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

isoforms during HCMV infection and interferon induction, with knockdown of TRIM25 resulting in decreased ZAPS and corresponding increased ZAPL expression. These results demonstrate for the first time that ZAP is a potent host restriction factor against large DNA viruses and that HCMV evades ZAP detection through suppression of CpG dinucleotides within the major immediate early transcripts. Furthermore, TRIM25 is required for efficient upregulation of the interferon inducible short isoform of ZAP through regulation of alternative splicing.

Introduction

Interferon (IFN) is a crucial first line of defence against viral infection and shapes the adaptive immune response by triggering release of cytokines and chemokines [1,2]. IFN expression is triggered by the recognition of pathogen-associated molecular patterns (PAMPs) [1]. These microbe-specific molecular structures are generally essential for the survival of the microbes, but fundamentally different from the host. Examples of PAMPs include peptidoglycans, liposaccharide (LPS) and pathogen specific nucleic acid motifs, such as double stranded RNA and unmethylated CpG sequences within DNA. Cells recognize PAMPs through pattern recognition receptors (PRRs) that trigger innate immune responses following recognition of the target. Families of PRRs include the membrane bound Toll-like receptors, C-type lectin receptors, the cytoplasmic NOD-like receptors and RIG-I like receptors [3]. Upon recognition of the specific PAMP during invasion by a foreign pathogen, PRRs trigger signaling cascades that lead to relocation of IRF3/IRF7 complexes and NF-kB into the nucleus, initiating expression of type I IFN. In turn, activation of the IFN receptor leads to up-regulation of hundreds of IFN stimulated genes (ISGs) that, together, establish an antiviral cellular environment [4].

Dinucleotide representation in RNA sequences have been investigated as a potential PAMP [5-9]. Plant and vertebrate genomes show significantly lower CpG and TpA frequencies than would be expected given their overall base composition. Lower CpG frequencies are thought to have arisen through deamination of methylated cytosines in nuclear DNA, resulting in CpG sequences mutating to TpG [10]. RNA and small DNA viruses of vertebrates have evolved a similar pattern of suppressed CpG dinucleotides and artificially increasing dinucleotide frequencies within their viral genomes through synonymous mutations results in considerable attenuation of virus replication [6-9].

A recent study identified the short form of the zinc-finger antiviral protein (ZAP) and its associated factor TRIM25 as responsible for recognition of high CpG frequencies in viral RNA

[11]. Mammalian ZAP is expressed in two major isoforms, ZAPS (short) and ZAPL (long) which are generated by differential splicing, with ZAPL encoding an additional catalytically inactive poly (ADP-ribose) polymerase (PARP) domain [12]. ZAP had previously been identified as a host antiviral factor and is capable of binding to viral RNA through a pocket created by the second of four zinc fingers within the RNA binding domain [13-15]. TRIM25 is an E3 ubiquitin ligase and member of the tripartite motif (TRIM) family, many of which have been associated with antiviral functions [16,17]. TRIM25 is required for ZAPS antiviral activity, although the precise mechanism by which TRIM25 contributes to ZAPS antiviral activity is not fully understood [18,19]. A focused siRNA screen against human ISGs showed that knockdown of ZAP or TRIM25 rescued the replication of a defective HIV construct with artificially raised CpG levels [11]. Immunoprecipitation of ZAPS and sequencing of associated RNA demonstrated that ZAPS directly interacts with high CpG regions of HIV RNA. Knockdown of ZAPS has also been shown to rescue echovirus 7 virus that was attenuated through artificially increased CpG levels [20].

While CpG dinucleotide levels have clearly been shown to impact the fitness of RNA and small DNA viruses [6-9], their role in host recognition of large DNA viruses is less clear.

