INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313-761-4700  800-521-0600
CHARACTERIZATION OF THE S1, S2, S3, AND DU OPEN READING FRAMES OF EQUINE INFECTIOUS ANEMIA VIRUS

by

WENDY K. STEAGALL

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Thesis Advisor: Dr. Susan L. Payne

Department of Molecular Biology and Microbiology
CASE WESTERN RESERVE UNIVERSITY
January, 1995
CASE WESTERN RESERVE UNIVERSITY

GRADUATE STUDIES

We hereby approve the thesis of

WENDY K. STEAGALL

candidate for the PHD

degree.*

(signed)  
(chair)  

(date) 10-21-94  

*We also certify that written approval has been obtained for any proprietary material contained therein.
Copyright © (1995) by

Wendy K. Steagall
CHARACTERIZATION OF THE S1, S2, S3, AND DU OPEN READING FRAMES OF EQUINE INFECTIOUS ANEMIA VIRUS

Abstract

by

WENDY K. STEAGALL

Equine Infectious Anemia Virus (EIAV) contains three small open reading frames, designated S1, S2, and S3, that may encode regulatory proteins. To determine the functions of these open reading frames, cDNA clones were isolated and characterized. Three classes of cDNA clones were obtained. The first class contains the leader region, S1, and S3, while the second class contains the leader region, S1, part of env and S3. The leader-S1 reading frame encodes Tat activity, while the env-S3 region encodes Rev activity. The third class of cDNA contains part of the gag gene fused to part of the env gene, and its function is undetermined. No cDNA clone was isolated that primarily encodes S2, and the function of S2 is unknown.

EIAV also encodes a dUTPase from the DU domain present in the pol gene. The inability of EIAV containing a 270 bp deletion in the DU domain (ΔDU) to replicate to wild type (WT) levels in primary equine macrophages has been described (Threadgill et al., 1993). In this dissertation, I describe the construction of a second dUTPase-deficient virus containing a single amino acid substitution in dUTPase (DUD71E). This virus showed a similar growth phenotype to ΔDU in macrophages, replicating to only 2% of WT levels. The blocks to viral replication of ΔDU and DUD71E in macrophages were also examined. Reverse transcription occurred normally in macrophages with full length viral DNA evident by 72 hours post-infection; however, the ΔDU EIAV incorporated uracil into the viral DNA, as shown by in vitro and in vivo assays, while WT EIAV did not. The viral DNA from the dUTPase-deficient viruses integrated into macrophage DNA, but at
levels 2- to 3-fold less than WT levels. Steady state levels of viral transcripts were severely decreased
in macrophages infected with dUTPase-deficient viruses as compared to WT EIAV. These results
suggest that the major block to dUTPase-deficient viral replication in macrophages occurs between
integration and viral transcription. A virally encoded dUTPase may be necessary to prevent
incorporation of uracil into viral DNA in nondividing macrophages, which may have low levels of this
enzyme.
This work is dedicated to D² - my models for excellence in research and life.
ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Susan Payne for her advice and guidance throughout the duration of this project. Her interest and insight in this endeavor are appreciated. I would also like to thank the members of my committee, Dr. Hsing-Jien Kung, Dr. Timothy Nilsen, Dr. Stuart LeGrice, and Dr. Pieter deHaseith, for their comments, interest, and advice over the years. Thanks also go to Phil Weiland, Debbie Threadgill, Karen Nelson, Jill Beam, and Kevin Kiger for their friendship in and out of the Molecular Biology department. Their comradeship and lively discussions of work and outside interests provided welcome relief when experiments failed to work. I would also like to thank my parents for their encouragement and support, both emotionally and financially, as I completed this project. Finally, I would like to thank my husband, Bob, and my daughter, Erika, for their love, laughter, and patience through the years of my graduate student career.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER ONE. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>History and Significance</td>
<td>1</td>
</tr>
<tr>
<td>Equine Infectious Anemia</td>
<td>2</td>
</tr>
<tr>
<td>Virus Structure and Genome</td>
<td>4</td>
</tr>
<tr>
<td>CHAPTER TWO. CHARACTERIZATION OF OPEN READING FRAMES S1, S2, AND S3</td>
<td>8</td>
</tr>
<tr>
<td>Introduction</td>
<td>8</td>
</tr>
<tr>
<td>EIAV Genome</td>
<td>8</td>
</tr>
<tr>
<td>Transcription Patterns</td>
<td>10</td>
</tr>
<tr>
<td>HIV-1</td>
<td>11</td>
</tr>
<tr>
<td>Tat</td>
<td>11</td>
</tr>
<tr>
<td>Vpu</td>
<td>14</td>
</tr>
<tr>
<td>Rev</td>
<td>16</td>
</tr>
<tr>
<td>Proposed Functions of S1, S2, and S3 Open Reading Frames of EIAV</td>
<td>19</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>19</td>
</tr>
<tr>
<td>Cells and Viruses</td>
<td>19</td>
</tr>
<tr>
<td>RNA Preparation</td>
<td>20</td>
</tr>
<tr>
<td>Northern Blot Analysis</td>
<td>20</td>
</tr>
<tr>
<td>PCR-Generated Probe</td>
<td>21</td>
</tr>
<tr>
<td>cDNA Synthesis</td>
<td>21</td>
</tr>
<tr>
<td>Southern Blot Analysis</td>
<td>22</td>
</tr>
<tr>
<td>cDNA Clones</td>
<td>22</td>
</tr>
<tr>
<td>RNase Protection Assays</td>
<td>25</td>
</tr>
<tr>
<td>Eukaryotic Expression Vectors</td>
<td>26</td>
</tr>
<tr>
<td>Bacterial Expression Vectors</td>
<td>26</td>
</tr>
<tr>
<td>Protein Expression</td>
<td>26</td>
</tr>
<tr>
<td>Transfections</td>
<td>27</td>
</tr>
<tr>
<td>CAT Assays</td>
<td>27</td>
</tr>
<tr>
<td>Stable S3 Cell Lines</td>
<td>27</td>
</tr>
<tr>
<td>Mutagenesis of the S3 Open Reading Frame</td>
<td>28</td>
</tr>
<tr>
<td>Reverse Transcriptase Assays</td>
<td>29</td>
</tr>
<tr>
<td>Results</td>
<td>31</td>
</tr>
<tr>
<td>Detection of Multiply-spliced Transcripts in EIAV-infected Cells</td>
<td>31</td>
</tr>
<tr>
<td>Temporal Analysis of RNA Production</td>
<td>33</td>
</tr>
<tr>
<td>Production of cDNA Clones and Southern Blot Analysis</td>
<td>35</td>
</tr>
<tr>
<td>Characterization of cDNA Clones</td>
<td>38</td>
</tr>
<tr>
<td>RNase Protection Analysis</td>
<td>49</td>
</tr>
<tr>
<td>Analysis of the Splicing Event of the env Transcript</td>
<td>53</td>
</tr>
<tr>
<td>Analysis of the Leader-S1 Open Reading Frame</td>
<td>55</td>
</tr>
<tr>
<td>Analysis of the S2 Open Reading Frame</td>
<td>60</td>
</tr>
<tr>
<td>Analysis of the env-S3 Open Reading Frame</td>
<td>62</td>
</tr>
</tbody>
</table>
### CHAPTER THREE. INCORPORATION OF URACIL INTO VIRAL DNA CORRELATES WITH REDUCED REPLICATION OF EIAV IN MACROPHAGES

**Introduction**
- Replication of Lentiviruses in Nondividing Cells
- dUTPase
- Viruses and dUTPase
- EIAV dUTPase

**Materials and Methods**
- Cells and Viruses
- Methotrexate Studies
- Mutagenesis
- dUTPase Assays
- PCR Analysis of the Production of Viral DNA
- Reverse Transcriptase Assays
- Uracil-DNA Glycosylase Assays
- Southern Blot Analysis
- RNA Preparation and Northern Blot Analysis

**Results**
- DUD71E Virus
- Production of Viral DNA
- Detection of Uracil in Viral DNA
- Integration
- Transcription
- Methotrexate Studies

**Discussion**
- DUD71E Virus
- Production of Viral DNA
- Utilization of dUTP
- Integration
- Transcription
- Methotrexate Studies
- Role of dUTPase in EIAV Replication
- Future Studies

**Summary**

**REFERENCES**
<p>| Figure 1.1 | Representation of the components of the EIAV virion | 6 |
| Figure 1.2 | Schematic representation of the EIAV genome | 7 |
| Figure 2.1 | Schematic representation of the genome of EIAV and the putative amino acid sequences of the S1, S2, and S3 open reading frames | 9 |
| Figure 2.2 | Schematic representation of the genome of HIV-1 | 12 |
| Figure 2.3 | Diagram of the stages of HIV-1 replication | 18 |
| Figure 2.4 | Synthesis of a single-stranded DNA probe that detects all EIAV RNA | 24 |
| Figure 2.5 | Mutagenesis of the S3 open reading frame without affecting the overlapping env reading frame | 30 |
| Figure 2.6 | Northern blot analysis of RNA from EIAV-infected cells | 32 |
| Figure 2.7 | Northern blot analysis of the temporal pattern of RNA production in infected FEA cells | 34 |
| Figure 2.8 | PCR amplification and Southern blot analysis of cDNAs from EIAV-infected FEK cells | 36 |
| Figure 2.9 | Schematic representation of the three classes of cDNA clones isolated from infected FEK cells | 39 |
| Figure 2.10 | Nucleotide and predicted amino acid sequence of pWS1 | 40 |
| Figure 2.11 | Nucleotide and predicted amino acid sequence of pWS2 | 42 |
| Figure 2.12 | Nucleotide and predicted amino acid sequence of pWS3 | 45 |
| Figure 2.13 | Nucleotide and predicted amino acid sequence of pWS4 | 48 |
| Figure 2.14 | RNase T2 protection analysis of pWS1 and pWS2 | 50 |
| Figure 2.15 | Schematic representation of the cDNA clone obtained from infected FEK cells to examine the splicing event used to generate the env mRNA | 54 |
| Figure 2.16 | Amino acid sequences of HIV-1 and EIAV Tat proteins | 56 |
| Figure 2.17 | Analysis of Tat activity of pSV-tat | 58 |
| Figure 2.18 | Induction of pHS2 | 61 |
| Figure 2.19 | Amino acid sequences of EIAV, visna virus, and HIV-1 Rev proteins | 63 |
| Figure 2.20 | Schematic representation of the mRNAs encoding visna virus Rev and EIAV putative Rev proteins | 64 |
| Figure 2.21 | Growth curve analysis of p19S3stop versus pSPEIAV19 | 65 |
| Figure 2.22 | Activation domains of HIV-1, visna virus, FIV, and EIAV Rev proteins | 78 |
| Figure 3.1 | Pathway of dTTP biosynthesis | 80 |
| Figure 3.2 | Pathway of DNA repair | 81 |
| Figure 3.3 | Alignment of amino acid sequences of dUTPases from man, E. coli, visna virus, and EIAV | 86 |
| Figure 3.4 | Schematic representation of the genome of EIAV | 88 |
| Figure 3.5 | Putative amino acid sequence of EIAV dUTPase. with motifs 1-5 as defined by McGeoch indicated | 90 |
| Figure 3.6 | Growth of ΔDU, DUD71E, and WT EIAV in macrophages | 100 |
| Figure 3.7 | Analysis of viral DNA synthesis by ΔDU and WT EIAV in macrophages and FEA cells | 104 |
| Figure 3.8 | Exogenous and endogenous reverse transcriptase assays | 107 |
| Figure 3.9 | Uracil-DNA glycosylase analysis of viral DNA from macrophages and FEA cells | 110 |
| Figure 3.10 | Southern blot analysis of viral DNA from macrophages and FEA cells | 112 |
| Figure 3.11 | Northern blot analysis of RNA from infected macrophages or FEA cells | 114 |
| Figure 3.12 | Schematic representation of the effect of methotrexate | 116 |
| Figure 3.13 | Growth of ΔDU and WT viruses in FEA cells treated with methotrexate | 117 |</p>
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Oligonucleotide Sequences</td>
<td>23</td>
</tr>
<tr>
<td>2.2</td>
<td>Splice Sites</td>
<td>72</td>
</tr>
<tr>
<td>2.3</td>
<td>Translation Initiation Sites</td>
<td>74</td>
</tr>
<tr>
<td>3.1</td>
<td>Oligonucleotide Sequences</td>
<td>94</td>
</tr>
<tr>
<td>3.2</td>
<td>Viral dUTPase Activity</td>
<td>99</td>
</tr>
</tbody>
</table>
CHAPTER ONE
INTRODUCTION

History and Significance Equine Infectious Anemia (EIA) is one of the most important infectious diseases of horses worldwide, affecting agriculture, sporting, and the military (Narayan and Clements, 1990). EIA was described in 1843 in the first account of what was later shown to be a naturally occurring lentiviral disease (Ligneau, 1843). In the early 1900s, Vallee and Carre assigned a filterable etiology to the disease and showed that transmission occurred with the transfer of blood (Vallee and Carre, 1904). The focus of EIA research in the 1950s and 1960s centered on diagnosis and control of the disease, and interest waned in the 1970s after the establishment of control programs and the development of a serodiagnostic test (Coggins et al., 1972; Coggins, 1986; Issel et al., 1986; reviewed in Carpenter and Alexandersen, 1992). In 1976, Charman classified the equine infectious anemia virus (EIAV) as a member of the retrovirus family (Charman et al., 1976), and evidence obtained in the 1980s placed EIAV in the lentivirus subfamily of retroviruses (Montelaro et al., 1988b; Stephens et al., 1986; Weiland et al., 1977; Yaniv et al., 1986), which also includes human immunodeficiency viruses 1 and 2 (HIV-1, HIV-2), visna virus, feline immunodeficiency virus (FIV), and caprine arthritis encephalitis virus (CAEV). This subfamily is characterized by a propensity for antigenic variation, high levels of extrachromosomal DNA in infected cells, establishment of persistent, lifelong infections, replication in cells of the immune system, a genomic organization distinct from other retroviruses, and complex mechanisms of gene regulation. Remarkably, the study of EIA and EIAV has pioneered several accomplishments in the characterization of animal viruses, including “the first assignment of a viral etiology to an animal disease, the development of the first commercial retrovirus diagnostic assay, the first demonstration of retrovirus transmission by insect vectors, the first description of antigenic variation during persistent retrovirus infections, and some of the earliest vaccine trials for preventing a retrovirus infection” (Montelaro, 1989).
In the 1980s, with the classification of EIAV as a lentivirus, interest in the virus grew again as the study of HIV and acquired immunodeficiency syndrome (AIDS) became more pressing. Unlike the slow progressive course of disease caused by most lentiviruses, EIA is a dynamic disease, with a rapid onset of clinical symptoms eventually resolving into inapparent infection. EIA is the only lentiviral disease in which the infected animal routinely brings virus replication under immunologic control despite viral antigenic variation. EIA and EIAV therefore represent a unique system for the investigation of the mechanisms of lentiviral pathogenesis, and although the disease and virus have been recognized for decades, many steps in the pathogenesis of EIAV have not been elucidated. In addition, the genome of EIAV is the smallest in size and perhaps the simplest in structure of the lentiviruses, which may simplify dissection of the functions of its gene products.

**Equine Infectious Anemia** Equine infectious anemia (EIA), also known as swamp fever, affects members of the family Equidae (Issel and Coggins, 1979). Acute EIA generally occurs 7 to 30 days after exposure to the virus and is characterized by fever and hemorrhages. There is massive replication of virus in, and destruction of, macrophages, and there may be greater than $10^9$ horse infective doses (HID/ml of virus in the whole blood and serum (Issel and Coggins, 1979). If the animal survives the acute stage of the disease, it may enter the chronic stage, which is characterized by cyclical episodes of weight loss, fever, anemia, and edema. The recurring febrile cycles initially take place every several weeks, and the frequency and severity of episodes gradually decline in time until the animal is asymptomatic. During the attacks of illness in the chronic phase, there may be greater than $10^3$ tissue culture infectious doses/ml in the serum; however, virus is undetectable in the quiescent periods (Issel and Coggins, 1979). The asymptomatic animal remains a carrier of the disease with at least 1 HID/ml in whole blood, and an episode of illness may be sparked by environmental stresses or by the administration of immunosuppressant drugs (Issel and Coggins, 1979). There is no cure for EIA.

EIAV is transmitted primarily by the interrupted feeding of hematophagous arthropods, such as the horse and deer flies of the family Tabanidae (Issel and Coggins, 1979). This is a purely mechanical transmission, as EIAV does not replicate in the flies. Since horse and deer flies cause
painful bites, they are more likely to be interrupted in their feeding and to carry infected blood in their mouthparts to a naive animal. The levels of EIAV in the blood of infected animals are generally proportional to the severity of the clinical disease, making transmission much more likely with the high viremia seen in the acute stage of EIA. Coggins has shown transmission of EIAV via a single horsefly interrupted from feeding on an animal with an acute case of EIA (Coggins, 1986). EIAV can also be transmitted via transplacental transfer or mechanically via contaminated medical paraphernalia. As there is no cure for EIA, the disease is controlled by the interruption of transmission: i.e. infected animals are isolated or euthanized; methods are employed to control insects.

While vaccines for EIA have been developed, they have met with limited success due to the antigenic variation of the virus. Each clinical episode of the disease is characterized by a variant of EIAV with new or altered antigenic determinants, such that plasma isolated from an infected animal can neutralize virus isolates recovered during earlier febrile episodes, while failing to neutralize those recovered during subsequent episodes (Kono, 1973; Montelaro et al., 1984; Montelaro et al., 1988a; Payne et al., 1987a; Payne et al., 1987b; Salinovich et al., 1986). Eventually, after a number of exposures to different variants, the immune system of the animal is able to bring the virus under control. Therefore, while an inactivated whole virus vaccine may confer protection against a homologous variant of virus, it does not protect against heterologous variants, although it may reduce virus replication and development of disease upon challenge with a heterologous variant (Issel et al., 1992).

EIAV replicates primarily in tissue macrophages and has been found in the liver, kidney, spleen, lymph nodes, and bone marrow (Kim and Casey, 1992; Rice et al., 1989; Sellon et al., 1992). It causes cytopathic effects in macrophages due to uncontrolled replication leading to a large number of budding particles per cell (Klevjer-Anderson et al., 1979). The study of EIAV has been hampered by the relatively few cell lines in which it will replicate in vitro and by the difficulty of obtaining sufficient numbers of primary equine leukocytes (in which the virus will replicate) on which to conduct studies. Due to the high frequency of infection with equine herpesvirus-2, it is difficult to maintain noninfected
donors for adequate leukocyte supply (Orrego et al., 1982). Furthermore, virus obtained from infected animals, such as the Wyoming strain, does not replicate readily in cultures other than the primary equine leukocyte cultures, in which the virus causes extensive cytopathic effects. In 1973, Malmquist et al. developed a strain of EIAV, designated prototype, that would replicate in equine fibroblasts and canine and feline cell lines by serial passage of the Wyoming strain first through equine macrophages, then through fetal equine spleen cells, and finally onto fetal equine dermal cultures (Malmquist et al., 1973). This prototype virus, however, could no longer cause disease in horses, although serial back-passage of the virus through horses did result in increased virulence and disease (Orrego et al., 1982). Therefore, there are the different selective pressures on the virus in vivo versus in vitro, with the result that strains of EIAV that replicate well in vitro tend not to be pathogenic in vivo. Because of these difficulties, a virulent molecular clone of EIAV has yet to be isolated, although infectious clones have been isolated and characterized (Payne et al., 1994; Whetter et al., 1990).

**Virus Structure and Genome** EIAV is spherical with a 110-120 nm diameter and a conical or tubular shaped core that is off-center in the virion (McGuire and Crawford, 1979, Weiland et al., 1977) (Figure 1.1). The enveloped virion is comprised of several hundred copies each of glycoproteins SU (or gp90; the surface glycoprotein) and TM (gp45; the transmembrane glycoprotein) embedded in the lipid envelope: 3000 copies of MA (p15; the matrix protein); several thousand copies each of CA (p26; the capsid or core protein), NC (p11; the nucleocapsid protein), and p9 (a protein of unknown function that is present in the virion core with CA); 5-10 copies of reverse transcriptase: integrase: protease: dUTPase; and an RNA dimer comprising the EIAV genome (Montelaro et al., 1982; Montelaro et al., 1988a).

The genome of EIAV (Figure 1.2) is 8.2 kb, smaller than the genome of other lentiviruses by about 1 kb, and contains the primary retroviral genes *gag*, *pol*, and *env*. The *gag* and *pol* reading frames overlap, while the *env* reading frame is non-overlapping. EIAV also contains three small open reading frames that have been designated S1, S2, and S3 (Kawakami et al., 1987; Rushlow et al., 1986), whose functions were unknown when this work began. There is also an open reading frame
between reverse transcriptase and integrase in the *pol* gene, which had been described as either a pseudoprotease or a dUTPase (Elder *et al.*, 1992; McClure *et al.*, 1987; McGeoch, 1990). Again, when this work started, the function of this open reading frame was unknown. This dissertation will discuss the characterization of these open reading frames and the work done to determine their role in the viral life cycle.
Figure 1.1. Representation of the components of the EIA virion: SU, the surface glycoprotein; TM, the transmembrane glycoprotein; MA, the matrix protein; CA, the capsid protein; NC, the nucleocapsid protein; p9, a protein of unknown function; RT, reverse transcriptase; and the RNA genome. The diagram is adapted from Montelaro et al., 1993.
FIGURE 1.2. Schematic representation of the EIAV genome. The *gag* gene encodes the matrix protein (MA), the capsid protein (CA), the nucleocapsid protein (NC), and p9. The *pol* gene includes protease (PR), reverse transcriptase (RT), the pseudoprotease/dUTPase region (DU), and integrase (IN). The *env* gene encodes the surface glycoprotein (SU) and the transmembrane glycoprotein (TM). There are also three small open reading frames designated S1, S2, and S3.
CHAPTER TWO
CHARACTERIZATION OF OPEN READING FRAMES S1, S2, AND S3

INTRODUCTION

EIAV Genome  Lentiviruses can be distinguished from oncogenic retroviruses by their genomic organization and complex mechanisms of gene regulation: in addition to the characteristic retroviral \textit{gag}, \textit{pol}, and \textit{env} genes, lentiviruses contain open reading frames that encode regulatory proteins. The genome of EIAV is 8.2 kb, making it about 1 kb smaller than that of other lentiviruses, and is thought to be genetically less complex than that of other lentiviruses. The EIAV genome is the smallest lentiviral genome examined to date, and the size differential is due primarily to the short U3 region of the long terminal repeat (LTR) (about 188 bases; 200-250 bases shorter than the U3 region of primate lentiviruses) and to the smaller \textit{pol-env} intergenic region in EIAV (about 190 bases; other lentiviruses contain about 1 kb in this region) (Montelaro et al., 1993). As the \textit{pol-env} intergenic region is the site of many of the accessory genes found in lentiviruses, it is not surprising that the genome of EIAV may be less complex than that of other lentiviruses. Only three extra open reading frames have been recognized in EIAV to date, while at least seven accessory genes in human immunodeficiency virus 1 (HIV-1) have been identified (Cullen and Greene, 1990; Furtado et al., 1991; Kawakami et al., 1987; Rushlow et al., 1986).

