IRES-mediated ribosome repositioning directs translation of a +1 overlapping ORF that enhances viral infection

†Craig H Kerr[1,2], †Qing S Wang[1], Kyung-Mee Moon[1,2], Kathleen Keatings[3], Douglas W Allan[3], Leonard J Foster[1,2], Eric Jan*[1]

[1]Department of Biochemistry and Molecular Biology, [2]Centre for High-Throughput Biology, [3]Department of Cellular and Physiological Sciences, University of British Columbia, Vancouver BC, V6T 1Z3 Canada

†contributed equally

*corresponding author: ej@mail.ubc.ca

keywords: ribosome, frameshift, dicistrovirus, drosophila, virus

ABSTRACT

RNA structures can interact with the ribosome to alter translational reading frame maintenance and promote recoding that result in alternative protein products. Here, we show that the internal ribosome entry site (IRES) from the dicistrovirus Cricket paralysis virus drives translation of the 0-frame viral polyprotein and an overlapping +1 open reading frame, called ORFx, via a novel mechanism whereby a subset of ribosomes recruited to the IRES bypasses downstream to resume translation at the +1-frame 13th non-AUG codon. A mutant of CrPV containing a stop codon in the +1 frame ORFx sequence, yet synonymous in the 0-frame, is attenuated compared to wild-type virus in a *Drosophila* infection model, indicating the importance of +1 ORFx expression in promoting viral pathogenesis. This work demonstrates a novel programmed IRES-mediated recoding strategy to increase viral coding capacity and impact virus infection, highlighting the diversity of RNA-driven translation initiation mechanisms in eukaryotes.

31 **Significance Statement**

32 Viruses use alternate mechanisms to increase the coding capacity of their viral
33 genomes. Here, we provide biochemical evidence that ribosomes recruited to the
34 dicistrovirus cricket paralysis virus IRES undergo a bypass event to direct translation of
35 a downstream +1 frame overlapping open reading frame, called ORFx. Mutations that
36 block ORFx expression inhibit +1 frame translation and infection in fruit flies. These
37 findings highlight the diversity of RNA-driven translation initiation mechanisms in
38 eukaryotes.

39

INTRODUCTION

The ribosome mediates translation involving decoding the open reading frame codon by codon through delivery of the correct aminoacyl-tRNAs to the ribosomal A site. This fundamental process occurs with high fidelity for proper gene expression in all species. However, mechanisms exist that can alter the translational reading frame, thus producing alternative protein products from a single RNA (1). In general, these mechanisms termed recoding, involve a specific RNA structure or element that interacts with the ribosome to cause the translating ribosome to shift reading frame by -1/+1/+2 allowing it to read through stop codons or bypass sequences and restart translation downstream (2-4). Study of these mechanisms has been enlightening; revealing key ribosome:RNA interactions that alter fundamental processes in the mechanics of ribosome decoding and reading frame maintenance. Importantly, recoding mechanisms are now appreciated as important regulatory processes that can impact the fate of protein expression in cells and viral infections (1, 5). Unlike these recoding mechanisms that involve an actively translating ribosome, the intergenic internal ribosome entry site (IRES) within a subset of dicistroviruses has the unusual property to directly recruit the ribosome and initiate translation from overlapping 0 and +1-frame codons to produce two distinct proteins (6). Here, we report a novel recoding mechanism and translational initiation pathway whereby a related dicistrovirus IRES directs the ribosome to initiate translation downstream.

Most eukaryotic mRNAs utilize a cap-dependent scanning mechanism involving >12 translation initiation factors to recruit the ribosome and initiate translation from an AUG start codon (7). Alternatively, an IRES, in general, is a structured RNA

63 element that facilitates 5' end-independent translation using subsets of translation
64 initiation factors, thus providing an advantage during viral infection or under times of
65 stress when cap-dependent translation is compromised (3-5).

66 Of the different classes of viral IRESs based on factor requirement and
67 mechanism, the intergenic IRES of the *dicistroviridae* family stands out as the most
68 streamlined using a unique mechanism where it directly binds 40S and 80S ribosomes
69 without the need for canonical initiation factors or initiator Met-tRNA$_i$ and initiates
70 translation from a non-AUG codon (8-12). The dicistrovirus IRES is composed of three
71 pseudoknots (PKI, II, and III) that separate into distinct domains; PKII and PKIII fold
72 independently to create the ribosome-binding domain while PKI mediates positioning of
73 the ribosome and establishes the translational reading frame (8, 10, 13). Structural
74 studies have indicated that PKII and PKIII form a compact core structure, and the PKI
75 region adopts a conformation that mimics an anti-codon:codon interaction that binds the
76 conserved core of the ribosome in the A site (14-17). From here, in an elongation factor
77 2 dependent manner, the IGR IRES pseudo-translocates to the ribosomal P site,
78 followed by aminoacyl-tRNA delivery to the A site and a second round of eEF2-
79 dependent pseudo-translocation of the IGR IRES to the E site of the ribosome (17-19).
80 Altogether, the IGR IRES acts as a complete RNA machine that supersedes initiation
81 factors and commandeers the ribosome, a strategy that is essential for viral protein
82 synthesis in dicistrovirus-infected cells (20).

83 In general, the dicistrovirus IGR IRESs are conserved at the structural, but not
84 sequence level and are classified into two sub-groups (termed Type I and II) based on
85 the presence of distinct structural elements; the main distinction comes from a larger

86 L1.1 loop and an additional stem-loop (SLIII) in Type II IRESs (21, 22). SLIII allows for
87 the PKI domain of Type II IRESs to mimic the global shape of a tRNA in addition to
88 assisting in reading frame selection and the larger L1.1 region functions to mediate 60S
89 recruitment (15, 18, 23). The domains of Type I and II IGR IRESs function similarly to
90 directly recruit 80S ribosomes and initiate translation (24-27). Recent high resolution
91 cryo-EM structures of the IGR IRES bound to the 80S ribosome have demonstrated that
92 the IRES initially binds in the A site (17, 19, 28). Movement of the IRES involves an
93 eEF2-dependent pseudo-translocation event where the ribosome rotates up to 10°
94 allowing PKI to move into the P site in an inchworm-like manner (28). This allows for the
95 non-AUG initiation codon of the IRES to be presented in the A site for the incoming
96 amino-acyl tRNA. The first pseudo-translocation event and delivery of the first amino-
97 acyl tRNA are the rate-limiting steps of initiation on the IGR IRES (29).

98 Biochemical, phylogenetic, and bioinformatics analyses have demonstrated that
99 a subset of Type II IGR IRESs can direct translation of a hidden +1 open reading frame
100 (ORF), termed ORFx, within ORF2 of the viral genome (6, 30). The functional role of
101 ORFx during viral infection remains elusive. Extensive mutagenesis of the PKI domain
102 of the *Israeli acute paralysis virus* (IAPV) IGR IRES has revealed that 0 and +1 frame
103 translation can be uncoupled, suggesting that the IGR IRES may adopt specific
104 conformations that govern the translational reading frame (18, 31). Generally, Type I
105 and II IGR IRESs are thought to operate similarly in mechanism. Specific domains
106 between the two types are functionally interchangeable (24). In the present study, we
107 investigate the capacity of other IGR IRESs from dicistroviruses to facilitate +1-frame
108 translation. We show that the IGR IRES from Cricket paralysis virus (CrPV) can

109 synthesize an ORFx protein using an unexpected mechanism that involves IRES-
110 mediated ribosome bypassing. Furthermore, we provide insight into the role of ORFx
111 during CrPV infection and show that mutants deficient in ORFx have impaired virulence
112 in adult flies, thus uncovering a novel viral recoding strategy that is essential for viral
113 infection.

114

115

116 **RESULTS**

117 **IGR IRES-dependent +1-frame translation is not conserved throughout**
118 ***Dicistroviridae***

119 We previously showed that a subset of dicistrovirus IGR IRESs can direct
120 translation in the 0- and +1-frames and that a base pair adjacent to the PKI domain is
121 important for initiation in the +1-frame (6, 31). The IGR IRESs are classified into two
122 types: Type I, and II, with the main difference being an extra SLIII within the PKI domain
123 of Type II IRES. Since the honeybee and fire ant viruses harbour Type II IGR IRESs
124 that can support +1-frame ORFx translation, we investigated whether Type I IRESs also
125 had the capacity for +1-frame translation. We first surveyed the other dicistrovirus IGR
126 IRESs and their downstream sequences for a potential base pair adjacent to the PKI
127 domain and an overlapping +1 open reading frame both *in silico* (Fig. S1). The IAPV
128 and ABPV IGR IRESs contain an adjacent U-G base pair whereas the SINV-1 and KBV
129 IRESs have a C-G base pair (Fig. S1). All but one of the IGR IRESs contains a potential
130 base pair adjacent to the PKI domain, with the majority of them possessing a potential
131 U-G base pair. As shown previously, the honeybee and fire ant dicistrovirus ORFx
132 proteins are approximately 94-125 amino acids in length (30). The predicted +1 ORFx
133 lengths of other dicistroviruses range from 1-53 amino acids in length. Apart from the
134 honeybee and fire ant dicistrovirus ORFx, the longest putative ORFx sequences are
135 found within the genomes of the CrPV and DCV at 53 amino acids in length.

