Genetic and molecular characterization of the Rev protein of equine infectious anemia virus

Michael Andrew Belshan
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd
Part of the Microbiology Commons, Molecular Biology Commons, and the Veterinary Medicine Commons

Recommended Citation
Belshan, Michael Andrew, "Genetic and molecular characterization of the Rev protein of equine infectious anemia virus " (1999). Retrospective Theses and Dissertations. 12443.
https://lib.dr.iastate.edu/rtd/12443
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.

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.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI
800-521-0600
Genetic and molecular characterization of the Rev protein of equine infectious anemia virus

by

Michael Andrew Belshan

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Co-Majors: Molecular, Cellular, and Developmental Biology; Microbiology
Major Professors: Susan Carpenter and F. Chris Minion

Iowa State University
Ames, Iowa
1999
This is to certify that the Doctoral dissertation of

Michael Andrew Belshan

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Co-Major Professor

Signature was redacted for privacy.

Co-Major Professor

Signature was redacted for privacy.

For the Co-Major Program

Signature was redacted for privacy.

For the Co-Major Program

Signature was redacted for privacy.

For the Graduate College
Dedication

For my wife, Anne Belshan, who walked almost all of this long journey with me.

For my parents, Marianne and James, who gave me the tools to succeed.

For my Grandma Marie, who gave me inspiration.

In loving memory of my Grandpa Jesse.
# TABLE OF CONTENTS

## ACKNOWLEDGMENTS

### ABSTRACT

<table>
<thead>
<tr>
<th>CHAPTER 1. GENERAL INTRODUCTION</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissertation Organization</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>2</td>
</tr>
<tr>
<td>Overall Goal</td>
<td>12</td>
</tr>
<tr>
<td>References</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 2. BIOLOGICAL CHARACTERIZATION OF REV VARIATION IN EQUINE INFECTIOUS ANEMIA VIRUS</th>
<th>23</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>23</td>
</tr>
<tr>
<td>Introduction, Results, and Discussion</td>
<td>23</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>32</td>
</tr>
<tr>
<td>References</td>
<td>33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 3. SEQUENCE VARIATION IN EQUINE INFECTIOUS ANEMIA VIRUS REV CORRELATES WITH VARIABLE STAGES OF CLINICAL DISEASE IN AN EXPERIMENTALLY INFECTED PONY</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>45</td>
</tr>
<tr>
<td>Introduction</td>
<td>46</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>48</td>
</tr>
<tr>
<td>Results</td>
<td>53</td>
</tr>
<tr>
<td>Discussion</td>
<td>58</td>
</tr>
<tr>
<td>References</td>
<td>62</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 4. BINDING OF EQUINE INFECTIOUS ANEMIA VIRUS REV TO AN EXON SPlicing ENHANCER MEDIATES ALTERNATIVE SPlicing AND NUCLEAR EXPORT OF VIRAL mRNA</th>
<th>81</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>81</td>
</tr>
<tr>
<td>Introduction</td>
<td>82</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>84</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

First and foremost, I would like to thank my major professor, Susan Carpenter. Words cannot fully express my gratitude. She has been a truly great mentor. She taught me how to be a scientist, pushed me to my reach my potential, but was wise enough to release me when I was ready to work on my own. Without hesitation, she always provided me with opportunities that expanded my horizons, and gave me numerous intangible benefits. Plus, whenever necessary, she was always ready for a round of golf.

I am grateful to all the members of Susan Carpenter's laboratory, both past and present. Special thanks are devoted to Anne Shoemaker, Sean Murphy, Prasith Baccam, Yvonne Wannemuehler, and Greg Park, without whom this research would not be complete. My thanks to Eric Vaughn for helping get me started and fostering my interest in molecular biology.

I am also grateful to my collaborators in the laboratories of Tom Hope and C. Martin Stoltzfus, especially Matt Harris and Patty Bilodeau.

Finally, I would like to acknowledge the guidance of my program of study committee: W. Alan Miller, Michael McCloskey, F. Chris Minion, and Michael Wannemuehler.
Members of the lentivirus subfamily of retroviruses are characterized as causing slow, chronic disease. Atypical of other lentiviruses, such as human immunodeficiency virus type 1, equine infectious anemia virus (EIAV) may produce a rapid, variable disease course in horses. Infected horses may undergo an acute episode of disease involving viremia, fever, and thrombocytopenia. Following this acute period, horses may resolve to an inapparent infection or suffer numerous recurrent cycles of viremia, fever, and thrombocytopenia. Horses which survive clinical episodes usually become inapparent carriers of the virus for life. Numerous virus and host factors contribute to the phenotypic manifestations of disease. These include, but are not limited to, the rate of virus replication, the host immune response, and genetic variation of the virus. Genetic variation in EIAV has been identified in a region overlapped by the genes encoding the transmembrane protein and the trans-regulatory protein Rev. Rev is a nucleocytoplasmic transport protein which regulates the expression of viral structural proteins and progeny RNA molecules during the late phase of virus replication. Therefore, factors which modulate Rev activity may result in changes in virus replication and ultimately contribute to virus pathogenesis in vivo. The long term goal of this research project is to determine the contribution of Rev variation to the manifestation of clinical disease during EIAV infection. To accomplish this goal, we first demonstrated that genetic variation modulates Rev activity in vitro. We then characterized variation within Rev in an experimentally-infected pony. Functional analysis of the dominant Rev variants at various time points during infection indicated that the variants dominant during the clinical periods possessed more nuclear export activity than those dominant in the acute or aclinical periods. To further understand the significance of Rev variation, we have mapped the cis- and trans-acting elements required for the nuclear export and alternative splicing activities of
Rev. Specifically, we mapped the alternative splicing domain of Rev and identified the RRE of EIAV. These studies provide strong evidence that variation in Rev is an important determinant of virus replication and therefore a significant factor of EIAV pathogenesis.
CHAPTER 1. GENERAL INTRODUCTION

Dissertation Organization

This dissertation describes a genetic and molecular characterization of the Rev protein of equine infectious anemia virus (EIAV). The genetic component describes the biological significance of variation in Rev both *in vitro* and *in vivo*. There is also a molecular examination of Rev function, with an emphasis on the mechanism by which Rev mediates alternative splicing during virus replication. The dissertation is in the alternative format and is divided into six chapters, four of which are presented as manuscripts. Chapter 1 is a general introduction to EIAV as well as a background of the molecular biology of the Rev family of proteins, with a specific emphasis on EIAV Rev. The introduction provides general information to enable the reader to understand the following chapters and to justify the significance of the research. Chapter 2 describes the finding that amino acid variation in Rev alters biological activity. This chapter was published in the *Journal of Virology* as a short form article. This work was done in collaboration with the laboratory of Thomas J. Hope, which provided the pERRE-All, pERRE-1, and pERRE-2 plasmids. All of the work in this manuscript was performed by myself except the sequence data in Fig. 2. Data from this chapter lays the groundwork for the studies presented in the following chapter. Chapter 3 presents *rev* sequence data from EIAV recovered from an experimentally-infected pony and demonstrates that variation in Rev may be an important factor contributing to clinical disease. This work was done in collaboration with Prasith Baccam and J. Lindsay Oaks. Prasith Baccam performed the sequencing and genetic analysis of the *rev* clones and J. Lindsay Oaks performed the experimental infection, provided the clinical data and serum samples, and performed the semi-quantitative RT-PCR for detection of virus load. This chapter has been organized into a manuscript for submission to the *Journal of Virology*. 
Chapter 4 details a molecular characterization of Rev-mediated alternative splicing. This work was done in collaboration with the laboratory of C. Martin Stoltzfus, which performed all the in vitro splicing assays. The work in this chapter has been organized into a paper that has been submitted to the journal *Molecular and Cellular Biology*. The data in this chapter presents the finding that Rev binding to an RNA element in EIAV exon 3 mediates both nuclear export of incompletely spliced viral mRNAs and exon 3 alternative splicing. Chapter 5 presents domain mapping studies of Rev by deletion analysis. The data in this chapter identifies domains of Rev required for activity and maps a region of the protein required for alternative splicing. This chapter will be combined with other work from the Carpenter lab and submitted in the future to the journal *Virology*. Following chapter 5 is a general conclusion of these studies and their implication for studies of EIAV and retroviruses in general. Within the conclusion, I also present my recommendations for future research on EIAV Rev.

**Introduction**

Equine infectious anemia virus (EIAV) is the etiological agent of a persistent, lifelong, variable disease in members of the family *Equidae*. The clinical signs of disease include, but are not limited to: Anemia, fever, thrombocytopenia, and weight loss (reviewed in 20). EIAV is a member of the *Lentiviridae* subfamily of the virus family *Retroviridae*. Other members of this family include the human, simian, and feline immunodeficiency viruses (HIV-1 and -2, SIV, and FIV, respectively), bovine immunodeficiency virus (BIV), caprine arthritis encephalitis virus (CAEV), and visna virus. EIAV is mechanically transmitted from host to host via interrupted feeding of hematophagous arthropods (e.g. common horse flies) or artificially by instruments containing contaminated blood products. Development of a practical diagnostic test, the agar gel immunodiffusion (AGID) test, and
subsequent eradication measures have lowered the disease prevalence to < 5% in the United States of America. However, the genetic and biological similarities of the virus to other members of the Lentiviridae and the dynamics and pathology of the variable disease course have shown EIAV to be an important tool for comparative virology, cell biology, and pathology.

Clinical disease

EIAV infects cells of the macrophage/monocyte lineage (8, 37). It is now clear that the virus replicates predominantly in differentiated monocytes, especially tissue macrophages (8). Undifferentiated monocytes appear to be susceptible to viral infection, but do not produce infectious virions. The state of monocyte differentiation also appears to be an important determinant of virus replication. Early differentiating monocytes produce low levels of virus whereas the highest levels of replication are seen in the tissue macrophages of the spleen and liver. In rare cases, EIAV has been found to replicate in other tissues such as the lung, brain, muscle, and kidney. Virus transport throughout the host is believed to be facilitated by infectious immune complexes (7). Immune complexes are normally cleared from sera by macrophages in the liver (Kupffer cells) and spleen. These cells represent highly susceptible targets of EIAV infection.

Infection of horses with EIAV can result in an acute febrile illness within one to two weeks, followed by a period of recurring cycles of fever and thrombocytopenia, possibly accompanied by anemia and weight-loss. Each clinical episode is associated with high levels of virus replication and the emergence of new antigenic variants of virus (20, 27). The frequency of the cycles declines over time, with the majority occurring within the first year postinfection. Horses which survive the early clinical episodes typically become clinically inapparent, seropositive carriers of the virus for life. Thus, the majority of EIAV-infected
horses are healthy, inapparent carriers. Despite the virus' name, anemia is not a typical sign of infection. The mechanism of anemia has been attributed to both immune-mediated clearance of complement-coated erythrocytes and general bone marrow suppression (47).

To date, there is a number of conflicting reports concerning the mechanism of EIAV-induced thrombocytopenia. Thrombocytopenia associated with HIV-1 infection has been attributed to an increase of platelet-associated antibody, abnormalities in megakaryocyte (MK) development and subsequent platelet production, and a general decrease in platelet survival (55). However, EIAV does not infect MKs and no depletion of MKs is associated with EIA disease (7, 45, 52). Indeed, one study observed an increase in the number of MK cells in the bone marrow of infected ponies, accompanied by an enlargement in the overall area and nuclear size of MKs (45). Platelet-associated IgM and IgG have been observed to increase with EIAV infection (7). Therefore, it has also been proposed that thrombocytopenia may result from clearance of EIAV-platelet immune complexes in sera. However, another study demonstrated that severe combined immunodeficiency (SCID) foals infected with EIAV still developed thrombocytopenia, suggesting that the loss of platelets is not the result of an immune-mediated mechanism (52). Furthermore, histological examination of the bone marrow of the infected SCID foals also failed to show any change in MK numbers or morphology suggesting that immune mediated destruction of MKs does not occur (52).

Studies examining platelet production and function have also been conflicting. Whereas one study observed a suppression of platelet production by an unknown mechanism (52), another study indicated platelet production was not compromised (45). The latter authors also showed platelets isolated from EIAV infected ponies had decreased aggregation responses in vitro. Therefore, they suggested that increased activation of platelets in vivo resulted in quicker removal of platelets from circulation (45).
**Virus structure**

Retroviruses are single-stranded, positive-sense RNA viruses which possess the ability to synthesize dsDNA from ssRNA using a RNA-dependent, DNA polymerase, or reverse transcriptase (RT), encoded within the *pol* gene of the virus. Other proteins encoded by the *pol* gene include the viral integrase, a dUTPase, and a protease. The relatively small genome of EIAV (approximately 8.3 kb) also contains other features common to all retroviruses including *gag* and *env* structural genes and flanking long terminal repeats (LTRs) (Fig. 1). The *gag* genes products comprise the inner structural proteins of the virion including the nucleocapsid, capsid, and matrix proteins. The *env* gene encodes a polyprotein that is proteolytically cleaved by cellular factors into the transmembrane (TM) and surface (SU) glycoproteins of the virion. Both the TM and SU play important roles in virus entry. The SU contains the key determinants for binding of the cell specific receptor(s), and the TM facilitates fusion of the virus with the host cell membrane.

![Diagram of EIAV genome and splicing patterns](image)

**FIG. 1.** Splicing patterns of EIAV. Structural proteins (Gag, Env) and viral enzymes (Pol) are translated from the unspliced and singly-spliced mRNAs. Both Tat and Rev are expressed from the triply-spliced, four exon mRNA. In the presence of Rev, a new multiply-spliced mRNA appears in which exon 3 is skipped, producing a three exon species that codes for only Tat. Expression of the two exon mRNA is also dependent upon the presence of Rev. This mRNA codes for Ttm, a protein of unknown function (3).
Virus particles contain a core structure comprised of two copies of the ssRNA viral genome associated with the nucleocapsid protein, integrase, dUTPase, and reverse transcriptase. The RNA-protein complex is surrounded by the 26 kDa capsid protein. Encompassing the capsid is the matrix layer which couples the protein core of the virus with the surrounding membrane. The envelope of the virus is acquired from the host cell membrane during the budding process. The 45 kDa TM protein is embedded in the membrane, and the 90 kDa SU protein is superficially associated with TM on the outside of the membrane. Several other non-structural, regulatory proteins are also encoded in the EIAV genome including Tat, Rev, and Ttm. These will be discussed in detail in the replication section. The LTRs flanking either side of the viral genome possess several factors important for viral replication. Included in the LTR is the transcription start site and the viral gene promoter containing numerous transcription factor binding sites including PU.1, Oct, and CREB motifs (32). The LTR also contains cis-acting sequences necessary for reverse transcription of viral RNA and dsDNA integration into the host cell genome.

**Replication/gene expression**

The virus life cycle begins with attachment to the host cell. Attachment is mediated through interactions between the surface envelope protein and one or more cellular receptors. The cellular protein(s) required for EIAV attachment and internalization have not been delineated. However, as mentioned above, EIAV is thought to enter host cells by a fusion mechanism common to other retroviruses. Upon entrance into the cell, the capsid core of the virion disassembles and reverse transcriptase converts the viral ssRNA to circular dsDNA. Upon completion of reverse transcription, a pre-integration complex forms containing the viral dsDNA associated with the integrase and matrix proteins. The complex migrates into the nucleus and the dsDNA integrates randomly into the host cell genome. Once the DNA is
inserted, transcription initiates, promoted by cellular transcription factors and RNA polymerase II.

Since all gene transcription initiates at a single site in the LTR, EIAV uses several mechanisms to differentially express genes, including ribosomal frameshifting, production of polyprotein precursors, and alternative splicing. As a complex retrovirus, EIAV possesses two stages of gene expression: An early, non-structural regulatory phase, and a late, structural phase. The two phases are differentiated by the patterns of viral mRNAs expressed in the cell. Initial gene expression is characterized by a fully spliced, bicistronic mRNA which gives rise to two regulatory proteins, Tat and Rev (6, 50) (Fig. 1). Tat is a viral transcriptional activator which binds a secondary structure near the transcriptional initiation site in the newly synthesized viral mRNA (called the TAR element) and increases the efficiency of transcription (12). The regulatory protein Rev regulates expression of the late class of viral mRNAs. These RNAs are incompletely spliced, code for the viral enzymes and structural proteins, and act as progeny RNA molecules. Hence, Rev is absolutely required for virus replication. In the presence of Rev, two alternatively spliced mRNAs also appear— a three exon species encoding Tat and a two exon mRNA encoding Ttm, a 27 kDa protein of unknown function (3) (Fig. 1). This Rev-mediated alternative splicing phenomenon is unique among retroviruses and will be discussed in a later section.

During the late phase of virus replication, the viral core proteins assemble in the cytoplasm of the cell and viral envelope proteins are processed and transported to the membrane of the host cell. Assembly occurs at the host cell membrane, and virions bud from the cell in a non-lytic fashion. Packaging of the pol gene products occurs as large Gag-Pol polyprotein precursors because the individual enzymes lack motifs necessary for incorporation into the virion. Cleavage of the polyprotein precursor occurs after the virus is in the cell-free state. Studies of protease deficient viruses have demonstrated that virions with
intact Gag-Pol polyproteins are unable to infect host cells. Hence, until proteolytic cleavage has occurred, virus particles are considered "immature."

**Rev-mediated nuclear export**

The simplest retroviruses, such as murine leukemia virus, regulate alternative splicing by using suboptimal splice sites. Therefore, numerous mRNAs are produced from the same pre-mRNA at a fixed rate, allowing for production of all proteins/RNAs necessary for viral replication. Complex retroviruses, such as EIAV and HIV-1, utilize Rev-like pathways to differentially express incompletely-spliced mRNAs coding for virion structural and enzymatic proteins and progeny RNA molecules (reviewed in 11). The mechanism of action of the prototype member of this family, HIV-1 Rev, has been well characterized. HIV-1 Rev, a nuclear shuttling protein, facilitates export of incompletely spliced RNAs by interacting with the viral pre-mRNA and exporting it before complete splicing can occur. HIV-1 Rev interacts with the pre-mRNA at a specific sequence called the Rev-responsive element (RRE) (53). The HIV-1 RRE is a complex multi-stem-loop structure located near the SU-TM cleavage site (9, 10, 29). After a single Rev protein binds the RRE, additional Rev proteins multimerize (38) forming a nuclear export signal that interacts with the CRM1 nucleoporin pathway to transport incompletely spliced RNAs from the nucleus (4, 13, 49, 51). This non-mRNA pathway is normally important in the nuclear export of U small nuclear RNAs (snRNAs) (13, 49). Rev export pathways have been identified in many retroviruses including FIV, BIV, visna virus, CAEV, bovine leukemia virus (BLV), and human T-cell leukemia virus (HTLV) (11). Functional homology exists between all Rev-like proteins although they typically share little sequence similarity. The most conserved region among these proteins is the leucine rich nuclear export signal (NES), important for recognition of the CRM1
pathway. Indeed, in several cases, the NES of different Rev-like proteins have been shown to be interchangeable (14, 30).

