown to be essential for the release of RRE-containing, Crm1 transported HIV-1 mRNAs from the nuclear periphery (439). In the absence of hRIP, these mRNAs are retained at the nuclear periphery and are not translated. DDX3 is a RNA helicase, localized to the cytoplasmic face of the nuclear pore, which is required for expression of RRE/Rev/Crm1 transported mRNAs (340). The data from these labs, along with the data presented in this chapter, are consistent with a model where the Rev/RRE complex, hRIP, DDX3 and the CBP20/80 bound cap interact at the nuclear periphery (directly or indirectly) to initiate translation. Sam68ΔC may interfere with this interaction and prevent binding of PABP, thereby inhibiting translation initiation (Chapter 3, Figure 4.11).

We have discovered that Sam68ΔC is able to use the marks of Crm1 mRNA export (Rev/RRE and CBP20/80) to specifically inhibit translation of Rev-dependent HIV-1 mRNAs. This observation is the first indication that mRNAs using different export pathways have distinct 5’cap compositions, and that these may render mRNAs susceptible to different control mechanisms. Eliciting the details of translation initiation for Rev-dependent HIV-1 mRNAs may lead us to discover further discrepancies from host mRNAs, and lead to new therapeutic approaches.
Figure 4.11. Model of spliced versus unspliced HIV-1 mRNA translation.
Chapter 5: Conclusions and Future Directions

Kim Marsh
5.1. Conclusions

Although nuclear export of HIV-1 9kb and 4kb mRNAs has long been recognized as occurring via a pathway distinct from that used by the bulk of the host mRNAs, translation of these HIV-1 mRNAs was long believed to occur in an identical manner to host mRNAs. In the preceding three chapters, I have shown evidence that the composition of incompletely spliced HIV-1 mRNPs is distinct from the bulk of Tap/p15 transported host mRNPs. We have used sensitivity to Sam68ΔC inhibition as a tool to distinguish between different HIV-1 mRNP compositions.

In this thesis, evidence for a distinct incompletely spliced HIV-1 mRNP and distinct translational regulation of this mRNP has been presented. The evidence falls into three major categories. First, RRE/Rev/Crm1 transported mRNAs are specifically translationally inhibited by Sam68ΔC, but the same mRNAs transported via the Tap/p15 export pathway are not translationally repressed (Figure 2.7, 2.9, 4.1). This experiment was carried out by inserting the CTE, to which Tap/p15 can directly bind, in place of the RRE. However, additional experiments showed that addition of the RRE to a Tap/p15 transported mRNA, did not make the mRNA Sam68ΔC sensitive (Figure 4.3). Together this data indicates that the Sam68ΔC sensitive mRNAs are not inhibited due to specific sequence elements to which Sam68ΔC can bind. As it is not a sequence element that defines the target, the data indicates that there is a distinct property that is targeted.

The second piece of evidence is that Sam68ΔC blocks translation of incompletely spliced HIV-1 mRNAs by blocking PABP1 binding at the poly(A) tail, or stripping PABP1 from the affected mRNAs. Experimental data indicates that the gross amount of
incompletely spliced HIV-1 mRNA available for translation, and the polyadenylation state of this mRNA, was unaffected (Figure 3.5, 3.6). Only PABP1 binding to the HIV-1 incompletely spliced mRNAs was affected by Sam68ΔC (Figure 3.7, 3.9). As PABP1 is an important translation factor, involved in the initiation of translation, this indicates that the HIV-1 incompletely spliced mRNAs must remodel their mRNP in a manner distinct from completely spliced HIV-1 mRNAs or host mRNAs.

