Human cytomegalovirus evades ZAP detection by suppressing CpG dinucleotides in the major immediate early genes.

Yao-Tang Lin*[1], Stephen Chiweshe*[1], Dominique McCormick[1], Anna Raper[1], Arthur Wickenhagen[4], Victor DeFillipis[3], Eleanor Gaunt[1], Peter Simmonds[3], Sam J Wilson[§4], Finn Grey[§1]

1. Division of Infection and Immunity, The Roslin Institute, University of Edinburgh, Easter Bush, Midlothian, UK.
2. Vaccine and Gene Therapy Institute, Oregon Health and Science University, Portland, Oregon, USA.
3. Peter Medawar Building for Pathogen Research, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom.
4. MRC-University of Glasgow Centre for Virus Research, Glasgow, United Kingdom

*Contributed equally to manuscript
[§]Co-corresponding authors

Abstract

The genomes of RNA and small DNA viruses of vertebrates display significant suppression of CpG dinucleotide frequencies. Artificially increasing dinucleotide frequencies results in substantial attenuation of virus replication, suggesting that these compositional changes may facilitate recognition of non-self RNA sequences. Recently, the interferon inducible protein ZAP, was identified as the host factor responsible for sensing CpG in viral RNA, through direct binding and possibly downstream targeting for degradation. Using an arrayed interferon stimulated gene expression library screen, we identified ZAPS, and its associated factor TRIM25, as direct inhibitors of human cytomegalovirus (HCMV) replication. Exogenous expression of ZAPS and TRIM25 significantly reduced virus replication while knockdown resulted in increased virus replication. HCMV displays a strikingly heterogeneous pattern of CpG representation with a specific suppression of CpGs within the IE1 major immediate early transcript which is absent in subsequently expressed genes. We demonstrated that suppression of CpG dinucleotides in the IE1 gene allows evasion of inhibitory effects of ZAP. During HCMV infection, expression of ZAP and TRIM25 are rapidly reduced, removing pressure to suppress dinucleotide frequencies in viral genes expressed after the immediate early genes, while acute virus replication and high levels of ZAP are mutually exclusive. Finally, we show that TRIM25 regulates alternative splicing between the ZAP short and long

38 isoforms during HCMV infection and interferon induction, with knockdown of TRIM25 resulting
39 in decreased ZAPS and corresponding increased ZAPL expression. These results
40 demonstrate for the first time that ZAP is a potent host restriction factor against large DNA
41 viruses and that HCMV evades ZAP detection through suppression of CpG dinucleotides
42 within the major immediate early transcripts. Furthermore, TRIM25 is required for efficient
43 upregulation of the interferon inducible short isoform of ZAP through regulation of alternative
44 splicing.
45
46 **Introduction**
47
48 Interferon (IFN) is a crucial first line of defence against viral infection and shapes the adaptive
49 immune response by triggering release of cytokines and chemokines [1,2]. IFN expression is
50 triggered by the recognition of pathogen-associated molecular patterns (PAMPs) [1]. These
51 microbe-specific molecular structures are generally essential for the survival of the microbes,
52 but fundamentally different from the host. Examples of PAMPs include peptidoglycans,
53 liposaccharide (LPS) and pathogen specific nucleic acid motifs, such as double stranded RNA
54 and unmethylated CpG sequences within DNA. Cells recognize PAMPs through pattern
55 recognition receptors (PRRs) that trigger innate immune responses following recognition of
56 the target. Families of PRRs include the membrane bound Toll-like receptors, C-type lectin
57 receptors, the cytoplasmic NOD-like receptors and RIG-I like receptors [3]. Upon recognition of
58 the specific PAMP during invasion by a foreign pathogen, PRRs trigger signaling cascades
59 that lead to relocation of IRF3/IRF7 complexes and NF-kB into the nucleus, initiating
60 expression of type I IFN. In turn, activation of the IFN receptor leads to up-regulation of
61 hundreds of IFN stimulated genes (ISGs) that, together, establish an antiviral cellular
62 environment [4].
63
64 Dinucleotide representation in RNA sequences have been investigated as a potential PAMP
65 [5-9]. Plant and vertebrate genomes show significantly lower CpG and TpA frequencies than
66 would be expected given their overall base composition. Lower CpG frequencies are thought
67 to have arisen through deamination of methylated cytosines in nuclear DNA, resulting in CpG
68 sequences mutating to TpG [10]. RNA and small DNA viruses of vertebrates have evolved a
69 similar pattern of suppressed CpG dinucleotides and artificially increasing dinucleotide
70 frequencies within their viral genomes through synonymous mutations results in considerable
71 attenuation of virus replication [6-9].
72
73 A recent study identified the short form of the zinc-finger antiviral protein (ZAP) and its
74 associated factor TRIM25 as responsible for recognition of high CpG frequencies in viral RNA

75 [11]. Mammalian ZAP is expressed in two major isoforms, ZAPS (short) and ZAPL (long) which
76 are generated by differential splicing, with ZAPL encoding an additional catalytically inactive
77 poly (ADP-ribose) polymerase (PARP) domain [12]. ZAP had previously been identified as a
78 host antiviral factor and is capable of binding to viral RNA through a pocket created by the
79 second of four zinc fingers within the RNA binding domain [13-15]. TRIM25 is an E3 ubiquitin
80 ligase and member of the tripartite motif (TRIM) family, many of which have been associated
81 with antiviral functions [16,17]. TRIM25 is required for ZAPS antiviral activity, although the precise
82 mechanism by which TRIM25 contributes to ZAPS antiviral activity is not fully understood [18,19].
83 A focused siRNA screen against human ISGs showed that knockdown of ZAP or TRIM25
84 rescued the replication of a defective HIV construct with artificially raised CpG levels [11].
85 Immunoprecipitation of ZAPS and sequencing of associated RNA demonstrated that ZAPS
86 directly interacts with high CpG regions of HIV RNA. Knockdown of ZAPS has also been
87 shown to rescue echovirus 7 virus that was attenuated through artificially increased CpG
88 levels [20].
89
90 While CpG dinucleotide levels have clearly been shown to impact the fitness of RNA and small
91 DNA viruses [6-9], their role in host recognition of large DNA viruses is less clear.
92
93 HCMV is a highly prevalent herpesvirus, persistently infecting between 30% and 100% of the
94 human population, correlating with socio-economic status[21]. HCMV remains an important
95 clinical pathogen accounting for more than 60% of complications associated with solid organ
96 transplant patients [22-24]. It is also the leading cause of infectious congenital birth defects
97 resulting from spread of the virus to neonates and has been linked to chronic inflammation
98 and immune aging [25-27]. However, infection is normally asymptomatic due to effective control
99 of virus replication by various arms of the immune system, including the interferon (IFN)
100 response [2]. Although HCMV has evolved multiple mechanisms to subvert and inhibit the
101 antiviral effects of IFNs, current evidence indicates they play a vital role in controlling
102 replication and pathogenesis. Individuals with mutations in key IFN signalling genes are
103 lethally susceptible to HCMV infections and recombinant IFN has been successfully used in
104 treating congenital HCMV and HCMV infection in AIDS patients. Furthermore, murine CMV is
105 more pathogenic in IFN knock-out mice than wildtype mice [2].
106
107 Infection with HCMV results in a robust IFN response. While viral attachment has been
108 reported to be sufficient to trigger the IFN response, detection of HCMV is likely to occur
109 through a multifactorial process [28]. The cellular sensors cGAS, IFI16 and ZBP1 have been
110 shown to play a role in detection of the HCMV genome, while Toll-like receptors (TLRs) are

111 also thought to be important [29-32]. Detection leads to activation of signalling proteins, including
112 IRF3, IRF7 and NFKB, resulting in up regulation of hundreds of ISGs [33-35]. Which of these
113 ISGs are directly responsible for inhibiting HCMV is poorly understood. While the initial ISG
114 response is robust, it is quickly shut down by the virus, resulting in a characteristic expression
115 profile of rapid induction at early time points (up to 24 hours post infection) followed by efficient
116 suppression of IFN regulated gene expression. This shut down is in part due to the expression
117 of the viral immediate early genes IE1 and IE2, both of which have been shown to curtail the
118 initial IFN response to HCMV [33,36,37].
119
120 To further dissect the role of the IFN response during HCMV infection, we used an arrayed
121 lentivirus expression library to identify ISGs that inhibit HCMV. We show that ZAP and TRIM25
122 can potently inhibit HCMV replication and that HCMV has evolved to suppress CpG
123 dinucleotides in the major immediate early transcript IE1 to evade detection by ZAP.
124
125 **Results**
126
127 *Arrayed ISG expression screening identified ZAPS and TRIM25 as direct inhibitors of HCMV*
128 *replication*
129
130 Systematic dissection of the 'antiviral state' using arrayed ISG expression libraries is an
131 effective method for identifying key components of the IFN response [38-42]. To identify ISGs
132 that inhibit HCMV, human primary fibroblast cells were transduced with 421 individual ISG-
133 encoding lentiviral vectors (or control empty vectors), in a 96-well plate format. Two days post
134 transduction, cells were infected at a multiplicity of infection (MOI) of three with TB40/E-GFP,
135 a BAC derived low passage HCMV strain containing an SV40 promoter driven eGFP cassette
136 inserted between the viral TRS1 and US34 genes [43]. Levels of GFP were monitored over a
137 seven-day period using a plate cytometer (Figure 1A). At seven days post infection,
138 supernatant was transferred to fresh untransduced fibroblast cells to determine virus
139 production levels. GFP levels were compared to the average signal for the respective plate.
140 While this approach can identify which ISGs have the capacity to inhibit HCMV replication,
141 many of the hits identified may inhibit virus replication by triggering IFN signalling through
142 overexpression alone, independently of HCMV infection. To identify ISGs more likely to
143 directly inhibit HCMV, as opposed to activation of IFN signalling, we performed a parallel
144 screen using IRF3 KO fibroblast cells generated by CRISPR/Cas9 editing [44]. Percentage GFP
145 levels were determined by comparing each ISG to the average GFP signal from each 96 well
146 plate to control for inter-plate variations. Figure 1B compares the percentage virus production
147 levels based on GFP signal for each ISG transduced well between wild type fibroblast cells