HCMV is a highly prevalent herpesvirus, persistently infecting between 30% and 100% of the human population, correlating with socio-economic status[21]. HCMV remains an important clinical pathogen accounting for more than 60% of complications associated with solid organ transplant patients [22-24]. It is also the leading cause of infectious congenital birth defects resulting from spread of the virus to neonates and has been linked to chronic inflammation and immune aging [25-27]. However, infection is normally asymptomatic due to effective control of virus replication by various arms of the immune system, including the interferon (IFN) response [2]. Although HCMV has evolved multiple mechanisms to subvert and inhibit the antiviral effects of IFNs, current evidence indicates they play a vital role in controlling replication and pathogenesis. Individuals with mutations in key IFN signalling genes are lethally susceptible to HCMV infections and recombinant IFN has been successfully used in treating congenital HCMV and HCMV infection in AIDS patients. Furthermore, murine CMV is more pathogenic in IFN knock-out mice than wildtype mice [2].

Infection with HCMV results in a robust IFN response. While viral attachment has been reported to be sufficient to trigger the IFN response, detection of HCMV is likely to occur through a multifactorial process [28]. The cellular sensors cGAS, IFI16 and ZBP1 have been shown to play a role in detection of the HCMV genome, while Toll-like receptors (TLRs) are

also thought to be important [29-32]. Detection leads to activation of signalling proteins, including IRF3, IRF7 and NFKB, resulting in up regulation of hundreds of ISGs [33-35]. Which of these ISGs are directly responsible for inhibiting HCMV is poorly understood. While the initial ISG response is robust, it is quickly shut down by the virus, resulting in a characteristic expression profile of rapid induction at early time points (up to 24 hours post infection) followed by efficient suppression of IFN regulated gene expression. This shut down is in part due to the expression of the viral immediate early genes IE1 and IE2, both of which have been shown to curtail the initial IFN response to HCMV [33,36,37].

To further dissect the role of the IFN response during HCMV infection, we used an arrayed lentivirus expression library to identify ISGs that inhibit HCMV. We show that ZAP and TRIM25 can potently inhibit HCMV replication and that HCMV has evolved to suppress CpG dinucleotides in the major immediate early transcript IE1 to evade detection by ZAP.

Results

Arrayed ISG expression screening identified ZAPS and TRIM25 as direct inhibitors of HCMV replication

Systematic dissection of the 'antiviral state' using arrayed ISG expression libraries is an effective method for identifying key components of the IFN response [38-42]. To identify ISGs that inhibit HCMV, human primary fibroblast cells were transduced with 421 individual ISG-encoding lentiviral vectors (or control empty vectors), in a 96-well plate format. Two days post transduction, cells were infected at a multiplicity of infection (MOI) of three with TB40/E-GFP, a BAC derived low passage HCMV strain containing an SV40 promoter driven eGFP cassette inserted between the viral TRS1 and US34 genes [43]. Levels of GFP were monitored over a seven-day period using a plate cytometer (Figure 1A). At seven days post infection, supernatant was transferred to fresh untransduced fibroblast cells to determine virus production levels. GFP levels were compared to the average signal for the respective plate. While this approach can identify which ISGs have the capacity to inhibit HCMV replication, many of the hits identified may inhibit virus replication by triggering IFN signalling through overexpression alone, independently of HCMV infection. To identify ISGs more likely to directly inhibit HCMV, as opposed to activation of IFN signalling, we performed a parallel screen using IRF3 KO fibroblast cells generated by CRISPR/Cas9 editing [44]. Percentage GFP levels were determined by comparing each ISG to the average GFP signal from each 96 well plate to control for inter-plate variations. Figure 1B compares the percentage virus production levels based on GFP signal for each ISG transduced well between wild type fibroblast cells