The third and final piece of evidence is that HIV-1 incompletely spliced mRNAs retain the nuclear cap binding complex, CBP20/80, during translation. The data shows that CBP80 is bound to both spliced and unspliced forms of pgTat mRNA, but eIF4E is only bound to the completely spliced, cleaved pgTat mRNA (Figure 4.5). Another explanation for the lack of eIF4E binding to incompletely spliced HIV1 mRNAs could be that their translation is IRES-dependent. Two IRES elements have been identified in HIV-1: the first is located in the 5’UTR from nucleotides 104-336 (349), and the second is located in the coding region of Gag (350). This explanation is not likely because the pgTat reporter does not contain either of the identified IRES elements, yet the unspliced pgTat mRNA is CBP80 and not eIF4E-bound. In the presence of HIV-1 provirus, CBP80 relocalizes from predominantly nuclear to a whole cell distribution (Figure 4.6). Also, CBP80 is associated with the polysomal fraction of HIV-1 reporter and HIV-1 provirus transfected cells (Figure 4.7, 4.9). This data provides a physical distinction between the two types of mRNPs being translated: incompletely spliced HIV-1 mRNPs versus completely spliced HIV-1 and host mRNPs. This indicates that incompletely spliced HIV-1 mRNAs are translated via an eIF4E-independent mechanism, and this could provide a target for novel anti-HIV-1 therapies.
There is a growing body of evidence from other labs, which indicates that
translation of incompletely spliced, Rev-transported, HIV-1 mRNAs occurs in a distinct
manner. The primary piece of evidence is Rev’s role not only in transport, but translation
of these HIV-1 mRNAs. Rev promotes the loading of gag-pol and env mRNA on to
polysomes (365). The authors show that although Rev elevates the amount of gag mRNA
in the cytoplasm 8-16 fold, the amount of Gag protein increases by over 800 fold (365).
In addition, experiments using cytoplasmically transcribed env mRNA show that Rev is
still able to enhance translation (366). Importantly, mRNA must contain the RRE for
Rev-induced translational stimulation, but the RRE alone is not sufficient, indicating the
requirement for additional cis-acting elements or trans-acting factors. The data presented
here indicates that one such trans-acting factor might be the CBP20/80 complex at the
5’end.

Other labs have identified additional trans-acting factors that may play a role in
the translation of incompletely spliced HIV-1 mRNAs. Human Rev-interacting protein
(hRIP) was recently shown to be essential for the release of RRE-containing, Crm1
transported, HIV-1 mRNAs from the nuclear periphery (439). In the absence of hRIP,
these mRNAs are retained at the nuclear periphery and are not translated. DDX3 is an
RNA helicase localized to the cytoplasmic face of the nuclear pore, which is required for
expression of RRE/Rev/Crm1 transported mRNAs (340). DDX3 is proposed to function
in analogous manner to hDbp5, the RNA helicase which facilitates unwinding of mRNAs
and removal of Tap/p15 as host mRNAs move from the nuclear pore into the cytoplasm
(128, 130-134). DDX3 and hRIP appear to play a role in the passage of the incompletely
spliced HIV-1 mRNAs from the nuclear pore to the cytoplasm, however it has been
difficult to distinguish this function from effects on translation. The data from other labs, along with the data presented in this thesis, are consistent with a model where the Rev/RRE complex, hRIP, DDX3 and the CBP20/80 bound cap interact at the nuclear periphery (directly or indirectly) to initiate translation of incompletely spliced HIV-1 mRNAs. Sam68ΔC may interfere with this interaction and prevent binding of PABP1, thereby inhibiting translation initiation.

There is a growing amount of evidence to suggest that subsets of mRNAs involved in similar biological processes are co-regulated, and that this is mediated by the proteins bound to elements in their UTRs (reviewed in (143-145)). These subsets of co-regulated mRNAs have been termed regulons. Regulons identified to date generally consist of a group of mRNAs containing a common cis-acting element and regulation via trans-acting factors that bind to the element. For example, one regulon identified consists of a subset of mRNAs containing AU-rich elements (AREs) in their 3’UTRs. The regulon is defined by the presence of an ARE in their 3’UTR, HuR binding to the ARE and co-regulation (138). HuR interacts with pp32 and APRIL, and subsequently Crm1, to mediate export of these mRNAs (138). Another regulon has been defined by the presence of an eIF4E-sensitivity element (4E-SE) in the 3’UTR (139-141). mRNAs containing the 4E-SE are all involved in cell cycle progression; are bound by eIF4E, not CBP20/80, in the nucleus; and are exported by Crm1.