136 To determine whether the other IGR IRESs can direct +1-frame translation, we
137 cloned each IGR IRES within the intergenic region of a dual luciferase bicistronic
138 construct (Fig. 1A). The Renilla luciferase (36 kDa; RLuc) monitors scanning-dependent

139 translation whereas firefly luciferase (FLuc) is translated by the IRES. To measure
140 reading frame selection, the firefly luciferase gene is fused in-frame either in the 0 or +1-
141 frame and translated proteins were monitored by incorporation of [^{35}S]-
142 methionine/cysteine. Additionally, we generated bicistronic reporter constructs that
143 contain a T2A 'stop-go' sequence, which allows quantitation of luciferase activity (32)
144 (Fig. S2). Since we previously showed robust IAPV IRES 0 and +1-frame translation (6),
145 we chose it as a benchmark to compare the +1-frame activity of the other IGR IRESs. In
146 general, all IGR IRESs can direct 0-frame translation to varying extents *in vitro*.
147 Normalizing to the IAPV 0-frame translational activity, the CrPV IGR IRES showed the
148 highest 0-frame translational activity (~170%) whereas the Mud crab dicistrovirus
149 (MCDV) IGR IRES had the lowest activity (~4%; Fig. 1A). By contrast, only a few IGR
150 IRESs can support +1-frame translation. The Type II IGR IRESs from Taura syndrome
151 virus (TSV) and MCDV did not support +1 frame translation suggesting that only a
152 subset of Type II IGR IRESs can direct 0- and +1-frame translation. Interestingly,
153 besides the IAPV IGR IRES, the BQCV and CrPV IRESs mediated +1-frame translation
154 above background levels, 80% and 5% of 0-frame translation, respectively. In summary,
155 only a subset of IGR IRESs can facilitate both 0 and +1-frame translation.
156

157 **CrPV +1-frame translation is IGR IRES-dependent and initiates downstream**

158 To explore +1-frame translation mechanisms further, we focused on the CrPV
159 IRES (Fig. 1B). We first determined whether the structural integrity of the CrPV PKI
160 domain is important for +1-frame translation. For these assays, we used a bicistronic
161 reporter construct that contains the CrPV IRES where a firefly luciferase gene was

162 subcloned into the +1-frame downstream of the IRES with CrPV nucleotides 6217-6387
163 (Accession KP974707.1), which includes the predicted +1-frame CrPV ORFx (65 kDa;
164 ORFx-FLuc). This dual luciferase reporter construct allows simultaneous monitoring of
165 scanning-dependent and CrPV IRES-mediated 0/+1-frame translation as a shortened 0-
166 frame protein (~11 kD) is also translated in addition to the +1-frame (Fig. 2A). Synthesis
167 of all three proteins is detected by incubating the bicistronic construct in a Sf21
168 translation extract in the presence of [^{35}S]-met/cys (Fig. 2B, lane 1) (6). As shown in Fig.
169 1A, CrPV IRES +1-frame ORFx translation is approximately 5% of 0-frame translation.
170 Mutating CC_{6214-5} to GG, which disrupts PKI base pairing and abolishes CrPV IRES
171 activity, resulted in negligible or diminished 0- and +1-frame translation whereas a
172 compensatory mutation that restores PKI base pairing rescued translation (10) (Fig. 2B,
173 lanes 2 and 3, Fig. S2), indicating that the integrity of the IRES and the PKI domain is
174 required for CrPV IRES +1-frame translation.

175 The adjacent U-G base pairing of the IAPV IRES is important for +1-frame
176 translation (6). CrPV also has the capacity to form a wobble base pair adjacent to the
177 PKI domain through nucleotides U_{6186} and G_{6217}, potentially directing translation from
178 the first +1-frame CUA leucine codon (Fig. 1B). To determine if this base pair is
179 necessary to drive CrPV IRES +1-frame translation, we mutated U_{6186} and G_{6217} to other
180 bases. Mutating U_{6186} to C or G led to an approximate 18%-23% reduction in +1 frame
181 activity and mutating G_{6217} to C, A, or U resulted in roughly a 39%, 24%, and 36%
182 reduction in +1-frame activity, respectively. Although each mutation reduced +1-frame
183 translational activity to some degree, none of the mutations abolished it (Fig. 2B, lanes
184 4-9). These results suggest that unlike with IAPV, base pairing between nucleotides

185 6186 and 6217 is not absolutely required for CrPV IRES +1-frame translation.

186 To determine the potential initiation site of CrPV ORFx, we systematically
187 replaced codons downstream of the IRES with a stop codon and monitored 0- and +1-
188 frame translation *in vitro* using the bicistronic reporter construct (Fig. 2B, 2C). Overall,
189 stop codons placed in the +1-frame did not significantly affect 0-frame translation,
190 indicating that IRES activity was not compromised (Fig. 2B, 2C). Replacing individual
191 codons between the 1st and the 12th +1-frame codon with a stop codon inhibited to
192 varying extents (between 36%-71% reduction compared to wild-type) but did not
193 completely abolish +1 frame translation (Fig. 2B, lanes 11-16; 2C). Conversely, +1-
194 frame translation was completely inhibited when the 13th +1-frame codon and codons
195 thereafter were replaced with a stop codon (Fig. 2B, lanes 15-16; 2C). Replacing both
196 the 3rd and 5th +1-frame codons with stop codons reduced +1-frame translation by 33%
197 but did not eliminate it, suggesting ribosome read-through did not occur. To address the
198 possibility that IRES translation initiates in the 0-frame and then a fraction of translating
199 ribosomes shift into the +1-frame, we inserted a stop codon in the 0-frame downstream
200 of the IRES. As expected, a stop codon in the 0-frame 1st to 4th codons downstream of
201 the IRES abolished 0-frame translation (Fig. 2B, lanes 18-21). However, the 0-frame
202 stop codon insertions reduced by 52-77% but did not eliminate +1 frame translation
203 indicating that ribosomes likely do not shift from the 0 to +1 reading frame. Furthermore,
204 introducing stop codons in both the 0 and +1 frames did not abolish +1-frame translation,
205 though inhibited +1-frame translation by ~80% (Fig. 2C, 0 S4/+1 S5). We noted that the
206 adjacent 14th codon is an AUG methionine. Mutating the AUG to a CGG (M14R) or
207 GUG (M14V) decreased but did not abolish +1-frame translation (Fig. 2C). Mutating the

208 13th AAA codon to UUU, AAG or GCU did not abolish +1-frame translation (Fig. 2C),
209 suggesting that there is flexibility in the codon identity for CrPV ORFx +1-frame
210 translation. In summary, the mutational analysis indicated that CrPV ORFx translation
211 requires an intact IRES and a +1-frame 13th nonstop codon.
212
213 **CrPV +1-frame translation requires 80S assembly and is edeine-insensitive**
214 The current data led us to generate two hypotheses: i) a subset of 40S subunits
215 recruited to the CrPV IRES scan downstream to start translation at the 13th codon
216 (scanning hypothesis) or ii) 40S or 80S ribosomes recruited to the IGR IRES bypass the
217 spacer region to the 13th codon (bypass hypothesis). To address the scanning
218 hypothesis, we took two approaches. First, we utilized mutants of the CrPV IRES in the
219 L1.1 loop that is known to be deficient in recruitment of the 60S subunit (15, 23). If
220 scanning is occurring, 40S subunits recruited to the L1.1 IRES may scan downstream to
221 the downstream +1-frame initiation codon. Reporter constructs harbouring mutations in
222 the L1.1 loop were deficient in 0 and +1-frame translation (Fig. 3A), suggesting that 60S
223 recruitment by the IGR IRES specifically is necessary for translation in the +1-frame.
224 Secondly, we utilized the translational inhibitor edeine to assess if scanning was
225 occurring (Fig. 3B). Edeine prevents the 40S ribosomal subunit from recognizing an
226 AUG start codon (33, 34). Both 0- and +1-frame translation were resistant to edeine
227 relative to that of scanning-dependent translation (Fig. 3B), thus suggesting that 40S
228 scanning is not involved in +1 frame translation.
229
230 **The integrity of the variable- loop region and pseudo-translocation of the IGR**

231 **IRES through the ribosome is required for ORFx expression**

232 If 80S ribosomes recruited to the IRES are indeed repositioning or "bypassing" to
233 the downstream 13th +1-frame codon, we next looked to investigate the rules which
234 govern this potential mechanism. Since the IGR IRES, which occupies the ribosomal A
235 site upon ribosome binding, undergoes pseudotranslocation to the ribosomal P site to
236 vacate the A site for delivery of the first aminoacyl-tRNA (17), we reasoned that the
237 ribosome bound to the CrPV IGR IRES must have an empty ribosomal P or A site in
238 order to reposition downstream and accommodate the +1 frame start codon. To address
239 this, we introduced mutations in the variable loop region (VLR), which has been shown
240 to disrupt the IGR IRES-mediated pseudotranslocation event (Fig. 1B) (35). Specifically,
241 shortening the length of the VLR by 2 or 3 nucleotides (Δ2 and Δ3, respectively) inhibits
242 the first pseudo-translocation event from the A site to the P site whereas altering the
243 identity of nucleotides A_{6204}, and $AA_{6208\text{-}9}$ to guanosines (G-rich) inhibited IRES
244 translocation from the P site to the E site (35). Interestingly, all three VLR mutants
245 decreased +1-frame activity (Fig. 4); both the G-rich and Δ3 mutants demonstrated little
246 to no activity, while the Δ2 mutant still exhibited roughly 50% activity to that of WT. This
247 result is consistent with previous data that the Δ2 mutant IRESs are still able to
248 accommodate a fraction of aminoacyl-tRNA in the A site (~25%), allowing translation to
249 occur (35). Altogether, these results indicate that the pseudotranslocation event of the
250 IGR IRES through the ribosome contributes to +1-frame translation downstream.

251

252 **The spacer region downstream of the IGR IRES is necessary for +1 frame**
253 **translation**

254 We next investigated whether the spacer region located between PKI of the IGR
255 IRES and the downstream AAA codon contributes to IRES-mediated ribosome bypass.
256 Our results showed that inserting stop codons in the 'spacer region' (Fig. 2A) between
257 the IRES PKI domain and the +1-frame 13th codon of ORFx inhibited but did not
258 completely abolish +1-frame translation, suggesting that an element within this spacer
259 region may promote +1-frame translation. We first addressed whether the CrPV spacer
260 region is sufficient to direct +1-frame translation by generating a chimeric construct
261 whereby the PSIV IRES is fused with the CrPV spacer region (Fig. 5). The PSIV IRES
262 directs strong 0-frame but no or relatively weak +1-frame translation (Fig. 1A). The
263 PSIV-CrPV spacer chimeric reporter resulted in 0-frame translation (Fig. 5A), indicating
264 that the spacer region does not affect the activity of the PSIV IRES. In contrast to the
265 PSIV IRES construct, the chimeric PSIV-CrPV reporter resulted in +1-frame translation,
266 implying that the CrPV spacer region is sufficient to drive +1-frame translation in the
267 presence of a functional IGR IRES (Fig. 5A).
268 To delineate whether there is a specific element within the spacer region that is
269 required for CrPV +1 frame translation, we systematically deleted from the 3' end of the
270 spacer region (Fig. 5B). Interestingly, deleting 3 to 27 nucleotides did not affect +1
271 frame translation. However, deleting 30 nucleotides and leaving seven nucleotides
272 adjacent to the IRES abolished +1 translation (Fig. 5B). Δ3, Δ6 and Δ27 mutants appear
273 to have much higher Fluc/Rluc activity when compared to WT, an observation that
274 requires further examination. In summary, these results suggest that the specific
275 sequences and its context within the spacer region, particularly the sequences
276 immediately downstream of the PKI domain, are important for mediating IRES-

277 dependent +1 frame translation.