The three open reading frames in EIAV have been designated S1, S2, and S3 (Figure 2.1). The S1 open reading frame lies completely in the \textit{pol-env} intergenic region; the S2 open reading frame begins in this intergenic region and ends overlapping the 5’ end of the \textit{pol} open reading frame, and the S3 open reading frame lies at the 3’ end of the \textit{env} gene, overlapping the transmembrane protein coding region. The S1 open reading frame is in phase with \textit{pol} and \textit{env} and could encode a protein of 50 amino acids, although there is no AUG in the reading frame to provide an initiation site for translation. The S2 reading frame could encode a 66 amino acid protein and contains an AUG as the second codon. The S3 reading frame could encode a protein of 135 amino acids, and while it contains no AUG for
<table>
<thead>
<tr>
<th></th>
<th>Kg</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**S1**

VLQLQEARPNYNHCQLCFLRSLGIDYLDASLRKKNKQRRLKAIIQQGGRQPQYL

**S2**

YMGLFGKVTSASASHSMGSQGESQPLLPNSQKNLSTVRQCFNLIVIMTVRTAWNRRKQETKK

**S3**

DPQGPLTEDQWCRVLQRQLPEEKIQRSTCIARRHLGPGPTQHTPSRRDRWIRGQLQAEVLQERLERWRIRGVQQAAKELGEVNRIGIWRELYFREDQRGDFSADGGYQRAQERLWGEQSSPRVLRPGDSSRRRKHL

Figure 2.1. Schematic representation of the genome of EIAV and the putative amino acid sequences of the S1, S2, and S3 open reading frames. Arrows represent potential splice donor or acceptor sites.
initiation of translation, there is a consensus splice acceptor site at the 5' end of the reading frame that could be utilized to splice on an AUG from an upstream region (Kawakami et al., 1987; Rushlow et al., 1986). These open reading frames probably encode proteins produced during EIAV as sera collected from infected horses react with recombinant proteins or peptides based on the sequences of these open reading frames (S. L. Payne, personal communication; Montelaro et al., 1993; Saman et al., 1990; Schiltz et al., 1992).

**Transcription Patterns** Lentiviruses produce three classes of alternatively spliced mRNAs in order to accomplish the differential gene expression necessary to encode multiple gene products from a single primary transcript. These classes include the full-length genomic RNA which encodes the Gag and Pol structural proteins and enzymes, the singly-spliced RNAs which encode the Env structural proteins and some of the accessory proteins, and the multiply-spliced transcripts which encode accessory proteins. HIV-1 has at least four splice donor sites and six splice acceptor sites, and an example of a transcript that contains each gene as the first open reading frame for translation has been found for all the genes except *pol*, whose precursor protein is translated through a ribosomal frameshift, and *env*, whose precursor protein is encoded by a bicistronic message that also contains the accessory protein, Vpu (reviewed in Cullen and Greene, 1990). There are also examples of different species of mRNA encoding the same protein: three transcripts encoding Tat, six encoding Rev, and three encoding Nef have been found in the HIV-1 system; the transcripts differ in splice sites utilized and noncoding exons included (Schwartz et al., 1990a).

Although the EIAV genome contains potential splice donor and splice acceptor sites (Figure 2.1) (Kawakami et al., 1987; Rushlow et al., 1986; Stephens et al., 1986), prior to 1990, it was not clear whether multiply-spliced transcripts were produced in EIAV infections. Rasty et al. examined the transcriptional patterns of EIAV in fetal equine kidney (FEK) cells and fetal donkey dermal (FDD) cells and detected a low abundance 1.5 kb transcript only in the infected FDD cells (Rasty et al., 1990). Furthermore, Northern blot analysis of RNA from infected equine kidney cells, feline cells, canine thymus cells, and primary horse macrophages revealed 1.5 kb transcripts only in the canine cells, and
these were at very low abundance (Stephens et al., 1990). However, there was evidence for EIAV regulatory genes, including a transactivating activity detected in EIAV-infected cells (Derse et al., 1987), and we decided to examine these open reading frames more closely to determine if they encoded regulatory proteins.

**HIV-1** Before continuing the discussion of EIAV accessory proteins, I will first present a review of the relevant HIV-1 proteins. HIV-1, by necessity, is the best studied lentivirus, and as such, provides a model for our studies of EIAV gene regulation. Seven accessory genes have been identified in the HIV-1 genome: tat, rev, nef, vpu, vif, vpr, and tev (Figure 2.2). Tat, rev, nef, and tev are encoded on multiply-spliced transcripts; vpu, vif, and vpr are encoded on singly-spliced transcripts (Arrigo et al., 1989; Cullen and Greene, 1990; Furtado et al., 1991). The Tat and Rev proteins are essential for viral replication, while the Nef, Vpu, Vif, Vpr, and Tev proteins are dispensable for viral replication in vitro. I will review the properties of the three HIV-1 proteins that may have their counterpart in EIAV: Tat, Vpu, and Rev.

**Tat** Tat is an essential protein encoded by two exons that yield an 86 amino acid protein that has been localized to the nucleus (Dayton et al., 1986; Fisher et al., 1986; Hauber et al., 1987; Sharp and Marciniak, 1989). The first exon can yield a fully active 72 amino acid form of Tat. There are three domains in the Tat protein: a cysteine-rich domain responsible for the formation of metal-linked dimers (Frankel et al., 1988), a basic region rich in arginine residues responsible for the nuclear localization and RNA binding properties of the protein, and an activation domain responsible for the function of the protein (Green and Loewenstein, 1988; Green et al., 1989).

Transactivation of the HIV-1 LTR was identified in 1985 by Sodroski et al., who saw an increase in chloramphenicol acetyl transferase (CAT) activity when a construct containing the CAT gene driven by the HIV-1 LTR was transfected into infected cells (Sodroski et al., 1985b). Two groups pinpointed the regions of the genome responsible for the transcriptional activation of the HIV-1 LTR and named the gene, tat (Arya et al., 1985; Sodroski et al., 1985a). Without tat, the basal rate of transcription mediated by host cellular factors is very low (Laspius et al., 1989). With tat, a greater than
Figure 2.2. Schematic representation of the genome of HIV-1. *Tat* encodes a transactivating activity; *rev* regulates the expression of the structural proteins; *vpu* is involved in virion release and *Env* precursor protein processing; *vif* is an infectivity factor; *vpr* encodes a protein which may increase the kinetics of viral replication; *nef* encodes a nonessential protein whose function has not been determined; and *tev* has both *tat* and *rev* activities.
100-fold increase in viral RNA and protein synthesis is seen, with a 20-fold increase in RNA and a 5-fold increase in the translational efficiency of this RNA (Sharp and Marciniak, 1989; similar results are seen in Cullen, 1986). The mechanism behind the transcriptional activation activity is not completely elucidated: some studies implicate an enhanced initiation of transcription (Gentz et al., 1989; Hauber et al., 1987; Rice and Mathews, 1988); others favor an anti-termination of transcription mechanism or Tat-mediated enhancement of RNA polymerase processivity (Kao et al., 1987; Kessler and Mathews, 1992; Marciniak et al., 1990; reviewed in Cullen and Greene, 1990 and Sharp and Marciniak, 1989); and others suggest a combination of the two effects (Bengal and Aloni, 1991; Laspia et al., 1989). Tat interacts with a 59 nucleotide RNA stem-loop structure that is located in the R region of the LTR and so is present at the 5' end of all viral mRNA species (Rosen et al., 1986; Wright et al., 1986). This stem-loop structure, termed TAR, must be located immediately downstream of the transcription initiation site and is orientation dependent (Muesing et al., 1987). The TAR region has been called an RNA enhancer, as the Tat protein interacts specifically with the TAR element in its RNA form, leading to transcriptional activation (Berkhout et al., 1989). Whether Tat then interacts directly with RNA polymerase II or another protein of the transcriptional complex is not known; however, cellular proteins have been identified that bind the TAR region or that interact with the Tat protein (Gatignol et al., 1991; Nelbock et al., 1990; Rounseville and Kumar, 1992). The role of the cellular proteins has yet to be elucidated.

Tat has also been implicated in relieving the inhibition of translation seen with TAR-containing mRNAs. It has been suggested that Tat provides a potentiation of translation of these mRNAs by somehow activating them in the nucleus before their transport to the cytoplasm for translation (Braddock et al., 1989; Braddock et al., 1990; Braddock et al., 1991). Inhibition of translation of TAR-containing mRNAs may also be due to the inaccessibility of the cap structure, thereby preventing the initiation of translation, or to the activation of the double-stranded RNA dependent kinase (ds1), which then phosphorylates the translation factor eIF-2 and prevents its recycling. Tat may be able to relieve these blocks to translation, but again the details of the mechanism
of action are unknown (Cullen, 1986; Edery et al., 1989; Gunnery et al., 1990; Parkin et al., 1988; Roy et al., 1990; Roy et al., 1991; SenGupta et al., 1990).

**Vpu** The Vpu protein is a nonessential 81 amino acid phosphoprotein (Strebel et al., 1989). The protein is encoded on a bicistronic mRNA with the Env precursor protein. As most of the other HIV-1 proteins are encoded on transcripts that position them as the first open reading frame, the fact that the Env precursor protein is produced from a bicistronic message with Vpu suggests there may be a need for coordinate expression and regulation of Vpu and Env (Schwartz et al., 1990b). The AUG of the Vpu open reading frame is in poor homology with the Kozak consensus sequence and may allow bypass by ribosomes to express Env (Schwartz et al., 1990b). Vpu is an integral membrane protein, with a perinuclear location in the cell, and contains an extremely hydrophobic amino-terminus and a hydrophilic carboxy-terminus (Strebel et al., 1989). It is capable of forming high molecule weight complexes composed of homo-oligomers (Maldarelli et al., 1993). While vpu is found in HIV-1, there is no vpu gene in the other primate lentiviruses, including human immunodeficiency virus 2 (HIV-2) and the simian immunodeficiency viruses (SIV) (Terwilliger et al., 1989).

While vpu is not an essential gene, mutations in vpu result in viral particles with altered replicative capacity: there is a 5- to 10-fold decrease in progeny virions after infection of T cells with a vpu-defective virus. This decrease in virion production is not associated with a decrease in cytopathicity or slower kinetics of replication, but is associated with impaired release of progeny (Klimkait et al., 1990) with up to 5-fold more Gag, Pol, and Env proteins found intracellularly upon infection with viruses containing a mutated vpu gene as compared to wild type (Strebel et al., 1988). Vpu increases the amount of virus released by altering the ratio of intracellular to extracellular viral proteins and by enhancing the rate at which the viral proteins are exported from infected cells (Terwilliger et al., 1989. Yao et al., 1992). To do this, Vpu may suppress virion budding on membranes other than the plasma membrane, or it may interact with other proteins or releasing factors that promote virion release (Klimkait et al., 1990; Strebel et al., 1989).
Study of Vpu function has revealed two different pathways of action: one involves an indirect effect on the processing of the Env precursor protein (gp160) to the surface glycoprotein (gp120) and the transmembrane protein (gp41), while the other involves a more global effect to enhance virus release. It has been found that the Env precursor protein interacts with the HIV-1 receptor, CD4, in the endoplasmic reticulum (ER) of infected cells. This gp160-CD4 complex prevents rapid processing of the Env precursor protein and results in accumulation of this protein intracellularly. The presence of the Vpu protein in the ER can enhance the processing of gp160 by destabilizing CD4. Vpu enhances the intrinsic turnover rate of CD4, thereby releasing gp160 from the complex and allowing Env processing to proceed (Willey et al., 1992a). Vpu does not require the presence of the Env protein to induce degradation of CD4; experiments in which CD4 was retained in the ER by the drug brefeldin A showed that this CD4 was also degraded in the presence of Vpu (Willey et al., 1992b). The actual mechanism of degradation of CD4 by Vpu has not been elucidated. It has been hypothesized that HIV-1 encodes a Vpu protein, while HIV-2 and SIV do not, because the affinity of the interaction between HIV-1 Env protein and CD4 is stronger than that seen with HIV-2 and SIV Env proteins and CD4. The HIV-1 complexes formed may be harder to destabilize without the help of Vpu than those seen with HIV-2 or SIV. Interestingly, Env processing in HIV-2-infected cells is extremely slow (Willey et al., 1992b).

Vpu has another role in the viral life cycle that is not dependent on CD4 or Env that involves particle release (Yao et al., 1992). Studies on HIV-1 pseudotyped with HIV-2 or amphotropic murine leukemia virus Env glycoproteins indicate that there is still a decrease in the release of the capsid protein in the absence of vpu even when HIV-1 env is absent (Geraghty and Panganiban, 1993). Vpu has also been found to enhance particle production from chimeric proviral clones containing gag genes from HIV-2, visna virus, or Moloney murine leukemia virus and env and vpu regions from HIV-1, indicating there may not be a specific interaction between Vpu and HIV-1 Gag, but that Vpu may have a general effect that enhances retroviral budding, perhaps through modification of a cellular pathway (Gottlinger et al., 1993). This could explain why Vpu response is variable upon infection of different
cell types; perhaps Vpu substitutes in some cells for a low-level cellular factor necessary for virion release (Gottlinger et al., 1993).

Rev Rev is a 116 amino acid phosphoprotein encoded on two exons (Feinberg et al., 1986; Hauber et al., 1988; Sodroski et al., 1986). It is an essential protein and was first localized to the nucleus, or more specifically, the nucleolus of infected cells (Cullen et al., 1988; Felber et al., 1989). More recently, studies have indicated Rev is also present in the cytoplasm and perinuclear zone (Arrigo et al., 1992; Kalland et al., 1994), and there is some evidence suggesting that Rev shuttles between the nucleus and cytoplasm, traveling with the nucleolar shuttle protein B23 from the cytoplasm to the nucleus and with the intron-containing RNA from the nucleus to the cytoplasm (Arrigo et al., 1992; Fankhauser et al., 1991; Kalland et al., 1994). There are two main domains in the protein: a leucine-rich activation domain and an arginine-rich region responsible for RNA binding and nuclear localization with flanking sequences responsible for multimerization (Malim and Cullen, 1991; Malim et al., 1989a; Malim et al., 1991; Olsen et al., 1990; Zapp et al., 1991). The phosphorylation of Rev has not been found to be important for its function (Cochrane et al., 1989).

Rev was identified in 1986 by Sodroski et al. who discovered that the open reading frame encoding Rev was necessary for the efficient synthesis of the Gag, Pol, and Env proteins. They hypothesized that Rev acted at a post-transcriptional level as there was no detectable difference in the amount and size distribution of viral mRNAs with and without Rev (Sodroski et al., 1986). Feinberg et al. also identified rev in 1986, but this group saw a difference in the pattern of RNA produced with and without Rev and concluded that Rev was responsible for the differential splicing of viral RNA transcripts (Feinberg et al., 1986). The details of Rev action are still a subject of ongoing debate.

HIV infection occurs in two stages: early in infection, regulatory proteins are made; later, structural proteins are produced (Figure 2.3) (Greene and Cullen, 1990; Kim et al., 1989). Rev is responsible for the switch between these stages. Early in infection, only multiply-spliced transcripts appear in the cytoplasm and are capable of being translated. These transcripts lack large regions of *gag* and *pol* (the 5' intron) and *env* (the 3' intron), and code for the regulatory proteins, Tat, Rev, and Nef.
Tat transactivates the LTR, leading to the production of more multiply-spliced transcripts. Once the level of Rev reaches a threshold (Pomerantz et al., 1992), it induces the cytoplasmic expression of incompletely spliced transcripts and allows their translation. As the full-length and singly-spliced messages encode the structural proteins of the virus, Rev is an essential protein for viral replication (Greene and Cullen, 1990).

Two types of cis-acting elements in the incompletely spliced transcripts are involved in the Rev response. The first elements, the cis-acting repression sequences (CRS), are located in both the 5' and 3' introns (Cochrane et al., 1991; Hadzopoulou-Cladaras et al., 1989; Maldarelli et al., 1991; Nasiouslas et al., 1994; Rosen et al., 1988; Schwartz et al., 1992). Transcripts with these sequences are retained in the nucleus. When fully spliced, however, the CRS are removed and the multiply-spliced RNA is able to leave the nucleus. The second cis-acting element, the Rev-responsive element (RRE), is also located in the 3' intron and forms a complex structure with several stem-loops (Hadzopoulou-Cladaras et al., 1989; Malim et al., 1989b; Rosen et al., 1988). When Rev is present, the interaction of Rev and the RRE overcomes the inhibition of the CRS, and incompletely spliced messages containing CRS can exit the nucleus.

While the result of Rev action, the production of structural proteins for virion assembly, is clear, how Rev operates is unclear. Studies have suggested Rev may allow the transport of incompletely spliced messages from the nucleus to the cytoplasm (Felber et al., 1989; Malim and Cullen, 1993; Malim et al., 1989b); may increase the half-life of unspliced mRNAs (Felber et al., 1989; Malim and Cullen, 1993); may disrupt spliceosomes or prevent their assembly on the transcripts, thereby preventing splicing and allowing release from the nucleus of incompletely spliced messages (Chang and Sharp, 1989; Kjems and Sharp, 1993; Kjems et al., 1991; Lu et al., 1990); may assist or permit the association of incompletely spliced messages with polysomes to allow translation (Arrigo and Chen, 1991; D'Agostino et al., 1992; Knight et al., 1987); or may perform any combination of these functions.
Figure 2.3. Diagram of the stages of HIV-1 replication. Low-level, basal transcription from the HIV-1 LTR results in production of a primary transcript, which, early in the life cycle, is fully spliced to form the multiply-spliced transcripts. These multiply-spliced transcripts encode the regulatory proteins, Tat and Rev. Tat activates transcription from the viral LTR, resulting in increased production of multiply-spliced mRNAs. When the level of Rev protein reaches a threshold, Rev induces a switch between the early, regulatory stage of replication to the late, structural stage by allowing the expression of incompletely spliced messages and their structural gene products. With the production of the structural proteins, progeny virions can assemble and mature. The diagram is adapted from Greene and Cullen, 1990.
Proposed Functions of S1, S2, and S3 Open Reading Frames of EIAV It has been shown previously that EIAV contains a transactivating activity (Derse et al., 1987), and a study by Sherman et al. and sequence alignments (Figure 2.16) suggest that EIAV Tat activity is encoded by the S1 open reading frame (Sherman et al., 1988). The S2 open reading frame shows little sequence homology to any of the HIV-1 accessory proteins, but the exon is in a similar genetic location to that of HIV-1 vpu. As S2 contains several serine and threonine residues, the protein may be phosphorylated as is HIV-1 Vpu (Montelaro et al., 1993). The S3 open reading frame shows limited sequence homology with HIV-1 Rev (Figure 2.19), and the reading frame itself is located in the EIAV genome in a position analogous to the location of visna virus rev in its genome (Tiley et al., 1990).

To characterize the expression of these open reading frames and to elucidate their functions, the transcriptional pattern of EIAV-infected cells was examined, and cDNA clones were isolated that contain these open reading frames. These experiments and the various assays employed to determine the functions of the cDNA clones are described below.

MATERIALS AND METHODS

Cells and Viruses The feline embryonic adenocarcinoma (FEA) cell line and primary fetal equine kidney (FEK) cells were maintained in Eagle’s minimal essential medium with Earle’s salts, supplemented with 5% fetal bovine serum, 25 U penicillin, 25 μg of streptomycin sulfate, and 0.5 μg of amphotericin B (Fungizone; GIBCO/BRL) per ml. The canine thymus (Cf2Th) cell line was maintained similarly, except that medium was supplemented with 10% fetal bovine serum. Each cell type has advantages and disadvantages in the study of EIAV. FEK cells, which are equine cells, have the disadvantage of being primary cells and therefore difficult to transfect. Both FEA and Cf2Th cells are cell lines and can be transfected; however, while EIAV replicates in both lines, these lines are not equine and results of experiments in these lines may not be directly applicable to equine cell infections.
Finally, it is very difficult to clonally isolate and grow FEA cells, while this procedure can be performed on C127Th cells.

The EIAV strain designated PV was employed in the cDNA cloning studies (Orrego et al., 1982), as this strain replicates well in tissue culture and retains its virulence in horses. Viral stocks prepared from the molecular clone pSPEIAV19 (Payne et al., 1994) were used to infect C127Th cells and FEA cells for Northern blot analysis. This molecularly cloned strain of virus replicates in vitro and in vivo, but is not virulent. Cells at 50% confluency were incubated with virus for 1 hour at 37°C to allow binding, then washed three times with media, and refed. Infections were monitored either by immunoblot analysis (Sambrook et al., 1989) or by detection of reverse transcriptase (RT) activity (see RT Assays, p.29).

**RNA Preparation** RNA was prepared by one of three methods: by the cesium chloride centrifugation method described by Sambrook (Sambrook et al., 1989), through the use of the FastTrack mRNA Isolation kit following the manufacturer's suggested protocol (Invitrogen), or by the method of Chomczynski and Sacchi as modified by Xie and Rothblum (Chomczynski and Sacchi, 1987; Xie and Rothblum, 1991). In the latter method, cells were lysed in 1.9 M guanidinium thiocyanate, 11.9 mM sodium citrate, pH 7.0, 95 mM sodium acetate, pH 4.0, and 0.34% β-mercaptoethanol. Water-saturated phenol and chloroform-isooamyl alcohol (24:1) were added to 47.5% and 10%, respectively, and lysates were vortexed and placed on ice for 20 minutes. Nucleic acid in the aqueous layer was precipitated with an equal volume of isopropanol, washed twice with 70% ethanol, and resuspended in DEPC-treated dH2O. The FastTrack mRNA Isolation kit specifically isolates poly A⁺ RNA; the other two methods yield total RNA.