Incompletely spliced HIV-1 mRNAs may represent a novel translational regulon. This regulon could be defined by the presence of the cis-acting element, RRE; RRE/Rev/Crm1 mediated nuclear export; the subset of proteins required for accessing the translational machinery (Rev, CBP20/80, hRIP, DDX3); and sensitivity to translational
inhibition by Sam68ΔC. If incompletely spliced HIV-1 mRNAs do indeed represent a novel translational regulon, this could be a target of future anti-HIV-1 therapeutics. However, to confirm that this is indeed a novel translational regulon a number of further experiments would be required.

5.2. Future Directions

5.2.1. Confirming eIF4E-independence and CBP20/80-dependence of incompletely spliced HIV-1 translation

To confirm that incompletely spliced HIV-1 mRNAs do indeed represent a novel translational regulon a number of further experiments would be required. We have proposed that this regulon is partially defined by the presence of CBP20/80, not eIF4E, at the 5’cap in the cytoplasm and throughout translation. This would imply that translation occurs in an eIF4E-independent manner. There are a number of experiments which could be used to test this hypothesis by exploiting the fact that eIF4E-dependent translation is subject to regulation by eIF4E-binding proteins (4E-BPs) and phosphorylation on serine-209 (reviewed in Chapter 1.4.1).

4E-BP1 binds eIF4E, blocking interaction with eIF4G, and inhibiting translation (211-213). Therefore, overexpression of 4E-BP1 should inhibit eIF4E-dependent translation, and if our hypothesis is correct, should not affect translation of incompletely spliced HIV-1 mRNAs. Hypophosphorylated 4E-BP1 binds very tightly to eIF4E, and dissociates upon phosphorylation (212, 214, 215). Therefore, overexpressed 4E-BP1 may be inactivated by cellular kinases. This could be overcome by overexpressing a mutant of 4E-BP1 which is naturally generated by caspase cleavage in cells undergoing apoptosis (216). Caspase cleavage removes the N-terminal 24 amino acids of 4E-BP1, the
truncated protein is hypophosphorylated and completely shuts off eIF4E-dependent, but not IRES-dependent, translation *in vivo* (216). Overexpression of 4E-BP1 or an N-terminal truncation mutant thereof, should inhibit translation of a control, eIF4E-dependent reporter but not affect translation of an incompletely spliced HIV-1 reporter construct. The control and HIV-1 reporters should be controlled by an inducible promoter to permit accumulation of 4E-BP1, or a mutant thereof, and HIV-1 Rev prior to translation of the reporters.

Two additional eIF4E family members have been identified: eIF4E-2 (4E-HP and 4E-LP) and eIF4E-3 (217, 218). Both family members can bind the 5’ cap, although with less affinity than eIF4E itself (217-219). However, neither can support translation (217-219). The *Drosophila* homologue of 4E-HP (d4E-HP) has been shown to inhibit translation of *caudal* mRNA, but the role of eIF4E-2 and eIF4E-3 in human translation has yet to be determined (220). 4E-HP affinity for the 5’ cap is enhanced by the activity of the ubiquitin-like molecule ISG15 (444). Modification of endogenous 4E-HP with ISG15 (ISGylation) is activated by interferon and genotoxic stress, but can be simulated by fusion of an ISG15-moiety at the N- or C-terminus of 4E-HP (444). To test that expression of incompletely spliced HIV-1 mRNAs is CBP20/80-dependent, we can overexpress 4E-HP and ISG15-4E-HP, and measure the effects on translation. We would expect that 4E-HP and ISG15-4E-HP would inhibit all cap-dependent translation, by competing for cap binding. Okumura et al were able to show that both 4E-HP and ISG15-4E-HP are able to inhibit cap-dependent translation *in vitro* (444). Therefore, translation of both a control eIF4E-dependent reporter and an incompletely spliced HIV-1 reporter
should be inhibited, while translation of a co-transfected IRES-dependent control should not be affected.