EIAV Rev is functionally homologous to HIV-1 Rev, but remains less well characterized. EIAV Rev is a highly charged, 165 amino acid protein translated from exons 3 and 4 of a multiply spliced, four exon, bicistronic mRNA which also encodes Tat (6, 35). The amino acid sequence of the MA-1 strain Rev bears little homology to any other Rev-like proteins (Fig. 2). The NES of Rev has been mapped to amino acids 31-55 (14), and domain swapping experiments have shown it to be interchangeable with the Rev proteins of HIV-1 and visna virus. A recent report has proposed that the nuclear localization signal of Rev is in the arginine rich C-terminus of the protein (17). The RNA binding, multimerization, and alternative splicing domains of EIAV Rev have not been identified.

**Figure 2.** Amino acid sequence of the Rev protein of the MA-1 strain of equine infectious anemia virus. Location of the nuclear export signal (NES) is underlined.

**Rev-mediated alternative splicing**

Pre-mRNA splicing is a complex process which occurs in a large RNA-protein complex called the spliceosome. Numerous protein factors associate with the snRNAs U1, U2, U4, U5, U6, and the pre-mRNA substrate which results in removal of each intron via two trans-esterification reactions (reviewed in (34, 48)). Assembly of the snRNAs and protein factors is a highly ordered process which requires precise recognition of splice site and
branch point sequences. The presence of suboptimal recognition sequences can result in the inefficient splicing of certain exons, or alternative splicing. Use of suboptimal splice sites is mediated in part by cis-acting RNA sequences that either enhance or repress recognition of a splice site by the spliceosome. Exon splicing enhancers (ESEs) and silencers (ESSs) have been described for many virus and cellular RNAs (2). ESEs are typically purine rich sequences embedded within alternatively spliced exons. The purine rich sequences mediate exon recognition through interactions with members of the SR protein family of splicing factors. SR proteins are both essential splicing factors and regulators of alternative splicing (reviewed in 15). As essential splicing factors, SR proteins are required for early steps of spliceosome assembly (44). SR proteins regulate alternative splicing by binding ESEs and recruiting essential splicing factors to suboptimal splice sites near ESE sequences resulting in inclusion of alternatively splice exons. The effect of SR proteins on alternative splicing has been shown to be dose dependent, further suggesting SR proteins are important regulators of alternatively spliced mRNAs.

All retroviruses utilize alternative splicing to differentially express multiple genes from a single pre-mRNA. The complex retroviruses mediate expression of incompletely spliced mRNAs via Rev export pathways. In addition to promoting nuclear export of incompletely spliced RNA, EIAV Rev also regulates exclusion of exon 3 of the multiply spliced RNA (Fig. 1) (31). In the presence of Rev, a new multiply-spliced mRNA, lacking exon 3, is produced. Exon 3 is flanked by a suboptimal splice acceptor and contains a purine rich, ESE-like sequence which has been shown to interact with the SR protein ASF/SF2 in vitro (16). This suggests that recognition of exon 3 may require SR protein-ESE interactions, and furthermore, that Rev mediated skipping of exon 3 may result from a disruption of such interactions.
Viral persistence

Viral pathogenesis and persistence over a life-long infection is dependent upon both host and viral factors. Genetic variation is one of the important mechanisms by which retroviruses persist. Variation in retroviruses results from errors produced during reverse transcription (19, 23, 36, 39). The reverse transcriptase enzyme is a low fidelity RNA-dependent, DNA polymerase. The continual misincorporation of nucleotides allows for rapid virus evolution in the constantly changing host environment. Typically, variation is scattered throughout the retrovirus genome. However, variation in EIAV is localized to three principle regions: The virus long terminal repeat (LTR), the SU protein (gp90), and a region overlapped by the TM protein (gp45) and the major exon of rev (1, 5, 32).

Genetic variation in structural proteins is a well-studied mechanism of EIAV persistence. The largest degree of variation and emergence of new antigenic variants occurs during periods of high viremia (20, 27). Antigenic variation has been accepted for a long time as an important mechanism of persistence. Variation of the surface envelope and transmembrane proteins can occur as point mutations, insertions, or duplications. These changes can alter antigenicity and allow evasion of host immune responses throughout infection (21, 25-27, 33, 40-43, 46, 54). Variation in the SU protein is recognized as one of the principle mechanisms of evasion of the immune response. Typically, the largest variation in SU occurs within the principal neutralization domain and a defined hypervariable region.

Recent studies in other retroviruses suggest that genetic variation in the virus LTR and other regulatory proteins may contribute significantly to persistence (5, 28, 32). Genetic variation which decreases the replication competence of a virus may play a role in evasion of the host immune response by down-regulating production of immunodominant structural proteins. The virus LTR is required for viral integration into the host chromosome and contains the promoter region and the only transcription start site for the virus. In vivo studies
have shown that nucleotide sequence changes in the EIAV LTR alter transcription factor binding sites (32). Altering, adding, or removing transcription factor binding sites might modulate overall levels of viral transcription and facilitate gene expression in cells other than differentiated monocytes. Hence, changes in transcription factor binding sites may modulate levels of virus gene expression during the course of infection and/or shift cell tropism (32).

Variation in the regulatory genes of complex retroviruses may provide yet another mechanism by which retroviruses persist. Variation in the HIV-1 rev, and another accessory gene, nef, has been correlated with attenuated viruses in long term non-progressor patients (LTNPs, patients who do not develop AIDS within ten years) (18, 22, 24). Variation in Rev may alter levels of virus replication by modulating the production of the late structural proteins which may act as a mechanism to evade the host immune response. This hypothesis is supported by the identification of Rev-attenuated phenotypes in asymptomatic HIV-1 infections and in LTNPs (22). Genetic variation in EIAV Rev may also be an important factor of disease pathogenesis. Extensive nucleotide substitutions have been found in the EIAV rev ORF both in vivo and in vitro (1, 28). Variation in EIAV Rev may modulate the production of virus structural proteins and hence, virus replication in vivo. Furthermore, Rev attenuation may be an important factor for long-term virus persistence in subclinical EIAV infections.

**Overall goal**

Rev post-transcriptionally regulates expression of incompletely spliced mRNAs coding for virus structural proteins and providing for progeny RNA molecules. Therefore, Rev is absolutely required for virus replication and factors which modulate Rev activity may result in changes in virus replication, and ultimately, disease pathogenesis in vivo. The long-term goal of this research is to determine if variation in EIAV Rev contributes to viral
pathogenecity during clinical disease. This does not, however, imply that variation in Rev is the sole or dominant regulator of virus replication in vivo. The importance of variation in the LTR and the SU protein has been well studied and recognized. Instead, we propose that variation in Rev may be an additional factor contributing to EIA pathogenesis. To achieve this goal, we proposed to accomplish the following:

1. Determine if variation in rev alters nuclear export activity. (Chapter 2)

2. Analyze rev variation in an experimentally infected animal to determine if variation correlates with clinical disease. (Chapter 3)

To fully grasp the significance of variation in rev, we must further understand the mechanism of Rev function. EIAV Rev has been shown to be functionally homologous to HIV-1 Rev. The HIV-1 Rev export pathway has been well characterized and therefore serves as a good model for EIAV Rev function. However, EIAV Rev shares little sequence homology with HIV-1 Rev, and at the beginning of these studies, only the NES of EIAV Rev had been mapped. Therefore, an additional goal of this research was to delineate the cis- and trans-acting factors necessary for Rev function. Specifically, we proposed to do the following:

1. Identify the EIAV Rev responsive element. (Chapters 2 and 4)

2. Delineate the mechanism of Rev-mediated alternative splicing. (Chapters 4 and 5)

3. Map the domains of Rev required for RNA nuclear export. (Chapter 5)
References


mRNA is dependent on multimer formation mediated through a basic stretch of amino acids. Genes Dev. 4:1357-1364.


CHAPTER 2. BIOLOGICAL CHARACTERIZATION OF REV VARIATION IN EQUINE INFECTIOUS ANEMIA VIRUS

A paper published in Journal of Virology

Michael Belshan¹, Matthew E. Harris², Anne E. Shoemaker¹, Thomas J. Hope³, and Susan Carpenter¹

ABSTRACT

Sequence analysis identified significant variation in the second exon of equine infectious anemia virus (EIAV) rev. Functional analysis indicated that limited amino acid variation in Rev significantly altered the export activity of the protein, but did not affect Rev-dependent alternative splicing. EIAV Rev can mediate export through two independent cis-acting Rev responsive elements, and differences among Rev variants were more pronounced when both RREs were present. Variation in Rev may be an important mechanism for regulation of virus replication in vivo and contribute to changes in clinical disease.

INTRODUCTION, RESULTS, AND DISCUSSION

Equine infectious anemia virus (EIAV) is a member of the lentivirus subfamily of retroviruses and possesses many of the characteristic features of that subfamily including a complex genome organization, tropism for cells of the monocyte/macrophage lineage, and establishment of a persistent, lifelong infection. Whereas lentivirus infections are typically characterized by a slow, chronic disease, EIAV can induce a rapid, variable disease course in horses. Horses which survive early clinical episodes carry a lifelong, persistent infection with low viral load. The rapid changes between clinical stages of disease which occur during EIAV infection provide for an excellent model for analyzing factors which contribute to lentivirus pathogenesis and persistence. One factor important in EIAV persistence and pathogenesis is genetic and antigenic variation. Genetic mutations in the viral env gene are associated with the

¹ Department of Veterinary Microbiology and Preventive Medicine, Iowa State University

² Infectious Disease Laboratory, The Salk Institute for Biological Studies
occurrence of antigenic variant viruses, and the role of antigenic variation in EIAV persistence has been extensively studied (14, 20, 29, 32, 35). Additional clusters of genetic variation are found in the virus long terminal repeat (LTR) and in the region of gp45/Rev overlap (1, 4, 21, 27). The biological significance of variation in these regions is not clear; however, genetic changes which alter levels of viral gene expression may be important factors in viral pathogenesis in vivo. EIAV replicates in cells of the monocyte/macrophage lineage (28) and the severity of clinical and pathological signs of disease is closely associated with levels of viral replication in these cells (3). Therefore, variation in viral regulatory elements may modulate overall levels of virus replication and contribute to changes in clinical disease course.

Lentiviruses utilize complex mechanisms to regulate virus replication. The regulatory protein Rev functions to direct the nuclear export of incompletely spliced viral RNAs encoding viral structural proteins during the late phase of virus replication. Rev-dependent RNA export pathways have been described for numerous lentiviruses (reviewed in 9), and the HIV-1 Rev-mediated RNA export pathway is the best characterized. HIV-1 Rev binds a secondary structure in the viral pre-mRNA called the Rev Responsive Element (RRE) (8, 24, 42), multimerizes (22, 23, 34, 41), then utilizes a non-mRNA nuclear export pathway to redirect movement of incompletely spliced RNA from the nucleus (2, 10, 11, 40). Discrete functional domains within the protein mediate the interactions of Rev with cellular proteins and viral RNA required for nuclear localization, RNA binding, multimerization and nuclear export.

EIAV Rev is a 165 amino acid protein translated from a bicistronic four exon mRNA coding for both Tat and Rev (7, 39). The nuclear export signal (NES) of EIAV Rev has been mapped to amino acids 31-55 (12), and domain swapping experiments have shown that the EIAV Rev NES can substitute for the HIV-1 or visna virus NES (2, 25, 30). Other functional domains of EIAV Rev have not yet been identified. EIAV Rev also has an additional, apparently unique function among complex retrovirus export proteins. Whereas HIV-1 Rev and HTLV-1 Rex inhibit the expression of both Tat and Rev specific mRNAs in facilitating the export of
incompletely spliced mRNAs, EIAV Rev specifically down-regulates its own production, independent of Tat, by promoting exon 3 skipping of the bicistronic mRNA (Figure 1) (13, 26). This mechanism allows for continuous production of Tat while Rev synthesis is limited. Although the mechanism by which the alternative splicing occurs has yet to be completely delineated, it has been proposed that binding of EIAV Rev to a RRE overlapping exon 3 interferes with SR protein-RNA or SR protein-snRNP interactions (13). The disruption of SR protein binding is thought to result in exon 3 exclusion (13). This multi-functional nature of EIAV Rev highlights its importance in regulation of virus gene expression and replication. As such, genetic variation in Rev may have added significance in vivo.

Rev is absolutely required for expression of lentivirus structural genes and production of new virus. Therefore, factors which modulate Rev activity and, consequently, alter levels of viral gene expression may be important in regulating virus replication in vivo. Rev-attenuated phenotypes have been identified during asymptomatic stages of HIV-1 infection, suggesting that variation in Rev could alter virus replication in vivo and contribute to the clinical outcome of infection (16, 18). It would be expected that viral phenotypes with highly competent Rev phenotypes would be present during periods of rapid virus replication; whereas, the production of "attenuated" or latent virus may be important for evasion of an active host immune response during periods of long-term persistence. Indeed, using multiple assays Hua et al. (16) showed that HIV-1 Rev clones obtained from asymptomatic infections were less functional than wild-type Rev. Restricted expression of viral structural genes is a common strategy of persistent viruses (33), and these findings suggest that variation in Rev may be an important factor in lentivirus pathogenesis. We had previously identified extensive nucleotide substitutions in the Rev ORF from virus obtained from a horse experimentally infected with EIAV (1). The co-existence of putative Rev-competent and Rev-deficient phenotypes suggested that variation in EIAV Rev may contribute to virus persistence through regulation of structural gene expression. The goal of the present study was to further characterize genetic variation in EIAV Rev and to
determine if variation in EIAV Rev altered biological activity.

**Genetic variation in EIAV Rev.** To further explore the potential role of Rev variation in EIAV pathogenesis, we analyzed Rev cDNAs obtained from cells inoculated with either the Th-1, Th-6, or MA-1 virus isolates (Fig. 2A,B), as well as additional EIAV Rev sequences available in Genbank (Fig. 2C). Th-1 and Th-6 are field-derived virus isolates of EIAV recovered during the first and sixth febrile cycles, respectively, of a horse experimentally inoculated with whole blood from an EIAV-seropositive, naturally infected horse (1, 6). MA-1 is a cell-culture adapted, avirulent virus derived from Th-1 by *in vitro* passage in equine dermal (ED) cells (5, 6). The analysis indicated a high degree of genetic variation in Rev exon 2. The sequences represent a variety of isolates and include virulent (Wyoming) and avirulent (MA-1) EIAV as well as *in vivo* (Th-1, Th-6) and *in vitro*-adapted (MA-1) virus. Some of the sequences were derived from a single proviral clone isolated by limiting dilution and thus are representative of a predominant virus (i.e. P3.21), whereas other sequences represent quasispecies obtained following PCR amplification and cloning of viral cDNA or proviral DNA (Th-1.51). Numerous amino acid substitutions were found in the NES and in a 71 amino acid region in the center portion of the exon. The changes included deletions as well as amino acid substitutions, some of which resulted in the appearance of premature stop codons. In many cases, identical substitutions were found to re-occur, or to occur at specific amino acids, regardless of the virus isolate. Examples include valine/alanine at amino acid 105 and isoleucine/arginine/asparagine at amino acid 113. In other cases, a single change was diagnostic of a particular virus isolate. For example, all of the MA-1 quasispecies contained a glycine at amino acid 39, whereas cDNA clones from the related Th-1 virus, or the unrelated Wyoming strain of virus have an aspartic acid at that location. Also, the glutamine-serine amino acids present at amino acids 134 and 135 were more frequent in MA-1 quasispecies, whereas the glycine to aspartic acid change at amino acid 115 was found only in the virulent Wyoming strain of EIAV. Within the NES, eighty percent of the variation occurred at amino acids reported to be necessary for Rev activity (25).
While it is possible that the observed variation in Rev is merely a consequence of random variation, or reflects selection for changes in the overlapping gp45 reading frame, recent findings suggest a role for Rev variation in the biology of EIAV in vivo. During successive febrile periods in a pony experimentally inoculated with EIAV, nucleotide and amino acid variation in EIAV Rev accumulated at approximately the same rate as that observed in gp90, and more frequently than was observed in gp45 or in the LTR (31). Taken together, these findings support the hypothesis that variation in EIAV Rev is biologically significant.