**Northern Blot Analysis** RNA samples were fractionated by electrophoresis in formaldehyde-containing agarose gels in 1xMOPS (0.02 M MOPS (3-[N-morpholino]propanesulfonic acid), 5 mM sodium acetate, 1 mM EDTA; pH 7.0) with buffer circulation. 2 μg of poly A⁺ RNA or 10-15 μg total RNA were analyzed. The gels were treated with 0.05 N NaOH for 30 minutes, 0.1 M Tris-Cl, pH 8.0
for 30 minutes, and transferred to nitrocellulose in 10xSSC (1xSSC is 0.15 M NaCl plus 0.015 M sodium citrate). The membranes were subjected to UV-crosslinking, baked at 80°C for 2 hours, prehybridized in 1% bovine serum albumin, 0.4 M sodium phosphate, pH 7.2, 15% formamide, 1 mM EDTA, pH 8.0, and 7% SDS, pH 7.2 at 65°C for 1-2 hours, and hybridized overnight at 65°C with 3x10⁶ cpm of probe. Probes used were either a ³²P-labeled actin fragment generated by random primed synthesis or a single-stranded DNA probe generated by run-off PCR in the presence of α-³²P-dCTP. This PCR-generated probe hybridizes to the region of the RNA 5' to the major splice donor site, thereby hybridizing to all EIAV mRNA species (see PCR-generated Probe, next section). Membranes were washed 30 minutes at 50°C in 2xSSC, 1% SDS, 30 minutes at 50°C in 0.2xSSC, 0.1% SDS, and 30 minutes at 65°C in 0.2xSSC, 0.1% SDS, followed by autoradiography.

PCR-Generated Probe The plasmid pWS1, which contains the leader region, S1, and S3, was linearized by digestion with SalI, extracted with phenol-chloroform-isoamyl alcohol (25:24:1), and ethanol precipitated. PCR reactions included approximately 5 ng of linearized pWS1, 0.2 µM oligo 280 (Table 2.1), dATP, dGTP, and dTTP, each at a final concentration of 167 µM, 100 µCi α-³²P-dCTP (specific activity of 3000 Ci/mmole), 10 mM Tris-Cl, pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 0.01% (w/v) gelatin, and 2.5 units TAQ DNA polymerase. The PCR profile included 1 minute of denaturation at 94°C, 2 minutes of annealing at 37°C, and 1 minute of extension at 72°C for 40 cycles. The probe was purified by centrifugation through G25 or G50 sephadex spin columns. This probe should hybridize to all viral messages as it contains sequences 5' to the major splice donor site (Figure 2.4).

cDNA Synthesis 1 µg poly A⁺ RNA was reverse transcribed by 200 units Moloney murine leukemia virus reverse transcriptase in 50 mM Tris-Cl, pH 8.3, 75 mM KCl, 3 mM MgCl₂, 10 mM DTT, 0.5 mM each dNTP, 421 pmoles primer (oligo 7754, 7841, or 5672; see Table 2.1), and 3 units RNasin for 1 hour at 37°C. This first-strand cDNA template was then added to a PCR reaction containing 0.2 µM primers (oligos 228 and 7754, 7841, or 5672; see Table 2.1), 1.5 mM MgCl₂, 100 mM Tris-Cl, pH 8.3,
50 mM KCI, 0.01% (w/v) gelatin, 0.2 mM each dNTP, and 2.5 units TAQ DNA polymerase. The reactions were denatured for 1 minute at 94°C, annealed for 2 minutes at 40-45°C, and elongated for 3.5 minutes at 72°C for 30 cycles.

**Southern Blot Analysis** PCR product or cDNA clones were analyzed by standard Southern blot techniques (Sambrook *et al.*, 1989). The membranes to be probed with oligos were hybridized overnight at 37-45°C in 5xSSC, 10x Denhardt's (1% Ficoll, 1% polyvinylpyrrolidone, 1% bovine serum albumin), 10 mM sodium phosphate, pH 7.0, 1% SDS, and 100 µg/ml denatured salmon sperm DNA with 5' ³²P-end-labeled oligos (oligo 278 [hybridizes to a region in the S2 open reading frame], 279 [hybridizes to a region in the S1 open reading frame], 287 [hybridizes to the LTR], or 7427 [hybridizes to a region in the S3 open reading frame]; see Table 2.1). These blots were washed three times for 20 minutes in 5xSSC, 0.1% SDS at 37°C or room temperature, followed by autoradiography. The membranes to be probed with ³²P-labeled proviral clone DNA were prehybridized in 6xSSPE (1xSSPE is 0.15 M NaCl, 10 mM sodium phosphate, and 1 mM EDTA; pH 7.4), 10x Denhardt's, 1% SDS, and 50 µg/ml denatured salmon sperm DNA, and hybridized in 6xSSPE, 1% SDS, 10% dextran sulfate, 100 µg/ml denatured salmon sperm DNA, and 3x10⁵ cpm/ml of a ³²P-labeled proviral clone generated by random primed synthesis. These blots were washed twice for 15 minutes in 6xSSPE, 0.5% SDS at room temperature, twice for 15 minutes in 1xSSPE, 0.5% SDS at 37°C, and once for 20 minutes in 0.1xSSPE, 0.5% SDS at 65°C, followed by autoradiography.

cDNA Clones cDNA generated by PCR was heated for 10 minutes at 95°C to inactivate TAQ DNA polymerase. The Klenow fragment of *E. coli* DNA polymerase I was used to repair any incomplete cDNAs, and the DNA was phosphorylated by T4 polynucleotide kinase and fractionated on low gelling point agarose gels. The DNA of interest was excised from the gel and purified with an Elutip column (Schleicher and Schuell) following the manufacturer's suggested protocol. This cDNA was ligated to SmaI-cut, phosphatased pBS+ (Stratagene), and used to transform DH5α cells (GIBCO/BRL). Plasmids containing sequences of interest were determined by Southern blot analysis using oligo 287 (Table 2.1) and sequenced by the dideoxy sequencing method using Sequenase (United States
<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>228</td>
<td>5’-CTG AAA AGG CCT TTG TA-3’</td>
</tr>
<tr>
<td>278</td>
<td>5’-TGA CTG TTG GGT AAT AGG GG-3’</td>
</tr>
<tr>
<td>279</td>
<td>5’-GAT TGC CTT CAG TCT TTG-3’</td>
</tr>
<tr>
<td>280</td>
<td>5’-TAA TCA ATT CCT GTC CTC CT-3’</td>
</tr>
<tr>
<td>287</td>
<td>5’-GCG TAG GAT CTC GAA CAG ACA AAC TAG AGA CAG GG-3’</td>
</tr>
<tr>
<td>5672</td>
<td>5’-GGG TGT AGA CCA TTT CAA AAT TAC TTC ATT-3’</td>
</tr>
<tr>
<td>7427</td>
<td>5’-GCC GCC TGT TGT A-3’</td>
</tr>
<tr>
<td>7754</td>
<td>5’-GCC TCT AAT ACA AAT CC-3’</td>
</tr>
<tr>
<td>7841</td>
<td>5’-GGC ATT GAT ACA TGA GA-3’</td>
</tr>
<tr>
<td>S2A</td>
<td>5’-GGG GTC GAC ATC GAC GGA AGA ATG GGA TTA TTT G-3’</td>
</tr>
<tr>
<td>S2B</td>
<td>5’-GGG AAG CTT CAT TTC TTG GTC-3’</td>
</tr>
<tr>
<td>S3stop</td>
<td>5’-GGG GGG GGG CCC TAA GGA TCT TAG GCG AAG AGG-3’</td>
</tr>
</tbody>
</table>
Figure 2.4. Synthesis of a single-stranded DNA probe that detects all EIAV RNA. Plasmid pWS1 was linearized by Sall and used as a template in a run-off PCR reaction that included oligo 280 (Table 2.1) and α-32P-dCTP. Oligo 280 binds to the 5' half of the major splice donor site (nucleotide 438), and PCR amplification using this oligo results in an antisense DNA strand that contains sequences 5' to the major splice donor site and that can hybridize to all viral transcripts. MSD, major splice donor site.
Biochemical).

**RNase Protection Assays** Most of the S3 region was deleted from the plasmid pWS1 (see Figure 2.9) by digestion with NruI and EcoRI, followed by religation, generating a plasmid designated pWS1A. Digestion of pWS1A by SalI, followed by *in vitro* transcription by T7 RNA polymerase in the presence of α-32P GTP yields an RNA probe (designated Probe A) of 478 bases that can anneal to RNA containing leader, S1, and S3 sequences. Similarly, most of the S3 region was deleted from the plasmid pWS2 (see Figure 2.9) by digestion with NruI and HincII and religation to generate a plasmid designated pWS2B. Digestion of pWS2B by EcoRI, followed by *in vitro* transcription by T3 RNA polymerase in the presence of α-32P GTP yields an RNA probe (designated Probe B) of 579 bases that can anneal to RNA containing leader, S1, part of env, and S3 sequences. The probes were gel-purified from denaturing polyacrylamide gels (Sambrook *et al.*, 1989). RNA samples were lyophilized, and RNase T2 hybridization buffer (80% formamide, 40 mM PIPES, pH 6.5, 400 mM NaCl, 1 mM EDTA, pH 8.0) and 100,000 cpm of probe were added to the pellet. The reaction was heated at 95°C for eight minutes and incubated overnight at 55°C. RNase T2 digestion buffer (300 mM NaCl plus 20 mM sodium acetate, pH 4.5) and 6 units RNase T2 were added to the reactions and incubated for 45 minutes at 37°C to digest unprotected RNA. SDS was added to a final concentration of 0.5% and proteinase K digestion occurred for 30 minutes at 55°C, followed by extractions with phenol-chloroform-isoamyl alcohol (25:24:1) and chloroform-isoamyl alcohol (24:1), and ethanol precipitation. Formamide loading buffer (80% formamide, 1 mM EDTA, pH 8.0, 0.1% bromophenol blue, 0.1% xylene cyanol) was added to the samples, and they were heated for five minutes at 95°C. Samples were then subjected to electrophoresis on 5% denaturing polyacrylamide gels, followed by autoradiography. Reactions contained 200 ng - 1 μg of either poly A+ RNA isolated from infected FEK cells or *in vitro* transcribed sense RNA (positive control) plus enough yeast tRNA to bring the total RNA sample to 50 μg. To generate the RNA used as a positive control, pWS1A was digested with
EcoRI and transcribed with T3 RNA polymerase, while pWS2B was digested with SphI, followed by transcription by T7 RNA polymerase. Negative control reactions contained 50 µg tRNA.

**Eukaryotic Expression Vectors** The vector pSV-tat was created by the digestion of pWS1 with Stul and NruI, followed by the isolation of the 435 bp fragment containing the leader region, S1, and the beginning of S3. This fragment was subcloned into the Stul-cut, phosphatased eukaryotic expression vector, pSVΔCAT, which contains the SV40 early promoter, SV40 origin of replication, and a polyadenylation site. pSV-S2 was created by subcloning the S2 open reading frame flanked by PvuII and Smal restriction sites into Hpal-cut, phosphatased pSVΔCAT. pSV-rev was constructed by the digestion of pWS3 with PvuII and HindIII, followed by ligation with Stul-HindIII-cut pSVΔCAT.

**Bacterial Expression Vectors** The S2 open reading frame was amplified in a standard PCR reaction with primers S2A and S2B (Table 2.1). The product was digested with Sall and HindIII, and ligated in a three-way ligation to the 1 kb band of Xbal-HindIII-digested pDS56/RBSII.6xHis and the 2.3 kb band of XbaI-Sall-digested pDS56/RBSII.6xHis to yield pH52. The pDS56/RBSII.6xHis vector is a bacterial expression vector that contains the lac promoter/operator region and is repressed in the presence of Lac repressor, which is supplied in high levels in trans from a second plasmid pDMI.1. Induction with isopropyl-β-D-thiogalactopyranoside (IPTG) is necessary for protein expression. pDS56/RBSII.6xHis also places six histidines at the amino terminus of the protein for ease of protein purification on a nickel nitrilotriacetate sephaose (NTA-Sepharose) column (Hochuli et al., 1987; Hochuli et al., 1988; LeGrice et al., 1990).

**Protein Expression** *Escherichia coli* (*E. coli*) cells containing pH52 were grown to an absorbence (A_{600}) of about 0.8 and were induced with 200 µg/ml IPTG for three hours (LeGrice et al., 1990). Bacteria were pelleted and lysed in 10 mM Tris-Cl, pH 8.0, 100 mM NaCl, 1 mM EDTA, pH 8.0, and 1% SDS. Protein expression was detected by Coomassie blue staining of SDS-PAGE gels or by standard Western blot analysis (Sambrook et al., 1989) using a polyclonal antibody that recognizes a peptide from the S2 open reading frame.
Transfections  Cells were transfected by one of three methods: by the use of Lipofectin, following the manufacturer's suggested protocol (GIBCO/BRL); by the calcium phosphate method (GIBCO/BRL), or by the use of liposomes made in the laboratory. To make liposomes, 5 mg dioleoyl-L-α-phosphatidylethanolamine was combined with 2 mg dimethyldioctadecylammonium bromide and lyophilized to dryness. The pellet was resuspended in water and sonicated for six minutes with pausing every 10 seconds to prevent overheating. The liposomes were aliquotted and stored at 4°C (Rose et al., 1991). For transfections, DNA was incubated with 30 μg liposomes for 15 minutes at room temperature and added to cells in media without fetal bovine serum for 3 hours or overnight at 37°C. Media with 20% fetal bovine serum was then added.

CAT Assays  FEA cells were transfected with pSV-tat and pEICAT using the Lipofectin reagent. pEICAT contains the gene for the chloramphenicol acetyltransferase (CAT) enzyme driven by the EIAV LTR (Dorn and Derse, 1988). After 48-72 hours, cells were harvested with trypsin and counted. Equal numbers of cells were then lysed in 0.25 M Tris-Cl, pH 7.5 by three cycles of freeze-thaw. Cellular debris was removed by centrifugation and the supernatant used for the CAT assay. The reactions were performed in scintillation vials and contained 2.5 mM chloramphenicol, 125 mM Tris-Cl, pH 7.5, 0.5 μCi 3H acetyl coenzyme A (1.4 Ci/m mole), and either cell extract, 125 mM Tris-Cl, pH 7.5 for a negative control, or CAT standards for a positive control. Reactions were overlaid with EconoFluor scintillation fluid and subjected to scintillation counting about every 20 minutes for two hours (Sambrook et al., 1989).

Stable S3 Cell Lines  The plasmids pSV-rev and pY3 were transfected into Cf2Th cells through the use of liposomes. pY3 contains the hygromycin resistance gene driven by the Moloney sarcoma virus LTR (Blochlinger and Diggelmann, 1984). Cells were selected with 0.1-0.2 mg/ml hygromycin. Cell lines resistant to hygromycin were analyzed by immunoprecipitation followed by Western blot analysis to determine if the Rev protein was being produced (Sambrook et al., 1989). Briefly, cells were lysed in triple-detergent lysis buffer (50 mM Tris-Cl, pH 8.0, 150 mM NaCl, 0.02% sodium azide, 0.1%
SDS, 100 µg/ml phenylmethylsulfonyl fluoride, 1 µg/ml aprotenin, 1% Nonidet P-40, 0.5% sodium deoxycholate), and the lysate incubated with a monoclonal antibody to epitope A (KEARDQEM: Montelaro et al., 1989) of the surface glycoprotein (gp90) overnight at 4°C in NET-gel buffer (50 mM Tris-Cl, pH 7.5, 150 mM NaCl, 0.1% Nonidet P-40, 1 mM EDTA, pH 8.0, 0.25% gelatin, 0.02% sodium azide). rProtein G agarose (GIBCO/BRL) was added to the lysates and incubated for 1 hour at 4°C. The complexes were washed twice with NET-gel buffer and once with 10 mM Tris-Cl, pH 7.5, 0.1% NP-40. Sample buffer (50 mM Tris-Cl, pH 6.8, 100 mM dithiothreitol, 2% SDS, 0.1% bromophenol blue, 10% glycerol) was added, and the samples were heated at 85°C for 10 minutes to release the antigen-antibody complexes from the rprotein G agarose. Samples were fractionated by SDS polyacrylamide gel electrophoresis and transferred to nitrocellulose using the Bio-Rad transfer unit. Standard Western blot analysis was performed using the epitope gp90 A monoclonal antibody and a horse radish peroxidase conjugate.

Mutagenesis of the S3 Open Reading Frame The env region of the infectious proviral clone, pER (Whettler et al., 1990), was subcloned into pBS+ using SphI and EcoRI sites, creating pBS-env. PCR was employed to produce DNA containing a single base substitution in the S3 open reading frame that creates a stop codon in the S3 frame while not affecting the overlapping env reading frame (Figure 2.5). PCR reactions contained 0.5 µM primers (M13 reverse primer [Stratagene] and S3stop [Table 2.1]) and EcoRI-linearized pBS-env DNA, and amplified a 1.8 kbp fragment that contained the mutation as the S3stop oligo contains the base change. This mutated fragment was re-introduced into pBS-env, creating pBS-env(stop), and the mutation confirmed by dideoxy sequencing. pBS-env(stop) was digested with SphI and EcoRI, and the mutated env fragment was ligated to the 5' end of another infectious proviral clone, pSPEIAV19 (Payne et al., 1994) to create p19S3stop. Again, the mutation was confirmed by sequencing, and the env gene was partially sequenced to confirm that there were no other mutations. Digestion of p19S3stop by restriction enzymes confirmed its overall structure.
p19S3stop was transfected into FEA cells by the calcium phosphate method and was monitored for viral replication by reverse transcriptase (RT) assays.

Reverse Transcriptase Assays RT assays were performed as described by Willey et al. (Willey et al., 1988). Reactions contained 0.25 optical density unit of poly(rA)-poly(dT)_{12-18}, 50 mM Tris-Cl, pH 7.8, 7.5 mM KCl, 2 mM dithiothreitol, 5 mM MgCl₂, 0.05% Nonidet P-40, and 0.5 μCi of ^32P-TTP (3000 Ci/mmole). Fifty μl of reaction cocktail was added to 10 μl culture supernatant (containing 1% Triton X-100) and incubated at 37°C for 1 hour. Reactions were terminated by spotting 5 μl of the reaction onto DE81 chromatography paper. The papers were washed twice for 5 minutes in 2xSSC, twice for 5 minutes in 95% ethanol, dried, and subjected to scintillation counting.
Figure 2.5. Mutagenesis of the S3 open reading frame without affecting the overlapping env reading frame. PCR, using the S3stop oligo (Table 2.1) which contains a thymine residue in place of a cytosine residue at base number 7219, results in mutant DNA that contains a stop codon as the tenth codon of the S3 reading frame, while keeping the leucine in the overlapping env frame unchanged.
RESULTS