148 and IRF3 KO cells. As virus production is the most relevant measurement for the full virus
149 replication cycle, our studies focused on the results from this screen (The data from the
150 primary replication screen is shown in supplemental figure 1A-C and supplemental table 1).
151 The blue box highlights ISGs that reduced HCMV virus production by more than 2-fold in
152 unmodified fibroblast cells, but were not substantially inhibitory in IRF3 KO fibroblast cells.
153 These IRF3-dependant inhibitory ISGs are highlighted in figure 1C and include cGAS, TLR3,
154 MyD88 and DDX60, all known to act through IRF3 signalling. Additional ISGs, not previously
155 known to act through IRF3, showed a similar pattern. The green box in figure 1B highlights
156 ISGs that inhibit HCMV virus production in both wild type fibroblast cells and IRF3 KO cells,
157 with figure 1D showing the relative virus production levels in each cell type. These IRF3-
158 independent ISGs include those that signal through IRF3 independent pathways (eg IRF7) or
159 are known to inhibit HCMV directly (IDO and RIPK2) validating the screening approach [45,46].
160 Interestingly, both ZAPS and TRIM25 inhibited HCMV virus production in an IRF3 independent
161 manner, indicating they may play a role in directly inhibiting the virus. ZAPS and TRIM25 have
162 recently been identified as host factors responsible for inhibiting HIV-1, Echo-7 and influenza
163 virus constructs with synthetically increased CpG levels [9,11,20]. The two co-factors have also
164 been reported to have antiviral activity against multiple viruses [47-51]. Based on this, we decided
165 to further characterise the role of ZAPS and TRIM25 in the inhibition of HCMV replication.
166
167 To independently confirm the results of the screen, wild type fibroblast cells and IRF3 KO cells
168 were transduced with independent lentiviral vector stocks expressing ZAPS, TRIM25, cGAS
169 or an empty vector control, and infected with TB40/E-GFP (MOI of 3). GFP fluorescence levels
170 were monitored for seven days (Figure 2A and B). As expected, efficient cGAS inhibition was
171 largely dependent on IRF3 expression whereas virus replication was potently inhibited by
172 ZAPS or TRIM25 in wild type and IRF3 KO cells. This confirmed our observations from the
173 initial screen and implicated ZAPS and TRIM25 as direct inhibitors of HCMV virus replication.
174 To confirm GFP reporter levels expressed by the virus accurately reflected virus replication,
175 plaque assays were performed in wild type fibroblast cells transduced with lentivirus
176 expressing either ZAPS or TRIM25 and compared to cells transduced with an empty control
177 lentivirus. The results confirm that expression of ZAPS and TRIM25 significantly reduced viral
178 replication (Figure 2C). To determine the effect of knockdown of ZAP and TRIM25 on HCMV
179 replication, fibroblast cells were transfected with siRNAs targeting ZAP, TRIM25 or a negative
180 control siRNA. Western blot analysis demonstrated efficient knockdown of ZAP and TRIM25
181 (Supplemental figure 2). Importantly, siRNA knockdown of ZAP or TRIM25 significantly
182 increased HCMV replication, with titres increased by as much as 100-fold following
183 knockdown of ZAP (Figure 2D and E). Thus, endogenous ZAP can inhibit HCMV and the
184 rescue of HCMV replication presented here is likely an underestimate, as siRNAs do not

185 remove all the endogenous protein and ZAPS expression would be substantially higher in IFN-
186 stimulated cells.
187
188 *Distinct patterns of dinucleotide frequencies in subfamilies of herpesvirus genomes*
189
190 Recent studies have identified ZAPS and TRIM25 as host factors involved in targeting and
191 inhibiting RNA with high CpG dinucleotide frequencies [9,11,20]. As such they convey evolutionary
192 pressure on RNA and small DNA viruses, resulting in suppression of CpG dinucleotide
193 sequences within their genomes. However, the effects of host recognition of high CpG
194 frequencies on larger DNA viruses, such as herpesviruses, has not been reported. Analysis
195 of herpesvirus genomes suggests distinct patterns of CpG dinucleotide composition
196 associated with the three subfamilies (alpha, beta and gamma) [52]. While alpha herpesviruses
197 such as HSV-1 show no suppression of CpG dinucleotides within their genomes
198 (Supplemental figure 3), substantial CpG suppression is seen throughout gamma herpesvirus
199 genomes, such as Epstein-Barr virus (EBV) (Supplemental Figure 4). Beta herpesviruses
200 demonstrate localised suppression of CpG sequences within the major immediate early
201 genes. These genes are the first viral transcripts to be expressed following infection and they
202 are critical for driving lytic replication of the virus and are thought to play a pivotal role during
203 the establishment, maintenance and reactivation from latency. Analysis of sixteen beta
204 herpesvirus genomes demonstrates extensive evolutionary conservation of suppressed CpG
205 dinucleotide sequences specifically associated with immediate early transcripts (Figure 3). In
206 contrast, there is no such pattern for the complementary GpC dinucleotide control,
207 (Supplemental figure 5). These results indicate that the earliest beta herpesvirus transcripts
208 have been selected for reduced CpG content, whereas these constraints do not extend to viral
209 transcripts expressed at later times during infection.
210
211 *IE1 expression is not affected by ZAP.*
212
213 To determine the effect of ZAP expression on viral genes, the open reading frames of IE1, IE2
214 and pp52, an early gene with a high CpG dinucleotide frequency, were cloned into expression
215 vectors. Cellular GAPDH tagged with T7, that has a low CpG frequency in line with mammalian
216 genomes, was cloned into the same expression vector to generate a negative control. These
217 plasmids were co-transfected with a plasmid expressing ZAPS. Expression of ZAPS had little
218 effect on T7-GAPDH (Figure 4A). Similarly, while expression of ZAPS profoundly reduced
219 pp52 expression by approximately 80% (Figure 4B) and IE2 expression to a lesser extent,
220 ZAPS had no effect on IE1 expression. These results are in line with what would be predicted
221 based on the CpG dinucleotide frequencies and confirm that suppression of CpG levels in IE1

222 allow this gene to evade the inhibitory effects of ZAPS. Interestingly, overexpression of ZAPL
223 shows the same effects as ZAPS, suggesting that the longer isoform has the same ability to
224 reduce expression of transcripts with high frequencies of CpG sequences (Figure 4C and D).
225
226 *IE1 expression is not affected by ZAP during HCMV infection*
227
228 As previously shown, overexpression of ZAPS profoundly inhibited virus replication (Figure
229 2A-C), and overexpression of ZAPS and ZAPL resulted in the reduced expression of HCMV
230 viral proteins encoded by genes with high CpG frequencies following transient transfection
231 experiments in HEK293T cells (Figure 4). To determine the effect of ZAPS on viral gene
232 expression, western blot analysis was performed on total protein samples from wild type
233 fibroblast cells transduced with lentivirus expressing ZAPS and infected two days later with
234 TB40E-GFP (MOI of 3) (Figure 5). As is the case with all herpesviruses, viral gene expression
235 occurs in a controlled temporal cascade with immediate early (IE) early (E) and late (L) gene
236 expression occurring in a sequential fashion. Total protein was harvested at the indicated time
237 points and levels of IE (IE1 and IE2), E (pp52) and L (pp28) proteins measured. Strikingly,
238 ZAPS had little to no effect on IE1 expression levels throughout the time course, but
239 substantially reduced IE2 levels and subsequent E and L viral gene expression. These results
240 indicate that in the context of viral infection, IE1 expression is unaffected by ZAPS, likely due
241 to highly suppressed CpG abundance in this transcript, enabling evasion of ZAPS.
242
243 *Endogenous ZAP is induced during HCMV infection but expression is mutually exclusive to*
244 *acute virus progression.*
245
246 As overexpression of ZAP profoundly attenuates HCMV replication, we next determined the
247 effects of endogenous ZAP during virus replication. Western blot analysis was performed on
248 total protein lysates harvested from wild type fibroblast cells infected with TB40/E-GFP (MOI
249 of 3), which results in close to 100% infection based on GFP expression. ZAPS expression
250 was robustly induced 24 hours post infection and although levels decreased over time, the
251 levels of endogenous ZAPS remained higher than in uninfected cells throughout the course
252 of the infection (Figure 6A). ZAPL expression was not induced but also decreased over the
253 course of the infection.
254
255 To determine whether ZAP expression levels correlated with HCMV infection at the individual
256 cell level, we employed confocal microscopy. Cells were co-stained for ZAP and viral IE1 or
257 IE2 expression in wild type fibroblast cells following infection with AD169, a laboratory adapted
258 strain of HCMV. This strain was used instead of TB40/E-GFP as it does not express GFP,

259 which would interfere with the fluorescent signal. Uniform, low levels of ZAP expression can
260 be seen throughout uninfected cells (Figure 6B). While almost all cells are IE1 positive
261 following infection with HCMV at an MOI of 3, a mixed population of cells demonstrate high
262 and low expression of ZAP. Multiple IE1 positive cells were observed with high levels of ZAP
263 expression, further demonstrating that IE1 expression is not affected by high ZAP levels. In
264 contrast IE2 expression and high levels of ZAP were mutually exclusive, suggesting
265 progression of virus replication is blocked in cells expressing high levels of ZAP. This is
266 confirmed by quantification of ZAP expression levels in IE1 and IE2 expressing cells (Figure
267 6C).
268
269 The existence of a mixed population of cells expressing high and low levels of ZAP, correlating
270 with progression of virus replication is further supported by sorting of infected cell populations.
271 Following infection with TB40/E-GFP at an MOI of 3, cells were sorted into high and low GFP
272 expressing populations, 24 hours post infection, then reseeded and total protein levels
273 determined at the indicated time points (Figure 7). Western blot analysis indicates that the low
274 GFP expressing population correlated with high levels of ZAPS expression and a failure in
275 progression of virus replication (although IE1 protein expression could still be detected). In
276 contrast, high GFP expressing populations corresponded with lower ZAPS expression and
277 high levels of viral protein production, consistent with efficient virus replication.
278
279 These data suggest that high levels of ZAP expression are mutually exclusive with the
280 successful progression of virus replication. In contrast, IE1 expression is unaffected,
281 consistent with suppressed CpG content facilitating evasion of ZAP detection.
282
283 *TRIM25 is required for efficient ZAPS induction following HCMV infection or IFN treatment*
284
285 TRIM25 is recognised as a key functional partner of ZAPS antiviral activity. However, the
286 mechanism by which TRIM25 augments ZAPS antiviral activity is unclear. It has been shown
287 that TRIM25 ubiquitinates ZAPS and itself through its E3 ubiquitin ligase function [18,19].
288 However, ZAPS antiviral activity does not seem to be dependent on TRIM25 ubiquitination. In
289 this study, we show that like ZAPS, TRIM25 is a potent, IRF3 independent inhibitor of HCMV
290 replication (Figure 1 and 2).
291
292 Strikingly, in addition to increasing viral protein expression, siRNA knockdown of TRIM25
293 results in a substantial reduction in ZAPS expression following infection with HCMV and a
294 corresponding increase in ZAPL expression (Figure 8A). As previously described, ZAPS and
295 ZAPL are expressed from the same primary transcript through differential splicing and only

296 ZAPS expression is considered to be stimulated by IFN treatment, indicating that differential
297 splicing is involved in the regulation of ZAPS and ZAPL levels (Figure 8B).
298
299 To determine whether the failure in ZAPS induction following TRIM25 knockdown was at the
300 RNA or protein level, ZAPS and ZAPL RNA levels were determined by RT-qPCR. Fibroblast
301 cells were transfected with a control siRNA or siRNA targeting TRIM25 and infected with
302 TB40/E-GFP. Total RNA was harvested 24 HPI and primers specific to ZAPS, ZAPL or to a
303 shared region of ZAP were used to measure transcript levels. While total levels of ZAP RNA
304 were unaffected by TRIM25 knockdown, levels of ZAPS were significantly reduced with a
305 corresponding increase in ZAPL levels (Figure 8C). This effect is not specific to HCMV
306 infection as the same observation can be seen in uninfected fibroblast cells and cells treated
307 with IFN (Figure 8D and E). Furthermore, ZAP induction, as well as induction of other well
308 characterised ISGs, was not reduced by TRIM25 knockdown ruling out a general failure in IFN
309 induction (Figure 8F and Supplemental figure 6). The results clearly show that efficient IFN
310 induced gene expression occurs despite TRIM25 knockdown, suggesting the effect is specific
311 to differential splicing of ZAP. Although these results do not rule out the possibility that TRIM25
312 regulates the antiviral function of ZAPS through direct protein-protein interaction, they explain
313 how TRIM25 contributes to the activity of ZAPS by regulation of ZAPS induction through
314 differential splicing.
315
316 **Discussion**
317
318 Associated pathologies and therapeutic potential of HCMV is dependent on the host immune
319 response against the virus. As IFNs shape the innate and adaptive responses to the virus,
320 understanding how the IFN response is regulated during HCMV infection and how the virus
321 subverts this response could have important implications for our understanding of diseases
322 associated with the virus as well as for the rational design of vaccines and cancer therapeutics.
323
324 Although it is clear that the IFN response is important during HCMV infection, which of the
325 hundreds of induced ISGs play a critical role is less clear. Here we present a systematic
326 analysis, determining the effect of over 400 individual ISGs on HCMV replication. By
327 combining screens in wild type and IRF3 knockout cells we were able to define specific
328 subsets of ISGs that were IRF3 independent and therefore more likely to represent direct
329 inhibitors of virus replication. These included IDO and RIPK2, which have previously been
330 identified as important antiviral factors during HCMV infection [45,46]. In addition, multiple novel
331 inhibitors were identified, including ZAP and TRIM25, which have previously been shown to
332 act in a coordinated antiviral fashion [18,19].

