

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

**Title**

Dual Upstream Open Reading Frames Control Translation of a Herpesviral Polycistronic mRNA

**Permalink**

<https://escholarship.org/uc/item/60h7n25r>

**Author**

Kronstad, Lisa Marie

**Publication Date**

2013

Peer reviewed|Thesis/dissertation

Dual Upstream Open Reading Frames Control Translation of  
a Herpesviral Polycistronic mRNA

By

Lisa Marie Kronstad

A dissertation submitted in partial satisfaction of the  
requirements for the degree of

Doctor of Philosophy

in

Microbiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Britt Glaunsinger, Chair  
Professor Kathleen Collins  
Professor Laurent Coscoy  
Professor Jennifer Doudna

Fall 2013



## Abstract

### Dual Upstream Open Reading Frames Control Translation of a Herpesviral Polycistronic mRNA

by

Lisa Marie Kronstad

Doctor of Philosophy in Microbiology

University of California, Berkeley

Professor Britt Glaunsinger, Chair

Kaposi's sarcoma-associated herpesvirus (KSHV) is the etiologic agent of multicentric Castleman's disease, primary effusion lymphoma and Kaposi's sarcoma. KSHV expresses a number of transcripts with the potential to generate multiple proteins, yet relies on the cellular translation machinery that is primed to synthesize only one protein per mRNA. Here we report that the viral transcript encompassing ORF35-37 is able to direct synthesis of both ORF35 and ORF36, while the 3'-proximal ORF37 is translated from a monocistronic transcript under separate transcriptional regulation. While ORF36 encodes the KSHV protein kinase and functions to phosphorylate and activate the anti-viral prodrug ganciclovir, ORF37 expresses SOX (shutoff and exonuclease), a viral protein responsible for promoting widespread degradation of host mRNAs. Surprisingly, ORF36 is robustly translated as a downstream cistron from the ORF35-37 polycistronic transcript in a cap-dependent manner. Two short, upstream open reading frames (uORFs) were identified within the 5' UTR of the polycistronic mRNA. While both uORFs function as negative regulators of ORF35, unexpectedly, translation of the downstream ORF36 gene occurs by a termination-reinitiation mechanism after translation of both uORF1 and uORF2. The sequence of the viral uORFs appears unimportant, and they instead function to position the translation machinery to favor translation of the downstream major ORF. Furthermore, the newly developed KSHV genetic system was used to introduce a mutation into the viral genome to disrupt uORF2 and this single nucleotide change leads to a dramatic drop in infectious viral titer. Positional conservation of uORFs within a number of related viruses suggests that this may be a common  $\gamma$ -herpesviral adaptation of a host translational regulatory mechanism. Thus, KSHV uses a host strategy generally reserved to repress translation to instead allow for the expression of an internal gene.

Dedicated to:

My Mother, Becky Kronstad - for unfailing support.

My Father, Jim Kronstad – for instilling within me a love of science.

My friend, Priscilla - for brief interludes of sunlight and swimming amidst the science.

My husband, Daniel Shore – for quite simply everything.

## TABLE OF CONTENTS

|                                                                                                                            |           |
|----------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>Chapter One: Introduction .....</b>                                                                                     | <b>1</b>  |
| <b>Overview.....</b>                                                                                                       | <b>1</b>  |
| <b>The Scanning Model of Translation Initiation .....</b>                                                                  | <b>1</b>  |
| <b>Alternative Mechanisms of Translation Initiation.....</b>                                                               | <b>1</b>  |
| Leaky Scanning.....                                                                                                        | 1         |
| Ribosome Shunting.....                                                                                                     | 2         |
| Reinitiation .....                                                                                                         | 2         |
| Internal Ribosome Entry Sites .....                                                                                        | 3         |
| <b>Herpesviridae .....</b>                                                                                                 | <b>4</b>  |
| <b>Kaposi's Sarcoma (KS) and KS-associated Herpesvirus (KSHV) .....</b>                                                    | <b>4</b>  |
| <b>Thesis Outline.....</b>                                                                                                 | <b>6</b>  |
| <b>Figures .....</b>                                                                                                       | <b>7</b>  |
| <b>Chapter Two: Characterization of Gene Expression from the KSHV <i>ORF34-37</i></b>                                      |           |
| <b>Locus .....</b>                                                                                                         | <b>11</b> |
| <b>Background .....</b>                                                                                                    | <b>11</b> |
| <b>Results.....</b>                                                                                                        | <b>12</b> |
| Identification of a <i>ORF37</i> -specific transcript .....                                                                | 12        |
| <i>ORF37</i> is <i>trans</i> -activated by RTA by a 142-nucleotide element within the <i>ORF37</i> promoter .....          | 12        |
| The <i>ORF35-37</i> mRNA supports translation of <i>ORF36</i> .....                                                        | 13        |
| <i>In vitro</i> translation of the <i>ORF34-37</i> and <i>ORF35-37</i> mRNAs.....                                          | 13        |
| Identification of a functionally bicistronic KSHV mRNA.....                                                                | 14        |
| Measurement of <i>ORF35</i> and <i>ORF36</i> protein half-lives .....                                                      | 14        |
| <i>ORF36</i> translation is 5' m <sup>7</sup> G cap-dependent.....                                                         | 14        |
| <b>Discussion.....</b>                                                                                                     | <b>16</b> |
| <b>Materials and Methods .....</b>                                                                                         | <b>17</b> |
| Plasmid constructs.....                                                                                                    | 17        |
| Cells, transfections and drug treatment.....                                                                               | 17        |
| Dual Luciferase assays .....                                                                                               | 18        |
| Western blots .....                                                                                                        | 18        |
| Northern blots.....                                                                                                        | 18        |
| 5' Rapid Amplification of cDNA Ends (5' RACE).....                                                                         | 18        |
| Immunofluorescence assay to verify lytic reactivation .....                                                                | 19        |
| <i>In vitro</i> translation.....                                                                                           | 19        |
| <b>Figures .....</b>                                                                                                       | <b>20</b> |
| <b>Chapter Three: Dual Upstream Open Reading Frames Control Translation of the <i>ORF35-37</i> Polycistronic mRNA.....</b> | <b>29</b> |
| <b>Background .....</b>                                                                                                    | <b>29</b> |
| <b>Results.....</b>                                                                                                        | <b>30</b> |
| The context of the <i>ORF35</i> start codon does not dramatically effect expression of <i>ORF36</i> ..                     | 30        |
| Identification of two translation competent short upstream open reading frames in the 5'UTR .....                          | 30        |
| Analysis of the influence of uORF1 on <i>ORF35</i> and <i>ORF36</i> expression.....                                        | 30        |
| A second uORF acts to repress <i>ORF35</i> translation and stimulate translation of <i>ORF36</i> .....                     | 31        |
| The <i>ORF36</i> start codon is accessed by linear scanning.....                                                           | 31        |

|                                                                                                                                         |           |
|-----------------------------------------------------------------------------------------------------------------------------------------|-----------|
| The length of uORF2 and the uORF2-ORF36 intercistronic distance is critical for expression of the 3'-proximal ORF36.....                | 32        |
| No <i>cis</i> -acting element within ORF36 is required for reinitiation and uORF2 does not function when supplied <i>in trans</i> ..... | 32        |
| uORF1 and uORF2 function synergistically to govern ORF35 and ORF36 translation initiation.....                                          | 33        |
| ORF36 expression occurs via reinitiation after uORF translation.....                                                                    | 34        |
| The ORF35 start codon is at least partially accessed by leaky scanning past AUG <sub>uORF1</sub> and AUG <sub>uORF2</sub> .....         | 34        |
| Ratio of ORF35 and ORF36 translational initiation.....                                                                                  | 35        |
| <b>Discussion</b> .....                                                                                                                 | <b>36</b> |
| <b>Materials and Methods</b> .....                                                                                                      | <b>39</b> |
| Plasmid constructs.....                                                                                                                 | 39        |
| Cells, transfections and drug treatment.....                                                                                            | 39        |
| Western blots .....                                                                                                                     | 40        |
| Northern blots.....                                                                                                                     | 40        |
| Graph Design and Statistical Analysis .....                                                                                             | 41        |
| Dual Luciferase assays .....                                                                                                            | 41        |
| <b>Figures</b> .....                                                                                                                    | <b>42</b> |
| <b>Chapter Four: Investigation of the role of uORFs in the context of KSHV</b>                                                          |           |
| <b>Infection</b> .....                                                                                                                  | <b>59</b> |
| <b>Background</b> .....                                                                                                                 | <b>59</b> |
| <b>Results</b> .....                                                                                                                    | <b>60</b> |
| Disruption of uORF2 alters ORF36 expression during lytic infection.....                                                                 | 60        |
| Mutation of uORF2 leads to a defect in the production of infectious viral particles.....                                                | 60        |
| Analysis of the phosphorylation status of eIF2 $\alpha$ during KSHV infection.....                                                      | 60        |
| Investigation of uORF mutants under conditions of eIF2 $\alpha$ phosphorylation.....                                                    | 61        |
| Conservation of uORFs within related $\gamma$ -herpesviruses .....                                                                      | 62        |
| <b>Discussion</b> .....                                                                                                                 | <b>63</b> |
| <b>Materials and Methods</b> .....                                                                                                      | <b>64</b> |
| Cells, transfections and drug treatment.....                                                                                            | 64        |
| Dual Luciferase assays .....                                                                                                            | 64        |
| BAC mutagenesis and DNA isolation.....                                                                                                  | 64        |
| Quantification of Infectious Virus .....                                                                                                | 64        |
| Western blots .....                                                                                                                     | 65        |
| Sequence Alignments.....                                                                                                                | 65        |
| Graph Design and Statistical Analysis .....                                                                                             | 65        |
| <b>Figures</b> .....                                                                                                                    | <b>66</b> |
| <b>Chapter Five: Perspectives</b> .....                                                                                                 | <b>75</b> |
| <b>Further Discussion and Future Directions</b> .....                                                                                   | <b>75</b> |
| <b>Concluding Remarks</b> .....                                                                                                         | <b>76</b> |
| <b>REFERENCES</b> .....                                                                                                                 | <b>77</b> |

## TABLE OF FIGURES

### Chapter One

|                                                                             |    |
|-----------------------------------------------------------------------------|----|
| <b>Figure 1.1.</b> Model of translation initiation.....                     | 7  |
| <b>Figure 1.2.</b> Alternative mechanisms of translation initiation.....    | 8  |
| <b>Figure 1.3.</b> Schematic of the KSHV <i>ORF34-37</i> genetic locus..... | 10 |

### Chapter Two

|                                                                                                                                   |    |
|-----------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Figure 2.1.</b> Identification and transcript start site mapping of ORF37.....                                                 | 20 |
| <b>Figure 2.2.</b> The ORF37 promoter is <i>trans</i> -activated by RTA.....                                                      | 21 |
| <b>Figure 2.3.</b> Mapping of the RTA-responsive element in the ORF37 promoter.....                                               | 22 |
| <b>Figure 2.4.</b> Efficient translation of ORF36 from the full-length ORF35-37<br>tricistronic transcript.....                   | 23 |
| <b>Figure 2.5.</b> <i>In vitro</i> translation of ORF34-37 and ORF35-37 transcripts.....                                          | 24 |
| <b>Figure 2.6.</b> The ORF35-37 mRNA is functionally bicistronic.....                                                             | 25 |
| <b>Figure 2.7.</b> Half-life analysis of ORF35 and ORF36.....                                                                     | 25 |
| <b>Figure 2.8.</b> Translation of ORF36 is independent of IRES activity and<br>dependent on the 5' mRNA m <sup>7</sup> G cap..... | 26 |

### Chapter Three

|                                                                                                                                                                              |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Figure 3.1.</b> The ORF35 start codon does not dramatically influence expression<br>of ORF36.....                                                                         | 42 |
| <b>Figure 3.2.</b> Two uORFs located in the 5'UTR engage the translation initiation<br>machinery.....                                                                        | 43 |
| <b>Figure 3.3.</b> Analysis of the contribution of uORF1 on ORF35 and ORF36<br>expression.....                                                                               | 44 |
| <b>Figure 3.4.</b> uORF2 mediates translational control of ORF35 and ORF36.....                                                                                              | 45 |
| <b>Figure 3.5.</b> The ORF36 start codon is accessed by linear scanning.....                                                                                                 | 46 |
| <b>Figure 3.6.</b> uORF2 length and the uORF2-ORF36 intercistronic distance are<br>important for expression of ORF36.....                                                    | 47 |
| <b>Figure 3.7.</b> No <i>cis</i> -acting element within ORF36 is required for reinitiation and<br>uORF2 does not function when supplied <i>in trans</i> .....                | 48 |
| <b>Figure 3.8.</b> uORF1 and uORF2 synergistically govern expression of ORF35 and<br>ORF36.....                                                                              | 50 |
| <b>Figure 3.9.</b> Toeprinting analysis to detect ribosome occupancy on AUG <sub>uORF1</sub> ,<br>AUG <sub>uORF2</sub> , AUG <sub>ORF35</sub> and AUG <sub>ORF36</sub> ..... | 51 |
| <b>Figure 3.10.</b> Ribosomes terminating translation of uORF1 and uORF2<br>reinitiate at the ORF36 start codon.....                                                         | 52 |
| <b>Figure 3.11.</b> The ORF35 start codon is partially accessed by leaky scanning<br>past the uORF1 and uORF2 start sites.....                                               | 53 |
| <b>Figure 3.12.</b> Ratio of ORF35 and ORF36 translation initiation.....                                                                                                     | 53 |
| <b>Figure 3.13.</b> Model of the mechanisms of translation initiation used to translate<br>ORF35, ORF36 and ORF37.....                                                       | 54 |
| <b>Figure 3.14.</b> The role of uORFs in translational control.....                                                                                                          | 55 |

## Chapter Four

|                                                                                                                                          |    |
|------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Figure 4.1.</b> Disruption of uORF2 alters ORF36 expression during lytic infection.....                                               | 66 |
| <b>Figure 4.2.</b> Construction of an FLAG epitope-tagged ORF35 BAC16 allows for direct detection of ORF35.....                          | 67 |
| <b>Figure 4.3.</b> Disruption of uORF2 results in a drop in infectious titer.....                                                        | 68 |
| <b>Figure 4.4.</b> Robust induction of eIF2 $\alpha$ phosphorylation during KSHV lytic replication.....                                  | 69 |
| <b>Figure 4.5.</b> Analysis of uORF1 and uORF2 mutants on ORF35 expression under conditions of eIF2 $\alpha$ phosphorylation.....        | 70 |
| <b>Figure 4.6.</b> Analysis of uORF1, uORF2 and uORF3 mutants on ORF36 expression under conditions of eIF2 $\alpha$ phosphorylation..... | 71 |
| <b>Figure 4.7.</b> uORF1 and uORF2 are conserved among select $\gamma$ -herpesviruses...                                                 | 72 |

## TABLES

|                                                                                                                                                                                                             |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Table 2.1.</b> List of oligonucleotide primers used in chapter two.....                                                                                                                                  | 28 |
| <b>Table 3.1.</b> List of oligonucleotide primers used in chapter three.....                                                                                                                                | 58 |
| <b>Table 4.1.</b> Analysis of the <100 nucleotides upstream of ORF35 in the genomes of $\gamma$ -herpesviruses with the conserved ORF34-37 genome arrangement deposited in the Virus Pathogen Resource..... | 73 |
| <b>Table 4.2.</b> List of oligonucleotide primers used in chapter four.....                                                                                                                                 | 74 |

## ACKNOWLEDGEMENTS

Funding was provided by a Burroughs Wellcome Fund Investigators in the Pathogenesis of Infectious Disease Award to Britt Glaunsinger and a National Sciences and Engineering Research Council of Canada (NSERC) fellowship to Lisa Kronstad.

I would like to thank Drs. Peter Sarnow, Yoshihiro Izumiya, Jae U. Jung, Zhi-Ming Zheng, Shelley Starck and Gregory Pari for their generous sharing of reagents and Drs. Rolf Renne, Shelley Starck, Brian McCarthy, Reed Shabman and Carolina Arias for technical assistance. I am grateful to all members of the Glaunsinger laboratory for helpful comments and suggestions throughout the years, especially Dr. Karen Clyde and Zoe Davis. I would particularly like to thank Kevin Brulois for design of the BAC mutagenesis system and the construction of the BAC16 uORF2 mutant. Thank you to my thesis committee Drs. Kathleen Collins, Laurent Coscoy and Jennifer Doudna and my qualifying exam committee Drs. Eva Harris, Laurent Coscoy, Greg Barton and Andrew Jackson for their wonderful support and guidance. Of course, none of this would have been possible without Dr. Britt Glaunsinger who gave me a chance; our many, many scientific discussions in her office over the years have been some of the most enjoyable moments of my graduate work.

# Chapter One: Introduction

## Overview

Gene expression is governed at the level of transcription, post-transcriptional processing, mRNA stability, translation, post-translational modifications and protein degradation. Viruses encode strategies to manipulate nearly every stage of gene expression, however by and large do not encode their own translational machinery and thus operate under the constraints of host protein synthesis [1,2]. Here we summarize the current understanding of the 5' to 3' scanning model of translation initiation that selects the 5'-proximal AUG and highlight a number of strategies that exist to bypass upstream start codons to enable downstream initiation. A lion's share of these unusual mechanisms of translation initiation have been discovered in viruses, likely because viral genomes have evolved to optimize restricted genomic space to allow for a high density of coding information. In this thesis we have focused on understanding how Kaposi's sarcoma-associated herpesvirus (KSHV) encodes transcripts with the potential to generate multiple proteins, yet relies on the cellular translation apparatus that is primed to synthesize only one protein per mRNA [3,4].

## The Scanning Model of Translation Initiation

Protein synthesis is generally regulated at the initiation stage, rather than during elongation or termination [5]. Eukaryotic initiation factors (eIFs) 1, 1A, 3 and 5 serve to prime the small ribosomal subunit (40S) for initiation, followed by binding to Met-tRNA<sub>i</sub> and eIF2-GTP to form the ternary complex (TC) [6] (**Figure 1.1; step 1**). Eukaryotic mRNAs possess a 5'-cap (m<sup>7</sup>GpppN) that is bound by the cap-binding eIF4F, a protein complex comprised of eIF4E, an initiation factor with specific affinity for the methylated 5' cap, eIF4A, a helicase which unwinds the RNA secondary structure of the 5' untranslated region (UTR) and eIF4G, which interacts with eIF4E, eIF4A and eIF3. (**Figure 1.1; step 2**). Ribosomal attachment is achieved through the 5' m<sup>7</sup>G cap-eIF4E-eIF4G-eIF3-40S chain of interactions, forming a 48S initiation complex that is competent to begin inspecting the 5'UTR in a 5' to 3' direction. The scanning ribosomes then initiate translation at (usually) the first AUG start codon, which generally serves as the sole site of productive initiation of protein synthesis (**Figure 1.1; step 3**) [5,7]. Met-tRNA<sub>i</sub> anti-codon base pairing with the AUG start codon triggers GTP hydrolysis and the release of eIF2-GDP (**Figure 1.1; step 4**), followed by 60S ribosomal subunit joining and the release of other initiation factors, concluding the initiation process and beginning the elongation stage (**Figure 1.1; step 5-6**) [6,8].

## Alternative Mechanisms of Translation Initiation

### Leaky Scanning

As a consequence of the 5'-polarity of the scanning pre-initiation complex, ribosomes generally do not engage start codons positioned internally within the mRNA and thus eukaryotic transcripts generally encode only one functional protein. For the majority of mRNAs the most 5'-proximal AUG is selected, however, strategies exist to

bypass upstream start codons to enable downstream initiation. For example, leaky scanning can occur if the nucleotides flanking the 5'-proximal AUG are not in the Kozak consensus sequence (ccRccAUGG), allowing the 40S ribosomal subunit to instead engage a downstream methionine codon [8-10]. The purines at the -3 and +4 positions likely contribute to start codon selection by stabilizing conformation changes that occur on codon-anticodon base pairing, by interacting with eIF2 $\alpha$  and with nucleotides AA<sub>1818-1819</sub> in helix 44 of 18S rRNA, respectively [11,12]. Notably, the G in the -4 position contributes to recognition particularly in the absence of an A in the -3 position [12]. The fidelity of initiation is further maintained by eIF1, which promotes recognition of the correct initiation codon by enabling 43S complexes to differentiate between non-AUG triplets and AUG flanked by a poor context [11,13,14]. Thus, if the first-AUG is flanked by a suboptimal context, it is poorly recognized by scanning ribosomes and can be bypassed in favor of a downstream AUG. Thus, leaky scanning allows for the production of two separately initiated proteins from a single mRNA (**Figure 1.2A**) [10].

### **Ribosome Shunting**

Ribosome shunting, first described in Cauliflower mosaic virus (CaMV), is a nonlinear scanning mechanism where ribosomes migrating from the 5' mRNA m<sup>7</sup>G cap are transferred or "shunted" from a 5'-donor to a 3'-acceptor site without inspection of the intervening sequence [15-18]. The CaMV 35S RNA leader sequence supports translation of the downstream cistron despite harboring extensive secondary structure and eight short ORFs. Several classical experiments found that two separate RNA molecules can reconstitute a RNA secondary structure competent to rescue shunting of the CaMV leader. Furthermore insertion of a complete reporter gene within the 5' UTR did not disrupt translation of the downstream cistron [16]. While the biochemical basis of shunting remains unclear, it is thought that the formation of the leader hairpin structure promotes shunting by bringing the donor and acceptor sites into close proximity [16,17]. Thus, despite the presence of intervening inhibitory sequences or RNA secondary structures, shunting grants access to ribosomes initially recruited to the 5' m<sup>7</sup>G cap to a downstream start codon (**Figure 1.2B**).

### **Reinitiation**

When an upstream AUG is followed shortly thereafter by an in-frame termination codon, ribosomes can reinitiate translation, albeit with reduced efficiency, at a downstream AUG [9,19]. When the 80S ribosome terminates translation of a short, upstream ORF (uORF), the 60S ribosomal subunit may be released (this has not been established biochemically) while the 40S subunit remains associated with the mRNA and resumes scanning and may initiate an additional round of translation at a downstream start codon (**Figure 1.2C**). Certain conditions must be met for reinitiation to occur. The first is a sufficiently short primary elongation event, possibly because certain initiation factors dissociate from the ribosomes progressively during elongation. Thus if the elongation phase is sufficiently brief, factors required for reinitiation may remain transiently associated with the 40S subunit and allow for a secondary initiation event. A second condition is the *de novo* acquisition of the eIF2-GTP-Met-tRNA<sub>i</sub> ternary complex, which is dependent on the length of the intercistronic region between the uORF and primary (p)ORF, with longer distances favoring reinitiation [9,10,19].

Notably, upstream start codons (uAUGs) are present in approximately half of human transcripts and are the most conserved trinucleotide across vertebrate species, consistent with their widespread functional role of reducing translation of the major ORF by 30-80% [9,20-23]. Four uORF characteristics correlate with greater inhibition of the downstream pORF: strong Kozak consensus sequence, evolution conservation, increased distance from the 5' mRNA m<sup>7</sup>G cap and multiple uORFs within the 5'UTR [21]. In the majority of characterized uORFs examples it is the act of translation rather than the peptide sequence that mediates their function, however in a few cases the ability of the uORF to repress downstream translation is dependent on the amino acid sequence of the encoded peptide [24-28]. For example, a uORF present in the 5' UTR of the human cytomegalovirus gp48 gene attenuates downstream translation in a sequence-dependent fashion, likely by delaying normal termination and preventing leaky scanning by the 40S ribosomal subunit to reach the downstream AUG [26]. However, in general, engagement of the translation apparatus rather than the translated product itself represses translation of the major ORF.

Recent evidence has emerged implicating eIF3 in the ability of ribosomes to reinitiate translation, partly due to the finding that eIF3 is not displaced by 60S subunit joining and is thus retained temporarily during translation [29-31]. Particularly intriguing was the discovery that reinitiation used to translate a Caliciviral bicistronic mRNA is stimulated by the presence of a ~70 nucleotide sequence element termed the 'termination upstream ribosomal binding site' (TURBS) [32,33]. Poyry et al. (2007) found that a motif within the TURBS sequence interacts with eIF3 and mutants defective in reinitiation showed reduced affinity for eIF3 binding [29,34]. Furthermore, mutational analysis and reciprocal base exchanges demonstrated that hybridization of mRNA and 18S rRNA was critical for reinitiation [32,35]. Thus, at least for certain mRNAs, strategies exist to enhance the efficiency of reinitiation events by tethering the post-terminating 40S ribosome to the mRNA through eIF3 or 18S/mRNA interactions [32-36].

The *GCN4* mRNA in *Saccharomyces cerevisiae* is a particularly well-studied example of translational control by reinitiation, with four uORFs regulating expression of the main open reading frame. Notably, this system can be reconstituted with only uORF1 and uORF4. Reinitiation after uORF1 occurs at a high rate however translation of uORF4 precludes further reinitiation events. Under non-nutrient limiting conditions, the ternary complex is abundant and ribosomes that resume scanning after uORF1 translation have sufficient time to charge with eIF2 and reinitiate at uORF4, preventing initiation at the *GCN4* start codon. However under starvation conditions, the scanning ribosomes bypass uORF4 in favor of the further downstream *GCN4* start site [37-39]. Starvation conditions in yeast gives rise to a pool of uncharged tRNAs that activate the Gcn2 protein kinase which then phosphorylates eIF2 $\alpha$ . Phosphorylated eIF2 $\alpha$  sequesters the guanine nucleotide exchange factor eIF2B, which can no longer catalyze the GDP-GTP exchange reaction, thus limiting the pool of ternary complexes [5,40]. Upregulation of the Gcn4 transcription factor by the limited ternary complex then stimulates transcription of a number of genes for amino acid biosynthesis and aminoacyl-tRNA synthetases [39,41]. Thus, uORF-mediated control of translation for the *GCN4* locus provides a sophisticated feedback loop to regulate amino acid levels within the cell.

### **Internal Ribosome Entry Sites**

Although the vast majority of mRNAs use a variation of the 5' m<sup>7</sup>G cap-dependent scanning mechanism, initiation of certain viral mRNAs is mediated by internal ribosome entry sites (IRES). IRESs are RNA elements that allow for end-independent ribosomal recruitment to an internal position within an mRNA (**Figure 1.2D**). First discovered in poliovirus, a number of viral IRESs have been discovered which use distinct mechanisms and interact with a range of eIFs and/or 40S subunits; this topic has been comprehensively reviewed by Jackson et al. [5,42-45]. Internal initiation is often used by viruses that have developed a strategy to shut down 5' m<sup>7</sup>G cap-dependent initiation in a host cell, thus allowing for an elegant escape strategy to continue translation of their own mRNAs that do not require a cap structure for protein synthesis.

Intriguingly, a number of cellular mRNAs have also been reported to harbor IRESs, although their existence remains shrouded in controversy [46-48]. The battleground of cellular IRES discovery is littered with false positives, victims of cryptic promoter and splice sites, while the stringency of controls required to definitively prove IRES activity remains hotly debated [49]. Furthermore, in contrast to many viral IRESs, the putative cellular IRES-containing mRNAs are capable of being translated by canonical scanning from the 5' mRNA m<sup>7</sup>G cap, highlighting the question of how the switch between 5' m<sup>7</sup>G cap-dependent and independent translation is regulated, a topic currently under intense investigation [49,50].

### **Herpesviridae**

Non-canonical mechanisms of translation control have generally been the domain of RNA viruses due to their compact genome, however this thesis work focuses on the translation initiation on transcripts expressed by KSHV, a large double stranded DNA virus of the Herpesviridae family. Herpesviruses consist of a nucleocapsid core surrounded by a mRNA and protein tegument layer and an enveloped membrane [51]. A hallmark of herpesviruses is their biphasic lifecycle of latency and lytic reactivation. During latency the viral genome is localized to the nucleus and is present as a circular molecule referred to as an episome; this genome expresses a restricted profile of genes [51]. Reactivation from latency leads to expression of the full complement of viral genes with temporal regulation and the production and egress of infectious viral particles [51].

The Herpesviridae is divided into the  $\alpha$ -,  $\beta$ -, and  $\gamma$ -herpesvirinae subfamilies formed on the basis of their cellular tropism, infectious characteristics and genomic conservation [52]. The eight human herpesviruses include examples in each subfamily. For example, the  $\alpha$ -herpesvirinae includes herpes simplex viruses 1 and 2 (HSV-1 and HSV-2) that cause oral and genital herpes, respectively, and varicella-zoster virus, the etiologic agent of chicken pox. The  $\beta$ -herpesvirinae includes cytomegalovirus (CMV), a major cause of birth defects, and human herpesviruses 6 and 7, the common causes of rashes in infants [53]. Finally, the  $\gamma$ -herpesvirinae encompasses Epstein-Barr virus (EBV/HHV-4) and Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8) that can cause certain lymphoproliferative disorders, in particular in immunocompromised individuals [54,55].

### **Kaposi's Sarcoma (KS) and KS-associated Herpesvirus (KSHV)**

Kaposi's sarcoma was first described in 1872 by Moritz Kaposi as a skin cancer affecting elderly men in Vienna, Austria [56]. In the early 1980's, the number of cases of

KS rose dramatically and served as the harbinger of the AIDS pandemic, caused by the shortly to be discovered HIV1 [57]. There are currently four recognized epidemiological forms of KS: 1) classic KS, the form recognized by Kaposi affecting elderly men of Mediterranean or Eastern European Jewish ancestry; 2) endemic KS, present in part of Eastern and Central Africa and associated with disseminated lymphadenopathy, often in children; 3) iatrogenic KS, which occurs in patients after organ transplant under immunosuppressive therapy; and 4) epidemic-KS, a major AIDS-defining neoplasm and presently one of the most common cancers in sub-Saharan countries [58-60].

Epidemiological studies strongly suggested that KS had an infectious etiology. Chang et al. used representational difference analysis to discover novel DNA sequences in AIDS-KS lesions that were homologous to  $\gamma$ -herpesviral capsid and tegument genes [61]. These DNA fragments proved to be the first glimpse of the KS-associated herpesvirus (KSHV/HHV-8) genome [61]. KSHV is now known to cause not only KS but the virus is also associated with Multicentric Castleman's disease (MCD) and Primary Effusion Lymphoma (PEL), two lymphotropic disorders [62,63].

KSHV has a ~165-kb dsDNA genome encoding an estimated 85 viral proteins [52,54,61,64]. The viral genes closely resemble those of their cellular counterparts in that they have canonical transcriptional promoters, consensus pre-mRNA splice sites and 3'-end formation signals. However, one notable departure from host genetic organization is the scarcity of poly(A) sites distributed throughout the viral genome, with a single signal often allocated to several consecutive ORFs [3,4]. These gene clusters give rise to viral transcripts with polycistronic coding potential, although in general only the 5'-proximal gene is translated on each mRNA [4,65,66]. Most genes are positioned as a 5' cistron by the use of multiple transcriptional start sites upstream of common poly(A) signals and/or alternative splicing [66,67]. Prior to the work presented in this thesis, the only described KSHV mechanism to enable translation of a 3'-proximal ORF was an IRES identified in the coding region of vCyclin (ORF72), which allows for expression of vFLIP (ORF71) [68-70].

## Thesis Outline

KSHV expresses a number of transcripts with the potential to generate multiple proteins, yet relies on the cellular translation machinery that is primed to synthesize only one protein per mRNA. The objective of the work presented in this thesis was to explore the transcriptional and translational regulation of a polycistronic mRNA that encompasses ORF35-37, the major RNA species expressed from the *ORF34-37* genetic locus [3] (**Figure 1.3**). In this work, we report the discovery that ORF37, which encodes the viral host shutoff and exonuclease protein SOX, is translated from a previously uncharacterized monocistronic mRNA under separate transcriptional regulation. Furthermore, it was found that the ORF35-37 transcript is able to direct the synthesis of two proteins and that the translational switch is regulated by two short upstream open reading frames (uORFs) in the native 5' untranslated region. uORFs are elements commonly found upstream of mammalian genes that function to interfere with unrestrained ribosomal scanning and thus repress translation of the major ORF. The sequence of the viral uORF appears unimportant, and instead functions to position the translation machinery in a location that favors translation of the downstream major ORF, via a reinitiation mechanism. Disruption of the second uORF in the KSHV bacterial artificial chromosome (BAC16) results in a dramatic drop in the production of infectious viral particles. This thesis work describes a novel functionally bicistronic viral mRNA that is translated via a unique adaption of ribosomal reinitiation and reveals how KSHV has pirated a host strategy generally reserved to dampen translation to instead allow for the expression of an internal gene.

## Figures

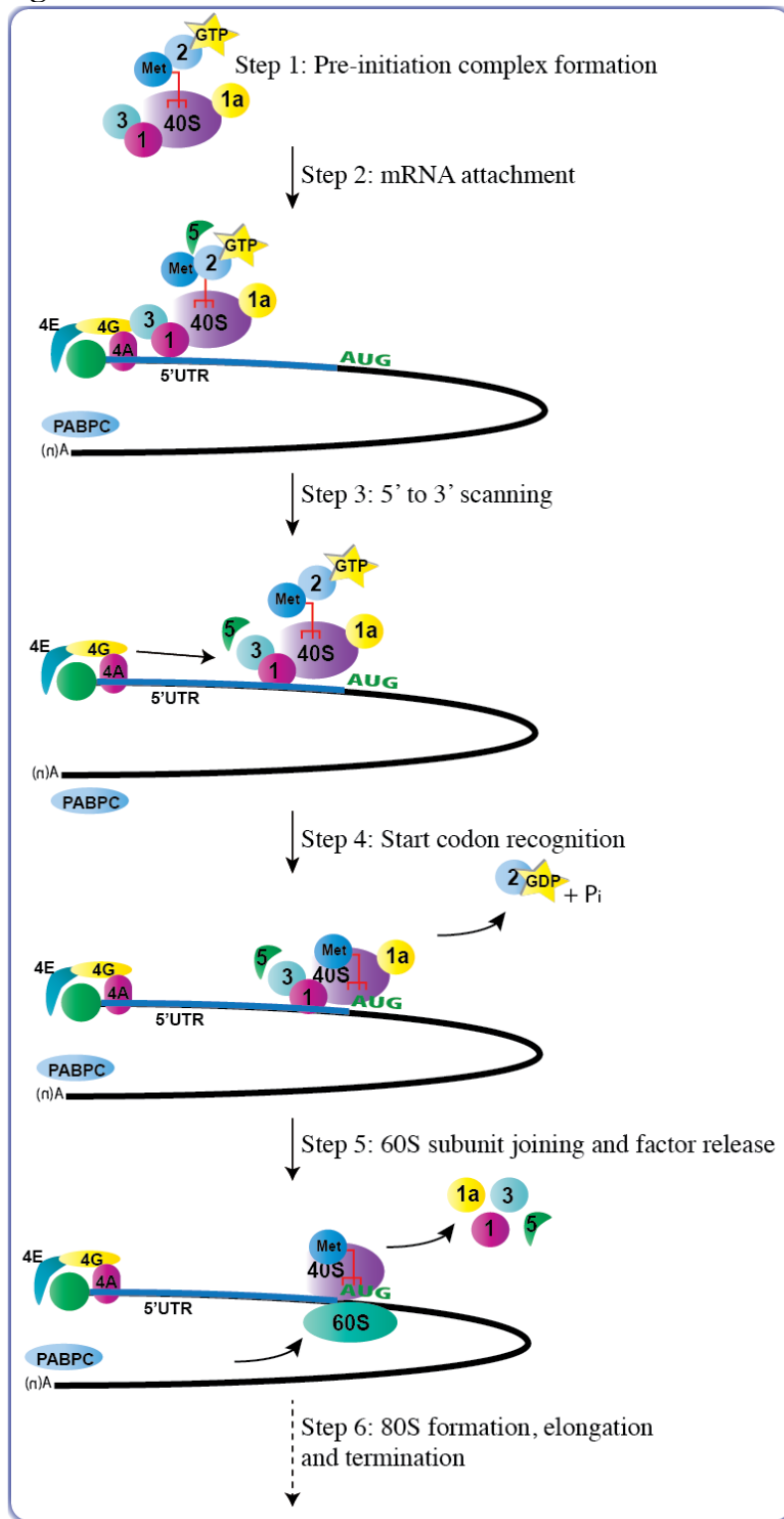


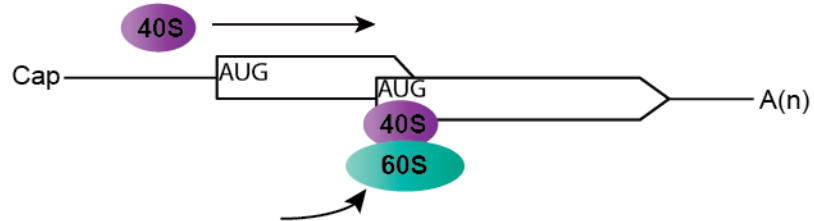
**Figure 1.1. Model of translation initiation.** **Step 1:** Initiation tRNA (Met-tRNA<sup>Met</sup><sub>i</sub>) with eIF2 and GTP forms the ternary complex associated with the 40S ribosomal subunit and with eIF1, eIF1A and eIF3 to comprise the 43S pre-initiation complex. **Step 2:** The

43S pre-initiation complex is recruited to the 5' m<sup>7</sup>G cap structure via eIF3 binding to eIF4G, a scaffolding protein that bridges the cap binding protein eIF4E and the RNA helicase eIF4A (combined: eIF4F). **Step 3:** The pre-initiation complex scans in a 5' to 3' direction aided by eIF4F that unwinds the cap-proximal mRNA in an ATP-dependent manner. **Step 4:** Recognition of initiation codon and 48S initiation complex formation switches the scanning complex from an open to a closed conformation leading to the displacement of eIF1 to allow eIF5-mediated hydrolysis of eIF2-bound GTP and the release of P<sub>i</sub>. **Step 5:** The 60S subunit joins the 48S complex and concomitantly displaces eIF2-GDP and other factor (eIF1, eIF3, eIF4B, eIF4F and eIF5) mediated by eIF5B forming the elongation-competent 80S complex [5]. **Step 6:** Peptide chain elongation and termination.