278

279 **+2-frame translation mediated by the CrPV IRES**

280 We hypothesized that if the ribosome is indeed bypassing to a specific 'landing
281 site' then it should be independent of frame, whereas if the ribosome merely begins
282 translation in the 0-frame before slipping into the +1-frame, then we expect to see no
283 ORFx translation. To address this, we inserted a series of nucleotides into our
284 bicistronic construct that shifts only the ORFx-Fluc into the +2-frame (Figure S3).
285 Specifically, we inserted either a single nucleotide or up to 7 proceeding the 13th AAA
286 codon. As a control, we inserted either 6 or 9 nucleotides in the same position that does
287 not introduce an additional frameshift. Insertion of a U creates a stop codon in the +1-
288 frame while an inserted C does not. To our surprise, we observed ORFx expression
289 with all insertions (Figure S3). These results suggest translation of ORFx is reading
290 frame-independent and indicates that the ribosome may be repositioned from the IRES
291 to the downstream 13th codon.

292

293 **CrPV +1-frame ORFx is expressed yet not required for infection in *Drosophila* S2**
294 **cells**

295 Our *in silico* and biochemical data indicate that ribosomes recruited to the CrPV
296 IRES may bypass downstream to the +1-frame translation from the 13th codon. The
297 CrPV ORFx is predicted to be 41 amino acids in length if ORFx is translated from the
298 +1-frame 13th codon (Fig. 6). To determine whether CrPV ORFx is synthesized during
299 infection, *Drosophila* S2 cells infected with CrPV (MOI 10) were harvested at 6 hours

300 post infection and lysed. Proteins were subsequently digested with trypsin and peptides
301 were analyzed by LC-MS/MS. We identified two peptides that correspond to CrPV
302 ORFx both of which were located downstream of the +1-frame 13th codon (Fig. 6).
303 Importantly, these peptides were not identified in mock-infected S2 cells. Thus, ORFx is
304 expressed in CrPV-infected S2 cells.

305 Given this, we sought to determine the influence of CrPV ORFx expression on
306 the outcome of viral infection. Using a recently developed CrPV infectious clone, termed
307 CrPV-2, we introduced mutations that would abolish ORFx expression (36). To this end,
308 we created two separate mutant clones: the +1-frame 12th codon (UUG_{6251-3}) was
309 altered to an amber stop codon (UAG; CrPV-S12) and the +1-frame 19th codon
310 (UUA_{6272-4}) changed to an ochre stop codon (UAA; CrPV-S19) (Fig. 7A). Both mutations
311 are synonymous in the 0-frame. Based on our translation data (Fig. 2), the +1-frame
312 S19 but not the S12 would inhibit ORFx expression.

313 First, we examined whether ORFx influences viral protein synthesis *in vitro*.
314 Incubation of *in vitro* transcribed CrPV-2 RNA in Sf21 translation extracts led to
315 synthesis and processing of all viral proteins as reported previously (Fig. S4A) (36).
316 While stop codons in the 0-frame of ORF1 or ORF2 inhibited synthesis of viral proteins,
317 both CrPV-S12 and CrPV-S19 RNAs resulted in viral protein synthesis that was
318 indistinguishable compared to wild-type CrPV -2 *in vitro*, demonstrating that CrPV ORFx
319 is not necessary for viral protein synthesis *in vitro* (Fig. S4B) (36).

320 We next assessed the viability of the CrPV-S12 and -S19 viruses in cell culture.
321 Harvested wild-type, mutant S12 or S19 CrPV-2 virus were used to infect naïve S2 cells
322 at a MOI 10 and 1 to follow the first round of infection and subsequent rounds of

323 infection, respectively. Infection with wild-type CrPV-2, CrPV-S12 or CrPV-S19 all
324 resulted in accumulation of viral proteins and RNA and shutdown of host translation in a
325 similar manner (Fig. S4C). Similarly, neither infection produced significantly different
326 titres between wild type CrPV-2 and either CrPV-S12 or CrPV-S19 at any time point,
327 apart from both mutant viruses resulting in higher titres than CrPV-2 after 24 hours post
328 infection (Fig. S4D). A similar result was observed with infections at a MOI 0.1 (data not
329 shown). Taken together, ORFx has no observable effect on the life cycle of CrPV in S2
330 cells.
331

332 **ORFx contributes to CrPV infection in adult flies and associates with membranes**

333 We addressed whether ORFx contributes to CrPV infection in adult fruit flies. To
334 test this, we injected adult flies intrathoracically with PBS, CrPV-2, CrPV-S12, or CrPV-
335 S19 and monitored fly mortality daily. Flies injected with CrPV-2 or CrPV-S12 exhibited
336 50% mortality by day 5 and 100% mortality by day 11 and 12 (Fig. 7B). By contrast, flies
337 injected with CrPV-S19 did not reach 50% mortality until day 9 and reached 100%
338 mortality at day 14 (Fig. 7B). These results demonstrate that ORFx contributes to CrPV
339 pathogenesis in adult flies.

340 To determine if the effect seen on CrPV pathogenesis is a result of defects in
341 viral replication, we measured viral titres and assessed viral protein levels in injected
342 adult flies at 5 days post infection. Interestingly, viral titres and viral protein levels
343 showed no significant differences between wild type CrPV-2, CrPV-S12, and CrPV-S19
344 (Fig. S5). Finally, using RT-PCR followed by sequencing, the S12 and S19 mutations
345 are stable during virus propagation in S2 cells (data not shown). In summary, our results

346 indicate that the defect in viral pathogenesis observed in the CrPV-S19 mutant virus is
347 not due to a defect in viral replication.

348 From our data, CrPV ORFx is predicted to be a 41 amino acid protein, mediated
349 by a ribosomes bypass mechanism from the 13th +1-frame AAA codon. Comparing the
350 sequence of ORFx from CrPV and other species shows no appreciable homology to
351 other proteins. Nevertheless, *in silico* topology predictions suggest that ORFx can adopt
352 an alpha helical transmembrane segment at its C-terminus (amino acids 22-39; Fig.
353 S6A). To examine ORFx function, we generated constructs containing either N- or C-
354 terminal HA-tagged ORFx. Transfection for 10 or 24 hours with the HA-tagged ORFx
355 constructs did not result in a dramatic decrease in S2 cell viability as measured by a
356 trypan blue exclusion assay (Fig. S6B). However, by 48 hours, transfection of ORFx-HA
357 and HA-ORFx led to a relatively minor reduction (13% and 15% decrease, respectively)
358 in cell viability, suggesting that ORFx expression is slightly toxic in S2 cells.
359 Immunoblotting for HA showed that HA-ORFx is expressed in S2 cells (Fig. S6C). To
360 determine the subcellular localization in S2 cells, we transfected the HA-tagged ORFx
361 constructs and monitored ORFx localization (pORFx-HA) by HA-antibody
362 immunofluorescence staining in comparison with cytoplasmic, ER, Golgi, and nuclear
363 marker protein antibodies. We also mutated two pairs of amino acids, LV and LI, to KK
364 and KR respectively, to disrupt the transmembrane domain (pORFx-TMmut-HA). Co-
365 staining showed that the wild-type HA-tagged ORFx overlaps mainly with ER protein
366 marker Calnexin, and partially overlaps with Golgi-associated protein, Golgin84 (Fig. 8,
367 S7, pORFx-HA). The wild-type HA-ORFx displayed little to no overlap with α-Tubulin
368 staining. By contrast, the mutant transmembrane HA-tagged ORFx staining showed no

369 overlap with Calnexin or Golgin84 (Fig. 8, S7, pORFx-TmMut-HA). Collectively, these
370 results suggest that ORFx associates with membranous organelles and potentially the
371 ER specifically. To examine this further, we used differential centrifugation to separate
372 subcellular components followed by immunoblotting of HA. As expected, tubulin is found
373 in the cytoplasmic fraction and the ER-associated KDEL protein and cytochrome C are
374 enriched in nuclei and ER fractions (Fig. S6C). Both HA-tagged ORFx are detected
375 within membranous fractions but not within the cytoplasmic fractions (Fig. S6C), further
376 supporting that ORFx resides within the membrane of cells that may be important for its
377 function.
378
379

380 **DISCUSSION**

381 Recoding mechanisms have illuminated diverse RNA structural elements that
382 interact with the ribosome to affect reading frame maintenance (2). In this study, we
383 have demonstrated a novel translation recoding mechanism by which an IRES may
384 promote ribosome repositioning at a downstream codon. In contrast to the IAPV IGR
385 IRES, which directs ribosome reading frame selection (6), the CrPV IGR IRES can
386 facilitate the expression of a downstream +1 overlapping frame, which we termed ORFx.
387 We provide extensive analysis using mutagenesis that translation of CrPV ORFx likely
388 occurs 37 nucleotides downstream at the 13th AAA (Lys) codon. Moreover, we show
389 that ORFx is expressed in CrPV-infected cells by mass spectrometry and that ORFx is
390 required for promoting CrPV pathogenesis in a *Drosophila* injection model. Our data
391 suggest a model whereby after 80S assembly on the CrPV IGR IRES, the majority of
392 the ribosomes translate in the 0-frame by delivery of the incoming Ala-tRNA^{Ala} to the 0-
393 frame GCU codon whereas a fraction of ribosomes bypasses or "slides" 37 nucleotides
394 downstream to direct translation at the +1-frame AAA (Lys) codon (Fig. S8). From our
395 results, the following rules appear to apply to +1-frame translation in CrPV: i) the PKI
396 must be intact, ii) both 40S and 60S ribosomal subunits are required to bind to the IRES,
397 iii) pseudotranslocation of the IRES through the ribosome is necessary, iv) the spacer
398 region between PKI and the 13th AAA codon is essential, and v) the nucleotide identity
399 of the spacer region is crucial for efficient +1 frame translation.