Two independent RREs can mediate EIAV Rev-dependent export. To assess the functional activity of Rev variants, we developed an in vitro nuclear export assay similar to that widely used in functional assays of other lentivirus Rev proteins (17, 25, 37). In other complex retroviruses, trans-activation of the Rev/Rex RNA export pathway occurs through an interaction with a single RRE (reviewed in 9). For the majority of lentiviruses, the RRE is located near the surface-transmembrane envelope region; two exceptions are FIV and HTLV-1, for which the RREs have been mapped near the 3' end of the genome (9,36,38). Surprisingly, previous studies have identified two cis-acting regions in EIAV which are able to act as RREs (15,26). However, specific binding of Rev has been shown with only one element overlapping the 3' end of the first rev exon (13). The exact location of the second RRE has not been identified and initial studies were performed to confirm the presence of two RREs. A pDMI38-derived reporter plasmid, pERRE-All, was constructed by insertion of the CAT gene and a region containing all of the putative EIAV RRE sequences into an intron flanked by HIV-1 splice sites (15, 17). Additional reporter plasmids were constructed containing EIAV regions previously shown to be able to act as RREs (26): pERRE-1 contains a short 5' RRE sequence overlapping the first Rev exon and pERRE-2 contains a major portion of the remaining downstream EIAV sequence present in pERRE-All (15). The locations of the EIAV sequence present in the reporter constructs is shown in Fig. 1B. For functional assays, 293 cells were seeded in triplicate at 1-5 x 10^5 cells/well in 6-well tissue culture dishes. The next day cells were
transfected with 0.2 μg of reporter plasmid, 0.2 μg of beta-galactosidase expression plasmid pCH110 (Pharmacia, Uppsala, Sweden) or pSV-βgal (Promega, Madison, WI) and 1 μg of an MA-1 Rev expression plasmid or empty vector. pUC19 DNA was added to bring the total amount of DNA transfected in each well to 2 μg. Cells were transfected using calcium phosphate co-precipitation and media was changed the next day. Two days posttransfection cells were harvested in phosphate-buffered saline (PBS) containing 5 mM EDTA, pelleted and resuspended in 300 μl 0.25 M Tris, pH 7.5, and lysed by three rounds of freeze-thawing. 50 μl of lysate was assayed for beta-galactosidase activity and these values were used to normalize lysates for CAT assays. Reaction volumes for CAT assays were equalized with 0.25 M Tris, pH 7.5, to a final volume of 92 μl and incubated at 37°C with 5 μl 20 mM acetyl coenzyme A and 3 μl of 50 mCi/mmole [14C]chloramphenicol. Unacetylated and acetylated forms were separated by thin-layer chromatography and quantified using a Molecular Dynamics phosphoimager (Sunnyvale, CA). The percentage of acetylation was calculated for each transfection, and the data represents the average acetylation and standard error of the mean for all experiments.

All three ERRE reporter plasmids were found to contain cis-acting elements able to mediate Rev-dependent RNA export (Fig. 3A). For purposes of comparison, results are presented as percent of activity found with pERRE-All, which is shown as 100%. pERRE-1, containing the RNA element shown by Gontarek et al. to interact with GST-Rev in vitro (13), produced the majority (52%) of activity found with pERRE-All. In contrast, assays with pERRE-2 resulted in only 17% of the activity seen with the pERRE-All vector. There was also a low level of transactivation of the background vector, pDM138. These findings confirm previous studies indicating that EIAV Rev can mediate nuclear export through two separate RRE elements (26). In addition, they provide quantitative results which indicate that the primary RRE is contained within ERRE-1 and encompasses Rev exon 1. Although both RREs appear to be required for maximum efficiency of the Rev-dependent export pathway, the significance of ERRE-2 as an important mediator of RNA export is questionable. The results presented here
suggest that ERRE-2 functions primarily as an enhancer of ERRE-1 rather than as an independent mediator of RNA export. Further studies are needed to more clearly ascertain the mechanism of the EIAV dual-RRE export pathway.

**Variation in Rev alters biological activity.** To determine whether amino acid variation in Rev resulted in differences in biological activity, eight variant Rev cDNAs, four from MA-1 infected ED cells and four from Th-1 infected macrophages, were cloned into the eukaryotic expression vector pCR3 (Invitrogen, Carlsbad, CA) for functional analysis. Western blot analysis using Rev-specific polyclonal antibodies confirmed that all variants expressed at similar levels (data not shown). Variant cDNAs were assayed in transient expression assays by co-transfection with Rev reporter plasmids containing ERRE-All, ERRE-1, or ERRE-2 as described above. The results indicated that amino acid changes in the Rev ORF significantly altered biological activity ($p < 0.0001$) when both RRE elements were present (Fig. 3B). Variants ranged from being inactive (A22) to producing activities greater than three-fold that of MARev (H21). EIAV Rev variants F22, H21, and 27D4 were significantly more active than all other variants. In general, the range of biological variation was greater in the Th-1 derived Revs, consistent with the greater degree of sequence diversity observed among those clones as compared to the MA-1 derived clones. The results are consistent with studies analyzing variation in HTV-1 Rev (16) and support the hypothesis that biological changes in Rev activity may have significance *in vivo*. In addition, our results demonstrate that variation in regions outside the NES can also alter nuclear export activity. Because all assays were done using a single RRE sequence, we cannot rule out the possibility that compensatory mutations in the RRE minimize the biological effects of Rev variation *in vivo*. Such analysis requires further mapping of the EIAV RREs.

Given the possibility of a synergistic interaction contributing to EIAV Rev-dependent export, we further characterized the effects of variation with the individual RREs by assaying the nine variants with both pERRE-1 and pERRE-2. Since the functional activity of the reporter
plasmids differs (Fig. 3A), the experiments with the separate reporter plasmids were performed using CAT assay conditions in which the acetylation of pcMARRev lysates was approximately 20%. Analysis of the individual Rev variants using the pERRE-1 reporter plasmid produced a similar pattern of activity relative to MARRev as was observed with ERRE-All (Fig. 3B,C). However, with the exception of A22, the differences among the variants were less pronounced than observed with the pERRE-All reporter plasmid. H21 was only two-fold more active than MARRev, while 27A2 and 27D2 were slightly more active relative to MARRev. Surprisingly, the pattern of variation seen with pERRE-All was abolished when the variants were assayed with pERRE-2 reporter plasmid (Fig. 3D). A22 was still inactive, but all other variants were more active than MARRev and not significantly different from each other. Together, these results indicate that genetic variation in Rev alters biological activity, and that the effects of Rev variation are enhanced in the presence of both RREs. The mechanism(s) by which EIAV Rev utilizes two separate RREs are unknown, and it is not clear why the differences among the variants are decreased when only one RRE element is present. As suggested above, the downstream RRE may function primarily as an enhancer element of nuclear export mediated by RRE-1. If so, the effects of variation in regions of Rev important for interaction between RRE-1 and RRE-2 may require the presence of both elements for observable differences in biological activity.

**Rev-dependent alternative splicing is not affected by variation.** The current model of EIAV Rev dependent alternative splicing proposes that binding of Rev to the ERRE-1 interferes with SR protein-RNA or SR protein-snRNP interactions to promote exon 3 skipping (13). The significance of this phenomenon in terms of virus replication is not known, although alternative splicing may play an important role in regulation of virus replication. Therefore, studies were undertaken to determine if the Rev sequence variants differed in splice-site utilization during EIAV mRNA processing. To analyze Rev-mediated alternative splicing patterns, we developed Cf2Th cell lines stably transfected with a Rev-defective (Cf2Th/51) or
Rev competent (Cf2Th/112) provirus by G418 selection (data not shown). Cf2Th/51 cells were then trans-complemented with the variant cDNAs. Cells were seeded at $2 \times 10^5$ cells/well in six-well tissue culture plates and transfected with 9 µg of Rev variant plasmid or empty vector DNA by liposome-mediated transfection (Boehringer Mannheim, Indianapolis, IN). Two days post transfection RNA was isolated and cDNAs were amplified by RT-PCR using primers which spanned all EIAV splice-donor and splice-acceptor sites (5' primer: (CGCAGACCTACCTGTTG (nt 354); 3' primer: TCTTCAGGTAACGACTGCC (nt 7301)) (Fig. 1). To allow for sensitive visualization of the splicing products, the 5' primer was [32P]end-labeled. RT-PCR was performed as described by the manufacturer (Perkin Elmer, Foster City, CA). PCR was performed with an annealing temperature of 55° and run for 25 cycles. DNA from the RT-PCR reactions was phenol-chloroform extracted, ethanol precipitated, and resuspended in 40 µl 0.1X TE. 10 µl of each reaction was electrophoresed through a denaturing (7 M urea) 5% polyacrylamide gel. Gels were fixed, dried, then exposed to film for visualization. As a control, Rev mRNAs were amplified from transfected plasmids using pCR3-specific primers (5' primer: ATACGACTCACTATAGGG; 3' primer: ATTTAGGTGACACTATAG). The results indicated both the 1,2,3,4 exon mRNA and the alternately spliced 1,2,4 exon mRNA were present in Cf2Th/112 cells and in all Cf2Th/51 cells trans-complemented with variant Rev cDNAs, including the functionally inactive A22 (Fig.4). In contrast, the 1,2,4 mRNA was not detected in Cf2Th/51 cells alone, or in cells trans-complemented with vector DNA. The failure to detect differences in alternative splicing patterns suggests that variation present in the cDNAs we examined is not important for Rev-mediated alternative splicing. The finding that A22 was similar to other variant cDNAs in splice site utilization indicates the nuclear export function of Rev is independent of exon 3 skipping. Therefore, these two functions most likely occupy separate domains in the protein, although both functions may require RRE binding.
Overall, our findings demonstrate that aa variation can enhance or attenuate EIAV Rev phenotype. Previous studies with HIV-1 have shown that variation within the HIV-1 Rev NES can alter biological phenotype, and that the observed changes in function were consistent in both \textit{in vitro} assays and studies of virus replication (16). We have demonstrated that variation in regions other than the NES can also modulate Rev activity \textit{in vitro}. Given the extent of Rev variation we have observed, our results suggest that variation in Rev may be an important mechanism of modulating levels of virus replication during the course of clinical disease. Indeed, Leroux et al. reported rapid variation in Rev during sequential febrile cycles of a pony experimentally inoculated with EIAV (21). Functional Rev is absolutely required for production of infectious virus and it might be expected that Rev-defective, or Rev-attenuated, genotypes would be rapidly selected against during replication \textit{in vivo}. However, factors which decrease Rev activity and decrease levels of viral gene expression may have a selective advantage and allow virus to persist in the face of an active host immune response. Further structural and functional analysis of \textit{in vivo} derived variants at different stages of clinical disease are needed to delineate the role of Rev in disease pathogenesis.

\section*{ACKNOWLEDGEMENTS}

We thank Yvonne Wannemuehler, Teresa A. Smith, and Mary Jane Long for technical assistance, Eric Vaughn for helpful discussions, and Wendy Maury and C. Martin Stoltzfus for critical review of this manuscript.

This work was supported in part by USDA grant 96-02102 (S.C.) and PHS grants AI30025 (S.C.) and AI35477 (T.H.). M.H. is supported by a National Science Foundation Graduate Research Fellowship.
REFERENCES


sequences during EIAV infection resemble changes reported for sequential isolates of HIV. Virology 161:321-331.


FIGURE LEGENDS

FIG. 1. Organization of the EIAV genome. (A) Known ORFs and predominant mRNAs isolated from virus-infected tissue culture cells (19). (B) Location of EIAV regions inserted into pDM138 CAT constructs (17). pERRE-All contains nucleotides 5280 to 7534, pERRE-1 contains a short 5' RRE sequence overlapping the first Rev exon (nucleotides 5280 to 5691), and pERRE-2 contains a major portion of the remaining downstream EIAV sequence present in pERRE-All (nucleotides 5837 to 7534) (15). All numbering of nucleotides in the present report is based on Kawakami et al. (19).

FIG. 2. Amino acid sequence alignments of Rev exon 2 (a.a. 31-165). (A) Amino acid sequence of exon 2 of MA-1 Rev (4) and amino acid sequences from cDNAs isolated from MA-1 infected ED cells (ME) and MA-1 infected HMC (MM). (B) Amino acid sequences of cDNAs isolated from Th-1 infected HMC (A22, B11, F22, H21) and viral DNAs isolated from an EIAV(+) horse at the first and sixth febrile cycles (Th-1 and Th-6 respectively) (1). Missing sequences are due to use of an internal 5' primer for PCR amplification. (C) In vivo Rev exon 2 sequences obtained from Genbank (Accession numbers: X63059, X16988, M14855, M18385, M18386, M18387, M18388, M87580, M93674).

FIG. 3. In vitro assays of EIAV Rev activity. Transfection experiments were performed as described in the text. Two days post-transfection cells were harvested and lysates normalized for the CAT reactions by beta-galactosidase assay. Percentage of acetylation is shown. Individual experiments included triplicate wells and the data shown represents the mean of at least three separate experiments. Error bars denote the standard error of the mean for all
experiments. (A) Rev can trans-activate through two discrete regions of EIAV. pcMAREv transactivation of CAT reporter plasmids pERRE-All, pERRE-1, pERRE-2, and pDM138. Each reporter plasmid contains the EIAV sequences shown in Fig. 1b, pDM138 is the background reporter plasmid. Transfections and CAT assays were performed as described in the text, except that lysates from wells with pERRE-All were diluted five-fold to allow for the comparison. (B) Transactivation of pDM138 CAT reporter plasmid pERRE-All by EIAV Rev variants showing that amino acid variation in Rev alters the biological activity of the protein. (C,D) Transactivation of pDM138 CAT reporter plasmids pERRE-1 (c) and pERRE-2 (d) by EIAV Rev variants indicating that the full effects of variation require both RRE regions. CAT assays were performed for each experiment with conditions that resulted in approximately 20% acetylation for the pcMAREv lysates. Therefore, although the activities of the reporter plasmids differs, all experiments appear on the same scale.

FIG 4. Amino acid variation does not alter Rev-dependent alternative splicing. Rev-defective cells were transfected with 9 μg of variant Rev plasmids. Total RNA was isolated and reverse transcribed using random hexamer primers. cDNA was amplified by PCR with EIAV primers specific for exon 1 and exon 4 using a 5' primer that was [32P]end-labeled. PCR products were isolated and separated by electrophoreses through a denaturing 5% polyacrylamide gel. The location of EIAV splicing products is shown. As a control, mRNAs from transfected plasmids were amplified using pCR3-specific primers flanking the Rev insert.
### A.

<table>
<thead>
<tr>
<th>MA-1</th>
<th>BASIC REGION</th>
</tr>
</thead>
<tbody>
<tr>
<td>ME27A1</td>
<td>N</td>
</tr>
<tr>
<td>ME27A2</td>
<td>N</td>
</tr>
<tr>
<td>ME27A3</td>
<td>N</td>
</tr>
<tr>
<td>ME27A4</td>
<td>N</td>
</tr>
<tr>
<td>ME27C1</td>
<td>N</td>
</tr>
<tr>
<td>ME27C2</td>
<td>N</td>
</tr>
<tr>
<td>ME27C3</td>
<td>N</td>
</tr>
<tr>
<td>ME27D1</td>
<td>N</td>
</tr>
<tr>
<td>ME27D2</td>
<td>N</td>
</tr>
<tr>
<td>ME27D3</td>
<td>N</td>
</tr>
<tr>
<td>ME27D4</td>
<td>N</td>
</tr>
</tbody>
</table>

### B.

<table>
<thead>
<tr>
<th>MA-1</th>
<th>BASIC REGION</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Y</td>
</tr>
<tr>
<td>B1</td>
<td>A</td>
</tr>
<tr>
<td>C1</td>
<td>N</td>
</tr>
<tr>
<td>Th-1.51</td>
<td>Y</td>
</tr>
<tr>
<td>Th-1.81</td>
<td>A</td>
</tr>
<tr>
<td>Th-1.112</td>
<td>A</td>
</tr>
<tr>
<td>Th-1.113</td>
<td>A</td>
</tr>
<tr>
<td>Th-1.114</td>
<td>A</td>
</tr>
<tr>
<td>Th-6.52</td>
<td>A</td>
</tr>
<tr>
<td>Th-6.54</td>
<td>A</td>
</tr>
<tr>
<td>Th-6.56</td>
<td>A</td>
</tr>
</tbody>
</table>

### C.

<table>
<thead>
<tr>
<th>MA-1</th>
<th>BASIC REGION</th>
</tr>
</thead>
<tbody>
<tr>
<td>P176</td>
<td>T</td>
</tr>
<tr>
<td>NS5</td>
<td>T</td>
</tr>
<tr>
<td>E14C2</td>
<td>T</td>
</tr>
<tr>
<td>P3.2-1</td>
<td>A</td>
</tr>
<tr>
<td>P3.2-2</td>
<td>A</td>
</tr>
<tr>
<td>P3.2-3</td>
<td>A</td>
</tr>
<tr>
<td>P3.2-5</td>
<td>A</td>
</tr>
<tr>
<td>WYO FD</td>
<td>S</td>
</tr>
<tr>
<td>TTAENR</td>
<td>E</td>
</tr>
</tbody>
</table>

---

Figure 2
Figure 3
Figure 4

Cf2Th/51 Cells Complemented with:

<table>
<thead>
<tr>
<th>pcDNA3</th>
<th>pcMARev</th>
<th>pcA22</th>
<th>pcB11</th>
<th>pcF22</th>
<th>pcH21</th>
<th>pc27A2</th>
<th>pc27D2</th>
<th>pc27D4</th>
<th>Cf2Th/112</th>
</tr>
</thead>
</table>

1,2,3,4 →
1,2,4 →
1,4 →
cDNA mRNAs →
CHAPTER 3. SEQUENCE VARIATION IN EQUINE INFECTIOUS ANEMIA VIRUS REV CORRELATES WITH VARIABLE STAGES OF CLINICAL DISEASE IN AN EXPERIMENTALLY INFECTED PONY.

A paper to be submitted to the Journal of Virology

Michael Belshan\textsuperscript{1}, Prasith Baccam\textsuperscript{1,2}, J. Lindsay Oaks\textsuperscript{3}, Brett A. Sponseller\textsuperscript{1}, Sean C. Murphy\textsuperscript{1}, James Cornette\textsuperscript{2} and Susan Carpenter\textsuperscript{1}

ABSTRACT

Here we examine genetic and biological variation in the regulatory protein Rev of equine infectious anemia virus (EIAV) throughout a dynamic disease course of an experimentally infected pony. The pony underwent a variable disease course, including an acute fever episode at 12 DPI, multiple recurrent fever episodes until 135 DPI, a prolonged subclinical period, and two late fever episodes. Viral RNA was isolated from the inoculum and sequential sera samples, and the rev exon 2/gp45 overlapping ORFs were amplified, cloned, and sequenced. Novel variants were found throughout infection and genetic analysis indicated that both the rev and gp45 ORFs were under selective pressure. The Rev variant predominant in the inoculum, R1, remained dominant during the acute period of infection (until 35 DPI). However, R1 was replaced by new dominant variants during the recurrent fever period (67-135 DPI). R1 re-emerged as the dominant variant during the period of subclinical infection. However, a new dominant variant, R93, was associated with the late fever episodes. In vitro analysis of Rev nuclear export activity indicated the dominant variants present during clinical periods were significantly more active than the variants dominant during the acute and subclinical periods. These results suggest Rev activity

\textsuperscript{1} Department of Microbiology and Preventive Medicine, Iowa State University
\textsuperscript{2} Department of Mathematics, Iowa State University
\textsuperscript{3} Department of Microbiology and Pathology, Washington State University
correlates with clinical disease and provide additional evidence that sequence variation in \textit{rev} may be a contributing factor of EIAV pathogenesis.