and IRF3 KO cells. As virus production is the most relevant measurement for the full virus replication cycle, our studies focused on the results from this screen (The data from the primary replication screen is shown in supplemental figure 1A-C and supplemental table 1). The blue box highlights ISGs that reduced HCMV virus production by more than 2-fold in unmodified fibroblast cells, but were not substantially inhibitory in IRF3 KO fibroblast cells. These IRF3-dependant inhibitory ISGs are highlighted in figure 1C and include cGAS, TLR3, MyD88 and DDX60, all known to act through IRF3 signalling. Additional ISGs, not previously known to act through IRF3, showed a similar pattern. The green box in figure 1B highlights ISGs that inhibit HCMV virus production in both wild type fibroblast cells and IRF3 KO cells, with figure 1D showing the relative virus production levels in each cell type. These IRF3-independent ISGs include those that signal through IRF3 independent pathways (eg IRF7) or are known to inhibit HCMV directly (IDO and RIPK2) validating the screening approach [45,46]. Interestingly, both ZAPS and TRIM25 inhibited HCMV virus production in an IRF3 independent manner, indicating they may play a role in directly inhibiting the virus. ZAPS and TRIM25 have recently been identified as host factors responsible for inhibiting HIV-1, Echo-7 and influenza virus constructs with synthetically increased CpG levels [9,11,20]. The two co-factors have also been reported to have antiviral activity against multiple viruses [47-51]. Based on this, we decided to further characterise the role of ZAPS and TRIM25 in the inhibition of HCMV replication.

To independently confirm the results of the screen, wild type fibroblast cells and IRF3 KO cells were transduced with independent lentiviral vector stocks expressing ZAPS, TRIM25, cGAS or an empty vector control, and infected with TB40/E-GFP (MOI of 3). GFP fluorescence levels were monitored for seven days (Figure 2A and B). As expected, efficient cGAS inhibition was largely dependent on IRF3 expression whereas virus replication was potently inhibited by ZAPS or TRIM25 in wild type and IRF3 KO cells. This confirmed our observations from the initial screen and implicated ZAPS and TRIM25 as direct inhibitors of HCMV virus replication. To confirm GFP reporter levels expressed by the virus accurately reflected virus replication, plaque assays were performed in wild type fibroblast cells transduced with lentivirus expressing either ZAPS or TRIM25 and compared to cells transduced with an empty control lentivirus. The results confirm that expression of ZAPS and TRIM25 significantly reduced viral replication (Figure 2C). To determine the effect of knockdown of ZAP and TRIM25 on HCMV replication, fibroblast cells were transfected with siRNAs targeting ZAP, TRIM25 or a negative control siRNA. Western blot analysis demonstrated efficient knockdown of ZAP and TRIM25 (Supplemental figure 2). Importantly, siRNA knockdown of ZAP or TRIM25 significantly increased HCMV replication, with titres increased by as much as 100-fold following knockdown of ZAP (Figure 2D and E). Thus, endogenous ZAP can inhibit HCMV and the rescue of HCMV replication presented here is likely an underestimate, as siRNAs do not

remove all the endogenous protein and ZAPS expression would be substantially higher in IFN-stimulated cells.

Distinct patterns of dinucleotide frequencies in subfamilies of herpesvirus genomes

Recent studies have identified ZAPS and TRIM25 as host factors involved in targeting and inhibiting RNA with high CpG dinucleotide frequencies [9,11,20]. As such they convey evolutionary pressure on RNA and small DNA viruses, resulting in suppression of CpG dinucleotide sequences within their genomes. However, the effects of host recognition of high CpG frequencies on larger DNA viruses, such as herpesviruses, has not been reported. Analysis of herpesvirus genomes suggests distinct patterns of CpG dinucleotide composition associated with the three subfamilies (alpha, beta and gamma) [52]. While alpha herpesviruses such as HSV-1 show no suppression of CpG dinucleotides within their genomes (Supplemental figure 3), substantial CpG suppression is seen throughout gamma herpesvirus genomes, such as Epstein-Barr virus (EBV) (Supplemental Figure 4). Beta herpesviruses demonstrate localised suppression of CpG sequences within the major immediate early genes. These genes are the first viral transcripts to be expressed following infection and they are critical for driving lytic replication of the virus and are thought to play a pivotal role during the establishment, maintenance and reactivation from latency. Analysis of sixteen beta herpesvirus genomes demonstrates extensive evolutionary conservation of suppressed CpG dinucleotide sequences specifically associated with immediate early transcripts (Figure 3). In contrast, there is no such pattern for the complementary GpC dinucleotide control, (Supplemental figure 5). These results indicate that the earliest beta herpesvirus transcripts have been selected for reduced CpG content, whereas these constraints do not extend to viral transcripts expressed at later times during infection.