400 A unique feature of this repositioning mechanism is that an intact IRES is
401 essential for +1-frame translation (Fig. 2). The current model is that the PKI domain of
402 the IRES occupies the ribosomal A site upon ribosome binding to the IRES, followed by
403 a pseudotranslocation event in order to vacate the A site to allow delivery of the first

404 aminoacyl-tRNA in the 0-frame (17). Given this model, it is difficult to envision how the
405 IRES can direct the ribosome to bypass 37 nucleotides to initiate translation at a +1-
406 frame AAA Lys codon. A potential clue comes from our mutagenesis analysis that
407 suggests that translocation of the CrPV IRES through the ribosome is a prerequisite for
408 downstream +1-frame translation (Fig. 4). This is intuitive as the PKI domain must
409 vacate the A site in order to allow translation in both the 0 and +1 frames. At this point, it
410 is unclear whether the CrPV IGR IRES ribosome repositioning event requires delivery of
411 an aminoacyl-tRNA to the A site of the ribosome prior to the ribosome sliding
412 downstream to initiate translation at the 13th +1 frame codon or whether a vacant
413 ribosome bypasses from the IRES to the initiating +1 frame codon. Determination of the
414 aminoacyl-tRNA(s) that are delivered during the initial CrPV IGR IRES
415 pseudotranslocation event should provide insights whether the ribosomal A site is
416 vacant prior to ribosome sliding downstream to the 13th +1 frame codon.
417 A recent *in vitro* study using single molecule fluorescence spectroscopy
418 demonstrated that the CrPV IGR IRES can facilitate +1-frame translation approximately
419 5% of the time compared to 0-frame translation whereby reading frame selection is
420 dictated by the kinetics of tRNA binding in the first 0- or +1-frame codon (12). Using a
421 more physiologically-related system, our results suggest that the CrPV IRES directs +1-
422 frame translation using a different mechanism. Based on our systematic stop codon
423 mutational analyses, the most likely scenario is that ribosomes recruited to the IRES
424 must bypass downstream to translate in the +1-frame (Fig. 2). In support of this,
425 insertion of a stop codon at the 12th +1-frame codon, but not the 19th, attenuates CrPV-
426 mediated death in a Drosophila injection model, thus providing biological significance of

427 CrPV +1-frame translation (Fig. 7). We also note that the downstream authentic ‘spacer’
428 sequence is necessary for +1-frame translation (Fig. 5), which is partially absent and/or
429 altered from previous studies, which may preclude detection of bypass (12).

430 The downstream spacer region appears to be a key feature that is necessary for
431 efficient CrPV IRES-mediated +1-frame synthesis (Fig. 5). Interestingly, the majority of
432 mutations within the spacer region causes a reduction ranging from 10-80% in the
433 amount of CrPV IRES-mediated +1-frame synthesis (Fig. 2), suggesting that nucleotide
434 or codon identity is crucial for triggering the bypass event. There are no obvious RNA
435 secondary structures within the spacer region, however, we cannot rule out any long
436 distance RNA:RNA interactions that may contribute to +1-frame translation.
437 Furthermore, sequences immediately downstream of the IRES PKI domain are
438 important for +1 frame translation (Fig. 5B). Investigation into how the spacer sequence
439 influences ribosomes bound to the IRES to start translation in the 0-frame or at the
440 downstream +1-frame initiating codon should shed light onto this atyptical ribosome
441 bypass mechanism.

442 How does the ribosome reposition to the downstream +1-frame initiation codon
443 after IRES binding? It is possible that the repositioning of the ribosomes occurs via
444 mechanism similar to that observed with prokaryotic 70S 'scanning' (37) or 70S 'sliding'
445 that occurs in coupled translational reinitiation (38, 39) (Fig. S8). Indeed, a related
446 dicistrovirus PSIV IRES can direct translation using a scanning-like mechanism in
447 prokaryotes, which may suggest a similar property observed with our studies on the
448 CrPV IRES (40). Moreover, it has been reported and proposed that energy-independent
449 scanning or diffusion of the ribosome or ribosomal subunits (i.e. phaseless wandering)

450 can occur to locate an AUG codon (39, 41-43)(Fig. S8). Nevertheless, our study shows
451 that CrPV IRES +1-frame translation requires 80S ribosome binding to the IRES (Fig.
452 3A) and is edeine-insensitive (Fig. 3B), thus we favour a model that 80S ribosomes
453 reposition to the 13th +1-frame codon. This warrants comparison with the translational
454 bypassing observed in *gene 60* of T4 bacteriophage (44). In *gene 60*, translating
455 ribosomes stall in a non-canonical rotated state at a 'take-off' Gly codon with a peptidyl-
456 $tRNA^{Gly}$, which dissociates from the anticodon, and 'lands' at a matching Gly codon 50
457 nucleotides downstream and allowing translation to resume. This bypass mechanism is
458 dependent on a post-translocation step requiring a downstream 5' ‘take-off’ stem-loop
459 structure and a nascent translated peptide (45, 46). Although CrPV does not have an
460 obvious RNA structure within the spacer region, it is possible that the highly complex
461 structure of the IGR IRES itself may contribute to downstream +1 frame translation,
462 especially given its dynamic nature during movement through the ribosome (Fig. S8)(28,
463 47). Moreover, it is known that the ribosome bound to the IGR IRES is in a rotated state
464 and that the first pseudotranslocation step is rate limiting, which may contribute to CrPV
465 IRES reading frame selection (17, 29). Finally, it is possible is that CrPV IRES bypass
466 could be occurring through an RNA looping event (48); the downstream RNA is brought
467 into close proximity with the 80S ribosome allowing it to transition to 13th AAA codon
468 potentially by an unknown protein factor or a long-range RNA:RNA interaction. Taken
469 together, it is likely that it is a combination of tRNA kinetics and conformational changes
470 of the IRES and the downstream spacer region that lead to bypassing, of which the
471 contributions of each element require further investigation.

472 The biological relevance of CrPV +1-frame translation was initially evidenced by

473 the detection of ORFx peptides in CrPV-infected S2 cells (Fig. 6). To our surprise,
474 disruption of ORFx synthesis by stop codon insertion in the +1-frame did not perturb
475 viral infection in tissue culture cells but showed retarded mortality in adult flies even
476 though viral load remained similar between wild type and mutant viruses (Fig. 7). CrPV
477 infection is thought to cause death through paralysis, subsequently leading to
478 dehydration or starvation of the host (49, 50). CrPV can infect several tissues in the fly
479 including the trachea, midgut, and central nervous system although the latter has not
480 been demonstrated directly (48, 50, 51). How ORFx may contribute to CrPV
481 pathogenesis is an outstanding question. Our results indicate that ORFx is membrane
482 associated (Fig. 8, S7) and does not contribute directly to viral replication but rather to
483 the pathogenesis of CrPV infection in fruit flies (Fig. 7, S4). Furthermore, expression of
484 ORFx is slightly cytotoxic in Drosophila cells, a property that may also contribute to
485 pathogenesis of CrPV infection. Future studies into the localization of ORFx, potential
486 interacting partners, and tissue tropism of wild-type versus mutant virus infection in the
487 fly should provide insights into the role of ORFx.

488 Viruses continue to surprise us with their ability to manipulate the ribosome in
489 remarkable ways. Here, in addition to the previous findings on the honey bee
490 dicistrovirus IRES, we have revealed another recoding mechanism utilizing an IRES,
491 thus highlighting the strong selection to increase the coding capacity of the dicistrovirus
492 genome. Furthermore, an IRES that can direct ribosome repositioning to facilitate the
493 translation of a hidden +1 overlapping ORFx adds to the growing list of diverse
494 pathways of ribosome translational recoding. It will be of considerable interest to
495 investigate whether other IRESs direct reading frame selection by a similar CrPV IGR

496 IRES-mediated ribosome repositioning mechanism. Ribosome repositioning or bypass
497 is not specific to bacteria or mitochondria (52, 53) but may be a more general
498 phenomenon in eukaryotes that initially thought.

499

500

501 **MATERIALS AND METHODS**

502 **Cell culture and virus.** *Drosophila* Schneider line 2 (S2; Invitrogen) cells were
503 maintained and passaged in Shield's and Sang medium (Sigma) supplemented with 10%
504 fetal bovine serum.

505 Propagation of CrPV in *Drosophila* S2 cells has been previously described (54).
506 CrPV-2 and mutant viruses were generated from *Drosophila* S2 cells using an adapted
507 protocol (55). Briefly, 5.0 x 10^7 S2 cells were transfected with *in vitro* transcribed RNA
508 derived from pCrPV-3 or mutant plasmids and incubated for 48h. Cells were dislodged
509 into the media, treated with 0.5% Igepal CA-630 (Nonidet P-40) and 0.1% 2-
510 mercaptoethanol, and incubated on ice for 10 min. Cell debris was cleared by
511 centrifugation at 13,800 RCF for 15 min at 4°C. Viral particles were then concentrated
512 by ultracentrifugation at 141,000 RCF for 2.5 h at 4°C. The pellet was resuspended in
513 PBS and sterilized through a 0.2 µM filter. Viral titres and yield were determined as
514 previously described (20). All viruses were sequence verified via RT-PCR with primers
515 directed against the CrPV IGR IRES.