333
334 ZAP was originally identified as an antiviral protein by screening a rat cDNA library for factors
335 that inhibited the replication of Moloney murine leukemia retrovirus [13]. Subsequent studies
336 have demonstrated that ZAP has antiviral activity against a range of RNA viruses, including
337 HIV, filoviruses, flaviviruses, coxsackievirus B3, influenza A virus, Newcastle disease virus
338 and Hepatitis B virus, a partially double stranded DNA virus [48-51,53-55]. Initial studies indicated
339 that ZAP antiviral activity was based on direct binding to viral RNA and degradation through
340 recruitment of exosome complex components and inhibition of translation [13,14,56]. Further
341 studies have suggested that ZAPS interacts with RIG-I to stimulate IFN expression through
342 IRF3 signalling [54]. This is consistent with the partial rescue of ZAPS inhibition of HCMV in
343 IRF3 KO cells observed in this study, suggesting a dual functionality of ZAPS in HCMV
344 antiviral activity, acting both as a pathogen sensing protein and as a direct antiviral factor.
345
346 Until recently, the sequence or sequence characteristics that defined ZAPS RNA binding
347 specificity were not known. However, a study using a focused siRNA screen identified ZAPS
348 and TRIM25 as host factors responsible for inhibition of HIV-1 constructs with artificially raised
349 CpG dinucleotide frequencies and it was shown that ZAPS specifically binds to RNA
350 sequences at CpG dinucleotide motifs [13]. The molecular mechanism for this binding has now
351 been established through X-ray crystallography of the ZAP RNA binding domain complexed
352 with a target CpG RNA, which demonstrates the second of four zinc fingers creates a highly
353 basic patch that is required for specific binding of CpG dinucleotides [57].
354
355 It had previously been suggested that mammalian RNA and small DNA viruses suppress CpG
356 dinucleotides to mimic the composition of their host genomic make up, thereby avoiding
357 recognition as foreign nucleic acid [6-8]. While CpG bias has previously been reported for
358 herpesviruses, this bias was attributed to potential methylation status rather than evasion of a
359 host antiviral response [52]. The pattern of CpG dinucleotide frequencies in herpesviruses is so
360 distinctive that, with few exceptions, viruses can be attributed to the main three subfamilies
361 based on the CpG dinucleotide frequency patterns within open reading frame (Supplemental
362 figures 1-3). The majority of alpha-herpesviruses demonstrate little or no CpG suppression,
363 while gamma-herpesviruses demonstrate substantial suppression across the genome. Beta-
364 herpesviruses display the most striking pattern of all, with suppression of CpG dinucleotides
365 linked to temporal gene expression and specifically restricted to the immediate early genes. It
366 is intriguing to speculate that the genomes of these large DNA viruses may have been so
367 dramatically moulded by a single host antiviral mechanism, although other factors may also
368 play a role.
369

370 The fact that CpG dinucleotide patterns are so uniform across the virus sub-families also
371 suggests a central underlying biological relevance. For example, alpha-herpesviruses are
372 mainly associated with latent infections in neuronal cells where the interferon response may
373 be limited due to immune privilege and constitutive expression levels of ZAP are low [20],
374 whereas beta-herpesviruses and gamma-herpesviruses are mainly associated with latency in
375 haematopoietic cells and are therefore under greater pressure to evade host cell IFN
376 responses [58]. Alternatively, alpha-herpesviruses may have evolved a robust mechanism that
377 directly counteracts ZAP antiviral activity, making it unnecessary to repress CpG
378 dinucleotides. This is supported by a previous study that showed HSV-1 was able to replicate
379 to wild type levels despite expression of a ZAP construct, although the construct expressed a
380 truncated version of rat ZAP, rather than full length human ZAP [47]. Further studies will be
381 required to determine whether the CpG dinucleotide frequency in herpesvirus sub-families is
382 directly related to ZAP expression and why these patterns are so distinctively associated with
383 the specific virus sub-families.
384
385 Multiple studies have linked TRIM25 with efficient ZAP antiviral activity, and, like ZAPS, our
386 screen identified TRIM25 as an IRF3 independent inhibitor of HCMV. TRIM25 is an E3 ligase
387 that catalyses ubiquitylation and ISGylation of target proteins [16,17]. TRIM25 contains a zinc ring
388 finger, B-box, coiled coil and PRY/SPRY domain. Both Takata *et al* and Li *et al* demonstrated
389 that TRIM25 is required for ZAP antiviral activity through siRNA screens [11,18], whereas Zheng
390 *et al* identified TRIM25 as a ZAP interacting factor through affinity purification [19]. These studies
391 demonstrated that TRIM25 interacts directly with ZAP through the PRY/SPRY domain and
392 ubiquitinates ZAP and itself. While ubiquitination requires binding of ZAP and TRIM25 to RNA
393 it does not seem to be directly necessary for ZAP antiviral activity [18,19,59]. TRIM25 also inhibits
394 influenza A virus in a ZAP independent manner through direct binding to the viral
395 ribonucleoprotein complex [60]. We show that TRIM25 expression potently inhibits HCMV
396 replication although further studies will be required to determine whether this occurs in a ZAP
397 dependent or independent manner.
398
399 In addition, our studies suggest that TRIM25 regulates the alternative splicing of ZAP. A recent
400 study demonstrated that alternative splicing of the ZAP primary transcript leads to at least four
401 different isoforms; ZAPS, ZAPM, ZAPL and ZAPXL [61], although our study focused on the
402 major isoforms ZAPS and ZAPL. ZAPL contains a catalytically inactive PARP-like domain at
403 the C terminus, that is missing from ZAPS. Both ZAPS and ZAPL have been reported to have
404 antiviral activity, although ZAPS was identified as the host factor responsible for CpG
405 recognition [13]. In our study we found that both ZAPS and ZAPL could target high CpG
406 transcripts. Following virus infection or IFN treatment ZAP is induced, however differential

407 splicing results in higher levels of induction of ZAPS compared to ZAPL. Following HCMV
408 infection, TRIM25 knockdown resulted in significantly lower levels of ZAPS protein and RNA
409 levels with a corresponding increase in ZAPL levels indicating a TRIM25 dependant change
410 in differential splicing of ZAP. Regulation of ZAP splicing by TRIM25 occurs in uninfected cells
411 and IFN stimulated cells, indicating it is not specific to HCMV but rather a fundamental aspect
412 of ZAP regulation. It is unclear at this point whether TRIM25 plays a direct role in splicing of
413 ZAP or whether the regulation occurs through an intermediate signalling pathway. TRIM25
414 has been shown to directly bind RNA through its PRY/SPRY domain [59], and plays a role in the
415 maturation of the microRNA let-7 [62], suggesting a more direct role in processing ZAP RNA is
416 possible. However further studies will be required to dissect its precise role [59].
417
418 Increasing the CpG levels within the IE1 coding region could be a viable approach for
419 generating safer live attenuated vaccines with reduced risk of reversion as the mutations could
420 be spread across the entire length of the IE1 gene. Increasing the CpG levels in IE1 could
421 also increase the immunogenicity of the virus by triggering more robust IFN responses, which
422 could improve HCMV vaccines as well as the use of HCMV as a vaccine vector and as a
423 therapeutic cancer vaccine. However, construction of such viruses may be challenging due to
424 the central role the major immediate early genes play in acute replication of the virus and the
425 potential for introduced mutations to disrupt the complex splicing events or regulatory regions
426 embedded within the coding region of IE1. Future experiments will be required to determine
427 the potential application of such viruses.
428
429 **Materials and Methods**
430
431 **Cell culture and viral infection**
432 Normal Human Dermal Fibroblasts (NHDF; Gibco) and IRF3 -/- cells were maintained in
433 Dulbecco's modified high glucose media (DMEM; Sigma) supplemented with 10% fetal bovine
434 serum (FBS; Gibco) and 1% penicillin-streptomycin (Invitrogen). IRF3 -/- cells were provided
435 by Victor DeFillipis. A low passage HCMV strain TB40E-GFP [43], which expresses GFP from
436 an SV40 promoter was used for arrayed ISG expression library screening, western blot
437 analysis, Real-Time PCR, and flow cytometry analysis. Laboratory adapted HCMV strain
438 AD169 was used for immunofluorescence experiments.
439
440 **Arrayed ISG lentivirus expression library screening**
441 The human lentivirus expression library has been previously described [42]. In brief the library
442 encodes 420 human ISG genes on the pSCRPSY backbone (KT368137.1). Normal human