Detection of Multiply-Spliced Transcripts in EIAV-infected Cells To begin our analysis of regulatory genes in EIAV, we first examined viral transcription to determine if EIAV, like other lentiviruses, produced three classes of transcripts: full-length, singly-spliced, and multiply-spliced. Fetal equine kidney (FEK) cells were infected with the PV strain of EIAV. This strain of virus was produced by serial back-passage of prototype virus through horses, and so is able to replicate in tissue culture while retaining its virulence in animals. FEK cells were chosen for these studies as infection of these equine cells results in the production of high titers of EIAV. Persistently infected FEK cells were lysed and Northern blot analysis was performed. As shown in Figure 2.6, three classes of transcripts were detected using the PCR-generated probe that hybridizes to all EIAV mRNAs. The 8.2 kb band represents the full-length message; the 3.0-4.2 kb band represents the singly-spliced message(s) (including env); and the 1.5-2.0 kb bands represent the multiply-spliced messages that presumably encode regulatory proteins. Similarly, persistent infection of the canine thymus (Cf2Th) cell line by a molecularly cloned strain of virus (pSPEIAV19) also results in three classes of viral transcripts (Figure 2.6). Therefore, a strain of virus that retains virulence in animals produces multiply-spliced mRNAs upon infection of FEK cells, as does a molecularly cloned virus upon infection of Cf2Th cells.
Figure 2.6. Northern blot analysis of RNA from EIAV-infected cells. RNA was prepared from infected FEK or Cf2Th cells, and Northern blot analysis was performed using a $^{32}$P-labeled PCR-generated probe which detects all viral transcripts. U. RNA from uninfected cells. I. RNA from infected cells; a. full-length RNA, b. singly-spliced RNA, c. multiply-spliced RNA.
Temporal Analysis of RNA Production  Studies of HIV-1 have uncovered two stages in HIV-1 replication: an early, regulatory stage in which only small, multiply-spliced mRNAs are produced, and a late, structural stage, in which the incompletely spliced messages necessary for the production of structural proteins are produced (Greene and Cullen, 1990; Kim et al., 1989). Rev, once it reaches a certain threshold level, has been shown to be responsible for the switch between stages (Pomerantz et al., 1992). To examine the temporal pattern of EIAV transcription, FEA cells were infected with virus produced from pSPEIAV19 and lysed at 0, 4, 8, 12, 16, 20, 24, and 48 hours post-infection (p.i.). Total RNA was isolated and Northern blot analysis performed using a $^{32}$P-labeled actin probe generated by random primed synthesis or a single-stranded $^{32}$P-labeled viral probe that should detect all mRNAs as it contains sequences 5' to the major splice donor site. Figure 2.7 shows the results of this analysis. The multiply-spliced messages are first detected 8-12 hours p.i., with the env message appearing by 12-16 hours p.i., and the full-length message by 16 hours p.i. The RNA detected at 4 hours p.i. may be from the input virus as it is somewhat degraded. While the levels of the multiply-spliced transcripts increase gradually over time, the levels of the env message and especially the full-length message appear to increase more dramatically after 20 hours p.i. This experiment shows that EIAV messages do appear in an ordered manner as do those of HIV-1; however, the temporal division in production of multiply-spliced versus incompletely-spliced transcripts is not as pronounced as it is in HIV-1 infection (Kim et al., 1989).
Figure 2.7. Northern blot analysis of the temporal pattern of RNA production in infected FEA cells. FEA cells were infected with virus produced from pSPEIAV19, and RNA was prepared at various hours after infection. Northern blot analysis was performed using a $^{32}$P-labeled actin fragment generated by random primed synthesis or a $^{32}$P-labeled viral probe generated by PCR that detects all viral transcripts. hrs. p.i., hours post-infection; a. full-length RNA, b. singly-spliced RNA, c. multiply-spliced RNA, d. actin RNA.
Production of cDNA Clones and Southern Blot Analysis  cDNA clones were isolated in order to determine the nature of the multiply-spliced mRNAs. RNA prepared from PV-infected FEK cells was reverse transcribed by Moloney murine leukemia virus reverse transcriptase using a primer that hybridized to the 3’ end of env (oligos 7841 or 7754 [see Table 2.1]), and this DNA was amplified by PCR (Figure 2.8A). The 5’ primer for PCR (oligo 228) hybridizes to the R region of the long terminal repeat (LTR), while the 3’ primer (either 7754 or 7841) hybridizes to the 3’ end of the env gene. Therefore, the potential to amplify the whole genome exists; however, it is more likely that only cDNAs representing the multiply-spliced messages will be amplified as these will probably be the only messages completely reverse transcribed or elongated in the PCR with the extension time used (3.5 minutes). Before cloning, the PCR products were fractionated on agarose gels and characterized by Southern blot analysis using four oligo probes (Figure 2.8B). Oligo 278 hybridizes to the S2 reading frame; oligo 279 hybridizes to the S1 reading frame; oligo 287 hybridizes to the R region of the LTR and so should detect all messages; and oligo 7427 hybridizes to the S3 open reading frame. Figure 2.8C shows the results of the Southern blot analysis: several bands are detected by each probe, ranging in size from 350 bp to 3.0 kbp.
A. Poly A⁺ RNA prepared from infected FEK cells was reverse transcribed into first-strand cDNA, and this cDNA was amplified by PCR using oligos 228 and 7754 or 7841. Oligo 228 hybridizes to the R region of the LTR, while oligos 7754 and 7841 hybridize to the 3' end of env. B. Southern blot analysis of the PCR product was performed using ³²P end-labeled oligos 287, 279, 278, or 7427. Oligo 287 binds to the LTR; 279 recognizes a part of the S1 open reading frame; 278 hybridizes to the 5' region of the S2 open reading frame which does not overlap env; and 7427 binds to the S3 open reading frame. C. Southern blot analysis. Markers are in kbp.
Characterization of cDNA Clones  PCR products were subcloned into pBS+ and potential positive clones (as determined by Southern blot analysis with a $^{32}$P-labeled proviral clone generated by random primed synthesis) were sequenced. Three different classes of cDNAs were isolated (Figure 2.9). The first class of cDNA is represented by pWS1, in which two splicing events occur. First, the 5' leader region is spliced in frame onto the S1 open reading frame (nucleotide 438 to 5114; using the numbering of the viral sequences available from Genbank Accession numbers M16575, M11337, K03334, M14855), creating an open reading frame of 252 base pairs, but without an AUG to initiate translation. Following the stop codon of S1, there is a second splicing event to the splice acceptor site 5' to S3 (nucleotide 5255 to 7214). Therefore, this cDNA contains two open reading frames, leader-S1 and S3; however, there is still no AUG for either frame (Figure 2.10). The leader region does contribute to S1 four codons that could potentially initiate translation, including two CUGs, a UUG, and an AUC (Figure 2.10). If protein synthesis began with the second CUG, which is in the most favorable context for translation based on comparison to the Kozak consensus sequence, the product would contain 75 amino acids.
Figure 2.9. Schematic representation of the three classes of cDNA clones isolated from infected FEK cells. pWS1 contains the leader sequence, S1, and S3. pWS2 and pWS3 contain leader, S1, part of env, and S3. pWS4 contains part of the gag gene spliced onto part of env.
CT GAA AAG GCC TTT GTA ATA AAT ATA ATT CTC TAC TCA GTC CCT GTC
  Glu Lys Ala Phe Val Ile Asn Ile Ile Leu Tyr Ser Val Pro Val
  47
TCT AGT TCG TCT CTA CGA GAT CCT ACA GYT GCC GCC CGA ACA GGG ACC
  Ser Ser Leu Ser Val Arg Asp Pro Thr Val Gly Ala Arg Thr Gly Thr
  95
TGA GAG GGG CGC AGA CCC TAC CGT CGT AAT CTC GAA GAT CTT AGG ATC
  *** Glu Gly Arg Pro Tyr Leu Leu Asn Leu Ala Asp Arg Arg Ile
  143
CCC GGG ACA GCA GAG GAG GCA TTA GAG AAG GCC GCC TGT GCT CCT
  Pro Gly Thr Ala Glu Asn Leu Gin Lys Ser Ser Gly Gly Val Pro
  191
GCG GAG ACA AGA GCC AAC TAC CAT TCT CAG 389
  Gly Gin Thr Asp Tyr His Cys Gin
  389
AGG ATG CAG GAA CTG AAC CCA AGA GCA TAC CAT TCT CAG 437
  Thr Cys Ile Ala Arg Pro Tyr Leu Arg Ala Ser Leu
  54
TGG TTT TGG CAA GTC CTC CCG CAG TCG TTA CCT GAA GAA AAA ATT CCA TCA CAA
  Thr Cys Ile Ala Arg Pro Tyr Leu Arg Ala Ser Leu
  479
CCT CGT CAT CAC CGT CGT ATG CGG ACA ATA CTA CAA CCA GAA
  Pro Ser Arg Arg Asp Arg Thr Ile Arg Gly Gin Ile Leu Gin Pro Glu
  527
GTC CAG GAA CGA CGT TAT GTA AGA AGG AGA ATA CTA CAA CCA GAA
  Val Leu Gin Glu Arg Leu Thr Arg Arg Asp Arg Arg Gin Ile Leu Gin Val Glu Gin Ala
  575
GCC AAA GAC CGT GAT GAA GTC ATT CCA GGC ATT AGA GCG TAT CAA
  Ala Lys Glu Leu Gly Val Asn Arg Gly Ile Thr Arg Arg Tyr Cys
  623
TTG CGA GAA GAC CGA CGG TCG GGG GAA GAA CTA TCA CCA AGT GCC CTT
  Phe Arg Glu Asp Gin Arg Gly Asp Phe Ser Ala Thr Gly Gly Tyr Gin
  671
AGA CCT CGG CTC CCG GCC GCC TAC CCT TGG GGC GCC TAT CAA
  Arg Ala Gin Glu Arg Leu Thr Gly Gin Ser Ser Pro Arg Val Leu
  719
AGA CGG CAA GAA CTC TGG GGG GAA CAA TCC TCA CCA AGT GCC CTT
  Arg Pro Gly Asp Ser Lys Arg Arg Arg Lys His Gin
  761
CTGGTCATTTAAACCCGCGAGAGACTCTCCCTATCCCTGGCTCGCGTTTCCCCTT
  ATG
  821
GCT ATT TGG GGG ACT AGT AAT TAT AGT AGG AGC CAT AGG CTG AAG CGG
  Ala Ile Leu Gly Thr Ser Asn Tyr Ser Arg Thr His Ser Arg Leu Thr
  869
ATT AGG TGG ACT CCC TTG TAT AAT AAG GAT TTG TAT TAG AGC
  Ile Thr Trp Thr Arg Cys Tyr Asn Lys Asp Leu Tyr
  912

Figure 2.10. Nucleotide and predicted amino acid sequence of pWS1. The splice sites are indicated by ↓, and the codons that may initiate translation are in bold italics.
The second class of cDNA is represented by pWS2 and pWS3, which represent mRNAs generated by the same splicing events, but were primed with different oligos for cDNA synthesis and PCR. pWS2 (primed with oligo 7841) is 87 bp longer than pWS3 (primed with 7754). This class of cDNA was generated through three splicing events: the leader was spliced onto the S1 open reading frame as in pWS1; after the S1 open reading frame, there is a second splice to a region in env; after 101 bp in env (nucleotides 5416 to 5516), there is a third splice to the splice acceptor site 5' to the S3 open reading frame (nucleotide 5516 to 7214) (Figure 2.9). This class of cDNA also contains two open reading frames: the leader-S1 frame seen in pWS1 and an env-S3 open reading frame. The leader-S1 region still has no AUG for translation; however, the env region in the env-S3 open reading frame contains two methionines (Figures 2.11 and 2.12). If protein synthesis began at the first AUG, which is in better consensus with the Kozak consensus sequence than the second AUG, the protein would contain 165 amino acids. pWS2 contains a deletion in S3 which switches the reading frame to env after 136 codons of env-S3, but contains an intact leader-S1 region (pWS2 contains an AA dinucleotide at base numbers 773 and 774 of the sequence shown in Figure 2.11 as compared to the CAG codon found at the similar position in pWS3 [nucleotides 772-774]). pWS3 contains two deletions in the leader-S1 region which result in premature termination of the putative leader-S1 protein, but encodes an intact env-S3 region (one of the three guanosines at nucleotides 316-318 of the leader [corresponding to two guanosines at nucleotides 90 and 91 of the sequence shown in Figure 2.12] is deleted, as is one of the three cytosines at nucleotides 5123-5125 of the S1 reading frame [resulting in two cytosines at nucleotides 221 and 222 of Figure 2.12]).
Figure 2.11. Nucleotide and predicted amino acid sequence of pWS2. The splice sites are indicated by ↓, and the codons that may initiate translation are in bold italics. The codon which marks the switch from the S3 reading frame to the env reading frame due to a deletion in pWS2 is in bold.
GCA GGC TAT GGA TTA CGT GCA CTC GCT GTT ATA ATA AGG ATT TGT ATT 1006
Ala Gly Tyr Gly Leu Arg Gly Leu Ala Val Ile Ile Arg Ile Cys Ile 214

AGA GGC TTA AAT TGT ATA TTT GAA ATA ATC AGA AAA ATG CTT GAT TAT 1054
Arg Gly Leu Asn Leu Ile Phe Glu Ile Ile Arg Lys Met Leu Asp Tyr 230

ATT GGA AGA GCT TTA AAT CCT GCC ACA TCT CAT GTA TCA ATG CC 1098
Ile Gly Arg Ala Leu Asn Pro Gly Thr Ser His Val Ser Met 244
Figure 2.12. Nucleotide and predicted amino acid sequence of pWS3. The splice sites are indicated by ↓, and the codons that may initiate translation are in bold italics. The two nucleotides which are missing from the pWS3 sequence are above the positions in the sequence where they belong.
The third class of cDNA is represented by pWS4, which was generated by one splicing event using a different 5' splice donor site in the gag gene (nucleotide 662). This region is spliced onto the 3' end of env at nucleotide 7478, to generate an open reading frame of at least 171 codons (if the sequence continues past the priming spot of oligo 7754 to the end of env, the predicted size of the protein product would be 32 amino acids longer [Figures 2.9 and 2.13]). The protein coded for by this message would contain 73 amino acids of the gag MA protein and the rest would be derived from the env TM protein.
CTGAAAGAGGCCCTTTGTAATAATATAATATTCTCT
ACTCATCCTCTCTCTCTACTTGGTCTGTGCAG
ATCCTACGCTGGCGGCCGAAACAGGGAACCTTGAGA
GGGCGCCAGACCTCTACCTGTTGAACCTGGCTGA
TCGTAGGATCCCAGGACAGCAAGAGAAGAC
ACAGAAGTCTCTCTGGAGGTGTTCCTTGCCAGAA
CACAGGAGGACAGGTAGATTGGAGACCCTTTGAGAC
MetGlyAspProLeuThrTrp
AGCAGGCTCTAAAGTATGATAGAGGTCGTAACGGGCTTGCAG
SerLysAlaLeuLysLeuGluValThrValGlnGlySerGln
AAAATTACTACTGOTAAATGTTGCGCTAATCTAATGACCTTA
LysLeuThrThrGlyAsnCysAspTrpAlaLeuSerLeuValAspLeu
TTTATTGATACTTCTTGAAGAAGGACCTGCTGCTGAGGAT
PheHisAspThrAsnPhenValLysGluLysAspTrpGlnLeuArgAsp
GTCATTCCAACCCTTGTGAACGTTACGACGCTGCTCAAGAA
ValIleProLeuLeuGluAspValThrGlnThrLeuSerGlyGinGlu
662↓17478
AGAGAGCTATTGCCGAGAAGACCACAAAGGAGATTCTCTGCT
ArgGluSerTyrIleSerGluLysThrGlyGluIleGlnPro
GGGGCCTCTAAGACAGCACAGAAGGCGCTGGGGGACCTCT
GlyAlaAlaIleAsnGluHisLysGlySerGlyAsnAsnPro
CACAGGCTCTTATACTGAGGATTGCAACCAGAAGGGAGA
ArgGlnGlySerLeuAspLeuGluIleArgSerGluGlyAsnIle
TATGACCTGTGATTAAAGCCTAAGAAGACTCCTCTGCTG
TyrAspCysIleLysAlaGlnGluGlyThrLeuAlaIleProCys
TGGGTATTCCCCTTAGCTGCTGGAGATACTGAGA
GlyGlyPheGluProLeuLeuLeuPheGlyLeuValIleValGly
GGGATAAGGGCTTAGTTGCAGCTGGTCTGTTGATTATAAGGATT
ArgIleAlaGlyTyrGlyLeuArgGlyLeuAlaValIleArgIle
TGTATTAGGCC
CysIleArgGly
729
71

Figure 2.13. Nucleotide and predicted amino acid sequence of pWS4. The splice sites are indicated by ↓, and the codon that may initiate translation is in bold italics.
**RNase Protection Analysis** In order to confirm the authenticity of the cDNA clones obtained by PCR, we performed RNase T2 protection assays. Poly A⁺ RNA isolated from infected FEK cells was incubated overnight with one of two ³²P-labeled antisense RNA probes generated by *in vitro* transcription. The duplexes were then subjected to RNase T2 digestion and analyzed by fractionation on denaturing polyacrylamide gels, followed by autoradiography. Probe A contains the leader region, S1, and part of S3, while probe B contains the leader region, S1, the 101 bases of env present in pWS2, and part of S3 (Figure 2.14A). Because all transcripts are generated from one primary transcript, analysis is complicated by the presence of the full-length and singly-spliced messages that can also bind probe. Probe A should protect a 443 base message if the message represented by the cDNA pWS1 exists in infected FEK cells. This band is present in RNase protection assays (Figure 2.14B, lane 4). The other bands represent protection of the other messages: the leader region, 213 bases; S1, 140 bases; leader-S1, 353 bases; part of env, 101 bases; and S3, 90 bases. *In vitro* transcribed RNA from pWS1 was used as a positive control (Figure 2.14B, lane 3); the protected fragment is 35 bases larger due to hybridization of plasmid sequences present in both the probe and the RNA. Probe B should protect a 544 base fragment if the RNA represented by pWS2 is present in infected FEK cells, and this band is present in RNase protection assays (Figure 2.14C, lane 10). Again, the extra bands represent protection of parts of the full-length and singly-spliced transcripts. *In vitro* transcribed RNA from pWS2 was used as a positive control (Figure 2.14C, lane 8); again, the protected fragment is larger due to plasmid sequences. These RNase T2 protection assays indicate that the messages containing the leader region spliced to S1 spliced to S3 and the leader region spliced to S1 spliced to env spliced to S3 do exist in FEK cells infected with EIAV.

pWS4 was also examined by RNase protection assays using a probe designed to confirm the splice site seen in this cDNA. Preliminary analysis of RNA from infected FEK cells indicates that this message is present (S. L. Payne, data not shown), but these studies were not repeated.
Figure 2.14. RNase T2 protection analysis of pWS1 and pWS2. A. Probe A hybridizes to the leader region, S1, and part of S3, and the presence of a 443 base fragment upon RNA digestion verifies the splicing events seen in pWS1. Probe B hybridizes to the leader region, S1, part of env, and part of S3, and the presence of a 544 base fragment verifies the splicing events seen in pWS2. These probes may also protect regions of the full-length and singly-spliced messages: the sizes of these predicted fragments are indicated. In vitro transcribed RNAs from pWS1 and pWS2 were employed as positive controls, and protected fragments with these samples will be slightly larger due to the presence of hybridizing plasmid sequences. B. and C. RNase protection analysis using Probe A (B) or Probe B (C). Poly A+ RNA, tRNA, or in vitro transcribed RNA was hybridized to the 32P-labeled probes overnight, and the resulting complexes subjected to RNase T2 digestion. The denatured fragments were fractionated on denaturing polyacrylamide gels, followed by autoradiography. Lane 1 contains Probe A alone; lanes 2-5 contain Probe A incubated with tRNA (lane 2), in vitro transcribed RNA from pWS1 (lane 3), poly A+ RNA from infected FEK cells (lane 4), and poly A+ RNA from uninfected FEK cells (lane 5). Lane 6 contains Probe B alone; lanes 7-10 contain Probe B incubated with tRNA (lane 7), in vitro transcribed RNA from pWS2 (lane 8), poly A+ RNA from uninfected FEK cells (lane 9), and poly A+ RNA from infected FEK cells. Markers are in nucleotides.
A.

<table>
<thead>
<tr>
<th></th>
<th>gag</th>
<th>pol</th>
<th>$S_1$</th>
<th>$S_2$</th>
<th>$S_3$</th>
<th>env</th>
</tr>
</thead>
<tbody>
<tr>
<td>pWS1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probe A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pWS2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probe B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Expected Sizes

- **Probe A**: 443 bases (+35 bases plasmid)
- **Probe B**: 544 bases (+35 bases plasmid)
- **leader**: 213 bases
  - **$S_1$**: 140 bases
  - **$S_2$**: 101 bases
  - **$S_3$**: 90 bases
- **leader-$S_1$**: 353 bases
Analysis of the Splicing Event of the env Transcript  To examine the splicing event necessary to create the singly-spliced env message and to see if this transcript utilizes the splice donor site at nucleotide 438, poly A+ RNA from infected FEK cells was reverse transcribed using oligo 5672 as a primer. The resulting first-strand cDNA was amplified with PCR, using oligos 228 and 5672, and the PCR product was cloned into pBS+ and sequenced. Oligo 5672 hybridizes in the env gene, close to the terminus of the S2 open reading frame, and was used to prime cDNA synthesis and for PCR as it should result in a cDNA clone that contains the splicing event seen in the production of env mRNA. A partial env cDNA was isolated as the entire 3.5 to 4.0 kb env message may not be easily reverse transcribed and efficiently amplified by PCR. Sequence analysis of several cDNA clones revealed that the cDNAs contain one splicing event: the leader region is spliced to S1 (nucleotide 438 to 5114) as seen in pWS1, pWS2, and pWS3 (Figure 2.15 and data not shown). As the splice donor site at nucleotide 438 is utilized in most of the cDNA clones isolated, this site has been designated the major splice donor site.
Figure 2.15. Schematic representation of the cDNA clone obtained from infected FEK cells to examine the splicing event of the env mRNA. Sequencing confirmed that the 5' splice donor site seen in pWS1, pWS2, and pWS3 is utilized in the env message to splice the leader sequence to S1 (data not shown).
Analysis of the Leader-S1 Open Reading Frame  While the leader region spliced onto S1 contains no AUG to initiate translation, it does contain four alternative initiator codons (Figure 2.10). To examine the function of this gene and its gene product, we subcloned the leader-S1 region from pWS1 into the eukaryotic expression vector pSVΔCAT to create pSV-tat. As the leader-S1 open reading frame shows homology to the HIV-1 tat gene (Figure 2.16), we tested our clone for Tat activity. pSV-tat and pEICAT were transfected into FEA cells, and chloramphenicol acetyltransferase (CAT) activity was measured. pEICAT contains the CAT gene driven by the EIAV LTR (Dorn and Derse, 1988). Transfection of pEICAT into uninfected FEA cells results in low CAT activity as the basal transcription rate from the EIAV LTR is very low. Transfection of pEICAT into EIAV-infected FEA cells results in high CAT activity, due to the Tat activity that is able to act in trans on the EIAV LTR (Derse et al., 1987; data not shown). As shown in Figure 2.17, transfection of pSV-tat plus pEICAT into FEA cells results in high levels of CAT activity: 40- to 50-fold higher CAT activity is seen than that produced with pSVΔCAT plus pEICAT. This result suggests that the leader-S1 region encodes a functional Tat gene, and that translation of this gene initiates at a codon other than AUG. Other groups have also cloned this gene and determined that it encodes Tat activity (Derse et al., 1989; Dorn et al., 1990; Noiman et al., 1991; Yaniv et al., 1989). The construct pSV-tat is currently being employed in the laboratory to study the transactivation potential of variant LTRs found in different strains of EIAV (S. L. Payne, personal communication).
Figure 2.16. Amino acid sequences of HIV-1 and EIAV Tat proteins. Domains of the proteins are marked. Region 1 represents the amino-terminal domain; region 2 contains the cysteine-rich domain; region 3 is the core domain; region 4 is the basic domain; and region 5 is the glutamine-rich carboxy-terminal domain. The figure is adapted from Carroll et al., 1991, and Derse et al., 1991.
HIV-1 | MEPVDRLEPWWKHPGSQPKTA
EIAV  | MADRRIPGTAEENLQKSSSGVPGQGNTGGQEARPN