## Cap-Dependent

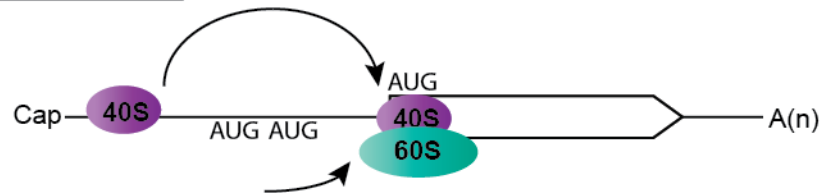
A

### Leaky Scanning



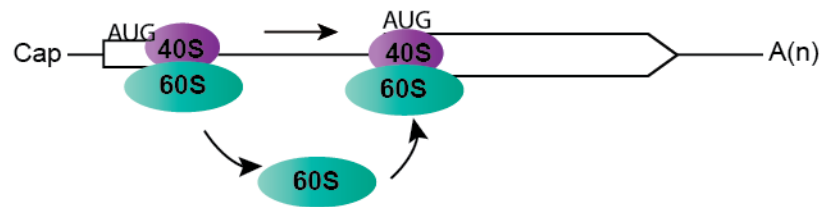
B

### Ribosome Shunting



C

### Reinitiation



## Cap-Independent

D

### Internal Ribosome Entry Site

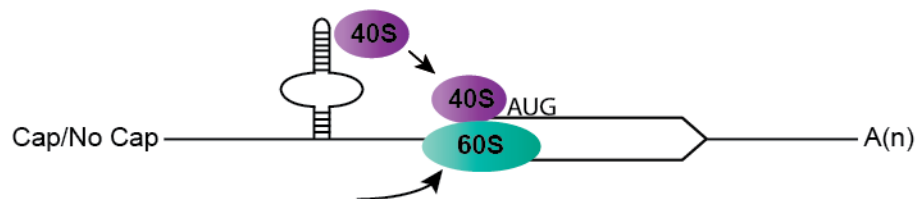


**Figure 1.2. Alternative mechanisms of translation initiation.** Diagrams include key structural features of mRNA including the 5' m<sup>7</sup>G cap (cap), the ORF denoted by an open arrow and the poly(A) tail (A<sub>n</sub>). (A) Leaky Scanning occurs when the scanning 40S ribosomes bypasses potential start codon in favor of a downstream initiation site, which can occur in three known circumstances: 1) the AUG is positioned within ≤10 nucleotides of the cap or; 2) if the start codon is flanked by a poor initiation context or; 3) if a non-AUG start codon is surrounded by an optimal context. (B) Ribosome shunting was first described on the cauliflower mosaic virus 35S RNA leader and is a poorly understood non-linear scanning mechanism where ribosomes hop over or “shunt” directly from a 5' donor to a 3' acceptor site without inspecting the codons of the intervening region [16,17,71-74]. (C) Reinitiation occurs when terminating ribosomes remain associated with the mRNA, continue to scan and initiate at a downstream start codon. Reinitiation requires the *de novo* recruitment of the ternary complex and generally occurs only after a short (s)ORF (<30 codons) and results in dampened expression of the primary (p)ORF. Reinitiation implies that certain initiation factors likely remain transiently associated with the 40S ribosome during elongation, allowing for a secondary initiation event on a single mRNA [9,19]. The ability of terminating ribosomes to resume scanning can be influenced by the RNA secondary structure surrounding the stop codon [75,76] or the function of the sORF-encoded peptide [26,27]. (D) Internal Ribosome Entry Sites (IRES) are RNA elements with extensive secondary structure that mediate 5' m<sup>7</sup>G cap-independent recruitment of ribosomes. IRESs can occur on capped or uncapped mRNAs. Depending on the type of IRES, the 40S ribosome is competent to scan to reach the AUG or is placed directly at the desired initiation site.

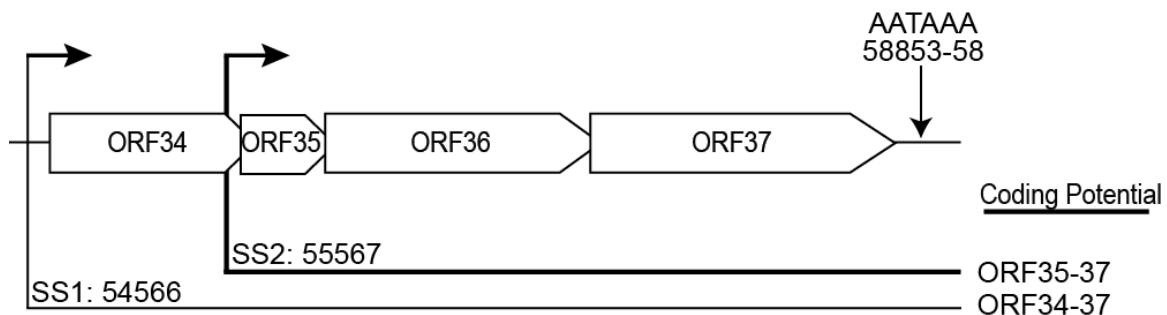


**Figure 1.3. Schematic of the KSHV *ORF34-37* genetic locus.** A schematic presentation of the *ORF34-37* genetic locus showing the previously identified *ORF34-37* and *ORF35-37* polycistronic mRNAs with thin and thick lines respectively [3]. Coding potentials are indicated on the right. Start sites (SS) are indicated for each transcript according to the nucleotide position described by Russo et al. [64]. The single poly(A) signal used by all four ORFs for transcription termination is shown.

## Chapter Two: Characterization of Gene Expression from the KSHV *ORF34-37* Locus

### Background

A previous report by Haque et al. identified two potential functionally polycistronic mRNAs that are transcribed from the KSHV *ORF34-37* genetic locus during lytic replication: a minor transcript encompassing ORFs 34, 35, 36, and 37 (ORF34-37) and a major transcript encompassing ORFs 35, 36 and 37 (ORF35-37) [3]. ORF34 has been shown to cause the degradation of hypoxia-inducible factor 1 $\alpha$  by a proteasome-dependent pathway [3,77]. The function of the protein product of the 5'-proximal ORF35 is ill defined, although it shares limited sequence similarity with the  $\alpha$ -herpesvirus UL14 gene product, which has described heat shock protein-like properties and functions to inhibit apoptosis during host cell infection [78,79]. The second gene, ORF36, encodes a serine/threonine kinase that activates the cellular c-Jun N-terminal kinase (JNK) signaling pathway and phosphorylates the viral transcriptional transactivator K-bZIP, two processes involved in the progression from early to late viral gene expression [80-82]. Moreover, ORF36 sensitizes KSHV-infected cells to ganciclovir, an anti-viral drug shown to reduce KSHV replication in cultured cells and in patients [83-86]. The 3'-proximal ORF37 expresses SOX (shutoff and exonuclease), a viral protein responsible for promoting widespread degradation of host mRNAs and also thought to assist in viral DNA replication and packaging [87-89]. However, the mechanism of translation initiation of the 5'-distal ORF36 and ORF37 proteins has remained unresolved [3,82].

In this chapter, we demonstrate that the ORF34-37 mRNA is functionally monocistronic, supporting translation of only ORF34. Moreover, we have identified and mapped a previously uncharacterized ORF37-specific transcript and show that this mRNA is *trans*-activated by the major viral latent to lytic switch protein RTA. Surprisingly, we find that no transcript exists where ORF36 is the 5'-proximal cistron and instead demonstrate that the ORF35-37 mRNA is functionally bicistronic, supporting translation of both ORF35 and ORF36. Finally, we demonstrate that the polycistronic locus lacks IRES activity, and that both proteins are expressed in a 5' m<sup>7</sup>G cap-dependent manner.

## Results

### Identification of a ORF37-specific transcript

The previously described tricistronic KSHV mRNA encompasses three partially overlapping open reading frames that are expressed with lytic kinetics (ORF35, 36, and 37) [3,90]. We thus sought to determine whether additional transcripts are produced from this locus in a B cell line (TREx BCBL1-RTA) that harbors KSHV in a latent state but can be stimulated to engage in lytic replication. We confirmed that drug treatment successfully reactivated the cell by immunofluorescence analysis using SOX as a specific marker of lytic replication, which is not expressed during latency (**Figure 2.1A**). RNA isolated from cells infected latently or lytically for 8-36 h was Northern blotted with riboprobes specific for ORF37. The ORF37 probe reacted with transcripts  $\geq 3.4$  kb and an additional  $\sim 1.7$  kb transcript that co-migrated with the control ORF37 monocistronic mRNA (**Figure 2.1B**). Analysis of transcription start sites by 5' RACE (**Figure 2.1C**), as well as comparisons with similar observations in a related  $\gamma$ -herpesvirus, further supported the presence of an ORF37 monocistronic transcript [91].

We next sought to evaluate directly whether the ORF35-37 transcript could support translation of ORF37 as a downstream gene. 293T cells were first transfected with a plasmid expressing the coding sequence of ORF35-37 downstream of the native viral 72-nucleotide 5' UTR, and lysates were Western blotted using polyclonal antisera specific for ORF37. The ORF37 protein was detected only in cells transfected with the monocistronic ORF37 plasmid (**Figure 2.1D**). Northern blotting of the mRNAs produced from each transfection confirmed that the transcripts were of the expected size and of equivalent abundance within the experiment (**Figure 2.1D**). Thus, ORF37 is most likely present as a silent cistron on the ORF35-37 polycistronic mRNA and is instead translated by the canonical m<sup>7</sup>G cap-dependent scanning mechanism from an independent monocistronic transcript (**Figure 2.1E: red line**).

### ORF37 is *trans*-activated by RTA by a 142-nucleotide element within the ORF37 promoter

The switch between latency and initiation of the lytic cycle is predominantly controlled by the expression of the replication and transcriptional activator (RTA) [92,93]. We thus sought to determine whether the promoter of the newly identified ORF37-transcript is *trans*-activated by RTA. The 1296 nts upstream of the ORF37 start codon (ORF37 promoter; ORF37p) was inserted upstream of firefly luciferase (fluc) in a plasmid backbone which harbors *Renilla* luciferase (*Rluc*) under separate transcriptional control to serve as an internal control for transfection. The ORF61 promoter (ORF61p) has been previously reported to be *trans*-activated by RTA and thus we inserted the ORF61p into the dual luciferase construct to serve as a positive control [94]. The ORF35 coding region served as a negative control. Either empty vector, ORF61p, ORF35 or ORF37p constructs were introduced with RTA tagged with V5 into either 293T cells by transfection or the BJAB B cell line by electroporation [95]. Both the ORF61 and ORF37 promoters conferred RTA-responsiveness in both 293T (**Figure 2.2A**) and BJAB cell lines (**Figure 2.2B**).

We next used deletion analysis of the ORF37 promoter to more closely delineate the RTA-responsive element by generating constructs with either the full length promoter (-1296) or with the regions of -1046, -896, -596, -296, or -154 nts from the ORF37

promoter, followed by co-electroporation with RTA into BJABs (**Figure 2.3A**). Full responsiveness was detected until the -154 construct where a dramatic drop in RTA-responsiveness was observed, indicating that the 142 nts between -296 and -154 are necessary for RTA *trans*-activation of the ORF37 promoter (**Figure 2.3A: red box and 2.3B**).

### **The ORF35-37 mRNA supports translation of ORF36**

We next searched for transcripts that harbor the ORF36 coding region, again by using the reactivation of the latent TREx BCBL1-RTA KSHV B cell line that harbors the ORF36 coding region. We confirmed that drug treatment successfully reactivated the cell by immunofluorescence analysis using SOX as a specific marker of lytic replication (**Figure 2.4A**). RNA isolated from cells infected latently or lytically for 8-36 h was Northern blotted with riboprobes specific for ORF36. In infected cells, the ORF36 probe recognized transcripts co-migrating with or larger than the polycistronic ORF35-37 mRNA but did not reveal any smaller, potentially monocistronic species (**Figure 2.4B**). Results from ORF36 5' rapid amplification of cDNA ends (RACE) experiments were in agreement with its transcript initiating upstream of ORF35 at nucleotide position 55567 as previously reported by Haque et al. (**Figure 2.4C**) [64].

We next sought to evaluate directly whether the ORF35-37 transcript could support translation of ORF36 as a downstream gene. 293T cells were first transfected with a plasmid expressing the coding sequence of ORF35-37 downstream of the native viral 72-nucleotide 5' UTR, and lysates were Western blotted using polyclonal antisera specific for ORF36. The ORF36 protein was readily translated from this polycistronic construct, whereas the ORF37 protein was detected only in cells transfected with the monocistronic ORF37 plasmid (**Figure 2.4D**). Northern blotting of the mRNAs produced from each transfection confirmed that the transcripts were of the expected size and of equivalent abundance across experiments (**Figure 2.4D**). Thus, in contrast to the general monocistronic character of cellular mRNAs, no transcript was detectable with ORF36 present at the 5'-proximal cistron and the ORF35-37 mRNA is capable of supporting translation of the internal ORF36 cistron (**Figure 2.4E; green line**).

### ***In vitro* translation of the ORF34-37 and ORF35-37 mRNAs**

The 5' RACE data investigating the transcriptional start site of ORF36 mapped to the start site upstream of ORF35, however a few clones also mapped to within the coding region of ORF35, perhaps representing a low abundance transcript where ORF36 is the 5' cistron. To directly address whether ORF36 is capable of being translated from the ORF35-37 mRNA, the ORF35-37 mRNA was *in vitro* transcribed and then used to program the *in vitro* translation system based on rabbit reticulate lysates (RRLs); ORF35, 36 and 37-specific mRNAs were used for size controls. A strong signal was detected the ORF35-37 mRNA (**Figure 2.5; lane 1**) which co-migrated with the band from the ORF36 monocistronic control, indicating that this mRNA supports translation of an ORF which is downstream from the 5' proximal ORF35. ORF35 was not detected in this experiment likely due to its small size and conditions used for electrophoresis (**Figure 2.5; lane 1 & 4**). Thus, ORF36 expression from the ORF35-37 polycistronic mRNA can be detected in both *in vivo* transfection experiments and *in vitro* translation extracts, despite being positioned downstream of ORF35.

We then used this system to determine which protein(s) were translated from the minor mRNA ORF34-37 species detected by Haque et al. [3]. We *in vitro* transcribed the

ORF34-37 mRNA and used it to program RRLs. We detected a strong signal in the ORF34-37 sample (**Figure 2.5; lane 2 & 6**), which co-migrated with the band from the ORF34 monocistronic control. However no other protein products were detected, indicating that this mRNA supports translation of the 5' ORF34 cistron but none of the 3' proximal ORFs. This result indicated that the ORF34-37 mRNA, while having the coding potential for four proteins, is functionally monocistronic *in vitro* (**Figure 2.5; lane 2-5**).

### **Identification of a functionally bicistronic KSHV mRNA**

ORF35 is conserved between the  $\alpha$ ,  $\beta$ , and  $\gamma$ -herpesvirus subfamilies but its function remains unknown and antibodies are not available to detect it in KSHV-infected cells [90]. ORF35 is predicted to encode a 151-amino acid protein, and its start site resides in a favorable Kozak context. However, as ORF35 was not detected in the *in vitro* translation assay, we considered the possibility that ORF35 is not translated, instead serving as a portion of the 5' UTR for ORF36. In order to directly compare the levels of ORF35 and ORF36 protein produced from the bicistronic construct, we engineered in-frame HA tags at the 5' or 3' end of each respective gene, maintaining the native viral 5' UTR (5' UTR HA-ORF35-ORF36-HA). Monocistronic versions of each HA-tagged gene were also generated as controls (5' UTR HA-ORF35, ORF36-HA). Importantly, Western blotting with HA antibodies revealed that the ORF35 protein is produced from both the monocistronic and bicistronic constructs (**Figure 2.6A**).

Although our data indicated that the ORF35-37 transcript is functionally bicistronic, it was still formally possible that ORF36 translation occurred from a low-abundance monocistronic transcript generated by a cryptic internal promoter or splice site(s) in the DNA plasmid. To address this possibility, we transfected cells directly with *in vitro* transcribed monocistronic or bicistronic mRNAs, and performed anti-HA Western blots to detect each protein (**Figure 2.6B**). Again, both ORF35 and ORF36 proteins were produced from the bicistronic 5' UTR HA-ORF35-ORF36-HA mRNA, as well as from the appropriate control monocistronic mRNA, confirming that this locus is functionally polycistronic.

### **Measurement of ORF35 and ORF36 protein half-lives**

Next, we sought to quantify the ratio of ORF35 to ORF36 expression. Protein accumulation is dependent on both the initiation rate of the ORF35 and ORF36 start codons in addition to the protein half-life of each protein. Thus, we sought to establish the relative stability of ORF35 and ORF36 by performing a half-life analysis using cycloheximide, a rapidly acting inhibitor of eukaryotic protein synthesis [96]. Cell extracts were prepared at the indicated times after a cycloheximide elongation block and residual ORF35 and ORF36 proteins levels were detected by immunoblotting with anti-HA antibodies (**Figure 2.7**). Surprisingly, the half-life of ORF35 was significantly shorter than that of ORF36 and thus direct comparison of the rates of translation initiation at the ORF35 and ORF36 start codon was unattainable by this method; this comparison will be addressed further in Chapter four.

### **ORF36 translation is 5' m<sup>7</sup>G cap-dependent**

The only other known example in KSHV of translation of a downstream ORF from a polycistronic mRNA occurs via an IRES [68-70]. We therefore used an

established dual luciferase assay to determine whether an IRES similarly resides upstream of ORF36. The dual luciferase construct consists of a 5'-proximal *Renilla* luciferase gene that can be constitutively translated via a m<sup>7</sup>G cap-dependent mechanism, followed by a 3'-distal firefly luciferase gene, which is not normally translated. The two genes are separated by a defective encephalomyocarditis virus ( $\Delta$ EMCV) to prevent translational read-through [42,97]. Sequences of interest are then inserted between the  $\Delta$ EMCV and the firefly luciferase gene, and IRES activity leads to the translation of firefly luciferase. Sequences encompassing ORF35, ORF35-36 or ORF34-36 as well as two known IRES elements (EMCV and KSHV ORF72) were cloned into the dual luciferase construct (**Figure 2.8A**). The capped and polyadenylated *in vitro* transcribed mRNA was electroporated into either latent (uninduced; UI) or lytically reactivated (Induced; I) TREx BCBL1-RTA cells (**Figure 2.8B**). The integrity of the mRNAs was verified by Northern blotting with a DNA firefly probe (**Figure 2.8C**). After 4 h, the ratio of firefly/*Renilla* luciferase activity was measured to determine whether IRES activity was detectable in the context of latent or lytic infection. Although both the EMCV and ORF72 control IRES elements supported translation of firefly luciferase, none of the sequences upstream of ORF36 possessed detectable IRES activity in either UI or I cells (**Figure 2.8D**).

We next sought to determine whether ORF36 translation was instead initiated via a 5' m<sup>7</sup>G cap-dependent mechanism by inserting a strong 40 nucleotide hairpin (Hp7;  $\Delta G = -61$  kcal/mol) after nucleotide 32 within the 72 nucleotide native 5' UTR of the 5' UTR HA-ORF35-ORF36-HA construct (**Figure 2.8E**) [98]. Stable hairpin structures ( $\Delta G < -30$  kcal/mol) present near the 5' m<sup>7</sup>G cap dramatically reduce translation initiation by stalling the pre-initiation complex [98]. Translation of both ORF35 and ORF36 was markedly reduced in the presence of Hp7 following either DNA or RNA transfection (**Figure 2.8F and G**). Thus, recognition of the 5' m<sup>7</sup>G cap and subsequent 40S scanning are critical for translation of both ORF35 and ORF36.

## Discussion

This chapter describes the transcriptional and translation analysis of the KSHV *ORF34-37* genetic locus. A previous report identified two potential functionally polycistronic mRNAs that are transcribed from the KSHV *ORF34-37* genetic locus during lytic replication: a minor transcript encompassing ORF34-37 and a major transcript encompassing ORF35-37 [3,82]. In this chapter we found that the minor ORF34-37 transcript is likely monocistronic, supporting translation of solely ORF34 with the three downstream ORFs likely translationally silent. A previously uncharacterized ORF37-specific transcript was also identified and likely accounts for ORF37 (SOX) protein expression during infection. The presence of a ORF37-specific mRNA is in agreement with the detection of an independent BGLF5 transcript, the EBV SOX homolog, by Northern blotting in an EBV-infected cell line [91]. Notably, this study used a BCBL-1 cell line that bears an inducible version of the KSHV major lytic transactivator RTA, whereas prior transcript mapping of this locus used the standard BCBL-1 cell line [3,99]. Given that ORF37 is transcriptionally *trans*-activated by RTA, this may have increased the abundance of the transcript to the level detected in the present study [100].

Several lines of evidence support the notion that ORF36 is expressed in a 5' m<sup>7</sup>G cap-dependent manner as a 3'-proximal cistron. No transcript of an appropriate size with ORF36 as the 5'-proximal cistron was detected in KSHV-infected cells, in agreement with the results of 5' RACE that indicated its transcription starts upstream of ORF35 [3]. In addition, ORF36 protein expression was detected from an *in vitro* transcribed ORF35-37 mRNA used to program rabbit reticulate lysates and after transfection of an *in vitro* transcribed bicistronic RNA transcript. Finally, interfering with scanning from the 5' mRNA m<sup>7</sup>G cap via insertion of a hairpin blocked ORF36 translation, consistent with our failure to detect IRES activity within the locus. This is in contrast with the sole functionally bicistronic KSHV mRNA described to date, where an IRES within the coding region of ORF72 allows for ORF71 expression in a 5' m<sup>7</sup>G cap-independent manner [68-70].

Aside from being bicistronic, translation from the KSHV ORF35-37 transcript is unusual in that the protein product of ORF36 is at least as robustly expressed as the 5' ORF35 despite the fact that the ORF35 start codon is in a favorable sequence context. This may be in part due a longer protein half-life. However it is clear that a further strategy is required to allow for the robust bicistronic potential of the ORF35-37 mRNA. These findings provide the first example of 5' m<sup>7</sup>G cap-dependent non-canonical translation in KSHV.

## Materials and Methods

### Plasmid constructs

pcDNA3.1(+)-*ORF34-37* and -*ORF35-37* was generated by PCR-amplifying the *ORF34-37* or *ORF35-37* genetic locus from the KSHV-BAC36 (kindly provided by G. Pari [101]) and cloning it into the EcoRI/NotI sites of pcDNA3.1(+) (Invitrogen). pcDNA3.1(+)-5' UTR-HA-*ORF35* was assembled in a two-step process starting with the addition of the N-terminal HA tag after the native start ATG (nucleotide sequence: GCTTACCCATACGATGTACCTGACTATGCG) to the coding sequence amplified from the KSHV genome as above, followed by an overlap extension PCR to insert the 72 nucleotide (nt) native 5' UTR. The final product was then inserted into the pcDNA3.1(+) EcoRI/NotI restriction sites. pcDNA3.1(+)-*ORF36* was constructed by PCR-amplification of the *ORF36* coding sequence or to add the in frame C-terminal HA tag (GCTTACCCATACGATGTACCTGACTATGCGTGA) followed by insertion into EcoRI/NotI restriction sites. pCDEF3-*ORF37* is described elsewhere [88]. HA-*ORF35-ORF36*-HA was amplified from the KSHV-BAC36 using primers with additional HA tag sequences and inserted into the EcoRI/NotI sites of pcDNA3.1(+). This was followed by scarless insertion of the native 5' UTR via two-step sequential overlap extension PCR [101]. A stable hairpin structure (Hp7 sequence: GGGGCGCGTGGTGGCGGCTGCAGC CGCCACCACGCGCCCC, [98]) was inserted into the 5' UTR at nucleotide position 55599. Two bicistronic, dual luciferase constructs, a negative control ( $\Delta$ EMCV; mutated IRES sequence) and a positive control ( $\Delta$ EMCV element + functional EMCV) were kindly provided by P. Sarnow (Stanford University) [42,97]. *ORF72*, *ORF34-36*, *ORF35-36* and *ORF35* PCR amplicons were inserted into the EcoRI restriction site downstream of the  $\Delta$ EMCV element and upstream of firefly luciferase. The  $\Delta$ EMCV construct was also used as the backbone to insert the *ORF61* and *ORF37* promoter and *ORF35* coding sequences. The RTA-V5 construct was a kind gift from D. Lukac (Rutgers New Jersey Medical School). The primers used to generate these constructs are listed in **Table 2.1**.

### Cells, transfections and drug treatment

Human embryonic kidney 293T cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) (Gibco). BJABs, the KSHV-negative B cell line, were maintained in RPMI (Gibco) supplemented with 10% FBS, penicillin (100 U/ml, Gibco) and streptomycin (100  $\mu$ g/ml, Gibco). TREx BCBL1-RTA [99] cells were maintained in RPMI supplemented with 10% FBS, L-glutamine (200  $\mu$ M, Invitrogen), penicillin (100 U/ml), streptomycin (100  $\mu$ g/ml) and hygromycin B (50  $\mu$ g/ml, Omega Scientific). To induce lytic reactivation of KSHV, TREx BCBL1-RTA cells were split to  $1 \times 10^6$  cells/ml and induced 24 h later with 2-O-tetradecanoylphorbol-13-acetate (TPA; 20 ng/ml, Sigma), doxycycline (1  $\mu$ g/ml) and ionomycin (500 ng/ml, Fisher Scientific) [102]. For half-life studies, cycloheximide (100  $\mu$ g/ml) was added to 293T cells 24 h post-transfection.

For DNA transfections, constructs (1  $\mu$ g/ml) were transfected into subconfluent 293T cells grown in 12-well plates, either alone or in combination with 0.1  $\mu$ g/ml GFP as a co-transfection control using Effectene reagent (Qiagen) or Lipofectamine 2000 (Invitrogen) following the manufacturer's protocols. For RNA transfections, 3  $\mu$ g/ml of mRNA *in vitro* transcribed using the mMessage mMachine kit (Ambion) and

polyadenylated with yeast poly(A) polymerase (Epicentre Technologies) was transfected into ~90% confluent 293T cells grown in 12-well plates using Lipofectamine 2000. TREx BCBL1-RTA or BJAB cells were transfected with 20 µg of DNA per 10<sup>7</sup> cells via electroporation (250 V, 960 µF) with a Gene Pulser II (Bio-Rad, Hercules, CA).

### Dual Luciferase assays

Luciferase activities were determined using the dual-luciferase assay system (Promega Cat#: E1960) and a bench-top luminometer according to the manufacturer's protocol. Fold RTA *trans*-activation was calculated by obtaining the firefly/*Renilla* activity ratios for each of constructs containing the ORF37 promoter panel of mutants or positive (ORF61 promoter) or negative (ORF35 coding region) and dividing them by the firefly/*Renilla* ratio obtained from the ΔEMCV negative control. IRES activity was calculated by obtaining the firefly/*Renilla* activity ratios for each of the constructs containing the putative IRES sequences or the positive controls and dividing them by the ratio obtained from the ΔEMCV negative control. The value of fold activation represents at least three independent experiments with triplicate samples in each electroporation. Error bars represent the standard deviation between replicates.

### Western blots

Protein lysates were prepared in RIPA buffer [50 mM Tris-HCl (pH8.0), 150 mM NaCl, 1% (v/v) Nonidet P-40, 0.5% (w/v) sodium deoxycholate, 0.1% (w/v) sodium dodecyl sulfate (SDS)] containing protease inhibitors (Roche), and quantified by Bradford assay. Equivalent quantities of each sample were resolved by SDS-PAGE, transferred to a polyvinylidene difluoride membrane and incubated with the following primary antibodies: mouse monoclonal HA (1:2000, Invitrogen), rabbit polyclonal ORF36 (1:5000, kindly provided by Y. Izumiya [82]), goat polyclonal horseradish peroxidase (HRP)-conjugated actin (1:500, Santa Cruz Biotechnology), rabbit polyclonal SOX J5803 (1:5000, [89]), or mouse monoclonal S6RP (1:1000, Cell Signaling) followed by incubation with HRP-conjugated goat anti-mouse or goat anti-rabbit secondary antibodies (1:5000 dilution) (Southern Biotechnology Associates).

### Northern blots

Total cellular RNA was isolated for Northern blotting using RNA-Bee (Tel-Test). The RNA was then resolved on 1.2-1.5% agarose-formaldehyde gels, transferred to Nytran nylon membranes (Whatman) and probed with <sup>32</sup>P-labeled DNA probes made using either the RediPrime II random prime labeling kit (GE Healthcare) or the Decaprime II kit (Ambion). Strand-specific riboprobes specific for ORF36 and ORF37 were synthesized using the Maxiscript T7 kit (Ambion) with <sup>32</sup>P-labelled UTP. The probes used for Northern blot analysis spanned the following regions according to the nucleotide positions described by Russo et al. [64]: ORF35 probe: 55639-56091, ORF36 full-length probe: 55991-57310: ORF36-specific probe: 56093-56805: and ORF37 probe: 57273-58733. Results in each figure are representative of at least three independent replicates of each experiment.

### 5' Rapid Amplification of cDNA Ends (5' RACE)

Total RNA was isolated from TREx BCBL1-RTA cells as described above. 5' RACE was carried out using the First Choice™ RLM-RACE Kit (Ambion, Austin, TX) as per the manufacturer's protocol. Briefly, 5 µg of total RNA was ligated to the RNA

adapter after treatment with calf intestinal phosphatase and tobacco acid pyrophosphatase followed by reverse transcription using ORF36- or ORF37-specific primers. PCR was then performed using a primer specific for the 5' RACE adapter and primers nested within ORF36 and ORF37. RACE products were purified from an agarose gel and cloned into TOPO-TA vector for sequencing with the M13F primer.

### **Immunofluorescence assay to verify lytic reactivation**

To induce lytic reactivation of KSHV, TREx BCBL1-RTA cells were split to  $2 \times 10^5$  cells/ml and induced 24 h later with 2-O-tetradecanoylphorbol-13-acetate (TPA; 20 ng/ml, Sigma), doxycycline (1  $\mu$ g/ml) and ionomycin (500 ng/ml, Fisher Scientific) [102]. Cells were harvested by spinning at 1.2K rpm and then fixed in 4% formaldehyde, incubated for 10 min in permeabilization buffer [1% triton x-100 (vol/vol) and 0.1% sodium citrate (w/vol)], then for 30 min in block buffer [1% triton x-100 (vol/vol), 0.5% tween-20 (vol/vol) and 3% BSA (w/vol)]. Cells were then incubated with rabbit polyclonal SOX J5803 (1:500 dilution) or rabbit polyclonal LANA #6 (1:1000) [166], follow by incubation with either Alexa Fluor 488 or 546-conjugated goat anti-rabbit secondary antibody (Invitrogen). Coverslips were mounted in DAPI-containing Vectashield mounting medium (Vector labs) to visualize nuclei by fluorescence microscopy.

### ***In vitro* translation**

Construct expression was confirmed by using T7-promoter *in vitro* transcribed mRNAs to program rabbit reticulocyte lysate translation extracts (Promega) containing [ $S^{35}$ ]-labeled methionine (Perkin Elmer). The synthesized protein was then electrophoresed on a 10% SDS-PAGE gel and subsequently fixed and dried. Autoradiography was performed to detect synthesized polypeptides with incorporated [ $S^{35}$ ]-labeled methionine.