IE1 expression is not affected by ZAP.

To determine the effect of ZAP expression on viral genes, the open reading frames of IE1, IE2 and pp52, an early gene with a high CpG dinucleotide frequency, were cloned into expression vectors. Cellular GAPDH tagged with T7, that has a low CpG frequency in line with mammalian genomes, was cloned into the same expression vector to generate a negative control. These plasmids were co-transfected with a plasmid expressing ZAPS. Expression of ZAPS had little effect on T7-GAPDH (Figure 4A). Similarly, while expression of ZAPS profoundly reduced pp52 expression by approximately 80% (Figure 4B) and IE2 expression to a lesser extent, ZAPS had no effect on IE1 expression. These results are in line with what would be predicted based on the CpG dinucleotide frequencies and confirm that suppression of CpG levels in IE1

allow this gene to evade the inhibitory effects of ZAPS. Interestingly, overexpression of ZAPL shows the same effects as ZAPS, suggesting that the longer isoform has the same ability to reduce expression of transcripts with high frequencies of CpG sequences (Figure 4C and D).

IE1 expression is not affected by ZAP during HCMV infection

As previously shown, overexpression of ZAPS profoundly inhibited virus replication (Figure 2A-C), and overexpression of ZAPS and ZAPL resulted in the reduced expression of HCMV viral proteins encoded by genes with high CpG frequencies following transient transfection experiments in HEK293T cells (Figure 4). To determine the effect of ZAPS on viral gene expression, western blot analysis was performed on total protein samples from wild type fibroblast cells transduced with lentivirus expressing ZAPS and infected two days later with TB40E-GFP (MOI of 3) (Figure 5). As is the case with all herpesviruses, viral gene expression occurs in a controlled temporal cascade with immediate early (IE) early (E) and late (L) gene expression occurring in a sequential fashion. Total protein was harvested at the indicated time points and levels of IE (IE1 and IE2), E (pp52) and L (pp28) proteins measured. Strikingly, ZAPS had little to no effect on IE1 expression levels throughout the time course, but substantially reduced IE2 levels and subsequent E and L viral gene expression. These results indicate that in the context of viral infection, IE1 expression is unaffected by ZAPS, likely due to highly suppressed CpG abundance in this transcript, enabling evasion of ZAPS.

Endogenous ZAP is induced during HCMV infection but expression is mutually exclusive to acute virus progression.

As overexpression of ZAP profoundly attenuates HCMV replication, we next determined the effects of endogenous ZAP during virus replication. Western blot analysis was performed on total protein lysates harvested from wild type fibroblast cells infected with TB40/E-GFP (MOI of 3), which results in close to 100% infection based on GFP expression. ZAPS expression was robustly induced 24 hours post infection and although levels decreased over time, the levels of endogenous ZAPS remained higher than in uninfected cells throughout the course of the infection (Figure 6A). ZAPL expression was not induced but also decreased over the course of the infection.

To determine whether ZAP expression levels correlated with HCMV infection at the individual cell level, we employed confocal microscopy. Cells were co-stained for ZAP and viral IE1 or IE2 expression in wild type fibroblast cells following infection with AD169, a laboratory adapted strain of HCMV. This strain was used instead of TB40/E-GFP as it does not express GFP,

which would interfere with the fluorescent signal. Uniform, low levels of ZAP expression can be seen throughout uninfected cells (Figure 6B). While almost all cells are IE1 positive following infection with HCMV at an MOI of 3, a mixed population of cells demonstrate high and low expression of ZAP. Multiple IE1 positive cells were observed with high levels of ZAP expression, further demonstrating that IE1 expression is not affected by high ZAP levels. In contrast IE2 expression and high levels of ZAP were mutually exclusive, suggesting progression of virus replication is blocked in cells expressing high levels of ZAP. This is confirmed by quantification of ZAP expression levels in IE1 and IE2 expressing cells (Figure 6C).