443 dermal fibroblast cells and IRF3-/- cells were seeded in 96-well plates a day before
444 transduction. Next day, cells reached 90-95% confluency and were transduced with the ISG
445 library. Transduced cells were incubated for 48 hours and then infected with GFP expressing
446 TB40E virus at an MOI of three. GFP intensity was monitored every 24 hours with Synergy
447 HT microplate reader (Biotek).
448
449 **siRNA transfection**
450 NHDFs were seeded in 6-well plates a day before siRNA transfection. Next day, cells reached
451 90-95% confluency and were transfected with siRNA twice (4 hours apart between first and
452 second transfections) using Lipofectamine RNAiMAX (Invitrogen) according to the
453 manufacturer's protocol. siRNAs used in this article are Human ON-TARGETplus siRNAs
454 against TRIM25 and ZC3HAV1 (4 individual siRNAs per gene; Dharmacon). Transfected
455 NHDFs were incubated for 48 hours and then infected with GFP expressing TB40E virus at
456 an MOI of three.
457
458 **Western blot analysis**
459 Cells were lysed with RIPA buffer (0.1% SDS, 1% Triton X-100, 1% deoxycholate, 5 mM
460 EDTA, 150 mM NaCl, and 10 mM Tris at pH 7.2) containing protease inhibitor cocktail (Roche).
461 Ten micrograms of the total lysate was separated in 10% SDS-polyacrylamide gels and
462 transferred to PVDF membranes (Millipore). Primary antibodies used in this paper are mouse
463 anti-CMV IE1/2 monoclonal antibody (MAB8131, Millipore), mouse anti-CMV pp52
464 monoclonal antibody (CH16, Santa Cruz Biotechnology), mouse anti-CMV pp28 monoclonal
465 antibody (CH19, Santa Cruz Biotechnology), rabbit anti-ZAP polyclonal antibody (PA5-31650,
466 Invitrogen), mouse anti-TRIM25 monoclonal antibody (BD Biosciences), rabbit anti-T7 tag
467 monoclonal antibody (D9E1X, Cell Signaling Technology), mouse anti-alpha tubulin
468 monoclonal antibody (DM1A, Abcam) and mouse anti-β-Actin monoclonal antibody (Abcam).
469 Blots were probed with primary antibody (1:500-1:5000) diluted in 5% dehydrated milk in Tris
470 Buffered Saline (TBS) and subsequently the HRP-conjugated secondary antibodies (Pierce)
471 at 1:5000. Blots were washed in TBS three times, incubated with chemiluminescent substrate
472 (SuperSignal West Pico; Thermo Scientific) according to the manufacturer's protocol, and
473 exposed in G:Box (Syngene) for visualization of bands.
474
475 **Vector Construction and transfection**
476 The expression vectors used in the article were made by cloning the coding sequences of
477 UL123 (IE1, TB40E), UL122 (IE2, TB40E), UL44 (pp52, TB40E), Human GAPDH with double
478 T7-tag at the N-terminus, into the pcDNA3.1 vector. The ZAP(L) and ZAP(S) expression

479 vectors were made by cloning the coding sequences of ZAP(L) and ZAP(S) into the pcDEF3.1
480 vector derived from pcDNA3.1 with the CMV promoter replaced by the EF1alpha promoter.
481
482 **Immunofluorescence**
483 Laboratory adapted HCMV strain AD169 infected cells were fixed in 4%
484 paraformaldehyde solution for 20 minutes and then permeabilized in Methanol:Acetone
485 solution (1:1) at -20°C for 7 minutes, and then blocked with 5% human serum in PBS for 30
486 minutes. Primary and secondary antibodies were diluted with 5% human serum in PBS. Cells
487 were washed with PBS after primary and after secondary antibody incubations. Primary
488 antibodies used in this paper are mouse anti-CMV IE2 monoclonal antibody (12E2, Santa
489 Cruz Biotechnology), mouse anti-CMV IE1/2 monoclonal antibody (MAB8131, Millipore), and
490 rabbit anti-ZAP polyclonal antibody (PA5-31650, Invitrogen) at 1:500. Alexa-fluor-647
491 conjugated goat anti-mouse or Alexa-fluor-488 conjugated goat-anti-rabbit IgG secondary
492 antibodies were diluted 1:1000. All images were acquired with Zeiss LSM 710 confocal
493 microscope fitted with 63X/1.4 oil-immersion objective lens.
494
495 **Real-Time PCR analysis**
496 Total RNA was isolated by using Trizol solution according to the manufacturer's protocol
497 followed by DNase (Turbo DNA-free kit, Ambion) treatment, and then reverse transcribed with
498 poly T primers using High Capacity cDNA Reverse Transcription Kit (Invitrogen). Real-Time
499 PCR was carried out using by Taqman assays with pre-designed gene-specific primer/probe
500 set (Applied Biosystems) on Rotor gene 3000 (Corbet Research). Custom primer/probe set
501 are TCCTCTCTCAGGATCTGTATGT, GGAGAGGAAGGAGTCAAAGATG, and 56-FAM/
502 ACCATCTAC/ZEN/CCATTGGCTCAAGCA/3IABkFQ for ZAP(L), and AGCATGGTGTGACT
503 GAAAGG, CTTCACAGCTGGAGAAGCTAAA, and 56-FAM/TCTGAAAGG/ZEN/GAAGTCTG
504 AGCGAGTCT/3IABkFQ for ZAP(S).
505
506 **Fluorescence-activated cell sorting (FACS)**
507 NHDFs were infected with GFP expressing TB40E virus at an MOI of three. 24 hours later,
508 the cells were trypsinized, and resuspended in PBS buffer. The resuspended cells were then
509 sorted into low GFP (parameter set using uninfected cells) or high GFP population by the BD
510 FACSAria IIIu cell sorter. Immediately after sorting, a small portion of the cells from each
511 population were collected and lysed with RIPA buffer, and the remaining low GFP and high
512 GFP cells were re-seeded into a 6-well plate respectively and incubated for another 24 to 72
513 hours before lysed with RIPA buffer.
514
515

516 **Analysis of CpG dinucleotides frequency in Herpesvirus genome**

517 CpG dinucleotide frequencies were determined using the program Composition scan in the
518 SSE Sequence editor v1.3 (PMID: 22264264). Results were expressed as ratios of observed
519 frequencies (fCpG) to those predicted by the frequencies of their component mononucleotides
520 (fC x fG). This ratio was normalised through computing constraints imposed by amino acid
521 coding on dinucleotide frequencies. For example, a methionine codon enforces the presence
522 of ApU and UpG dinucleotides, while glycine codons requires a GpG dinucleotide, as well as
523 GpN, Corrected dinucleotide ratios are therefore based on observed to expected frequencies
524 coding-enforced dinucleotides are excluded. All accession data is listed in supplemental table
525 2.

526

527 **Acknowledgments**

528 This project was funded by the MRC (MR/N001796/1), the BBSRC (BBS/E/D/20002172) and
529 through the Principal's Career Development PhD scholarship from the University of
530 Edinburgh. We would like to acknowledge help and technical assistance from Robert Fleming
531 and Graeme Robertson of the Roslin Bioimaging and flow cytometry facility.

532

533 **Competing interests**

534 There are no competing interests.

535

536 **References**

537

538 1 Ivashkiv, L. B. & Donlin, L. T. Regulation of type I interferon responses. *Nat Rev*
539 *Immunol* **14**, 36-49, doi:10.1038/nri3581 (2014).
540 2 Amsler, L., Verweij, M. & DeFilippis, V. R. The tiers and dimensions of evasion of the
541 type I interferon response by human cytomegalovirus. *J Mol Biol* **425**, 4857-4871,
542 doi:10.1016/j.jmb.2013.08.023 (2013).
543 3 Kumar, H., Kawai, T. & Akira, S. Pathogen recognition by the innate immune system.
544 *Int Rev Immunol* **30**, 16-34, doi:10.3109/08830185.2010.529976 (2011).
545 4 Shaw, A. E. *et al.* Fundamental properties of the mammalian innate immune system
546 revealed by multispecies comparison of type I interferon responses. *PLoS Biol* **15**,
547 e2004086, doi:10.1371/journal.pbio.2004086 (2017).
548 5 Atkinson, N. J., Witteveldt, J., Evans, D. J. & Simmonds, P. The influence of CpG and
549 UpA dinucleotide frequencies on RNA virus replication and characterization of the
550 innate cellular pathways underlying virus attenuation and enhanced replication. *Nucleic*
551 *Acids Res* **42**, 4527-4545, doi:10.1093/nar/gku075 (2014).
552 6 Karlin, S., Doerfler, W. & Cardon, L. R. Why is CpG suppressed in the genomes of
553 virtually all small eukaryotic viruses but not in those of large eukaryotic viruses? *J Virol*
554 **68**, 2889-2897 (1994).
555 7 Rima, B. K. & McFerran, N. V. Dinucleotide and stop codon frequencies in single-
556 stranded RNA viruses. *J Gen Virol* **78 (Pt 11)**, 2859-2870, doi:10.1099/0022-1317-78-
557 11-2859 (1997).

558 8 Simmonds, P., Xia, W., Baillie, J. K. & McKinnon, K. Modelling mutational and
559 selection pressures on dinucleotides in eukaryotic phyla--selection against CpG and
560 UpA in cytoplasmically expressed RNA and in RNA viruses. *BMC Genomics* **14**, 610,
561 doi:10.1186/1471-2164-14-610 (2013).
562 9 Gaunt, E. *et al.* Elevation of CpG frequencies in influenza A genome attenuates
563 pathogenicity but enhances host response to infection. *Elife* **5**, e12735,
564 doi:10.7554/eLife.12735 (2016).
565 10 Coulondre, C., Miller, J. H., Farabaugh, P. J. & Gilbert, W. Molecular basis of base
566 substitution hotspots in Escherichia coli. *Nature* **274**, 775-780, doi:10.1038/274775a0
567 (1978).
568 11 Takata, M. A. *et al.* CG dinucleotide suppression enables antiviral defence targeting
569 non-self RNA. *Nature* **550**, 124-127, doi:10.1038/nature24039 (2017).
570 12 Kerns, J. A., Emerman, M. & Malik, H. S. Positive selection and increased antiviral
571 activity associated with the PARP-containing isoform of human zinc-finger antiviral
572 protein. *PLoS Genet* **4**, e21, doi:10.1371/journal.pgen.0040021 (2008).
573 13 Gao, G., Guo, X. & Goff, S. P. Inhibition of retroviral RNA production by ZAP, a
574 CCCH-type zinc finger protein. *Science* **297**, 1703-1706, doi:10.1126/science.1074276
575 (2002).
576 14 Guo, X., Carroll, J. W., Macdonald, M. R., Goff, S. P. & Gao, G. The zinc finger
577 antiviral protein directly binds to specific viral mRNAs through the CCCH zinc finger
578 motifs. *J Virol* **78**, 12781-12787, doi:10.1128/JVI.78.23.12781-12787.2004 (2004).
579 15 Meagher, J. L. *et al.* Structure of the zinc-finger antiviral protein in complex with RNA
580 reveals a mechanism for selective targeting of CG-rich viral sequences. *Proc Natl Acad*
581 *Sci U S A* **116**, 24303-24309, doi:10.1073/pnas.1913232116 (2019).
582 16 Zou, W. & Zhang, D. E. The interferon-inducible ubiquitin-protein isopeptide ligase
583 (E3) EFP also functions as an ISG15 E3 ligase. *J Biol Chem* **281**, 3989-3994,
584 doi:10.1074/jbc.M510787200 (2006).
585 17 Khan, R., Khan, A., Ali, A. & Idrees, M. The interplay between viruses and TRIM
586 family proteins. *Rev Med Virol* **29**, e2028, doi:10.1002/rmv.2028 (2019).
587 18 Li, M. M. *et al.* TRIM25 Enhances the Antiviral Action of Zinc-Finger Antiviral
588 Protein (ZAP). *PLoS Pathog* **13**, e1006145, doi:10.1371/journal.ppat.1006145 (2017).
589 19 Zheng, X. *et al.* TRIM25 Is Required for the Antiviral Activity of Zinc Finger Antiviral
590 Protein. *J Virol* **91**, doi:10.1128/JVI.00088-17 (2017).
591 20 Odon, V. *et al.* The role of ZAP and OAS3/RNAseL pathways in the attenuation of an
592 RNA virus with elevated frequencies of CpG and UpA dinucleotides. *Nucleic Acids*
593 *Res* **47**, 8061-8083, doi:10.1093/nar/gkz581 (2019).
594 21 Fields, B. N. K. D. M. H. P. M. *Fields virology*. (Wolters Kluwer Health/Lippincott
595 Williams & Wilkins, 2013).
596 22 Kotton, C. N. CMV: Prevention, Diagnosis and Therapy. *Am J Transplant* **13 Suppl 3**,
597 24-40; quiz 40, doi:10.1111/ajt.12006 (2013).
598 23 Kowalsky, S., Arnon, R. & Posada, R. Prevention of cytomegalovirus following solid
599 organ transplantation: a literature review. *Pediatr Transplant* **17**, 499-509,
600 doi:10.1111/petr.12118 (2013).
601 24 Bruminhent, J. & Razonable, R. R. Management of cytomegalovirus infection and
602 disease in liver transplant recipients. *World journal of hepatology* **6**, 370-383,
603 doi:10.4254/wjh.v6.i6.370 (2014).
604 25 Ballard, R. A., Drew, W. L., Hufnagle, K. G. & Riedel, P. A. Acquired cytomegalovirus
605 infection in preterm infants. *Am J Dis Child* **133**, 482-485 (1979).