Although 4E-HP and ISG15-4E-HP have been shown to inhibit translation \textit{in vitro}, it is possible that they will not measurably affect cap-dependent translation \textit{in vivo}. However, \textit{in vitro} systems will probably not be adequate for the study of incompletely spliced HIV-1 mRNA translation. This is because the composition of the incompletely spliced HIV-1 mRNP in the cytoplasm is moulded by the mRNA export pathway, and it would be very difficult to reconstitute this in an \textit{in vitro} system. According to our hypothesis, translation of incompletely spliced HIV-1 mRNAs occurs in a cap-dependent manner. We have shown preliminary evidence that CBP20/80 is bound to the cap of these mRNAs throughout translation, and that the exchange for eIF4E does not occur. However, this does not mean that in an \textit{in vitro} system, optimized for translation, translation would not occur via eIF4E-dependent means.

5.2.2. \textit{Identification of host mRNAs translated in a CBP20/80-dependent manner}

We have proposed that the incompletely spliced HIV-1 mRNA regulon is defined by the presence of CBP20/80 at the 5’cap, nuclear export via the RRE/Rev/Crm1 pathway, and susceptibility to translational inhibition by Sam68\textDeltaC. If this regulon is exclusive to incompletely spliced HIV-1 mRNAs it makes it an ideal target for anti-HIV-1 therapeutics. However, if some host mRNAs are similarly regulated, it is important to identify them before pursuing this anti-viral strategy.

First it is important to identify whether there are host mRNAs that are translated in a CBP20/80-, not eIF4E-, dependent manner. To identify these mRNAs, we could perform anti-CBP80 or anti-eIF4E pulldowns on whole cell lysates. Alternately, it may
be preferable to use only the cytoplasmic or polysomal fractions for input to the immunoprecipitations. The co-immunoprecipitated mRNAs could then be identified by microarray. Comparison of the mRNAs associated with CBP80 versus eIF4E should identify three populations of mRNAs. First, there may be a subset of mRNAs which are associated with eIF4E and not CBP80. For example, cyclin D1 mRNA contains an eIF4E-sensitivity element and is bound by eIF4E, not CBP20/80, in the nucleus (139, 141). Second, another subset of mRNAs may be associated with both CBP20/80 and eIF4E. One would hypothesize that this would be the bulk of host mRNAs. These mRNAs are CBP20/80-bound in the nucleus and after transport to the cytoplasm, however after the pioneer round of translation would be eIF4E-bound. The third and final subset of mRNAs would be associated with CBP20/80 and not eIF4E. As a positive control, incompletely spliced HIV-1 mRNAs should be transfected into the host cells and should be identified as a part of this population.

The population of mRNAs identified as CBP80-bound, and not eIF4E-bound will be of further interest to us. These host mRNAs may potentially be translated in a CBP20/80-dependent, eIF4E-independent manner. One would presume that there would be a large number of false positives, due to the fact that the amount of eIF4E is limiting, and mRNAs may be present in the cytoplasm but not be actively translated. Mammalian mRNA processing bodies (P-bodies) are the cytoplasmic foci which contain pools of untranslated mRNAs (reviewed in (445)). mRNAs localized to P-bodies may be degraded or return to translation. Evidence shows that mRNAs localized to P-bodies are associated with eIF4E but not eIF4G or ribosomal proteins (446, 447). Therefore, even mRNAs which are not being actively translated should still be immunoprecipitated with eIF4E. To
eliminate false positives, the susceptibility of the identified mRNAs to 4E-BP1 and 4E-HP inhibition can be measured. As explained in section 5.2.1., if the mRNAs are translated in a CBP20/80-dependent manner they will not be affected by 4E-BP1, but will be inhibited by 4E-HP. If translation of any host mRNAs is identified as CBP20/80-dependent it could represent a novel host mRNA regulon.