The existence of a mixed population of cells expressing high and low levels of ZAP, correlating with progression of virus replication is further supported by sorting of infected cell populations. Following infection with TB40/E-GFP at an MOI of 3, cells were sorted into high and low GFP expressing populations, 24 hours post infection, then reseeded and total protein levels determined at the indicated time points (Figure 7). Western blot analysis indicates that the low GFP expressing population correlated with high levels of ZAPS expression and a failure in progression of virus replication (although IE1 protein expression could still be detected). In contrast, high GFP expressing populations corresponded with lower ZAPS expression and high levels of viral protein production, consistent with efficient virus replication.

These data suggest that high levels of ZAP expression are mutually exclusive with the successful progression of virus replication. In contrast, IE1 expression is unaffected, consistent with suppressed CpG content facilitating evasion of ZAP detection.

TRIM25 is required for efficient ZAPS induction following HCMV infection or IFN treatment

TRIM25 is recognised as a key functional partner of ZAPS antiviral activity. However, the mechanism by which TRIM25 augments ZAPS antiviral activity is unclear. It has been shown that TRIM25 ubiquitinates ZAPS and itself through its E3 ubiquitin ligase function [18,19]. However, ZAPS antiviral activity does not seem to be dependent on TRIM25 ubiquitination. In this study, we show that like ZAPS, TRIM25 is a potent, IRF3 independent inhibitor of HCMV replication (Figure 1 and 2).

Strikingly, in addition to increasing viral protein expression, siRNA knockdown of TRIM25 results in a substantial reduction in ZAPS expression following infection with HCMV and a corresponding increase in ZAPL expression (Figure 8A). As previously described, ZAPS and ZAPL are expressed from the same primary transcript through differential splicing and only

ZAPS expression is considered to be stimulated by IFN treatment, indicating that differential splicing is involved in the regulation of ZAPS and ZAPL levels (Figure 8B).

To determine whether the failure in ZAPS induction following TRIM25 knockdown was at the RNA or protein level, ZAPS and ZAPL RNA levels were determined by RT-qPCR. Fibroblast cells were transfected with a control siRNA or siRNA targeting TRIM25 and infected with TB40/E-GFP. Total RNA was harvested 24 HPI and primers specific to ZAPS, ZAPL or to a shared region of ZAP were used to measure transcript levels. While total levels of ZAP RNA were unaffected by TRIM25 knockdown, levels of ZAPS were significantly reduced with a corresponding increase in ZAPL levels (Figure 8C). This effect is not specific to HCMV infection as the same observation can be seen in uninfected fibroblast cells and cells treated with IFN (Figure 8D and E). Furthermore, ZAP induction, as well as induction of other well characterised ISGs, was not reduced by TRIM25 knockdown ruling out a general failure in IFN induction (Figure 8F and Supplemental figure 6). The results clearly show that efficient IFN induced gene expression occurs despite TRIM25 knockdown, suggesting the effect is specific to differential splicing of ZAP. Although these results do not rule out the possibility that TRIM25 regulates the antiviral function of ZAPS through direct protein-protein interaction, they explain how TRIM25 contributes to the activity of ZAPS by regulation of ZAPS induction through differential splicing.

Discussion

Associated pathologies and therapeutic potential of HCMV is dependent on the host immune response against the virus. As IFNs shape the innate and adaptive responses to the virus, understanding how the IFN response is regulated during HCMV infection and how the virus subverts this response could have important implications for our understanding of diseases associated with the virus as well as for the rational design of vaccines and cancer therapeutics.