\textbf{INTRODUCTION}

Equine infectious anemia virus (EIAV) possesses many of the characteristic features of the \textit{Lentiviridae} subfamily of retroviruses, including a complex genome organization, tropism for cells of the monocyte/macrophage lineage, and establishment of a persistent, lifelong infection. Many lentivirus infections are characterized by a slow, chronic disease, but infection of horses with EIAV can result in an acute, dynamic disease course. The characteristic infection with EIAV involves recurring cycles of viremia, fever, and thrombocytopenia, with a possible development of anemia (reviewed in 19). Fever cycles are typically irregular and may last for up to one year, despite the onset of a neutralizing antibody response. Horses which survive these initial clinical episodes become life-long, inapparent carriers of the virus.

Virus persistence over a life-long infection is dependent upon both host and virus factors. Genetic variation in lentiviruses, the result of errors produced by reverse transcriptase, is a well-studied mechanism of lentivirus persistence (18, 22, 33, 37). Variation of the surface envelope protein, due to mutations, insertions, or duplications within the principal neutralization domain or hypervariable region, can alter antigenicity and allow evasion of host immune responses (20, 24-26, 32, 38-41, 44, 51). Studies examining the immune responses of EIAV infected horses indicate both cellular and humoral immune responses contribute to suppression of viral replication during inapparent stages of infection (12, 44). However, no subset of the immune response has been shown to specifically correlate with control of virus replication during long term persistence (12). Therefore, other host and virus factors may be important determinants of virus persistence and pathogenesis.
Lentiviruses utilize complex mechanisms to regulate virus replication. Variation in regulatory elements which alters the level of virus gene expression and subsequent replication may play an important role in disease. Genetic variation has been previously described in the second exon of the EIAV regulatory protein Rev (1, 3). Rev facilitates the nuclear export of the incompletely spliced viral RNAs during the late phase of virus replication. These viral RNAs encode the structural proteins and provide progeny RNA molecules. Rev-dependent RNA export pathways have been described for numerous lentiviruses (reviewed in 8, 15), and the human immunodeficiency virus type 1 (HIV-1) Rev export pathway is the best characterized. HIV-1 Rev binds a structure in the viral pre-mRNA called the Rev-responsive element (RRE) (50), multimerizes (36), then utilizes the CRM1 nuclear export pathway to redirect movement of incompletely spliced RNA from the nucleus (5, 9, 45, 46). Discrete functional domains within the protein mediate the interactions of Rev with cellular proteins and viral RNA required for nuclear localization, RNA binding, multimerization and nuclear export. EIAV Rev is a 165 amino acid protein functionally homologous to HIV-1 Rev (10). The nuclear export signal of EIAV Rev has been mapped to amino acids 32-55 (10) and has been shown to be able to substitute for the HIV-1 or visna virus effector domains (10, 28). A putative nuclear localization signal has been mapped to the arginine rich C-terminus of Rev (13).

Rev is absolutely required for expression of structural genes and production of new virus. Therefore, factors which modulate Rev activity and, consequently, alter levels of viral gene expression may be important in regulating virus replication in vivo. Rev-attenuated phenotypes have been identified during asymptomatic stages of HIV-1 infection suggesting that variation in Rev could alter virus replication levels and contribute to the clinical outcome of infection (16, 21). We had previously identified extensive nucleotide substitutions in the second exon of EIAV rev in virus obtained from an experimentally infected horse (1). The
coexistence of putative Rev-competent and Rev-deficient phenotypes suggested that variation in Rev may contribute to virus persistence through regulation of virus replication. Subsequent in vitro analysis indicated that limited amino acid variation in Rev could alter biological activity (3), further suggesting variation in Rev may contribute to EIAV persistence and pathogenesis by modulating virus replication during disease.

The present study characterizes genetic and biological variation in the second exon of EIAV rev during a long-term infection of an experimentally-infected pony. Sequence analysis indicates a constant, but dynamic evolution of rev throughout disease. An in vitro assay used to measure the activity of Rev variants throughout disease demonstrates that Rev variants dominant in the recurrent and late fever periods displayed significantly more nuclear export activity than the variants dominant in the inoculum and subclinical periods of disease. These findings indicate that variation in rev correlates with clinical disease, and strongly argues that variation in rev is an additional factor contributing to the pathogenesis of disease caused by EIAV.

**MATERIALS AND METHODS**

**Experimental infection.** The experimental infection of pony 524 has been previously reported (35). The pony was infected with $10^3$ horse infectious doses of the highly virulent Wyoming strain of EIAV (EIAV_{yo}). Physical examinations, rectal temperatures, hemograms, and platelet counts were performed daily during clinical episodes and intermittently during chronic clinical disease or subclinical infection. Serum and plasma were collected, processed, and stored at −80°C until analyzed.

**Virus neutralization assay.** Assays were performed using equine dermal (ED) cells maintained in Dulbecco's modified eagle medium (DMEM) supplemented with 20% fetal bovine serum, penicillin, and streptomycin as previously described (7). ED cells were seeded
one day prior to infection at 10^5 cells/well in a 6-well dish. Serum from pony 524 was heat
inactivated at 60°C to destroy complement and serially diluted two-fold in supplemented
media. Five hundred focus forming units of EIAV_{WSU-5} in a 100 μl volume was incubated
with 100 μl of serially diluted pony 524 serum at 37°C for 1 h. Duplicate wells of ED cells
were inoculated with 100 μl of virus-serum mixture in the presence of 8 μg/ml polybrene,
and the following day the media was changed. Cells were incubated for an additional 72 h,
fixed with 100% methanol, and immunocytochemistry performed using convalescent anti-
EIAV horse sera to detect foci of virus-infected cells (7). Results are expressed as the serum
neutralization titer, defined as the reciprocal of the highest serum dilution that gave a 80%
reduction in foci as compared with preimmune and negative control serum.

**Isolation of viral RNA.** EIAV_{wy} inoculum was obtained as a serum sample collected
from an experimentally-infected foal at 7 days post-infection (DPI). To obtain viral RNA,
100 μl of EIAV_{wy} inoculum or pony 524 serum was pelleted from by centrifugation at
93,000 x g for 1 h at 4°C. Viral RNA was isolated from these pellets by guanidine
thiocyanate lysis and acid phenol-chloroform extraction using a commercially available kit
(Ambion, Austin, TX) and resuspended in 24 μl RNase-free glass distilled water containing
0.1 mM EDTA.

**Amplification and sequence analysis of viral variants.** RNA samples were DNase
treated using the methods of Huang et al. (17). Briefly, two units of DNase I (Ambion, TX)
were added to 3 μl of viral RNA, 20 mM MgCl_2, 1 mM of each dNTP, 1X PCR buffer II
(Perkin-Elmer, Branchburg, NJ), 20 units of RNase inhibitor, and 2.5 mM of random
hexamers in a total volume of 20 μl. The reaction was incubated at 37°C for 30 min and
heated to 75°C for 5 min to inactivate the DNase. After cooling to 4°C, 50 units of Moloney
murine leukemia virus reverse transcriptase were added. Reactions were incubated at 42°C
for 45 min, heated to 99°C for 5 min to inactivate the reverse transcriptase, and then cooled
to 5°C for 5 min. The reaction was brought up to 100 µl in 1X PCR buffer II with 2 mM MgCl₂, 0.2 mM of each dNTP, 2.5 units Taq polymerase (Perkin Elmer), and 1 µM of each primer. Primers were designed based on conserved regions flanking the second exon of rev (Table 1). PCR amplification conditions consisted of 37 cycles of denaturation at 94°C for 2 min, annealing at 50°C for 1 min, and extension at 72°C for 1 min. The initial and final cycles contained a prolonged extension at 72°C for 5 min. 2 µl of PCR product was TA-cloned into pGEM-T vectors as recommended by the manufacturer (Promega, Madison, WI) and transformed into Escherichia coli DH-5α. Positive clones were identified by colony blot hybridization using a subgenomic fragment of rev labeled with [³²P]dTTP by random primed labeling (Roche Molecular Biochemical, Indianapolis, IN). Plasmids were isolated from positive clones using a commercially available kit (Promega), and the rev inserts were sequenced bi-directionally with vector primers flanking the insert. All sequencing was performed by the Iowa State University DNA Synthesis and Sequencing Facility using an automated DNA sequencer. Sequences were aligned by MacVector and AssemblyLIGN software (Oxford Molecular, Beaverton, OR). To estimate the RT-PGR error rate, viral mRNA was isolated and the second rev exon amplified from a clonal cell line stably transfected with an EIAV proviral DNA clone. The amplicons were cloned and 17 individual clones were sequenced. The error rate of the RT-PGR amplification procedure was determined to be 0.025%, or 3 substitutions per 11,991 base pairs sequenced.

Construction of Rev Expression Vectors. rev variants were subcloned into pcRevWT, a Rev expression vector previously described as pcH21 (3). This plasmid contains a Rev cDNA in the pCR3 background (Invitrogen, Carlsbad, CA). pcRevWT was digested with the restriction endonuclease ApaI to remove rev exon 2. The cut plasmid was gel purified, dephosphorylated, and ligated with the exon 2 sequence variants. Two methods, based upon the direction of the inserts in pGEM-T, were employed to move variant exon 2
sequences into the pcRevWT background. Inserts in the "forward" direction were digested with \textit{ApaI}, gel purified, and ligated into the digested pcRevWT background. Inserts in the "reverse" orientation were PCR amplified with the EM7160 primer and a 3' EIAV primer containing an \textit{ApaI} site (underlined): 5'-CTGGGCCCTCATAAATGTTCCTCCTTCGC. The PCR products were purified, digested with \textit{ApaI}, and ligated into the pcRevWT background. All clones were verified for directionality by restriction mapping and confirmed by sequencing.

\textbf{Statistical analyses.} A two-dimensional chi-square test was used to test the null hypothesis that the frequency distribution of observed variants and the choice of primer pairs was independent. The chi-square statistic was calculated, and the p-value was determined for each contingency table. The statistical threshold used in this study was 95% confidence. Subset pairwise analyses were also performed on the data, and the results were interpreted with Bonferroni's correction (31). For the Monte Carlo simulation (43) to compare the four primer pairs, the chi-square distance was calculated between the four sets of simulated data. The data was simulated 10,000 times, and experimentally observed chi-square distances were compared with the 10,000 simulated distances to find the percentile ranking, which correlates to the p-value. For the pairwise analyses of primer pairs, 10 variants were randomly created for each set of the two simulated primer pairs. The rest of the simulation was the same as described immediately above, except the Kullback distance (4) was calculated rather than chi-square distance.

\textbf{Quantification of virus load.} RNA quantification standards were synthesized by \textit{in vitro} transcription from a linearized plasmid containing a 450-base pair fragment of EIAV \textit{gag}, (pEIAp26.1). The RNA was purified by extraction with Trizol Reagent, treated with 6 U of DNase I (Ambion) for 45 min at 37°C to remove plasmid DNA, and then extracted again with Trizol Reagent. Final RNA preparation was quantified by spectrophotometry, aliquoted,
and stored at -80°C. The absence of contaminating plasmid DNA was verified by PCR without reverse transcription.

To create standards, known copy numbers of RNA standards were assayed in triplicate as described below; these included $2 \times 10^7$, $2 \times 10^8$, $2 \times 10^9$, $2 \times 10^4$, and $2 \times 10^3$ copies of RNA. Reaction products were visualized in ethidium bromide-stained agarose gels, and their densities quantified using a commercial digital imaging system (IS1000, Alpha Innotech). The mean density for each copy number was plotted against the log_{10} of the copy numbers to construct a standard curve.

The one tube, semi-quantitative RT-PCR protocol was modified from that described by Hamel et al (11). The reaction buffer contained 50 mM KCl, 10 mM Tris (pH 9.0), 0.1% Triton X-100, and 2.0 mM MgCl₂, and each reaction contained 0.4 mM deoxyribonucleotide mix (Roche Molecular Biochemicals), 1 U of RNase inhibitor (Roche Molecular Biochemicals), 20 U of Moloney Murine Leukemia Virus reverse transcriptase (SuperScript II, Gibco BRL, Rockville, MD), 1.5 U of Taq DNA polymerase (Gibco BRL), and 90 pmoles of the forward primer (5'- ACTACTTGGGTGAATACCAT) and 90 pmoles of the reverse primer (5'- TCTGCCTAAACTGATCAAAA) in a final volume of 25 µl. These oligonucleotides prime both reverse transcription as well as the subsequent PCR reaction, and amplify a 322 base pair segment of the EIAV capsid protein gene (p26) (nt 897-1199). Amplification was performed as follows: 42°C for 40 min, DNase I inactivation at 95°C for 3 min, and 40 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 30 s. Each RT-PCR run included positive and negative controls. Complete removal of all DNA from samples was confirmed by performing a duplicate reaction without reverse transcriptase. Following amplification, 13 µl of the PCR reaction was analyzed by electrophoresis through 2.0% agarose and visualized by ethidium bromide staining and ultraviolet light. Serum samples from each time point were assayed in duplicate, and their mean density values were used to
calculate copy numbers of viral RNA from the standard curve. RT-PCR reactions for standards and samples were performed simultaneously using a reagent-master mix, and were analyzed simultaneously in a single agarose gel.

**CAT Assays.** 293 cells were maintained in Dulbecco’s modified eagle medium supplemented with 10% fetal calf serum and penicillin/streptomycin. Transient transfections and CAT assays were performed using 293 cells as previously described (3). Briefly, 1 μg of either pcDNA3 or Rev variant plasmid was transfected by calcium phosphate co-precipitation with 0.2 μg of pERRE-All reporter plasmid, 0.2 μg of pCH110 (Amersham Pharmacia Biotech, Piscataway, NJ), and 0.6 μg of pUC19. pERRE-All, a derivative of pDM138 containing EIAV nt 5280-7534, has been described previously (3). Each experiment also included a sham group which contained no reporter plasmid, but an additional 0.2 μg of pUC19. Two days posttransfection, cells were harvested, resuspended in 0.3 ml 0.25 M Tris (pH 7.5), lysed by freeze/thawing, and assayed for β-galactosidase activity to normalize CAT assays for transfection efficiency. Normalized lysates were assayed for CAT activity in a 0.1 ml volume with 3 μl [14C]chloramphenicol and 1 mM acetyl CoA. Acetylated products were separated by thin layer chromatography and the percentage of acetylation determined by quantification by phosphorimager. Experiments were performed in triplicate and results summarize a minimum of nine independent transfections. Statistical significance was determined by analysis of variance.

**RESULTS**

**Clinical profile of EIAV\textsubscript{wyo}-infected pony 524.** To accurately reflect the diversity of an *in vivo* infection, pony 524 was inoculated with a highly virulent Wyoming strain of EIAV (EIAV\textsubscript{wyo}), which has been maintained by serial *in vivo* passage (35). This inoculum represents a heterogeneous population of virus, similar to a natural infection. After
inoculation, pony 524 experienced a clinical disease course characterized by recurring fever cycles interspersed with afebrile periods ranging from days to months (Fig. 1A). The initial acute episode was characterized by a bi-phasic febrile response and thrombocytopenia from 10 to 22 days postinfection (DPI). The pony was then afebrile until 39 DPI. Thereafter, pony 524 underwent a period of seven recurrent fever episodes, which ranged from two to six days in length. The last fever episode ended after day 135, and pony 524 remained afebrile except for two late fever episodes at days 565 and 799. Febrile episodes were defined as rectal temperature greater than or equal to 101°F. Thrombocytopenia (platelet count less than 1 x 10^5 platelets/µl) was also observed concurrent with the initial fever cycles. Surprisingly, after the initial fever cycles subsided, pony 524 remained persistently thrombocytopenic throughout the experiment. Except for the fever and thrombocytopenia, no other clinical signs were observed.

Semi-quantitative RT-PCR was used to monitor virus levels at various DPI (Fig. 1B). Consistent with previous reports, high virus loads were found in the initial fever period at 12 DPI, but fell to a nearly undetectable level during the afebrile interval at 35 DPI. The highest virus load, 9.08 x 10^7 copies RNA/ml, was detected during the recurrent febrile period at 67 DPI. Levels of plasma viral RNA decreased somewhat in the subclinical period and were barely detectable at 754 DPI. However, virus rebounded to 1.38 x 10^5 copies/ml at 800 DPI. Serum neutralization assays to EIAVwsu-5, a cell culture derived strain of EIAVwyo (34) indicated that broadly reacting neutralizing antibodies appeared after the chronic fever cycles subsided (201 DPI) (Fig. 1C). Pony 524 maintained strong neutralizing antibody to EIAVwsu-5 throughout the remainder of the infection.

**Quasispecies distribution of the inoculum and statistical analysis of sampling procedure.** The overall goal of these studies was to examine rev variation at successive stages of clinical disease. A potential problem with current methods used to characterize viral
variants is the possible selection of a sub-population of variants that may not be reflective of the population present in vivo. To confirm that the selected primer pairs did not bias the distribution of observed variants, statistical analysis was performed on the distribution of variants obtained by RT-PCR amplification of the inoculum. These sequences were amplified using four different sets of primer pairs, designed from conserved regions flanking exon 2 of rev (Table 1). A total of 61 individual clones, at least ten from each primer pair, were sequenced from the EIAV_wyo inoculum. Thirty-eight different genotypes and 25 different Rev amino acid variants, designated R1 through R25, were observed (Fig. 2). Five genotypes were observed at a frequency greater than one, and two clones, R1 and R2, accounted for 30% and 25% of the sample population, respectively. These results indicated that the distribution of rev genotypes in the inoculum was not random.