HIV-1 | CTNCYCKKKCFHCQVCFITKALGISYY
EIAV  | YHCQLCFLRSLGIDY

HIV-1 | GRKKRRQRRRASHNQSQTHQASLSKQ
EIAV  | LDA SLRKNNKQRLKAXQQGQQPOQYLL

region 1
region 2
region 3
region 4
region 5
Figure 2.17. Analysis of Tat activity of pSV-tat. FEA cells were cotransfected with pSV-tat and pEICAT. After 72 hours, cells were lysed and examined for CAT activity. Data shown is from one experiment; other experiments gave similar results. pSV2CAT is a eukaryotic expression plasmid containing the CAT gene driven by the SV40 early promoter and was used as a positive control. pSV = pSVΔCAT.
Analysis of the S2 Open Reading Frame  No mRNA capable of specifically encoding the S2 open reading frame was cloned, although the singly-spliced env transcript contains the S2 open reading frame and can produce S2 protein in vitro (Schiltz et al., 1992). To study the function of S2, we subcloned the S2 region into the eukaryotic expression vector pSVΔCAT to create pSV-S2 and into the bacterial expression vector, pDS56/RBSII.6xHis to create pH52. Protein expression from pH52 was confirmed by a small-scale induction of E. coli containing this plasmid, as outlined in Materials and Methods, followed by electrophoresis on an SDS-PAGE gel and Coomassie blue staining, or by Western blot analysis using a polyclonal antibody raised to an S2 peptide. Figure 2.18 shows that a protein of the expected size (approximately 7 kDa) is produced in the induced cultures. Both the plasmid pSV-S2 and purified S2 protein (produced by S. L. Payne) were then sent to our collaborators at North Carolina State University for further study (I will discuss their results below, p.75).
Figure 2.18. Induction of pHS2. *E. coli* cultures containing pHS2 were induced with IPTG for three hours and lysed. Samples were fractionated on 15% SDS-PAGE gels and either stained with Coomassie blue (A) or examined by Western blot analysis using an antibody raised to an S2 peptide (B). The S2 protein is indicated by the arrow. Markers are in kDa. U., uninduced. I., induced.
Analysis of the env-S3 Open Reading Frame  It was postulated that the env-S3 region encodes EIAV Rev as it shows limited amino acid homology to HIV-1 Rev in the arginine-rich region responsible for RNA binding and nuclear localization (Figure 2.19), and the positions of its exons are in analogous genetic locations to that of visna virus rev (Tiley et al., 1990) (Figure 2.20). It has also been shown that cDNAs isolated from EIAV-infected single cell clones that display a rev-defective phenotype contain mutations in the S3 reading frame (Stephens et al., 1990). To determine the function of env-S3, we wished to create a mutation in the S3 open reading frame in the context of an infectious proviral clone and to examine viral replication from this mutated clone. If the env-S3 open reading frame encodes EIAV Rev, there should be no viral growth from the mutated clone, and it should be possible to complement this mutation by co-transfection of a eukaryotic expression vector that expresses env-S3. We wanted to perform these experiments in the context of a proviral clone rather than constructs containing portions of the genome, as studies performed on HIV-1 subgenomic constructs sometimes yielded misleading results. To this end, a single base mutation was made in the S3 region (Figure 2.5) in the context of an infectious proviral clone, creating p19S3stop. This mutation creates a stop codon in the beginning of the S3 open reading frame without changing the amino acid in the overlapping env reading frame. Restriction digest analysis of p19S3 stop and sequence analysis of most of the env gene of this clone reveal no other mutation. Transfection of p19S3stop into FEA cells resulted in little viral replication as measured by reverse transcriptase activity (Figure 2.21). This result would be expected with the truncation of an essential protein; however, we have been technically unable to complement this phenotype with co-transfection of pSV-rev and so cannot conclusively conclude that S3 is responsible for Rev activity. Technical problems included the low basal transcription rate of the EIAV LTR, lower Tat-activated transcription in FEA cells, and low transfection efficiencies. These problems lead to levels of viral products (RNA and proteins) undetectable by standard methods and result in undetectable levels of viral replication.
<table>
<thead>
<tr>
<th>EIAV</th>
<th>MAKSKE</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visna</td>
<td>MAKSKE</td>
<td>5</td>
</tr>
<tr>
<td>HIV-1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EIAV</th>
<th>ARDQEMNL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visna</td>
<td>SKPSRTTWRDMEPPLETWNQVLQELV</td>
</tr>
<tr>
<td>HIV-1</td>
<td>32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EIAV</th>
<th>KESKEEKRRNDW KKIDPQG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visna</td>
<td>KRQQQDEEEQQGGLVSGLQAASKADQ</td>
</tr>
<tr>
<td>HIV-1</td>
<td>MAGRSGDSEELRTVRLIKLLYQS</td>
</tr>
<tr>
<td></td>
<td>25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EIAV</th>
<th>PLESNQWCRVLRQRSLPGEKIPSQT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visna</td>
<td>IYT</td>
</tr>
<tr>
<td>HIV-1</td>
<td>DPPP</td>
</tr>
<tr>
<td></td>
<td>29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EIAV</th>
<th>LGPGPTQQP</th>
<th>NRRDRWIRGQI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visna</td>
<td>GNSGDRSTGGIGGKTKKKRGWYKW</td>
<td></td>
</tr>
<tr>
<td>HIV-1</td>
<td>NPEGTRQRARRRRRWERQ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>85</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EIAV</th>
<th>LQPEVLQERL</th>
<th>EWRIRGVQQAA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visna</td>
<td>LRKLRAEREKNIPSQFYPDMESNMGME</td>
<td></td>
</tr>
<tr>
<td>HIV-1</td>
<td>RQIHSISERILSTYLGRSAEPVP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>106</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EIAV</th>
<th>KELGEVNRGIEWRELYFREDQRGDFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visna</td>
<td>NLTLETQLEDNALYNPATHIGDMA</td>
</tr>
<tr>
<td>HIV-1</td>
<td>LQLPLLRLTLDCNEDCGTSGSTQGV</td>
</tr>
<tr>
<td></td>
<td>134</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EIAV</th>
<th>AWGGYQRERLWGEQSSPRVLRPGDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visna</td>
<td>MDGREWMEWRESAQ</td>
</tr>
<tr>
<td>HIV-1</td>
<td>GSPQILVESPTVLESGTKE</td>
</tr>
<tr>
<td></td>
<td>148</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EIAV</th>
<th>KRKKHLL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visna</td>
<td>KKKRKGGLSQRNTAYPGK</td>
</tr>
<tr>
<td>HIV-1</td>
<td>156</td>
</tr>
</tbody>
</table>

---

**Figure 2.19.** Amino acid sequences of EIAV, visna virus, and HIV-1 Rev proteins. Homologous regions are in bold italics.
Figure 2.20. Schematic representation of the mRNAs encoding visna virus Rev and EIAV putative Rev proteins. Both Rev proteins would be encoded by exons 3 and 4 of their respective messages (Tiley et al., 1990).
Figure 2.21. Growth curve analysis of p19S3stop versus pSPEIAV19. p19S3stop, which contains a single base mutation in S3 that creates a stop codon, and a wild type proviral clone, pSPEIAV19, were transfected into FEA cells, and culture supernatants were collected and analyzed for virus by reverse transcriptase assays.
To attempt to complement the S3 mutation in another manner, stable Cf2Th cell lines producing env-S3 protein were created via co-transfection of pSV-rev and pY3 and selection with hygromycin. Low levels of putative Rev protein were detected by a monoclonal antibody that recognizes the epitope gp90A of Env. As there are two possible AUGs at which the translation of env-S3 may begin, this Western blot analysis indicates that in vitro, env-S3 begins at the first AUG, which is in better homology with the Kozak consensus sequence (data not shown). Transfection of these lines with p19S3stop did not result in virus replication, but it is possible that the failure to complement results from levels of env-S3 that are too low to allow viral replication. HIV-1 Rev must reach a certain threshold level in order to have effect (Pomerantz et al., 1992).

DISCUSSION

To characterize the S1, S2, and S3 open reading frames of EIAV, we have examined the pattern of transcription of virus in infected cells. cDNA clones representing both the multiply-spliced messages and the singly-spliced env message have been isolated and sequenced. Functional studies of the proteins encoded by the small open reading frames have been initiated.

Transcription Patterns We have examined, by Northern blot analysis, the steady state levels of viral transcripts in four cell types: FEK cells (Figure 2.6), Cf2Th cells (Figure 2.6), FEA cells (Figure 2.7), and primary equine macrophages (Figure 3.11, p.114). The three classes of transcripts, full-length (about 8 kb), singly-spliced (about 4 kb), and multiply-spliced (about 1.5 kb), common to lentiviruses were detected in all four types of cells upon infection with either the PV strain of virus or virus produced from pSPEIAV19. Both of these strains of virus retain the ability to replicate in tissue culture and in the animal. Some early studies of EIAV transcription suggested that multiply-spliced messages were either not present or were at low abundance during EIAV infection of FEK cells (Rasty et al., 1990) or of FEK cells, FEA cells, Cf2Th cells, or primary horse macrophages (Stephens et al., 1990). The ability to detect an abundance of multiply-spliced transcripts in our studies may reflect the use of a single-stranded probe in the Northern blot analysis, which may be better able to detect low
abundance messages than the double-stranded DNA fragments used as probes in the studies by Rasty et al. and Stephens et al.

Studies by Dorn et al. and Noiman et al. corroborate our findings of multiply-spliced viral transcripts in infected cells. Dorn et al. examined EIAV-infected FEA cells with results similar to ours: three classes of viral messages were detected (Dorn et al., 1990). Noiman et al. examined the transcriptional pattern of infected Cfl2Th at great depth, using restriction fragments, riboprobes, or oligonucleotides as probes (Noiman et al., 1990b). This group detected a genomic transcript (8.2 kb), a low abundance 5.0 kb transcript which was detected only intermittently, a env transcript (4.0 kb), a 2.0 kb transcript, and a 1.8 kb transcript. Analysis using probes from different regions of the genome suggested that the 5.0 kb transcript may be generated through a single splicing event joining the leader region to an area in the 3' end of the pol gene. The 2.0 and 1.8 kb transcripts are probably generated by multiple splicing events. The 2.0 kb message hybridized with probes from the leader region, the 5' region of the S2 reading frame, and part of env. The 1.8 kb message hybridized with probes from the leader region, the S1 reading frame, and the S3 reading frame. This last transcript is suggestive of the cDNA clone pWS1 that also contains these regions of the genome. The results of Northern blot analysis of RNA from infected cells from our and other laboratories suggest multiply-spliced messages are a part of the EIAV replication cycle.

The temporal pattern of RNA production was also examined, as HIV-1 has two stages of RNA production (early and late) which are regulated by the HIV-1 Rev protein. Northern blot analysis of RNA prepared from infected FEA cells (Figure 2.7) indicates EIAV transcripts appear in an ordered manner with multiply-spliced messages appearing first, followed by singly-spliced and then full-length mRNAs. This order is the same as seen in HIV-1 infection (Kim et al., 1989). Early and late stages of EIAV RNA production in infected FEA cells are not as pronounced as they are in HIV-1 infections; however, this result may be cell-type specific. Examination of RNA production in the natural host cell of the virus, macrophages, may yield different results.
Examination of cDNA Clones  cDNA clones were generated from RNA prepared from infected FEK cells to characterize the sequences of and the splicing events that produce the multiply-spliced messages. Three different classes of cDNAs were isolated (Figure 2.9). The first, represented by pWS1 (Figure 2.10), is generated by two splicing events: the leader region is spliced to the S1 reading frame; after the stop codon of S1, there is a second splice to the S3 reading frame. This produces two open reading frames: leader-S1(tat) and S3. The message lacks, however, an AUG to initiate translation of either reading frame, although the leader region contributes four possible alternative initiator codons to the S1 open reading frame. If translation began at the second CUG (as indicated in Figure 2.10), the resulting protein would be 75 amino acids long. The splicing events necessary to generate this clone have been confirmed by RNase protection analysis of RNA from infected cells (Figure 2.14B). Three other groups have also reported the isolation of similar cDNAs, one using FEA cells (Dorn et al., 1990), one using C2Th cells (Noiman et al., 1991), and one using an acutely infected horse (Beisel et al., 1993). No differences in predicted amino acid sequence after the first possible initiator codon were found in the leader-S1 sequences of pWS1 and the published sequences (Dorn et al., 1990; Noiman et al., 1991).

The second class of cDNA clone isolated is generated by three splicing events: the leader region is spliced to S1; after the termination codon of S1, there is a splice to a region in the 5' end of env; and after 101 bases of env, there is a third splice to S3 (Figure 2.9). There are two open reading frames in these cDNAs: leader-S1(tat) and env-S3 (rev). Again, while there is no AUG for translational initiation in the leader-S1 open reading frame, the leader region contains four alternative codons at which to begin translation. The env region of the env-S3 reading frame contributes 30 codons and two AUGs to this open reading frame. and if translation began at the first AUG, the resulting protein would be 165 amino acids long. This class of cDNA is represented by two clones, pWS2 and pWS3. pWS2, as shown in Figure 2.11, contains a deletion in the S3 region that switches the reading frame back to env and probably encodes an inactive env-S3 protein. pWS3, as shown in Figure 2.12, contains two base pair deletions which result in the premature termination and probable
inactivation of the leader-S1 protein. The env-S3 reading frame of pWS3 is intact. The authenticity of these clones was confirmed by RNase protection analysis (Figure 2.14C). cDNA clones similar to pWS2 and pWS3 have also been reported. Stephens et al., examining cDNA clones from Cf2Th cells, reported cDNAs which contained the splicing events noted here and had mutations causing frameshifts from the S3 reading frame to the env reading frame (Stephens et al., 1990). Four other groups isolated similar cDNAs, but with intact env-S3 regions as in pWS3 (Beisel et al., 1993; Noiman et al., 1990a; Noiman et al., 1991; Rosin-Arbesfeld et al., 1993). Comparison of the predicted amino acid sequence of env-S3 of pWS3 to the same region of published cDNAs showed similar proteins with six (Noiman et al., 1990a) or seven (Rosin-Arbesfeld et al., 1993) amino acid differences.

The third class of cDNA clone isolated from infected FEK cells is represented by pWS4, which is generated by one splicing event and contains 73 amino acids of the gag MA protein fused to 98 amino acids of the env TM protein (Figures 2.9 and 2.13). Preliminary RNase protection analysis using a probe based on the splicing event in pWS4 suggests that this message is produced in infected FEK cells (S. L. Payne, data not shown); however, similar cDNA clones have not been reported. A published cDNA clone closest in sequence contains the leader region spliced directly to the splice acceptor site 5' to the S3 open reading frame. This message encodes a protein, designated Ttm, which initiates translation at the leader initiator codon and contains the carboxy-terminus of the transmembrane protein (Beisel et al., 1993). Further characterization of pWS4 was not performed.

Upon re-examination of the Southern blot analysis performed on the cDNAs amplified by PCR (Figure 2.8), it is possible to assign some of the hybridization bands to an isolated and sequenced cDNA clone. It is interesting to note, however, that the oligo 27S detects a 1.0 kbp band. This oligo hybridizes to the 5' region of the S2 open reading frame, and as yet, no cDNA clone representing multiply-spliced mRNA containing this region has been isolated. Noiman et al. also detected a small multiply-spliced mRNA (2.0 kb) using an oligo specific to the 5' region of S2 (Noiman et al., 1990b). It is tempting to speculate that there is a small multiply-spliced message that primarily encodes S2; however, the only message that has been isolated that is capable of producing a protein from the S2
open reading frame is the potentially tricistronic env message (see Translation of Multi-cistronic Messages, p.71).

**Major Splice Donor Site** To confirm the splicing event that generates the singly-spliced message encoding the Env precursor protein, cDNA clones were isolated that were primed in env and contain the splice junction (Figure 2.15). Sequence analysis of several cDNA clones indicated that the env mRNA is generated by a splice from nucleotide 438 in the leader region to nucleotide 5114 in S1 (data not shown). This is one of the same splice junctions seen in pWS1, pWS2, and pWS3, and was also confirmed by another group (Schiltz et al., 1992). Because the splice donor site at nucleotide 438 is utilized in most of the transcripts isolated, it has been termed the major splice donor site. A summary of the splice donor and acceptor sites used to produce EIAV messages is presented in Table 2.2.

**Translation of Multi-cistronic Messages** As all of the transcripts produced by EIAV contain the leader region with its alternative initiator codons for translation, all viral transcripts are necessarily polycistronic. It was originally postulated that non-AUG initiation codons were present in the published cDNAs because these cDNAs were isolated from tissue culture-adapted strains of EIAV. It was possible that the utilization of non-AUG codons for translation initiation arose during virus adaptation and attenuation in cell culture. However, as shown in Figures 2.10 and 2.11, the cDNAs isolated from cells infected with the PV strain of virus, which is capable of causing disease in horses, also lack an AUG for the translation of tat. Beisel et al. confirmed these studies by the isolation of cDNA clones from horses acutely infected with the Wyoming strain of virus (Beisel et al., 1993), and Derse et al. have shown that the replacement of the CUG that initiates translation with an AUG results in no significant increase in the level of transactivation in transient expression assays (Carroll and Derse, 1993; Derse et al., 1989). Studies have determined that the second CUG of the leader region (Figure 2.10) is the initiator codon (Carroll and Derse, 1993), and examination of the translation of the bicistronic mRNA encoding Tat and Rev (with a splicing pattern similar to pWS2 and pWS3) implicates a leaky scanning mechanism by ribosomes of the tat CUG that allows translation of the downstream Rev protein (Carroll and Derse, 1993). Schiltz et al. examined the translational potential
Table 2.2. Splice Sites

<table>
<thead>
<tr>
<th>Consensus splice donor</th>
<th>C/A A G ↓ G T A/G A G T</th>
</tr>
</thead>
<tbody>
<tr>
<td>major splice donor (nt. 438)</td>
<td>C A G G T A A G A</td>
</tr>
<tr>
<td>S1 splice donor (nt. 5255)</td>
<td>A A G G T T T G A</td>
</tr>
<tr>
<td>env splice donor (nt. 5516)</td>
<td>T A G G T A T G T</td>
</tr>
<tr>
<td>gag splice donor (nt. 662)</td>
<td>G A G G C C T T T</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Consensus splice acceptor</th>
<th>(T/C)_N C/T A G ↓ G</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 splice acceptor (nt. 5114)</td>
<td>A G T A T T G T T G C A G G</td>
</tr>
<tr>
<td>env splice acceptor (nt. 5416)</td>
<td>A T A A T A A T G A C A G T</td>
</tr>
<tr>
<td>env splice acceptor (nt. 7478)</td>
<td>C A T T T G A G A G A G A</td>
</tr>
<tr>
<td>S3 splice acceptor (nt. 7214)</td>
<td>C T C T C G C C T A A G A</td>
</tr>
</tbody>
</table>

Nucleotides matching the consensus sequence are underlined.
of the tricistronic env mRNA, which contains the leader region, S1, S2, and env. in wheat germ extracts and saw the production of all three proteins (Schiltz et al. 1992). Again, a leaky scanning mechanism is suggested as both the CUG of the leader region and the AUG of the S2 region are in weak Kozak sequence context as compared to the AUG of env. A summary of translation initiation sites is presented in Table 2.3.

**Tat** The leader-S1 open reading frame encodes Tat activity, as shown in Figure 2.17, and is able to increase the level of CAT activity produced from an EIAV LTR promoter by about 50-fold. Other groups have also determined that the leader-S1 region encodes Tat activity (Derse et al., 1989; Dorn et al., 1990; Noiman et al., 1991; Yaniv et al., 1989), and it has been shown that the essential regions of the protein are contained in the S1 domain (Dorn et al., 1990). The 29 amino acids from the leader region are dispensable for Tat function. As shown in Figure 2.16, HIV-1 Tat lacks the cysteine-rich domain of HIV-1 Tat that has been implicated in the formation of metal-linked dimers. EIAV Tat does have, however, the other regions common to lentiviral Tat proteins (as designated by Carroll et al., 1991), including the amino terminus (comprised mostly of leader region-derived amino acids), core region, basic region, and carboxy terminus (Carroll et al., 1991; Dorn et al., 1990). Both the basic and carboxy terminus regions are necessary to form the promoter recognition domain of EIAV Tat, and it has been proposed that the basic region binds the TAR RNA, while the carboxy region stabilizes this interaction (Carroll et al., 1991). The core region may be the activation domain of EIAV Tat, as shown by studies on EIAV/HIV-1 Tat chimeric proteins in which the EIAV core domain fused to the HIV-1 RNA recognition domain is able to transactivate an HIV-1 LTR-CAT construct (Derse et al., 1991). The EIAV TAR element has been localized to nucleotides +1 to +25 and is comprised of an imperfect inverted repeat that folds into a stem-loop structure (Carvalho and Derse, 1991).
Table 2.1. Translation Initiation Sites

<table>
<thead>
<tr>
<th>Kozak Consensus</th>
<th>A/G X X A U G G</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 (tat)</td>
<td>U A C C U G U</td>
</tr>
<tr>
<td></td>
<td>C U G U G G A</td>
</tr>
<tr>
<td></td>
<td>A A C C U G G *</td>
</tr>
<tr>
<td></td>
<td>A G G A U C C</td>
</tr>
<tr>
<td>S2</td>
<td>U A U A U G G</td>
</tr>
<tr>
<td>Env-S3 (rev)</td>
<td>A G C A U G G *</td>
</tr>
<tr>
<td></td>
<td>G A A A U G A</td>
</tr>
<tr>
<td>Env</td>
<td>A A C A U G G</td>
</tr>
<tr>
<td>Gag/pWS4</td>
<td>A A G A U G G</td>
</tr>
</tbody>
</table>

Nucleotides matching the consensus sequence are underlined. The codon at which translation could initiate is in bold, while the sequences at which translation has been shown to initiate for the Tat and Rev proteins are starred (*).
S2 As no cDNA clone capable of primarily encoding S2 was isolated, the genomic S2 open reading frame was subcloned into eukaryotic and bacterial expression vectors. S2 protein was shown to be produced upon induction of cultures containing the bacterial expression plasmid (Figure 2.18), and both the eukaryotic expression plasmid, pSV-S2, and the S2 protein were sent to our collaborators in North Carolina, who have initiated studies on S2.

Although there have been hints that the S2 open reading frame may be encoded on multiply-spliced mRNAs (Figure 2.8; Noiman et al., 1990b), the only message isolated to date capable of producing S2 protein is the tricistronic env message, which encodes Tat, S2, and Env proteins (Schiltz et al., 1992). It is interesting to note that HIV-1 Vpu protein is encoded on a bicistronic message that also encodes Env, and as yet, no mRNA capable of encoding HIV-1 Vpu alone has been isolated. It is postulated that coordinate regulation of HIV-1 Vpu and Env proteins may be necessary for proper viral assembly and release. Perhaps coordinate regulation of S2 and EIAV Env proteins is also necessary.

Our collaborators in North Carolina (Maureen T. Flaherty, Stephanie T. Perry, and Frederick J. Fuller) have created mutations in the S2 open reading frame in the context of an infectious proviral clone that prematurely truncate the protein, while leaving the overlapping env reading frame intact. Transfection of these mutated clones into Cf2Th cells results in the production of virus, and infection of equine dermal cells and equine macrophages with these viruses also yields virion production, but at a rate slightly slower than WT. Examination of infected macrophages by electron microscopy indicates that the viral particles produced from the mutated S2 viruses are the same morphologically as those resulting from WT infection, and there appears to be no defect in viral budding from the cells. The S2 protein was localized to the nucleus and to a perinuclear location by immunoperoxidase staining of macrophages. Whether the S2 open reading frame encodes EIAV Vpu activity or some other activity has not been determined, and assignment of a function to S2 may have to wait for the isolation of a virulent molecular clone of EIAV, followed by studies of the effect of S2 mutations in this clone in horses. Since S2 is apparently dispensable for viral replication in primary equine macrophage cultures, and since horses produce antibodies to S2-based peptides and recombinant proteins (S. L. Payne,
personal communication; Schiltz et al., 1992), the protein may play a role in viral infection that can only be elucidated by studying its function in vivo.