Although it is clear that the IFN response is important during HCMV infection, which of the hundreds of induced ISGs play a critical role is less clear. Here we present a systematic analysis, determining the effect of over 400 individual ISGs on HCMV replication. By combining screens in wild type and IRF3 knockout cells we were able to define specific subsets of ISGs that were IRF3 independent and therefore more likely to represent direct inhibitors of virus replication. These included IDO and RIPK2, which have previously been identified as important antiviral factors during HCMV infection [45,46]. In addition, multiple novel inhibitors were identified, including ZAP and TRIM25, which have previously been shown to act in a coordinated antiviral fashion [18,19].

ZAP was originally identified as an antiviral protein by screening a rat cDNA library for factors that inhibited the replication of Moloney murine leukemia retrovirus [13]. Subsequent studies have demonstrated that ZAP has antiviral activity against a range of RNA viruses, including HIV, filoviruses, flaviviruses, coxsackievirus B3, influenza A virus, Newcastle disease virus and Hepatitis B virus, a partially double stranded DNA virus [48-51,53-55]. Initial studies indicated that ZAP antiviral activity was based on direct binding to viral RNA and degradation through recruitment of exosome complex components and inhibition of translation [13,14,56]. Further studies have suggested that ZAPS interacts with RIG-I to stimulate IFN expression through IRF3 signalling [54]. This is consistent with the partial rescue of ZAPS inhibition of HCMV in IRF3 KO cells observed in this study, suggesting a dual functionality of ZAPS in HCMV antiviral activity, acting both as a pathogen sensing protein and as a direct antiviral factor.

Until recently, the sequence or sequence characteristics that defined ZAPS RNA binding specificity were not known. However, a study using a focused siRNA screen identified ZAPS and TRIM25 as host factors responsible for inhibition of HIV-1 constructs with artificially raised CpG dinucleotide frequencies and it was shown that ZAPS specifically binds to RNA sequences at CpG dinucleotide motifs [13]. The molecular mechanism for this binding has now been established through X-ray crystallography of the ZAP RNA binding domain complexed with a target CpG RNA, which demonstrates the second of four zinc fingers creates a highly basic patch that is required for specific binding of CpG dinucleotides [57].

It had previously been suggested that mammalian RNA and small DNA viruses suppress CpG dinucleotides to mimic the composition of their host genomic make up, thereby avoiding recognition as foreign nucleic acid [6-8]. While CpG bias has previously been reported for herpesviruses, this bias was attributed to potential methylation status rather than evasion of a host antiviral response [52]. The pattern of CpG dinucleotide frequencies in herpesviruses is so distinctive that, with few exceptions, viruses can be attributed to the main three subfamilies based on the CpG dinucleotide frequency patterns within open reading frame (Supplemental figures 1-3). The majority of alpha-herpesviruses demonstrate little or no CpG suppression, while gamma-herpesviruses demonstrate substantial suppression across the genome. Beta-herpesviruses display the most striking pattern of all, with suppression of CpG dinucleotides linked to temporal gene expression and specifically restricted to the immediate early genes. It is intriguing to speculate that the genomes of these large DNA viruses may have been so dramatically moulded by a single host antiviral mechanism, although other factors may also play a role.

The fact that CpG dinucleotide patterns are so uniform across the virus sub-families also suggests a central underlying biological relevance. For example, alpha-herpesviruses are mainly associated with latent infections in neuronal cells where the interferon response may be limited due to immune privilege and constitutive expression levels of ZAP are low [20], whereas beta-herpesviruses and gamma-herpesviruses are mainly associated with latency in haematopoietic cells and are therefore under greater pressure to evade host cell IFN responses [58]. Alternatively, alpha-herpesviruses may have evolved a robust mechanism that directly counteracts ZAP antiviral activity, making it unnecessary to repress CpG dinucleotides. This is supported by a previous study that showed HSV-1 was able to replicate to wild type levels despite expression of a ZAP construct, although the construct expressed a truncated version of rat ZAP, rather than full length human ZAP [47]. Further studies will be required to determine whether the CpG dinucleotide frequency in herpesvirus sub-families is directly related to ZAP expression and why these patterns are so distinctively associated with the specific virus sub-families.