516 **LC-MS/MS Analysis**. Cell pellets harvested from CrPV-infected cells at 6 hpi were
517 solubilized in 1% sodium deoxy cholate and 50 mM NH_4HCO_3. Protein concentrations
518 were determined via BCA assay (Thermo). Proteins (100 µg) were reduced (2 µg DTT,
519 37°C, 30 min) and alkylated (5 µg iodoacetamide, RT, 20 min). Samples were digested
520 with trypsin overnight at room temperature. Peptides were acidified with 1% TFA to pH
521 <2.5 and the precipitated deoxycholate was remove via centrifugation. Peptides were
522 desalted and concentrated on C18 STAGE-tips, eluted in 80% acetonitrile, 0.5% acetic
523 acid, and dried in a vacuum concentrator (Eppendorf)(Rappsilber, Ishihama, & Mann,

524 2003). Samples were resuspended in 20% acetonitrile and 0.1% formic acid before
525 loading on an Agilent 6550 mass spectrometer.
526 Data were searched using MaxQuant (v1.5.3.30)(56). Parameters included:
527 carbamidomethylated cysteine (fixed), methionine oxidation (variable), glutamine and
528 asparagine deamidation (variable), and protein N-terminal acetylation (variable); trypsin
529 specific; maximum 2 missed cleavages; 10 ppm precursor mass tolerance; 0.05 Da
530 fragment mass tolerance; 1% FDR; +1 to +7 charge states; and common contaminants
531 were included. Both the *Drosophila* and CrPV protein databases used were the most
532 recent annotations downloaded from UniProt (www.uniprot.org).
533 **Fly stocks and viral injections.** Flies (*Isogenic* w^{1118}; Bloomington Drosophila Stock
534 Center) were maintained on standard cornmeal food at 25°C and 70% humidity with a
535 12 h light-dark cycle. Freshly eclosed virgin males and females were separated and
536 collected in groups of 10 each. Flies (10 males and 10 females) were injected with 200
537 nL of PBS, CrPV-2, CrPV-S12, or CrPV-S19 (5000 FFU) using a PV830 PicoPump
538 (World Precision Instruments) and transferred to standard food. Mortality was monitored
539 daily.
540
541

ACKNOWLEDGEMENTS

We thank John Atkins and the Jan lab for insightful discussions. Kathy Pan and Seonghoon Lee contributed a subset of data in Fig. 1A and 2B, respectively. This study was supported by a Canadian Institutes of Health Research (PJT-148761) to E.J. and to D.A. (MOP-130517), an NSERC Discovery Grant to L.F. and an NSERC CGS-D fellowship to C.K.

549 **REFERENCES**

550

551 1. Baranov PV, Atkins JF, & Yordanova MM (2015) Augmented genetic decoding: global,
552 local and temporal alterations of decoding processes and codon meaning. *Nature*
553 *reviews. Genetics* 16(9):517-529.
554 2. Dinman JD (2012) Mechanisms and implications of programmed translational
555 frameshifting. *Wiley interdisciplinary reviews. RNA* 3(5):661-673.
556 3. Firth AE & Brierley I (2012) Non-canonical translation in RNA viruses. *J Gen Virol* 93(Pt
557 7):1385-1409.
558 4. Jan E, Mohr I, & Walsh D (2016) A Cap-to-Tail Guide to mRNA Translation Strategies in
559 Virus-Infected Cells. *Annual review of virology* 3(1):283-307.
560 5. Au HH & Jan E (2014) Novel viral translation strategies. *Wiley interdisciplinary reviews.*
561 *RNA* 5(6):779-801.
562 6. Ren Q*, et al.* (2012) Alternative reading frame selection mediated by a tRNA-like domain
563 of an internal ribosome entry site. *Proc Natl Acad Sci U S A* 109(11):E630-639.
564 7. Jackson RJ, Hellen CU, & Pestova TV (2010) The mechanism of eukaryotic translation
565 initiation and principles of its regulation. *Nature reviews. Molecular cell biology*
566 11(2):113-127.
567 8. Wilson JE, Pestova TV, Hellen CU, & Sarnow P (2000) Initiation of protein synthesis
568 from the A site of the ribosome. *Cell* 102(4):511-520.
569 9. Jan E, Kinzy TG, & Sarnow P (2003) Divergent tRNA-like element supports initiation,
570 elongation, and termination of protein biosynthesis. *Proc Natl Acad Sci U S A*
571 100(26):15410-15415.
572 10. Jan E & Sarnow P (2002) Factorless ribosome assembly on the internal ribosome entry
573 site of cricket paralysis virus. *J Mol Biol* 324(5):889-902.
574 11. Pestova TV & Hellen CU (2003) Translation elongation after assembly of ribosomes on
575 the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator
576 tRNA. *Genes Dev* 17(2):181-186.
577 12. Petrov A, Grosely R, Chen J, O'Leary SE, & Puglisi JD (2016) Multiple Parallel Pathways
578 of Translation Initiation on the CrPV IRES. *Mol Cell* 62(1):92-103.
579 13. Nishiyama T*, et al.* (2003) Structural elements in the internal ribosome entry site of
580 Plautia stali intestine virus responsible for binding with ribosomes. *Nucleic Acids Res*
581 31(9):2434-2442.
582 14. Costantino DA, Pfingsten JS, Rambo RP, & Kieft JS (2008) tRNA-mRNA mimicry drives
583 translation initiation from a viral IRES. *Nat Struct Mol Biol* 15(1):57-64.
584 15. Pfingsten JS, Costantino DA, & Kieft JS (2006) Structural basis for ribosome recruitment
585 and manipulation by a viral IRES RNA. *Science* 314(5804):1450-1454.
586 16. Schuler M*, et al.* (2006) Structure of the ribosome-bound cricket paralysis virus IRES
587 RNA. *Nat Struct Mol Biol* 13(12):1092-1096.
588 17. Fernandez IS, Bai XC, Murshudov G, Scheres SH, & Ramakrishnan V (2014) Initiation
589 of Translation by Cricket Paralysis Virus IRES Requires Its Translocation in the
590 Ribosome. *Cell*.
591 18. Au HH*, et al.* (2015) Global shape mimicry of tRNA within a viral internal ribosome entry
592 site mediates translational reading frame selection. *Proc Natl Acad Sci U S A*
593 112(47):E6446-6455.
594 19. Muhs M*, et al.* (2015) Cryo-EM of ribosomal 80S complexes with termination factors
595 reveals the translocated cricket paralysis virus IRES. *Mol Cell* 57(3):422-432.
596 20. Kerr CH, Ma ZW, Jang CJ, Thompson SR, & Jan E (2016) Molecular analysis of the
597 factorless internal ribosome entry site in Cricket Paralysis virus infection. *Scientific*
598 *reports* 6:37319.

599 21. Nakashima N & Uchiumi T (2009) Functional analysis of structural motifs in
600 dicistroviruses. *Virus Res* 139(2):137-147.
601 22. Jan E (2006) Divergent IRES elements in invertebrates. *Virus Res* 119 (1):16-28.

602 23. Jang CJ, Lo MC, & Jan E (2009) Conserved element of the dicistrovirus IGR IRES that
603 mimics an E-site tRNA/ribosome interaction mediates multiple functions. *J Mol Biol*
604 387(1):42-58.
605 24. Jang CJ & Jan E (2010) Modular domains of the Dicistroviridae intergenic internal
606 ribosome entry site. *RNA* 16(6):1182-1195.
607 25. Hatakeyama Y, Shibuya N, Nishiyama T, & Nakashima N (2004) Structural variant of the
608 intergenic internal ribosome entry site elements in dicistroviruses and computational
609 search for their counterparts. *RNA* 10(5):779-786.
610 26. Cevallos RC & Sarnow P (2005) Factor-independent assembly of elongation-competent
611 ribosomes by an internal ribosome entry site located in an RNA virus that infects
612 penaeid shrimp. *J Virol* 79(2):677-683.
613 27. Hertz MI & Thompson SR (2011) In vivo functional analysis of the Dicistroviridae
614 intergenic region internal ribosome entry sites. *Nucleic Acids Res* 39(16):7276-7288.
615 28. Abeyrathne PD, Koh CS, Grant T, Grigorieff N, & Korostelev AA (2016) Ensemble cryo-
616 EM uncovers inchworm-like translocation of a viral IRES through the ribosome. *eLife* 5.
617 29. Zhang H, Ng MY, Chen Y, & Cooperman BS (2016) Kinetics of initiating polypeptide
618 elongation in an IRES-dependent system. *eLife* 5.
619 30. Firth AE, Wang QS, Jan E, & Atkins JF (2009) Bioinformatic evidence for a stem-loop
620 structure 5'-adjacent to the IGR-IRES and for an overlapping gene in the bee paralysis
621 dicistroviruses. *Virol J* 6:193.
622 31. Ren Q, Au HH, Wang QS, Lee S, & Jan E (2014) Structural determinants of an internal
623 ribosome entry site that direct translational reading frame selection. *Nucleic Acids Res*
624 42(14):9366-9382.
625 32. Wang QS, Au HH, & Jan E (2013) Methods for studying IRES-mediated translation of
626 positive-strand RNA viruses. *Methods* 59(2):167-179.
627 33. Dinos G*, et al.* (2004) Dissecting the ribosomal inhibition mechanisms of edeine and
628 pactamycin: the universally conserved residues G693 and C795 regulate P-site RNA
629 binding. *Mol Cell* 13(1):113-124.
630 34. Kozak M & Shatkin AJ (1978) Migration of 40 S ribosomal subunits on messenger RNA
631 in the presence of edeine. *J Biol Chem* 253(18):6568-6577.
632 35. Ruehle MD*, et al.* (2015) A dynamic RNA loop in an IRES affects multiple steps of
633 elongation factor-mediated translation initiation. *eLife* 4.
634 36. Kerr CH*, et al.* (2015) The 5' untranslated region of a novel infectious molecular clone of
635 the dicistrovirus cricket paralysis virus modulates infection. *J Virol* 89(11):5919-5934.
636 37. Yamamoto H*, et al.* (2016) 70S-scanning initiation is a novel and frequent initiation mode
637 of ribosomal translation in bacteria. *Proc Natl Acad Sci U S A* 113(9):E1180-1189.
638 38. Nomura M, Gourse R, & Baughman G (1984) Regulation of the synthesis of ribosomes
639 and ribosomal components. *Annu Rev Biochem* 53:75-117.
640 39. Adhin MR & van Duin J (1990) Scanning model for translational reinitiation in eubacteria.
641 *J Mol Biol* 213(4):811-818.
642 40. Colussi TM*, et al.* (2015) Initiation of translation in bacteria by a structured eukaryotic
643 IRES RNA. *Nature* 519(7541):110-113.
644 41. Sarabhai A & Brenner S (1967) A mutant which reinitiates the polypeptide chain after
645 chain termination. *J Mol Biol* 27(1):145-162.
646 42. Skabkin MA, Skabkina OV, Hellen CU, & Pestova TV (2013) Reinitiation and other
647 unconventional posttermination events during eukaryotic translation. *Mol Cell* 51(2):249-
648 264.