## Figures

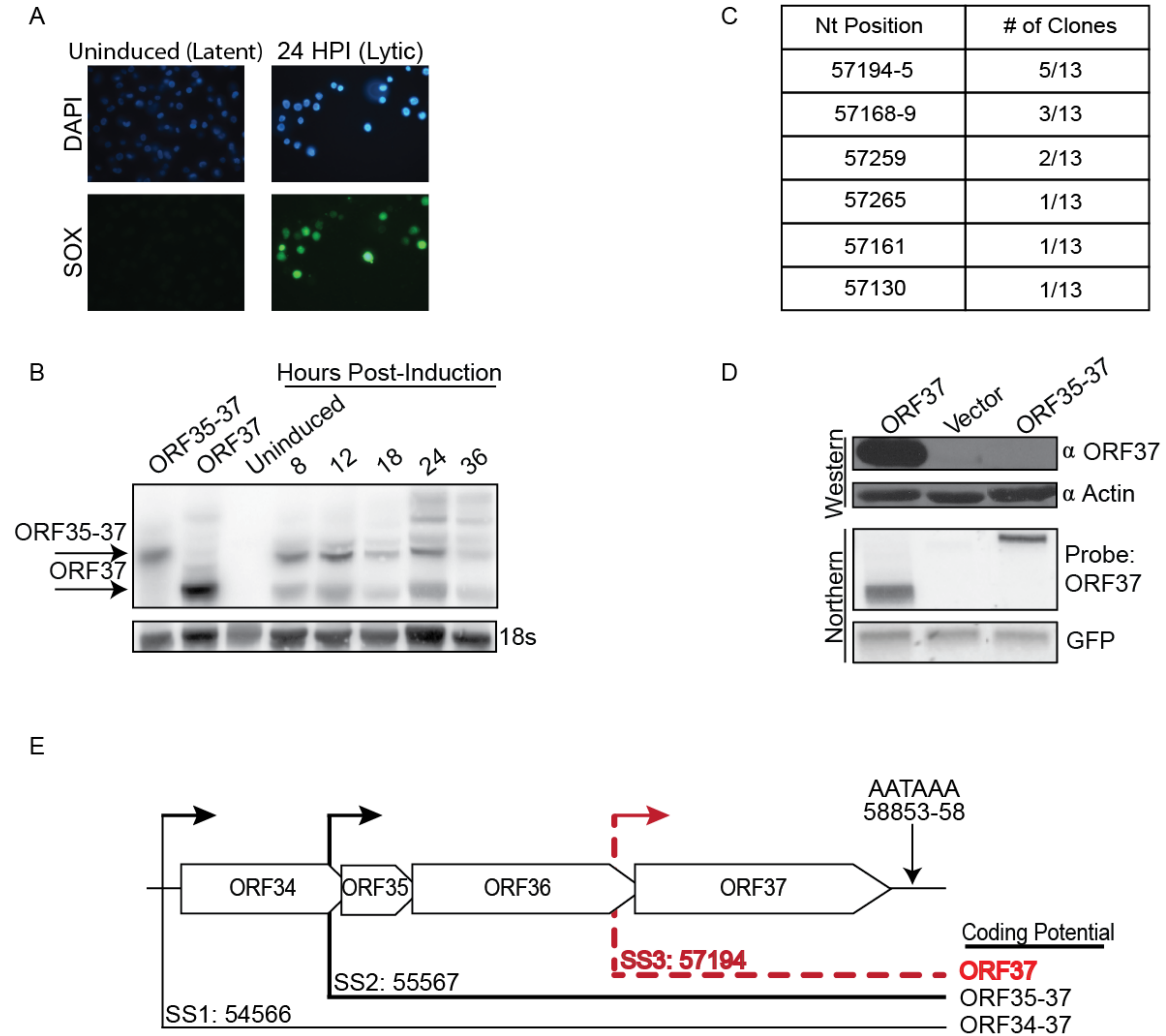


**Figure 2.1. Identification and transcript start site mapping of ORF37.** (A) Lytic reactivation was confirmed by performing immunofluorescence assay on either latent or lytic (24 h post-induction) cells using SOX as a marker of lytic reactivation. (B) TREx BCBL1-RTA cells were mock treated (latent) or lytically reactivated for the indicated times. RNA was then isolated and Northern blotted with a  $^{32}$ P-labeled ORF37 strand-specific riboprobe. (C) Total RNA was isolated from TREx BCBL1-RTA cells and 5' RACE was carried out using the First Choice™ RLM-RACE Kit (Ambion, Austin, TX) as per the manufacturer's instructions. PCR was carried out using the adapter primer and two rounds of nested primers specific for ORF37. (D) 293T cells were transfected with the indicated plasmid, and total protein and RNA were isolated 24 h later. Protein lysates were resolved by SDS-PAGE and detected by Western blot with antibodies against ORF37. Actin served as a loading control. To verify transcript integrity, RNA was Northern blotted with  $^{32}$ P-labeled ORF37 DNA probe or with a probe against the GFP co-transfection control. (E) Schematic of the previously uncharacterized ORF37

monocistronic transcript indicated by a dotted line in red with the most prominent (5/13) transcription start site identified by 5'-RACE.

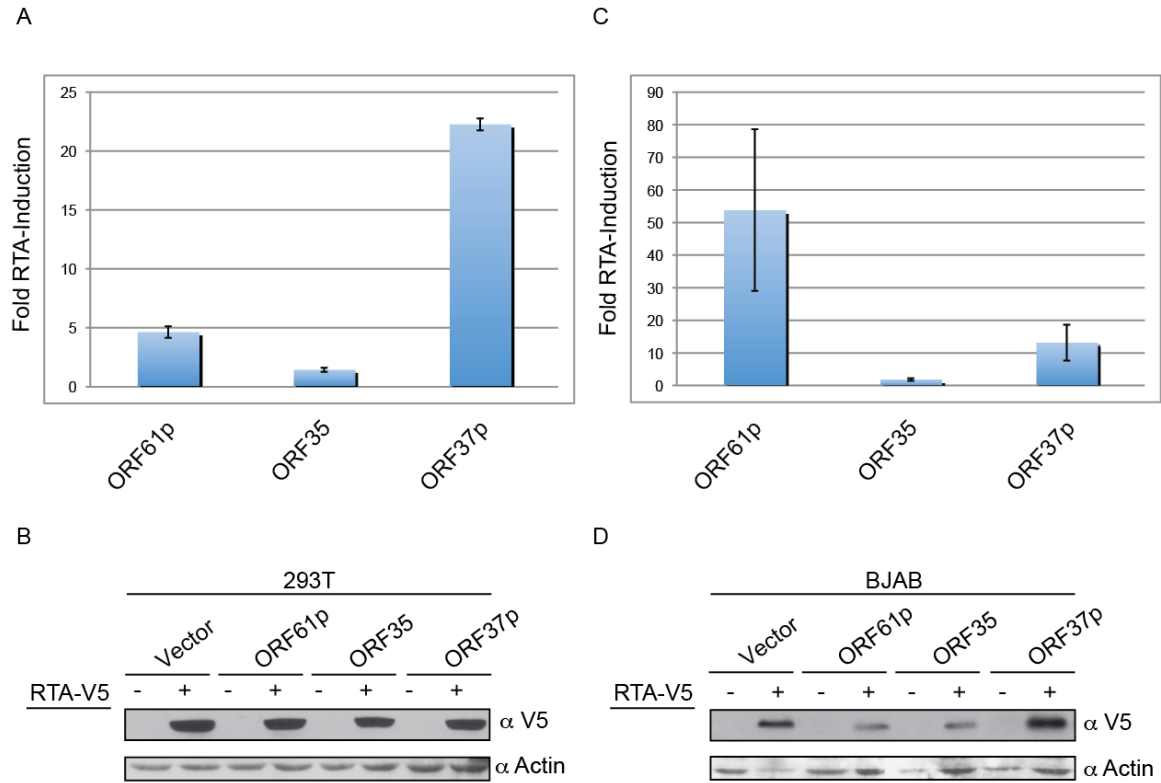


**Figure 2.2. ORF37 promoter is *trans*-activated by RTA.** Dual luciferase constructs with either no sequence (Vector), the ORF61 promoter (ORF61p), ORF35 coding region or the ORF37 promoter (ORF37p) inserted upstream of firefly luciferase were co-transfected with RTA-V5 into 293T cells (A) or BJAB cells (C). The fold induction of each construct by RTA was plotted, with the error bars representing standard deviation of the results of three independent experiments. (B, D) Protein lysates were resolved by SDS-PAGE and detected by Western blot with antibodies against V5 to detect RTA levels. Actin served as a loading control.

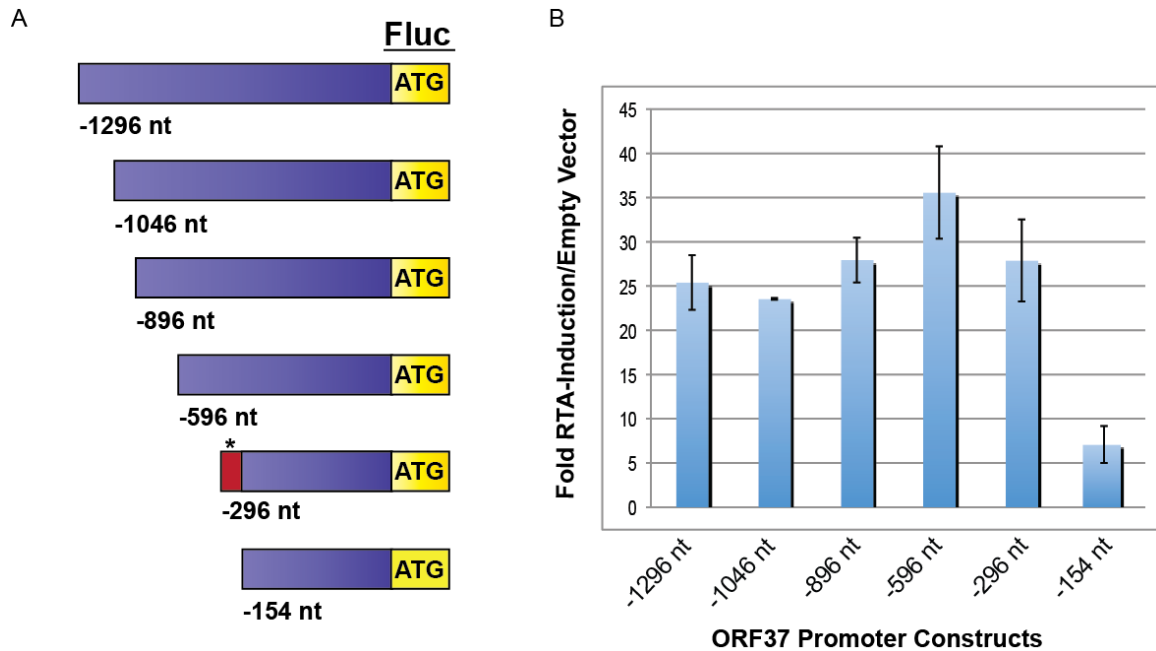


**Figure 2.3. Mapping of the RTA-responsive element in the ORF37 promoter.** (A) Schematic representation of serial deletions of the putative ORF37 promoter fused upstream of firefly in a dual luciferase construct. A 1296 nucleotide fragment upstream of the ORF37 ATG start codon was cloned in front of the firefly reporter gene. Deletions were engineered from the 5' end, with the number indicating the nucleotide position upstream of the translation initiation site. The red zone indicated by a star reflects the deleted region that results in a drop in RTA-responsiveness. (B) RTA activation of the ORF37 promoters with serial deletions. Reporter constructs with indicated promoter deletions were used to transfect 293T cells with either empty vector or RTA expression vector. The fold induction of each construct by RTA was plotted, with the error bars representing standard deviation of the results of three independent experiments.

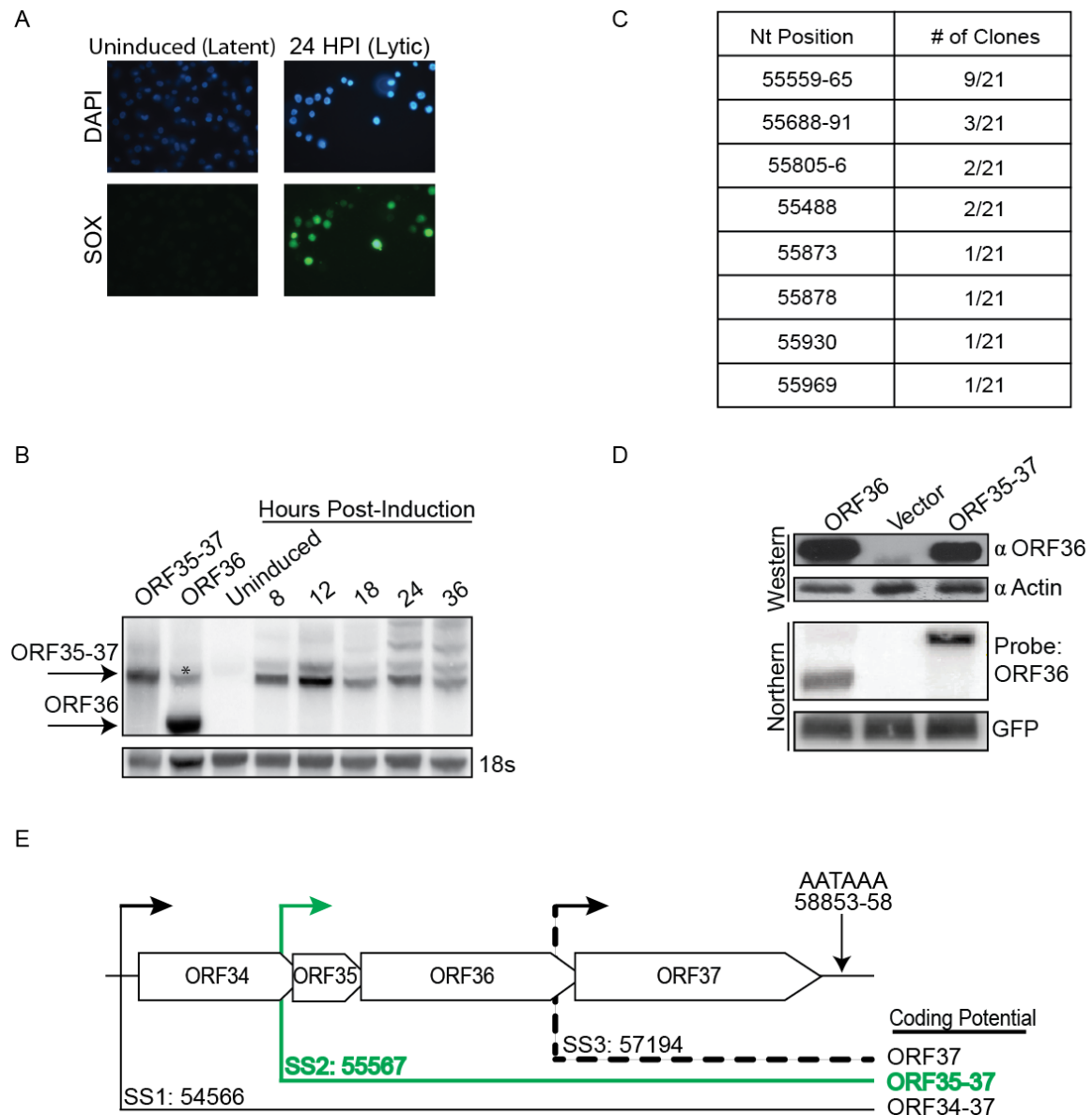


**Figure 2.4. Efficient translation of ORF36 from the full-length ORF35-37 tricistronic transcript.** (A) Lytic reactivation was confirmed by performing immunofluorescence assay on either latent or lytic (24 h post-induction) cells using SOX as a marker of lytic reactivation. (B) TREx BCBL1-RTA cells were mock treated (latent) or lytically reactivated for the indicated times. RNA was then isolated and Northern blotted with a  $^{32}$ P-labeled ORF36 strand-specific riboprobe. An additional higher molecular weight 293T-specific cross-reacting band was also detected in the ORF36 control lane, denoted by \*. (C) Total RNA was isolated from TREx BCBL1-RTA cells and 5' RACE was carried out using the First Choice™ RLM-RACE Kit (Ambion, Austin, TX) as per the manufacturer's instructions. PCR was carried out using the adapter primer and two rounds of nested primers specific for ORF36. (D) 293T cells were transfected with the indicated plasmid, and total RNA and protein were isolated 24 h later. Protein lysates were resolved by SDS-PAGE and detected by Western blot with antibodies

against ORF36. Actin served as a loading control. To verify transcript integrity, RNA was Northern blotted with  $^{32}\text{P}$ -labeled ORF36 DNA probes or with a probe against the GFP co-transfection control. (E) A schematic presentation of the *ORF34-37* genetic locus showing the previously identified ORF34-37 and ORF35-37 (green) polycistronic mRNAs with thin and thick lines respectively. The ORF37-specific transcript is denoted as a dotted line. Coding potentials are indicated on the right. Start sites (SS) are indicated for each transcript according to the nucleotide position described by Russo et al. [64]. The single poly(A) signal used by all four ORFs for transcription termination is shown.

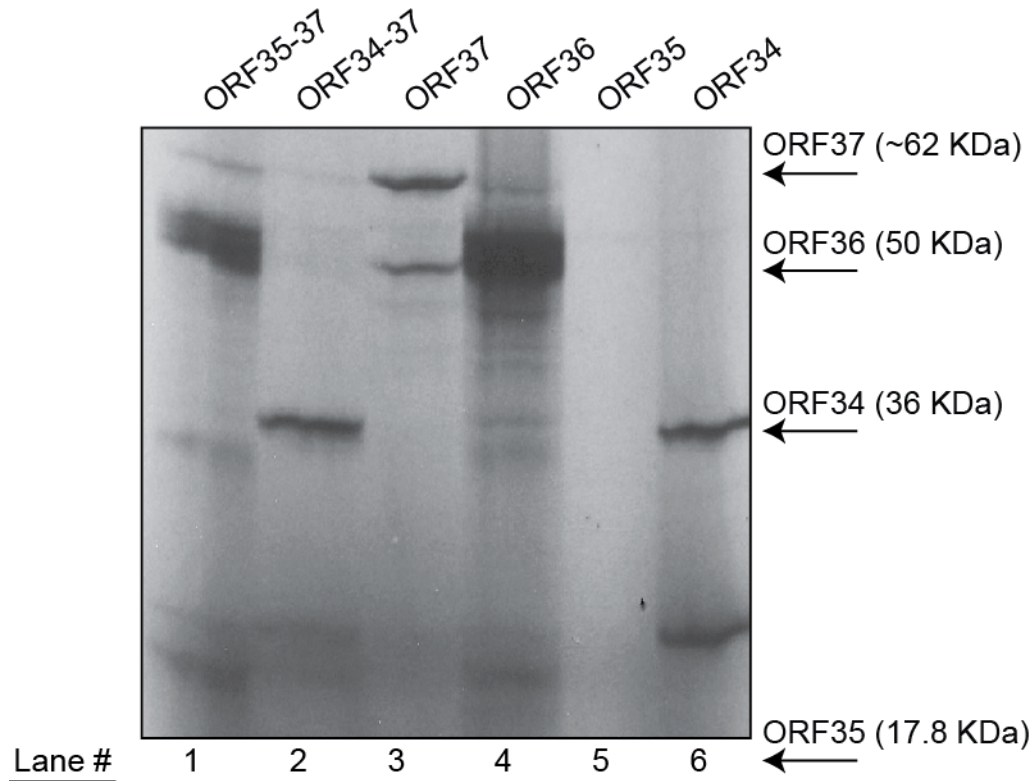


**Figure 2.5. *In vitro* translation of ORF34-37 and ORF35-37 transcripts.** The full length ORF34-37 or ORF35-37 and each of the ORF34, ORF35, ORF36 and ORF37 size controls were first *in vitro* transcribed by T7 RNA polymerase and then used to program rabbit reticulate lysates in the presence of  $[\text{S}^{35}]$ -labeled methionine. Protein lysates were then resolved by SDS-PAGE and visualized by autoradiography on x-ray film.

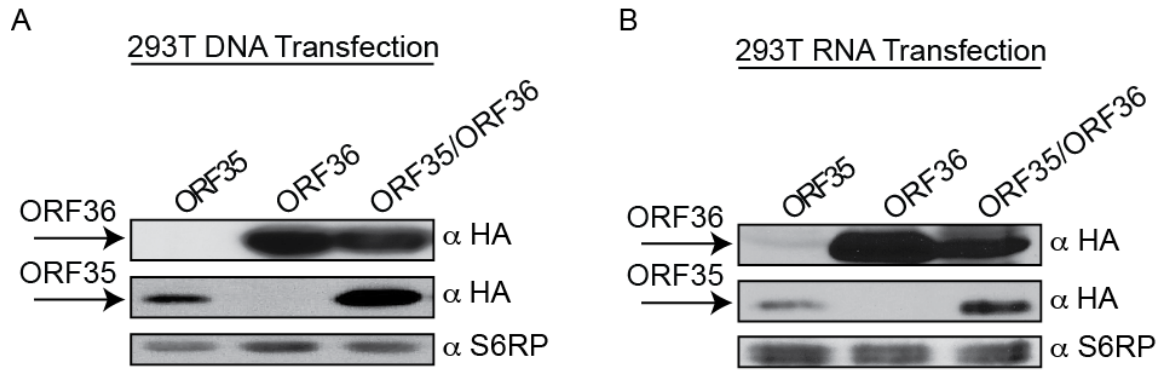


**Figure 2.6. The ORF35-37 mRNA is functionally bicistronic.** (A) Western blot analysis of 293T cells transfected with either N-terminally HA-tagged ORF35 with the native 5' UTR (ORF35), C-terminally HA tagged ORF36 (ORF36) or the full length 5' UTR HA-ORF35-ORF36-HA (ORF35/ORF36) DNA constructs. Equivalent amounts of protein lysates were resolved by SDS-PAGE and proteins were detected with anti-HA antibodies. (B) 293T cells were transfected with the indicated *in vitro* transcribed capped and polyadenylated RNA. Protein lysates were prepared from cells harvested 4 h post-transfection, resolved by SDS-PAGE and proteins were detected with anti-HA antibodies. The ribosomal protein S6RP served as a loading control for both experiments.

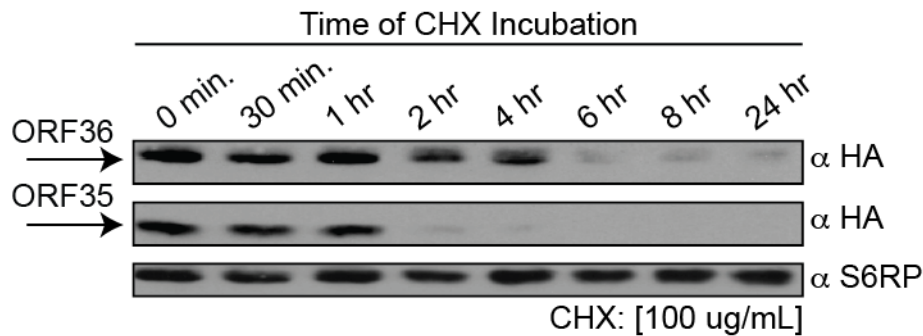


**Figure 2.7. Half-life analysis of ORF35 and ORF36.** 293T cells were transfected with the 5' UTR HA-ORF35-ORF36-HA construct. After 24 hr, the transfected 293T cells were treated with the translation elongation inhibitor cycloheximide at a concentration of 100  $\mu$ g/mL. Protein lysates were harvested at the indicated time points and immunoblotted with anti-HA antibodies. The ribosomal protein S6RP served as a loading control.

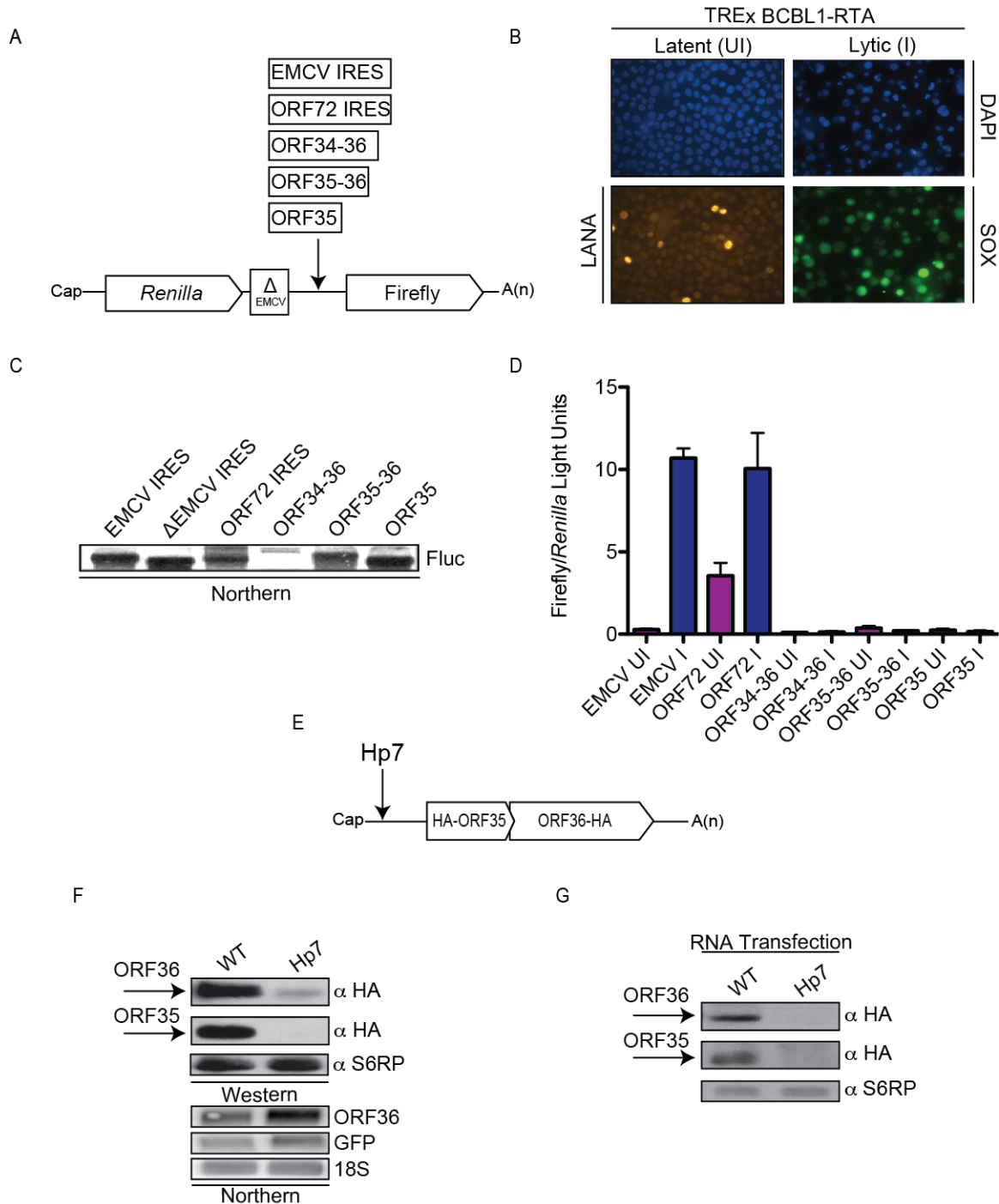


**Figure 2.8. Translation of ORF36 is independent of IRES activity and dependent on the 5' mRNA m<sup>7</sup>G cap.** (A) Diagram of dual luciferase transcripts. (B) Lytic reactivation was confirmed by performing immunofluorescence assays on either latent (uninduced; UI) or lytic (24 h post-induction; I) cells using LANA as a marker of latency (left panel) or SOX as a marker of lytic reactivation (right panel). (C) *In vitro* transcribed mRNAs were examined by Northern blot analysis using a <sup>32</sup>P-labeled fluc DNA probe to confirm the presence of full length RNA species. (D) The indicated *in vitro* transcribed,

polyadenylated transcripts were electroporated into lytically reactivated TREx BCBL1-RTA cells. A dual luciferase assay was performed 4 h post-electroporation to determine the relative levels of firefly and *Renilla* luciferase activity. The experiment was performed in triplicate; error bars represent the standard deviation between replicates. (E) Schematic of 5' UTR HA-ORF35-ORF36-HA containing a  $\Delta G = -61$  kcal/mol hairpin (Hp7) inserted after nucleotide position 32 in the native 5' UTR. (F) 293T cells were transfected with the indicated WT or Hp7 plasmid shown in (E), and equivalent amounts of protein lysates were resolved by SDS-PAGE and Western blotted with anti-HA antibodies. The ribosomal protein S6RP served as a loading control. RNA samples were examined by Northern blot analysis with a  $^{32}\text{P}$ -labeled ORF36 DNA probe. GFP served as a co-transfection control. 18S rRNA was used as a loading control. (G) 293T cells were transfected with the indicated *in vitro* transcribed capped and polyadenylated RNA as shown in (E). Protein lysates were prepared from cells harvested 4 h post-transfection, resolved by SDS-PAGE and proteins were detected with anti-HA antibodies. The ribosomal protein S6RP served as a loading control.

**Table 2.1. List of oligonucleotide primers used in chapter two.**

| <b>Primer Name</b>                  | <b>Primer Sequence 5'-3'</b>                                             |
|-------------------------------------|--------------------------------------------------------------------------|
| <b>ORF61p F</b>                     | GGG CCC GAA TTC CCG CCT CCC CCA CTC AGG CCT TT                           |
| <b>ORF61p R</b>                     | GGG CCC GAA TTC GAC CTT ACA GAA ACA CAG TCC AG                           |
| <b>ORF34-37 F</b>                   | GC GAA TTC GGG ACA GTG TCG CGT GAA TGT C                                 |
| <b>ORF35-37 F</b>                   | GCGAATTCGTCCAACCTACGGGCGACTATCT                                          |
| <b>ORF35-37 R</b>                   | GAGCTCGCGGCCGCAATGCGGGATATGGGGATTGC                                      |
| <b>HA-ORF35 F</b>                   | GGGCCCCGAATTCATGGCTTACCCATACGATGTACCTGACTATGCGGACT<br>CAACCAACTCTAAAAGAG |
| <b>ORF35 R</b>                      | GAGCTCGCGGCCGCTTAGGGAGTTTCAGGGCACAC                                      |
| <b>ORF36 F</b>                      | GGGCCCCGAATTCATGGAGAGGAGACCCCCACTC                                       |
| <b>ORF36 R</b>                      | GAGCTCGCGGCCGCTCAGAAAACAAGTCCGCGGGTG                                     |
| <b>ORF36-HA R</b>                   | GAGCTCGCGGCCGCTCACGCATAGTCAGGTACATCGTATGGG                               |
| <b>ORF36 probe F</b>                | GCGGCCGCGCTACCCGGATTTTCAGAGAGAC                                          |
| <b>ORF36 probe R</b>                | GAATTCTAACACTGGAAGAGGACGCAAGA                                            |
| <b>Firefly probe F</b>              | CATATCGAGGTGGACATCACT                                                    |
| <b>Firefly probe R</b>              | GGTAGATGAGATGTGACGAACG                                                   |
| <b>Hp7</b>                          | GAATTCGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGCGCCCCGG<br>TACC                 |
| <b>ORF72 F</b>                      | GGGCCCCGAATTCATGGCAACTGCCAATAACCC                                        |
| <b>ORF72 R</b>                      | GGGCCCCGAATTCGGTGCCGGCTTGTATATGTGA                                       |
| <b>ORF34-36 F – Dual Luciferase</b> | GCGAATTCGGGACAGTGTCGCGTGAATGT                                            |
| <b>ORF34-36 R – Dual Luciferase</b> | GAATTCTCAGAAAACAAGTCCGCGGG                                               |
| <b>ORF35-36 F – Dual Luciferase</b> | GGGCCCCGAATTCATGGACTCAACCAACTCTAAAAGAG                                   |
| <b>ORF35-36 R – Dual Luciferase</b> | GAATTCTCAGAAAACAAGTCCGCGGG                                               |
| <b>ORF35 F – Dual Luciferase</b>    | GGGCCCCGAATTCATGGACTCAACCAACTCTAAAAGAG                                   |
| <b>ORF35 R – Dual Luciferase</b>    | GGGCCCCGAATTCATGGAGTTTCAGGGCACA                                          |

\*F: Forward

\*R: Reverse

## Chapter Three: Dual Upstream Open Reading Frames Control Translation of the ORF35-37 Polycistronic mRNA

### Background

Translation initiation of eukaryotic mRNAs is dependent on the 5' mRNA m<sup>7</sup>G cap and proceeds by ribosomal scanning until recognition of an AUG codon in a favorable context [6,8]. As a consequence of the translation machinery not engaging start codons at internal positions within the mRNA, eukaryotic transcripts generally encode only one functional protein. For the majority of mRNAs the most 5'-proximal AUG is selected, however strategies exist to bypass upstream start codons to enable downstream initiation. For example, leaky scanning can occur if the nucleotides flanking the 5'-proximal AUG are not in the Kozak consensus sequence (ccRccAUGG), allowing the 40S ribosomal subunit to instead engage a downstream methionine codon [8,9]. Alternatively, when an upstream AUG is followed shortly thereafter by an in-frame termination codon, ribosomes can reinitiate translation, albeit with reduced efficiency, at a downstream AUG. These upstream open reading frames (uORFs) presumably permit translation of a downstream gene because factors necessary for initiation have not yet dissociated during the short elongation period. Notably, uORFs are common regulatory elements in eukaryotic transcripts, and generally function to reduce translation of the major ORF [9,20]. Additional, albeit rare, examples of internal ORF translation also exist, for example after ribosome shunting over a highly structured upstream sequence [16,71,74,103], or upon direct 40S recruitment via internal ribosome entry sites (IRESs) [42,44,104-106].

In this chapter, we describe how two uORFs regulate the relative expression levels of the ORF35 and ORF36 proteins from the ORF35-37 transcript. While translation of ORF35 likely occurs after leaky scanning past the two short, overlapping uORFs located in the 5'-untranslated region (UTR), translation of ORF36 occurs via a reinitiation mechanism after engagement both uORFs. Thus, KSHV uses a host strategy normally reserved to repress translation of the major ORF to instead permit expression of a 3'-proximal cistron on a viral polycistronic mRNA.

## Results

### The context of the ORF35 start codon does not dramatically effect expression of ORF36

It is notable that ORF36 protein production is robust given that its translation requires the pre-initiation complex to bypass the relatively strong Kozak context surrounding the ORF35 start codon (AgaAUGG) and to scan through 424 nucleotides of upstream sequence. To determine whether the context of the ORF35 start codon influences the expression of ORF36, we mutated the preferred nucleotide (A) at position -3 to the least preferred nucleotide (U) (**Figure 3.1A**). As expected, ORF35 expression was reduced, however unexpectedly this mutation did not significantly alter ORF36 expression, arguing against a pure leaky scanning mechanism to explain ribosomal access to the ORF36 start site (**Figure 3.1B**). Direct transfection with *in vitro* transcribed mRNAs confirmed that this result was not due to induction of an alternative promoter (**Figure 3.1C**). Thus, the relative strength of the ORF35 start site does not dramatically influence ORF36 translation, suggesting that there is an alternative mechanism in place that disfavors initiation at the 5' gene.

### Identification of two translation competent short upstream open reading frames in the 5'UTR

We searched for features of the ORF35-37 transcript sequence that might contribute to translational start site selection. Within the native 72-nucleotide 5' UTR we noticed two short upstream ORFs (uORFs). The first 9 codon uORF, dubbed uORF1, spans KSHV nts 55603 to 55603 and has an AUG residing in a relatively weak Kozak context (CguAUGA) [64]. The second 11 codon uORF (uORF2) spans KSHV nts 55627-55660, resides in an intermediate context (AccAUGA) and overlaps with both the 3' end of uORF1 and the ORF35 start codon (**Figure 3.2A**).

We therefore sought to determine whether these uORFs were indeed recognized by the translation machinery. Due to their small size, uORF-generated peptides tend to be highly unstable and are very difficult to detect. To circumvent this problem, we made a single nucleotide change in each uORF to place them in frame with ORF35 lacking its AUG ( $\Delta$ 35), thereby generating uORF-ORF35 fusions (**Figure 3.2B**). Thus, restoration of ORF35 expression is a direct readout translation initiation from the uORF start codon. In both cases, the uORF fusions restored ORF35 expression to levels corresponding to the relative strength of the Kozak context flanking the start codon of each uORF (**Figure 3.2C**) [8,12].

### Analysis of the influence of uORF1 on ORF35 and ORF36 expression

uORFs are common in mammalian transcripts and generally act to repress translation of the downstream ORF. If a uORF is present, translation of the major ORF occurs either by a relatively inefficient termination-reinitiation event or by context-dependent leaky scanning past the uORF start codon [20]. To determine the contribution of uORF1 towards ORF35 and ORF36 translation, we mutated the uORF1 start site ( $\Delta$ 1). ORF35 expression was elevated in the  $\Delta$ 1 mutant in the DNA transfection, yet repressed in the RNA transfection (**Figure 3.3A-D**). We re-positioned the HA tag at the 5' end of ORF35 and found that this did not alter the translational regulation of the DNA transfection by showing similar results upon repositioning of the HA tag internally within

ORF35 (**Figure 3.3E and F**). Thus, ORF35 expression undergoes modest negative regulation or positive regulation by ribosomal engagement at the uORF1 start codon depending on whether DNA or RNA transfection was performed, although in either case disruption of uORF1 does not appear to influence ORF36 expression. This discrepancy will be examined in more detail later in this chapter.

### **A second uORF acts to repress ORF35 translation and stimulate translation of ORF36**

In the uORF2-ORF35 fusion, the uORF2 fusion abrogated expression of ORF36 (**Figure 3.2C**). The uORF2 start codon is in a more favorable Kozak context than that of uORF1, and disruption of the uORF2 AUG (AUG → UUG;  $\Delta 2$ ) or weakening the Kozak context of its start codon (KCS2 wkn) increased ORF35 translation and severely decreased translation of ORF36 in both DNA and RNA transfection experiments (**Figure 3.4A-D**). Notably, the  $\Delta 2$  mutant was designed to ensure the uORF1 stop codon remains intact, permitting the independent analysis of uORF1 and uORF2. Unlike uORF1, uORF2 therefore plays a dominant role in regulating expression of both genes in this polycistronic mRNA.