Multiple studies have linked TRIM25 with efficient ZAP antiviral activity, and, like ZAPS, our screen identified TRIM25 as an IRF3 independent inhibitor of HCMV. TRIM25 is an E3 ligase that catalyses ubiquitylation and ISGylation of target proteins [16,17]. TRIM25 contains a zinc ring finger, B-box, coiled coil and PRY/SPRY domain. Both Takata *et al* and Li *et al* demonstrated that TRIM25 is required for ZAP antiviral activity through siRNA screens [11,18], whereas Zheng *et al* identified TRIM25 as a ZAP interacting factor through affinity purification [19]. These studies demonstrated that TRIM25 interacts directly with ZAP through the PRY/SPRY domain and ubiquitinates ZAP and itself. While ubiquitination requires binding of ZAP and TRIM25 to RNA it does not seem to be directly necessary for ZAP antiviral activity [18,19,59]. TRIM25 also inhibits influenza A virus in a ZAP independent manner through direct binding to the viral ribonucleoprotein complex [60]. We show that TRIM25 expression potently inhibits HCMV replication although further studies will be required to determine whether this occurs in a ZAP dependent or independent manner.

In addition, our studies suggest that TRIM25 regulates the alternative splicing of ZAP. A recent study demonstrated that alternative splicing of the ZAP primary transcript leads to at least four different isoforms; ZAPS, ZAPM, ZAPL and ZAPXL [61], although our study focused on the major isoforms ZAPS and ZAPL. ZAPL contains a catalytically inactive PARP-like domain at the C terminus, that is missing from ZAPS. Both ZAPS and ZAPL have been reported to have antiviral activity, although ZAPS was identified as the host factor responsible for CpG recognition [13]. In our study we found that both ZAPS and ZAPL could target high CpG transcripts. Following virus infection or IFN treatment ZAP is induced, however differential

splicing results in higher levels of induction of ZAPS compared to ZAPL. Following HCMV infection, TRIM25 knockdown resulted in significantly lower levels of ZAPS protein and RNA levels with a corresponding increase in ZAPL levels indicating a TRIM25 dependant change in differential splicing of ZAP. Regulation of ZAP splicing by TRIM25 occurs in uninfected cells and IFN stimulated cells, indicating it is not specific to HCMV but rather a fundamental aspect of ZAP regulation. It is unclear at this point whether TRIM25 plays a direct role in splicing of ZAP or whether the regulation occurs through an intermediate signalling pathway. TRIM25 has been shown to directly bind RNA through its PRY/SPRY domain [59], and plays a role in the maturation of the microRNA let-7 [62], suggesting a more direct role in processing ZAP RNA is possible. However further studies will be required to dissect its precise role [59].

Increasing the CpG levels within the IE1 coding region could be a viable approach for generating safer live attenuated vaccines with reduced risk of reversion as the mutations could be spread across the entire length of the IE1 gene. Increasing the CpG levels in IE1 could also increase the immunogenicity of the virus by triggering more robust IFN responses, which could improve HCMV vaccines as well as the use of HCMV as a vaccine vector and as a therapeutic cancer vaccine. However, construction of such viruses may be challenging due to the central role the major immediate early genes play in acute replication of the virus and the potential for introduced mutations to disrupt the complex splicing events or regulatory regions embedded within the coding region of IE1. Future experiments will be required to determine the potential application of such viruses.