A chi-square analysis testing the null hypothesis that the frequency distribution of the variants was independent of the primer pairs did not detect any statistically significant differences at either the nucleotide (p=0.31) or amino acid level (p=0.35). Six pairwise chi-square analysis using Bonferroni's correction (31) also failed to detect statistical significance at either the nucleotide level or the protein level. The lack of significant bias due to primer pair choice suggests the observed quasispecies population is reflective of the in vivo population. However, the sample sizes in the majority of cells in the contingency tables were lower than desired for chi-square analysis. Therefore, a Monte Carlo computer simulation (43) was performed to evaluate the validity of our chi-square analysis (Table 2). The comparison of all four primer pairs produced p-values of 0.30 and 0.36 at the nucleotide and protein levels, respectively. Simulations checking the pairwise chi-square analyses also produced p-values similar to those obtained by the chi-square analysis. The results of these simulations corroborated the chi-square analysis and further indicated that the selected PCR primer pairs did not bias the data obtained. Therefore, the frequency and distribution of the
Rev variants obtained by RT-PCR can be considered representative of the virus genotypes present in vivo.  

**Genetic changes in rev/gp45.** To analyze Rev variation throughout infection, RNA was isolated from sera samples at various stages of clinical disease including: Acute febrile (12 DPI), first afebrile interval (35 DPI), recurrent febrile (67, 89, and 118 DPI), afebrile (201, 289, 385, and 437 DPI), and late febrile (754 and 800 DPI) periods. Twenty-two to twenty-six rev exon 2/gp45 clones were amplified from each time point using primer pairs 2 and 3. Sequences from each time point represent data from two independent RT-PCR reactions. Genetic analysis indicated the transition (Ts) to transversion (Tv) ratio was high at each time point (Table 3), consistent with observations of other retroviruses (23). The overall Ts/Tv ratio for the region was 6.7 and application of the Tajima test (49) confirmed that variation in this region was due to selective pressure (p < 0.01). The ratio of non-synonymous (Kn) to synonymous (Ks) mutations was greater than one in both ORFs at each time and the overall Kn/Ks ratio was 1.54 for rev and 2.15 for gp45. Together, these data indicate that both rev and gp45 were under selective pressure throughout infection. Novel variants were detected at each time point; however, the appearance of new variants was more rapid during clinical periods. The greatest number of new variants for both rev and gp45 was observed at 67, 89, 118, and 201 DPI, corresponding to the middle and end of the recurrent fever period. Conversely, fewer novel genotypes were observed at 289, 385, 437, and 754 DPI, periods of sub-clinical infection. These results demonstrate that changes in the frequency and distribution in both ORFs correlated with the different stages of clinical disease.  

**Rev variant distribution throughout clinical disease.** There were marked changes in the frequency and distribution of the Rev aa variants obtained from successive stages of clinical disease (Fig. 3). Overall, variation was observed throughout the exon with the
exception of aa 89-99 and conserved changes in the C-terminus (e.g. R to K). In nine of the eleven time points there was a predominant variant that was identical to the consensus sequence for that time point. Indeed, the majority of variants from a particular time point were generally similar to the dominant variant. The predominant Rev variant in the inoculum, R1, also occurred at the greatest frequency in the samples from the acute fever episode (12 DPI) and subsequent afebrile period (35 DPI) suggesting that the inoculum phenotype was maintained early in infection. Several changes in the distribution of Rev variants occurred during the recurrent fever period. Notably, R1 was not detected during the recurrent fever period and two of the three time points had new dominant variants- R32 at 67 DPI and R53 at 118 DPI. No variant predominated at 89 DPI, however the consensus sequence was similar to the consensus sequences at 67 and 118 DPI. The change in Rev frequency and distribution during the recurrent febrile period suggests a shift occurred in selective pressure. At 201 DPI, the first serum sample analyzed after the febrile episodes subsided, R1 was observed again, but no variant was dominant. R1 became the major variant at 289 DPI and predominated throughout the remainder of the afebrile period (385 and 437 DPI). Hence, it appeared that another change in selective pressure occurred that favored R1. After the late fever episode at 565 DPI, a new variant, R93, arose as the dominant variant at 754 and 800 DPI. Although R1 was still observed at 754 DPI, it was not detected at 800 DPI suggesting another shift in pressure, similar to the recurrent period, that selected against R1. The dynamics of Rev evolution are highlighted by the disappearance and reappearance of the R1 variant which dominated early after infection and in the afebrile period, but was replaced by novel variants during both the recurrent and late febrile episodes. This clearly demonstrates that different selective pressures were present at distinct stages of clinical disease.
The activity of dominant Rev variants correlates with clinical disease. To assess if there was a correlation between Rev activity and state of clinical disease, the exon 2 cDNAs of the Rev variants dominant throughout the infection were cloned into a eukaryotic expression vector containing exon 1 of rev and their nuclear export activity determined using a transient transfection assay with a pDM138 reporter vector, pERRE-All (Fig. 4) (3). All assays were normalized to R1, the variant dominant in the inoculum, the acute fever period (12 DPI), first afebrile interval (35 DPI), and the afebrile period (289, 385, and 437 DPI). R2, a variant which contained a single aa change from R1 and was generally observed at the same time points, exhibited a similar level of activity to R1. Interestingly, all other variants assayed showed significantly higher levels of nuclear export activity as compared to R1 and R2 (p < 0.001). These included variants dominant or co-dominant at 67, 89, 118, 754 and 800 DPI, time points during or near clinical episodes. All of these variants exhibited approximately 150% activity relative to R1, except R17, which displayed 125% of the activity of R1. Together, these results demonstrate that Rev variants with higher activity are associated with periods of clinical disease (Fig. 5), indicating that variation in Rev may be an important factor of EIA disease pathogenesis.

DISCUSSION

Antigenic variation has long been recognized as an important mechanism of retrovirus persistence and pathogenesis. Recent studies suggest that variation in viral regulatory elements, including the LTR and accessory genes, also contributes to clinical disease (3, 6, 16, 27, 30). We previously observed a high degree of genetic variation in the rev exon 2/gp45 region of EIAV (3, 27) and demonstrated that limited amino acid variation in Rev alters biological activity (3). To delineate the significance of rev variation in vivo, we have examined sequence variation in an experimentally-infected animal. Discrete genetic and
biological changes were observed in Rev during sequential stages of an EIAV infection and genetic analysis demonstrated that the rev ORF evolved throughout infection. The evolution of rev quasispecies was highest during clinical periods, concomitant with higher levels of virus replication. Furthermore, changes in the frequency and identity of dominant Rev variants correlated with changes in clinical disease. The nuclear export activity of the Rev variants dominant during the recurrent febrile episodes was significantly higher than the activity of the variants predominant during the acute febrile or afebrile periods. Together, these findings argue that variation in rev may be a contributing factor of EIA disease pathogenesis.

Viremia has long been associated with clinical episodes in EIAV-infected animals (19). The abrogation of recurrent clinical episodes is typically co-incident with the maturation of the host immune response and decline in virus replication (12). Thus, we would expect that an absence of immune selection in the recurrent febrile period should favor an increase in virus load, whereas immune pressure in the afebrile period should result in a decrease of virus load. Although the appearance of broadly acting neutralizing antibody in pony 524 at 201-289 DPI was associated with the resolution of the febrile period, we did not observe the expected drop in virus load. Because Rev regulates virus replication, the observed increase in virus load from 35 to 67 DPI is consistent with the replacement of R1 by variants with significantly higher levels of nuclear export activity during the recurrent febrile period. However, viremia remained high until late in the afebrile stage of disease (754 DPI), despite the re-emergence of the less active R1 at 201 DPI. Therefore, it is not surprising that we did not observe variants less active than R1 in this period. Overall, changes in the activity of the dominant Rev variants did not correlate with virus load as well as with clinical disease in pony 524. It is possible that Rev variants less active than R1 might
have been detected if pony 524 had resolved to a "typical" subclinical state in which virus load decreased to near undetectable levels.

It is generally accepted that sequence variation in overlapping reading frames is much more constrained than variation in sequences with a single ORF. However, a previous report examining human immunodeficiency virus type 1 (HIV-1) variation in the \textit{rev/gp41} overlapping ORFs suggested that despite such constraints, both genes evolved independently (29). Our data would suggest the opposite for the EIAV \textit{rev/gp45} overlap. Several lines of evidence support this hypothesis. First, the evolution of both ORFs was surprisingly similar throughout infection. The rate of appearance of new \textit{rev} and \textit{gp45} variants was most frequent in both ORFs during periods of high virus replication. Second, out of 146 nt variants, there was a similar number of total Rev and gp45 aa variants (101 and 105, respectively). Third, the overall diversity (number of aa variants at each time point) of \textit{rev} and \textit{gp45} was approximately the same throughout infection. Finally, with the exception of 35 DPI, all dominant Rev variants were associated with a single gp45 variant. The only evidence that would suggest a divergent evolution is the fact that at seven of the eleven time points, there was a higher increase in the number of new \textit{rev} variants than \textit{gp45} variants. Therefore, we argue that the evolution of \textit{rev} and \textit{gp45} is convergent, and therefore we cannot rule out that changes in \textit{gp45} also contribute to the selective advantage of the observed Rev variants. Hence, variation in both regions may be an important determinant of virus pathogenicity \textit{in vivo}.

The underlying assumptions of the RT-PCR method are that all viral quasispecies are amplified with equal efficiency and the resulting proportions are representative of the starting population. While these assumptions have generally been accepted, some reports suggest that PCR amplification might result in selective amplification of individual templates, especially in reactions containing a heterogeneous population of templates (2, 42, 47, 48). Previous
studies have suggested that interactions between primers and templates might alter the efficiency of template amplification and thus alter the sample population (2, 14, 48), or that primer pairs may systematically bias PCR amplification (2, 42, 48). However, these studies have generally involved 2 or 3 different templates of known concentrations. The complex nature of viral quasispecies in vivo can further complicate sequencing studies. Our statistical analysis of the inoculum overcomes such complexities and shows that the primer pairs we used for PCR amplification did not alter the frequency of observed variants. The risk of bias may have been reduced by designing primers to conserved target sequences and using primers with similar GC-content and optimal annealing temperatures (Table 1). Taken together, our results suggest that randomly sampled viral populations obtained by RT-PCR can reflect in vivo populations. This offers confidence in the characterization of viral populations present in vivo, and strengthens interpretations of studies involving the role of virus variation in immune evasion and disease progression.

Many previous in vivo studies of EIAV variation have used a cloned, cell-culture propagated virus as an inoculum. Such an inoculum typically represents a homogenous population of virus. While such studies have been useful in identifying genetic changes that arise during an in vivo infection, our intention was to better reflect a naturally occurring infection in an experimentally-infected animal. Therefore, the inoculum used was obtained from a donor pony undergoing its first fever cycle and represented a heterogeneous population of virus instead of a single clone. Although this method required a more extensive analysis, we believe it provided an accurate representation of the complex nature of EIAV evolution throughout disease. The heterogeneous nature of the inoculum, in combination with the extensive genetic analysis at sequential time points, allowed us to fully evaluate the genetic and biological phenotype of Rev variants that were selected at variable stages of clinical disease. It is clear from this analysis that genetic and phenotypic changes in Rev
correlated with changes in clinical stages of disease. The corresponding changes in the
activity of the dominant Rev variants during the febrile periods provides strong evidence that
variation in Rev can confer a selective advantage in vivo.

REFERENCES


toward the wild-type k-ras and p53 sequences: implications for PCR detection of

72:4421-4426.

Cambridge.

novel cellular cofactor for the rev/rex class of retroviral regulatory proteins. Cell
82:485-494.

Identification of a hypervariable region in the long terminal repeat of equine infectious


### Table 1. Primer pairs used for amplification of EIAV Rev sequences

<table>
<thead>
<tr>
<th>Primer Pair</th>
<th>Primer Name *</th>
<th>Primer Sequence (5’-3’)</th>
<th>GC content(%)</th>
<th>O.A.T. b (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 EM 7066</td>
<td>CAATTTGGCACAATCCATGA</td>
<td>40</td>
<td>55.7</td>
<td></td>
</tr>
<tr>
<td>1 EM 7674C</td>
<td>GCGAGAGTTCCCTTCTTGCC</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 EM 7066</td>
<td>CAATTTGGCACAATCCATGA</td>
<td>40</td>
<td>55.7</td>
<td></td>
</tr>
<tr>
<td>2 EM 7680C</td>
<td>GGAATTCCGAGAGTTCCCTTCTTGCC</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 EM 7160</td>
<td>GGATGGGGAGCTTCCATTAT</td>
<td>45</td>
<td>55.5</td>
<td></td>
</tr>
<tr>
<td>3 EM 7674C</td>
<td>GCGAGAGTTCCCTTCTTGCC</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 EM 7160</td>
<td>GGATGGGGAGCTTCCATTAT</td>
<td>45</td>
<td>55.5</td>
<td></td>
</tr>
<tr>
<td>4 EM 7680CR1</td>
<td>GGAATTCCGAGAGTTCCCTTCTTGCC</td>
<td>50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a* Indicates the primer is on the complementary (anti-sense) strand, and R1 indicates the presence of an EcoRI restriction site (underlined) at the 3’ end.

bOptimal annealing temperature for PCR amplification.
Table 2. Pairwise analyses testing independence of primer pairs and the frequency distribution of observed quasispecies*

<table>
<thead>
<tr>
<th>Primer pairs analyzed</th>
<th>Nucleotide Variants</th>
<th>Protein Variants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CS^a MC^b</td>
<td>CS^a MC^b</td>
</tr>
<tr>
<td>1 and 2</td>
<td>0.34 0.14</td>
<td>0.65 0.68</td>
</tr>
<tr>
<td>1 and 3</td>
<td>0.25 0.42</td>
<td>0.27 0.22</td>
</tr>
<tr>
<td>1 and 4</td>
<td>0.65 0.44</td>
<td>0.27 0.25</td>
</tr>
<tr>
<td>2 and 3</td>
<td>0.36 0.46</td>
<td>0.29 0.25</td>
</tr>
<tr>
<td>2 and 4</td>
<td>0.49 0.27</td>
<td>0.61 0.64</td>
</tr>
<tr>
<td>3 and 4</td>
<td>0.17 0.13</td>
<td>0.07 0.05</td>
</tr>
</tbody>
</table>

^aP-values derived from chi-square analysis.
^bP-values derived from Monte Carlo simulation.

*Using Bonferroni's correction, these p-values are compared to 0.0083 for interpretations at the 95% confidence level. Thus, none of the pairwise analyses is statistically significant.

Table 3. Genetic Analysis of rev and gp45 ORFs

<table>
<thead>
<tr>
<th>DPI</th>
<th>IN</th>
<th>12</th>
<th>35</th>
<th>67</th>
<th>89</th>
<th>118</th>
<th>201</th>
<th>289</th>
<th>385</th>
<th>437</th>
<th>754</th>
<th>800</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total Clones:</td>
<td>61</td>
<td>24</td>
<td>23</td>
<td>23</td>
<td>24</td>
<td>24</td>
<td>25</td>
<td>22</td>
<td>26</td>
<td>23</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>Rev</td>
<td># Variants</td>
<td>25</td>
<td>7</td>
<td>11</td>
<td>13</td>
<td>18</td>
<td>11</td>
<td>16</td>
<td>8</td>
<td>12</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>New Variants</td>
<td>25</td>
<td>6</td>
<td>7</td>
<td>9</td>
<td>14</td>
<td>8</td>
<td>12</td>
<td>3</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Kn/Ks^1</td>
<td>2.2</td>
<td>6</td>
<td>7</td>
<td>4.3</td>
<td>4.5</td>
<td>3</td>
<td>3</td>
<td>2.2</td>
<td>2.8</td>
<td>4</td>
<td>3.5</td>
<td>1.6</td>
</tr>
<tr>
<td>gp45</td>
<td># Variants</td>
<td>30</td>
<td>7</td>
<td>10</td>
<td>14</td>
<td>16</td>
<td>11</td>
<td>15</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>New Variants</td>
<td>30</td>
<td>5</td>
<td>4</td>
<td>11</td>
<td>12</td>
<td>7</td>
<td>10</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Kn/Ks^1</td>
<td>3.5</td>
<td>*2</td>
<td>11</td>
<td>6</td>
<td>4</td>
<td>3.2</td>
<td>8.3</td>
<td>6.5</td>
<td>5.3</td>
<td>14</td>
<td>8.5</td>
<td>11.1</td>
</tr>
</tbody>
</table>

^1Ratio of non-synonymous to synonymous mutations in that ORF
^2No non-synonymous changes present in gp45 at that time point
FIGURE LEGENDS

FIG. 1. Clinical and serological profile of pony 524. (A) Clinical profile of EIAV_{wyo} infected pony 524. Platelet counts (gray line, primary y axis) and periods of fever (columns, secondary y axis) were monitored as described in Materials and Methods. Fever was defined as rectal temperature greater than or equal to 101°F. (B) Virus load, as measured by copies of RNA/ml serum, was determined by semi-quantitative RT-PCR for the days indicated by open circles. (C) EIAV_{wSU-5} neutralizing antibody (VN) titers of pony 524. VN titers were determined as described in Materials and Methods from serum samples obtained at 0, 11, 35, 67, 89, 118, 201, 289, 498, 754, and 800 DPI.

FIG. 2. Amino acid distribution of Rev sequences in EIAV_{wyo} inoculum. Twenty-five variants were obtained from 61 total clones sequenced. The frequency (N) and variant identity (V) are given. The aa sequence of the dominant variant (R1) is given, and the sequence of all other quasispecies are below: (.) Indicates an identical aa to R1 at that position and (-) indicates a -1 frameshift mutation.

FIG. 3. Alignment of Rev aa sequences obtained from sera of pony 524 at 12, 35, 67, 89, 118, 201, 289, 385, 437, 754, and 800 DPI. The frequency (N) and variant identity (V) are given for each time point, as well as the total number of clones sequenced (below the N column). Dominant variants are highlighted in bold. Sequences are compared to R1, which is listed at the top of each time point. For individual sequences (.) indicates an identical aa to R1 at that position, (*) indicates a premature stop codon in the rev ORF, and (+) indicates a +1 frameshift.
FIG. 4. *In vitro* biological activity of Rev variants from pony 524. Rev exon 2 variants were inserted into a eukaryotic expression vector and transient transfection assays were performed in 293 cells as described in the Materials and Methods. CAT assays were normalized to the activity of R1 and the results are presented as a percentage of activity compared to R1. Error bars denote the standard error of the mean.