Although a few rare uORFs have been found to function in a sequence-dependent manner [24-28], for most characterized uORFs it is the act of translation rather than the peptide sequence that mediates their function. The fact that several amino acids of uORF2 sequence are altered in the construct bearing the HA tag at the 5' end of ORF35 is in agreement with the amino acid sequence of uORF2 not being the primary determinant of its activity. Indeed, rebuilding the uORF2 mutants into a construct in which the HA tag was moved to an internal position in ORF35 yielded indistinguishable results (**Figure 3.4E and F**).

### **The ORF36 start codon is accessed by linear scanning**

Translation reinitiation at the internal ORF36 start codon could occur either after linear scanning of the 40S complex through the 332-nucleotide intercistronic region between uORF2 and ORF36 or through shunting of the complex past this sequence and its subsequent positioning proximal to ORF36. To distinguish between these possibilities, two strong hairpins (Hp7) that impede scanning were inserted within the 5'-proximal or 3'-proximal coding region of ORF35 (**Figure 3.5A**). If the 40S ribosomal subunit were shunted past these internal sequences, one or both of the hairpins (depending on the location of the shunting sites) should not compromise ORF36 translation [16,17]. However, we observed a significant reduction in ORF36 expression in the presence of either hairpin, arguing that the 40S complex scans in a linear fashion through ORF35 (**Figure 3.5B**).

One potential caveat is that the insertion of the hairpins might dramatically alter the RNA folding landscape, disrupting a secondary structure required for shunting. To exclude this possibility, the single natural methionine codon present within the coding region of ORF35 was mutated to an arginine (**Figure 3.5C; Midmut**). If this internal sequence were bypassed via shunting after uORF2 termination, the natural start codon should not be able to compete with the ORF36 AUG for the pre-initiation complex. However, we found that ORF36 expression was increased from the MidMut construct, arguing against a shunting mechanism and further suggesting that this methionine normally engages a fraction of the scanning ribosomes before they can reach the ORF36

start codon (**Figure 3.5D**). Collectively, these data support a model in which the preferential recognition of uORF2 diverts ribosomes past the ORF35 start codon, whereupon they scan in a linear fashion and reacquire the pre-initiation complex before reinitiating translation at a downstream start codon.

### **The length of uORF2 and the uORF2-ORF36 intercistronic distance is critical for expression of the 3'-proximal ORF36**

Translation of a major ORF following engagement at a uORF generally occurs via a termination-reinitiation event. The length of a uORF is important for reinitiation, as it is thought that due to the short duration of uORF translation some of the translation initiation accessory factors have not yet dissociated prior to termination at the uORF stop codon [20]. In this regard, translation of the downstream ORF decreases dramatically if the time required to complete translation of the uORF is increased, for example by increasing the ORF length or inserting secondary structure to stall the ribosome [107,108]. Therefore, we reasoned that if ORF36 translation initiates using the same 40S ribosomal subunit involved in translation of uORF2, then artificially elongating uORF2 should inhibit ORF36 expression. This experiment was performed on the construct backbone with the ORF35 HA tag located internally to mimic the wild type length of uORF2. Indeed, extension of uORF2 from 11 to 64 codons (uORF2-64 aa) resulted in a drop in ORF36 expression (**Figure 3.6A-C**).

The rate-limiting step of reinitiation is postulated to be the re-acquisition of the pre-initiation complex (eIF2-GTP-Met-tRNA<sub>i</sub>) during ribosomal scanning, and thus a sequence of sufficient length must be present downstream of the uORF for this to occur [9,20]. We therefore evaluated how the distance between the uORF2 stop codon and the subsequent start codon influenced reinitiation within the viral mRNA. Start codons in a favorable Kozak context were inserted at two positions between the uORF2 stop codon and the ORF36 start site. We hypothesized that start codons located close to uORF2 would not be as efficiently recognized, and therefore they would not inhibit ORF36 expression. However, more distally located start codons should better engage the initiation machinery, thereby preventing translation from occurring at the authentic ORF36 start site. In agreement with this prediction, a start codon positioned 16 nucleotides downstream of uORF2 did not strongly inhibit ORF36 expression, whereas a methionine positioned 246 nucleotides after termination of uORF2 severely compromised ORF36 expression (**Figure 3.6D and E**). These data support the conclusion that engagement of the ORF36 start codon is dependent on the reacquisition of the pre-initiation complex after termination of uORF2 translation.

### **No *cis*-acting element within ORF36 is required for reinitiation and uORF2 does not function when supplied *in trans***

Next, we asked whether additional *cis*-acting elements within ORF36 are required for its translation after uORF2 engagement. For this experiment, we replaced the ORF36 gene with a GFP reporter (**Figure 3.7A**). GFP protein was expressed robustly as a downstream gene from this construct, arguing against a requirement for an element within ORF36 for its translation (**Figure 3.7B; lane 3**). Similar to our results with ORF36, disruption of uORF2 compromised expression of GFP (**Figure 3.7B: lane 3 vs. lane 4**), supporting a uORF2-dependent mechanism as the primary pathway enabling translation of a downstream gene from this locus.

Although our data indicate that uORF2 functions in *cis* to allow for the translation of a downstream cistron by a termination-reinitiation mechanism, it remained formally possible that uORF2 could also exert its function in *trans* on a separate mRNA molecule [109,110]. To address this possibility, we co-transfected cells with either the WT or  $\Delta 2$  bicistronic 5' UTR HA-ORF35-ORF36-HA and a reporter construct with the ORF36 coding region replaced with the *Renilla* luciferase gene with uORF2 intact (36-rluc) or disrupted ( $\Delta 2$  36-rluc) (**Figure 3.7C**). Immunoblot analysis served as a read-out of ORF36 and thus uORF2 translation from the bicistronic 5' UTR HA-ORF35-ORF36-HA construct in each sample (**Figure 3.7D**). As expected, expression of *Renilla* luciferase was abrogated in the  $\Delta 2$  36-rluc compared to 36-rluc when co-transfected with empty vector, however co-transfection with either WT or  $\Delta 2$  5' UTR HA-ORF35-ORF36-HA construct did not restore expression of *Renilla* luciferase (**Figure 3.7E**). Thus, uORF2 does not function in *trans* to rescue expression of a downstream cistron on an mRNA molecule lacking uORF2.

### **uORF1 and uORF2 function synergistically to govern ORF35 and ORF36 translation initiation**

The above analysis of translation initiation of the ORF35-37 polycistronic mRNA used Western blotting of lysates from cells transfected with plasmid DNA or *in vitro* transcribed RNA containing the 5'UTR HA-ORF35-ORF36-HA [111]. To more quantitatively assess the translational control exerted by both uORF1 and uORF2, reporter constructs (5'UTR-uORF2-*Rluc* and 5'UTR-ORF35 $\Delta$ 98-*Rluc*) were generated with the ORF35 or ORF36 coding region, respectively, replaced with the *Renilla* luciferase gene. These reporter constructs allow for the initiation at each ORF (AUG<sub>ORF35</sub> and AUG<sub>ORF36</sub>) to be monitored separately but from the authentic upstream sequence environment (**Figure 3.8A and C**). This plasmid backbone also harbors a firefly luciferase under the control of an independent promoter to provide an internal control of transfection efficiency.

Earlier in the chapter we investigated the independent roles of uORF1 and 2, we therefore sought to determine their combined effect on translation of the ORF35-37 polycistronic mRNA. In this background, similar to experiments performed with the native 5'UTR HA-ORF35-ORF36-HA, disruption of uORF2 ( $\Delta 2$ ) led to increased ORF35 translation and severely decreased translation of ORF36 (**Figure 3.8B and D**). However, disruption of AUG<sub>uORF1</sub> ( $\Delta 1$ ) in the dual luciferase background did not significantly influence expression of either ORF35 or ORF36 (**Figure 3.8B and D**). However, to our surprise, disruption of both AUG<sub>uORF1</sub> ( $\Delta 1$ ) and AUG<sub>uORF2</sub> ( $\Delta 2$ ) led to a synergistic increase in ORF35 expression and decrease in ORF36 expression (**Figure 3.8B and D**).

Collectively, these data suggest a model where with AUG<sub>uORF1</sub> disrupted, ribosomes continue to scan and alternatively initiate at AUG<sub>uORF2</sub>, with a similar net result of ORF35 repression and reinitiation at ORF36 (**Figure 3.8B and D; lane 1 vs. 2**). However when the AUG<sub>uORF2</sub> is additionally disrupted, the ribosomes that would have been captured by AUG<sub>uORF2</sub> now initiate at AUG<sub>ORF35</sub> and can no longer reinitiate at ORF36. This would lead to a synergistic increase in ORF35 expression and a corresponding drop in ORF36 levels (**Figure 3.8B; lane 3 vs. 4**). Thus, it would appear that uORF1 and uORF2 are both repressive elements of ORF35, although uORF1 requires the presence of uORF2 to exert this function. Finally, these data suggest that a

certain fraction of ribosomes terminating after uORF1 translation, similar to uORF2, reinitiate translation at the AUG<sub>ORF36</sub>; reinitiation likely does not occur at the AUG<sub>ORF35</sub> due to an insufficient uORF1-ORF35 intercistronic distance.

### **ORF36 expression occurs via reinitiation after uORF translation**

We next sought to directly assess whether ribosomes that translate uORF1 and/or uORF2 access AUG<sub>ORF36</sub> by a termination-reinitiation mechanism using a toeprinting approach [112-114]. The toeprinting assay uses reverse transcriptase to extend a fluorescently labeled primer on a mRNA template in translation extracts, with the size of the cDNA reflecting the location of the bound ribosome [115]. Toeprinting requires the 80S ribosome to stall at the AUG initiator codon, which is accomplished by combining cycloheximide (CHX) and sparsomycin (SPR), two potent translation elongation inhibitors [96,116]. An automated DNA sequencing machine then resolves the fluorescently labeled cDNAs [115,117]. Upon performing this experiment, a ribosome triplet peak was indeed observed 12-15 nts downstream of AUG<sub>uORF1</sub>, AUG<sub>uORF2</sub>, AUG<sub>ORF35</sub> and AUG<sub>ORF36</sub> reflecting ribosome stalling in the initiation codon in the ribosome P-site (**Figure 3.9A and B; +CHX/SPR**) [118]. We confirmed that the predicted peak indeed mapped to the start codon of interest by performing toeprinting on mRNAs lacking AUG<sub>uORF1</sub> ( $\Delta 1$ ), AUG<sub>uORF2</sub> ( $\Delta 2$ ), AUG<sub>ORF35</sub> ( $\Delta 35$ ) and AUG<sub>ORF36</sub> ( $\Delta 36$ ). As expected, the peak at each start codon was abolished in the corresponding  $\Delta$ AUG transcript (**Figure 3.9C-F**). Parallel controls included the incubation of the mRNAs in RRL lacking cycloheximide and sparsomycin (-CHX/SPR) and in the absence of RRL (-RRL).

After confirming that ribosomes could be detected at each start codon, we then compared the toeprint intensity at the AUG<sub>ORF36</sub> when elongation inhibitors were added to extracts prior to adding the mRNA template ( $T_0$ ) or added after the translation of the mRNA was underway ( $T_{15}$  or  $T_{30}$ ). Elongation inhibitor treatment at  $T_0$  reveals where ribosomes first initiate translation while drug treatment after translation has been allowed to proceed ( $T_{15}$  or  $T_{30}$ ) indicates the position of both primary and reinitiation events (**Figure 3.10A**) [37]. Therefore, to directly assess whether a termination-reinitiation mechanism is indeed the biochemical basis whereby the ribosomes gain access to the ORF36 start codon, the peak at the AUG<sub>ORF36</sub> was compared between the  $T_0$  and  $T_{15/30}$  conditions either with AUG<sub>uORF1</sub> and AUG<sub>uORF2</sub> intact (WT) or disrupted ( $\Delta 1/\Delta 2$ ) (**Figure 3.10B-C**). Indeed, the ribosome occupancy at AUG<sub>ORF36</sub> increased from  $T_0$  to  $T_{15}$  to  $T_{30}$  in the WT construct although not in the  $\Delta 1/\Delta 2$  construct (**Figure 3.10D**). As expected, analysis of the ribosome occupancy at AUG<sub>ORF35</sub> revealed that in the  $\Delta 1/\Delta 2$  mutant ORF35 expression was elevated in each toeprint condition over WT (**Figure 3.10E**). These data support the conclusion that engagement of the ORF36 start codon is dependent on ribosomes that have previously translated uORF1/2.

### **The ORF35 start codon is at least partially accessed by leaky scanning past AUG<sub>uORF1</sub> and AUG<sub>uORF2</sub>.**

uORF1 and 2 synergistically repress translation of ORF35 indicating that the majority of if not all the ribosomes that translate uORF1 do not reinitiate at AUG<sub>ORF35</sub>, likely due to the inadequate intercistronic distance between AUG<sub>uORF1</sub> and AUG<sub>ORF35</sub>. Thus, it is the question arises why does ORF35 expression persist despite the presence of two upstream repressive elements? We therefore sought to determine whether translation

initiation at AUG<sub>ORF35</sub> was initiated from ribosomes that have leaky scanned past both AUG<sub>uORF1</sub> and AUG<sub>uORF2</sub>. To address this possibility, the -3, -2, -1 and +4 unfavorable kozak consensus sequence flanking AUG<sub>uORF1</sub> was mutated to the preferred context (CguAUGA → AccAUGG; KCS1 enh) and the intermediate context flanking AUG<sub>uORF2</sub> was enhanced at the -4 position from A to G (AccAUGA → AccAUGG; KCS2 enh) (**Figure 3.11A and C**). We then assessed how these mutations influences the expression of either ORF35 or ORF36 (**Figure 3.11B and D**). Enhancing the AUG<sub>uORF1</sub> and AUG<sub>uORF2</sub> contexts either independently or in combination led to a repression of ORF35, however intriguingly did not abolish expression to background levels ( $\Delta 35$ ) (**Figure 3.11B**). This data is consistent with the pre-initiation complex initiating with higher fidelity at AUG<sub>uORF1</sub> and AUG<sub>uORF2</sub>, allowing for fewer ribosomes to leaky scan past the uORFs in favor of AUG<sub>ORF35</sub>. In agreement with this model, both the KCS1 and KCS2 mutants led to an increase in ORF36 expression, although only KCS2 was significant, likely due to increase recognition of the uORF start codons and thus reinitiation at AUG<sub>ORF36</sub> (**Figure 3.11D**).

### Ratio of ORF35 and ORF36 translational initiation

Finally, we sought to quantify the ratio of ORF35 to ORF36 expression. Protein accumulation is dependent on both the initiation rate of the ORF35 and ORF36 start codons in addition to the half-life of each protein. To bypass the complication that different ORF35 and ORF36 have different half-lives (**Figure 2.7**), we again used the 5'UTR-uORF2-*Rluc* and 5'UTR-ORF35 $\Delta$ 98-*Rluc* constructs to eliminate the distinct half-lives (**Figure 3.12A**). We then evaluated directly the ratio of *Renilla* to firefly luciferase from each construct and found that the ratio of translation initiation at AUG<sub>ORF35</sub> to AUG<sub>ORF36</sub> is 0.5125:1.0 (**Figure 3.12B**). Thus, the ORF35-37 mRNA integrates leaky scanning past uORF1 and uORF2 to allow the expression of the 5'-proximal ORF35 in combination with a synergistic termination-reinitiation mechanism after both uORF1 and uORF2 to allow for translation of the 5'-proximal ORF36. This mechanism results in fairly balanced initiation rates for both proteins, though ORF36 accumulates to higher levels due to its increased protein stability.

## Discussion

Collectively our data indicate that the ORF36 start codon is accessed via a reinitiation event after translation of uORF1 and uORF2, allowing the ribosomes to bypass the favorable ORF35 start codon (**Figure 3.13**). In support of this model, lengthening uORF2 to decrease the efficiency of reinitiation abrogated ORF36 expression. The contribution of uORF1 was revealed by generating a uORF1/2 double mutant which leads to a synergistic increase of ORF35 expression and a corresponding drop in ORF36 levels. Furthermore, when translation elongation inhibitors were added to arrest translation elongation in RRLs prior to mRNA addition, ribosomes collected primarily at the upstream start codons yet comparatively little at AUG<sub>ORF36</sub>. However, when the drugs were added after allowing translation to proceed for a short time period, the toeprint intensity increased in a uORF1/2-dependent manner. Here we provide both genetic and biochemical evidence that a termination-reinitiation mechanism is required for ribosomes to translation ORF36, revealing a rare example of uORFs enhancing translation of a downstream major ORF.

To date, the only described short uORF that enables access to the start codon of a downstream gene in a polycistronic transcript was identified in hepatitis B virus (HBV). The HBV uORF, dubbed C0, weakly inhibits the 5'-proximal C ORF while stimulating translation of the 3'-proximal J and P proteins [71,119]. However, the termination-reinitiation event described for HBV may be facilitated by a shunting mechanism, as non-linear scanning was found to occur in the homologous region in the related duck hepatitis B virus [120]. This appears not to be the case for ORF36 because insertion of strong hairpins within the coding region upstream strongly compromises ORF36 expression, suggesting that the ribosomes are scanning continuously from the 5' mRNA m<sup>7</sup>G cap to the ORF36 start codon.

uORFs are common features found in the 5' UTRs of many eukaryotic mRNAs and are widely recognized as *cis*-regulatory element that reduce translation of the major ORF [21]. Initiation at the primary ORF instead occurs by leaky scanning or a low-efficiency reinitiation event, which is in agreement with the function of uORF1 as a negative regulator of ORF35 [20,21,121]. A few cases have been described in which the ability of the uORF to repress downstream translation is dependent on the amino acid sequence of the encoded peptide [24-28]. For example, a uORF present in the 5' UTR of the human cytomegalovirus gp48 gene attenuates downstream translation in a sequence-dependent fashion, likely by delaying normal termination and preventing leaky scanning by the 40S ribosomal subunit to reach the downstream AUG [26]. However, in general, engagement of the translation apparatus rather than the translated product itself represses translation of the major ORF. Indeed, regulation of the ORF35-37 transcript appears independent of the uORF peptide sequence because the 5' HA-tagged construct had 5 of the 11 uORF2 codons mutated yet still functioned to permit translation of ORF36. Moreover, uORFs in homologous regions of the genome in related  $\gamma$ -herpesviruses lacked amino acid conservation.

Factors that influence the ability of a terminating ribosome to resume scanning remain poorly understood. The size of the uORF is a considerable restriction on reinitiation in eukaryotes as reinitiation only occurs following translation of a short 5'cistron [10]. This is supported by the many examples of mRNAs that contain silent 3' cistrons, including ORF37 in the ORF35-37 transcript, as well as numerous other KSHV polycistronic mRNAs [4,10,66,94]. One notable exception is found within cauliflower mosaic virus, where a viral protein was discovered to promote reinitiation of ribosomes that had translated full-length 5' cistron, possibly by preventing the loss of eIF3 from the ribosomes during the initial translation event [30,122]. No hard rule exists to specify the size of a uORF that is permissive for reinitiation, perhaps because the nucleotide length is not the only variable which determines the duration of elongation. A study that introduced a pseudoknot that slows elongation, without altering the uORF length, found that reinitiation was abolished [9]. On the ORF35-37 mRNA, lengthening uORF2 from 11 to 43 codons did not dramatically alter reinitiation efficiency, however when the uORF was further lengthened to 64 aa, expression of ORF36 was compromised.

It has been shown using chimeric preproinsulin mRNAs that efficient reinitiation progressively improves upon lengthening the intercistronic sequence up to 79 nucleotides [19]. Sufficient intercistronic sequence length is thought to be necessary to allow time for the scanning 40S ribosomal subunit to reacquire eIF2-GTP-Met-tRNA<sub>i</sub> prior to encountering the downstream start codon, although at what point the sequence length becomes inhibitory is not known [20,108]. In the context of the viral ORF35-37 transcript, the ribosome is able to reinitiate translation with a high frequency despite scanning 361 or 332 nucleotides after terminating translation of uORF1 or uORF2, respectively, indicating that intergenic regions significantly longer than 79 nucleotides still enable reinitiation.

Interestingly, a prior report identified a translational enhancer element within the tricistronic S1 mRNA of avian reovirus that functions to increase expression of a downstream cistron. This occurs as a consequence of sequence complementarity to 18S rRNA, which is reminiscent of the prokaryotic Shine-Dalgarno sequence [123,124]. A similar strategy of having 18S rRNA complementarity within a bicistronic mRNA was also found to enhance the ability of the minor calicivirus capsid protein VP2 to be translated by reinitiation [33,36]. Whether enhancer elements exist in the KSHV uORF-ORF36 intercistronic region to facilitate translation at the downstream cistron remains to be determined. This study did not formally rule out that an *cis* sequence element in the intercistronic uORF2-ORF36 region served to enhance the efficiency of reinitiation. Although analysis of the intervening sequence revealed one 6-nucleotide motif that was complementary to 18S rRNA, disruption of this motif did not alter the levels of ORF36 expression (data not shown). Reinitiation at the ORF36 start codon occurred both in a transient transfection and *in vitro* system, ruling out that a viral protein was required for reinitiation, although not for enhancement. However, no essential reinitiation element exists downstream of the ORF36 start codon, as replacement of these sequences with GFP does not block its translation. This is distinct from the termination-reinitiation mechanism described for certain retrotransposons, which require complex downstream secondary structures [125].

We have established that ribosomes are able to translate uORF1 and 2, terminate and then reinitiate at ORF36. Yet, despite the presence of two repressive uORFs

expression of ORF35 persists at levels ~0.5 of ORF36. We suspected that ribosomes were simply able to leaky scanning past uORF1 and uORF2 to gain access to the ORF35 start codon. In support of this model weakening the context surrounding the uORF2 start codon enhanced ORF35 expression, suggesting that the ORF35 start site is primarily reached by ribosomes that have bypassed the AUG of uORF2. However, enhancing both uORF Kozak consensus sequences to the ideal context did not abrogate ORF35 to background levels [8]. Intriguingly, it has been previously reported that two AUG's in close proximity ( $\leq 15$  nts) modifies leaky scanning such that initiation is no longer strictly sequential with 5'-3' polarity, but competitive [126-128]. Initiation at such closely spaced start codons is thought to occur by ~15 nucleotide backward excursions during the 5'-3' ribosome scanning process, although this proposed mechanism remains shrouded in controversy [128]. Ten nucleotides separate the AUG<sub>uORF2</sub> and AUG<sub>ORF35</sub>, thus the possibility exists that expression of ORF35 is in part due to its close proximity to the uORF2 start codon. Indeed, repositioning AUG<sub>uORF2</sub> further upstream compromises expression of ORF35 (data not shown), however stepwise insertions to increase the distance between AUG<sub>uORF2</sub> and AUG<sub>ORF35</sub> paired with detailed toeprinting analysis would be required to further test this hypothesis. Thus, the work in this chapter has extended the known roles of uORFs, depicted in **Figure 3.14A-F**, to include allowing for the translation of a bicistronic mRNA (**Figure 3.14G**). Whether the close proximity of the uORF to the 5'-cistron is an absolute requirement for bicistronic potential of this mRNA remains to be experimentally verified.

## Materials and Methods

### Plasmid constructs

To construct 5' UTR-ORF35iHA-ORF36-HA, a backbone plasmid consisting of 5'UTR ORF35-ORF36-HA was first generated by PCR-amplification from the KSHV-BAC36 with HA tag sequences solely for ORF36 and inserted into the EcoRI/NotI sites of pcDNA3.1(+). This construct was then linearized by inverse PCR at nucleotide position 55795 followed by ligation-independent cloning using InFusion (Clontech) with primers consisting of an HA tag flanked by 15 base pair regions of vector overlap. For the 5' UTR HA-ORF35 $\Delta$ 96-HA-GFP construct, HA-GFP was inserted between the NotI/XbaI restriction sites in pcDNA3.1(+), and the 5' UTR-HA-ORF35  $\Delta$ 96 fragment was then inserted between the EcoRI/NotI restriction sites upstream of HA-GFP.

psiCHECK-2-5'UTR-*Rluc* was generated by first PCR-amplifying the 5'UTR and first 20 nucleotides of ORF35 from pcDNA3.1(+)-ORF35-37 construct described previously as well as 15 base pair regions of vector overlap [111]. The psiCHECK-2 backbone was then linearized by inverse PCR immediately prior to the *Rluc* start codon followed by ligation-independent cloning using InFusion (Clontech). psiCHECK-2-5'UTR-ORF35 $\Delta$ 98-*Rluc* was generated by PCR-amplifying the 5'UTR and first 352 nucleotides of ORF35 from pcDNA3.1(+)-ORF35-37 construct described previously as well as 15 base pair regions of vector overlap [111]. The psiCHECK-2 backbone was then linearized by inverse PCR immediately prior to the native ORF36 start codon followed by ligation-independent cloning using InFusion (Clontech). The primers used to generate these constructs are listed in **Table 3.1**.

Where specified, parental plasmids were subjected to site-directed mutagenesis using the QuikChange kit (Stratagene) as per the manufacturer's protocol. The context of the ORF35 start codon was weakened by mutating the wild type AgaAUGG to UgaAUGG (35 KCS wkn). The uORF1 fusion to  $\Delta$ 35 was generated by mutating the uORF1 stop codon UGA to UGG (uORF1- $\Delta$ 35). The uORF2 fusion to  $\Delta$ 35 was generated by deleting one nucleotide (A) located immediately prior to the ORF35 start codon (uORF2- $\Delta$ 35). uORF1 and uORF2 mutants (designated  $\Delta$ 1 and  $\Delta$ 2) were generated by substituting the AUG start codon with AGA or UGA, respectively. Two codons within in the ORF35 coding region were converted to AUGs in a strong context: (1) AccAACU to AccAUGG and (2) AauUUUG to AauAUGG. The native AUG residing at location 55779-81 within the ORF35 coding region was mutated to an AGA (MidMut). uORF2 was lengthened from 11 to 64 codons by mutating the first UAA stop codon to AGA, the second UAA stop codon to CAA, the third UGA stop codon to CGA, and the fourth and fifth UAG stop codon to CAG, resulting in the use of the next downstream stop codon (uORF2-64 aa). The ORF35 and ORF36 start codons were disrupted by mutating the wild type AUG to AGA ( $\Delta$ 35 and  $\Delta$ 36). The uORF1 Kozak context was strengthened by mutating the wild-type CguAUGA to AccAUGG (KCS1 enh). The uORF2 Kozak context was weakened by mutating the wild-type AccAUGA to UuuAUGA (KCS2 wkn) and strengthened to AccAUGG (KCS2 enh). The uORF1 stop codon is disrupted by the KCS2 enh mutated, it was therefore re-introduced immediately

3' from the original stop codon in the KCS2 enh and KCS1/2 enh constructs (TCA → TGA).

### **Cells, transfections and drug treatment**

Human embryonic kidney 293T cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) (Gibco).

For DNA transfections, constructs (1 µg/ml) were transfected into subconfluent 293T cells grown in 12-well plates, either alone or in combination with 0.1 µg/ml GFP as a co-transfection control using Effectene reagent (Qiagen) or Lipofectamine 2000 (Invitrogen) following the manufacturer's protocols. For RNA transfections, 3 µg/ml of mRNA *in vitro* transcribed using the mMessage mMachine kit (Ambion) and polyadenylated with yeast poly(A) polymerase (Epicentre Technologies) was transfected into ~90% confluent 293T cells grown in 12-well plates using Lipofectamine 2000. TREx BCBL1-RTA cells were transfected with 20 µg of DNA per 10<sup>7</sup> cells via electroporation (250 V, 960 µF) with a Gene Pulser II (Bio-Rad, Hercules, CA).

### **Western blots**

Protein lysates were prepared in RIPA buffer [50 mM Tris-HCl (pH8.0), 150 mM NaCl, 1% (v/v) Nonidet P-40, 0.5% (w/v) sodium deoxycholate, 0.1% (w/v) sodium dodecyl sulfate (SDS)] containing protease inhibitors (Roche), and quantified by Bradford assay. Equivalent quantities of each sample were resolved by SDS-PAGE, transferred to a polyvinylidene difluoride membrane and incubated with the following primary antibodies: mouse monoclonal GFP (1:2000, BD Biosciences), mouse monoclonal HA (1:2000, Invitrogen), rabbit polyclonal ORF36 (1:5000, kindly provided by Y. Izumiya [82]), goat polyclonal horseradish peroxidase (HRP)-conjugated actin (1:500, Santa Cruz Biotechnology), rabbit polyclonal SOX J5803 (1:5000, [89]), rabbit polyclonal ORF57 (1:5000, kindly provided by Z. Zheng [129]), rabbit polyclonal LANA #6 (1:1000) or mouse monoclonal S6RP (1:1000, Cell Signaling) followed by incubation with HRP-conjugated goat anti-mouse or goat anti-rabbit secondary antibodies (1:5000 dilution) (Southern Biotechnology Associates).

### **Northern blots**

Total cellular RNA was isolated for Northern blotting using RNA-Bee (Tel-Test). The RNA was then resolved on 1.2-1.5% agarose-formaldehyde gels, transferred to Nytran nylon membranes (Whatman) and probed with <sup>32</sup>P-labeled DNA probes made using either the RediPrime II random prime labeling kit (GE Healthcare) or the Decaprime II kit (Ambion). Results in each figure are representative of at least three independent replicates of each experiment.

### **Synthesis of mRNAs**

Plasmid DNA (psiCHECK-2) was linearized with XhoI and used as templates for transcription by T7 RNA polymerase (mMessage mMachine T7; Ambion) to yield each mRNAs. Transcription reactions contained m<sup>7</sup>GTP cap analog (Ambion) to yield naturally capped mRNAs.

### **Primer extension-inhibition (toeprinting) analysis to detect translation initiation**

## **complexes**

Toeprinting assays were carried out according the method described by Gould et al. [115]. The DNA oligonucleotide 5'- TTC TCG GCG TGC TTC TCG GAA T -3' was used as the reverse primer positioned within *Rluc*. The primer was labeled with the fluorophore 5-carboxyfluorescein (5-FAM) (Life Technologies) for fragment analysis assay. To anneal to primer and mRNA, twenty picomoles of primer and 500 ng mRNA were combined in 50 nM Tris-HCl, pH 7.5, and heated to 68°C for two minutes and cooled to 37°C for eight minutes in a total volume of 10 µl.

The ribosome binding reactions used micrococcal nuclease-treated rabbit reticulocyte lysate (RRL) (Promega). Reaction mixtures were assembled on ice in a total volume of 25 µl containing 50% (v/v) reticulocyte lysates, 500 µg/mL cycloheximide (CHX) (Sigma), 200 µM sparsomycin (SPR), 2 mM DTT, 20 µM amino acid mixture minus methionine, 20 µM amino acid mixture minus leucine (Promega) and 1.5 U/µl Ribolock RNase inhibitor (Thermo Scientific). Template mRNA with the annealed 5'- FAM primer (0.5 ng/reaction) were added to allow initiation complex assembly at 30°C for 20 minutes post-drug addition. CHX/SPR were either pre-incubated at 30°C for 5 minutes to allow the drugs to interact with the translational machinery ( $T_0$ ) or 15 or 30 minutes after the mRNA:primer complex had been added to the translation extract ( $T_{15}$  or  $T_{30}$ ).

The reverse transcriptase primer inhibition reaction was carried out in a total volume of 20 µL containing 4 µl of the ribosome binding reaction or 0.5 ul annealed mRNA/primer, 5X Superscript II Buffer, 5 mM dithiothreitol (DTT), 0.8 mM each dNTP, 1.5 U/µl Ribolock RNase inhibitor and 5 U/µl SUPERSCRIPT® II RNase H- reverse transcriptase (Invitrogen) and placed at 30°C for 35 min. Primer extension products were extracted by adding an equal volume of phenol:CHCl<sub>3</sub> pH 8, followed by precipitation with 1/10th the volume of 3 M NaOAc (pH 5.2) and 2.5 volumes of EtOH. cDNAs were mixed with 9.5 ul HiDi formamide (Genetic Analysis Grade from Applied Biosystems Inc.) and 0.5 ul ROX 500 (Gel Company; fragment sizes: 70, 80, 90, 100, 120, 140, 160, 180, 190, 200, 240, 260, 280, 300, 320, 340, 360, 380, 400, 425, 450, 475, 490, 500). Fragment analysis products and were heated to 95°C for 3 minutes to denature, then flash cooled on wet ice for 2 minutes. Samples were then analyzed on a 96-capillary 3730xl DNA Analyzer system with a POP7 polymer and 50 cm array. Injection voltage: 1.5kV. Injection time: 15 seconds. Run temperature: 63°C (Applied Biosystems). Toeprint peaks correspond to the inhibition of reverse transcription of the RNA template by ribosomes with the initiation codons in their P sites and are located 15-17 nts downstream of base A of the start codon [118]. Fluorescent cDNA analysis was performed using PeakScanner v2.0 software (ABI).

## **Graph Design and Statistical Analysis**

Graphs were designed using either excel or the GraphPad Prism Software version 5 all and statistical analysis performed using the GraphPad Prism Software version 5. Data were analyzed for statistical significance using the Student's two-tailed unpaired *t*-test.

## **Dual Luciferase assays**

Luciferase activities were determined using the dual-luciferase assay system (Promega Cat#: E1960) and a bench-top luminometer according to the manufacturer's

protocol. The *Rluc*/Fluc value represents at least three independent experiments with triplicate samples in each transfection. Error bars represent the standard deviation between replicates.

## Figures

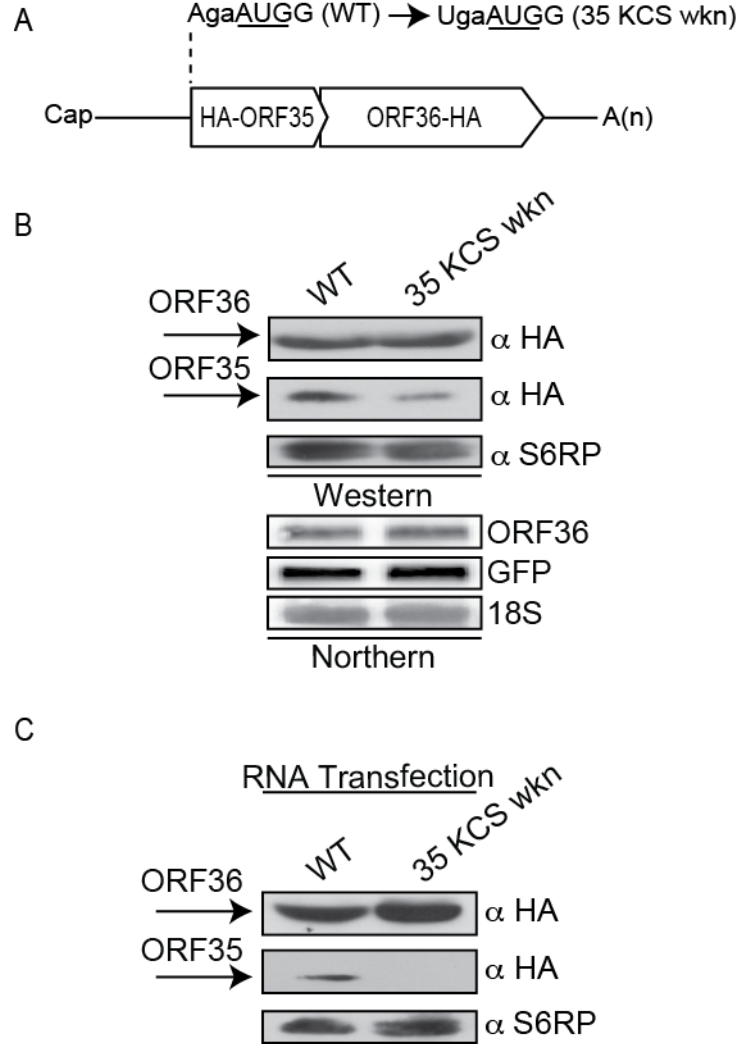


**Figure 3.1. The ORF35 start codon does not dramatically influence expression of ORF36.** (A) Schematic of 5' UTR HA-ORF35-ORF36-HA indicating the nucleotide mutated to weaken the Kozak context flanking the ORF35 AUG (35 KCS wkn). (B) 293T cells were transfected with the indicated WT or 35 KCS wkn plasmid shown in (A) and equivalent amounts of protein lysates were resolved by SDS-PAGE and Western blotted with anti-HA antibodies. The ribosomal protein S6RP served as a loading control. RNA samples were examined by Northern blot analysis with a <sup>32</sup>P-labeled ORF36 DNA probe. GFP served as a co-transfection control. 18S rRNA was used as a loading control. (C) 293T cells were transfected with the indicated *in vitro* transcribed capped and polyadenylated RNA. Protein lysates were harvested 4 h post-transfection,

resolved by SDS-PAGE and detected with anti-HA antibodies. The ribosomal protein S6RP served as a loading control.