606 26 Griffith, D. M. *et al.* Systemic inflammation after critical illness: relationship with
607 physical recovery and exploration of potential mechanisms. *Thorax* **71**, 820-829,
608 doi:10.1136/thoraxjnl-2015-208114 (2016).
609 27 Jergovic, M., Contreras, N. A. & Nikolich-Zugich, J. Impact of CMV upon immune
610 aging: facts and fiction. *Med Microbiol Immunol* **208**, 263-269, doi:10.1007/s00430-
611 019-00605-w (2019).
612 28 Juckem, L. K., Boehme, K. W., Feire, A. L. & Compton, T. Differential initiation of
613 innate immune responses induced by human cytomegalovirus entry into fibroblast cells.
614 *J Immunol* **180**, 4965-4977, doi:10.4049/jimmunol.180.7.4965 (2008).
615 29 Paijo, J. *et al.* cGAS Senses Human Cytomegalovirus and Induces Type I Interferon
616 Responses in Human Monocyte-Derived Cells. *PLoS Pathog* **12**, e1005546,
617 doi:10.1371/journal.ppat.1005546 (2016).
618 30 Lio, C. W. *et al.* cGAS-STING Signaling Regulates Initial Innate Control of
619 Cytomegalovirus Infection. *J Virol* **90**, 7789-7797, doi:10.1128/JVI.01040-16 (2016).
620 31 Diner, B. A., Lum, K. K., Toettcher, J. E. & Cristea, I. M. Viral DNA Sensors IFI16
621 and Cyclic GMP-AMP Synthase Possess Distinct Functions in Regulating Viral Gene
622 Expression, Immune Defenses, and Apoptotic Responses during Herpesvirus Infection.
623 *MBio* **7**, doi:10.1128/mBio.01553-16 (2016).
624 32 DeFilippis, V. R., Alvarado, D., Sali, T., Rothenburg, S. & Fruh, K. Human
625 cytomegalovirus induces the interferon response via the DNA sensor ZBP1. *J Virol* **84**,
626 585-598, doi:10.1128/JVI.01748-09 (2010).
627 33 Browne, E. P., Wing, B., Coleman, D. & Shenk, T. Altered cellular mRNA levels in
628 human cytomegalovirus-infected fibroblasts: viral block to the accumulation of
629 antiviral mRNAs. *J Virol* **75**, 12319-12330, doi:10.1128/JVI.75.24.12319-12330.2001
630 (2001).
631 34 DeFilippis, V. R. *et al.* Interferon regulatory factor 3 is necessary for induction of
632 antiviral genes during human cytomegalovirus infection. *J Virol* **80**, 1032-1037,
633 doi:10.1128/JVI.80.2.1032-1037.2006 (2006).
634 35 Yurochko, A. D., Kowalik, T. F., Huong, S. M. & Huang, E. S. Human cytomegalovirus
635 upregulates NF-kappa B activity by transactivating the NF-kappa B p105/p50 and p65
636 promoters. *J Virol* **69**, 5391-5400 (1995).
637 36 Taylor, R. T. & Bresnahan, W. A. Human cytomegalovirus immediate-early 2 gene
638 expression blocks virus-induced beta interferon production. *J Virol* **79**, 3873-3877,
639 doi:10.1128/JVI.79.6.3873-3877.2005 (2005).
640 37 Paulus, C., Krauss, S. & Nevels, M. A human cytomegalovirus antagonist of type I
641 IFN-dependent signal transducer and activator of transcription signaling. *Proc Natl*
642 *Acad Sci U S A* **103**, 3840-3845, doi:10.1073/pnas.0600007103 (2006).
643 38 Feng, J. *et al.* Interferon-Stimulated Gene (ISG)-Expression Screening Reveals the
644 Specific Antibunyaviral Activity of ISG20. *J Virol* **92**, doi:10.1128/JVI.02140-17
645 (2018).
646 39 Rihn, S. J. *et al.* TRIM69 Inhibits Vesicular Stomatitis Indiana Virus. *J Virol* **93**,
647 doi:10.1128/JVI.00951-19 (2019).
648 40 Dittmann, M. *et al.* A serpin shapes the extracellular environment to prevent influenza
649 A virus maturation. *Cell* **160**, 631-643, doi:10.1016/j.cell.2015.01.040 (2015).
650 41 Kane, M. *et al.* Identification of Interferon-Stimulated Genes with Antiretroviral
651 Activity. *Cell host & microbe* **20**, 392-405, doi:10.1016/j.chom.2016.08.005 (2016).
652 42 Schoggins, J. W. *et al.* A diverse range of gene products are effectors of the type I
653 interferon antiviral response. *Nature* **472**, 481-485, doi:10.1038/nature09907 (2011).

654 43 Umashankar, M. *et al.* A novel human cytomegalovirus locus modulates cell type-
655 specific outcomes of infection. *PLoS Pathog* **7**, e1002444,
656 doi:10.1371/journal.ppat.1002444 (2011).
657 44 Sali, T. M. *et al.* Characterization of a Novel Human-Specific STING Agonist that
658 Elicits Antiviral Activity Against Emerging Alphaviruses. *PLoS Pathog* **11**, e1005324,
659 doi:10.1371/journal.ppat.1005324 (2015).
660 45 Kapoor, A., Forman, M. & Arav-Boger, R. Activation of nucleotide oligomerization
661 domain 2 (NOD2) by human cytomegalovirus initiates innate immune responses and
662 restricts virus replication. *PLoS One* **9**, e92704, doi:10.1371/journal.pone.0092704
663 (2014).
664 46 Zimmermann, A. *et al.* Checks and balances between human cytomegalovirus
665 replication and indoleamine-2,3-dioxygenase. *J Gen Virol* **95**, 659-670,
666 doi:10.1099/vir.0.061994-0 (2014).
667 47 Bick, M. J. *et al.* Expression of the zinc-finger antiviral protein inhibits alphavirus
668 replication. *J Virol* **77**, 11555-11562, doi:10.1128/jvi.77.21.11555-11562.2003 (2003).
669 48 Chiu, H. P. *et al.* Inhibition of Japanese encephalitis virus infection by the host zinc-
670 finger antiviral protein. *PLoS Pathog* **14**, e1007166, doi:10.1371/journal.ppat.1007166
671 (2018).
672 49 Liu, C. H., Zhou, L., Chen, G. & Krug, R. M. Battle between influenza A virus and a
673 newly identified antiviral activity of the PARP-containing ZAPL protein. *Proc Natl*
674 *Acad Sci U S A* **112**, 14048-14053, doi:10.1073/pnas.1509745112 (2015).
675 50 Muller, S. *et al.* Inhibition of filovirus replication by the zinc finger antiviral protein. *J*
676 *Virol* **81**, 2391-2400, doi:10.1128/JVI.01601-06 (2007).
677 51 Zhu, Y. *et al.* Zinc-finger antiviral protein inhibits HIV-1 infection by selectively
678 targeting multiply spliced viral mRNAs for degradation. *Proc Natl Acad Sci U S A* **108**,
679 15834-15839, doi:10.1073/pnas.1101676108 (2011).
680 52 Honess, R. W. *et al.* Deviations from expected frequencies of CpG dinucleotides in
681 herpesvirus DNAs may be diagnostic of differences in the states of their latent genomes.
682 *J Gen Virol* **70 (Pt 4)**, 837-855, doi:10.1099/0022-1317-70-4-837 (1989).
683 53 Mao, R. *et al.* Inhibition of hepatitis B virus replication by the host zinc finger antiviral
684 protein. *PLoS Pathog* **9**, e1003494, doi:10.1371/journal.ppat.1003494 (2013).
685 54 Hayakawa, S. *et al.* ZAPS is a potent stimulator of signaling mediated by the RNA
686 helicase RIG-I during antiviral responses. *Nat Immunol* **12**, 37-44, doi:10.1038/ni.1963
687 (2011).
688 55 Li, M. *et al.* Zinc finger antiviral protein inhibits coxsackievirus B3 virus replication
689 and protects against viral myocarditis. *Antiviral Res* **123**, 50-61,
690 doi:10.1016/j.antiviral.2015.09.001 (2015).
691 56 Guo, X., Ma, J., Sun, J. & Gao, G. The zinc-finger antiviral protein recruits the RNA
692 processing exosome to degrade the target mRNA. *Proc Natl Acad Sci U S A* **104**, 151-
693 156, doi:10.1073/pnas.0607063104 (2007).
694 57 Meagher, J. L. *et al.* Structure of the zinc-finger antiviral protein in complex with RNA
695 reveals a mechanism for selective targeting of CG-rich viral sequences. *Proc Natl Acad*
696 *Sci U S A*, doi:10.1073/pnas.1913232116 (2019).
697 58 Fields, B. N., Knipe, D. M. & Howley, P. M. *Fields' virology*. 5th ed. / editors-in-chief,
698 David M. Knipe, Peter M. Howley ; associate editors, Diane E. Griffin ... [et al.]. edn,
699 (Wolters Kluwer/Lippincott Williams & Wilkins, 2007).
700 59 Choudhury, N. R. *et al.* RNA-binding activity of TRIM25 is mediated by its
701 PRY/SPRY domain and is required for ubiquitination. *BMC Biol* **15**, 105,
702 doi:10.1186/s12915-017-0444-9 (2017).