Rev To examine the function of the env-S3 open reading frame, we created p19S3stop, a mutated proviral clone which contains a single base substitution that prematurely terminates the S3 open reading frame without affecting the overlapping env reading frame (Figure 2.5). Upon transfection into FEA cells, this clone was unable to replicate (Figure 2.21), suggesting that the S3 reading frame is essential for viral replication. As we were unable to complement this mutation due to technical reasons that resulted in very low expression of viral products, it remained unclear whether the phenotype of no progeny production was truly a result of the S3 mutation.

The env-S3 region was subcloned into a eukaryotic expression vector to create pSV-rev. and stably transfected C2Th cell lines producing env-S3 protein were developed using hygromycin selection. Env-S3 protein production in these lines was confirmed by immunoprecipitation followed by Western blot analysis (data not shown). A monoclonal antibody that recognizes the gp90 epitope A was utilized in these studies. This epitope is located between the two AUGs of the env sequence, and recognition of the env-S3 protein by this antibody suggests that translation of this protein begins at the first AUG in this cell line.

Other groups have confirmed the env-S3 reading frame encodes Rev activity. Three groups have demonstrated Rev activity using reporter plasmids containing the CAT gene and the EIAV env gene as a source of the EIAV RRE (Fridell et al., 1993; Mancuso et al., 1994; Rosin-Arbesfeld et al., 1993). In the absence of Rev, CAT activity is low, but if Rev is present, CAT activity increases. Martarano et al. demonstrated the Rev activity of the S3 open reading frame more formally by using wild type and rev-defective proviral clones (Martarano et al., 1994). Rev was found to be necessary for the production of Gag proteins, and Northern blot analysis revealed that the presence of Rev increased the steady state levels of unspliced messages in the nucleus and promoted their export to the cytoplasm. While this group was able to complement the S3 mutations in the proviral clones with a eukaryotic vector encoding env-S3 with regard to RNA and Gag protein production, no
complementation resulting in virus replication was demonstrated. Martarano et al. also localized the EIAV RRE and found that two RREs were present in the env gene: one near the vicinity of the env exon of the tat-rev message and one near the 5' end of the S3 open reading frame. Each RRE retained the ability to function in the Rev-response; however, the Rev-response was more efficient if both RREs were present. To date, all other lentiviruses contain only one RRE. A novel regulation of alternative splicing by Rev was also detected that involves the regulation of the inclusion of the env exon seen in the tat-rev message. In the absence of Rev, the predominant multiply-spliced message produced is the four exon transcript encoding Tat and Rev. In the presence of Rev, the three exon message encoding Tat is predominantly produced. Alternative splicing resulting in different multiply-spliced messages as a function of Rev activity has not been reported for any other of the lentiviral systems. It is possible that the presence of two RREs correlates with this novel activity of EIAV Rev.

The activation domain of EIAV Rev has been localized to amino acids 32-55 (Figure 2.22) and is strikingly different in sequence than the leucine-rich activation domains of HIV-1 Rev or visna virus Rev. This polar domain is able to substitute for the activation domains of HIV-1 Rev and visna Rev, as shown by studies on chimeric Rev proteins, suggesting that EIAV Rev may interact with the same cellular factor as HIV-1 Rev or visna Rev but at a different site or that EIAV Rev may interact with a different factor in the same pathway affected by HIV-1 Rev and visna Rev (Fridell et al., 1993; Mancuso et al., 1994). Feline immunodeficiency virus (FIV) Rev has also been shown to contain a polar activation domain that can substitute for the HIV-1 Rev activation domain (Mancuso et al., 1994). Deletion of the six carboxy-terminal amino acids of EIAV Rev abolished Rev activity. As this region is rich in basic amino acids, it may be part of the RNA binding domain or nuclear localization signal of EIAV Rev (Martarano et al., 1994).

It should now be possible to show viral replication upon complementation of a rev mutation in a proviral clone with a eukaryotic expression vector expressing Rev. While Martarano et al. showed that the env-S3 reading frame encoded Rev activity by examining mRNA and protein production, there was never a formal demonstration of viral progeny production upon complementation of a rev- proviral
Leucine-rich activation domains

HIV-1       L P P L E R L T L D
Visna       M V G M E N L T L E

Polar activation domains

FIV         K A F K K M M T D L E D R F R K L F G S P S K D E Y T
EIAV        P Q G P L E S D Q W C R V L R Q S L P E E K I P

Figure 2.22. Activation domains of HIV-1, visna virus, FIV, and EIAV Rev proteins (Fridell et al., 1993; Mancuso et al., 1994).
clone with Rev expressed in trans (Martarano et al., 1994). Now that studies on Tat activity in
different cell lines and on the transfection efficiencies of the cell lines that EIAV can infect have been
completed (data not shown; S. L. Payne, Frederick Fuller, David Derse personal communications), we
can chose a cell line in which to perform the complementation that has the highest basal transcription
level, a high Tat activated transcription level, high transfection efficiency, and can be clonally grown -
the canine osteosarcoma cell line, D17. The low level expression of viral proteins from the weak EIAV
LTR promoter, coupled with the low transfection efficiencies of the cell lines we have used, have made
it impossible to achieve high enough levels of Tat and Rev proteins to allow detectable progeny
production. Transfection of D17 cells with p19S3stop, pSV-rev, and pSV-tat and/or the transfection of
Rev and Tat producing cell lines with p19S3stop should lead to viral replication, as measured by cell-
free reverse transcriptase activity, if the Rev activity supplied in trans complements the S3 mutation.
As Martarano et al. determined that env-S3 encoded Rev activity in D17 cells and were able to detect
RNA and protein production by standard Northern and Western blot analyses of these cells (Martarano
et al., 1994), transfected D17 cells should be capable of expressing high levels of viral proteins,
resulting in measurable virion production upon successful complementation of the rev mutation.

An interesting aspect of EIAV Rev function is its association with two RREs and its novel
regulation of alternative splicing involving the multiply-spliced transcripts. Future studies should
examine the Rev protein more closely to determine if there are specific amino acids or domain(s)
responsible for the alternative splicing activity. Similarly, the role of each RRE in alternative splicing
should be examined by mutational studies. Finally, it may be interesting to design a study that would
utilize both the polar activation domain (as in EIAV and FIV Rev proteins) and the leucine-rich
activation domain (as in the primate lentiviruses and visna virus Rev proteins) as probes to isolate the
cellular protein(s) with which Rev interacts. As the two types of activation domains are
interchangeable (Fridell et al., 1993; Mancuso et al., 1994), it is probable that the different Rev
proteins either affect the same cellular pathway or interact with the same cellular protein at different
sites. Identification of the cellular pathway affected by Rev could lead to new insights about cellular RNA splicing, transport, and association with polysomes.

**Summary** We have characterized the small open reading frames, S1, S2, and S3, of EIAV by examining their production and analyzing their function. Northern blot analysis has confirmed the presence of multiply-spliced messages in EIAV-infected cells, and cloning and sequencing of cDNAs representing these messages have resulted in the identification of three such mRNAs. The first contains two open reading frames, leader-S1 and S3, and has been shown to encode Tat activity. The second contains two open reading frames, leader-S1 and env-S3, and encodes Tat and Rev activity. The third contains one open reading frame, and its function is unknown.
CHAPTER THREE

INCORPORATION OF URACIL INTO VIRAL DNA CORRELATES WITH REDUCED REPLICATION OF EIAV IN MACROPHAGES

INTRODUCTION

Replication of Lentiviruses in Nondividing Cells. Replication in nondividing cells is a hallmark of the lentivirus group of retroviruses. Unlike oncogenic retroviruses, which require host cell proliferation, or more specifically passage through mitosis, to integrate their viral DNA and to produce progeny, lentiviruses can productively infect cells arrested in the cell cycle (Bukrinsky et al., 1992; Bukrinsky et al., 1993; Clements and Payne, 1994; Klevjer-Anderson et al., 1979; Lewis and Emerman, 1994; Lewis et al., 1992; Li et al., 1993; Roe et al., 1993; Varmus et al., 1977; Weinberg et al., 1991). The properties of lentiviruses that enable them to produce progeny in these cells have yet to be completely elucidated; however, studies on HIV-1 have implicated a viral core protein (Lewis et al., 1992) and have also uncovered active, ATP-based import of the HIV-1 preintegration complex into the nucleus, thus bypassing the need for the transient disassembly of the nuclear envelope upon passage through mitosis by the infected cell (Bukrinsky et al., 1992). A nuclear localization signal has been pinpointed in the HIV-1 matrix protein, which is a component of the preintegration complex, and mutations in this signal abrogated viral replication and integration of HIV-1 DNA in nondividing cells, while permitting these functions in dividing cells (Bukrinsky et al., 1993). Similar mechanisms for replication in nondividing cells may be true of all lentiviruses, although similar studies have not yet been reported for the nonprimate lentiviruses. Some nonprimate lentiviruses, such as visna virus, EIAV, and feline immunodeficiency virus (FIV), have been shown to encode dUTPase activity (Clements and Payne, 1994; Elder et al., 1992; Threadgill et al., 1993; Wagaman et al., 1993), and a study from our laboratory (Threadgill et al., 1993) has indicated that this activity plays an important role in the replication of EIAV in nondividing macrophages.
dUTPases are ubiquitous cellular enzymes that play a key role in nucleotide biosynthesis. They promote the hydrolysis of dUTP, generating dUMP, a precursor in the dTTP biosynthetic pathway (Figure 3.1), and keep the cellular ratio of dUTP to dTTP low (less than or equal to $10^{-5}$ in mammalian cells [Richards et al., 1986]). As most DNA polymerases can utilize both dUTP and dTTP (Bessman et al., 1958; El-Hajj et al., 1988; Focher et al., 1990; Mosbaugh, 1988; Shlomai and Kornberg, 1978), maintenance of the correct balance of deoxynucleotide pools, and of a low dUTP:dTTP ratio, is very important for fidelity in DNA replication. Uracil in DNA can arise either by misincorporation of dUTP by DNA polymerase into nascent DNA or by cytosine deamination in mature DNA. To repair the DNA, uracil-DNA glycosylase removes the uracil base, leaving an apyrimidinic site in the DNA which is then repaired through the actions of apyrimidinic endonuclease, DNA polymerase, and DNA ligase (Figure 3.2). If the level of dUTP in the cell is high, a vicious circle of excision and repair of the DNA can result, leading to strand breaks, strand exchanges, and eventually cell death due to chromosomal aberrations and overlapping gaps in the DNA (Richards et al., 1986).

*Escherichia coli* (*E. coli*) with dUTPase mutations (*dur* or *sof*) transiently accumulate small DNA fragments, probably due to this repair process, and show a higher than normal frequency of recombination (Tye et al., 1977; Tye et al., 1978). Uracil in DNA can also have an effect on DNA conformation and sequence specific binding of proteins to the DNA, as uracil lacks the methyl group of thymine (Richards et al., 1986). The substitution of thymine by uracil has been shown to decrease binding of several proteins to their specific sites in *in vitro* assays, including Fos-Jun complexes on the TPA responsive element (Risse et al., 1989), HeLa cell nuclear proteins to the cAMP responsive element (Verri et al., 1990), herpes origin binding protein to its origin of replication (site I of Ori*) (Focher et al., 1992), *E. coli* RNA polymerase on the $\lambda$ P$_R$ promoter (Dubendorff et al., 1987), and the Lac repressor to the lac operator (Goeddel et al., 1977).

The crystal structure of the *E. coli* dUTPase indicates it is a trimer (Cedergren-Zeppezauer et al., 1992), instead of a tetramer as first proposed (Shlomai and Kornberg, 1978). The *E. coli* dUTPase is a metalloenzyme containing Zn$^{2+}$ and has a Km for dUTP of 12 $\mu$M (Shlomai and Kornberg, 1978).
Figure 3.1. Pathway of dTTP biosynthesis. dUTPase is necessary to hydrolyze dUTP to provide dUMP for dTMP synthesis by thymidylate synthetase. The diagram is adapted from Kornberg, 1980.
Figure 3.2. Pathway of DNA repair. Uracil is excised from DNA by uracil-DNA glycosylase, leaving an apyrimidinic site in the DNA which is repaired through the actions of apyrimidinic endonuclease, DNA polymerase, and DNA ligase. The diagram is adapted from Stryer, 1981.
while mammalian dUTPases tend to be dimers or trimers with a preference for Mg$^{2+}$ and Km's for dUTP on the order of 0.1-8.0 μM (Caradonna and Adamkiewicz, 1984; Giroir and Deutsch, 1987; Hokari and Sakagishi, 1987; Mahagaokar et al., 1980b; Spector and Boose, 1983; Williams and Cheng, 1979). McGeoch has studied the amino acid sequences of a variety of dUTPases and proposed that the proteins are composed of five conserved motifs (Figure 3.3) (McGeoch, 1990). Amino acids from motifs 1-4 are thought to be critical for activity, while motif 5 may perform a subunit contact function (Cedergren-Zepezeauer et al., 1992; Threadgill et al., 1993).

**Viruses and dUTPase** Expression of cellular dUTPases correlates with the state of cellular differentiation: i.e. high in undifferentiated, dividing cells, low in terminally differentiated, nondividing cells (Duker and Grant, 1980; Mahagaokar et al., 1980a; Pardo and Gutierrez, 1990; Pri-Hadash et al., 1992; Spector and Boose, 1983). Thus, a virally encoded dUTPase activity could assist in the establishment of viral replication in nondividing host cells. Herpesviruses, poxviruses, some type B and type D retroviruses, and some lentiviruses have been suggested or shown to encode dUTPases (Broyles, 1993; Elder et al., 1992; Koppe et al., 1994; McGeoch, 1990; Pyles et al., 1992; Wagaman et al., 1993; Williams, 1984; Williams and Parris, 1987). While viral dUTPase activity is apparently dispensable for herpesvirus growth in cultured cells (Caradonna and Cheng, 1981; Fisher and Preston, 1986), mutation of the dUTPase in herpes simplex virus type 1 (HSV-1) results in reduced neurovirulence, neuroinvasiveness, and reactivation from latency in vivo (Pyles et al., 1992). Similarly, HSV-1 encodes a viral uracil-DNA glycosylase, which has been found to be dispensable in vitro, but essential in vivo for full neurovirulence, neuroinvasiveness, and reactivation from latency (Pyles and Thompson, 1994). Thus, HSV-1-encoded dUTPase and uracil-DNA glycosylase are apparently essential for full viral function in terminally differentiated, nondividing neuronal cells. While the poxvirus-encoded dUTPase has not yet been shown to be essential for viral replication (Broyles, 1993), a poxvirus-encoded uracil-DNA glycosylase has been shown to be essential for viral replication (Stuart et al., 1993). As poxviruses replicate in the host cell cytoplasm and rapidly shutoff most
Figure 3.3. Alignment of amino acid sequences of dUTPases from man, *E. coli*, visna virus, and EIAV.

Motifs 1-5 as defined by McGeoch are indicated (McGeoch. 1990).
cellular functions, a virally encoded dUTPase and uracil-DNA glycosylase may be necessary to prevent misincorporation of dUTP into viral DNA.

Among the lentiviruses, dUTPases have been identified in EIAV, visna virus, caprine arthritis-encephalitis virus, and FIV, and are encoded as part of the pol gene (Figure 3.4) (Elder et al., 1992; Clements and Payne, 1994; Threadgill et al., 1993; Wagaman et al., 1993). The target cell for each of these viruses is the macrophage, except for FIV which replicates in both macrophages and T lymphocytes. Interestingly, the dUTPase domain is missing from the primate lentiviruses, HIV-1, HIV-2, and the simian immunodeficiency viruses, which infect CD4+ T lymphocytes and macrophages. The kinetics of HIV-1 replication in macrophages differs from that seen in T lymphocytes, as titers of HIV-1 progeny produced from infected macrophages are generally lower than that seen with infected T cells. While infected T cells produce higher titers of virus for a short period of time and are killed by HIV-1 infection, infected macrophages are not killed and produce low levels of virus over a longer period of time (Collman and Nathanson, 1992). It is possible that EIAV dUTPase is required for replication in equine macrophages in order to produce the high levels of infectious virus seen during the febrile episodes of EIA, while HIV-1 does not need to produce such high titers of virus from macrophages and therefore can do without a virally encoded dUTPase. It has also been suggested that HIV-1 borrows dUTPase activity from the endogenous HERV-K retrovirus, which has been suggested to contain the pseudoprotease/dUTPase region (Elder et al., 1992; McClure et al., 1987; McIntosh et al., 1992). While it is not known if the putative dUTPase gene is expressed from HERV-K in T cells or macrophages, it is tempting to speculate that if it is expressed, HIV-1 may utilize this enzyme during viral replication.

**EIAV dUTPase** To confirm the identity of the EIAV dUTPase domain and to examine the effects of a mutation in EIAV dUTPase on viral growth, a 270 bp in-frame deletion was generated in the infectious proviral clone pSPEIAV19 (designated WT) (Payne et al., 1994) to produce the mutated proviral clone ΔDU (Figure 3.5). This deletion encompassed motifs 1-4 (as defined by McGeoch, 1990) of the putative dUTPase, and therefore should inactivate the enzyme. Examination of the gene products
Figure 3.4. Schematic representation of the genome of EIAV. The *pol* gene encodes protease (PR), reverse transcriptase (RT), a pseudoprotease/dUTPase (DU), and integrase (IN).
produced in E. coli from a pol construct containing the dUTPase deletion indicated that the deletion had not grossly disrupted protein processing or expression. A DNA-dependent DNA polymerase activity assay showed activity associated with the 66 kDa subunit of reverse transcriptase, indicating that the deletion did not affect the DNA polymerase activity of EIAV reverse transcriptase. The WT and ΔDU proviral clones were transfected into FEA cells, and virus collected. There was no significant difference in the amount or rate of progeny production between the two viruses upon transfection into FEA cells. Semi-purified virus was analyzed for dUTPase activity: under conditions in which the WT virus showed 96% conversion of dUTP to dUMP, the ΔDU virus showed only 17% conversion (which was no higher than the background levels). This result confirmed that the open reading frame between reverse transcriptase and integrase in the EIAV pol gene encodes a dUTPase, and that deletion of motifs 1-4 results in an inactive enzyme.

Infection of FER cells, which are undifferentiated, actively dividing cells, with WT and ΔDU viruses resulted in no significant difference in viral growth. However, the ΔDU virus replicated poorly, with less than 1% of WT levels of progeny produced, in equine macrophages, which are the primary host cell of EIAV and are terminally differentiated, nondividing cells. No cytopathic effects were seen upon infection of macrophages with ΔDU, while the WT virus caused extensive cell death. These results suggested EIAV dUTPase is required for efficient viral replication in macrophages, which may have very low levels of dUTPase to lend in trans to the ΔDU virus. These results have been published (Threadgill, D. S., W. K. Steagall, M. T. Flaherty, F. J. Fuller, S. T. Perry, K. E. Rushlow, S. F. J. LeGrice, and S. L. Payne. 1993. Characterization of equine infectious anemia virus dUTPase: growth properties of a dUTPase-deficient mutant. Journal of Virology 67:2592-2600).

In the present studies, we describe the growth phenotype of a second dUTPase-deficient virus (DUD71E), which contains a point mutation in the proposed active site of dUTPase, and present an investigation of the block(s) to virus replication in macrophages. We have looked at several steps in viral replication that could be affected by the loss of dUTPase activity, including viral DNA
Figure 3.5. Putative amino acid sequence of EIAV dUTPase, with motifs 1-5 as defined by McGeoch indicated (McGeoch, 1990). The amino acids deleted from dUDU are underlined, and the aspartic acid that was mutated to a glutamic acid in DUD71E is in bold italics (amino acid number 71).
production, integration, and transcription. Most of these results have been submitted for publication (Steagall, W. K., M. D. Robek, S. T. Perry, F. J. Fuller, and S. L. Payne. Incorporation of uracil into viral DNA correlates with reduced replication of equine infectious anemia virus in macrophages.).

MATERIALS AND METHODS

Cells and Viruses The feline cell line, FEA, and fetal equine kidney (FEK) cells were maintained as described in Chapter 2. Equine macrophages were prepared from whole blood by centrifugation through Ficoll-Hypaque and maintained in RPMI 1640 plus 10% fresh autologous horse serum (Threadgill et al., 1993). Virus stocks were prepared by transfection of molecular clones into FEA cells using the calcium phosphate precipitation method (GIBCO/BRL). Clarified supernatants from transfected FEA cultures were used as stocks for infection of FEA cells, FEK cells, and equine macrophages (Threadgill et al., 1993).

Methotrexate Studies FEK cells were grown in Eagle's minimal essential medium with Earle's salts and supplemented with 5% fetal bovine serum, antibiotics, and fungizone as described above, plus either 100 μM hypoxanthine; 100 μM hypoxanthine, 10 μM methotrexate, and 5 μM thymidine; or 100 μM hypoxanthine, 10 μM methotrexate, 5 μM thymidine, and 0.1 mM, 1.0 mM, or 5.0 mM deoxyuridine (Ingraham et al., 1986). Cells at 50% confluency were incubated with WT or ΔDU virus for 1 hour at 37°C, then were washed and refed with the proper supplements. Viral replication was monitored by reverse transcriptase assays.

Mutagenesis A KpnI-PstI fragment containing the dUTPase coding domain from clone pSPEIAV19 (Payne et al., 1994) was subcloned into pBS+ (Stratagene), and a point mutation introduced using the Doubletake Double-Stranded Mutagenesis Kit (Stratagene) following the manufacturer's suggested protocol. The mutagenic primer for the DUD71E mutation is 27-mer of the sequence 5'-TGT ATA TCC TTC TTC AAT TAT TCC TTC-3' (noncoding strand), with the base change in bold (a T to A change in the coding strand at base number 4120 of the EIAV proviral sequence [Genbank Accession
numbers M16575, M11337, K03334, and M14855). This represents an aspartic to glutamic acid substitution at amino acid number 71 of the putative amino acid sequence of the dUTPase (Figure 3.5). DNA sequence analysis confirmed the presence of the mutation, and the mutated dUTPase region was substituted into a plasmid containing the 5' end of the proviral clone, pSPEIAV19. This mutated 5' end was ligated via a unique SphI site to the 3' end of the proviral clone pSPEIAV19, and the ligation mix used to transfect FEA cells.