649 43. Abaeva IS, Pestova TV, & Hellen CU (2016) Attachment of ribosomal complexes and
650 retrograde scanning during initiation on the Halastavi arva virus IRES. *Nucleic Acids Res*
651 44(5):2362-2377.
652 44. Wills NM*, et al.* (2008) Translational bypassing without peptidyl-tRNA anticodon
653 scanning of coding gap mRNA. *EMBO J* 27(19):2533-2544.
654 45. Chen J*, et al.* (2015) Coupling of mRNA Structure Rearrangement to Ribosome
655 Movement during Bypassing of Non-coding Regions. *Cell* 163(5):1267-1280.
656 46. Samatova E, Konevega AL, Wills NM, Atkins JF, & Rodnina MV (2014) High-efficiency
657 translational bypassing of non-coding nucleotides specified by mRNA structure and
658 nascent peptide. *Nature communications* 5:4459.
659 47. Pisareva VP, Pisarev AV, & Fernandez IS (2018) Dual tRNA mimicry in the Cricket
660 Paralysis Virus IRES uncovers an unexpected similarity with the Hepatitis C Virus IRES.
661 *eLife* 7.
662 48. Paek KY*, et al.* (2015) Translation initiation mediated by RNA looping. *Proc Natl Acad*
663 *Sci U S A* 112(4):1041-1046.
664 49. Reinganum C (1975) The isolation of cricket paralysis virus from the emperor gum moth,
665 Antheraea eucalypti Scott, and its infectivity towards a range of insect species.
666 *Intervirology* 5(1-2):97-102.
667 50. Chtarbanova S*, et al.* (2014) Drosophila C virus systemic infection leads to intestinal
668 obstruction. *J Virol* 88(24):14057-14069.
669 51. Bonning BC & Miller WA (2010) Dicistroviruses. *Annu Rev Entomol* 55:129-150.
670 52. Lang BF*, et al.* (2014) Massive programmed translational jumping in mitochondria. *Proc*
671 *Natl Acad Sci U S A* 111(16):5926-5931.
672 53. Herr AJ, Gesteland RF, & Atkins JF (2000) One protein from two open reading frames:
673 mechanism of a 50 nt translational bypass. *EMBO J* 19(11):2671-2680.
674 54. Garrey JL, Lee YY, Au HH, Bushell M, & Jan E (2010) Host and viral translational
675 mechanisms during cricket paralysis virus infection. *J Virol* 84(2):1124-1138.
676 55. Krishna NK, Marshall D, & Schneemann A (2003) Analysis of RNA packaging in wild-
677 type and mosaic protein capsids of flock house virus using recombinant baculovirus
678 vectors. *Virology* 305(1):10-24.
679 56. Cox J & Mann M (2008) MaxQuant enables high peptide identification rates,
680 individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.
681 *Nature biotechnology* 26(12):1367-1372.

682

683

684

685 **FIGURE LEGENDS**

686 **Fig. 1. Biochemical analysis of potential +1-frame translation by Dicistrovirus IGR**
687 **IRESs. A) IGR IRES mediated +1-frame translation using an *in vitro* translation**
688 **assay.** The translation of firefly luciferase (FLuc), which is fused in the 0-frame or +1-
689 frame, is driven by the individual IGR IRES within the bicistronic reporter construct.
690 Linearized reporter constructs were incubated in Sf21 *in vitro* translation insect cell
691 extract at 30 °C for 2 hours in the presence of [^{35}S] methionine/cysteine. Translation of
692 FLuc and RLuc was monitored by phosphorimager analysis after resolving on a 16%
693 SDS PAGE. On the bottom panel, the ratio of FLuc/RLuc are quantified and normalized
694 to the IAPV 0-frame translation. Shown are averages from at least three independent
695 biological experiments (± SD). **B) The secondary structure of CrPV IGR IRES.** The
696 ΔPKI, ΔPKII, ΔPKIII, L1.1A, and L1.1B mutants are shown. Comp denotes
697 compensatory mutations combining ΔPKI CC/GG and ΔPKI GG/CC. DM - double
698 mutation of ΔPKI and ΔPKII. TM - triple mutation of ΔPKI, ΔPKII, and ΔPKIII. The
699 potential U6186/G6617 base pair to direct +1 frame translation from a CUA leucine
700 codon.

701

702 **Fig. 2. Translation of +1 frame CrPV ORFx is IGR IRES-dependent and initiates**
703 **downstream of the PKI region. A) Schematic of mutations introduced downstream**
704 **of the IGR IRES**. The PKI region and downstream 'spacer' sequence of CrPV IGR
705 IRES are shown. A series of mutations and stop codons were introduced either on the
706 0-frame or +1-frame. **B) Analysis of the IGR IRES translation *in vitro*.** Linearized
707 reporter constructs are incubated in Sf21 translation extract at 30°C for 2 hours in the

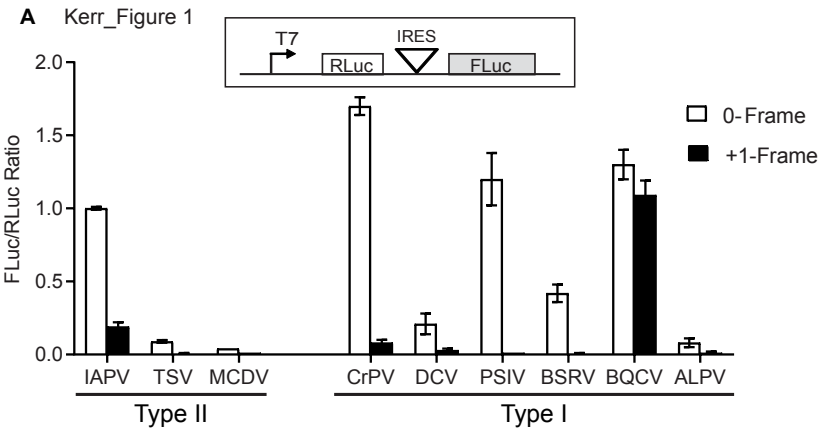
708 present of [^{35}S]-methionine/cysteine. Translation of FLuc and RLuc was monitored by
709 autoradiography after resolving on a 16% SDS PAGE. A representative gel from at least
710 three independent biological experiments is shown. * = p-value < 0.05. **C) Quantitation**
711 **of IRES-mediated translation.** Translation of FLuc and RLuc monitoring FLuc/RLuc
712 enzymatic activity were quantified and normalized to wild-type (WT) IAPV IRES. RLuc
713 monitors scanning-dependent translation acting as an internal control. RLA- relative
714 light units. Shown are averages from at least three independent biological experiments
715 (± SD). * = p-value < 0.05.
716
717 **Fig. 3. Translation of CrPV ORFx is dependent on 80S binding and scanning-**
718 **independent. A) Effect of L1.1 mutation on +1 frame translation.** Comparison of
719 luciferase activities between WT IGR IRES and L1.1 mutants (L1.1A and L1.1B) that do
720 not bind the 60S subunit. The value of RLuc and FLuc are normalized to WT. **B) +1-**
721 **frame translation is relatively insensitive to edeine.** CrPV IRES-mediated 0- and +1-
722 frame translation is monitored in the absence or presence of 2 μM edeine. The ΔPKI
723 mutants in both 0- and +1-frame are used as controls for IGR IRES dependent
724 translation. The luciferase activities of RLuc and FLuc on the left panel are normalized
725 to the WT of the 0-frame, whereas the luciferase activities of RLuc and FLuc on the right
726 panel are normalized to the WT of the +1-frame. Shown are averages from at least
727 three independent biological experiments (± SD).
728
729 **Fig. 4. Pseudotranslocation is necessary for +1-frame translation.** Constructs
730 containing mutations in the variable loop region (VLR) of the IGR IRES known to disrupt

731 pseudotranslocation were monitored for translation activity. Luciferase activities were
732 compared between WT IGR IRES and VLR mutants after *in vitro* translation assays. "x"
733 denotes deleted nucleotides within VLR. Shown are averages from at least three
734 independent biological experiments (± SD).
735
736 **Fig. 5. The downstream sequence of CrPV IGR IRES promotes +1-frame**
737 **translation. A) Chimeric IRES displays +1-frame translation activity.** The PSIV IGR
738 IRES, which shows no +1-frame translation activity (Fig. 1 and Fig. 3A), is combined
739 with the downstream sequence of CrPV IGR IRES PKI region. The luciferase activities
740 are normalized to the wild-type CrPV constructs in both +1- and 0-frame. The PSIV-
741 CrPV chimera in the 0-frame is used as control to demonstrate that PSIV IGR IRES
742 functions under this context. **B) Truncation of the spacer region impairs +1-frame**
743 **translation from the CrPV IGR IRES.** Schematic displaying truncations within the
744 spacer region between the PKI domain and the +1 frame start site is shown above. *In*
745 *vitro* translation assay of the CrPV IRES bicistronic construct with the respective
746 mutations were monitored by luciferase activities. Both RLuc and FLuc activities are
747 normalized to that of WT constructs. Results are shown as a normalized ratio of
748 FLuc/RLuc activity (left) and as separated RLuc and Fluc activities (right) Shown are
749 averages from at least three independent biological experiments (± SD). RLA = relative
750 luciferase activity

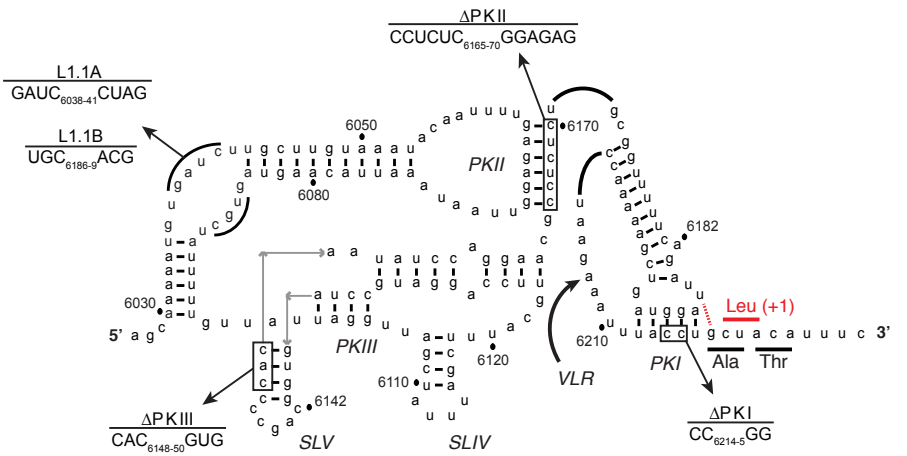
751 **Fig. 6. ORFx is expressed in CrPV infected S2 cells.** The predicted +1 frame amino
752 acid sequence is shown above. Residues that are italicized represent the amino acid
753 sequence if initiation occurred adjacent to the IGR IRES, whereas the bolded K lysine