703 60 Meyerson, N. R. *et al.* Nuclear TRIM25 Specifically Targets Influenza Virus
704 Ribonucleoproteins to Block the Onset of RNA Chain Elongation. *Cell Host Microbe*
705 **22**, 627-638 e627, doi:10.1016/j.chom.2017.10.003 (2017).
706 61 Li, M. M. H. *et al.* Characterization of Novel Splice Variants of Zinc Finger Antiviral
707 Protein (ZAP). *J Virol* **93**, doi:10.1128/JVI.00715-19 (2019).
708 62 Choudhury, N. R. *et al.* Trim25 Is an RNA-Specific Activator of Lin28a/TuT4-
709 Mediated Uridylation. *Cell Rep* **9**, 1265-1272, doi:10.1016/j.celrep.2014.10.017
710 (2014).
711

Figure 1

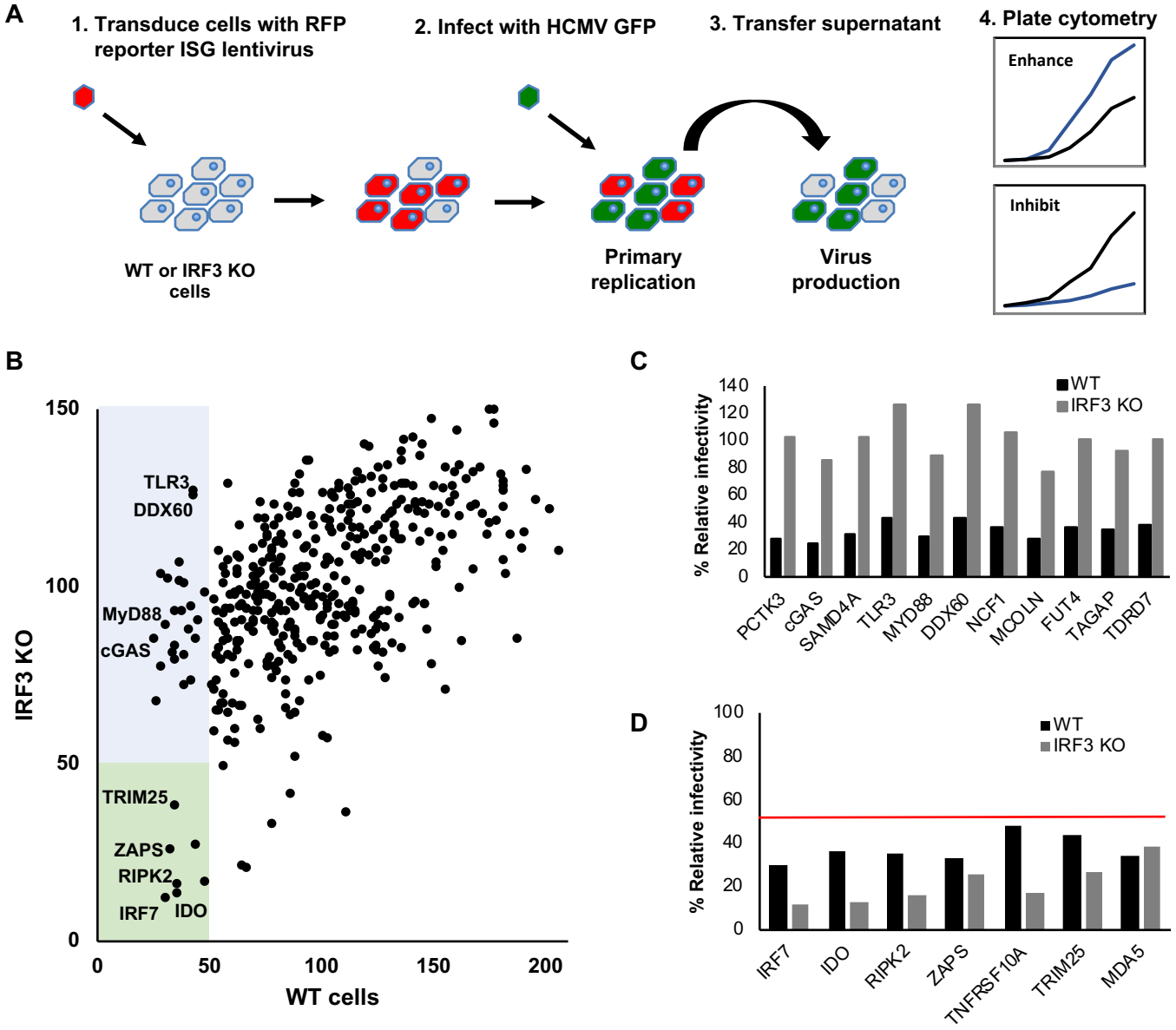


Figure 1. Arrayed ISG expression screening identified ZAPS and TRIM25 as direct inhibitors of HCMV replication. (A) Wild type and IRF3 KO fibroblast cells were seeded in 96-well plates and transduced with arrayed ISG lentivirus library. Cells were infected at 48 hours post-transduction with TB40/E-GFP (MOI of 3). GFP levels were monitored over a seven-day period to measure primary replication. Seven-days post infection supernatant was transferred to untransduced cells to measure virus production (B) Direct comparison of relative HCMV virus production for each individual ISG between wild type and IRF3 KO fibroblast cells. Blue box = ISGs that reduced virus production by more than 2-fold in wild type cells, but not in IRF3 KO cells. The green box = ISGs that reduced virus production by more than 2-fold in both wild type and IRF3 KO cells. Relative primary replication and virus production levels of HCMV for the ISGs in the blue box (C) and green box (D) are shown.

Figure 2

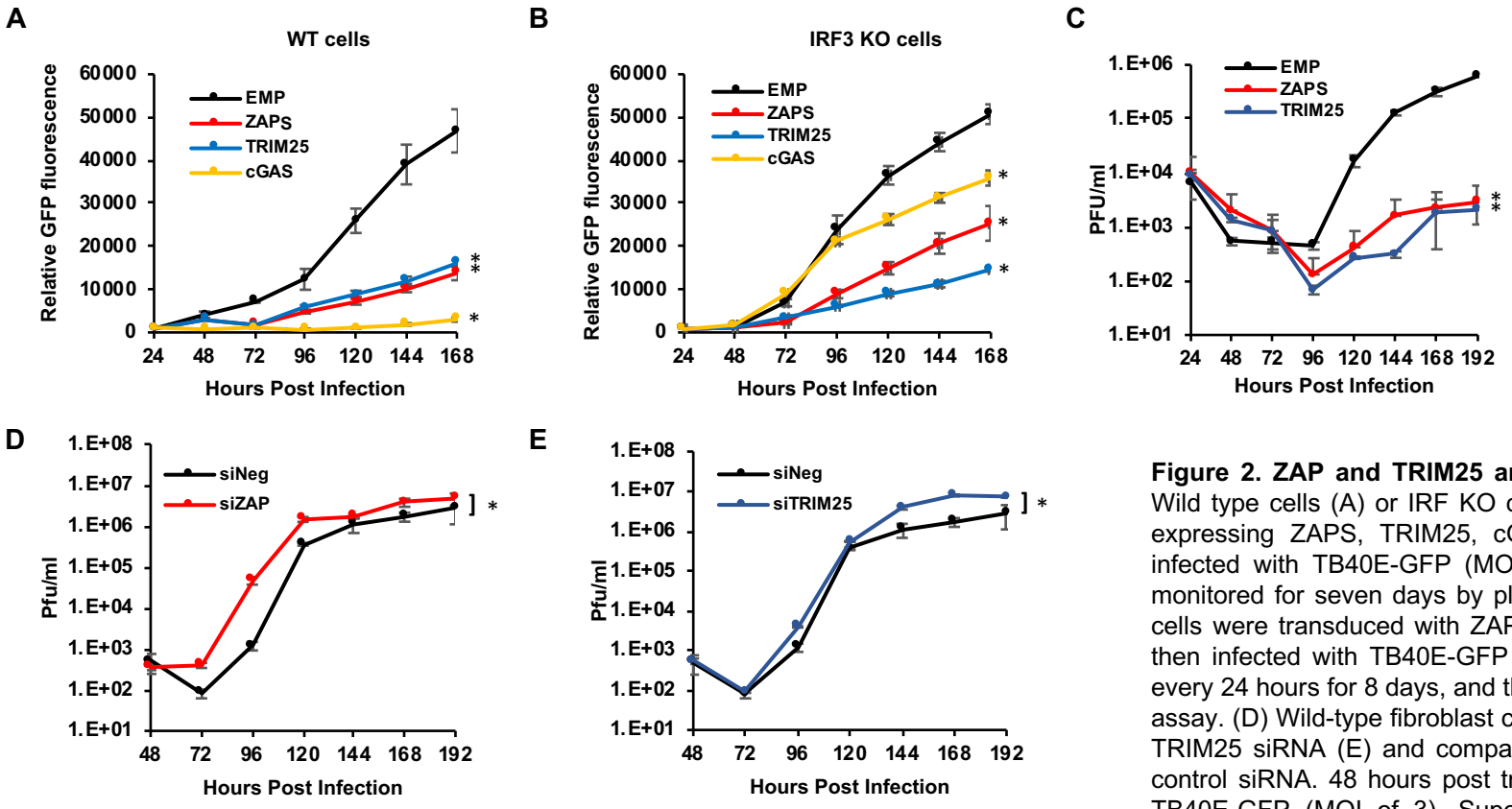


Figure 2. ZAP and TRIM25 are host restriction factors for HCMV. Wild type cells (A) or IRF KO cells (B) were transduced with lentivirus expressing ZAPS, TRIM25, cGAS or an empty control vector, and infected with TB40E-GFP (MOI of 3). GFP fluorescence levels were monitored for seven days by plate fluorometry. (C) Wild-type fibroblast cells were transduced with ZAPS, TRIM25 or an empty vector control, then infected with TB40E-GFP (MOI of 3). Supernatant was collected every 24 hours for 8 days, and the viral titres were determined by plaque assay. (D) Wild-type fibroblast cells were transfected with ZAP siRNA or TRIM25 siRNA (E) and compared to cells transfected with a negative control siRNA. 48 hours post transfection, the cells were infected with TB40E-GFP (MOI of 3). Supernatant was collected and virus titres determined by plaque assay. N = 2. * p-value < 0.05 based on two-way ANOVA.