FIG. 5. Variation in EIAV Rev correlates with clinical disease. A summary of (A) the activity of dominant Rev variants and (B) rectal temperature of pony 524 at various DPI. On days with no dominant variant (i.e. 89 and 201 DPI), the relative activity was determined as a weighted average of the most frequently occurring variants.
Figure 1
Figure 2
Figure 3
Figure 4
Figure 5
CHAPTER 4. BINDING OF EQUINE INFECTION ANEMIA VIRUS REV TO AN EXON SPlicing ENHANCER MEDIATES ALTERNATIVE SPlicing AND NUCLEAR EXPORT OF VIRAL MRNAs

A paper submitted to the journal *Molecular Cellular Biology*

Michael Belshan¹, Patricia Bilodeau², Gregory S. Park¹, C. Martin Stoltzfus², and Susan Carpenter¹

ABSTRACT

In addition to facilitating the nuclear export of incompletely spliced viral mRNAs, equine infectious anemia virus (EIAV) Rev regulates alternative splicing of the third exon of the *tat/rev* mRNA. In the presence of Rev, this exon of the bicistronic RNA is skipped in a fraction of the spliced mRNAs. In this report, the cis-acting requirements for exon 3 usage were correlated with sequences necessary for Rev binding and transport of incompletely spliced RNA. The presence of a purine-rich exon splicing enhancer (ESE) was required for exon 3 recognition, and the addition of Rev inhibited exon 3 splicing. GST-Rev bound to probes containing the ESE, and mutation of GAA repeats to GCA within the ESE inhibited both exon 3 recognition in RNA splicing experiments and GST-Rev binding in vitro. These results suggest that Rev regulates alternative splicing by binding at or near the ESE to block SR protein-ESE interactions. A 57 nt sequence containing the ESE was sufficient to mediate Rev dependent nuclear export of incompletely spliced RNAs. Mutation of the ESE significantly reduced Rev export activity indicating the ESE functions as a Rev responsive element (RRE). Together, these results demonstrate that EIAV Rev mediates exon 3 exclusion through protein-RNA interactions required for efficient export of incompletely spliced viral RNAs.

¹ Department of Veterinary Microbiology and Preventive Medicine, Iowa State University
² Department of Microbiology, University of Iowa
INTRODUCTION

Retroviruses utilize a variety of mechanisms to differentially express numerous proteins from relatively small genomes which possess a single transcriptional start site. These mechanisms include, but are not limited to, the use of polyprotein precursors, ribosomal frameshifting (19), alternate start codons (7), bicistronic mRNAs (7), and alternative splicing (28). Alternative splicing allows for the production of multiple viral mRNAs from a single RNA precursor. The simplest retroviruses, such as murine leukemia virus, express only two mRNAs, an unspliced RNA which serves as both mRNA for the Gag and Pol proteins and as progeny viral RNA, and a singly spliced mRNA which encodes the \textit{env} gene. In contrast, complex retroviruses, such as human immunodeficiency virus type 1 (HIV-1) produce at least 20 different mRNAs, including several multiply spliced RNAs that encode small regulatory proteins (28).

In all retroviruses, alternative splicing requires the presence of suboptimal splice sites, allowing for expression of several mRNAs from a single pre-RNA. Regulation of suboptimal splice sites is a complex process mediated in part by \textit{cis}-acting RNA sequences that can either enhance or repress recognition of a splice site by the splicing machinery. Exon splicing enhancers (ESEs) and silencers (ESSs) have been described for many virus and cellular RNAs (2, 22). ESEs typically are purine rich sequences embedded within alternatively spliced exons. The purine rich sequences mediate exon recognition through interactions with members of the SR protein family of splicing factors. SR proteins are both essential splicing factors and regulators of alternative splicing (reviewed in 17). Binding of SR proteins to an ESE recruits essential splicing factors to suboptimal splice sites near ESE sequences resulting in exon inclusion of alternatively spliced exons.

In addition to the above mechanisms, complex retroviruses utilize Rev-like pathways to differentially regulate expression of incompletely spliced RNAs encoding for virion
structural and enzymatic proteins and progeny RNA molecules (reviewed in 10). The prototype member of this family, HIV-1 Rev, binds to the viral pre-mRNA at a specific sequence called the Rev-responsive element (RRE) (9, 34), multimerizes (23, 33), and facilitates export of incompletely spliced RNAs from the nucleus via a nucleoporin pathway distinct from that used by cellular mRNAs (14,15). Equine infectious anemia virus (EIAV) Rev is functionally homologous to HIV-1 Rev (16), but is less well characterized. EIAV Rev is a 165 amino acid protein translated from exons 3 and 4 of a multiply spliced, 4-exon, bicistronic mRNA which also encodes the trans-activating protein, Tat (Fig. 1) (7). In addition to promoting nuclear export of incompletely spliced RNA, EIAV Rev also regulates inclusion of exon 3 of the multiply spliced RNA. In the presence of Rev, a multiply spliced mRNA lacking exon 3 is produced (21). Exon 3 is flanked by a suboptimal splice acceptor and contains a purine rich, exon splicing enhancer-like motif which has been shown to interact with the SR protein ASF/SF2 (18). This has suggested that Rev-mediated skipping of exon 3 may result from a disruption of ESE-SR protein interactions (18). We previously mapped a RRE of Rev to a 534 nt region containing exon 3 (4). This suggested the possibility that Rev mediates exon 3 skipping by binding at or near the purine rich sequence to disrupt SR protein interactions necessary for exon 3 recognition.

Here, we further delineate the role of Rev in exon 3 alternative splicing. Our results indicate that the purine rich sequence in exon 3 is required for the utilization of the exon 3 splice acceptor, confirming the presence of an ESE within exon 3. RNA gel mobility shift assays and nuclear export assays demonstrate that Rev binds to the ESE, and that this binding facilitates RNA export. Together, these results indicate the exon 3 ESE is a RRE of EIAV. Therefore, Rev mediates exon 3 alternative splicing by binding the viral pre-mRNA and interfering with SR protein – ESE interaction.
MATERIALS AND METHODS

PCR and Plasmid construction. All plasmid constructs were confirmed by sequence analysis (Iowa State University DNA Synthesis and Sequencing Facility). DNA templates for splicing substrates were amplified from p33k, a subclone of the p26 EIAV proviral clone previously described (6). Unless otherwise indicated, PCR reactions were performed as directed by the manufacturer (Perkin Elmer, Foster City, CA) using 1 μM primers. Standard PCR reactions consisted of 25 cycles of one minute denaturation at 94°C, one minute annealing at 50°C, and one minute extension at 72°C, followed by an addition cycle with a prolonged, five minute extension. All DNA templates for splicing substrates used a common 5’ primer, CTGAAGGCAATCCAACAAGG; and individual 3’ primers to generate the substrates shown in figure 2A. The 3’ primers used and the region of EIAV amplified were: CTCTCTATGATAAGCTTC, EIAV nt 5233-5793; CCAGTAGTTCCTGCTAAGCA, nt 5233-5573; TTTCCACCAGTCATTTCTTC, nt 5233-5535; CAGGGTCAATTCTTGGTCT, nt 5233-5490. All nucleotide numbering is based on that of Kawakami et al. (20). After PCR, fragments were TA cloned into the pGEM-T easy vector as directed by the manufacturer (Promega, Madison, WI).

The expression plasmid pRevWT was described previously as pCH21 (4). pDM138, pERE-All (EIAV nt 5280-7534), and pERE-1 (nt 5280-5834) have also been previously described (4). To construct pERE-1A, primers containing a ClaI restriction site were used to amplify EIAV nt 5281-5795. ERRE-1A 5’ primer: GGATCGATTAGATATGGGATTATT; 3’ primer: GGATCGATCTCTCTATGATAAGCTTC (ClaI sites are underlined). The minRRE sequence (EIAV nt 5485-5540) was synthesized as complementary oligonucleotides with ClaI extensions on the 5’ and 3’ ends. The oligonucleotides were heated at 95°C for 5 min and annealed by slow cooling. The fragment was phosphorylated, then ligated with pDM138.
The pGST-Rev expression vector contains a cDNA cloned in frame into the BamH1 site of the GST fusion vector pGEX-3X (Amersham Pharmacia Biotech, Piscataway, NJ).

ESE mutants were constructed by PCR-Ligation-PCR mutagenesis according to the methods described by Ali et al. (1) using internal primers designed with the specified mutations shown in figure 4A. The two regions were amplified with Vent DNA polymerase (New England Biolabs, Beverly, MA). The 3' fragment was phosphorylated, ligated with the 5' fragment, and 2 µl of the ligation reaction was PCR amplified with outer primers described above to amplify EIAV nt 5233-5793. Amplicons were cloned into pGEM-T. To construct the mutant pDM138 constructs, the pERRE-1A primers described above were used to PCR amplify the respective mutants from the pGEM-T background and clone into the Clal restriction site of pDM138.

**Synthesis of RNA substrates.** The plasmids containing the splicing substrates were digested with SpeI (GIBCO BRL, Rockville, MD) to create linearized templates for transcription of RNA splicing substrates. In vitro run-off RNA transcripts labeled with
[32P]UTP (Amersham Pharmacia Biotech) were generated as previously described (2). DNA templates for RNA binding analysis were amplified by PCR from p33k using 5' primers containing a T7 promoter site (a diagram of the substrates is in Fig. 3A). The primers used for the substrates were: 5’ primer:

TAATACGACTCACTATAGGGAGGAACAGCATGGCAGAATCG, 3’ primer:
TTTCCACCAGTCATTTCCTTTC (RREp1, nt 5443-5535); 5’ primer:
TAATACGACTCACTATAGGGAGGTGAGAGATACTTAAAG, 3’ primer:
CCACCAAAAGTATCTCTCC (RREp2, nt 5489-5589); 5’ primer:
TAATACGACTCACTATAGGGAGGTGAGGTGGAGAATAGG, 3’ primer:
CCCTATATAATGTTGCTG (RREp3, nt 5523-5622); 5’ primer:
TAATACGACTCACTATAGGGAGGCGAGGAAGAAGACGAG, 3’ primer:
CCTGCTAAGCATAACAGA (RRe4, nt 5458-5565). The T7 promoter is underlined in the 5' primers. Amplified DNA was phenol:chloroform extracted, ethanol precipitated, and resuspended in RNase-free dH2O. The RRe5 DNA fragment was synthesized as two complementary oligonucleotides containing the T7 promoter attached to EIAV nt 5485–5540. Complementary DNA fragments were combined at equal molar amounts, heated at 95° C for 5 min then slow cooled to anneal.

Expression/Purification of GST-Rev. BL21 Escherichia coli transformed with the pGST-Rev expression vector was grown overnight at 1/10 final culture volume in NZY broth containing 0.1 mg/ml ampicillin. The next day, cells were brought up to final volume, grown an additional 3 hours, then expression of GST-Rev induced with 1 mM IPTG for 5 h. After induction, cells were washed 3 times and resuspended in 50 mM Tris (pH 8.0), 50 mM NaCl (TN buffer). Cells were lysed by sonication and the supernatant clarified by centrifugation at 10,000 xg. GST-Rev was purified by binding to glutathione sepharose 4B beads (Amersham Pharmacia Biotech) overnight and washed 3X with TN buffer. The fusion protein was eluted with 15 mM reduced glutathione in 50 mM Tris, pH 8.0, concentrated with a 30 kDa MWCO filter concentrator (Millipore, Bedford, MA), and dialyzed against TN buffer. Protein expression was confirmed by SDS-PAGE and immunoblotting with convalescent anti-EIAV antisera or anti-GST antisera (Amersham Pharmacia Biotech) which detected expression of the GST-Rev fusion protein and several minor bands, including GST alone (data not shown). In some cases, the fusion protein was digested with 4 units/bead bed volume of Factor Xa protease (Amersham Pharmacia Biotech) while bound to glutathione sepharose beads. Excess GST and GST-Rev fusion protein were removed with the glutathione sepharose beads, and the supernatant, containing cut Rev protein, concentrated and dialyzed as described above.

In vitro splicing, and gel electrophoresis. Splicing reactions were carried out as previously described (2). In brief, approximately 8 fmol of EIAV RNA substrates were
incubated for 2 hours at 30°C with 60% (v/v) nuclear extract in Dignam's buffer D (13), 20 mM creatine phosphate, 3 mM MgCl₂, 0.8 mM ATP and 2.6% (w/v) polyvinyl alcohol. In some experiments, EIAV Rev protein was diluted in buffer D (13) and added to the splicing reactions at the indicated concentrations. RNAs were analyzed on 4% polyacrylamide gels containing 7 M urea.

**RNA binding Assays and gel electrophoresis.** RNA–protein interactions were performed in 1X RNA binding buffer containing 10 mM HEPES-KOH (pH 7.5), 100 mM KCl, 1 mM MgCl₂, 0.5 mM EDTA, 1 mM DTT, 50 μg/μl *E. coli* tRNA, and 10% glycerol. RNA was *in vitro* transcribed in the presence of [³²P]UTP as described above. 100 ng to 2 mg of GST or GST-Rev fusion protein was incubated with approximately 1 x 10⁶ cpm RNA probe on ice for 15 minutes. The reactions were directly loaded onto a 8% native 100 mM tris-glycine polyacrylamide gel (37.5:1 acrylamide:bis cross-linking ratio) which had been pre-run 1 hour. The samples were electrophoresed an additional 3 hours. The gel was fixed in 20% ethanol, 10% acetic acid for 15 minutes, dried, and exposed to X-ray film with an intensifying screen.

**CAT Assays.** 293 human embryonic kidney cells were maintained in Dulbecco’s modified eagle medium supplemented with 10% fetal calf serum and penicillin/streptomycin. Transient transfections and CAT assays were performed using 293 cells as previously described (4). Briefly, 1 μg of either pcDNA3 (Invitrogen, Carlsbad, CA) or pRevWT was transfected by calcium phosphate co-precipitation with 0.2 μg of pDM138 reporter plasmid, 0.2 μg of pCH110, and 0.6 μg of pUC19. Each experiment also included a sham group which contained no reporter plasmid, but an additional 0.2 μg of pUC19. Two days posttransfection, cells were harvested, resuspended in 0.3 ml 0.25 M Tris (pH 7.5), lysed by freeze/thawing, and assayed for β-galactosidase activity to normalize CAT assays for transfection efficiency. Normalized lysates were assayed for CAT activity using 3 μl
RESULTS

The purine rich sequence is required for exon 3 recognition. Exon 3 of the bicistronic, four exon, equine infectious anemia virus (EIAV) mRNA contains a purine rich sequence which resembles an exon splicing enhancer (ESE). Previous reports showed that the SR protein ASF/SF2 cross-links to the ESE-like sequence in vitro and suggested the ESE-like sequence may enhance exon 3 inclusion during pre-mRNA splicing (18). To investigate the cis-acting requirements for exon 3 inclusion, we constructed a series of DNA templates to generate radiolabelled RNA substrates for in vitro splicing. All substrates contained the exon 2 splice donor, the intervening intron, exon 3, and downstream sequences. Nested 3' deletions were made to identify splicing enhancer sequences present within or downstream of exon 3 (Fig. 2A). Splicing of radiolabelled substrates was assayed in vitro using HeLa cell nuclear extracts, which include all SR proteins. All constructs containing the purine rich sequence were spliced (Fig. 2B, lanes 3-5), whereas no splicing was observed using the substrate lacking the purine rich sequence (Fig. 2B, lane 2). This is consistent with the hypothesis that the purine sequence functions as an ESE and is required for exon 3 inclusion in the multiply spliced four exon transcript. Taken together with previous work (18), this suggests exon 3 recognition requires ASF/SF2 interactions at the ESE. The addition of as little as 100 ng of Rev to a splicing reaction containing the largest splicing substrate inhibited exon 3 recognition (Fig. 3). This confirms earlier, in vivo observations of Rev-mediated
changes of alternative splicing, and indicates that Rev is the only viral protein necessary for exon 3 skipping.

**Rev binds the ESE.** Disruption of ASF/SF2-ESE interaction by Rev could occur either by direct interaction of Rev with ASF/SF2 or by binding of Rev to a region of the viral pre-mRNA near the ESE. In previous work, we identified an RRE region spanning exon 3 (nt 5280 to 5834) (4). This favors a mechanism wherein Rev-RRE interactions disrupt ASF/SF2 binding at the ESE. To examine whether Rev binds at or near the ESE, we generated a series of RNA probes and tested for Rev-RNA interactions by RNA gel mobility shift assays using a bacterially expressed and purified GST-Rev fusion protein. The location of the RNA probes relative to exon 3 and the ESE is shown in figure 4A. GST-Rev bound to exon 3 probes RREP1 and RREP2, which both contain the ESE; however, no binding was observed with the GST negative control (Fig. 4B, lanes 1 to 4). Minor binding was observed with RREP3 (Fig. 4B, lane 6), which has sequences immediately downstream of the ESE (nt 5523 to 5622), but lacks the purine rich region. The binding site was further delineated to a 57 nt region of viral RNA using two smaller ESE-containing probes, RREP4 and RREP5. GST-Rev interacted with both probes (Fig. 4B, lanes 8 and 11), further suggesting that Rev binds at or near the ESE. To confirm the specificity of binding, gel shift assays were performed with RREP2 in the presence of excess cold competitor RREP2 or RREP3 (Fig. 4C). Excess cold RREP2 inhibited GST-Rev binding (Fig. 4C, lanes 2 and 3), whereas no inhibition was observed with RREP3 (Fig. 4C, lanes 4 and 5), demonstrating that the binding of GST-Rev to the ESE-containing RREP2 is specific. No slower migrating bands, indicative of Rev multimerization, were observed in any of the RNA binding analysis, although multimerization was readily observed with HIV-1 Rev when used as a positive control (data not shown). The minor bands in figure 4C represent GST-Rev and degraded by-products. Overall, these results suggest that Rev specifically interacts with the viral RNA at or near the ESE.
Mutagenesis of the ESE reduces exon 3 splicing and Rev binding. The finding that GST-Rev bound to a 57 nt region containing the ESE suggested that a Rev-RNA interaction was directly competing with ASF/SF2 for binding at the ESE. If so, there should be similar sequence requirements for exon 3 recognition and Rev binding. The ESE contains two purine rich sequences (designated A and B), which include seven GAA repeats (Fig. 5A). GAA repeats have been shown to be important for ASF/SF2 recognition of ESE sequences in other systems (reviewed in 17). Therefore, we constructed five ESE mutants in the largest splicing construct and the RREp4 RNA probe fragment which contained various GAA to GCA mutations (Fig. 5A). The mutant templates were tested for in vitro splicing and GST-Rev binding. Mutation of all GAA motifs (mutAll) or the B purine stretch (mutB) resulted in a decrease in both exon 3 in vitro splicing (figure 5B, lanes 3 and 5) and in GST-Rev binding (figure 5C, lanes 2 and 4). Mutation of only two of the four GAA repeats in the B purine stretch resulted in a more modest reduction in both in vitro splicing and GST-Rev binding (Fig. 5B, lanes 6 and 7 and 5C, lanes 5 and 6). Mutation of GAA repeats in the A region (mutA) did not appear to significantly affect either exon 3 splicing or GST-Rev binding in vitro (Fig. 5B, lane 4 and 4C, lane 3) suggesting the B purine stretch alone contains cis-acting sequences necessary for exon 3 recognition and GST-Rev binding. The finding that each mutant had comparable effects in both assays suggests similar requirements in the ESE for both exon 3 recognition and Rev binding. These results support a model of Rev inhibition of splicing through direct competition with SR proteins for binding at the ESE.