754 residue denotes the position of the 13th amino acid. *Drosophila* S2 cells were infected
755 with CrPV (MOI 10). Proteins were extracted, digested with trypsin and subjected to LC-
756 MS/MS analysis. Peptide fragment spectra were searched against a *Drosophila* uniprot
757 database plus CrPV proteins. Two peptides (highlighted) were detected from the trypsin
758 digestion of S2 cell lysate at 6 hours post infection. Individual fragment ions are
759 annotated in the spectra and in the sequence representation.
760
761 **Fig. 7. Pathogenesis of a CrPV mutant lacking ORFx is attenuated in adult**
762 ***Drosophila melanogaster* flies. A) Schematic of mutations in the CrPV infectious**
763 **clone, CrPV-2.** The downstream nucleotide sequence of CrPV-2 IGR IRES and its
764 potential amino acid sequence of +1-frame ORFx are shown. The ORF2-Stop mutant
765 depletes the synthesis of 0-frame viral structural proteins. Mutants S12 and S19 place a
766 stop codon on the +1-frame with a synonymous mutation in the 0-frame. Note that
767 CrPV-2 (Accession: KP974706) sequence has a $C_{6279}A$ mutation (denoted with a black
768 arrowhead) compared to the original CrPV (Accession: AF218039) sequence. Residues
769 that are italicized, underlined and in bold are the peptides identified by mass
770 spectrometry in Fig. 6. **B)** Adult flies (Iso w^{1118}; 10 male, 10 female) were injected
771 intrathoracically with 5000 FFU of CrPV-2, CrPV-S12, CrPV-S19, or PBS. Subsequently,
772 flies were flipped onto standard media and survival was monitored daily. Shown is a
773 graph representing the average from three separate biological experiments (±SEM).
774
775 **Fig. 8. Subcellular localization of CrPV ORFx.** S2 cells were transfected with a
776 construct expressing C-terminally HA-tagged ORFx or a mutant version of ORFx in the
777 transmembrane region and incubated for 48 h. Following incubation, cells were fixed,

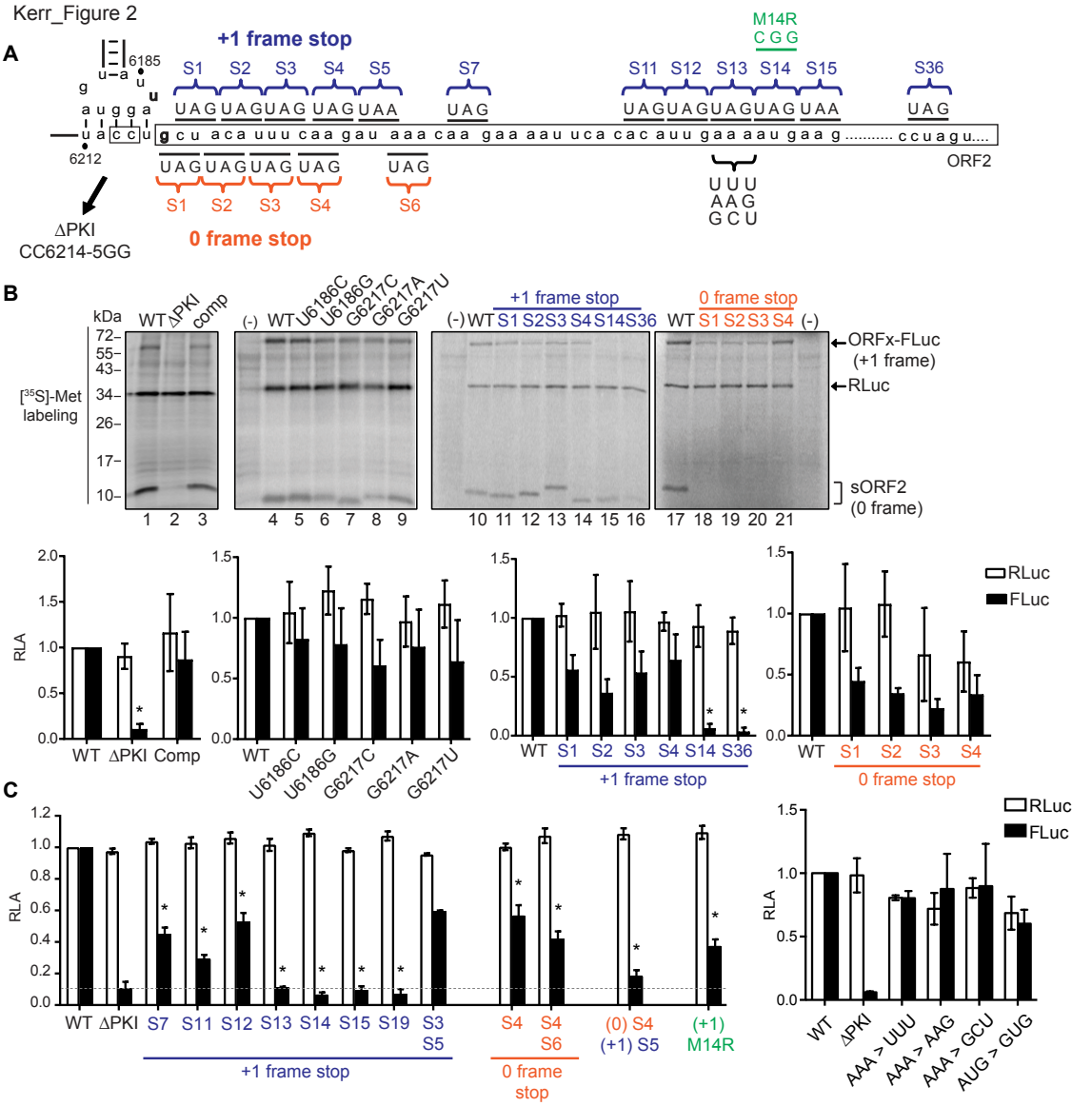
778 permeabilized and co-stained with HA antibody and antibodies against the ER (Calnexin
779 99A) and cytoplasm (alpha-Tubulin). Shown are representative Z-stack micrographs
780 from three independent experiments. Scale bars represent 15 μm.
781
782

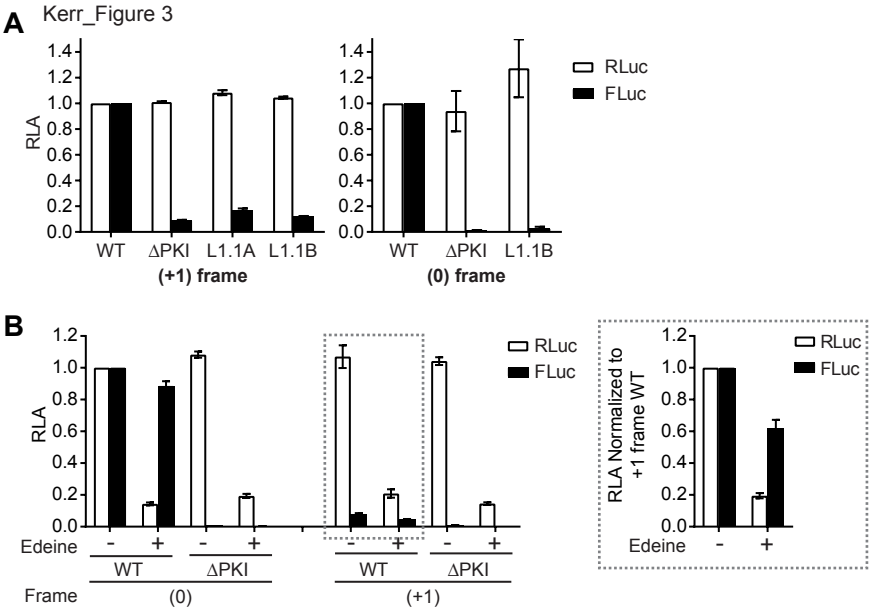


B Cricket paralysis virus IGR IRES

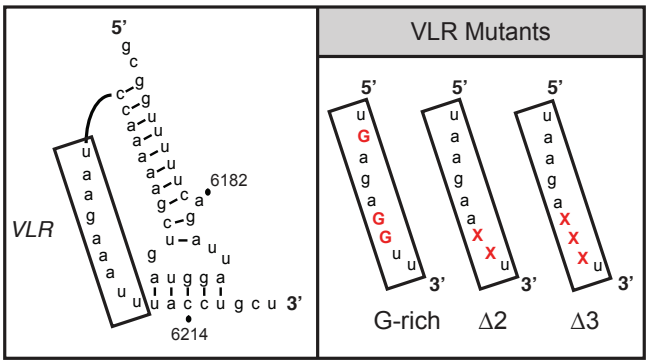


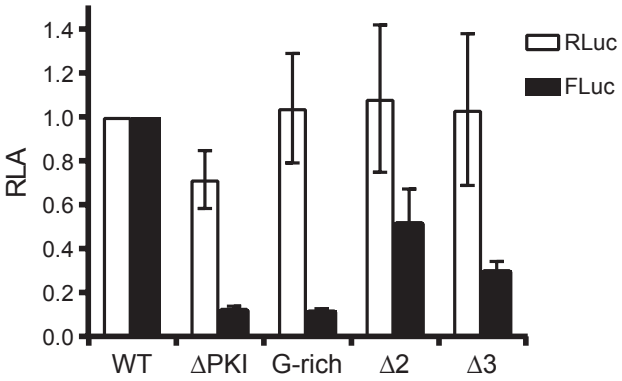
Kerr_Figure 2




A
Kerr_Figure 3
RLA
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
WT
ΔPKI
L1.1A
L1.1B
(+1) frame
(0) frame
RLuc
FLuc
B
Edeine
Frame
(0)
(+1)
RLA Normalized to +1 frame WT