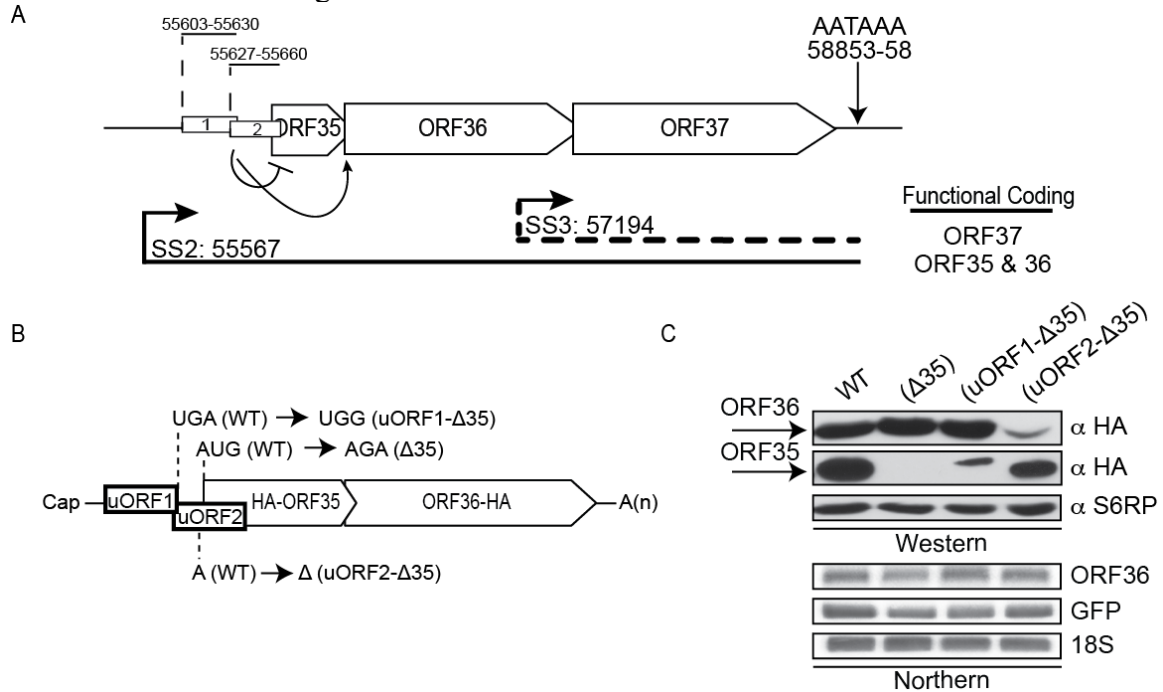


**Figure 3.2. Two uORFs located in the 5'UTR engage the translation initiation machinery.** (A) Schematic of the ORF35-37 genetic locus with the position of two uORFs indicated. (B) The ORF35 start codon mutant (AUG→AGA; Δ35) and uORF fusion reporter RNAs are depicted schematically. uORF1-Δ35 has the uORF1 stop codon disrupted (UGA→UGG) while uORF2-Δ35 has one nucleotide deleted from uORF2 to shift the reading frame +1 (A→Δ). (C) 293T cells were transfected with the indicated wild-type or mutant plasmids, and 24 h post-transfection protein lysates were resolved by SDS-PAGE and Western blotted with anti-HA antibodies. The ribosomal protein S6RP served as a loading control. RNA samples were examined by Northern blot analysis with a <sup>32</sup>P-labeled ORF36 DNA probe. GFP served as a co-transfection control. 18S rRNA was used as a loading control.

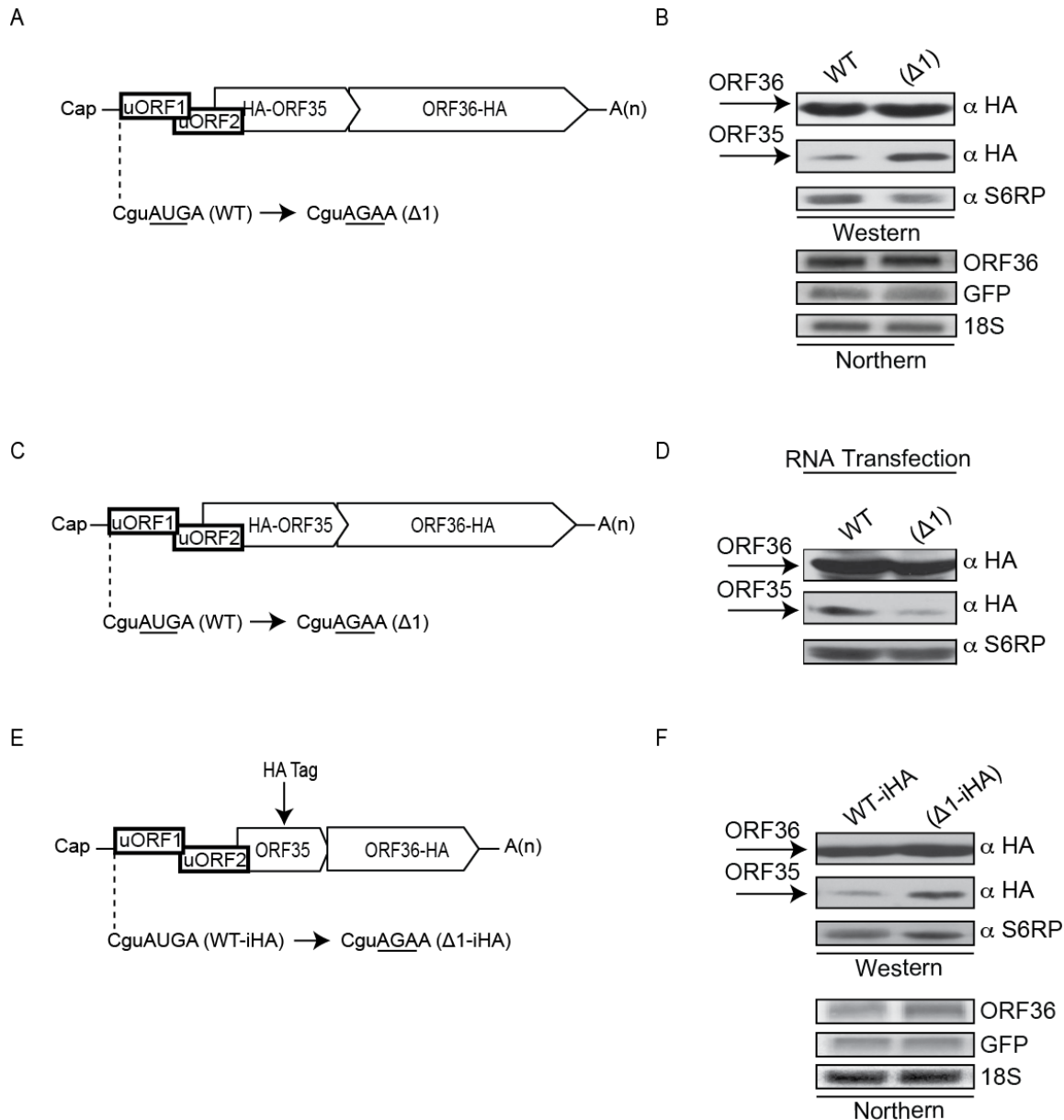


**Figure 3.3. Analysis of the contribution of uORF1 on ORF35 and ORF36 expression.** (A, C) Schematic representation of uORF1 indicating the nucleotides mutated to disrupt the uORF1 AUG (Δ1). (B, F) 293T cells were transfected with the indicated wild-type or mutant plasmids, and 24 h post-transfection protein lysates were resolved by SDS-PAGE and Western blotted with anti-HA antibodies. The ribosomal protein S6RP served as a loading control. RNA samples were examined by Northern blot analysis with a <sup>32</sup>P-labeled ORF36 or GFP DNA probe. GFP served as a co-transfection control in B and F. 18S rRNA was used as a loading control. (D) 293T cells were transfected with the indicated *in vitro* transcribed capped and polyadenylated RNA. Protein lysates were harvested 4 h post-transfection, resolved by SDS-PAGE and detected with anti-HA antibodies. The ribosomal protein S6RP served as a loading control. (E) Schematic representation of the uORF1 mutations introduced into a

construct with the native 5' UTR-ORF35-ORF36-HA with an HA tag positioned internally and in-frame with ORF35 (WT-iHA).

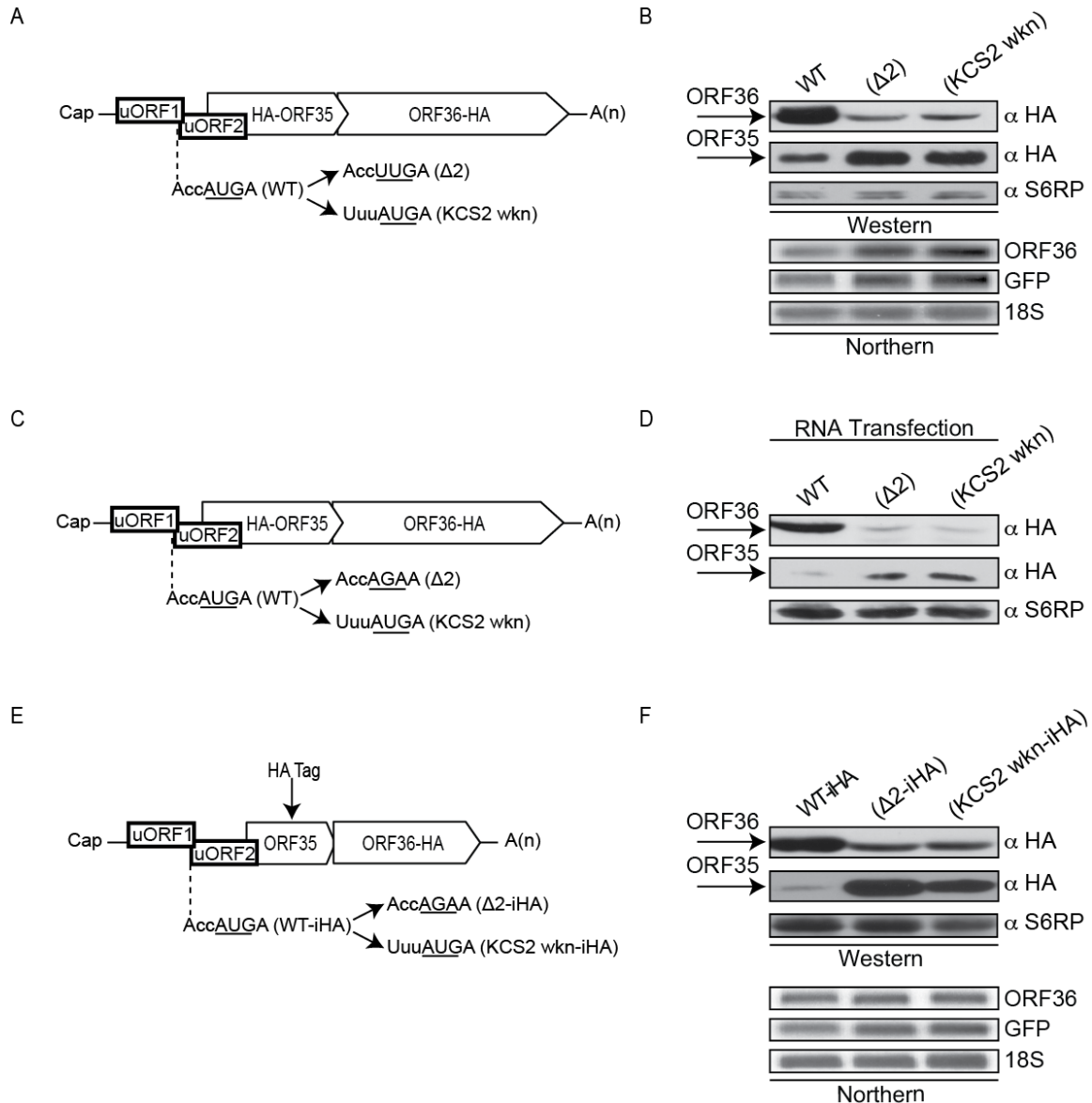


**Figure 3.4. uORF2 mediates translational control of ORF35 and ORF36.**

(A, C) Diagram indicating the nucleotide mutations used to disrupt ( $\Delta 2$ ) or weaken (KCS2 wkn) the context of the uORF2 start codon. (B, F) 293T cells were transfected with the indicated wild-type or mutant plasmids, and 24 h post-transfection protein lysates were resolved by SDS-PAGE and Western blotted with anti-HA antibodies. The ribosomal protein S6RP served as a loading control. RNA samples were examined by Northern blot analysis with a  $^{32}$ P-labeled ORF36 DNA probe. GFP served as a co-transfection control. 18S rRNA was used as a loading control. (D) 293T cells were transfected with *in vitro* transcribed capped and polyadenylated RNA to compare the wild type bicistronic mRNA with the uORF2 start codon mutants. Protein lysates were harvested 4 h post-transfection, resolved by SDS-PAGE and detected with anti-HA

antibodies. The ribosomal protein S6RP served as a loading control. (E) Schematic representation of the uORF2 mutations introduced into a construct with the native 5' UTR-ORF35-ORF36-HA with an HA tag positioned internally and in-frame with ORF35 (WT-iHA).

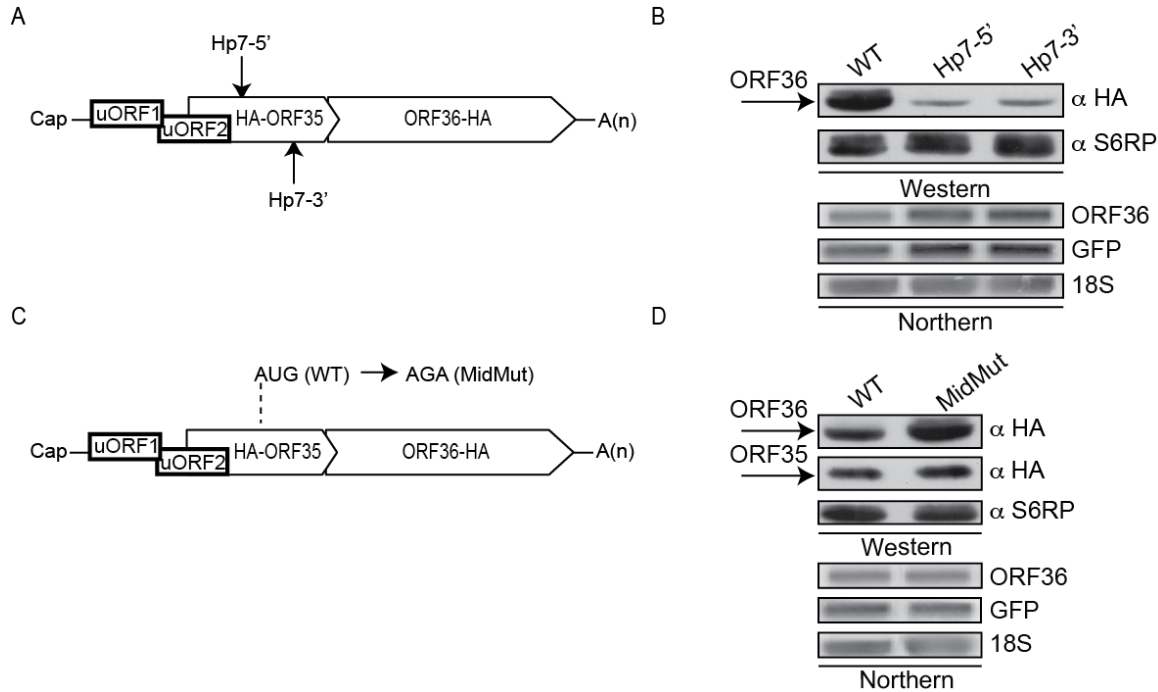


**Figure 3.5. The ORF36 start codon is accessed by linear scanning.** (A) Schematic of the wild-type 5' UTR-HA-ORF35-ORF36-HA construct showing the location of the Hp7 insertion into the 5' or 3'-proximal region of the ORF35 coding region. (B, D) 293T cells were transfected with the indicated wild-type or mutant plasmids, and 24 h post-transfection protein lysates were resolved by SDS-PAGE and Western blotted with anti-HA antibodies. The ribosomal protein S6RP served as a loading control. RNA samples were examined by Northern blot analysis with a  $^{32}$ P-labeled ORF36 DNA probe. GFP served as a co-transfection control. 18S rRNA was used as a loading control. (C) Schematic of the wild-type 5' UTR-HA-ORF35-ORF36-HA construct showing the location of the native AUG within the ORF35 codon region which has been mutated to AGA to generate the MidMut construct.

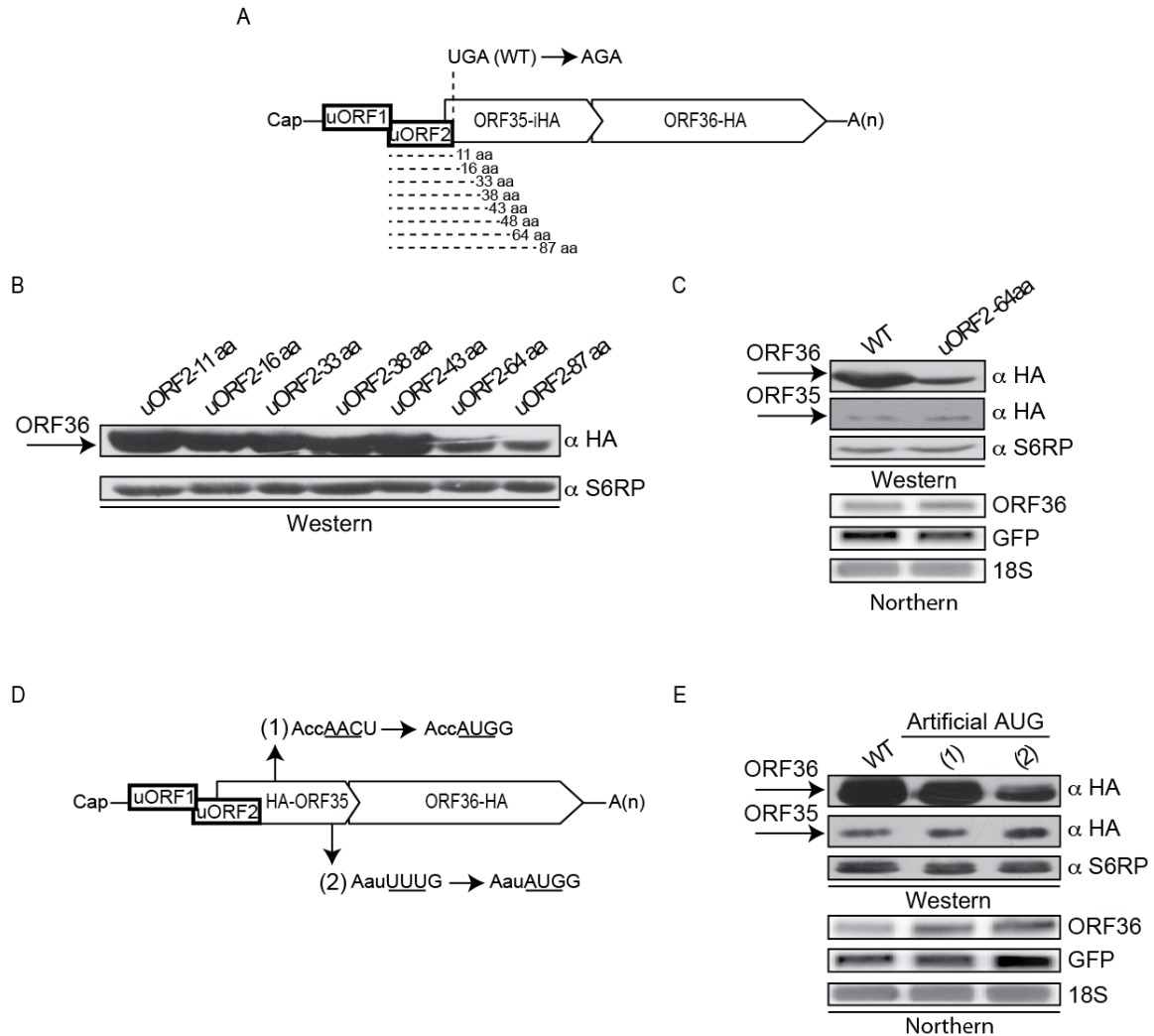


**Figure 3.6. uORF2 length and the uORF2-ORF36 intercistronic distance are important for expression of ORF36.** (A) Schematic of the artificial elongation of uORF2. The uORF2 stop codon and the six subsequent in-frame stop codons were mutated, artificially lengthening uORF2 incrementally from its native 11 to 16, 33, 38, 43, 48, 64 or 84 amino acids. (B) 293T cells were transfected with the indicated wild-type or mutant plasmids, and 24 h post-transfection protein lysates were resolved by SDS-PAGE and Western blotted with anti-HA antibodies. The ribosomal protein S6RP served as a loading control. (C) A side-by-side comparison of the WT and uORF2-64 aa construct by Western blot as described in (B) as well as by Northern blot analysis with a  $^{32}$ P-labeled ORF36 DNA probe. GFP served as a co-transfection control. 18S rRNA was used as a loading control. (D) Schematic of AUG insertions at two locations in the ORF35 coding region, placed out of frame with ORF36. All AUGs were designed to have the two dominant Kozak consensus sequence nucleotides (A at -3 and G at +4). (E) 293T cells were transfected with the indicated wild-type or mutant plasmids, and 24 h post-transfection protein lysates were resolved by SDS-PAGE and Western blotted with anti-HA antibodies. The ribosomal protein S6RP served as a loading control. RNA samples were examined by Northern blot analysis with a  $^{32}$ P-labeled ORF36 DNA probe. GFP served as a co-transfection control. 18S rRNA was used as a loading control.

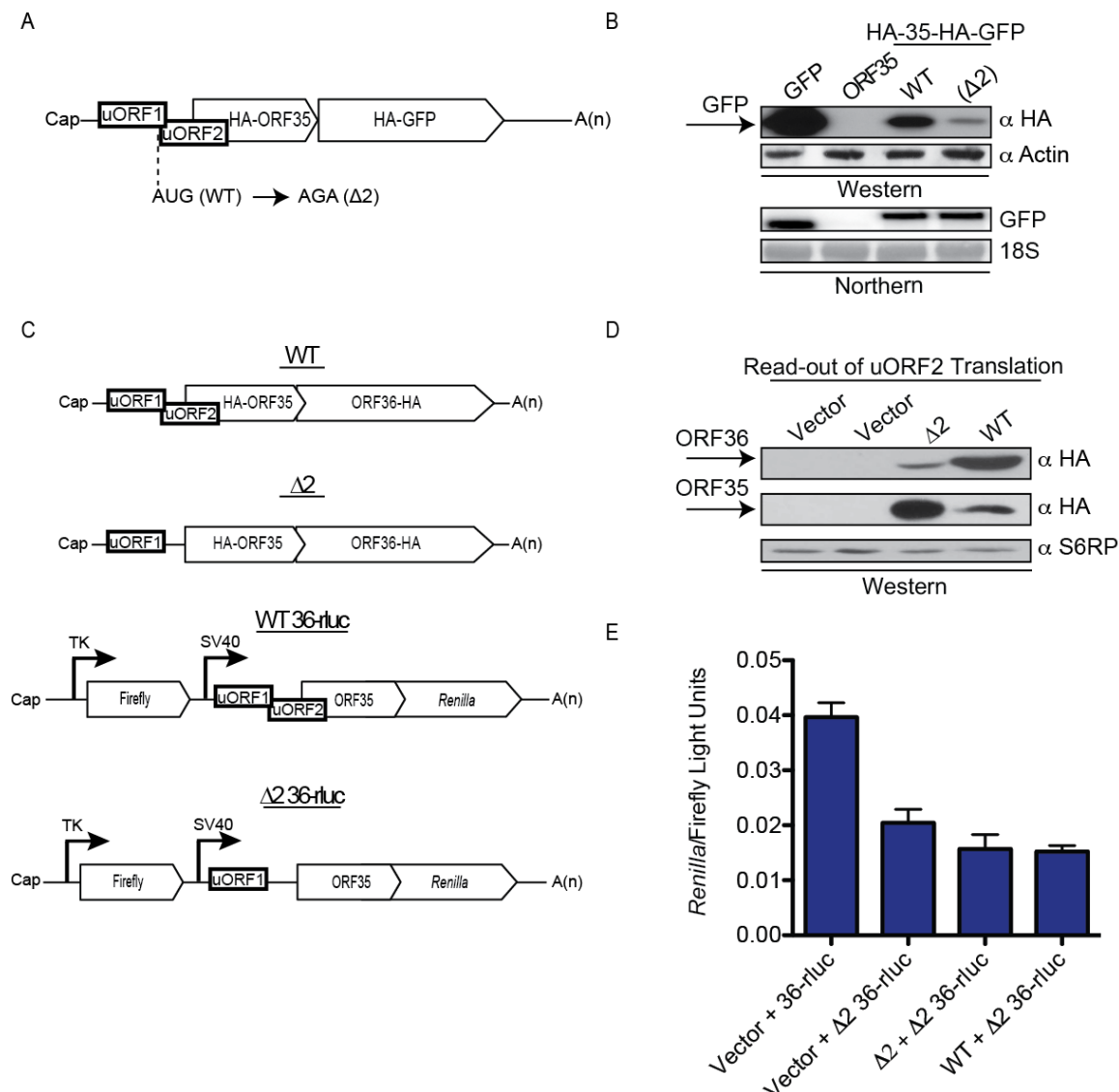


**Figure 3.7. No *cis*-acting element within ORF36 is required for reinitiation and uORF2 does not function when supplied *in trans*.** (A) Schematic of the bicistronic plasmid in which the ORF36 coding region was replaced with GFP. Because ORF36 partially overlaps with ORF35, this required truncating the C-terminus of ORF35. The uORF2 AUG mutation to AGA is also shown. (B) 293T cells were transfected with the indicated wild-type or mutant plasmids, and 24 h post-transfection protein lysates were resolved by SDS-PAGE and Western blotted with anti-HA antibodies. Actin served as a loading control. RNA samples were examined by Northern blot analysis with a  $^{32}\text{P}$ -labeled GFP DNA probe. 18S rRNA was used as a loading control. (C) Schematic of the bicistronic plasmid with either uORF2 intact (WT) or disrupted ( $\Delta 2$ ). Diagram of the dual luciferase construct that harbors firefly luciferase and the native 5'UTR of the ORF35-37 mRNA with the uORF1 and uORF2 (36-rluc) or uORF2 disrupted ( $\Delta 2$  36-

rluc) regulatory elements followed by the sequence of the first 352 nucleotides of ORF35 inserted upstream of *Renilla* luciferase, thus *Renilla* initiates from the native ORF36 start codon. (D) The indicated plasmid combinations were co-transfected into 293T cells with either 36-rluc (lane 1) or  $\Delta 2$  36-rluc (lanes 2-4). 24 h post-transfection protein lysates were resolved by SDS-PAGE and Western blotted with anti-HA antibodies. The ribosomal protein S6RP served as a loading control. (E) The indicated constructs were co-transfected into 293T cells and a dual luciferase assay was performed 24 h post-transfection to determine the relative levels of firefly and *Renilla* luciferase activity. The experiment was performed in triplicate; error bars represent the standard deviation between replicates.

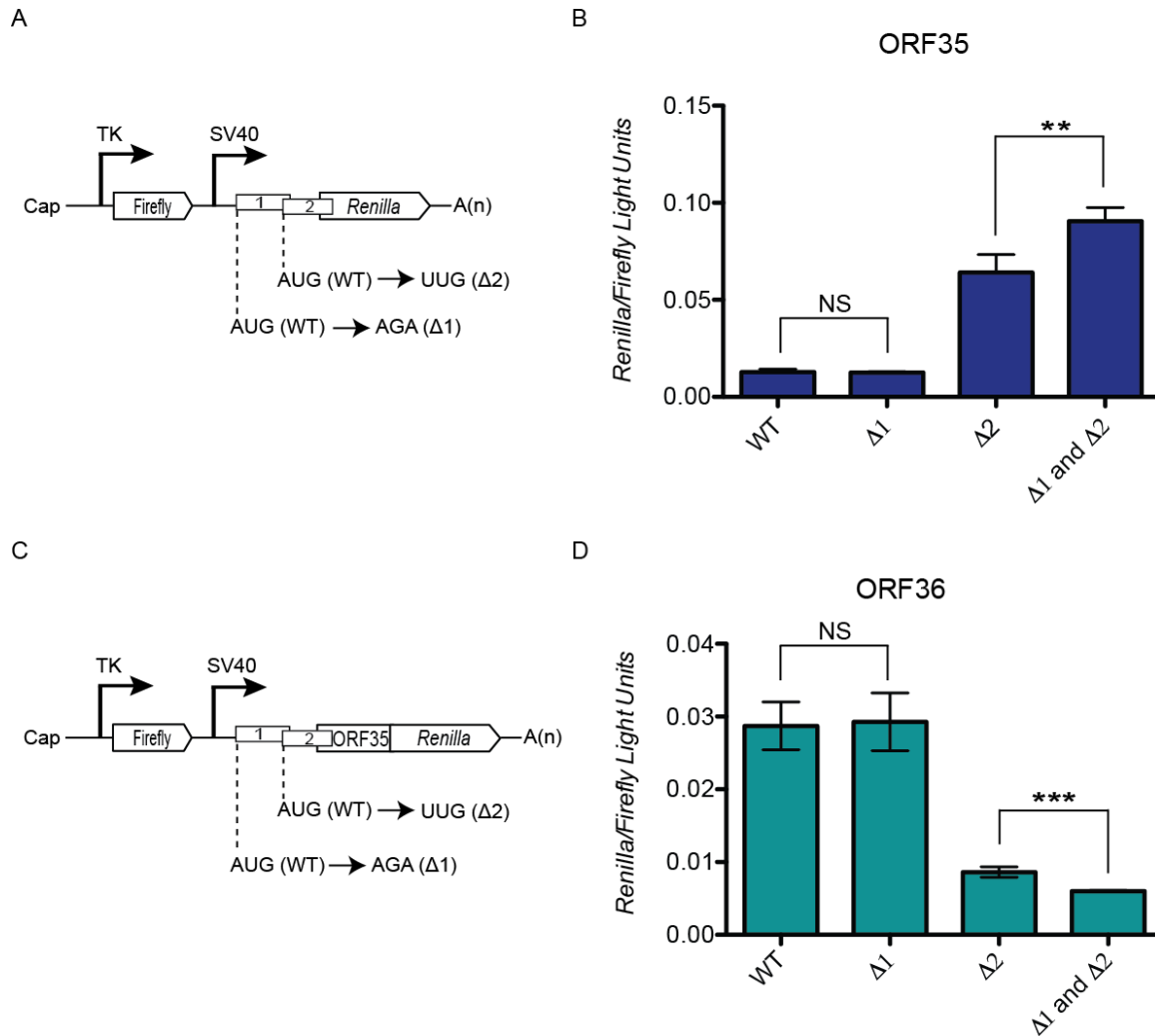


**Figure 3.8. uORF1 and uORF2 synergistically govern expression of ORF35 and ORF36.** (A, C) Diagram indicating the nucleotide mutations used to disrupt uORF1 ( $\Delta 1$ ) and/or uORF2 ( $\Delta 2$ ) within the ORF35 and ORF36 dual luciferase reporter constructs. (B, D) The indicated constructs were transfected into 293T cells and a dual luciferase assay was performed 24 h post-transfection to determine the relative levels of *Renilla* and firefly luciferase activity. The experiment was performed in triplicate; error bars represent the standard deviation between replicates. Statistical significance was evaluated with a two-tailed unpaired *t* test (N.S. > 0.05; \*\**P* < 0.001; \*\*\**P* < 0.0001).

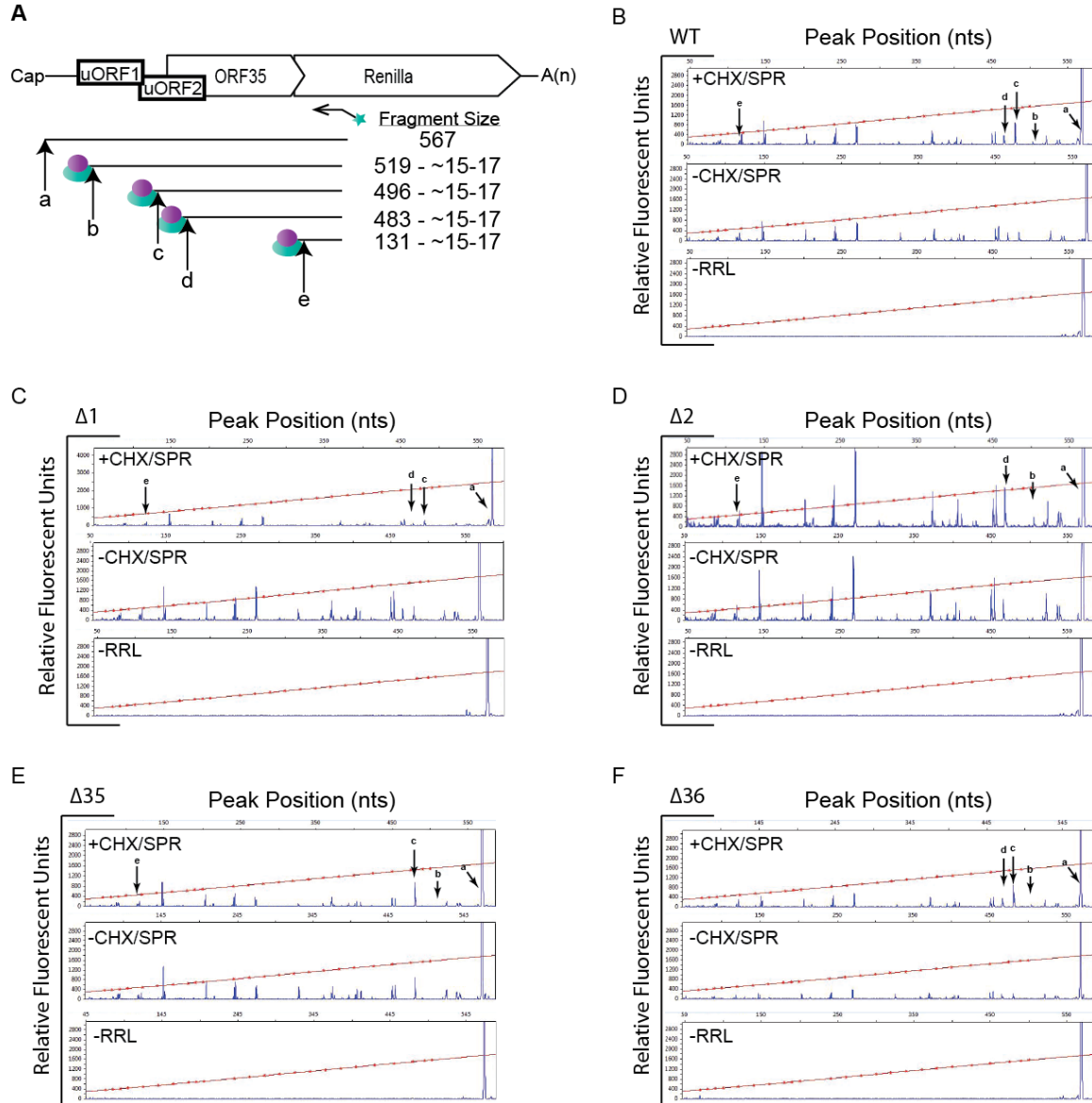


**Figure 3.9. Toeprinting analysis to detect ribosome occupancy on AUG<sub>uORF1</sub>, AUG<sub>uORF2</sub>, AUG<sub>ORF35</sub> and AUG<sub>ORF36</sub>.** (A) Schematic indicating the predicted fragment sizes produced by either the full-length reverse transcriptase reaction (a; FL) or ribosome occupancy at AUG<sub>uORF1</sub> (b), AUG<sub>uORF2</sub> (c), AUG<sub>ORF35</sub> (d), or AUG<sub>ORF36</sub> (e) in the fluorescent toeprinting assay. Electropherograms of WT (B), Δ1 (C), Δ2 (D), Δ35 (E), Δ36 (F). Toeprinting assays using the fluorescently labeled DNA oligonucleotide 5' – (FAM) – TTCTCGGCGTGCTTCTCGGAAT annealed 148 nucleotides downstream of the AUG<sub>ORF36</sub> (starred line with arrow). The toeprint reactions were performed in the presence of absence of cycloheximide and sparsomycin (+CHX/SPR or –CHX/SPR, respectively). A control reaction was carried out in the absence of rabbit reticulocyte lysates (-RRL). Primer extension products were suspended in a mix containing ROX 500 molecular weight marker and subjected to capillary electrophoresis. Size standard are shown for each electropherograms and sequence position is indicated in nucleotides. The

position of each AUG is shown with an arrow with the corresponding letter code from (A).