5.2.3. Analysing anti-HIV-1 therapeutic potential of Sam68ΔC based strategies

The translation of incompletely spliced HIV-1 mRNAs, in addition to being CBP20/80-dependent, is also inhibited by Sam68ΔC. If Sam68ΔC is to be used as a basis for anti-HIV-1 therapeutics, it is important to determine what host mRNA translation will be affected by its activity. If the translation of a subset of host mRNAs is CBP20/80-dependent, these would be the logical mRNAs to initially test for susceptibility to Sam68ΔC inhibition. However, it does not necessarily follow that this host regulon is the same as the incompletely spliced HIV-1 mRNA translational regulon. Sam68ΔC sensitivity would most easily be assessed by co-transfection of the host proteins of interest and Sam68ΔC, followed by western blots to measure effects on translation. If there are no antibodies to the endogenous proteins available, the proteins could be Flag- or HA-tagged, and the effects of Sam68ΔC on their expression measured.

It is possible that no host mRNAs will be identified as being translated in a CBP20/80-dependent manner. Although these mRNAs would be more likely to be inhibited by Sam68ΔC, it will be important to assess the effects of Sam68ΔC on global host protein expression. We have already shown that Tap/p15 mRNA expression is not generally inhibited, however a subset of mRNAs may be affected. We have shown that Sam68ΔC does not alter the mRNA abundance or polyadenylation status of a target
mRNA, therefore basic microarray based strategies would not be useful. However, Sam68ΔC inhibits PABP1 binding to a target mRNA. Therefore to assess the global affect of Sam68ΔC on host translation, PABP1 immunoprecipitations in the presence and absence of Sam68ΔC could be performed. The immunoprecipitated mRNAs could then be analysed by microarray. Those host mRNAs which are depleted in the immunoprecipitates from Sam68ΔC-transfected cells are likely candidates for Sam68ΔC translational inhibition. As controls, incompletely spliced HIVR1 mRNAs should be depleted in the PABP1 immunoprecipitates from Sam68ΔC transfected cells and actin mRNA should be unaffected (see Figures 3.7, 3.9). Any hits generated in this manner would be verified by co-transfection of the proteins of interest and Sam68ΔC, followed by western blots to measure effects on translation.

If there are host mRNAs translationally inhibited by Sam68ΔC, it may mean this anti-HIV-1 strategy is not viable as a therapeutic. Inhibition of host mRNA translation may be an indication that Sam68ΔC has off-target effects which would have deleterious effects if introduced into a patient. However, the targeting of the incompletely spliced HIV-1 regulon, in a Sam68ΔC-independent manner, may still be a viable option for the development of anti-HIV-1 therapeutics. These strategies will depend on the specific inhibition of CBP20/80-, and RRE/Rev/Crm1-dependent translation.

5.2.4. A deeper look at translation in the HIV-1 transfected cell

As reviewed in Chapter 1.5.4., HIV-1 mRNA translation is entirely dependent on the translation machinery of the host cell. For HIV-1 to replicate successfully, it must usurp the host translation machinery such that the viral mRNA is translated in large quantity. However, this must be balanced with the requirement for host cell viability,
which means that host cellular mRNAs must continue to be translated. There is evidence that HIV-1 translation can occur through a cap-independent, IRES-dependent mechanism. Two IRES elements, which are active during G2/M phase of the cell cycle have been identified in HIV-1 (349, 350). Additionally, it has been discovered that the HIV-1 protease is able to cleave PABP1 and eIF4GI but not eIF4GII (367-370). In this thesis, evidence has been presented that the translation of the incompletely spliced HIV-1 mRNAs is regulated through a CBP20/80-containing mRNP. It would be interesting to explore how the IRES- and CBP20/80-dependent modes of translation are affected by protease cleavage of PABP1 and eIF4GI. This research might elucidate which modes of translation are used at different stages in the viral life cycle.

Cleavage of PABP1 and eIF4GI by HIV-1 protease decreases cap-dependent translation 2-10 fold, but does not result in a complete shutoff of cap-dependent translation (369, 370). As previously stated, it would be difficult to manipulate in vitro translation studies on incompletely spliced HIV-1 mRNAs, due to a dependence on Rev-mediated nuclear export. Therefore, an in vivo study using Rev-responsive, incompletely spliced HIV-1 reporters should be used. Reporters with and without the HIV-1 IRES should be used. As controls, an eIF4E-dependent reporter and an HIV-1 reporter with a non-cell cycle dependent IRES (such as poliovirus 2A) should be used. Using RNAi we can specifically knock-down PABP1 and/or eIF4GI and measure the effects on translation of the various reporter constructs. Knock-down of eIF4GI should not completely abrogate translation as the other eIF4G isoform, eIF4GII, could mediate cap-dependent translation (369, 370, 448). There is evidence that cleavage of eIF4GI inhibits translation of newly synthesized mRNAs, but has no effect on mRNAs already engaged
with the translational machinery (448). The authors propose that intact eIF4GI is required for the pioneer round of translation. Knock-down of PABP1 would probably completely inhibit cap- and IRES-dependent translation (434).