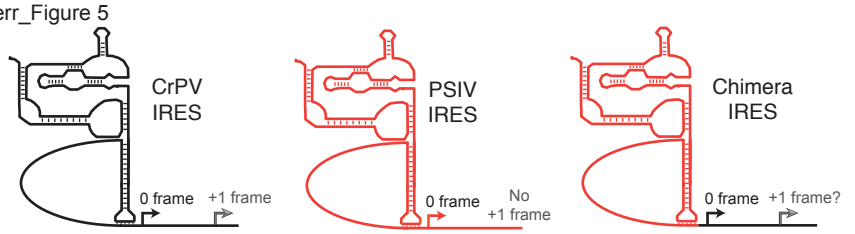
Kerr_Figure 4

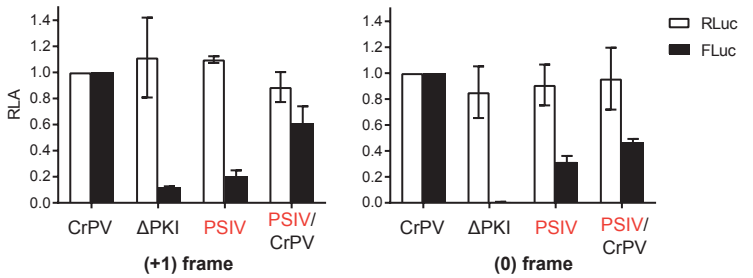




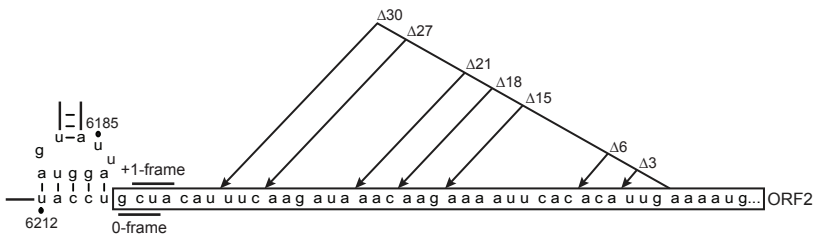
Kerr_Figure 5

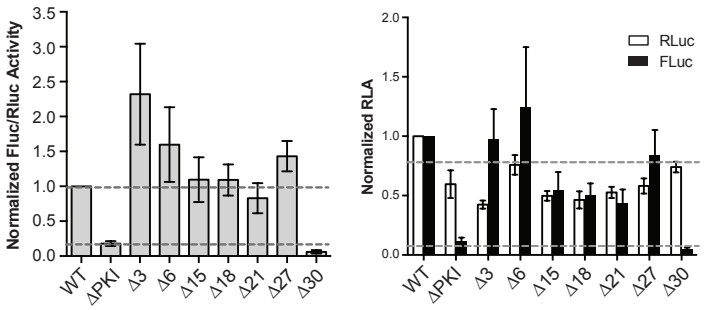
A





B

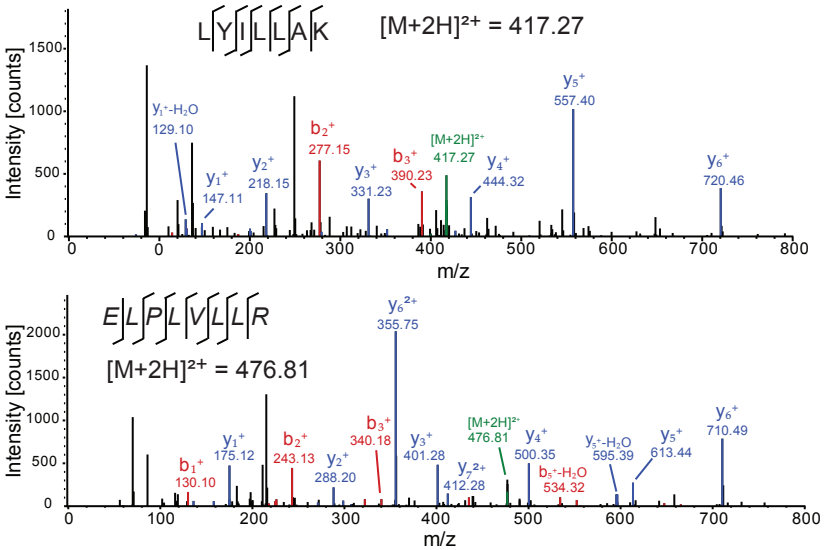




Kerr_Figure 6

Peptide #1 Peptide #2

Predicted Sequence: N - **K**MKIKDLCQNRKK **LYILLAK ELPLVLLR** SLISLIFQLIIW - **C**



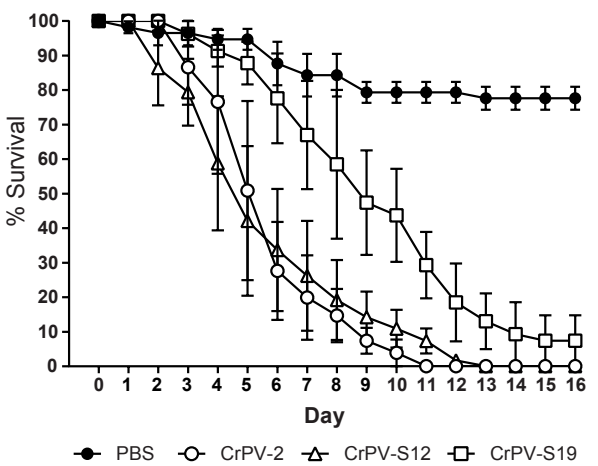
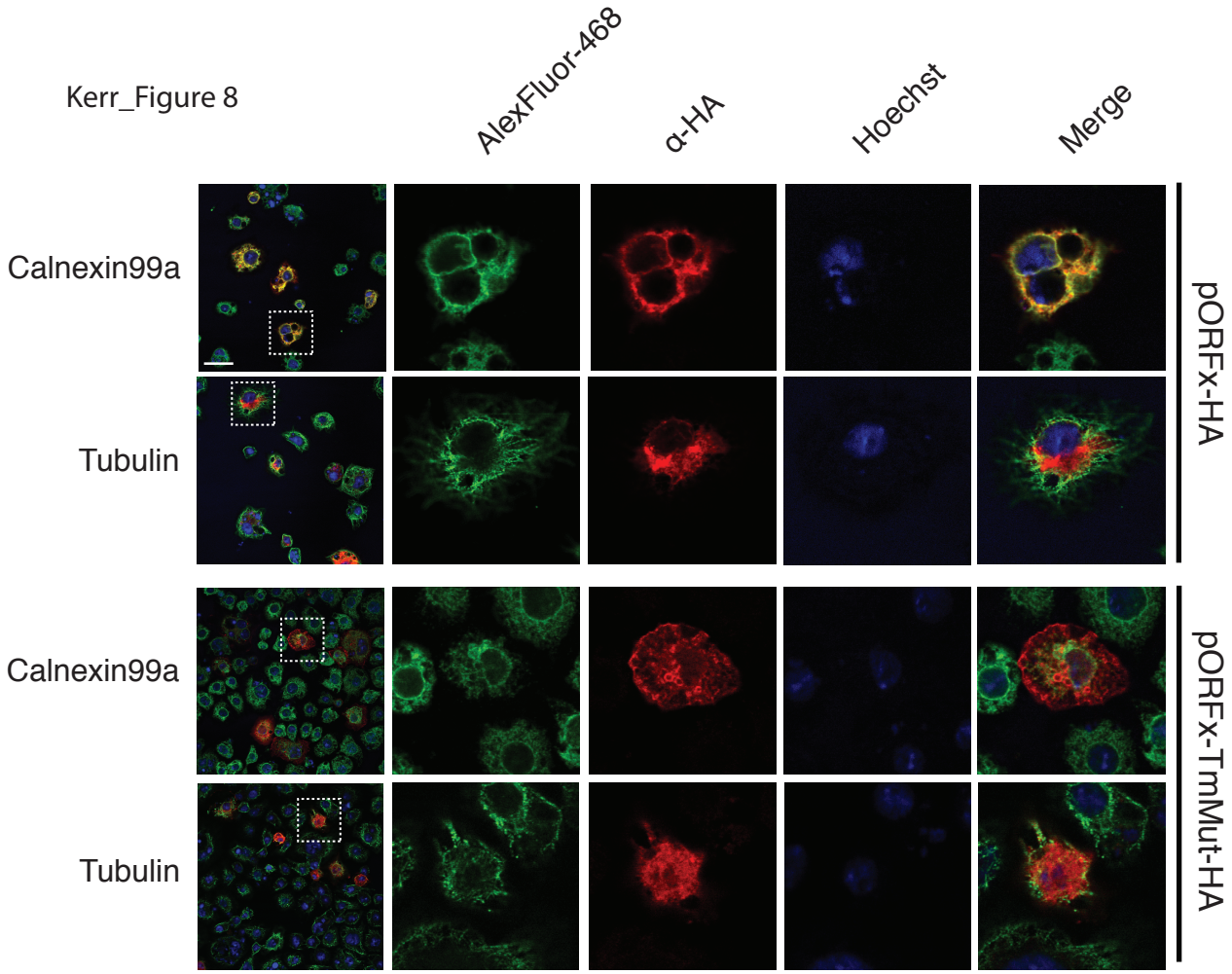
A Kerr_Figure 7

```
                                0 frame
                                ORF2-stop                  S12 (+1 frame)
                                  uaa                      uag
                                                            14
(+1) frame        L  H  F  K  I  N  K  K  I  H  T  L  K  M  K  I  K  D
 (0) frame       A  T  F  Q  D  K  Q  E  N  S  H  I  E  N  E  D  K  R
          UACCU gcuacauuucaagauaaacaagaaaauucacacauugaaaaugaagauaaaaga
           PKI  6218            6232                  6251

          S19 (+1 frame)
          uaa                                                34
(+1) frame  L  C  Q  N  R  K  K  L  Y  I  L  L  A  K  E  L  P  L  V  L
 (0) frame  L  M  S  E  Q  K  E  I  V  H  F  V  S  E  G  V  T  P  S  T
          cuuaugucagaacagaaagaaauuguacauuuuguuagcgaaggaguuaccccuaguacu
          6272                                              6317

                                                      51
(+1) frame  L  R  S  L  I  S  L  I  F  Q  L  I  I  W  T
 (0) frame  T  A  L  P  D  I  V  N  L  S  T  N  Y  L  D  M
          acugcgcucccugauaucguuaaucuuucaacuaauuaucuggacauga
          6332                                  6368        6379
```

B



Kerr_Figure 8