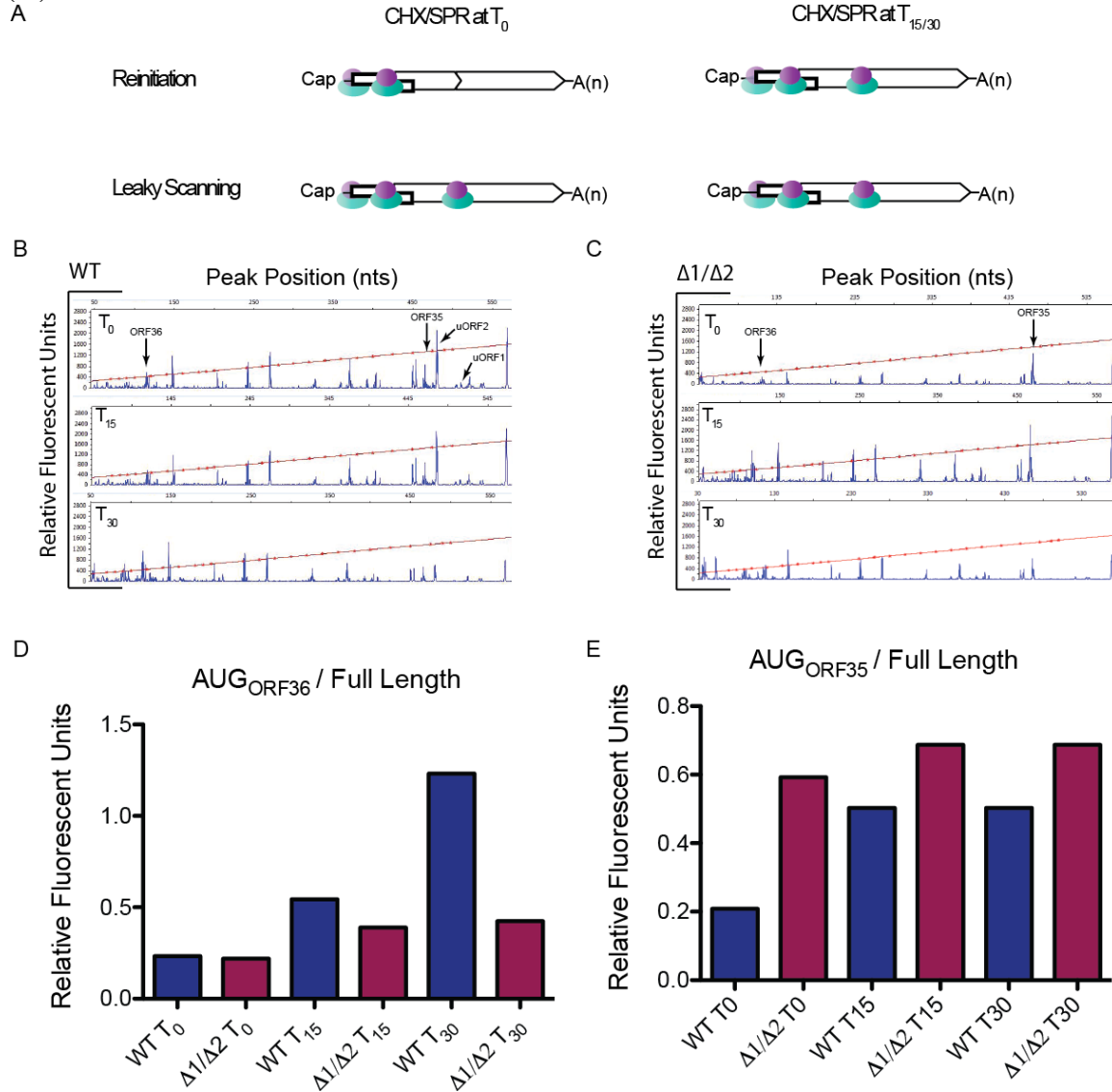


**Figure 3.10. Ribosomes terminating translation of uORF1 and uORF2 reinitiate at the ORF36 start codon.** (A) Diagram depicting how the toeprinting assay is used to differentiate leaky scanning from reinitiation events. Addition of cycloheximide (CHX) and sparsomycin (SPR) at  $T_0$  will arrest ribosomes at their first initiation event on the mRNA while adding it at steady-state translation ( $T_{15/30}$ ) will stall ribosomes both at their primary and subsequent reinitiation events. (B) Toeprinting electropherograms generated from fragment analysis data of WT (B) or  $\Delta 1/\Delta 2$  (C) either with CHX/SPR added at  $T_0$  or  $T_{15}$  or  $T_{30}$  using the 3' *Renilla* primer for the primer extension step. (D) The collected data were processed with PeakScanner v2.0 software and followed by graphical analysis representing the toeprint intensity calculated by the area under the curve of the ribosomal triplet relative to the full-length product.

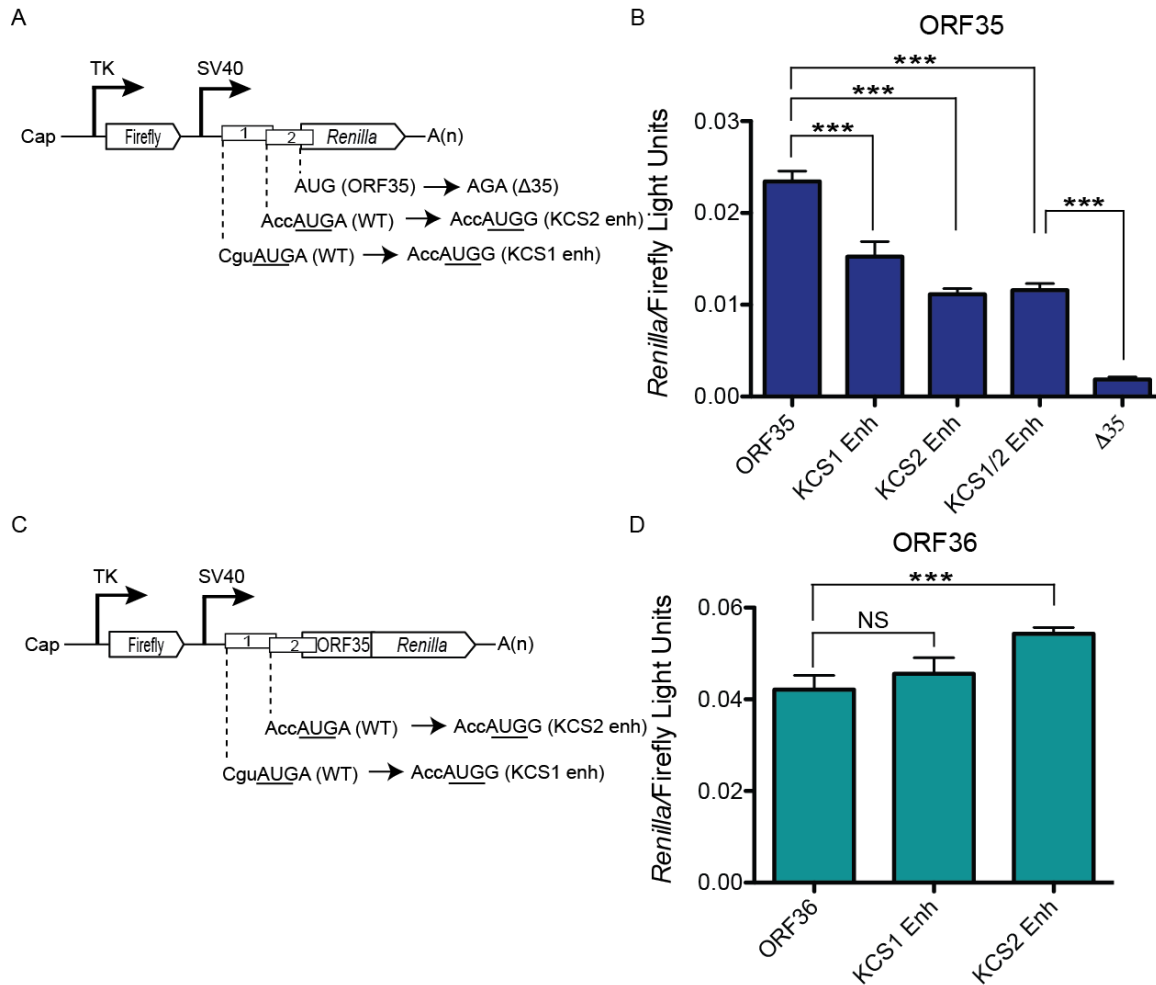


**Figure 3.11. The ORF35 start codon is at least partially accessed by leaky scanning past AUG<sub>uORF1</sub> and AUG<sub>uORF2</sub>.** (A, C) Diagram indicating the nucleotide mutations used to enhance the kozak consensus sequence of the uORF1 (KCS1 enh) and/or the uORF2 (KCS2 enh) start codons within the ORF35 and ORF36 dual luciferase reporter constructs. (B, D) The indicated constructs were transfected into 293T cells and a dual luciferase assay was performed 24 h post-transfection to determine the relative levels of firefly and *Renilla* luciferase activity. Experiment was performed in triplicate; error bars represent the standard deviation between replicates. Statistical significance was evaluated with a two-tailed unpaired *t* test (NS; non-significant, \*\*\**P* < 0.0001).

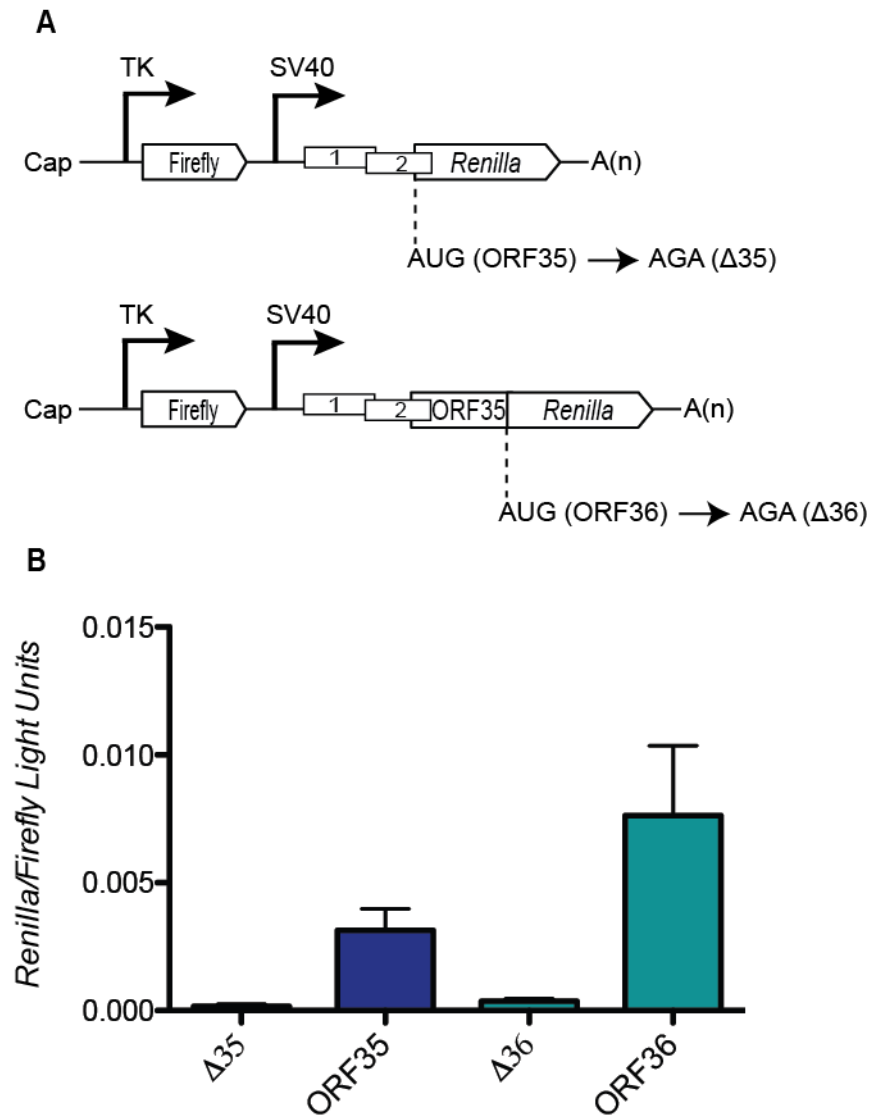


**Figure 3.12. Ratio of ORF35 and ORF36 translation initiation.** (A) Diagram of the dual luciferase constructs indicating the nucleotide mutations used to disrupt AUG<sub>ORF35</sub> (Δ35) and AUG<sub>ORF36</sub> (Δ36). (B) The indicated constructs were transfected into 293T cells and a dual luciferase assay was performed 24 h post-transfection to determine the relative levels of firefly and *Renilla* luciferase activity. The experiment was performed in triplicate; error bars represent the standard deviation between replicates.

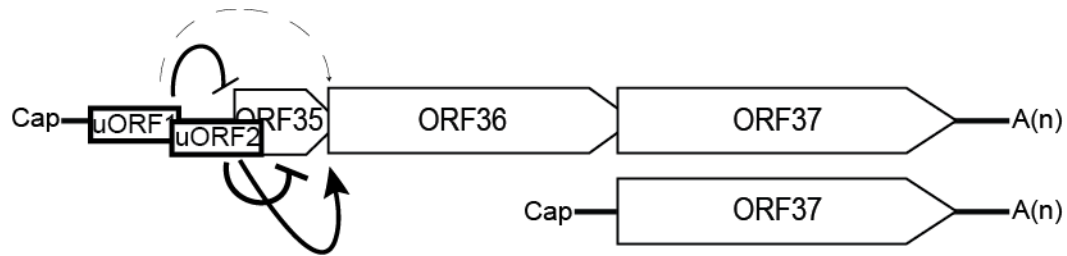
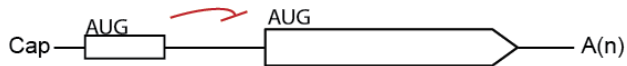
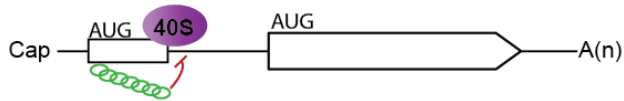


**Figure 3.13. Model of the mechanisms of translation initiation used to translate ORF35, ORF36 and ORF37.** ORF35 translation is repressed synergistically by uORF1 and uORF2. ORF36 is translated from a termination-reinitiation event after translation of both uORF1 and uORF2. ORF37 is translated from an ORF37-specific transcript generated from a promoter nested within the coding region of ORF36.

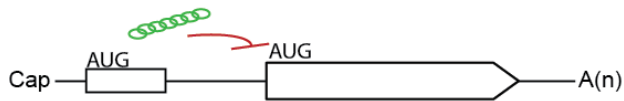
A - Sequence-independent uORF; inhibition of pORF



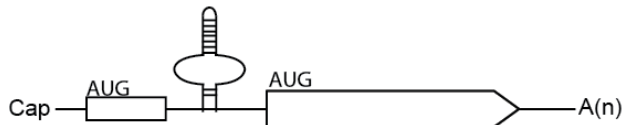
B - Sequence-dependent uORF; peptide functions in *cis* to stall ribosomes



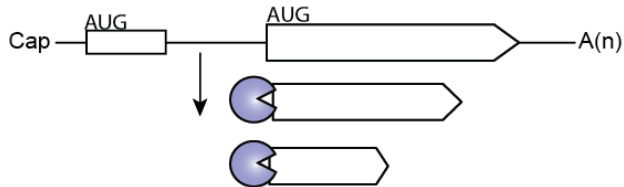
C - Sequence-dependent uORF; peptide functions in *trans* to inhibit pORF



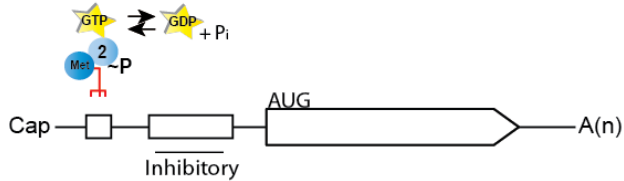
D - uORF and IRES or Shunt in the 5'UTR cooperate for pORF translation



E - Nonsense-mediated Decay



F - Bypass of inhibitory uORF during eIF2 $\alpha$  phosphorylation



G - uORF positioned to allow expression of a bicistronic mRNA



**Figure 3.14. The role of uORFs in translational control.** (A) The majority of uORFs function in *cis* in a sequence-independent manner to repress translation of the primary (p)ORF. (B) uORFs can function in a sequence-dependent in *cis* to dampen translation of the primary ORF by encoding rare codons or by forming a stable secondary structure causing the ribosome to stall (i.e. methionine synthase [130], UCP2 [131]). (C) Certain uORFs function in *trans* to repress translation of the pORF [132]. (D) uORFs and IRES or shunting elements or ribosome can cooperate to allow for expression of the pORF. (E) A certain percentages of uORFs confer susceptibility to nonsense-mediated decay. (F) Under conditions where eIF2 $\alpha$  is phosphorylated, terminating ribosomes require additional scanning time to re-acquire the ternary complex essential for reinitiation, thus bypassing a uORF which is normally inhibitory under non-stress conditions, thus favoring translation of the pORF [37-39]. (G) Positioning of the uORF2 (and non-critically uORF1) present on the KSHV ORF35-37 mRNA allows for the expression of a downstream cistron as the ribosomes that translate the uORF are able to reinitiate and translate ORF36 [111].

**Table 3.1. List of oligonucleotide primers used in chapter three.**

| <b>Primer Name</b>                        | <b>Primer Sequence 5'-3'</b>                                           |
|-------------------------------------------|------------------------------------------------------------------------|
| <b>ORF35 R <math>\Delta</math>96 R</b>    | GGGCCCCGCGGCCGCTTAATTCTCTTCCAGCGCATCTAGG                               |
| <b>HA-GFP F</b>                           | GGGCCCCGCTAGCATGGCTTACCCATACGATGTACCTGACTATGCGG<br>TGAGCAAGGGCGAGGAGCT |
| <b>HA-GFP R</b>                           | GGGCCCTCTAGATTACTTGTACAGCTCGTCCATGC                                    |
| <b>5'UTR-<i>Rluc</i> F</b>                | TATAGGCTAGCCACCGTCCAACCTACGGGCGACTATCT                                 |
| <b>5'UTR-uORF2 -<i>Rluc</i> R</b>         | CACCTTGGAAGCCATTTTAGAGTTGGTTGAGTCCATTCTCC                              |
| <b>5'UTR ORF35<math>\Delta</math>98 R</b> | CACCTTGGAAGCCATTCTCTTCCAGCGCATCTAGG                                    |
| <b>psiCHECK-2 iPCR F</b>                  | ATGGCTTCCAAGGTGTACGAC                                                  |
| <b>psiCHECK-2 iPCR R</b>                  | GGTGGCTAGCCTATAGTGAGTCGTATT                                            |

\*F: Forward

\*R: Reverse

## Chapter Four: Investigation of the role of uORFs in the context of KSHV Infection

### Background

KSHV genetics have been historically beset by limitations with mammalian cell lines. That is, while many adherent lines in culture can be latently infected following exposure to KSHV virions, few cell lines can be lytically reactivated without the use of pleiotropic chemicals [133,134]. Therefore analysis of mutants in the context of the native viral life cycle of immediately early, delayed early and late viral gene expression has been severely restricted [54,55,135]. Thus, the experimental approach in chapters two and three relied on transient transfection or *in vitro* systems to investigate the role of uORF1 and uORF2 in the translational control of the ORF35-37 mRNA.

A critical advance in the field of KSHV genetics was the generation of the iSLK cell lines which harbor a doxycycline-inducible RTA transgene, which when expressed ectopically can initiate the lytic reactivation of KSHV [135]. Further headway was made when in 2012, Brulois et al. reported the development of a KSHV genetic system using bacterial artificial chromosome (BAC) technology capable of modifying the viral genome [136]. This novel KSHV BAC clone, termed BAC16, is derived from the rKSHV.219 virus that originates from the KSHV- and EBV-infected JSC1 primary effusion lymphoma cells. Thus, a BAC16 mutant can be transfected into the iSLK cell line and, following lytic induction, be analyzed in the context of the viral life cycle.

In this chapter, we use the combined KSHV genetic system with the RTA-inducible SLK cell line to confirm that uORF2 regulates expression of ORF36-encoded viral protein kinase. Furthermore, we engineered a FLAG epitope-tagged ORF35, allowing for the detection of this protein for the first time during infection. Surprisingly, we report that in the context of viral replication the introduction of a single nucleotide change to disrupt uORF2 in the KSHV genome leads to a dramatic defect in the production of infectious particles. Lastly, we investigate whether the phosphorylation state of eIF2 $\alpha$ , a variable critical for reinitiation efficiency, is modulated during KSHV infection and whether the phosphorylation status governs expression of ORF35 or ORF36 in a uORF-dependent manner. Analysis of homologous genetic loci from additional  $\gamma$ -herpesviruses similarly revealed the presence of dual short upstream ORFs (uORFs), suggesting this may be a conserved mechanism of translation initiation among these viruses.

## Results

### Disruption of uORF2 alters ORF36 expression during lytic infection

To confirm that uORF2 regulates ORF36 expression during lytic KSHV infection, we engineered a uORF2 point mutant (BAC16- $\Delta$ 2; ATG $\rightarrow$ TTG) and a revertant mutant rescue (BAC16- $\Delta$ 2-MR; TTG $\rightarrow$ ATG) within the recently described KSHV BAC16 (**Figure 4.1A and B**) [136]. BAC16-WT, BAC16- $\Delta$ 2 and BAC16- $\Delta$ 2-MR were transfected into iSLK-PURO cells bearing a doxycycline-inducible RTA expression system to enable lytic reactivation (**Figure 4.1C**) [135]. Immunoblot analysis using polyclonal anti-sera specific for ORF36 revealed that while ORF36 was readily detectable at 48 hours post-lytic reactivation in cells infected with WT or the mutant rescue virus, deletion of the uORF2 start codon severely compromised ORF36 expression (**Figure 4.1D**). In contrast, the uORF2 mutation had no effect on the levels of the KSHV latent protein LANA or the lytic protein ORF57, confirming its specificity for ORF36 (**Figure 4.1D**). Thus, uORF2 plays a critical role in enabling expression of the ORF36-encoded viral protein kinase during lytic KSHV infection.

No functional antibodies exist to detect ORF35 during KSHV infection. Therefore to confirm ORF35 expression during lytic KSHV infection, we engineered an in frame FLAG epitope tag within the coding region of ORF35 at nucleotide position 55796 – 820 in the BAC16-WT (**Figure 4.2A and B**). A N- or C-terminal tag could not be inserted because this would either disrupt the coding region of uORF2 or the N-terminus of ORF36, respectively. BAC16-WT-iFLAG was transfected into iSLK-PURO cells bearing a doxycycline-inducible RTA expression system to enable lytic reactivation, as described above [135]. Immunoblot analysis using polyclonal anti-sera specific for FLAG revealed that ORF35 was readily detectable at 96 hours post-lytic reactivation (**Figure 4.3C**). This result marks the first time the viral ORF35 protein has been detected in context of lytic infection.

### Mutation of uORF2 leads to a defect in the production of infectious viral particles

Next, in collaboration with Kevin Brulois and Jae Jung at USC, we investigated whether the observed drop in ORF36 expression influenced the production of infectious viral progeny. At 48 h post-induction of the lytic cycle, cell-free supernatants from the BAC16-WT, BAC16- $\Delta$ 2 and BAC16- $\Delta$ 2-MR iSLK-PURO cell lines were transferred to 293A cells and infectious virus was either visualized by GFP fluorescence (**Figure 4.3A**) or quantified by flow cytometry analysis of GFP expression at 24 h post-infection (**Figure 4.3B**). Perhaps due to differences in episome copy number, the number of infectious virus particles was higher in the BAC16- $\Delta$ 2-MR compared to BAC16-WT (**Figure 4.3B**). However, there was a 4.92-fold decrease in infectious viral particles in the BAC16- $\Delta$ 2 compared to BAC16-WT and 36.65-fold drop in the BAC16- $\Delta$ 2 compared to the BAC16- $\Delta$ 2-MR. Thus, uORF2 plays a critical role in enabling expression of the ORF36-encoded viral protein kinase during KSHV lytic infection, which is important for the production of infectious virus particles.

### Analysis of the phosphorylation status of eIF2 $\alpha$ during KSHV infection

The translation of the primary ORF on certain mRNAs that harbor uORFs is enhanced when eIF2 $\alpha$  is phosphorylated (e.g. *ATF4*, *CHOP*, *GCN4*). It is therefore tempting to speculate whether the reinitiation efficiency of uORFs located on the ORF35-

37 mRNA may be governed by the eIF2 $\alpha$  phosphorylation status of the infected cell [137-140]. Notably, PKR, one of the protein kinases responsible for phosphorylation of eIF2 $\alpha$ , is activated by dsRNA produced during viral infection. While a single report on the status of the KSHV-host eIF2 $\alpha$  arms race found that the latently-expressed Interferon Regulatory Factor 2 (vIRF-2) inhibits PKR [141], whether KSHV modulates eIF2 $\alpha$  phosphorylation during lytic replication has not been established. Therefore we asked whether the status of eIF2 $\alpha$  phosphorylation is modified in lytically reactivated iSLK.219 cells, a KSHV-positive cell line [135]. Indeed, we found that beginning at 60 h post-induction, eIF2 $\alpha$  was robustly phosphorylated (**Figure 4.4A**). Importantly, no change in total eIF2 $\alpha$  was observed and the change in eIF2 $\alpha$  phosphorylation status was not observed in the matched KSHV-negative iSLK.H cell line (**Figure 4.4A and B**). However, whether this result holds true in physiologically relevant cell lines and whether viral RNAs or proteins modulate this response remains to be explored [142].

### Investigation of uORF mutants under conditions of eIF2 $\alpha$ phosphorylation

To address the possibility that eIF2 $\alpha$  phosphorylation may regulate translation of ORF35 and ORF36 in a uORF-dependent fashion, we first sought to establish a cell-culture assay. Treatment of cells with chemicals, such as thapsigargin (TG) can trigger eIF2 $\alpha$  phosphorylation and a general inhibition of host cell protein synthesis [137,143,144]. We therefore sought to confirm that exposure of 293T cells to TG indeed triggers eIF2 $\alpha$  phosphorylation in cells transfected with our constructs of interest (**Figure 4.5 and 4.6A and B**). Indeed in cells transfected with our either 5'UTR-uORF2-*Rluc* or 5'UTR-ORF35 $\Delta$ 98-*Rluc* bicistronic constructs, which serve as reporter for ORF35 or ORF36, respectively, treatment with TG triggered eIF2 $\alpha$  phosphorylation (**Figure 4.5 and 4.6A and B**). Furthermore, treatment with TG resulted in a decrease in translation of the control firefly luciferase reporter, as predicted (**Figure 4.5C and 4.6C**). Thus, we have established an assay to investigate whether translation of ORF35 or ORF36 is modulated by eIF2 $\alpha$  phosphorylation in a uORF-dependent manner.

An additional 9-codon short ORF, dubbed uORF3, is located within the uORF1/2-ORF36 intercistronic region, thus one could speculate that depending on the phosphorylation of eIF2 $\alpha$ , terminating ribosomes could either reinitiate at this internal AUG or bypass it in favor of the more 5'-distal AUG<sub>ORF36</sub>, reminiscent of uORF1 and uORF4 on the *S. cerevisiae* *GCN4* locus (see Introduction). The main hypothesis driving this series of experiments was that under conditions of eIF2 $\alpha$  phosphorylation, ORF36 translation would perhaps be enhanced in that the reinitiating ribosomes would be more likely to bypass AUG<sub>uORF3</sub> in favor of AUG<sub>ORF36</sub>. We were uncertain whether eIF2 $\alpha$  phosphorylation would modulate initiation at AUG<sub>ORF35</sub> as our data indicates the ORF35 start codon is not accessed by reinitiation. Single ( $\Delta$ 1 &  $\Delta$ 2) or double ( $\Delta$ 1 &  $\Delta$ 2) on the ORF35 reporter construct and single ( $\Delta$ 1,  $\Delta$ 2, &  $\Delta$ 3), double ( $\Delta$ 1 &  $\Delta$ 2,  $\Delta$ 1 &  $\Delta$ 3,  $\Delta$ 2 &  $\Delta$ 3) or triple ( $\Delta$ 1 &  $\Delta$ 2 &  $\Delta$ 3) mutants on the ORF36 reporter construct were analyzed by dual luciferase assay. However under the conditions of this assay, no significant uORF-dependent translational control on ORF35 (**Figure 4.5C-E**) or ORF36 (**Figure 4.6C-E**) in DMSO vs. TG treated cells was observed. One notable observation is that while the WT or  $\Delta$ 1 ORF35 *Rluc*/Fluc ratio is not abrogated during eIF2 $\alpha$  phosphorylation, in the  $\Delta$ 2 and  $\Delta$ 1 &  $\Delta$ 2 mutant, expression of ORF35 is decreased (**Figure 4.5E**). Thus, while no clear effect of eIF2 $\alpha$  phosphorylation was observed on ORF36 expression, these data

suggest a model whereby ORF35 expression is maintained by the presence of uORF2 under conditions of cellular stress, such as those that arise during late lytic infection.

### **Conservation of uORFs within related $\gamma$ -herpesviruses**

We examined whether the loci analogous to KSHV *ORF35-37* in several additional  $\gamma$ -herpesviruses also possessed uORFs within their 5' UTRs. Indeed, we identified two 6-12 codon uORFs within the predicted 5' UTR of the locus in Epstein Barr virus (EBV), herpesvirus saimiri (HaSV-2) and ateline herpesvirus 3 (AtHV-3), and one 11 codon uORF in good context within the 5' UTR of the rhesus rhadinovirus (RRV) locus (**Figure 4.7A and B**). The fact that the uORF positioning but not the coding sequence is conserved supports the hypothesis that their regulatory contribution relies on their ability to engage translation complexes, rather than the actual peptide produced. Interestingly, in all cases where two uORFs are present, the first uORF is within a weaker Kozak context than the second uORF, which overlaps the start codon of each ORF35 homolog (EBV BGLF3.5, SaHV-2 ORF35, AtHV-3 ORF35 and RRV ORF35). Furthermore, 8 of the 9 ORF35 homologs examined contain < 2 internal methionine codons, as would be predicted if a termination-reinitiation mechanism was used to translate the downstream gene (**Table 4.1**). Thus, the conservation of uORFs at this genetic locus suggests that using uORFs to enable expression of a 3'-proximal gene may be a conserved strategy for translational control among these viruses. However, whether these loci indeed encode a functional polycistronic mRNA and are regulated by a similar uORF-based mechanism remains to be experimentally verified.

## Discussion

In this chapter, we used the newly developed KSHV genetic system to demonstrate that uORF2 is critical for the expression of the ORF36-encoded viral protein kinase during KSHV lytic infection. Unexpectedly, the single nucleotide mutation engineered to disrupt the uORF2 start codon led to a striking defect in the production of infectious viral particles. Whether this defect is due to a loss in ORF36 expression or the over-expression of ORF35 is unknown. Due to their translation from a single mRNA, siRNAs can not be used to specifically target ORF35 or ORF36, and thus it remains a challenge to dissect out the precise cause of the  $\Delta 2$ -specific drop in infectious titer. One approach would be to over-express ORF36 in the context of the BAC16- $\Delta 2$  virus and ask whether this rescues the defect in viral titer. Alternatively, it would be interesting whether over expression of ORF35 in the context of WT KSHV infection independently compromises the production of infectious viral particles. However, the possibility exists that the uORFs serve to confer a tight level of regulation to ensure that neither ORF35 nor ORF36 is synthesized at deleterious levels during infection. In EBV, a viral mutant that over-produces BGLF4 (the ORF36 homolog) exhibited a defect in viral replication [91].

In the absence of independent transcriptional regulation, we considered the possibility that ORF35 and ORF36 expression levels are subject to translational control. The extent of eIF2 $\alpha$  phosphorylation is known to regulate the efficiency of reinitiation by limiting ternary complex formation and allowing for the bypass of certain AUGs. Therefore, we first investigated whether phosphorylation of eIF2 $\alpha$  is modulated during KSHV lytic infection. We observed a dramatic enhancement of eIF2 $\alpha$  phosphorylation in the KSHV-positive iSLK.219 cell line that occurred with approximate delayed-early kinetics. We thus developed an ORF36 reporter assay under conditions of either physiological or enhanced eIF2 $\alpha$  phosphorylation and asked whether stress modulates the levels of ORF36. However, no straightforward uORF-dependent phenotype was observed on ORF36 expression under changing conditions of eIF2 $\alpha$  phosphorylation.

Intriguingly, while ORF35 levels were not altered by an increase in eIF2 $\alpha$  phosphorylation, expression of ORF35 was specifically abrogated by stress when uORF2 was disrupted. This result is reminiscent, yet distinct, of what is observed for both the mammalian *CHOP* mRNA and the EBOV polymerase (L) mRNA where phosphorylation of eIF2 $\alpha$  facilitates ribosomal bypass of the inhibitory upstream ORF to enhance translation of the primary ORF [140,145]. However, we lack an understanding of the biochemical basis by which enhanced eIF2 $\alpha$  phosphorylation may allow for the bypass of the inhibitory uORF, presenting a challenge for further investigation [140]. Therefore, whether eIF2 $\alpha$  phosphorylation indeed modulates uORF and/or ORF35 recognition remains ambiguous. Determining if and how this non-canonical mechanism of translational control influences the KSHV lifecycle will be an important future endeavor.