Figure 3

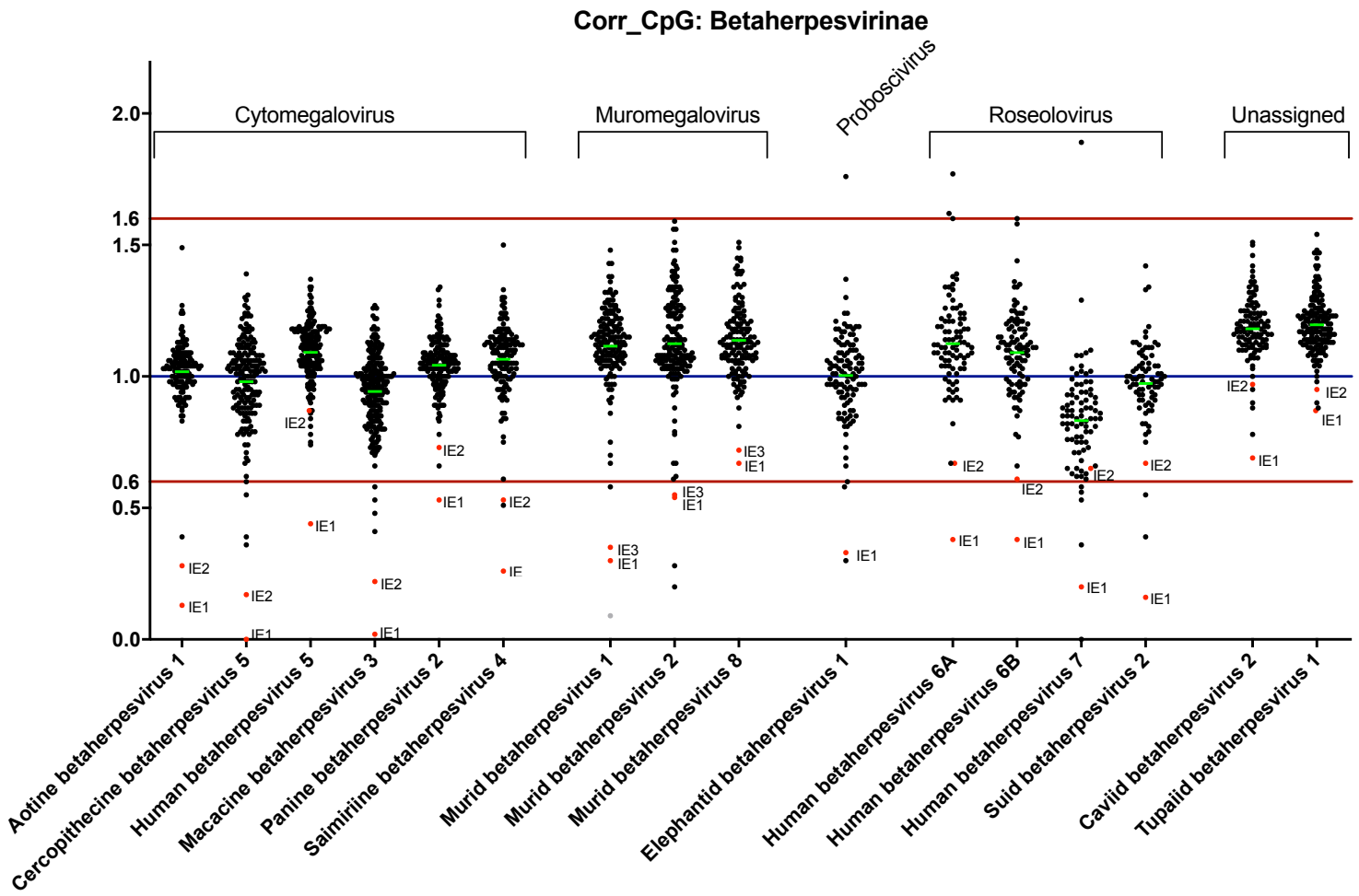


Figure 3. Specific suppression of CpG nucleotides within the immediate early genes of beta-herpesviruses. The CpG content of annotated open reading frames from 16 beta-herpesvirus genomes are shown, following normalization for length and GC content. A corrected CpG ratio of one reflects the expected number of CpGs based on GC content of a transcript.

Figure 4

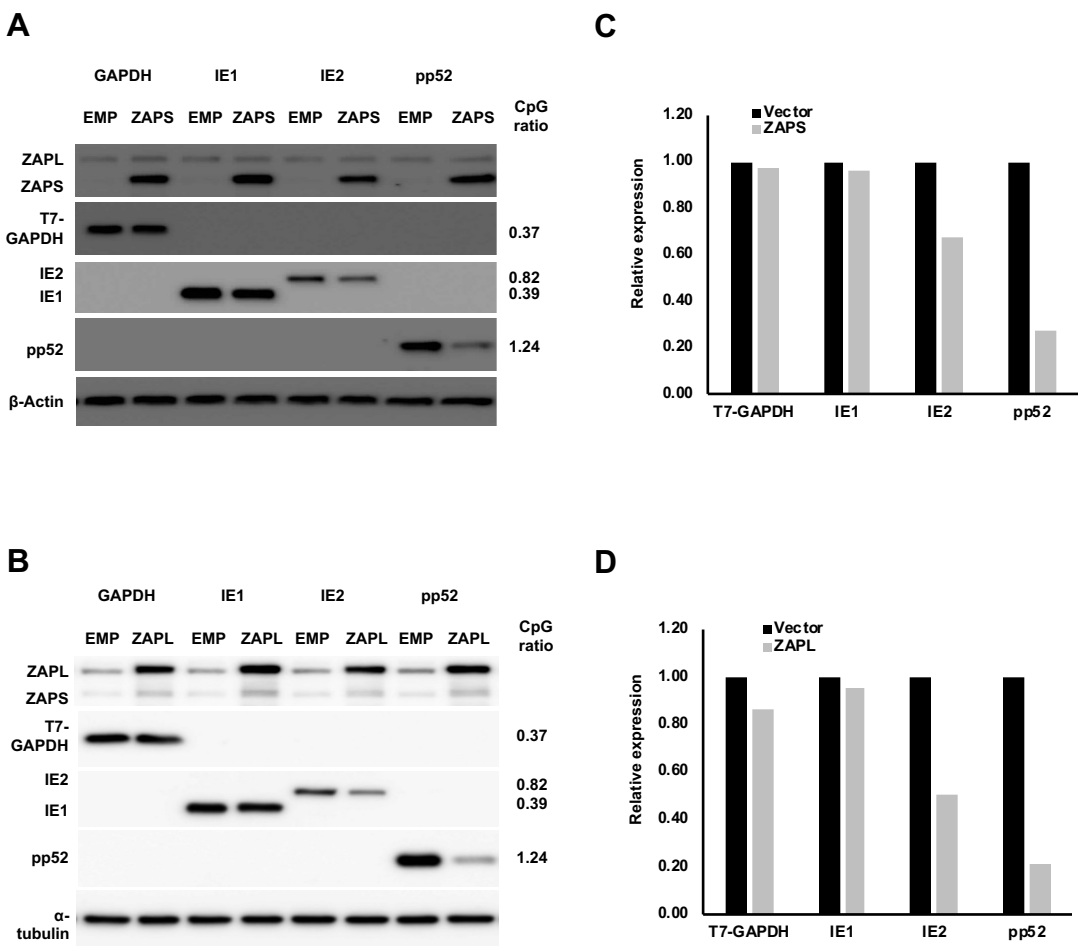


Figure 4. ZAP targets HCMV genes with high CpG dinucleotide frequencies. (A) 293T cells were co-transfected with plasmids expressing ZAPS (A), ZAPL (B) or a control empty vector with plasmids expressing HCMV genes with varied CpG dinucleotide frequencies (IE1, IE2 and pp52). A plasmid expressing T7-tagged GAPDH was included as a control with low CpG sequence content. CpG frequencies are indicated beside each gel. Expression levels of HCMV genes and T7-GAPDH were determined by western blot analysis. Expression levels for proteins co-expressed with ZAPS (C) and ZAPL (D) were quantified using Image J software and normalized to the empty construct control samples.

Figure 5

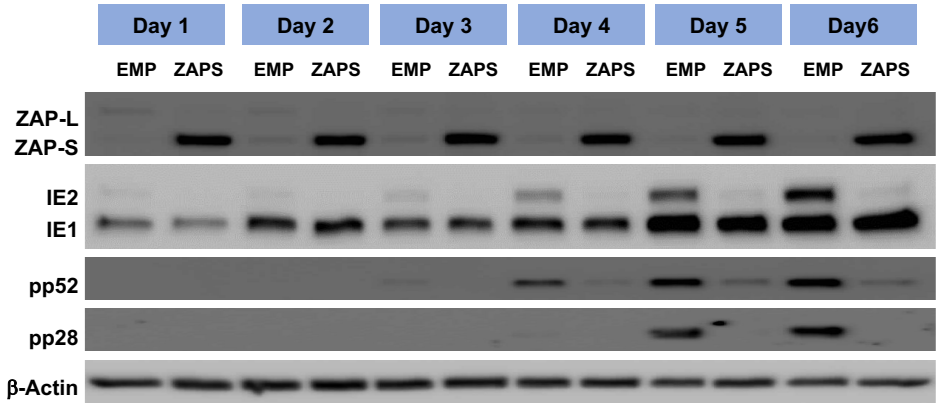
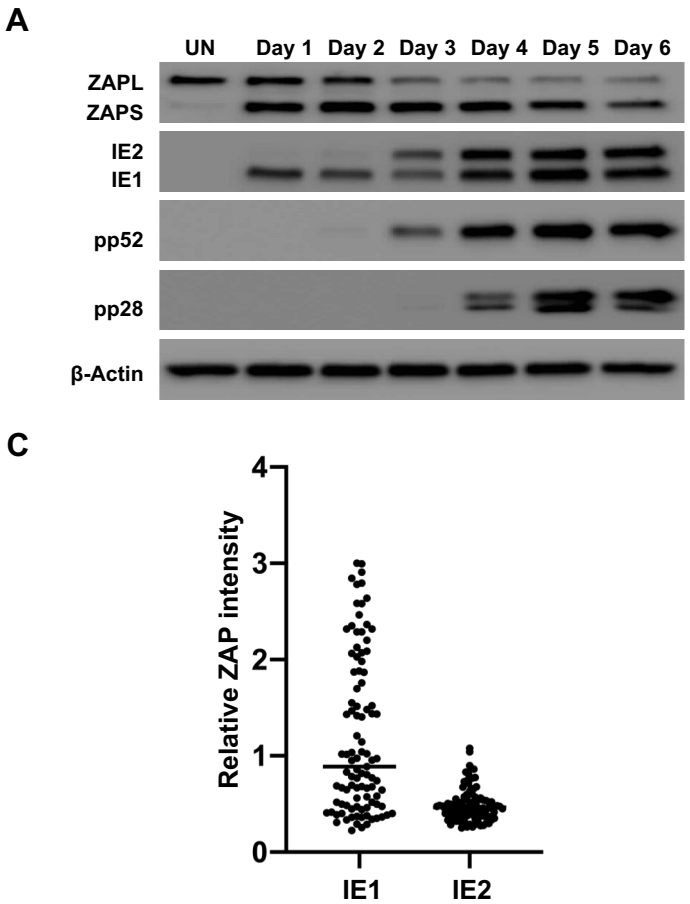


Figure 5. IE1 expression is not affected by ZAP. Wild type fibroblast cells were transduced with ZAPS or empty vector control lentiviruses and then infected with TB40E-GFP 48 hours post transduction. Total protein lysates were harvested every 24 hours and the expression levels of the viral proteins of each major kinetic class of HCMV were determined by western blot analysis. The result demonstrates that ZAPS overexpression leads to significant reduction of IE2 expression and downstream early and late viral proteins, whereas IE1 expression remains unaffected.