**dUTPase Assays** Virus-containing culture supernatants were underlayered with 20% glycerol and centrifuged for 2 hours at 25,000 rpm in an SW41 rotor. The resulting pellet was resuspended in 50 mM Tris-Cl, pH 8.0. Protein concentrations were determined with the Bio-Rad protein assay reagent, and equal amounts of protein were assayed for dUTPase activity as described by Williams (Threadgill et al., 1993; Williams, 1984). Reaction mixes contained 50 mM Tris-Cl, pH 8.0, 2 mM β-mercaptoethanol, 1 mM MgCl₂, 0.1% (wt/vol) bovine serum albumin, 2 mM p-nitrophenylphosphate, a total of 0.1 mM dUTP ([³¹H]dUTP at 50 μCi/μmole), and 3 μg protein. The reaction mixes were incubated at 37°C for 20 minutes, and reactions were terminated by spotting 20 μl onto DE81 filter paper and immediate washing in a solution of 4 M formic acid, 1 mM ammonium formate. Filters were further processed as described previously (Williams and Cheng, 1979).

**PCR Analysis of the Production of Viral DNA** Wild type and ΔDU viral stocks were treated with 2 μg/ml RNase-free DNase (Worthington) to remove any DNA from lysed cells in the stocks that could confound results obtained with the polymerase chain reaction (PCR). 90% confluent FEA cells were incubated with 1.2x10⁷ cpm of virus (as measured by reverse transcriptase activity) in the presence of 10 μg/ml polybrene. After 1 hour at 37°C, the virus was removed and the cells washed three times with media and refed. Equine macrophages were incubated with 500,000 cpm of virus (MOI about 1) for 1 hour at 37°C, washed, and refed. Following the methods of Zack et al. (Zack et al., 1990; see also Arrigo et al., 1989 and Zack et al., 1992), cells were lysed in urea lysis buffer (4.7 M urea, 1.3% SDS, 0.23 M NaCl, 0.67 mM EDTA, pH 8.0, 6.7 mM Tris-Cl, pH 8.0) at various times. Lysates were
extracted with phenol-chloroform (25:24:1 phenol:chloroform:isoamyl alcohol) and ethanol precipitated. Progress of reverse transcription at each of the time points was analyzed by PCR using the primers listed in Table 3.1. PCR reactions included approximately 1.8 μg DNA (amounts of DNA were first normalized by PCR amplification of a portion of the α-globin gene [data not shown]). 2.5 mM dNTPs, 5xPCR buffer (250 mM NaCl, 125 mM Tris-Cl, pH 8.0, 25 mM MgCl₂, 500 μg/ml bovine serum albumin), 50 ng cold primer, 30 ng 5'-3²P-labeled primer, and 2.5 units TAQ DNA polymerase (Boehringer Mannheim). Twenty-five cycles of PCR were performed (1 minute at 91°C for denaturation, 2 minutes at the appropriate temperature for annealing and extension). PCR products were analyzed by electrophoresis on 4% or 8% nondenaturing polyacrylamide gels, followed by autoradiography.

**Reverse Transcriptase Assays** Reverse transcriptase assays to quantitate viral replication in FEK cells (as shown in Figure 3.13) have been described in Chapter 2. Reverse transcriptase activity in equine macrophage culture supernatants (as shown in Figure 3.6) was determined by the method of Gregerson *et al.* (Gregerson *et al.*, 1988) with [³H]TTP. Macrophages were maintained in 12-well plates, and infections of approximately 10⁵ cells were performed in duplicate. At each time point assayed, virus was pelleted from 1 ml of culture supernatant, and the virus pellet was resuspended in 10 μl of buffer containing 50 mM Tris-Cl, pH 8.3, 20 mM dithiothreitol, and 0.25% Triton X-100 (Threadgill *et al.*, 1993).

In the exogenous and endogenous reverse transcriptase assays shown in Figure 3.8, culture supernatants were clarified by low speed centrifugation and virus was collected by centrifugation through 20% glycerol at 25,000 rpm for 2 hours in an SW28 rotor. Viral pellets were resuspended in 50 mM Tris-Cl, pH 8.0, and protein concentrations measured using the Bio-Rad protein assay reagent.

Exogenous reverse transcriptase (RT) assays contained 0.4 mg/ml virus, 0.25 optical density unit of poly(rA) poly(dT)₁₂-₁₈, 50 mM Tris-Cl, pH 8.0, 7.5 mM KCl, 2 mM DTT, 5 mM MgCl₂, 0.05% NP-40, 0.02% Triton X-100, 50 μM dUTP or dTTP, and 1.52 μM [³H]dUTP or [methyl-
<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>209</td>
<td>5'-GAG TCC CTT CTC TGC TGG GCT GAA AAG-3'</td>
</tr>
<tr>
<td>287</td>
<td>5'-GCG TAG GAT CTC GAA CAG ACA AAG TAG AGA</td>
</tr>
<tr>
<td></td>
<td>CAG GG-3'</td>
</tr>
<tr>
<td>311</td>
<td>5'-CGG GCG CCA GCG TAG GA-3'</td>
</tr>
<tr>
<td>351</td>
<td>5'-CCC CCC CCG TCG AGC GTA GGG TCT GCG CC-3'</td>
</tr>
<tr>
<td>1439</td>
<td>5'-TTG GCT CCC AGA AAT AGT ATA TAC A-3'</td>
</tr>
<tr>
<td>4519</td>
<td>5'-GCA TGT ATG TAT CCT GAA TTT GAC-3'</td>
</tr>
<tr>
<td>TNF A</td>
<td>5'-GCA AAG CTT AAG GGT CTC AC-3'</td>
</tr>
<tr>
<td>TNF B</td>
<td>5'-GCA AAG GCT CTT GAT GGC AG-3'</td>
</tr>
</tbody>
</table>
Reactions were incubated at 37°C for one hour, and two aliquots spotted on DE81 paper. One aliquot was washed twice in 2xSSC (1xSSC is 0.15 M NaCl plus 0.015 M sodium citrate), twice with 95% ethanol, and dried. Samples were counted in the presence of scintillation fluid (Cytoscant). Competition reactions contained 1.52 μM \(^3\text{H}\)dTTP and a total of 52 μM dUTP plus dTTP.

Endogenous RT assays contained 0.5 mg/ml virus, 0.1 M Tris-Cl, pH 8.0, 5 mM magnesium acetate, 50 mM NaCl, 10 mM DTT, 0.02% Triton X-100, 200 μM each dATP, dCTP, dGTP, and 38 μM dUTP or dTTP, and 1.52 μM \(^3\text{H}\)dUTP or [methyl-\(^3\text{H}\)]dTTP (Rice and Coggins, 1979). Reactions were incubated at 42°C for one hour and then treated as above. Competition reactions contained 1.52 μM \(^3\text{H}\)dTTP and a total of 40 μM dUTP plus dTTP.

Exogenous RT assays in which recombinant dUTPase (Robek et al., manuscript in preparation) was added in trans were performed essentially as described above except that the level of dUTPase activity present in WT virions was determined and recombinant dUTPase (in amounts yielding the units of activity indicated in Figure 3.8C) was added to RT assays containing ΔDU virions.

Uracil-DNA Glycosylase Assays Infected FEA cells or macrophages were lysed in extraction buffer (10 mM Tris-Cl, pH 8.0, 0.1 M EDTA, pH 8.0, 0.5% SDS), and DNA was prepared by the method of Sambrook (Sambrook et al., 1989). Reactions included 1.8 μg DNA, 2.5 mM dNTPs, 5xPCR buffer, 30 - 50 ng of each primer, 1 mM DTT, and 1.0 - 2.5 units uracil-DNA glycosylase (Boehringer Mannheim or Perkin Elmer). Primers to amplify a portion of the tumor necrosis factor gene (TNF A and TNF B) (Table 3.1) were added to provide an internal control for the PCR. Reactions were incubated for one hour at 37°C, heated 10 minutes at 95°C, and cooled on ice. 2.5 units of TAQ DNA polymerase were added, and PCR amplification proceeded as described above. PCR products were either analyzed by electrophoresis on 4% nondenaturing polyacrylamide gels (in which case one primer had been end-labeled with \(^{32}\text{P}\)-ATP prior to PCR) followed by autoradiography or on 1% agarose gels and transferred to nitrocellulose for Southern blot analysis. 3x10^5 cpm/ml each of a \(^{32}\text{P}\)-labeled 4.5 kb BamHI-BamHI fragment of an EIAV proviral clone containing gag and pol sequences generated
by random primed synthesis and 5'-32P-labeled TNF B oligo were used to simultaneously probe the membranes.

**Southern Blot Analysis** DNA, prepared as for the uracil-DNA glycosylase experiments, was fractionated by electrophoresis on 0.7% agarose gels and transferred to nitrocellulose by standard methods (Sambrook *et al.*, 1989). Membranes were subjected to UV-crosslinking, baked at 80°C for 2 hours, and prehybridized for 1-2 hours in 6xSSPE, 10xDenhardt's (1% Ficoll, 1% polyvinylpyrrolidone, 1% bovine serum albumin), 1% SDS, and 50 μg/ml denatured salmon sperm DNA. Blots were hybridized overnight in 6xSSPE, 1% SDS, 10% dextran sulfate, 100 μg/ml denatured salmon sperm DNA, and 2x10^6 cpm/ml of a 32P-labeled probe generated by random primed synthesis of a full-length EIAV proviral clone. Membranes were washed as follows: 2 times for 15 minutes at room temperature in 6xSSPE, 0.5% SDS, 2 times for 15 minutes at 37°C in 1xSSPE, 0.5% SDS, and once for 20 minutes at 68°C in 0.1xSSPE, 0.5% SDS, followed by autoradiography.

**RNA Preparation and Northern Blot Analysis** RNA was prepared either by the cesium chloride centrifugation method described by Sambrook (Sambrook *et al.*, 1989) or by the method of Chomczynski and Sacchi as modified by Xie and Rothblum (Chomczynski and Sacchi, 1987; Xie and Rothblum, 1991). In the latter method, infected FEA cells or macrophages were lysed in 1.9 M guanidinium thiocyanate, 11.9 mM sodium citrate, pH 7.0, 95 mM sodium acetate, pH 4.0, and 0.34% β-mercaptoethanol. Water-saturated phenol and chloroform-isoamyl alcohol (24:1) were added to 47.5% and 10%, respectively, and lysates were vortexed and placed on ice for 20 minutes. Nucleic acid in the aqueous layer was precipitated with an equal volume of isopropanol, washed twice with 70% ethanol, and resuspended in DEPC-treated dH2O.

Samples were fractionated by electrophoresis in formaldehyde-containing agarose gels in 1xMOPS (0.02 M MOPS (3-[N-morpholino]propanesulfonic acid), 5 mM sodium acetate, 1 mM EDTA; pH 7.0) with buffer circulation. The gels were treated with 0.05 N NaOH for 30 minutes, 0.1 M Tris-Cl, pH 8.0 for 30 minutes, and transferred to nitrocellulose in 10xSSC. The membranes were
subjected to UV-crosslinking, baked at 80°C for 2 hours, prehybridized in 1% bovine serum albumin, 0.4 M sodium phosphate, pH 7.2, 15% formamide, 1 mM EDTA, pH 8.0, and 7% SDS, pH 7.2 at 65°C for 1-2 hours, and hybridized overnight at 65°C with 3x10^6 cpm of probe. Probes used were either a 32P-labeled actin fragment generated by random primed synthesis or a single-stranded DNA probe generated by run-off PCR in the presence of α-32P-dCTP. This PCR-generated probe hybridizes to the region of the RNA 5' to the major splice donor site, thereby hybridizing to all EIAV mRNA (see Chapter 2). Membranes were washed 30 minutes at 50°C in 2xSSC, 1% SDS, 30 minutes at 50°C in 0.2xSSC, 0.1% SDS, and 30 minutes at 65°C in 0.2xSSC, 0.1% SDS, followed by autoradiography.

RESULTS

We have examined several stages in the viral life cycle that could be affected by the loss of dUTPase activity, including reverse transcription, integration, and transcription, using two dUTPase-deficient viruses. ΔDU EIAV, which contains a 270 bp deletion in the dUTPase domain, has been previously described (Threadgill et al., 1993). The second virus, DUD71E, contains a single amino acid substitution in the dUTPase domain and was constructed to verify the phenotype described for the deletion mutant. In the current studies, we have analyzed virus replication during the initial round(s) of infection and prior to the onset of extensive cytopathic effects in the wild type (WT) infected cultures. Replication was also characterized in the dividing feline cell line, FEA cells, in which the dUTPase-deficient virus replicates as well as WT virus.

DUD71E Virus To verify the results of previous studies on ΔDU EIAV (Threadgill et al., 1993) and for the studies described here, we created a proviral clone with a single base substitution in the dUTPase gene. Site-directed mutagenesis was performed as described in Materials and Methods to create an aspartic acid to glutamic acid substitution at amino acid number 71 of the dUTPase protein (Figure 3.5), and the resulting viral clone was designated DUD71E. Based on the crystal structure of the E. coli dUTPase, the aspartic acid at amino acid 71 of the EIAV dUTPase is predicted to be in the
active site of the enzyme, and a conserved amino acid substitution was made to verify this prediction without causing gross alteration of the protein structure.

DUD71E DNA was transfected into FEA cells for the production of viral stocks. dUTPase assays were performed on semi-purified virions to determine the effect of the single amino acid change on dUTPase activity. As shown in Table 3.2, neither ΔDU nor DUD71E has significant dUTPase activity as compared to WT virus. In other studies, we have demonstrated that recombinant DUD71E dUTPase has no detectable dUTPase activity using standard assay conditions (Robek et al., manuscript in preparation).

Virus containing the DUD71E mutation was used to infect macrophages to verify that the growth inhibition previously seen in ΔDU-infected macrophages was due solely to the lack of dUTPase activity and not due to unintended effects caused by the 270 bp deletion. As seen in Figure 3.6, the replication kinetics of DUD71E and ΔDU viruses in macrophages are essentially identical and are severely impaired as compared to WT virus (less than 2% WT levels of virus produced at 7 days p.i.).
Table 3.2. Viral dUTPase Activity

<table>
<thead>
<tr>
<th>Virus</th>
<th>dUTPase activity(^a) (% conversion)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>96.0</td>
</tr>
<tr>
<td>ΔDU</td>
<td>17.0</td>
</tr>
<tr>
<td>DUD71E</td>
<td>20.5</td>
</tr>
</tbody>
</table>

\(^a\) Values represent percent conversion of dUTP to dUMP plus PP, and are the averages of two experiments.
Figure 3.6. Growth of ΔDU, DUD71E, and WT EIAV in macrophages. Equivalent amounts of cell-free virus (based on RT units) were used to infect primary equine macrophage cultures. and virus replication was monitored by RT assays. DU=ΔDU.
Production of Viral DNA  The kinetics of viral DNA production by WT and ΔDU viruses were examined in both FEA cells and macrophages to determine if viral DNA synthesis is blocked during ΔDU virus infection. Cells were incubated with equivalent amounts of virus, as measured by reverse transcriptase activity, and lysed at various times post-infection. The polymerase chain reaction (PCR) was then used to analyze the progress of viral DNA production, using primer pairs designed to distinguish between the different stages of viral DNA synthesis (Figure 3.7A and Table 3.1). Primer pair 209 + 287 amplifies the R-U5 region of the long terminal repeat (LTR), which is the first region synthesized as minus strand DNA. As shown in Figure 3.7B, the first products of DNA synthesis are detected by 0 to 6 hours post-infection (p.i.) in macrophages infected with the ΔDU virus. Minus strand DNA synthesis is clearly detected by 24 hours p.i. in macrophages infected with either the ΔDU or WT virus. The detection of viral DNA at 0 hours p.i. in the ΔDU-infected macrophages may represent detection of partial reverse transcripts which have been shown to be present in HIV-1 virions (Lori et al., 1992; Trono, 1992). The primer pair 209 + 311 amplifies DNA containing the R region of the LTR and the primer binding site (PBS) and therefore detects the initiation and first synthesis of plus strand DNA. Primers 3439 + 4519 determine the progress of minus strand synthesis through the dUTPase coding region. Both products are detected by 24 hours p.i. in macrophages infected with either virus type. Finally, primers 209 + 351 amplify only completed or nearly completely viral DNA: this DNA is detected 24 to 72 hours p.i. in macrophages infected with either virus. The viral DNA in ΔDU-infected macrophages appears more abundant at all time points than DNA from WT-infected macrophages even after normalization of total DNA amounts by PCR amplification of a portion of the α-globin gene (data not shown): this may be due to the ability of the ΔDU virion to utilize dUTP as well as dTTP (see Figure 3.8) which could facilitate viral DNA production. The same analysis was carried out in FEA cells. Viral DNA synthesis was considerably faster in FEA cells than in macrophages, with first minus strand synthesis seen as early as 1 hour p.i. with both ΔDU and WT viruses. Initiation of plus strand synthesis occurred by 2 hours p.i. in FEA cells with both viruses: viral
DNA had been synthesized through the dUTPase region by 3 to 4 hours p.i.; and completed viral DNA was detected by 4 hours p.i. Therefore, although viral DNA synthesis is slower in macrophages than in FEA cells, viral DNA production initiates and proceeds analogously in ΔDU- and WT-infected cells. This result is also confirmed by Southern blot analysis which shows similar amounts of unintegrated, linear viral DNA at 48 (data not shown) and 72 hours p.i. (Figure 3.10A) in both ΔDU- and WT-infected macrophages. There is no significant difference in rate of viral DNA production between the ΔDU and WT viruses to account for the failure of the ΔDU virus to replicate in macrophages.
Figure 3.7. Analysis of viral DNA synthesis by ΔDU and WT EIAV in macrophages and FEA cells.