## Materials and Methods

### Cells, transfections and drug treatment

The iSLK-PURO, iSLK.H and iSLK.219 cell lines were maintained in DMEM supplemented with 10% FBS, penicillin (100 U/ml, Gibco) and streptomycin (100 µg/ml, Gibco) [135,136]. To induce lytic reactivation of KSHV, iSLK-PURO cells were treated with doxycycline (1 µg/ml, BD Biosciences) and sodium butyrate (1 mM, Sigma). Human embryonic kidney 293T and 293A cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) (Gibco). eIF2 $\alpha$  phosphorylation in 293T cells was stimulated using thapsigargin treatment (TG, Sigma Cat# T9033).

For BAC transfections and reconstitution, ~70% confluent iSLK-PURO cells were grown in a 24-well plate followed by transfection with 500 ng of BAC DNA via FuGENE 6 (Promega), after 6 h, a further 500 ng BAC DNA was transfected with Effectene, following the manufacturer's protocols and subsequently selected with 800 µg/ml hygromycin B to establish a pure population. iSLK-PURO-BAC16 cells were then induced with doxycycline (1 µg/mL) and sodium butyrate (1 mM) to enter the lytic cycle of KSHV replication.

### Dual Luciferase assays

Luciferase activities were determined using the dual-luciferase assay system (Promega Cat#: E1960) and a bench-top luminometer according to the manufacturer's protocol. The *Rluc/Fluc* value represents at least three independent experiments with triplicate samples in each transfection. Error bars represent the standard deviation between replicates.

### BAC mutagenesis and DNA isolation

The KSHV BAC16 was modified as described previously [136] using a two-step scarless red recombination system [146]. Briefly, BAC16 was introduced in GS1783 *E. coli* strain by electroporation (0.1 cm cuvette, 1.8 kV, 200 $\Omega$  25µF). A linear DNA fragment encompassing a kanamycin resistance expression cassette, an I-SceI restriction site and flanking sequence derived from KSHV genonic DNA was generated by PCR and subsequently electroporated into GS1783 *E. coli* harboring BAC16 and transiently expressing *gam*, *bet* and *exo*. Integration of the Kan<sup>R</sup>/I-SceI cassette was verified by PCR and restriction enzyme digestion of the purified BAC16 DNA by pulse field gel electrophoresis using the MidRange I PFG marker (New England BioLabs; Cat #: N3551S) and 1 kb plus DNA ladder (Life Technologies Cat #: 10787-026). The second recombination event between the duplicated sequences resulted in the loss Kan<sup>R</sup>/I-SceI cassette and the seamless recirculation of the BAC16 DNA, yielding kanamycin-sensitive colonies that were screened by replica plating. BAC16 DNA was purified from chloramphenicol-resistant colonies using the NucleoBond<sup>®</sup> 100 (Machery-Nagel) as per the manufacturer's instructions. Primers used for BAC mutagenesis can be found in Table 4.2.

### Quantification of Infectious Virus

Cell-free supernatant was harvested from iSLK-PURO cells at 48 h post-induction and

used to inoculate  $2 \times 10^5$  293A cells seeded in a 12-well plate. Immediately after inoculation, the plate was centrifuged at 2000 g for 45 min at 30°C and incubated at 37°C in 5% CO<sub>2</sub> for 1 h, subsequently the inoculum was removed and replaced with fresh media. After 24 h, cells were collected and washed once with cold phosphate-buffered saline. The percentage of GFP-positive cells was detected using a FACS Cantoll (BD Bioscience, San Jose, CA). Infectious units (IU) are expressed as the number of GFP-positive cells in each well at the time of analysis. The experiment was performed in triplicate; error bars represent the standard deviation between replicates.

### **Western blots**

Protein lysates were prepared in RIPA buffer [50 mM Tris-HCl (pH8.0), 150 mM NaCl, 1% (v/v) Nonidet P-40, 0.5% (w/v) sodium deoxycholate, 0.1% (w/v) sodium dodecyl sulfate (SDS)] containing protease inhibitors (Roche), and quantified by Bradford assay. Protein lysates of phosphorylated proteins were prepared in cell lysis buffer [10 mM Tris-HCl (pH7.4), 100 mM NaCl, 1 mM Sodium Fluoride, 20 mM Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub>, 1% (v/v) Triton-X 100, 0.5% (w/v) sodium deoxycholate, 10% glycerol, 0.1% (w/v) sodium dodecyl sulfate (SDS)] containing protease inhibitors (Roche). Equivalent quantities of each sample were resolved by SDS-PAGE, transferred to a polyvinylidene difluoride membrane and incubated with the following primary antibodies: rabbit polyclonal ORF36 (1:5000, kindly provided by Y. Izumiya [82]), rabbit polyclonal LANA #6 (1:1000) [166], [89]), rabbit polyclonal eIF2 $\alpha$  (1:1000 Cell Signaling cat #9722), rabbit polyclonal phospho-eIF2 $\alpha$  (Ser51) (1:1000 Cell Signaling cat #9721), rabbit polyclonal ORF57 (1:5000, kindly provided by Z. Zheng [129], rabbit polyclonal FLAG (Sigma Cat #: F7425) or mouse monoclonal S6RP (1:1000, Cell Signaling) followed by incubation with HRP-conjugated goat anti-mouse or goat anti-rabbit secondary antibodies (1:5000 dilution) (Southern Biotechnology Associates). Phospho-specific antibodies were diluted in 5% (w/v) BSA, 1X TBS and 0.1% Tween-20.

### **Sequence Alignments**

The uORF1 and uORF2 alignments were generated from data obtained from the NIAID Virus Pathogen Database and Analysis Resource (ViPR) online through the web site at <http://www.viprbrc.org>.

### **Graph Design and Statistical Analysis**

Graphs were designed using either excel or the GraphPad Prism Software version 5 all and statistical analysis performed using the GraphPad Prism Software version 5. Data were analyzed for statistical significance using the Student's two-tailed unpaired *t*-test.

## Figures

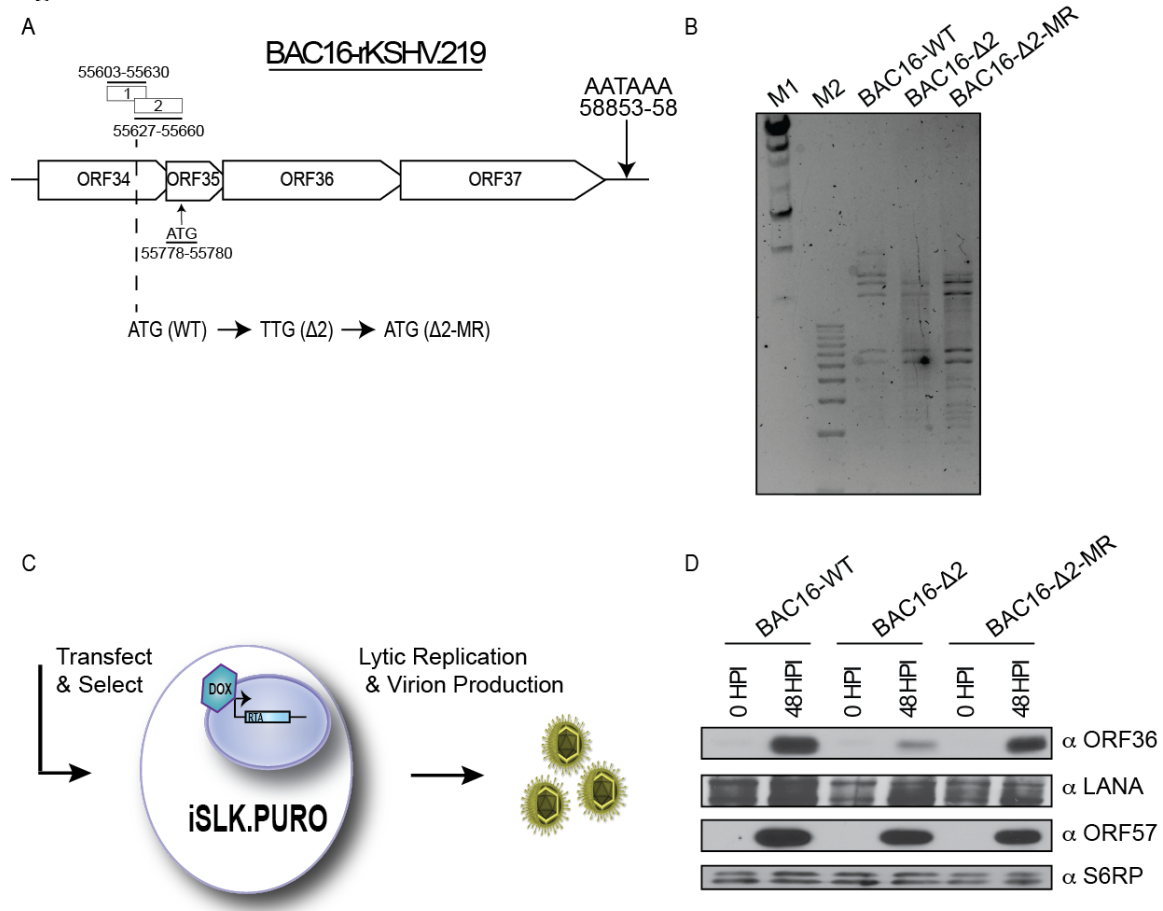


**Figure 4.1. Disruption of uORF2 alters ORF36 expression during lytic infection.**

(A) Schematic of the KSHV BAC16 indicating the nucleotide mutations used to disrupt (Δ2) or restore (Δ2-MR) the uORF2 start codon. (B) BAC16 WT, uORF2 mutant (BAC16-Δ2), or mutant rescue (BAC16-Δ2-MR) DNA was isolated from GS1783 *Escherichia coli*, digested with *NheI* and subjected to pulse-field gel electrophoresis. M1: MidRange I PFG marker (NEB) and M2: 1 Kb marker (Biorad). The BAC16 WT, Δ2 and Δ2-MR were generated by Kevin Brulois. (C) Schematic of the protocol to introduce KSHV BAC16 into the iSLK-PURO cells which bears a doxycycline-inducible RTA expression system to enable lytic reactivation [135]. (D) iSLK-PURO cells stably harboring the WT KSHV BAC16, a uORF2 mutant BAC16 (BAC16-Δ2), or a mutant rescue BAC16 (BAC16-Δ2-MR) were either untreated or lytically reactivated for 48 h. Protein lysates were Western blotted with antibodies against ORF36, the viral latent protein LANA and a viral lytic protein ORF57. The ribosomal protein S6RP served as a loading control.

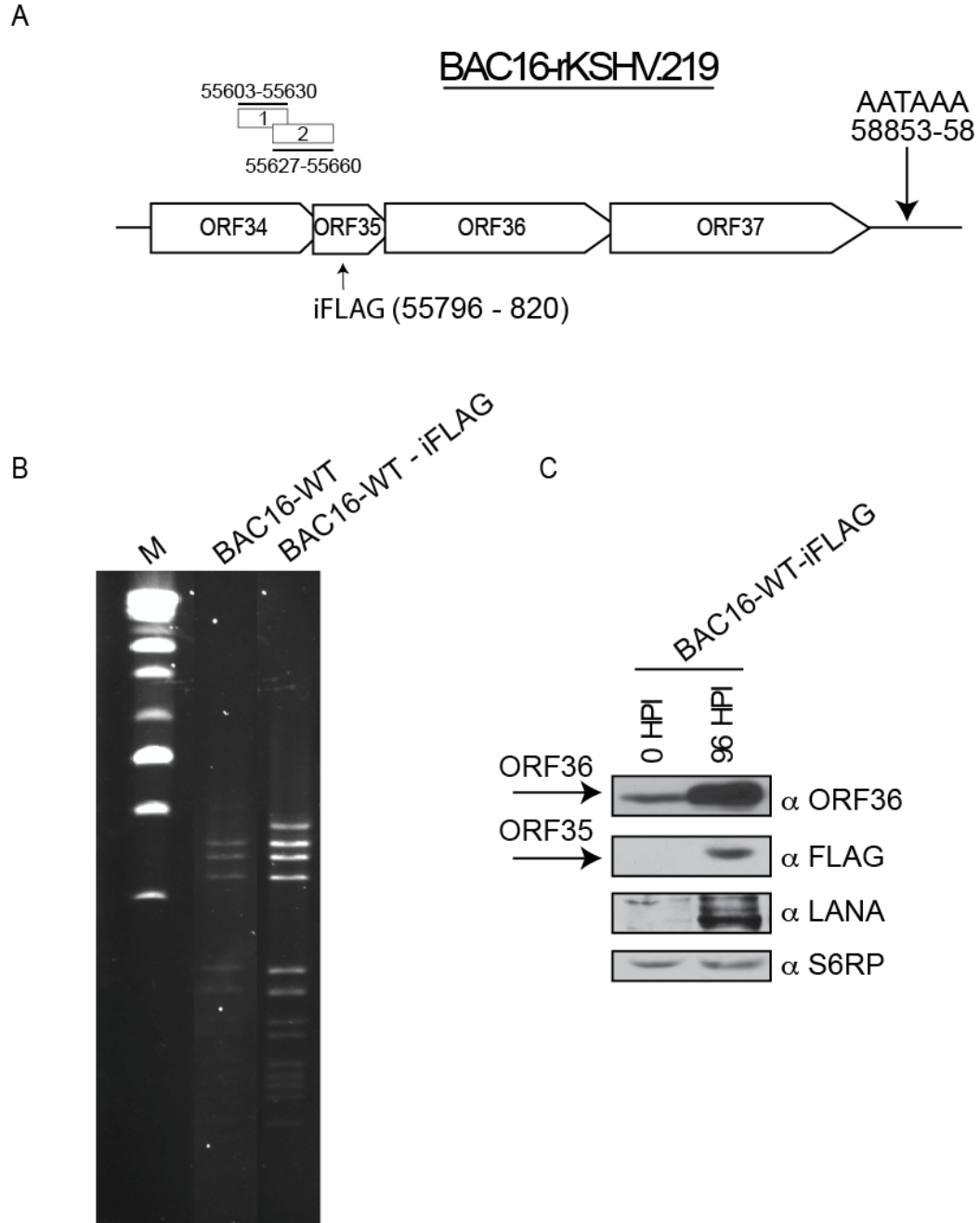


**Figure 4.2. Construction of an FLAG epitope-tagged ORF35 BAC16 allows for direct detection of ORF35.** (A) Schematic of the KSHV BAC16 indicating the position of the internal FLAG epitope tag (iFLAG). (B) BAC16 WT or WT-iLFLAG DNA was isolated from GS1783 *Escherichia coli*, digested with *NheI* and subjected to pulse-field gel electrophoresis. M, MidRange I PFG marker (New England Biolabs). (C) iSLK-PURO cells stably harboring the WT KSHV BAC16 with an internally tagged FLAG epitope-tagged ORF35 were either untreated or lytically reactivated for 96 h. Protein lysates were Western blotted with antibodies against ORF36, FLAG, or the viral latent protein LANA. The ribosomal protein S6RP served as a loading control.

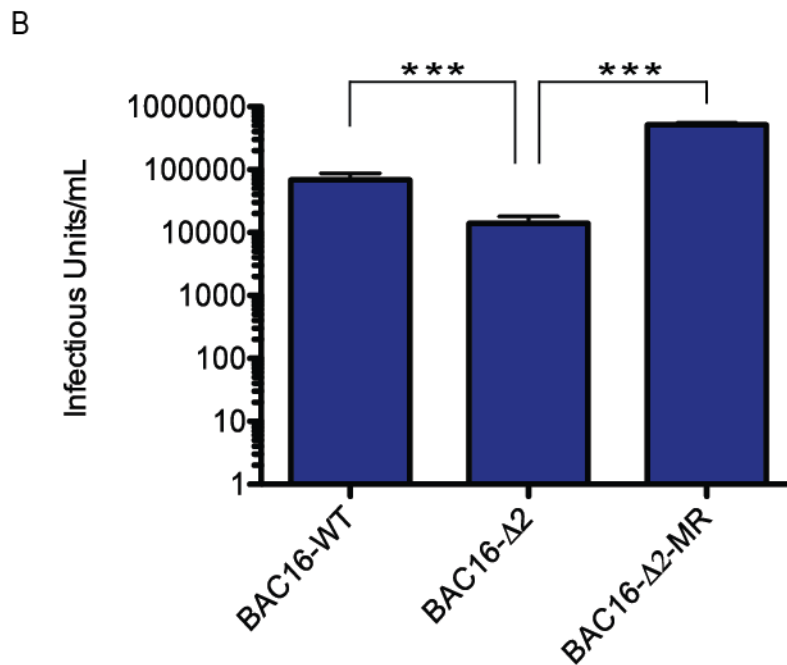
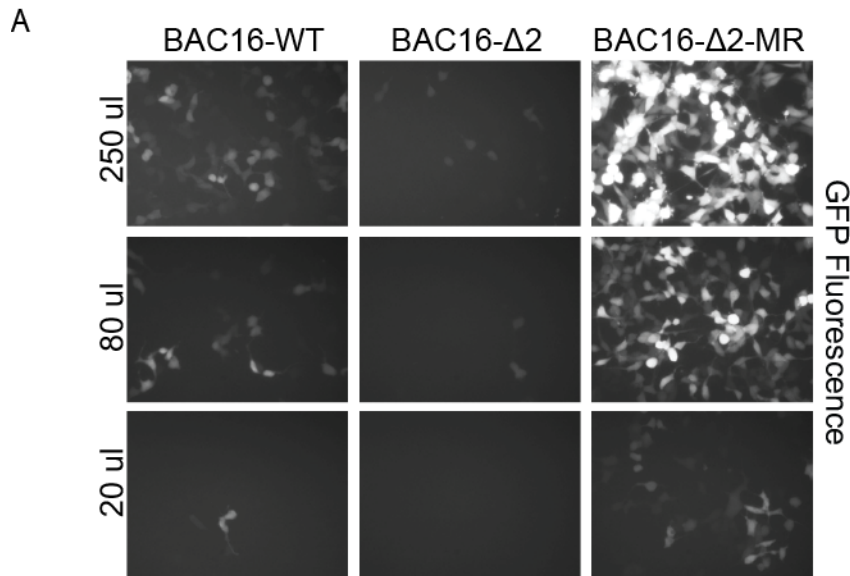


**Figure 4.3. Disruption of uORF2 results in a drop in infectious titer.** (A) Viral supernatants after 48 h post-lytic reactivation from the indicated iSLK-PURO BAC16-WT, -Δ2 and -Δ2-MR cell lines were used to infect  $1 \times 10^5$  2933A cells for 1 hr. 24 h post-infection, 293A cells were visualized by GFP fluorescence. (B) Infectious units were quantified from viral supernatants harvested from the indicated iSLK-PURO BAC16-WT, -Δ2 and -Δ2-MR cell lines after 48 h post-lytic reactivation by flow cytometry (see Materials and Methods). The titer reflects the total number of GFP+ cells in each well divided by the volume of each inoculum. This experiment was performed in triplicate by our collaborator Kevin Brulois; error bars represent the standard deviation

between replicates. Statistical significance was evaluated with a two-tailed unpaired  $t$  test ( $***P < 0.0001$ ).

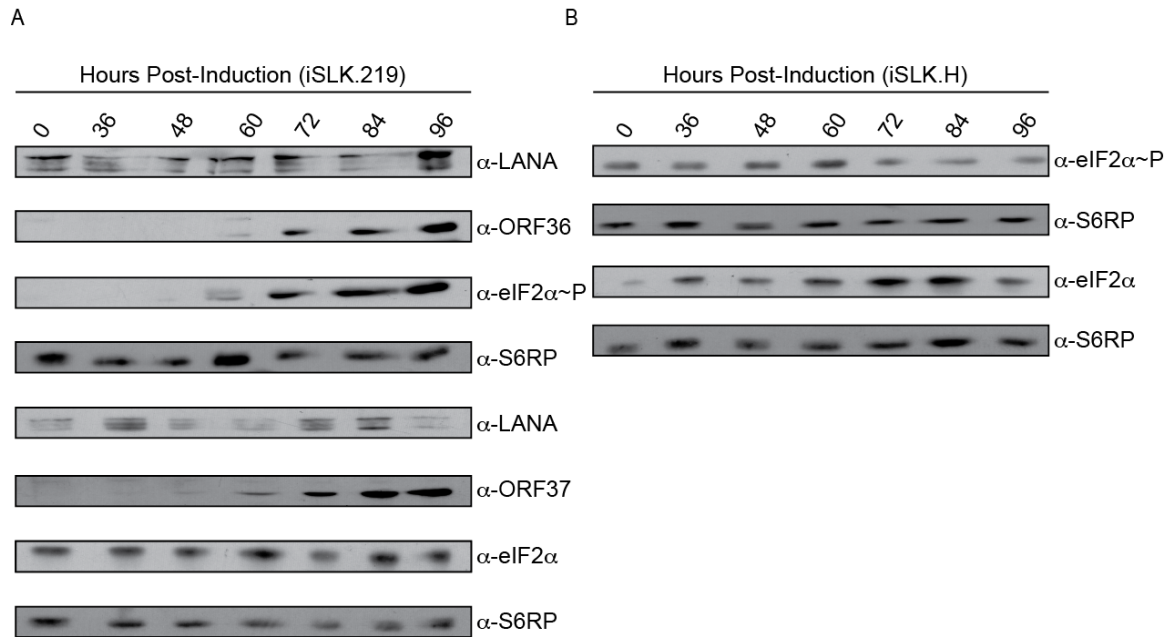


**Figure 4.4. Robust induction of eIF2α phosphorylation during KSHV lytic replication.** (A) iSLK.219 or (B) iSLK.H cells were lytically reactivated with doxycycline (1  $\mu$ g/ml) for the indicated times. Total protein was harvested and cell lysates were resolved by SDS-PAGE and detected by Western blot with antibodies against LANA, ORF36, ORF37, total eIF2α or phospho-eIF2α (Ser51). The ribosomal protein S6RP served as a loading control.

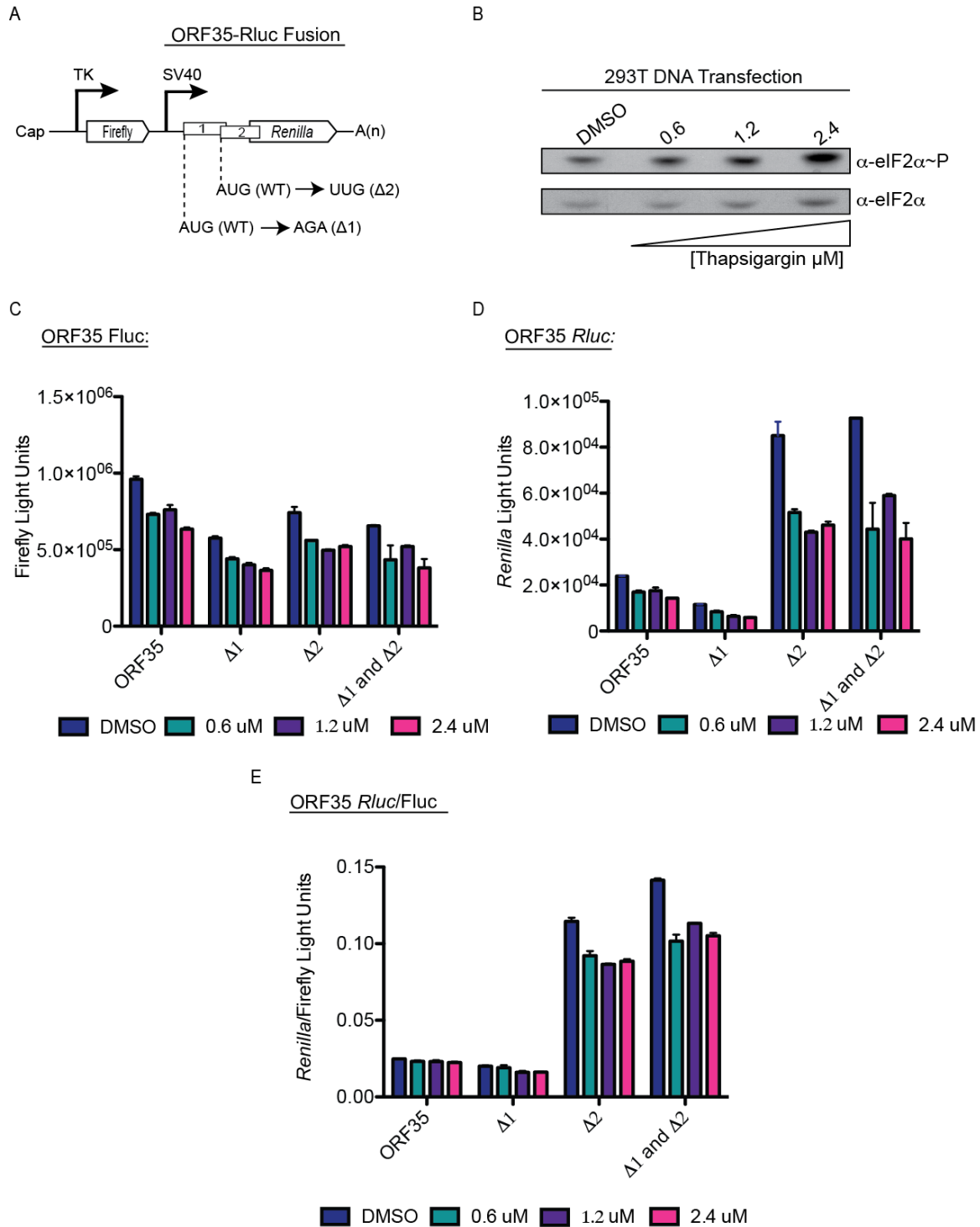


**Figure 4.5. Analysis of uORF1 and uORF2 mutants on ORF35 expression under conditions of eIF2α phosphorylation.** (A) Diagram indicating the nucleotide mutations used to disrupt uORF1 (Δ1) and/or uORF2 (Δ2) within the ORF35 dual luciferase reporter construct. (B-D) 293T were transfected with the indicated construct. At 24 h post-transfection, cells were treated with either vehicle or TG and a dual luciferase assay was performed 10 h post-treatment. (B) Raw firefly values. (C) Raw *Renilla* values. (D)

*Renilla* normalized to firefly values. The experiment was performed in triplicate; error bars represent the standard deviation between replicates.

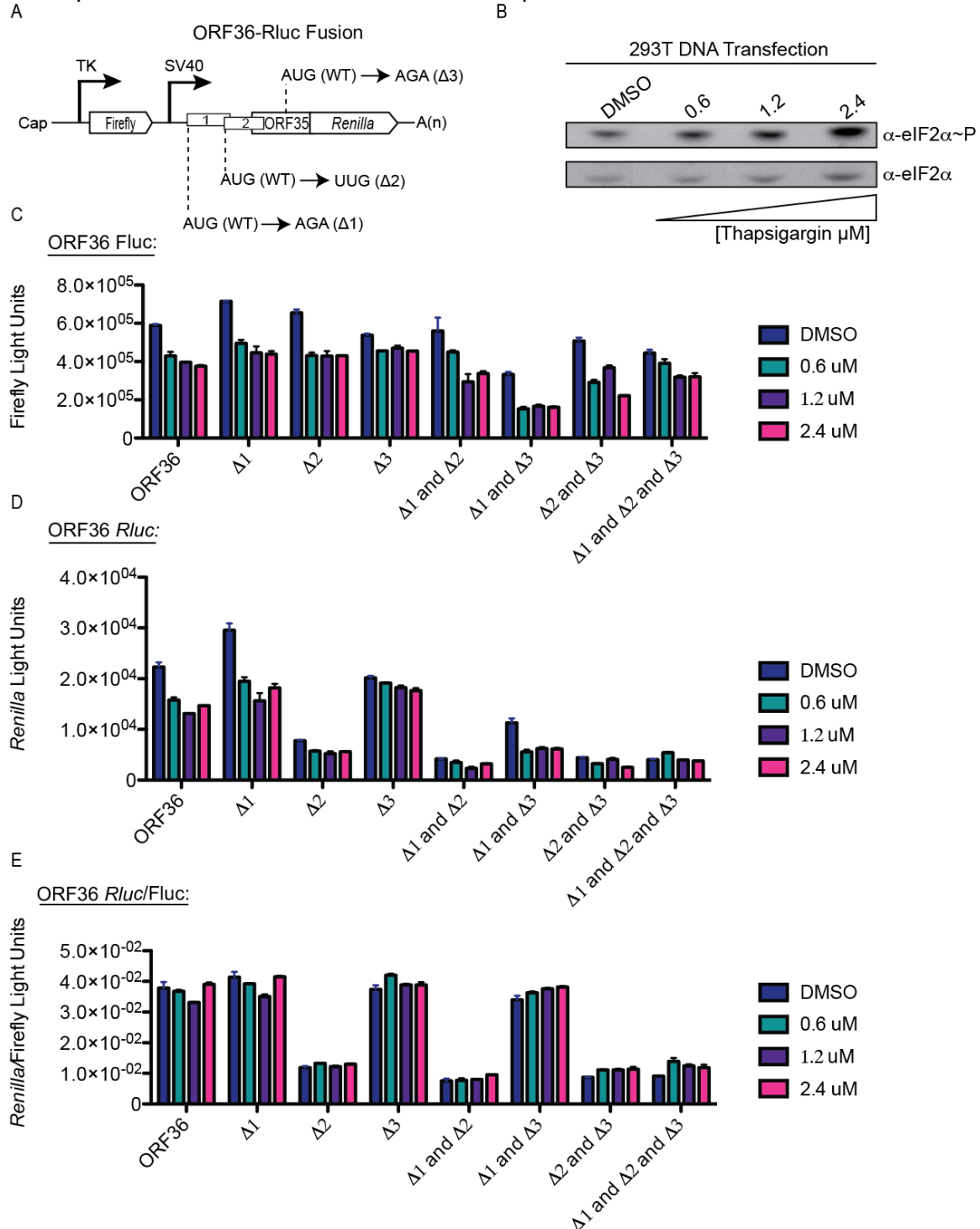
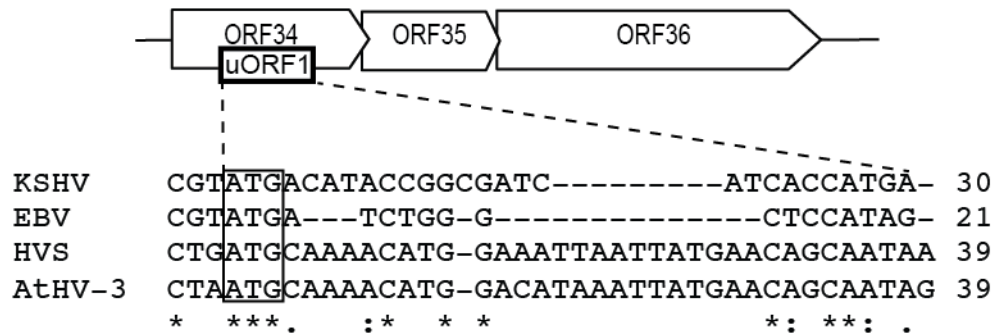


**Figure 4.6. Analysis of uORF1, uORF2 and uORF3 mutants on ORF36 expression under conditions of eIF2 $\alpha$  phosphorylation.** (A) Diagram indicating the nucleotide mutations used to disrupt uORF1 ( $\Delta 1$ ) and/or uORF2 ( $\Delta 2$ ) and/or uORF3 ( $\Delta 3$ ) within the ORF36 dual luciferase reporter construct. (B) 24 h post-transfection with the ORF36

dual luciferase reporter, 293T cells were treated with either DMSO as a vehicle control or increasing doses of TG. 6 h post-treatment, total protein was harvested and lysates were resolved by SDS-PAGE and detected by Western blot with antibodies against total eIF2 $\alpha$  or phospho-eIF2 $\alpha$  (Ser51). (C-E) 293T were transfected with the indicated construct. At 24 h post-transfection, cells were treated with either vehicle or TG and a dual luciferase assay was performed 10 h post-treatment. (C) Raw firefly values. (D) Raw *Renilla* values. (E) *Renilla* normalized to firefly values. The experiment was performed in triplicate; error bars represent the standard deviation between replicates.

A



B

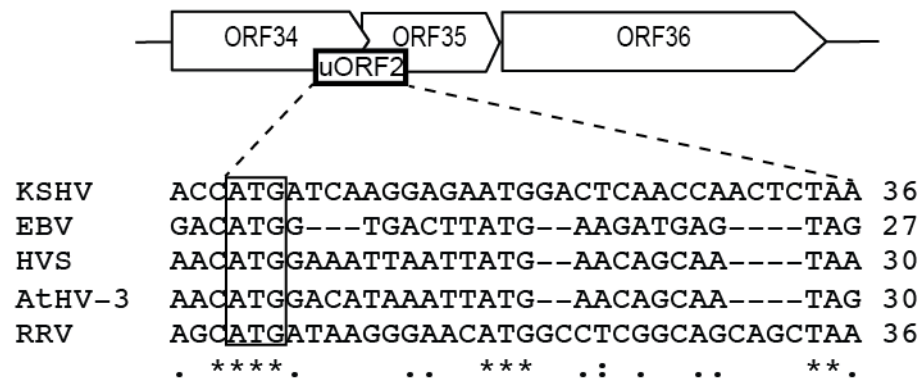


**Figure 4.7. uORF1 and uORF2 are conserved among select  $\gamma$ -herpesviruses.** Alignment using ClustalW2 of (A) uORF1 or (B) uORF2 from KSHV, EBV, HVS, AtHV-3 and RRV. Consensus nucleotides are indicated (three: asterisk; two: dot). uORF length is indicated on the right, and the uORF start codons are boxed.

**Table 4.1. Analysis of the <100 nucleotides upstream of ORF35 in the genomes of  $\gamma$ -herpesviruses with the conserved ORF34-37 genome arrangement deposited in the Virus Pathogen Resource.** A representative strain of each  $\gamma$ -herpesvirus deposited in the Virus Pathogen Database and Analysis Resource that retains the arrangement of *ORF34-37* genetic locus was included in the sequence analysis. The region upstream of the ORF35 start codon ( $\leq 100$  nucleotides) was used as an arbitrary prediction of 5'UTR length. The number of internal AUG codons represents those located between the uORF2 stop codon and the start codon of ORF36 within each respective mRNA.

| Name                              | Strain  | GenBank Accession # | uORF1      | uORF2        | # of internal AUG codons |
|-----------------------------------|---------|---------------------|------------|--------------|--------------------------|
| HHV-8 (KSHV)                      | GK18(P) | NC_009333           | Yes, weak* | Yes, strong* | 2                        |
| HHV-4 (EBV)                       | AG876   | NC_009334           | Yes, weak* | Yes, strong* | 0                        |
| Saimiriine herpesvirus 2 (SaHV-2) | C488    | AJ410493            | Yes, weak* | Yes, strong* | 2                        |
| Ateline herpesvirus 3 (AtHV-3)    | 73      | NC_001987           | Yes, weak* | Yes, strong* | 2                        |
| Macacine herpesvirus 5 (RRV)      | 17577   | NC_003401           | No         | Yes, strong* | 1                        |
| Murid herpesvirus 4               | WM8     | GQ169129            | No         | Yes, weak*   | 4                        |
| Equid herpesvirus 2               | 86/87   | NC_001650           | No         | No           | 0                        |
| Ovine herpesvirus 2               | UNKNOWN | DQ198083            | No         | No           | 0                        |
| Alcelaphine herpesvirus 1         | C500    | AF005370            | No         | No           | 1                        |

\*weak: Lacks a purine at the -3 position of the Kozak consensus sequence.

\*strong: Presence of a purine at the -3 position of the Kozak consensus sequence.