Figure 6



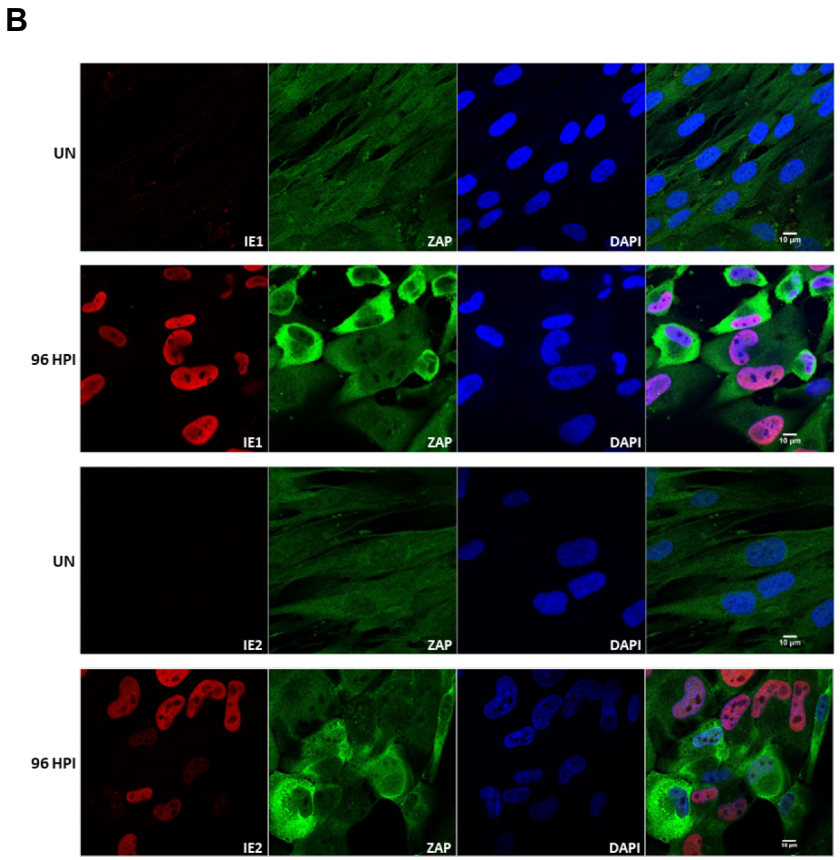


Figure 6. ZAP expression is reduced in IE2 positive cells. (A) Wild type fibroblast cells were infected with TB40E-GFP at an MOI of 3. Total protein lysates were harvested every 24 hours and the expression levels of ZAP and the viral proteins of each major kinetic class of HCMV were determined by western blot analysis. It demonstrates that ZAPS expression was robustly induced 24 hours post infection and remained higher than in uninfected cells throughout the course of the infection. (B) Wild type fibroblast cells were infected with HCMV 96 hours post infection, cells were fixed and stained with ZAP along with either IE1 or IE2. ZAP expression levels in IE1 or IE2 expressing cells were determined by confocal microscopy. Nuclei are stained with DAPI. (C) ZAP expression levels in 100 IE1-expressing cells and 100 IE2-expressing cells at 96 HPI were quantified and normalized to ZAP expression levels in uninfected cells from the confocal images using ZEN blue software.

Figure 7

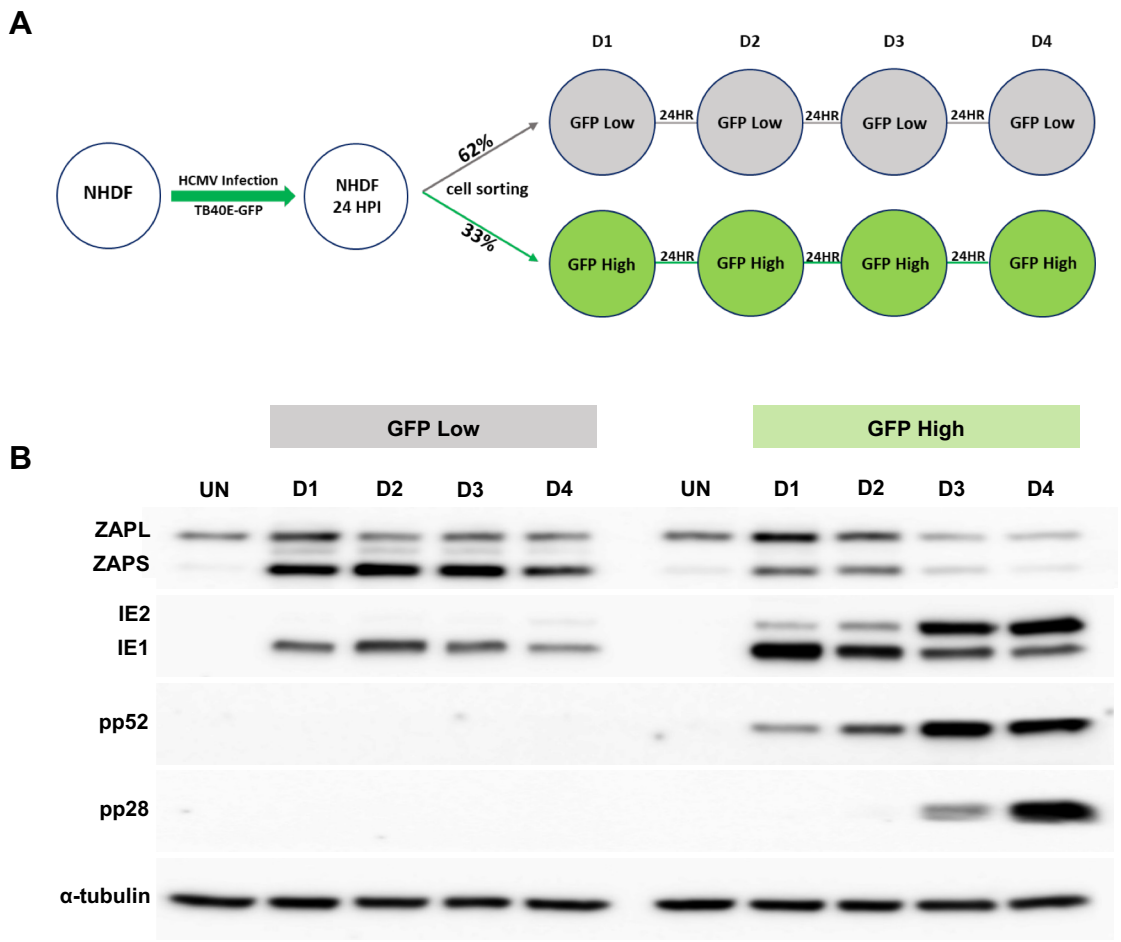


Figure 7. High ZAP expression correlates with failure in HCMV acute replication. (A) Diagram of cell sorting following TB40E-GFP infection. 24 hours post infection, fibroblast cells were sorted into low and high GFP expressing populations, and then re-seeded. Total protein lysates were harvested at the indicated time points. (B) Western blot analysis shows that the low GFP expressing population correlated with high levels of ZAPS expression and a failure in progression of virus replication, although IE1 expression can still be detected. In contrast, the high GFP expressing population correlated with low ZAPS expression and high levels of viral protein production, consistent with successful virus replication.

Figure 8

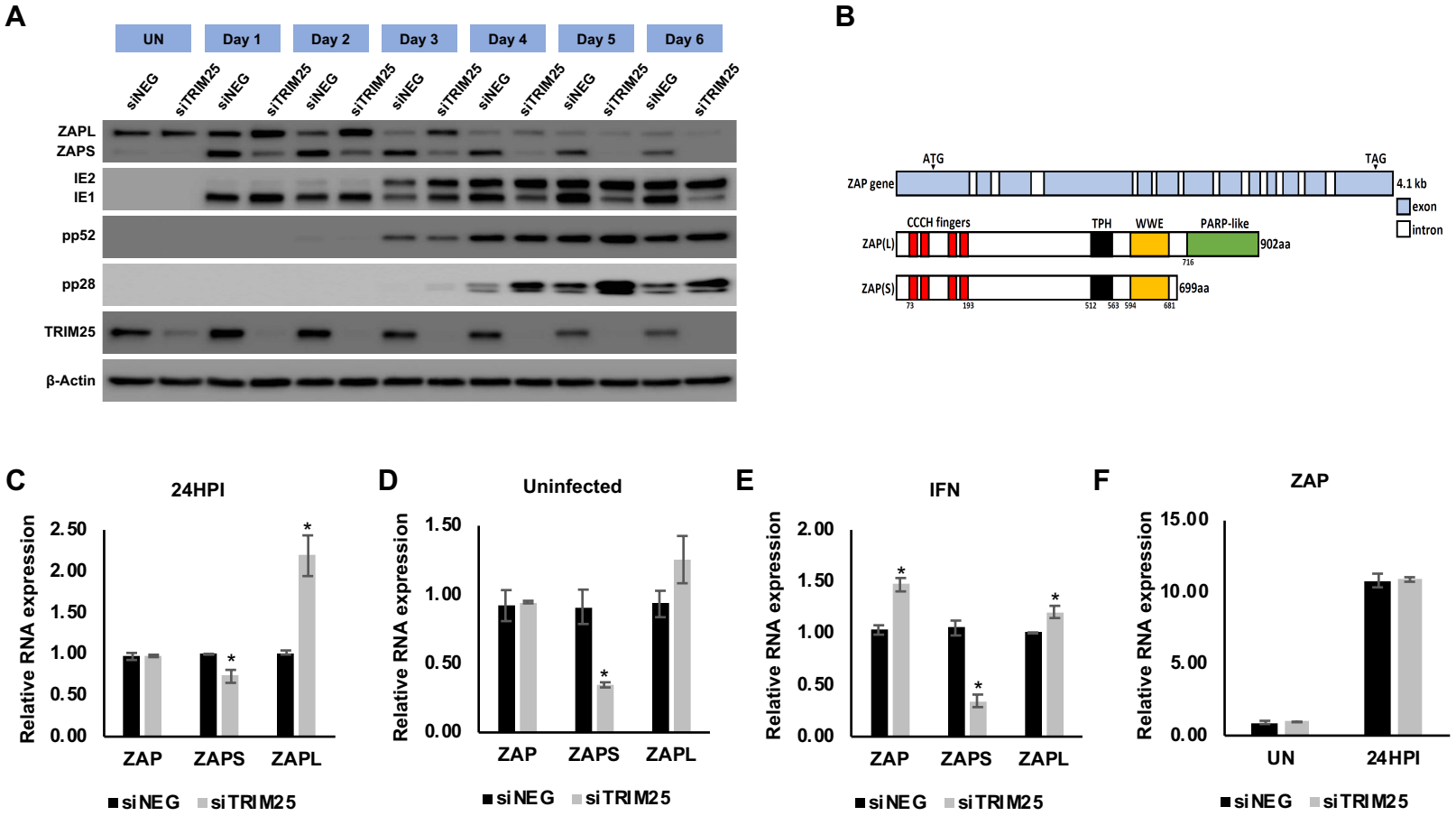


Figure 8. TRIM25 regulates differential splicing of ZAP. (A) Wildtype fibroblast cells were transfected with TRIM25 siRNA or negative control siRNA and infected with TB40/E-GFP 48 hours post transfection (MOI of 3). Total protein lysates were harvested every 24 hours and expression levels of ZAPS and ZAPL and viral proteins were determined by western blot analysis. The result demonstrates that knockdown of TRIM25 leads to substantial reductions in ZAPS expression and corresponding increase in ZAPL expression. (B) Diagram of the genomic structure of human ZAP gene showing the two major isoforms, ZAPS and ZAPL. Wildtype fibroblast cells were transfected with TRIM25 siRNA or negative control siRNA and infected with TB40E-GFP at an MOI of 3 (C), mock infected (D) or treated with IFN-α (E), 48 hours post transfection. Total RNA was harvested 24 hours later and primers specific to ZAPS, ZAPL or a shared exon were used to determined transcript levels by quantitative RT-qPCR analysis. Levels were normalised to GAPDH and compared to RNA levels from cells transfected with the control siRNA. (F) Levels of ZAP were compared between uninfected and infected cells showing no impact of TRIM25 knockdown on ZAP induction. N=2, * p-value < 0.05 based on student T test.