The HIV-1 protease does not cause degradation of eIF4GI and PABP1, there are stable cleavage products which probably retain some activity (368, 370). Therefore, RNAi knockdown of PABP1 should be combined with overexpression of the 3 PABP1 fragments. Similarly, RNAi knockdown of eIF4GI should be combined with overexpression of the 3 major eIF4GI fragments. Measuring the effects of these knockdowns of the various control and HIV-1 reporter constructs may give us an idea of the temporal control of HIV-1 translation throughout the virus life cycle. For instance, if HIV-1 proteolytic cleavage of PABP1 and eIF4GI specifically inhibits eIF4E-dependent translation and not CBP20/80-dependent translation it may be a means through which HIV-1 mRNAs are preferentially translated. If cleavage of PABP1 and eIF4GI inhibits all cap-dependent translation, but not IRES-dependent translation, it may indicate that this cleavage occurs only at the very late stages of viral infection.

In summary, the data in this thesis indicates that incompletely spliced HIV-1 mRNAs may belong to distinct translational regulon. This regulon is defined by RRE/Rev/Crm1 nuclear export, CBP20/80-dependent translation and sensitivity to Sam68ΔC inhibition. Data from other labs indicate that this regulon is also regulated by the trans-acting factors hRIP and DDX3 (340, 439). This regulon could become a target for new anti-HIV-1 therapeutic approaches. Four sets of experiments which arise from these results have been outlined in this chapter. First, the eIF4E-independence and
CBP20/80-dependence of incompletely spliced HIV-1 mRNA translation must be confirmed. Second, it is important to identify whether or not any host mRNAs are translated in a CBP20/80-dependent manner. This is interesting on two fronts: it could reveal a novel host regulon, and if host mRNAs are regulated in this manner it may make this HIV-1 regulon a less attractive target for anti-HIV-1 therapeutics. Third, the anti-HIV-1 therapeutic potential of Sam68ΔC based strategies should be assessed. Fourth, it would be interesting to know how CBP20/80-dependent HIV-1 translation fits in with HIV-1 IRES-dependent translation and HIV-1 protease cleavage of the eIF4GI and PABP1 translation factors.
References


89. Fong N, Bentley DL. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: Different functions for different segments of the CTD. Genes Dev. 2001 Jul 15;15(14):1783-95.


150. Gradi A, Svitkin YV, Imataka H, Sonenberg N. Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host


159. Demeshkina N, Repkova M, Vemyaminova A, Graifer D, Karpova G. Nucleotides of 18S rRNA surrounding mRNA codons at the human ribosomal A, P, and E sites: A crosslinking study with mRNA analogs carrying an aryl azide group at either the uracil or the guanine residue. RNA. 2000 Dec;6(12):1727-36.


270. O'Reilly MM, McNally MT, Beemon KL. Two strong 5' splice sites and competing, suboptimal 3' splice sites involved in alternative splicing of human immunodeficiency virus type 1 RNA. Virology. 1995 Nov 10;213(2):373-85.


286. Ashe MP, Pearson LH, Proudfoot NJ. The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. EMBO J. 1997 Sep 15;16(18):5752-63.


381. Poole E, Strappe P, Mok HP, Hicks R, Lever AM. HIV-1 gag-RNA interaction occurs at a perinuclear/centrosomal site; analysis by confocal microscopy and FRET. Traffic. 2005 Sep;6(9):741-55.


386. Lukong KE, Richard S. Sam68, the KH domain-containing superSTAR. Biochim Biophys Acta. 2003 Dec 5;1653(2):73-86.