**Table 4.2. List of oligonucleotide primers used in chapter four.**

| <b>Primer Name</b>             | <b>Primer Sequence 5'-3'</b>                                                                                        |
|--------------------------------|---------------------------------------------------------------------------------------------------------------------|
| <b>BAC16 Δ2 F</b>              | TATCTAATCATCCCATCGTATGACATACCGGCGATCATCACCTTGATCA<br>AGGAGAATGGACTCAAAGGATGACGACGATAAGTAGGG                         |
| <b>BAC16 Δ2 R</b>              | TAAACTCTCTTTTAGAGTTGGTTGAGTCCATTCTCCTTGATCAAGGTGAT<br>GATCGCCGGTATGTCAAACCAATTAACCAATTCTGATTAG                      |
| <b>BAC16 ORF35<br/>iFLAG F</b> | GAAGCAGTTTCAGGACGCACAGCATGCCGTCAGGGCCCACGACTACAA<br>GGACGACGATGACAAGGGTGCCTGAAGCGCGAAGCAGGATGACGAC<br>GATAAGTAGGG   |
| <b>BAC16 ORF35<br/>iFLAG R</b> | GTACCAGAGTCCCAGCTCGGCTTCGCGCTTCAGTGCACCCTTGTCATC<br>GTCGTCCTTGTAGTCGTGGGCCCTGACGGCATGCTAACCAATTAACCAA<br>TTCTGATTAG |

\*F: Forward

\*R: Reverse

## Chapter Five: Perspectives

### Further Discussion and Future Directions

The compact nature of viral genomes has resulted in the evolution of specialized strategies to maximize their coding capacity. Accordingly, many viral mRNAs do not conform to the one protein per mRNA cellular paradigm and require specialized mechanisms to subvert the translational constraints of the host. Examples of such mechanisms include the production of subgenomic RNAs, translation of a large polyprotein that is cleaved into multiple proteins, ribosomal frameshifting, leaky scanning and ribosomal shunting [147].

In this thesis we describe a novel bicistronic mRNA encoded by the  $\gamma$ -herpesvirus KSHV. In other characterized examples of viral and cellular translation via a reinitiation mechanism, expression of the downstream gene is significantly tempered as a consequence of ribosomal engagement at an upstream start codon [26,36,148-150]. Aside from being bicistronic, translation from the KSHV ORF35-37 transcript is unusual in that the protein product of ORF36 is expressed more robustly than the 5' ORF35 despite the fact that the ORF35 start codon is in a favorable sequence context. We reveal that ORF36 translation occurs by a unique adaptation of ribosomal reinitiation after translation of two overlapping uORFs in the native 5'UTR. These findings provide the first example of cap-dependent non-canonical translation in KSHV and illustrate a novel strategy to translate polycistronic mRNA.

A key question that arises from this work is how the finely tuned strategy of translational control executed by uORF1 and uORF2 on the ORF35-37 mRNA impacts the KSHV lytic lifecycle. One possibility is that ORF35 and ORF36 are required at different points during lytic infection and that during the course of viral replication, conditions arise that favor translation of one protein versus the other. A detailed time course of the temporal cascade of gene expression using the internally FLAG epitope-tagged KSHV BAC16 virus and anti-ORF36 antibodies would reveal if these two proteins are translated at different points during the viral lifecycle.

We have learned from uORF-mediated regulation of the *S. cerevisiae GCN4* that whether a downstream AUG is recognized by reinitiation is tightly regulated by the availability of the ternary complex which is in turn controlled by the level of eIF2 $\alpha$  phosphorylation [37-39]. Strikingly, infection with an array of diverse viruses gives rise to dsRNA which activates PKR, one of several kinases that mediate phosphorylation of eIF2 $\alpha$ , which in turn globally inhibits viral and host protein synthesis [151]. Perhaps unsurprisingly, a recurrent weapon in the host-virus molecular arms race is prevention of eIF2 $\alpha$  phosphorylation. These strategies have been well characterized for a number of viruses including human hepatitis C virus and HSV-1 [152-154]. Thus it is tempting to speculate that the reinitiation efficiency after uORF1 and uORF2 translation may be modulated by ternary complex availability, which is in turn governed by the eIF2 $\alpha$  phosphorylation status of the infected cell. Our initial analysis of eIF2 $\alpha$  during KSHV infection detected a robust enhancement of eIF2 $\alpha$  phosphorylation with approximate delayed-early kinetics. However, whether this result holds true in other cell lines and if viral RNAs or proteins modulate this response remains to be explored. Furthermore, no straightforward evidence of the phosphorylation state of eIF2 $\alpha$  modulating ORF36

expression was found in our cell culture model of eIF2 $\alpha$  phosphorylation. However this system does not perfectly mimic conditions during infections; perhaps the timing or level of eIF2 $\alpha$  phosphorylation found during infection was not faithfully recapitulated in our cell culture system. Therefore, once the viral proteins that modulate eIF2 $\alpha$  phosphorylation during KSHV infection have been delineated, it will be compelling to examine whether the uORF1 and uORF2 configuration on the ORF35-37 polycistronic mRNA stimulate or repress translation of ORF35 or ORF36 in a manner dependent on the phosphorylation status of eIF2 $\alpha$ .

Notably, many of the non-canonical translation initiation strategies originally discovered in viruses were subsequently found to occur on certain cellular mRNAs [16,49,74]. It is clear that the cellular translation apparatus is capable of using a uORF-dependent strategy to allow for the translation of a viral bicistronic mRNA, perhaps this mechanism is also used on certain cellular mRNAs to allow for bicistronic translation. Intriguingly  $\geq 4000$  uORFs exist that overlap with the primary ORF in the human transcriptome, it would be interesting to examine whether the position of certain of these uORFs function to expand to coding capacity of cellular mRNAs or give rise to protein isoforms with distinct biological properties [21,155].

### **Concluding Remarks**

Viruses, with the notable exception of Mimiviruses, generally do not encode genes involved in translation and thus operate under the constraints of host protein synthesis. Thus, strong selection is exerted on viral genomes to evolve a broad range of non-canonical translation strategies to maximize their coding capacity [1,2]. Such mechanisms as translation of a large polyprotein that is cleaved into multiple proteins, IRESs, and non-canonical 5' m<sup>7</sup>G cap-dependent strategies have been largely the dominion of RNA viruses due to the extreme compact nature of the viral genome. However, in this report we describe a novel functional bicistronic mRNA encoded from the genome of a large, dsDNA virus. Perhaps unsurprisingly, the mechanism of translational control of the KSHV ORF35-37 mRNA does not conform to either the cellular paradigm of one protein per mRNA or those non-canonical strategies described for RNA viruses. Instead, this thesis describes a novel adaptation of reinitiation after two uORFs, a mechanism generally used to repress translation of a primary ORF, to instead allow for the translation of an internal gene. Future work on the translational control of this mRNA may provide further insight into the uORF-mediated regulation of ORF35 and ORF36 during the viral lifecycle or lead to the discovery of uORF-dependent cellular bicistronic mRNAs.

## REFERENCES

1. Colson P, Fournous G, Diene SM, Raoult D (2013) Codon usage, amino acid usage, transfer RNA and amino-acyl-tRNA synthetases in Mimiviruses. *Intervirology* 56: 364-375.
2. Arslan D, Legendre M, Seltzer V, Abergel C, Claverie JM (2011) Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proc Natl Acad Sci U S A* 108: 17486-17491.
3. Haque M, Wang V, Davis DA, Zheng ZM, Yarchoan R (2006) Genetic organization and hypoxic activation of the Kaposi's sarcoma-associated herpesvirus ORF34-37 gene cluster. *J Virol* 80: 7037-7051.
4. Zheng ZM (2003) Split genes and their expression in Kaposi's sarcoma-associated herpesvirus. *Rev Med Virol* 13: 173-184.
5. Jackson RJ, Hellen CU, Pestova TV (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* 11: 113-127.
6. Gebauer F, Hentze MW (2004) Molecular mechanisms of translational control. *Nat Rev Mol Cell Biol* 5: 827-835.
7. Hinnebusch AG (2006) eIF3: a versatile scaffold for translation initiation complexes. *Trends Biochem Sci* 31: 553-562.
8. Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44: 283-292.
9. Kozak M (2001) Constraints on reinitiation of translation in mammals. *Nucleic Acids Res* 29: 5226-5232.
10. Kozak M (2002) Pushing the limits of the scanning mechanism for initiation of translation. *Gene* 299: 1-34.
11. Pisarev AV, Kolupaeva VG, Pisareva VP, Merrick WC, Hellen CU, et al. (2006) Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. *Genes Dev* 20: 624-636.
12. Kozak M (1997) Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. *Embo J* 16: 2482-2492.
13. Pestova TV, Kolupaeva VG (2002) The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev* 16: 2906-2922.
14. Pestova TV, Borukhov SI, Hellen CU (1998) Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nature* 394: 854-859.
15. Futterer J, Hohn T (1991) Translation of a polycistronic mRNA in the presence of the cauliflower mosaic virus transactivator protein. *Embo J* 10: 3887-3896.
16. Futterer J, Kiss-Laszlo Z, Hohn T (1993) Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA. *Cell* 73: 789-802.
17. Dominguez DI, Ryabova LA, Pooggin MM, Schmidt-Puchta W, Futterer J, et al. (1998) Ribosome shunting in cauliflower mosaic virus. Identification of an essential and sufficient structural element. *J Biol Chem* 273: 3669-3678.

18. Ryabova LA, Hohn T (2000) Ribosome shunting in the cauliflower mosaic virus 35S RNA leader is a special case of reinitiation of translation functioning in plant and animal systems. *Genes Dev* 14: 817-829.
19. Kozak M (1987) Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes. *Mol Cell Biol* 7: 3438-3445.
20. Morris DR, Geballe AP (2000) Upstream open reading frames as regulators of mRNA translation. *Mol Cell Biol* 20: 8635-8642.
21. Calvo SE, Pagliarini DJ, Mootha VK (2009) Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc Natl Acad Sci U S A* 106: 7507-7512.
22. Iacono M, Mignone F, Pesole G (2005) uAUG and uORFs in human and rodent 5'untranslated mRNAs. *Gene* 349: 97-105.
23. Neafsey DE, Galagan JE (2007) Dual modes of natural selection on upstream open reading frames. *Mol Biol Evol* 24: 1744-1751.
24. Mize GJ, Ruan H, Low JJ, Morris DR (1998) The inhibitory upstream open reading frame from mammalian S-adenosylmethionine decarboxylase mRNA has a strict sequence specificity in critical positions. *J Biol Chem* 273: 32500-32505.
25. Lovett PS, Rogers EJ (1996) Ribosome regulation by the nascent peptide. *Microbiol Rev* 60: 366-385.
26. Cao J, Geballe AP (1995) Translational inhibition by a human cytomegalovirus upstream open reading frame despite inefficient utilization of its AUG codon. *J Virol* 69: 1030-1036.
27. Cao J, Geballe AP (1996) Coding sequence-dependent ribosomal arrest at termination of translation. *Mol Cell Biol* 16: 603-608.
28. Luo Z, Sachs MS (1996) Role of an upstream open reading frame in mediating arginine-specific translational control in *Neurospora crassa*. *J Bacteriol* 178: 2172-2177.
29. Poyry TA, Kaminski A, Jackson RJ (2004) What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? *Genes Dev* 18: 62-75.
30. Park HS, Himmelbach A, Browning KS, Hohn T, Ryabova LA (2001) A plant viral "reinitiation" factor interacts with the host translational machinery. *Cell* 106: 723-733.
31. Szamecz B, Rutkai E, Cuchalova L, Munzarova V, Herrmannova A, et al. (2008) eIF3a cooperates with sequences 5' of uORF1 to promote resumption of scanning by post-termination ribosomes for reinitiation on GCN4 mRNA. *Genes Dev* 22: 2414-2425.
32. Luttermann C, Meyers G (2009) The importance of inter- and intramolecular base pairing for translation reinitiation on a eukaryotic bicistronic mRNA. *Genes Dev* 23: 331-344.
33. Meyers G (2007) Characterization of the sequence element directing translation reinitiation in RNA of the calicivirus rabbit hemorrhagic disease virus. *J Virol* 81: 9623-9632.
34. Poyry TA, Kaminski A, Connell EJ, Fraser CS, Jackson RJ (2007) The mechanism of an exceptional case of reinitiation after translation of a long ORF reveals why

- such events do not generally occur in mammalian mRNA translation. *Genes Dev* 21: 3149-3162.
35. Luttermann C, Meyers G (2007) A bipartite sequence motif induces translation reinitiation in feline calicivirus RNA. *J Biol Chem* 282: 7056-7065.
  36. Meyers G (2003) Translation of the minor capsid protein of a calicivirus is initiated by a novel termination-dependent reinitiation mechanism. *J Biol Chem* 278: 34051-34060.
  37. Gaba A, Wang Z, Krishnamoorthy T, Hinnebusch AG, Sachs MS (2001) Physical evidence for distinct mechanisms of translational control by upstream open reading frames. *EMBO J* 20: 6453-6463.
  38. Hinnebusch AG (1993) Gene-specific translational control of the yeast GCN4 gene by phosphorylation of eukaryotic initiation factor 2. *Mol Microbiol* 10: 215-223.
  39. Hinnebusch AG (2005) Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol* 59: 407-450.
  40. Asano K, Krishnamoorthy T, Phan L, Pavitt GD, Hinnebusch AG (1999) Conserved bipartite motifs in yeast eIF5 and eIF2Bepsilon, GTPase-activating and GDP-GTP exchange factors in translation initiation, mediate binding to their common substrate eIF2. *Embo J* 18: 1673-1688.
  41. Delforge J, Messenguy F, Wiame JM (1975) The regulation of arginine biosynthesis in *Saccharomyces cerevisiae*. The specificity of argR- mutations and the general control of amino-acid biosynthesis. *Eur J Biochem* 57: 231-239.
  42. Johannes G, Carter MS, Eisen MB, Brown PO, Sarnow P (1999) Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. *Proc Natl Acad Sci U S A* 96: 13118-13123.
  43. Pelletier J, Kaplan G, Racaniello VR, Sonenberg N (1988) Cap-independent translation of poliovirus mRNA is conferred by sequence elements within the 5' noncoding region. *Mol Cell Biol* 8: 1103-1112.
  44. Pelletier J, Sonenberg N (1988) Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* 334: 320-325.
  45. Pelletier J, Sonenberg N (1989) Internal binding of eucaryotic ribosomes on poliovirus RNA: translation in HeLa cell extracts. *J Virol* 63: 441-444.
  46. Kozak M (2007) Lessons (not) learned from mistakes about translation. *Gene* 403: 194-203.
  47. Kozak M (2005) A second look at cellular mRNA sequences said to function as internal ribosome entry sites. *Nucleic Acids Res* 33: 6593-6602.
  48. Mokrejs M, Masek T, Vopalensky V, Hlubucek P, Delbos P, et al. (2010) IRESite--a tool for the examination of viral and cellular internal ribosome entry sites. *Nucleic Acids Res* 38: D131-136.
  49. Jackson RJ (2013) The current status of vertebrate cellular mRNA IRESs. *Cold Spring Harb Perspect Biol* 5.
  50. Pyronnet S, Pradayrol L, Sonenberg N (2000) A cell cycle-dependent internal ribosome entry site. *Mol Cell* 5: 607-616.
  51. Pellet PE, and B. Roizman (2007) *The Family Herpesviridae: A Brief Introduction*. . Fields Virology, 5th Edition Lippincott Williams & Wilkins, Philadelphia 2: 2479-2499.

52. Mesri EA, Cesarman E, Boshoff C (2010) Kaposi's sarcoma and its associated herpesvirus. *Nat Rev Cancer* 10: 707-719.
53. Hollier LM, Grissom H (2005) Human herpes viruses in pregnancy: cytomegalovirus, Epstein-Barr virus, and varicella zoster virus. *Clin Perinatol* 32: 671-696.
54. Ganem D (2006) KSHV infection and the pathogenesis of Kaposi's sarcoma. *Annu Rev Pathol* 1: 273-296.
55. Ganem D (2010) KSHV and the pathogenesis of Kaposi sarcoma: listening to human biology and medicine. *J Clin Invest* 120: 939-949.
56. Kaposi M (1872) Idiopathisches multiples Pigmentsarkom der Haut. . *Arch Dermatol Syph* 4.
57. Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, et al. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220: 868-871.
58. Wabinga HR, Parkin DM, Wabwire-Mangen F, Mugerwa JW (1993) Cancer in Kampala, Uganda, in 1989-91: changes in incidence in the era of AIDS. *Int J Cancer* 54: 26-36.
59. Parkin DM, Sitas F, Chirenje M, Stein L, Abratt R, et al. (2008) Part I: Cancer in Indigenous Africans--burden, distribution, and trends. *Lancet Oncol* 9: 683-692.
60. Gallo RC (1998) The enigmas of Kaposi's sarcoma. *Science* 282: 1837-1839.
61. Chang Y, Cesarman E, Pessin MS, Lee F, Culpepper J, et al. (1994) Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* 266: 1865-1869.
62. Soulier J, Grollet L, Oksenhendler E, Cacoub P, Cazals-Hatem D, et al. (1995) Kaposi's sarcoma-associated herpesvirus-like DNA sequences in multicentric Castleman's disease. *Blood* 86: 1276-1280.
63. Cesarman E, Chang Y, Moore PS, Said JW, Knowles DM (1995) Kaposi's sarcoma-associated herpesvirus-like DNA sequences in AIDS-related body-cavity-based lymphomas. *N Engl J Med* 332: 1186-1191.
64. Russo JJ, Bohenzky RA, Chien MC, Chen J, Yan M, et al. (1996) Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8). *Proc Natl Acad Sci U S A* 93: 14862-14867.
65. Nador RG, Milligan LL, Flore O, Wang X, Arvanitakis L, et al. (2001) Expression of Kaposi's sarcoma-associated herpesvirus G protein-coupled receptor monocistronic and bicistronic transcripts in primary effusion lymphomas. *Virology* 287: 62-70.
66. Majerciak V, Yamanegi K, Zheng ZM (2006) Gene structure and expression of Kaposi's sarcoma-associated herpesvirus ORF56, ORF57, ORF58, and ORF59. *J Virol* 80: 11968-11981.
67. Talbot SJ, Weiss RA, Kellam P, Boshoff C (1999) Transcriptional analysis of human herpesvirus-8 open reading frames 71, 72, 73, K14, and 74 in a primary effusion lymphoma cell line. *Virology* 257: 84-94.
68. Low W, Harries M, Ye H, Du MQ, Boshoff C, et al. (2001) Internal ribosome entry site regulates translation of Kaposi's sarcoma-associated herpesvirus FLICE inhibitory protein. *J Virol* 75: 2938-2945.
69. Grundhoff A, Ganem D (2001) Mechanisms governing expression of the v-FLIP gene of Kaposi's sarcoma-associated herpesvirus. *J Virol* 75: 1857-1863.

70. Bieleski L, Talbot SJ (2001) Kaposi's sarcoma-associated herpesvirus vCyclin open reading frame contains an internal ribosome entry site. *J Virol* 75: 1864-1869.
71. Cao F, Tavis JE RNA elements directing translation of the duck hepatitis B Virus polymerase via ribosomal shunting. *J Virol* 85: 6343-6352.
72. Chappell SA, Dresios J, Edelman GM, Mauro VP (2006) Ribosomal shunting mediated by a translational enhancer element that base pairs to 18S rRNA. *Proc Natl Acad Sci U S A* 103: 9488-9493.
73. Pooggin MM, Ryabova LA, He X, Futterer J, Hohn T (2006) Mechanism of ribosome shunting in Rice tungro bacilliform pararetrovirus. *RNA* 12: 841-850.
74. Yueh A, Schneider RJ (2000) Translation by ribosome shunting on adenovirus and hsp70 mRNAs facilitated by complementarity to 18S rRNA. *Genes Dev* 14: 414-421.
75. Grant CM, Hinnebusch AG (1994) Effect of sequence context at stop codons on efficiency of reinitiation in GCN4 translational control. *Mol Cell Biol* 14: 606-618.
76. Miller PF, Hinnebusch AG (1989) Sequences that surround the stop codons of upstream open reading frames in GCN4 mRNA determine their distinct functions in translational control. *Genes Dev* 3: 1217-1225.
77. Haque M, Kousoulas KG (2013) The Kaposi's sarcoma-associated herpesvirus ORF34 protein binds to HIF-1alpha and causes its degradation via the proteasome pathway. *J Virol* 87: 2164-2173.
78. Yamauchi Y, Wada K, Goshima F, Takakuwa H, Daikoku T, et al. (2001) The UL14 protein of herpes simplex virus type 2 translocates the minor capsid protein VP26 and the DNA cleavage and packaging UL33 protein into the nucleus of coexpressing cells. *J Gen Virol* 82: 321-330.
79. Yamauchi Y, Daikoku T, Goshima F, Nishiyama Y (2003) Herpes simplex virus UL14 protein blocks apoptosis. *Microbiol Immunol* 47: 685-689.
80. Hamza MS, Reyes RA, Izumiya Y, Wisdom R, Kung HJ, et al. (2004) ORF36 protein kinase of Kaposi's sarcoma herpesvirus activates the c-Jun N-terminal kinase signaling pathway. *J Biol Chem* 279: 38325-38330.
81. Park J, Lee D, Seo T, Chung J, Choe J (2000) Kaposi's sarcoma-associated herpesvirus (human herpesvirus-8) open reading frame 36 protein is a serine protein kinase. *J Gen Virol* 81: 1067-1071.
82. Izumiya Y, Izumiya C, Van Geelen A, Wang DH, Lam KS, et al. (2007) Kaposi's sarcoma-associated herpesvirus-encoded protein kinase and its interaction with K-bZIP. *J Virol* 81: 1072-1082.
83. Kedes DH, Ganem D (1997) Sensitivity of Kaposi's sarcoma-associated herpesvirus replication to antiviral drugs. Implications for potential therapy. *J Clin Invest* 99: 2082-2086.
84. Casper C, Krantz EM, Corey L, Kuntz SR, Wang J, et al. (2008) Valganciclovir for suppression of human herpesvirus-8 replication: a randomized, double-blind, placebo-controlled, crossover trial. *J Infect Dis* 198: 23-30.
85. Cannon JS, Hamzeh F, Moore S, Nicholas J, Ambinder RF (1999) Human herpesvirus 8-encoded thymidine kinase and phosphotransferase homologues confer sensitivity to ganciclovir. *J Virol* 73: 4786-4793.

86. Martin DF, Kuppermann BD, Wolitz RA, Palestine AG, Li H, et al. (1999) Oral ganciclovir for patients with cytomegalovirus retinitis treated with a ganciclovir implant. Roche Ganciclovir Study Group. *N Engl J Med* 340: 1063-1070.
87. Covarrubias S, Gaglia MM, Kumar GR, Wong W, Jackson AO, et al. Coordinated destruction of cellular messages in translation complexes by the gammaherpesvirus host shutoff factor and the Mammalian exonuclease xrn1. *PLoS Pathog* 7: e1002339.
88. Glaunsinger B, Ganem D (2004) Lytic KSHV infection inhibits host gene expression by accelerating global mRNA turnover. *Mol Cell* 13: 713-723.
89. Glaunsinger B, Chavez L, Ganem D (2005) The exonuclease and host shutoff functions of the SOX protein of Kaposi's sarcoma-associated herpesvirus are genetically separable. *J Virol* 79: 7396-7401.
90. Masa SR, Lando R, Sarid R (2008) Transcriptional regulation of the open reading frame 35 encoded by Kaposi's sarcoma-associated herpesvirus. *Virology* 371: 14-31.
91. Feederle R, Mehl-Lautscham AM, Bannert H, Delecluse HJ (2009) The Epstein-Barr virus protein kinase BGLF4 and the exonuclease BGLF5 have opposite effects on the regulation of viral protein production. *J Virol* 83: 10877-10891.
92. Dourmishev LA, Dourmishev AL, Palmeri D, Schwartz RA, Lukac DM (2003) Molecular genetics of Kaposi's sarcoma-associated herpesvirus (human herpesvirus-8) epidemiology and pathogenesis. *Microbiol Mol Biol Rev* 67: 175-212, table of contents.
93. Sun R, Lin SF, Gradoville L, Yuan Y, Zhu F, et al. (1998) A viral gene that activates lytic cycle expression of Kaposi's sarcoma-associated herpesvirus. *Proc Natl Acad Sci U S A* 95: 10866-10871.
94. Wang SS, Chen LW, Chen LY, Tsai HH, Shih YC, et al. (2010) Transcriptional regulation of the ORF61 and ORF60 genes of Kaposi's sarcoma-associated herpesvirus. *Virology* 397: 311-321.
95. Bu W, Carroll KD, Palmeri D, Lukac DM (2007) Kaposi's sarcoma-associated herpesvirus/human herpesvirus 8 ORF50/Rta lytic switch protein functions as a tetramer. *J Virol* 81: 5788-5806.
96. Oen H, Pellegrini M, Cantor CR (1974) Peptidyl transferase inhibitors alter the covalent reaction of BrAcPhe-tRNA with the E. coli ribosome. *FEBS Lett* 45: 218-222.
97. Chen CY, Sarnow P (1995) Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* 268: 415-417.
98. Kozak M (1989) Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol Cell Biol* 9: 5134-5142.
99. Nakamura H, Lu M, Gwack Y, Souvlis J, Zeichner SL, et al. (2003) Global changes in Kaposi's sarcoma-associated virus gene expression patterns following expression of a tetracycline-inducible Rta transactivator. *J Virol* 77: 4205-4220.
100. Bu W, Palmeri D, Krishnan R, Marin R, Aris VM, et al. (2008) Identification of direct transcriptional targets of the Kaposi's sarcoma-associated herpesvirus Rta lytic switch protein by conditional nuclear localization. *J Virol* 82: 10709-10723.

101. Zhou FC, Zhang YJ, Deng JH, Wang XP, Pan HY, et al. (2002) Efficient infection by a recombinant Kaposi's sarcoma-associated herpesvirus cloned in a bacterial artificial chromosome: application for genetic analysis. *J Virol* 76: 6185-6196.
102. Arias C, Walsh D, Harbell J, Wilson AC, Mohr I (2009) Activation of host translational control pathways by a viral developmental switch. *PLoS Pathog* 5: e1000334.
103. Pestova TV, Kolupaeva VG, Lomakin IB, Pilipenko EV, Shatsky IN, et al. (2001) Molecular mechanisms of translation initiation in eukaryotes. *Proc Natl Acad Sci U S A* 98: 7029-7036.
104. Jackson RJ (2005) Alternative mechanisms of initiating translation of mammalian mRNAs. *Biochem Soc Trans* 33: 1231-1241.
105. Tsukiyama-Kohara K, Iizuka N, Kohara M, Nomoto A (1992) Internal ribosome entry site within hepatitis C virus RNA. *J Virol* 66: 1476-1483.
106. Wilson JE, Powell MJ, Hoover SE, Sarnow P (2000) Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. *Mol Cell Biol* 20: 4990-4999.
107. Luukkonen BG, Tan W, Schwartz S (1995) Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. *J Virol* 69: 4086-4094.
108. Donze O, Damay P, Spahr PF (1995) The first and third uORFs in RSV leader RNA are efficiently translated: implications for translational regulation and viral RNA packaging. *Nucleic Acids Res* 23: 861-868.
109. Parola AL, Kobilka BK (1994) The peptide product of a 5' leader cistron in the beta 2 adrenergic receptor mRNA inhibits receptor synthesis. *J Biol Chem* 269: 4497-4505.
110. Pendleton LC, Goodwin BL, Solomonson LP, Eichler DC (2005) Regulation of endothelial argininosuccinate synthase expression and NO production by an upstream open reading frame. *J Biol Chem* 280: 24252-24260.
111. Kronstad LM, Brulois KF, Jung JU, Glaunsinger BA (2013) Dual short upstream open reading frames control translation of a herpesviral polycistronic mRNA. *PLoS Pathog* 9: e1003156.
112. Hartz D, McPheeters DS, Gold L (1989) Selection of the initiator tRNA by *Escherichia coli* initiation factors. *Genes Dev* 3: 1899-1912.
113. Ringquist S, Cunningham P, Weitzmann C, Formenoy L, Pleij C, et al. (1993) Translation initiation complex formation with 30 S ribosomal particles mutated at conserved positions in the 3'-minor domain of 16 S RNA. *J Mol Biol* 234: 14-27.
114. Spedding G, Gluick TC, Draper DE (1993) Ribosome initiation complex formation with the pseudoknotted alpha operon messenger RNA. *J Mol Biol* 229: 609-622.
115. Gould PS, Bird H, Easton AJ (2005) Translation toeprinting assays using fluorescently labeled primers and capillary electrophoresis. *Biotechniques* 38: 397-400.
116. Monro RE, Celma ML, Vazquez D (1969) Action of sparsomycin on ribosome-catalysed peptidyl transfer. *Nature* 222: 356-358.
117. Kozak M (1998) Primer extension analysis of eukaryotic ribosome-mRNA complexes. *Nucleic Acids Res* 26: 4853-4859.

118. Kozak M, Shatkin AJ (1977) Sequences and properties of two ribosome binding sites from the small size class of reovirus messenger RNA. *J Biol Chem* 252: 6895-6908.
119. Chen A, Kao YF, Brown CM (2005) Translation of the first upstream ORF in the hepatitis B virus pregenomic RNA modulates translation at the core and polymerase initiation codons. *Nucleic Acids Res* 33: 1169-1181.
120. Sen N, Cao F, Tavis JE (2004) Translation of duck hepatitis B virus reverse transcriptase by ribosomal shunting. *J Virol* 78: 11751-11757.
121. Kozak M (1991) An analysis of vertebrate mRNA sequences: intimations of translational control. *J Cell Biol* 115: 887-903.
122. Ryabova LA, Pooggin MM, Hohn T (2006) Translation reinitiation and leaky scanning in plant viruses. *Virus Res* 119: 52-62.
123. Racine T, Duncan R Facilitated leaky scanning and atypical ribosome shunting direct downstream translation initiation on the tricistronic S1 mRNA of avian reovirus. *Nucleic Acids Res* 38: 7260-7272.
124. Scheper GC, Voorma HO, Thomas AA (1994) Basepairing with 18S ribosomal RNA in internal initiation of translation. *FEBS Lett* 352: 271-275.
125. Kojima KK, Matsumoto T, Fujiwara H (2005) Eukaryotic translational coupling in UAAUG stop-start codons for the bicistronic RNA translation of the non-long terminal repeat retrotransposon SART1. *Mol Cell Biol* 25: 7675-7686.
126. Matsuda D, Dreher TW (2006) Close spacing of AUG initiation codons confers dicistronic character on a eukaryotic mRNA. *Rna* 12: 1338-1349.
127. Williams MA, Lamb RA (1989) Effect of mutations and deletions in a bicistronic mRNA on the synthesis of influenza B virus NB and NA glycoproteins. *J Virol* 63: 28-35.
128. Kozak M (1995) Adherence to the first-AUG rule when a second AUG codon follows closely upon the first. *Proc Natl Acad Sci U S A* 92: 2662-2666.
129. Majerciak V, Pripuzova N, McCoy JP, Gao SJ, Zheng ZM (2007) Targeted disruption of Kaposi's sarcoma-associated herpesvirus ORF57 in the viral genome is detrimental for the expression of ORF59, K8alpha, and K8.1 and the production of infectious virus. *J Virol* 81: 1062-1071.
130. Col B, Oltean S, Banerjee R (2007) Translational regulation of human methionine synthase by upstream open reading frames. *Biochim Biophys Acta* 1769: 532-540.
131. Hurtaud C, Gelly C, Bouillaud F, Levi-Meyrueis C (2006) Translation control of UCP2 synthesis by the upstream open reading frame. *Cell Mol Life Sci* 63: 1780-1789.
132. Rabadan-Diehl C, Martinez A, Volpi S, Subburaju S, Aguilera G (2007) Inhibition of vasopressin V1b receptor translation by upstream open reading frames in the 5'-untranslated region. *J Neuroendocrinol* 19: 309-319.
133. Bechtel JT, Liang Y, Hvidding J, Ganem D (2003) Host range of Kaposi's sarcoma-associated herpesvirus in cultured cells. *J Virol* 77: 6474-6481.
134. Vieira J, O'Hearn P, Kimball L, Chandran B, Corey L (2001) Activation of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) lytic replication by human cytomegalovirus. *J Virol* 75: 1378-1386.

135. Myoung J, Ganem D (2011) Generation of a doxycycline-inducible KSHV producer cell line of endothelial origin: maintenance of tight latency with efficient reactivation upon induction. *J Virol Methods* 174: 12-21.
136. Brulois KF, Chang H, Lee AS, Ensser A, Wong LY, et al. (2012) Construction and manipulation of a new Kaposi's sarcoma-associated herpesvirus Bacterial artificial chromosome clone. *J Virol*.
137. Wek RC, Jiang HY, Anthony TG (2006) Coping with stress: eIF2 kinases and translational control. *Biochem Soc Trans* 34: 7-11.
138. Vattam KM, Wek RC (2004) Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci U S A* 101: 11269-11274.
139. Hinnebusch AG (1997) Translational regulation of yeast GCN4. A window on factors that control initiator-trna binding to the ribosome. *J Biol Chem* 272: 21661-21664.
140. Palam LR, Baird TD, Wek RC (2011) Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance CHOP translation. *J Biol Chem* 286: 10939-10949.
141. Burysek L, Pitha PM (2001) Latently expressed human herpesvirus 8-encoded interferon regulatory factor 2 inhibits double-stranded RNA-activated protein kinase. *J Virol* 75: 2345-2352.
142. Sturzl M, Gaus D, Dirks WG, Ganem D, Jochmann R (2013) Kaposi's sarcoma-derived cell line SLK is not of endothelial origin, but is a contaminant from a known renal carcinoma cell line. *Int J Cancer* 132: 1954-1958.
143. Mounir Z, Krishnamoorthy JL, Wang S, Papadopoulou B, Campbell S, et al. (2011) Akt determines cell fate through inhibition of the PERK-eIF2alpha phosphorylation pathway. *Sci Signal* 4: ra62.
144. Schmitt E, Naveau M, Mechulam Y (2010) Eukaryotic and archaeal translation initiation factor 2: a heterotrimeric tRNA carrier. *FEBS Lett* 584: 405-412.
145. Shabman RS, Hoenen T, Groseth A, Jabado O, Binning JM, et al. (2013) An upstream open reading frame modulates ebola virus polymerase translation and virus replication. *PLoS Pathog* 9: e1003147.
146. Tischer BK, Smith GA, Osterrieder N (2010) En passant mutagenesis: a two step markerless red recombination system. *Methods Mol Biol* 634: 421-430.
147. Gale M, Jr., Tan SL, Katze MG (2000) Translational control of viral gene expression in eukaryotes. *Microbiol Mol Biol Rev* 64: 239-280.
148. Horvath CM, Williams MA, Lamb RA (1990) Eukaryotic coupled translation of tandem cistrons: identification of the influenza B virus BM2 polypeptide. *Embo J* 9: 2639-2647.
149. Powell ML, Napthine S, Jackson RJ, Brierley I, Brown TD (2008) Characterization of the termination-reinitiation strategy employed in the expression of influenza B virus BM2 protein. *Rna* 14: 2394-2406.
150. Ahmadian G, Randhawa JS, Easton AJ (2000) Expression of the ORF-2 protein of the human respiratory syncytial virus M2 gene is initiated by a ribosomal termination-dependent reinitiation mechanism. *Embo J* 19: 2681-2689.

151. Hershey JWBaM, W.C. (2000) Pathway and mechanism of initiation of protein synthesis. In *Translational Regulation of Gene Expression* (Hershey, JWB et al, eds),: 33-38.
152. Gale M, Jr., Blakely CM, Kwieciszewski B, Tan SL, Dossett M, et al. (1998) Control of PKR protein kinase by hepatitis C virus nonstructural 5A protein: molecular mechanisms of kinase regulation. *Mol Cell Biol* 18: 5208-5218.
153. Poppers J, Mulvey M, Khoo D, Mohr I (2000) Inhibition of PKR activation by the proline-rich RNA binding domain of the herpes simplex virus type 1 Us11 protein. *J Virol* 74: 11215-11221.
154. Schneider RJ, Mohr I (2003) Translation initiation and viral tricks. *Trends Biochem Sci* 28: 130-136.
155. Wethmar K, Smink JJ, Leutz A Upstream open reading frames: molecular switches in (patho)physiology. *Bioessays* 32: 885-893.