The ESE can function as a RRE to mediate RNA nuclear export. In other complex retroviruses, Rev functions to regulate the export of incompletely spliced RNAs via interaction with the viral pre-mRNA at a specific sequence called the Rev responsive element (RRE). We had previously used the pDM138 reporter system to preliminarily map the RRE of EIAV to a region which overlapped exon 3 (ERRE-1, nt 5280-5834) (4). However, this
fragment possessed only 52% of the activity of a reporter containing a much larger fragment of EIAV (ERRE-All) suggesting further downstream sequences enhanced Rev-mediated export. The RNA binding data given above suggested that the functional sequence within ERRE-1 was the ESE sequence. To test this, we constructed a pDM138 reporter plasmid, minRRE, which contains only 57 nt of EIAV (RRe5, nt 5485–5540), spanning the ESE and the remainder of exon 3. Transient transfection assays in 293 cells demonstrated the minRRE reporter produced levels of CAT activity comparable to ERRE-1 (Fig. 6A), but only 35% of the activity observed with ERRE-All. This indicates minRRE contains the functional RRE in ERRE-1, however additional elements outside ERRE-1 may be required for full export activity.

To confirm the ESE is the RRE within minRRE, we introduced the GAA to GCA mutations used for in vitro splicing and RNA binding assays (Fig. 5A) into a reporter vector containing the same sequences present in the largest splicing substrate (Fig. 2A). This vector, ERRE-1A, although 41 nt shorter than ERRE-1, exhibited similar levels of activity as ERRE-1 (Fig. 6A). In all cases, mutation of the GAA repeats in the ESE significantly reduced Rev dependent nuclear export activity (p < 0.01) (Fig. 6B). The greatest reduction in activity was seen in constructs containing mutations of all seven GAA repeats (mutAll) or the three repeats in the A purine stretch (mutA). The reduction in activity in mutA indicates this region, while not necessary for GST-Rev binding, is required for RNA nuclear export. Mutations in the B region (mutB, B12, B34) also significantly reduced activity (p < 0.01). Therefore, we conclude that the B purine stretch functions in GST-Rev binding, exon 3 inclusion, and Rev-dependent nuclear export. Together, these results indicate the ESE sequence acts as an RRE and that Rev mediates alternative splicing by binding at or near the ESE to disrupt ASF/SF2 interactions.
DISCUSSION

These studies demonstrate that the purine rich sequence within exon 3 of equine infectious anemia virus (EIAV) is both an exon splicing enhancer (ESE) and a functional Rev responsive element (RRE). The use of an ESE as an RRE is unprecedented among complex retroviruses. A prior study demonstrated that the SR protein ASF/SF2 cross-links in vitro with exon 3 and suggested that the purine rich sequence was an ESE (18). Consistent with this previous data, we show the purine rich sequence is required for exon 3 recognition in splicing reactions in vitro. Taken together, these data suggest ASF/SF2 is the SR protein required for efficient splicing of exon 3. The ability of a 57 nt sequence containing the ESE to interact in vitro with GST-Rev and act as a functional RRE in a heterologous nuclear export assay system demonstrates that this region contains a functional RRE. Mutation of GAA nucleotide repeats in the ESE reduced GST-Rev binding, exon 3 splicing in vitro, and nuclear export of ESE containing pre-mRNA. Therefore, both exon 3 splicing and Rev-mediated RNA export have similar cis-acting RNA requirements. We propose that Rev-mediated nuclear export requires binding at or near the ESE. This results in skipping of exon 3 through an inhibition of ASF/SF2 - ESE interactions required for recognition of exon 3 by the host cell splicing machinery.

It is interesting that mutation of the 5' purine stretch (mutA) decreased nuclear export but appeared to have little effect on GST-Rev binding in vitro. Studies with HIV-1 Rev also indicate that sequences in the HIV-1 RRE are required for nuclear export but not RNA binding (24, 25). While these observations have not been fully explained, it is likely that RNA secondary structure plays a role. Secondary structure is a key determinant for HIV-1 binding, multimerization, and function (8, 9, 11, 12, 23, 24, 26). Since no biochemical data is available to date on the structure of the EIAV ESE/RRE, it is not possible to determine the structural effects of the mutations used in our study. Our data would suggest that mutation of
the 5' purine stretch does not affect the primary binding site of Rev, but may alter distant structures required for Rev-mediated nuclear export. It has been demonstrated that HIV-1 Rev multimerization occurs only after binding to a primary site on the RNA and other regions of the RNA are important for secondary binding (9, 11, 31). However, we were unable to observe Rev multimerization in our RNA binding assays, including those assays containing the ESE. Therefore, it remains unclear why mutA exhibited reduced activity with no apparent defect in RNA binding. It is also possible that binding of host cell proteins to the RRE may be required to facilitate Rev export activity. Further studies will be necessary to delineate the role of this purine region in Rev-mediated nuclear export.

Interactions of Rev-like proteins with SR proteins have been demonstrated in other complex retroviruses. SR proteins have been shown to bind the HIV-1 RRE in a Rev-dependent manner (27). The same study also demonstrated that excess exogenous ASF/SF2 could produce a dose-dependent inhibition of HIV-1 Rev function in vivo. We have previously reported an inhibition of EIAV replication in activated macrophages associated with a delay in the appearance of incompletely spliced RNAs (30). The data presented here would suggest this inhibition may be a result of competition of ASF/SF2 with Rev for binding at the ESE. In support of this hypothesis, preliminary data in our laboratory has shown that excess ASF/SF2 provided in trans can inhibit Rev function in transient transfection assays and reduce virus particle production from a proviral containing cell line (5). Together, these results suggest the inhibition in activated macrophages may be due to an increase in the level of SR proteins. It is known that expression of SR proteins varies in cells at different states of activation and differentiation (17, 29, 32), including an increased expression of the SR protein SRp30c in activated T-cells (29). However, little is yet known about the phenotype of SR proteins in monocyte cells. Also, our data cannot rule out the
possibility that in addition to competing for binding at the RRE, Rev may also inhibit ASF/SF2 function via protein-protein interactions.

Whether alternative splicing of exon 3 is merely a consequence of Rev-mediated nuclear export or plays an important role in virus replication has not been determined. The RREs of most complex retroviruses are located near the SU-TM cleavage site or in the 3' end of env. The location of EIAV RRE in the 5' env is unique, and may be explained by the requirement of the ESE for exon 3 inclusion. The env mRNA is spliced using the exon 2 splice acceptor (Fig. 1). A singly spliced mRNA using the exon 3 splice acceptor has not been observed in infected cells, and would encode a truncated Env protein lacking the signal peptide. A singly spliced mRNA using the exon 4 splice acceptor is observed, which produces a truncated transmembrane protein from the alternate start codon present in exon 1. Therefore, the use of an ESE as an RRE may function to silence recognition of exon 3 to eliminate another singly spliced transcript. Although a number of retroviruses utilize cis-acting sequences such as ESEs to take advantage of cellular mechanisms of alternative splicing, EIAV appears to be the only retrovirus to encode a trans-acting protein that directly modulates SR-mediated alternative splicing. EIAV was the first described lentivirus and is smaller and genetically less complex than the other lentiviruses. It is possible that the EIAV Rev-ESE interaction may represent a transitional step in the evolution of the Rev-Rex pathway utilized by most complex retroviruses. Interestingly, previous work in our laboratory and others suggested that EIAV may possess two separate RREs (4, 21). In the current study, reporter constructs containing the 57 nt ESE region showed significantly reduced activity as compared to the ERRE-All reporter construct containing a larger portion of the env gene (Fig. 6A). However, a second RRE element has not been identified, nor is it clear that such an element can function independently of the ESE to mediate export of viral
pre-mRNAs. Additional studies will help to fully understand the biological and evolutionary significance of the EIAV Rev-mediated export pathway.

ACKNOWLEDGEMENTS

We thank Yvonne Wannemuehler for technical assistance, Tom Hope for plasmids pERRE-All, pERRE-1, and pDM138, and Sean Murphy for statistical analysis.

This work was supported by funds from the Carver Grant Trust (S.C.), an Iowa State University - University of Iowa Interinstitutional grant in Biomedical Sciences (S.C. and C.M.S.), USDA grant 96-358204-3847 (S.C.), and PHS grant AI36073 from the National Institute of Allergy and Infectious Disease (C.M.S.).

REFERENCES


FIGURE LEGENDS

**Fig. 1.** Organization and splicing patterns of EIAV. Schematic of EIAV genome with ORFs. The tat ORFs are indicated with a 't', the first exon of rev with a 'r' (location in genome is indicated by the shaded region). Splicing patterns and genes expressed are indicated. The tm ORF encodes a truncated transmembrane protein of unknown function (3).

**Fig. 2.** Exon 3 splicing requires the purine rich sequence. (A) Diagram of RNA substrates used for *in vitro* splicing showing the location of exons 2 and 3. All substrates contain the exon 2 splice donor and exon 3 splice acceptor. Approximate location of purine rich sequence is indicated by the hatched box. (B) After incubation for 2 h in HeLa cell splicing extracts, RNA products were electrophoresed through 4% polyacrylamide gels and visualized by autoradiography. Location of spliced and unspliced products is shown.

**Fig. 3.** Rev inhibits exon 3 splicing. GST-Rev or GST was added at the indicated concentrations to the splicing reactions. Location of the splicing products is indicated.

**Fig. 4.** GST-Rev binds the ESE in exon 3. (A) Location of RNA probes, relative to exon 3, used in RNA gel mobility shift assays. The purine rich ESE sequence is indicated by the hatched box. (B) After incubation with GST or GST-Rev, radiolabelled RNAs were electrophoresed through 8% native polyacrylamide gels. Location of GST-Rev-RNA complexes is indicated with an arrow. (C) Competition assays were performed with either 0.5 or 1 µg of the indicated excess cold competitor RNAs. Competitors were mixed with GST-Rev 10 min prior to the addition of radiolabelled probe.
**Fig. 5.** *In vitro* splicing and RNA binding of ESE mutants. (A) Sequence of two purine stretches (designated A and B) in exon 3. GAA repeats were mutated to GCA in largest splicing construct (Fig. 2A) and RNA probe RREp4 (Fig. 4A). (B) *In vitro* splicing analysis of mutant ESE constructs. Location of splicing products is indicated. (C) RNA gel mobility shift assays detecting GST-Rev binding to the mutant probes.

**Fig. 6.** EIAV ESE can function as an RRE. (A) pDM138-derived reporter vectors containing various regions of the EIAV genome. Transient transfections and CAT assays were performed as described in Materials and Methods. The results are presented as the percentage acetylation. Experiments were performed in triplicate and the results represents at least nine independent transfections. Error bars denote the standard error of the mean. (B) ESE mutations indicated in figure 5A were also introduced in the ERRE-1A reporter vector and assayed for CAT activity in the presence or absence of Rev as described above.
Figure 1
Figure 2
A. Pu-A  Pu-B
Wild-type GCAGAAATCTAAGAGAAAGAAAAGAAAAAAGAGAGAAA
mutAll .C...C...C...C...C...C...C...C...C...C...C...
mutA .C...C...C................................................................
mutB .................................................................C...C...C...C...C...
mutB12 .................................................................C...C...C...C...
mutB34 .................................................................C...C...C...C...

B. a>S — cn

\[ \begin{array}{c}
506/517 - 396 - 344 - 298 - 220 - 201 - 154 - \\
1 - 2 - 3 - 4 - 5 - 6 - 7
\end{array} \]

C. GST-Rev +  Wild-type  mutAll  mutA  mutB  mutB12  mutB34

\[ \begin{array}{c}
1 - 2 - 3 - 4 - 5 - 6
\end{array} \]

Figure 5
A.

B.

Figure 6
CHAPTER 5. MAPPING THE ALTERNATIVE SPLICING DOMAIN OF EQUINE INFECTIOUS ANEMIA VIRUS REV

A selection from a paper that will be submitted to the journal Virology

Michael Belshan, Sean C. Murphy, and Susan Carpenter

INTRODUCTION

Complex retroviruses (e.g. HIV-1, EIAV) are classified as exhibiting temporal regulation of gene expression during infection of a cell. The early phase is characterized by the presence of small, multiply-spliced mRNAs encoding virus regulatory proteins (Fig. 1). For EIAV, these proteins include Tat and Rev. Tat, a transcription transactivator, increases the level of viral RNA transcription up to 1000-fold. Tat acts by binding a structure near the transcription initiation site in the newly synthesized viral mRNA (called the TAR element) and increasing the processivity of transcription (7). The regulatory protein Rev post-transcriptionally regulates the expression of incompletely spliced viral mRNAs that code for virion structural proteins, envelope glycoproteins, and provide for progeny RNA molecules. Upon reaching a threshold concentration in the host cell, Rev facilitates a switch from the early gene expression to late gene expression by directing the nuclear export of the incompletely spliced viral RNAs. In the absence of Rev, virus replication does not occur.

Rev-dependent RNA export pathways have been described for numerous retroviruses (reviewed in 8), and the HIV-1 Rev-mediated RNA export pathway is the best characterized. HIV-1 Rev enters the nucleus, binds a structure in the viral pre-mRNA called the Rev Responsive Element (RRE) (6, 14, 24), multimerizes (15, 16, 20, 23), then utilizes a non-mRNA nuclear export pathway to redirect movement of incompletely spliced RNA from the nucleus (4, 9, 10, 22). Discrete functional domains within the protein mediate the interaction of Rev with both cellular proteins and the viral RNA. These domains include nuclear
localization, RNA binding, multimerization and nuclear export.

EIAV Rev is a highly charged, 165 amino acid protein translated from a bicistronic four exon mRNA coding for both Tat and Rev (Figure 1) (5, 21). Though functionally homologous to HIV-1 Rev, EIAV Rev shares little amino acid homology. The nuclear export signal (NES) of EIAV Rev has been mapped to amino acids 31-55 (11), and domain swapping experiments have shown that this region of EIAV Rev can substitute for the HIV-1 or visna virus NES (4, 17, 19). A putative nuclear localization signal has also been mapped to the arginine rich C-terminus of the protein (13).

EIAV Rev also has an additional, apparently unique function among complex retrovirus Rev-like proteins. Whereas HIV-1 Rev and human T-cell leukemia virus-1 Rex inhibit the expression of both Tat and Rev specific mRNAs in facilitating the export of incompletely spliced mRNAs, EIAV Rev specifically down-regulates its own production, independent of Tat, by promoting exon 3 skipping of the bicistronic mRNA (Fig. 1) (12, 18). This mechanism allows for continuous production of Tat while Rev synthesis is reduced. Inclusion of exon 3 in the mRNA requires the interactions of SR proteins with an exon splicing enhancer (ESE) within the exon sequence (3, 12). Rev mediates exon 3 skipping by binding to a RRE overlapping exon 3, which interferes with SR protein-ESE interactions required for exon 3 inclusion. The multi-functional nature of EIAV Rev highlights its importance in regulation of virus gene expression and replication. As such, genetic and biological variation in Rev may have added significance in vivo.

Here we present a functional mapping of EIAV Rev by deletion analysis. In vitro assays of Rev activity revealed that the protein is highly sensitive to deletions. Analysis of Rev mediated alternative splicing indicated that amino acids 75-127 of the protein are important for exon 3 skipping. The ability of a NES-deficient Rev mutant to still mediate alternative splicing demonstrated that these two functions of Rev are independent.
RESULTS

Construction/Expression of Rev deletion mutants. A series of deletion mutants of EIAV Rev was constructed to map domains essential for Rev activity (Fig. 2A). Deletions were constructed by PCR-ligation-PCR mutagenesis (1) and cloned into the eukaryotic expression vector pCR3.1. Deletions were confirmed by sequence analysis, then protein expression confirmed by Western blot analysis of transiently transfected cells using convalescent anti-EIAV horse sera (Fig. 2B). The Western blot results indicated that all cDNAs express Rev and would be usable in in vitro assays. However, not all mutants expressed Rev at comparable levels. All those with alternate start codons (e.g. Δ1-11, Δ1-19, Δ1-30) appeared to express Rev at lower levels. Changes in apparent protein levels could have also resulted from decreased antibody reactivity to proteins with deletions (e.g. Δ101-113).