(A) Schematic representation of viral DNA production and the primer pairs used to distinguish the steps in reverse transcription. Primers 209 + 287 amplify the R-U5 region of the long terminal repeat (LTR) and thereby detect the initial synthesis of minus strand DNA. Primers 209 + 311 amplify DNA containing the R region of the LTR and the primer binding site (PBS) and so detect the initial synthesis of plus-strand DNA. Primer pair 3439 + 4519 determines the progress of minus strand synthesis through the dUTPase coding region, while primer pair 209 + 351 amplifies only completed or nearly completed viral DNA. The right side of the figure shows the predicted result of PCR amplification with the various primer pairs, where a "+" indicates amplification is expected. (B) PCR analysis of viral DNA from infected macrophages and FEA cells. Cells were incubated with equivalent amounts of either ΔDU or WT EIAV (based on RT units) and lysed as described in Materials and Methods at various times post-infection. DNA was amplified with the various primer pairs, and products were analyzed by electrophoresis on nondenaturing polyacrylamide gels, followed by autoradiography. Amounts of DNA were first normalized by PCR amplification of part of the α-globin gene (data not shown). The product seen upon amplification of ΔDU DNA with the primer pair 3439 + 4519 is 270 bp smaller than that seen with the WT DNA as the amplification occurs across the dUTPase coding region where the deletion was made. hrs. p.i., hours post-infection.
Detection of Uracil in Viral DNA  To determine if the EIAV reverse transcriptase (RT) could utilize uracil in the presence or absence of dUTPase, we performed both *in vitro* and *in vivo* studies. First, we examined the ability of RT to incorporate uracil in exogenous and endogenous RT assays. In the exogenous reaction, semi-purified virions were incubated with the exogenous template, \( poly(rA)_{12-18} \), and either \([^{3}H]dUTP \) or [methyl-\(^{3}H\)]dTTP. As shown in Figure 3.8A, \( \Delta DU \) virus incorporated \([^{3}H]dUTP \) into the exogenous template (incorporating an average of 1123 pmoles dUTP over three trials as compared to 1072 pmoles dTTP in control reactions), while the WT virus did not appreciably use dUTP (53 pmoles dUTP versus 1554 pmoles dTTP). In the endogenous reaction, the incorporation of dUTP or dTTP into the viral RNA template was measured. Again, the \( \Delta DU \) virus was capable of incorporating dUTP into a DNA product (35 pmoles dUTP versus 23 pmoles dTTP), while the WT virus was not (3 pmoles dUTP versus 91 pmoles dTTP). Competition experiments which contained \([^{3}H]dUTP \) and varying ratios of dTTP to dUTP were also performed using \( \Delta DU \) virions. As shown in Figure 3.8B, the EIAV RT, similar to other DNA polymerases (El-Hajj *et al.*, 1988; Focher *et al.*, 1990; Mosbaugh, 1988), does not show a strong preference for dTTP over dUTP. Finally, we determined the level of dUTPase activity present in WT virions and added recombinant dUTPase to the \( \Delta DU \) virions *in trans*. As shown in Figure 3.8C, we see little incorporation of dUTP into the template even at levels of dUTPase activity 5- to 40-fold lower than the level found in WT virions. This experiment suggests that the dUTPase activity in the WT virus keeps local concentrations of dUTP low so the deoxynucleotide is not available for the RT, which does not show a strong preference for dTTP as compared to dUTP, and also suggests that an intimate relationship between the dUTPase and other viral protein(s) may not be required.
Figure 3.8. (A) Exogenous and endogenous reverse transcriptase assays to determine if EIAV reverse transcriptase can utilize dUTP. ΔDU and WT virions were incubated with [\(^{3}H\)]dUTP or [methyl-\(^{3}H\)]dTTP and a poly(rA) poly(dT)\(_{12,18}\) substrate (exogenous reaction) or the viral RNA template (endogenous reaction). Aliquots were spotted on DE81 chromatography paper, washed, and subjected to scintillation counting. (B) Exogenous and endogenous competition reactions to determine if ΔDU RT has a preference for dTTP over dUTP. ΔDU virions were incubated with the templates as in (A) and various ratios of dTTP to dUTP. Aliquots were treated as in (A). (C) Exogenous RT assays containing ΔDU virions and recombinant dUTPase added in trans. WT virus contains approximately 1000 units of dUTPase activity. Samples were treated as in (A).
Since the ΔDU RT does incorporate uracil into DNA in *in vitro* assays, we examined viral DNA for the presence of uracil. DNA, prepared from infected FEA cells or infected macrophages, was incubated with uracil-DNA glycosylase (UDG), which specifically removes the uracil base from DNA, leaving apyrimidinic sites. The samples were heated, causing breaks in the DNA at the apyrimidinic sites. PCR was then performed to see if amplification of the DNA was abrogated by cleavage of the template. As shown in Figure 3.9, PCR amplification of ΔDU viral DNA from macrophages treated with UDG is decreased as compared to an untreated sample (compare lanes 2 and 1), while amplification of WT viral DNA remains the same regardless of UDG treatment (lanes 4 and 3). Similar results were seen when this assay was performed on several samples of DNA from several different macrophage infections (data not shown). The amplification of a portion of the tumor necrosis factor (TNF) gene serves as an internal control for the PCR, and while the intensity of this band does decrease in this experiment with UDG treatment, the viral band disappears with UDG treatment. As the primer pair used to amplify viral sequences amplifies a region of the dUTPase gene, the product seen in the ΔDU samples reflects the 270 bp deletion, yielding an 833 bp product, as compared to the 1103 bp product for the WT samples. A similar treatment of DNA from ΔDUΔ71E-infected macrophages also shows the presence of uracil in viral DNA (Figure 3.9): viral PCR product is not detectable when the DNA is first treated with UDG (lane 6), while a viral product is seen without treatment (lane 5). UDG treatment and subsequent PCR amplification of DNA prepared from ΔDU-infected FEA cells gave a signal similar to that seen with the untreated sample (Figure 3.9, lanes 9 and 10). The TNF internal control was not used for the FEA samples as the primers are based on equine sequence and do not amplify feline DNA. The results of UDG treatment, although not quantitative, strongly suggest the presence of uracil in viral DNA products synthesized by dUTPase-deficient viruses in macrophages.
Figure 3.9. Uracil-DNA glycosylase (UDG) analysis of viral DNA from macrophages and FEA cells. DNA from infected cells was treated with UDG, heated for 10 minutes at 95°C, and amplified by PCR with primers 3439 and 4519. Primers TNF A and TNF B were added to the reactions to provide an internal control for the UDG treatment and PCR. The product seen upon amplification of ΔDU DNA with the primer pair 3439 + 4519 is 270 bp smaller than that seen with the WT DNA as the amplification occurs across the dUTPase coding region where the deletion was made.
Integration The next step in the viral life cycle examined was integration of the viral DNA into the host chromosome. It is possible that viral DNA containing uracil would be subject to host repair mechanisms that would prevent integration. Southern blot analysis on DNA prepared from infected cells was performed to assay for integrated provirus. Integration of WT and ΔDU DNA occurs in both cell types, as can be seen by the presence of hybridization signal in the region of chromosomal DNA on the Southern blots (Figure 3.10). Quantitation by densitometry performed on the Southern blots reveals a slight decrease in integration of ΔDU viral DNA in macrophages as compared to WT viral DNA: the ratio of the integrated signal to the unintegrated signal is an average of 2.4 times lower with the ΔDU virus (with 5 sets of ΔDU and WT DNA analyzed). DUD71E viral DNA also integrates into macrophage DNA (Figure 3.10); again, quantitation by densitometry (of the Southern blot in Figure 3.10) indicates this level of integration is 2.3-fold lower than that of WT DNA at 72 hours p.i. To confirm the results of the Southern blot analysis of macrophage DNA and to rule out the possibility that the signal seen in the chromosomal band was due to trapping of unintegrated viral DNA in the genomic DNA, we designed a PCR-based assay (based on the assay in Lewis et al., 1992 and Lewis and Emerman, 1994) to examine integration of the ΔDU viral DNA. Using this method, we also detected integrated proviral DNA in macrophages infected with either the ΔDU or WT viruses (data not shown). Therefore, the ΔDU viral DNA does integrate into macrophage DNA, although it may do so less efficiently than the WT virus. However, this decrease in integration may not be sufficient to explain the overall decrease in viral replication seen with the dUTPase-deficient virus in macrophages.
Figure 3.10. Southern blot analysis of viral DNA from macrophages (A) and FEA cells (B). DNA from cells infected with ΔDU, DUD71E, or WT EIAV was analyzed by standard Southern blot techniques using a $^{32}$P-labeled proviral clone generated by random primed synthesis. a. integrated viral DNA. b. unintegrated viral DNA. c. nonspecific band. hrs p.i., hours post-infection.
Transcription  Steady state levels of viral RNA were examined by Northern blot analysis. It is possible that the presence of uracil in viral DNA may disrupt transcription of viral RNA or that host-mediated repair processes during or after integration may result in provirus that is no longer able to support transcription due to point mutations or recombination. RNA was prepared from infected macrophages and FEA cells, separated by electrophoresis on formaldehyde-containing agarose gels, and probed with actin-specific and viral-specific probes. As shown in Figure 3.11, viral transcripts are detected in both cell types; however, the steady state levels of the ΔDU transcripts are decreased as compared to WT. Quantitation by densitometry performed on a Northern blot from macrophages (shown in Figure 3.11A) indicates about a 25-fold decrease in the levels of ΔDU multiply-spliced messages as compared to that of WT. Levels of singly-spliced and full-length messages are even more severely decreased (approximately 85-fold and 311-fold, respectively; all samples were normalized by comparison of actin signals). This is a minimum estimate, as other Northern blots do not show detectable levels of ΔDU transcripts although the amount of actin hybridization in these samples is comparable to that seen with RNA from WT-infected macrophages (data not shown). Quantitation by densitometry of viral transcripts from FEA cells indicates a three-fold difference in levels of ΔDU messages as compared to that of WT (in analysis of two sets of RNA samples) (Figure 3.11B). Northern blot analysis of RNA from DUD71E-infected macrophages also indicates a decrease in the steady state levels of viral transcripts as compared to those of WT (data not shown). The decrease in the steady state levels of viral transcripts from macrophages infected with dUTPase-deficient viruses as compared to the levels seen in WT-infected macrophages likely accounts for the decreased virus production in this cell type.
Figure 3.11. Northern blot analysis of RNA from infected macrophages (A) or FEA cells (B). RNA was prepared from ΔDU- or WT-infected cells and analyzed by Northern blot analysis using either a viral-specific probe (top panels) or an actin-specific probe (bottom panels). a. full-length viral RNA. b. singly-spliced viral RNA. c. multiply-spliced viral RNA. d. actin RNA.
Methotrexate Studies As large numbers of macrophages are difficult to obtain, it is difficult to perform studies that require many cells as starting material, such as the measurement of dUTPase and dUTP levels. Macrophages, as terminally differentiated, nondividing cells, probably have very low levels of cellular dUTPase and therefore, higher ratios of dUTP to dTTP. Presumably, the ΔDU virus is unable to replicate in an environment containing high dUTP/dTTP levels and no dUTPase, either virally or cellurally derived. To attempt to mimic the intracellular environment of equine macrophages in an undifferentiated, dividing cell type that is more readily available, we studied viral replication in FEK cells treated with methotrexate, hypoxanthine, thymidine, and deoxyuridine. Studies have shown that treatment of cells with the proper combination of these reagents results in a dramatic increase in the level of cellular dUTP due to the inhibition of dihyrofolate reductase by methotrexate and the subsequent depletion of methylene-tetrahydrofolate, which is required as a methyl source for dTMP production (Figure 3.12) (Goulian et al., 1980a; Goulian et al., 1990b; Ingraham et al., 1986). As shown in Figure 3.13, addition of methotrexate, thymidine, and deoxyuridine to the cell culture results in at least a two-fold drop in the levels of replication of the ΔDU virus as compared to the WT virus and as much as a five-fold decrease with the addition of 5.0 mM deoxyuridine. While we have not yet quantitated the levels of dUTP in these cells, the failure of the ΔDU virus to replicate as well as WT in cells maintained under conditions shown to increase dUTP in other cell types (Goulian et al., 1980a; Goulian et al., 1980b; Ingraham et al., 1986; Sedwick and Laszlo, 1981) suggests that the higher levels of cellular dUTP and higher dUTP/dTTP ratios may account for the difference in replicative abilities of the two viruses.
Figure 3.12. Schematic representation of the effect of methotrexate. Methotrexate is a competitive inhibitor of dihydrofolate reductase, thereby blocking the regeneration of tetrahydrofolate and indirectly preventing the synthesis of dTMP from dUMP. The diagram is adapted from Stryer, 1981.
Figure 3.13. Growth of ΔDU and WT viruses in FEK cells treated with methotrexate. FEK cells treated with hypoxanthine (Hx), with hypoxanthine, methotrexate, and thymidine (HAT), or with HAT and different concentrations of deoxyuridine (dUrd), were infected with equivalent amounts of cell-free virus (based on RT units), and virus replication was monitored by RT assays. Data in this graph are from the eighth day post-infection. Control cells had no supplementation of media. DU=ΔDU.
DISCUSSION

DUD71E  The EIAV genome contains a region encoding a dUTPase, whose activity has been shown previously to be important for viral replication in the natural host cell, equine macrophages (Threadgill et al., 1993). Another dUTPase-defective virus, DUD71E, has been created to expand our studies of the function of the EIAV dUTPase, and the block(s) to the replication of dUTPase-defective viruses in macrophages have been investigated. The DUD71E virus was constructed to address concerns that the phenotype of poor replication of ΔDU virus in macrophages may be an artifact of the deletion itself, such as an unforeseen effect on RNA splicing or polyprotein processing. The replacement of an aspartic acid with a glutamic acid in the putative active site dUTPase results in a significant reduction in dUTPase activity as shown in Table 3.2. Analysis of viral replication in macrophages shows that DUD71E performs similarly to the ΔDU virus with replication levels at about 2% that of WT virus (Figure 3.6). Therefore, we conclude that the phenotype of poor viral replication in macrophages seen with the ΔDU virus is in fact due to the loss of dUTPase activity and not due to secondary effects of the deletion.

Production of Viral DNA  As a first step toward elucidating the block(s) to viral replication of dUTPase-deficient viruses in macrophages, the synthesis of viral DNA was examined by PCR in order to determine if ΔDU viral DNA production was blocked. As seen in Figure 3.7, both the ΔDU and WT viruses show similar patterns of DNA production in nondividing macrophages and in dividing FEA cells. In fact, viral DNA from ΔDU-infected macrophages appears to be more abundant than DNA from WT-infected macrophages at all time points. This increase in abundance may be a consequence of the ΔDU virion's ability to incorporate dUTP and dTTP into viral DNA: perhaps reverse transcription can proceed more rapidly with the use of both deoxynucleotides. Complete viral DNA can be detected in macrophages by 72 hours post-infection and by 4 hours post-infection in FEA cells after infection with either ΔDU or WT viruses. This difference in the time course of viral DNA production between cell types is probably not due to a difference in quantity of input virus, as a slightly
higher multiplicity of infection was used with macrophages than FEA cells. Instead, it may reflect differences in deoxynucleotide pools between actively dividing (FEA) and highly differentiated nondividing cells (macrophages). Studies on HIV-1 replication in macrophages indicate that production of viral DNA in these cells is slower than in stimulated peripheral blood lymphocytes and H9 cells, presumably due to the low levels of deoxynucleotide triphosphate substrates in the nondividing macrophages (Clements and Payne, 1994; Collin and Gordon, 1994; O'Brien et al., 1994). Southern blot analysis (Figure 3.10) also confirms that viral DNA synthesis by the ΔDU virus is proceeding at a pace similar to that of the WT virus in macrophages. As similar amounts of unintegrated viral DNA are seen at 72 and 120 hours p.i. Therefore, the block to ΔDU replication in macrophages is not due to a lag or block in viral DNA production.

**Utilization of dUTP** As one of the functions of a cellular dUTPase is to maintain a low ratio of dUTP to dTTP and thereby indirectly prevent DNA polymerases from incorporating uracil into nascent DNA, we decided to examine whether the EIAV RT could utilize dUTP in the production of viral DNA. *In vitro* RT assays in which the ΔDU virions were provided with [³H]dUTP indicate that the EIAV RT can utilize dUTP as well as TTP (Figure 3.8A-B). Studies in which recombinant dUTPase protein was added *in trans* to the ΔDU virions (thereby restoring WT function) (Figure 3.8C) suggest that the dUTPase protein is present to keep the local concentration of dUTP low to prevent incorporation of uracil into viral DNA by the EIAV RT, similar to the situation found with cellular dUTPases and DNA polymerases. This result also suggests that an intimate association between dUTPase and other viral proteins (i.e. RT) is not required. Examination of ΔDU or DUD71E viral DNA from macrophages by treatment of the DNA with uracil-DNA glycosylase followed by PCR amplification indicates this DNA contains uracil (Figure 3.9), while ΔDU DNA from FEA cells and WT DNA from either macrophages or FEA cells contain little to no uracil. These *in vitro* and *in vivo* studies indicate that the EIAV RT is capable of incorporating uracil into viral DNA. The most plausible explanation for higher levels of uracil in ΔDU viral DNA in macrophages than in FEA cells is that macrophages, as nondividing cells,
may have lower levels of dUTPase activity (and a higher ratio of dUTP to dTTP) than the proliferating FEA cells. Presumably, introduction of dUTPase activity into macrophages would restore WT activity to the dUTPase-deficient virus and decrease the amount of uracil in the viral DNA; however, these experiments are technically difficult due to the low transfection efficiency of macrophages, and thus are not feasible at this time.

**Integration** Integration was the next step in viral replication examined. As shown in Figure 3.10, the ΔDU and DUD71E proviruses integrate into macrophage DNA. Quantitation by densitometry indicates, however, that there may be a 2- to 3-fold decrease in the ratio of DNA integrated to that unintegrated in infections with the dUTPase-deficient viruses as compared to WT. Equivalent amounts of unintegrated viral DNA suggest that is is not an artifact of viability or infectivity of input virus. Both WT and ΔDU viruses integrated to an equivalent level in persistently infected FEA cells, indicating that the deletion did not affect viral integrase function. While this decrease in the amount of integrated viral DNA may contribute to the phenotype of growth inhibition by dUTPase-deficient viruses in macrophages, it may not be sufficient to account for the level of inhibition of viral replication seen.

**Transcription** Finally, the steady state levels of viral RNA were examined by Northern blot analysis.

As shown in Figure 3.11, the steady state level of ΔDU transcripts is decreased as compared to that of the WT virus in both cell types. In FEA cells, quantitation by densitometry indicates a decrease of about 3-fold in all viral transcripts; however, we cannot rule out that this difference is due to a slight difference in virus input or viability. In macrophages, quantitation by densitometry of the Northern blot shown in Figure 3.11A reveals about a 25-fold decrease in the level of multiply-spliced transcripts, an 85-fold decrease in singly-spliced transcripts, and a 311-fold decrease in the full-length message. The greater decrease in levels of singly-spliced and full-length messages probably reflects decreased levels of Tat and Rev proteins resulting from the lower level of multiply-spliced transcripts and is therefore indirectly due to the defect in dUTPase activity. Two other Northern blots on infected
macrophage RNA were performed (data not shown) in which no ΔDU viral transcripts were detected; therefore, the numbers above are minimal estimates of the decrease of the steady-state levels of viral transcripts. Northern blot analysis of DUD71E RNA from infected macrophages yields similar results, with a decreased level of viral transcripts as compared to WT. Therefore, these experiments suggest that one of the blocks to dUTPase-defective viral replication in macrophages is a severe decrease in the steady state levels of viral transcripts.

**Methotrexate Studies** We postulated that the dUTPase-defective viruses replicate poorly in macrophages because these cells, as terminally differentiated, nondividing cells, may contain low levels of dUTPase activity to lend *in trans* to these viruses. These cells likely have higher ratios of dUTP/dTTP than that found in dividing cells, where dUTPase is expressed coincident with other enzymes involved in deoxynucleotide biosynthesis and DNA synthesis. However, the measurement of the levels of dUTPase and dUTP in macrophages is technically very difficult due to the small numbers of cells available for the experiments. In order to test the growth of dUTPase-defective EIAV in a cell type that can be readily obtained and probably contains a high ratio of dUTP/dTTP, methotrexate studies were performed. It has been shown that the correct combination of methotrexate, hypoxanthine, thymidine, and deoxyuridine can increase the levels of dUTP in fibroblasts from undetectable levels (less than or equal to 0.3 femtoles/10^6 cells) to 139 pmoles/10^6 cells (using 100 μM hypoxanthine, 10 μM methotrexate, 5 μM thymidine, and 10 mM deoxyuridine; Ingraham *et al.*, 1986). This increase in dUTP results in a dUTP/dTTP ratio of about 1, while under normal conditions, the ratio is estimated at less than or equal to 10^5 (Richards *et al.*, 1986). We applied this method to FEK cells and infected them with ΔDU or WT virus. As shown in Figure 3.13, the addition of methotrexate, hypoxanthine, thymidine, and deoxyuridine to the FEK cells during infection results in at least a two-fold reduction in the growth of the ΔDU virus as compared to the WT virus and a five-fold reduction with the highest amount of deoxyuridine added. The lower level of viral growth of both viruses at the higher concentrations of deoxyuridine presumably results from the death of the cells, which are also affected
by the methotrexate treatment which has been shown to result in the incorporation of dUMP into the cellular DNA, leading to the accumulation of low molecular weight DNA (Goulian et al., 1980a; Sedwick and Laszlo, 1981). While we have not yet quantitated the amount of dUTP in the treated cells, it is likely that the level of dUTP was elevated as described in other studies. The decrease in replication of the ΔDU virus in the treated cells suggests that a cellular environment with a high ratio of dUTP/dTTP is not conducive to the growth of the dUTPase-defective virus. Future studies will attempt to measure the levels of dUTPase and dUTP in the treated HEK cells.

Role of dUTPase in EIAV Replication Our findings from the Southern and Northern blot analyses suggest the major block to dUTPase-defective viral replication in macrophages occurs between integration and viral transcription. The experiments performed here indicate viral DNA is integrated in macrophages infected with dUTPase-deficient viruses; however, we have not yet determined if the proviral DNA is intact or has suffered recombination or point mutations, which are a potential result of cellular repair of DNA containing uracil. Results of a single experiment in which UDG treatment and PCR amplification of integrated DNA were performed suggest uracil is also present in ΔDU integrated DNA (data not shown). It is possible that the presence of a viral dUTPase may be necessary to prevent the accumulation of point mutations and hyperrecombination which can result from repair processes on DNA containing uracil, as previously suggested by Elder et al. (Elder et al., 1992). It is also possible that transcription is decreased due to the presence of uracil in the viral LTR which may prevent binding of the appropriate transcription factors. Uracil, when substituted for thymine, has been shown to disrupt protein-DNA interactions in vitro reactions in a variety of systems, including Fos-Jun complexes on the TPA responsive element (Risse et al., 1989), HeLa cell nuclear proteins to the cAMP responsive element (Verri et al., 1990), herpes origin binding protein to its origin of replication (site I of Ori3) (Focher et al., 1992), E. coli RNA polymerase on the λ PR promoter (Dubendorff et al., 1987), and the Lac repressor to the lac operator (Goeddel et al., 1977).
Future Studies  The characterization of dUTPase function in macrophages continues with the examination of viral DNA produced in ΔDU-infected macrophages to determine if this DNA contains more mutations than DNA from WT-infected cells. Small regions of the EIAV genome will be amplified by PCR, cloned into pBS+, and sequenced. The difference in mutation levels between the ΔDU and WT clones should suggest if the presence of uracil in the ΔDU DNA leads to an increased rate of mutation and/or hyperrecombination, as is seen in dut E. coli (Tye et al., 1977; Tye et al., 1978). This study will determine if the bulk of the viral DNA in ΔDU-infected macrophages contains mutations; however, the bulk of viral DNA in EIAV infections is unintegrated DNA and is not transcribed. Should a higher level of mutation be found in ΔDU DNA, it may be necessary to show that this is also true of integrated ΔDU DNA. We have developed a PCR-based assay to determine if viral DNA is integrated into the host genome, and with this assay in a preliminary experiment, we have shown that uracil may also be present in integrated ΔDU DNA (data not shown). This experiment should be repeated to determine if uracil is present in the integrated ΔDU DNA, and then the PCR-based assay can be modified to allow cloning of LTRs from integrated DNA. These LTRs can be sequenced to determine if mutations are present. This approach will only allow examination of the LTRs; however, the detection of mutations in other portions of the genome would require isolation of proviruses from genomic libraries.

If the mutation levels seen in ΔDU DNA are the same as that in WT DNA, it is possible that the presence of uracil in the DNA has an effect on the ability of transcription factors to bind to the LTR. Again, it would be necessary to repeat the experiment showing that uracil is present in ΔDU LTRs from integrated DNA to determine if uracil is present to affect transcription. In vivo footprinting techniques could be utilized to determine if a transcription factor is present on the WT LTR that is absent from the ΔDU LTR in macrophages (Demarchi et al., 1993; Mueller et al., 1988; Mueller and Wold, 1989).
Finally, it would be interesting to determine the effect of the ΔDU mutation on disease progression in the horse. As low levels of infectious virus are produced in macrophage cultures infected with ΔDU, it is possible that an animal infected with ΔDU could become at least an asymptomatic carrier. These studies are not yet possible as there is no virulent clone of EIAV in which to construct the dUTPase mutation.

Summary We have examined the function of the dUTPase domain of EIAV by examining the replication of dUTPase-defective viruses in macrophages. These viruses replicate to less than 2% of WT levels in these terminally differentiated, nondividing cells, while replicating to WT levels in actively dividing cells. After examination of several steps in the viral life cycle, the major block to replication in macrophages appears to occur between integration and transcription. The viral DNA was shown to contain uracil, and the presence of the uracil may be responsible for the severe decrease in steady-state levels of viral RNA. EIAV may encode a dUTPase to prevent incorporation of uracil into viral DNA by reverse transcriptase, as dUTPase may be at low levels in the nondividing host cell of EIAV, the macrophages.
REFERENCES


McGeoch, D. J. 1990. Protein sequence comparisons show that the “pseudoproteases” encoded by poxviruses and certain retroviruses belong to the deoxyuridine triphosphatase family. Nucleic Acids Res. 18:4105-4110.


Verri, A., P. Mazzarello, G. Biamonti, S. Spadari, and F. Focher. 1990. The specific binding of nuclear protein(s) to the cAMP responsive element (CRE) sequence (TGACGTCA) is reduced by the misincorporation of U and increased by the deamination of C. Nucleic Acids Res. 18:5775-5780.