Functional analysis of Rev deletion mutants. To determine the effects of the various deletions on Rev nuclear export activity, we functionally assayed the Rev mutants with a reporter construct, pERRE-All, in which the CAT gene and a region containing putative EIAV RRE sequences are flanked by HIV-1 splice sites (2, 14). Since Rev inhibits splicing, the production of CAT is proportional to the activity of the Rev. Functional assays were performed using transient expression assays as previously described (2). Briefly, 293 human embryonic kidney cells were seeded in triplicate in 6-well tissue culture dishes and transfected with mutant Rev cDNA or empty vector, reporter plasmid, and beta-galactosidase reporter plasmid. Two days posttransfection cells were harvested, resuspended in 300 μl Tris buffer, and lysed by three cycles of freeze-thawing. 50 μl of lysate was assayed for beta-galactosidase activity and these values were used to normalize lysates for CAT assays. CAT assays were performed using [14C]chloramphenicol and unacetylated and acetylated forms were separated by thin-layer chromatography and quantified using a phosphorimager.
(Molecular Dynamics). The percentage of acetylation was calculated for each transfection, and the average acetylation and standard error of the mean were calculated for all experiments.

To use a more relevant biological method to measure Rev mutant activity, we developed a virus complementation assay. This method utilizes a canine fetal thymus (Cf2Th) cell line stably transfected with a Rev-defective EIAV provirus (Cf2Th/51 cells) (2). Cf2Th/51 cells were trans-complemented with the Rev mutants and the release of virus particles into the cell supernatant was measured using a \[^{32}\text{P} \] -based assay for the detection of reverse transcriptase activity. Cf2Th/51 cells were transfected with each Rev mutant and three days post-transfection 100 µl of supernatant sampled. To assay reverse transcriptase activity, 10 µl of the supernatant was lysed and reverse transcriptase activity measured by the incorporation of \[^{32}\text{P} \] TTP onto a poly(rA) template. Incorporation was measured using a phosphorimager. Assays were performed in triplicate and standard error of the mean calculated.

Results from both assays are shown in figure 3. Overall, the analysis indicated that Rev is highly sensitive to deletion mutagenesis. Almost all of the mutants exhibited decreased activity in comparison to wild-type Rev. Only one mutant (Δ131-143) retained wild-type levels of activity in both assays. Certain mutants (Δ1-19, Δ1-30, Δ164-165) may have decreased activity due to poor expression. However, other mutants (Δ75-82, Δ144-165) expressed at higher levels and still remained non-functional, indicating that the loss of activity was likely due to deletion of an important functional domain. Loss of activity did not consistently correlate with the observed expression levels from Western blots. The C-terminus of the protein contains a highly conserved KRRRK motif which has been proposed to be the nuclear localization signal (NLS) of Rev (13). Deletion of this region or the center portion (aa 75-127) of the protein proved highly deleterious, and therefore demonstrates both
domains are important for function. The center region, aa 75-127, contains two basic regions and is predicted to be α-helical in structure. Therefore, it is likely that this region contains motifs important for RNA binding and multimerization.

**Amino acids 75-127 are required for Rev mediated alternative splicing.** The current model of EIAV alternative splicing proposes that exon 3 skipping is mediated by disruption of an interaction between SR-proteins and an exon splicing enhancer (ESE) sequence in exon 3 by the binding of Rev to an RRE in the same region (3, 12). To determine the ability of each deletion mutant to promote exon 3 skipping, splicing patterns were determined in Cf2Th/51 cells trans-complemented with each of the mutants. Cf2Th/51 cells were transfected with each mutant and total cellular RNA was isolated three days post-transfection. $[^{32}P]$End-labeled RT-PCR was employed to detect the presence/absence of the alternative spliced mRNA (Fig. 4). Trans-complementation with wild-type Rev resulted in approximately 50% alternative splicing versus a positive control cell line (Cf2Th/112) containing a Rev-competent provirus (approximately 17% alternative splicing) (Fig. 4).

The results demonstrated that the center portion of the protein (aa 75-127) is required for exon 3 skipping. This region is also required for RNA export (Fig. 3) and therefore may contain either the RNA binding or multimerization domain. Interestingly, the deletion of the highly basic C-terminus (aa 144-165) which contains a proposed NLS of Rev (13) was deleterious to RNA export activity, but did not adversely affect alternative splicing. However, it is likely that the small size of the protein permitted passive diffusion into the nucleus. Deletion of the NES of Rev (aa 32-55) did not adversely affect alternative splicing. This indicates Rev-mediated RNA export is independent of alternative splicing, although both functions require RNA binding. Taken together, these results validate the currently proposed model of EIAV alternative splicing and more clearly define the importance of alternative splicing in terms of EIAV replication.
DISCUSSION

Here we demonstrate that, in addition to the nuclear export signal (NES), two regions (aa 75-127 and 145-165) of equine infectious anemia virus (EIAV) Rev are necessary for biological activity. In addition, aa 75-127 are required for Rev to mediate exon 3 alternative splicing. The majority of the deletion mutants exhibited reduced activity in vitro, indicating that EIAV Rev is highly sensitive to deletion mutagenesis. Only one region (aa 131-144) could be deleted and still maintain wild-type levels of function.

One of the two areas of the protein especially sensitive to deletions (aa 145-165) has recently been proposed to contain the nuclear localization signal of Rev (13). Consistent with that finding, deletion mutants in this region were non-functional. Surprisingly however, deletion of this highly basic, C-terminus of the protein, did not inhibit Rev mediated alternative splicing. It is likely that the small size of the mutant (~16 kDa) allowed it to passively diffuse into the nucleus and still promote exon 3 skipping. Further experiments will be necessary to confirm this hypothesis.

Consideration of the mechanism by which EIAV Rev-mediates alternative splicing provides further insights into the possible function of aa 75-127. Exon 3 skipping results from a disruption of SR protein-ESE interactions when Rev binds a RRE in exon 3. Given that alternative splicing results from RRE binding and that mutants deleted in aa 75-127 are deficient in this function, it is likely that the RNA binding and multimerization domains are localized within this region of the protein. Further in vitro studies to identify the RNA binding domain of Rev will more clearly delineate the mechanisms of Rev-mediated RNA export and alternative splicing. However, given the broad, deleterious effects of our deletion analysis, more sensitive analysis, such as site-directed mutagenesis, may be necessary to define specific amino acids requirements of Rev function.


Figure 1. Splicing patterns of EIAV. Structural proteins (Gag, Env) and viral enzymes (Pol) are translated from the unspliced and singly-spliced mRNAs. Both Tat and Rev are encoded on the triply spliced, four exon mRNA. In the presence of Rev a new multiply-spliced mRNA appears in which exon 3 is skipped. This mRNA codes for Tat alone. The two exon mRNA codes for Ttm, a protein of unknown function.
Figure 2. (A) Deletion mutants of EIAV Rev for functional analysis. (B) Expression of deletion mutants. 1x10^5 293 cells were transfected with 5 μg of mutant plasmid. Two days posttransfection, cells were harvested and a Western blot performed. Rev mutants were detected using anti-EIAV polyclonal horse sera.
Figure 3. In vitro activity of Rev deletion mutants. (A) Transactivation of Rev reporter plasmid ERRE-All by deletion mutants. ERRE-All contains the CAT gene and an EIAV RRE flanked by HIV-1 splice sites. Transfections were made into 293 cells and normalized by beta-galactosidase activity from co-transfected pCH110. Rev activity is measured as a production of CAT. Mutants are transfected in triplicate per experiment and results are from three experiments, standard error of the mean is shown. (B) Complementation of a Rev-defective provirus by deletion mutants as measured by cell-supernatant reverse transcriptase (RT) activity. 1 x 10^5 cells of Cf2Th/51 cells were transfected with 2 µg of each mutant plasmid. Three days posttransfection, 10 µl of supernatant was assayed for RT activity using [32P]TTP. RT activity was quantified by phosphoimager and the standard error of the mean is shown.
Figure 4. Alternative splicing in cells *trans*-complemented with Rev deletion mutants. Cf2Th/51 cells, containing a Rev-defective provirus, were transfected with Rev deletion mutants. Total RNA was isolated and [{}^{32}P]end-labeled RT-PCR performed using primers within exon two and four (see Fig. 1). DNA was extracted and separated through a denaturing polyacrylamide gel. The gel was dried and bands were quantified using a phosphorimager and the percentage of alternatively spliced message is shown. Cf2Th/112 cells, a positive control cell line, contain a Rev-competent provirus.
CHAPTER 6. GENERAL CONCLUSION

The complex retroviruses are defined not only by a complicated genome, but also by a pattern of gene expression dependent upon the action of viral regulatory proteins. The use of regulatory proteins creates a mechanism through which variation can modulate virus replication and provides a potential factor of pathogenicity. EIAV is unique both as a lentivirus and as a complex retrovirus which makes the virus an excellent tool for comparative studies. EIAV is one of the simplest complex retroviruses. Whereas HIV-1 has a 9.7 kb genome and expresses more than 20 mRNAs with nine gene products, the EIAV genome is a modest 8.3 kb, has five mRNAs, and seven gene products. Lentiviruses are typically characterized as causing slow, chronic diseases which may take years to progress. EIAV however, can induce a rapid, variable disease course almost immediately upon infection of a horse. These are just some of the factors make EIAV an appropriate model for analyzing virus persistence and pathogenesis.

Variation in EIAV Rev Correlates with Disease

Sequence analysis of EIAV cDNAs from both in vivo and in vitro contexts indicated the second exon of rev contains a hypervariable region. Sequence variation in retroviruses is attributed to errors in nucleotide incorporation by reverse transcriptase (5, 7, 9, 11). These errors are generally randomized throughout the genome and then selected during virus replication. We found the variation in Rev to be localized to two primary regions- the activation domain and a stretch of amino acids bordered by QERL repeats (Chapter 2, Fig. 2). It is possible that immune selection may account for some of the variation in Rev since the major rev exon (exon 4) overlaps with the ORF of gp45, the transmembrane protein of the virion. While our results do not rule out a role for immune selection of certain variant
genotypes, they argue strongly that the effect of variation in this region on Rev function must also be considered as a plausible selective force \textit{in vivo}. Further support for the role of Rev variation in EIAV persistence is found in the analysis of Rev variation during successive stages of clinical disease.

Alterations in patterns of clinical disease during an infection result from changes in either the host immune response and/or changes in virus replication. Prior to these studies, we reported that a heterogeneous mixture of Rev genotypes was present in the first febrile cycle of a horse acutely infected with EIAV (1). Functional analysis of these genotypes indicated the presence of both Rev-competent and Rev-defective phenotypes. Variation which restricts expression of virus structural genes may contribute to EIAV evasion of the host immune response during periods of clinical quiescence. Studies with HIV-1 have suggested that variation in Rev phenotype may contribute to the replication potential of the virus, and hence viral pathogenesis (6). Indeed, in our \textit{in vivo} studies of \textit{rev} variation, we observed that highly active Rev-phenotypes predominated during periods of clinical disease and less active Rev phenotypes predominated during the acute and subclinical stages of disease. Restricted expression of viral structural genes is a common strategy of persistent viruses (10). Therefore, less active Rev phenotypes may be important for evasion of an active host immune response during periods of long-term persistence. Findings from this work provide strong evidence that variation in Rev is an important factor in both EIAV persistence and pathogenicity.

\textbf{Identification of the EIAV RREs}

In other lentiviruses, \textit{trans}-activation of the Rev/Rex RNA export pathway occurs through an interaction with a single RRE (reviewed in 3). For the majority of lentiviruses the RRE is located near the surface-transmembrane envelop region; two exceptions are FIV and
HTLV-1, for which the RREs have been mapped near the 3' end of the genome (3, 12, 13). The studies in Chapter 2 confirm previous reports that EIAV Rev is unique among other lentiviruses in that it \textit{trans}-activates through two discrete regions in the viral RNA (2, 8). The studies in Chapter 4 further mapped the 5' element to a 57 nt region within the first \textit{rev} exon. This is the most 5' RRE mapped among lentiviruses. However, similar to ERRE-1, the reporter construct containing the 57 nt ESE region showed significantly reduced activity as compared to the ERRE-All reporter construct containing a larger portion of the \textit{env} gene. The second RRE element has not been more specifically identified, nor is it clear that such an element can function independently of the 5' RRE to mediate export of viral pre-mRNAs. Furthermore, the results demonstrated the combined activities of two reporter plasmids containing the individual RREs was only 69% of the activity of the reporter plasmid containing both elements in the \textit{in vivo} context. This suggests that a synergistic interaction between both \textit{cis}-acting regions is required for optimal export activity. Additional studies will be necessary to fully understand the biological and evolutionary significance of the two RRE-containing EIAV Rev-mediated export pathway.

Interestingly, the presence of both RREs significantly enhanced the differences in biological activity observed among the Rev variants in the studies of Chapter 2. The basis for this synergism is not clear. The sequence of ERRE-1 overlaps the Rev exon 3 sequences present in the variant cDNAs and it is possible that the Rev expression plasmids contain a \textit{cis}-acting element capable of binding to Rev. Therefore, the variant expression plasmids provided \textit{in-trans} may transcribe RNAs which compete with the reporter plasmids for Rev binding, thereby suppressing the full expression of CAT. If the ERRE sub-elements differed in their affinity for Rev binding, it could also explain the finding that the sum of the activities of the individual reporter plasmids was only 69% of the activity of the reporter containing both elements. Computer modeling predicts different RNA secondary structures for the RRE
present in Rev cDNAs versus the RREs in the reporter plasmids and it is not obvious that the variant cDNAs could act as a competitive inhibitor of ERRE-1. Clearly, more detailed biochemical analysis of the EIAV *cis*-acting sequences is needed.

**EIAV Rev-mediated alternative splicing**

In facilitating the export of incompletely spliced mRNAs, HIV-1 Rev and HTLV-1 Rex inhibit the expression of both Tat and Rev specific mRNAs. The EIAV Rev export pathway is unique from HIV-1 and HTLV-1 in that EIAV Rev specifically down-regulates its own production, independent of Tat, by promoting exon 3 skipping of the bicistronic mRNA (4, 8). This mechanism allows for continuous production of Tat while Rev synthesis is limited. The analysis of the multiply spliced RNAs from both Rev-defective and Rev-competent cell lines in Chapter 2 confirms that alternative splicing is a Rev-dependent phenomenon. The prior model of EIAV Rev dependent alternative splicing proposed that binding of Rev to the RRE overlapping exon 3 interferes with SR protein-RNA or SR protein-snRNP interactions to promote exon 3 skipping (4). Consistent with this model, amino acid variation which altered the RNA export function of the protein did not alter the alternative splicing, and a variant deficient for RNA export retained the ability to promote skipping of the third exon. These results suggested the two functions most likely occupy separate domains in the protein, although both functions may require RRE binding. The multifunctional nature of EIAV Rev highlights its importance in terms of regulation of virus gene expression and replication. As such, genetic and biological variation in Rev may have added significance *in vivo*.

The studies in Chapter 4 demonstrate that the purine rich sequence within exon 3 of equine infectious anemia virus (EIAV) is both an exon splicing enhancer (ESE) and a functional Rev responsive element (RRE). The use of an ESE as an RRE is unprecedented
among complex retroviruses. A prior study demonstrated that the SR protein ASF/SF2 cross-links \textit{in vitro} with exon 3 and suggested that the purine rich sequence was an ESE (4). Consistent with this previous data, we demonstrated that the purine rich sequence is required for exon 3 recognition in splicing reactions \textit{in vitro}. Taken together, these data suggest ASF/SF2 is the SR protein required for efficient splicing of exon 3. The ability of a 57 nt sequence containing the ESE to interact \textit{in vitro} with GST-Rev and act as a functional RRE in a heterologous nuclear export assay system demonstrates this region contains a functional RRE. The mutation studies of GAA nucleotide repeats in the ESE confirmed both exon 3 splicing and Rev-mediated RNA export have similar \textit{cis}-acting RNA requirements. Therefore, we propose that Rev-mediated nuclear export requires binding at or near the ESE and results in skipping of exon 3 through an inhibition of ASF/SF2-ESE interactions required for recognition of exon 3 by the host cell splicing machinery.

\textbf{Future studies}

Together, these studies comprise a comprehensive genetic and molecular characterization of EIAV Rev, providing insights into the pathogenesis of the virus and the molecular mechanisms of virus replication. Many questions remain however, both to confirm some of the data presented here and to complete the functional studies of Rev. In Chapter 2 we identified two independent RREs of EIAV. The 5' RRE was further delineated in Chapter 4 to a 57 nt region, however this RRE had severely reduced activity in comparison to the ERRE-AU reporter. This suggests that the 3' RRE is necessary for full function. Since that publication, little characterization of the 3' RRE has been done. Further studies mapping the exact location of this RRE are required to fully delineate the role of the 3' RRE and its importance in the EIAV Rev export pathway.
The exhaustive analysis of pony 524 described in Chapter 3 provides strong evidence that variation in Rev contributes to EIAV persistence and pathogenesis. One weakness of this work is that the analysis was performed in a single animal undergoing a particular form of disease. EIAV can also induce an acute disease which results in death within several weeks, or inapparently infected horses. I would recommend examining \textit{rev} variation in these other types of experimentally infected animals as a mechanism to confirm our findings in pony 524 and further define the significance of Rev variation in EIAV pathogenesis.

We clearly delineated the mechanism of Rev-mediated exon 3 skipping in Chapter 4. Since SR proteins interact at the RRE, they potentially can inhibit Rev function. A previous study in our laboratory, mentioned in the discussion of Chapter 3, demonstrated a down-regulation of Rev activity in activated macrophages. This suggests that such an inhibition may result from an upregulation of SR proteins during activation of macrophages. However, little is known about the phenotype of SR proteins in macrophages. Therefore I recommend further studies to determine the role of SR proteins in EIAV replication. These studies should begin by determining if excess SR proteins can inhibit EIAV replication. If this occurs, then it would be important to confirm the inhibition results from a down-regulation of Rev activity. Finally, the phenotype of SR proteins should be studied in quiescent and activated macrophages to determine what changes in SR proteins are associated with activation state.

The functional mapping of Rev has just begun. The deletion mutants described in Chapter 5 are a good tool to further identify the domains required for Rev activity. Indeed, the studies described in Chapter 4 would suggest that the alternative splicing domain mapped in Chapter 5 is actually the RNA binding domain of Rev. This can be readily determined using the deletion mutants in RNA gel mobility shift assays. However, it would appear that either smaller deletions or site-directed mutagenesis will be necessary to more finely identify
any domain. RNA binding studies should also delineate the multimerization domain.
Following confirmation of the nuclear localization signal, Rev will be completely mapped.

References