Materials and Methods

Cell culture and viral infection

Normal Human Dermal Fibroblasts (NHDF; Gibco) and IRF3 -/- cells were maintained in Dulbecco's modified high glucose media (DMEM; Sigma) supplemented with 10% fetal bovine serum (FBS; Gibco) and 1% penicillin-streptomycin (Invitrogen). IRF3 -/- cells were provided by Victor DeFillipis. A low passage HCMV strain TB40E-GFP [43], which expresses GFP from an SV40 promoter was used for arrayed ISG expression library screening, western blot analysis, Real-Time PCR, and flow cytometry analysis. Laboratory adapted HCMV strain AD169 was used for immunofluorescence experiments.

Arrayed ISG lentivirus expression library screening

The human lentivirus expression library has been previously described [42]. In brief the library encodes 420 human ISG genes on the pSCRPSY backbone (KT368137.1). Normal human

dermal fibroblast cells and IRF3-/- cells were seeded in 96-well plates a day before transduction. Next day, cells reached 90-95% confluency and were transduced with the ISG library. Transduced cells were incubated for 48 hours and then infected with GFP expressing TB40E virus at an MOI of three. GFP intensity was monitored every 24 hours with Synergy HT microplate reader (Biotek).

siRNA transfection

NHDFs were seeded in 6-well plates a day before siRNA transfection. Next day, cells reached 90-95% confluency and were transfected with siRNA twice (4 hours apart between first and second transfections) using Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's protocol. siRNAs used in this article are Human ON-TARGETplus siRNAs against TRIM25 and ZC3HAV1 (4 individual siRNAs per gene; Dharmacon). Transfected NHDFs were incubated for 48 hours and then infected with GFP expressing TB40E virus at an MOI of three.

Western blot analysis

Cells were lysed with RIPA buffer (0.1% SDS, 1% Triton X-100, 1% deoxycholate, 5 mM EDTA, 150 mM NaCl, and 10 mM Tris at pH 7.2) containing protease inhibitor cocktail (Roche). Ten micrograms of the total lysate was separated in 10% SDS-polyacrylamide gels and transferred to PVDF membranes (Millipore). Primary antibodies used in this paper are mouse anti-CMV IE1/2 monoclonal antibody (MAB8131, Millipore), mouse anti-CMV pp52 monoclonal antibody (CH16, Santa Cruz Biotechnology), mouse anti-CMV pp28 monoclonal antibody (CH19, Santa Cruz Biotechnology), rabbit anti-ZAP polyclonal antibody (PA5-31650, Invitrogen), mouse anti-TRIM25 monoclonal antibody (BD Biosciences), rabbit anti-T7 tag monoclonal antibody (D9E1X, Cell Signaling Technology), mouse anti-alpha tubulin monoclonal antibody (DM1A, Abcam) and mouse anti-β-Actin monoclonal antibody (Abcam). Blots were probed with primary antibody (1:500-1:5000) diluted in 5% dehydrated milk in Tris Buffered Saline (TBS) and subsequently the HRP-conjugated secondary antibodies (Pierce) at 1:5000. Blots were washed in TBS three times, incubated with chemiluminescent substrate (SuperSignal West Pico; Thermo Scientific) according to the manufacturer's protocol, and exposed in G:Box (Syngene) for visualization of bands.

Vector Construction and transfection

The expression vectors used in the article were made by cloning the coding sequences of UL123 (IE1, TB40E), UL122 (IE2, TB40E), UL44 (pp52, TB40E), Human GAPDH with double T7-tag at the N-terminus, into the pcDNA3.1 vector. The ZAP(L) and ZAP(S) expression

vectors were made by cloning the coding sequences of ZAP(L) and ZAP(S) into the pcDEF3.1 vector derived from pcDNA3.1 with the CMV promoter replaced by the EF1alpha promoter.

Immunofluorescence

Laboratory adapted HCMV strain AD169 infected cells were fixed in 4% paraformaldehyde solution for 20 minutes and then permeabilized in Methanol:Acetone solution (1:1) at -20^{o}C for 7 minutes, and then blocked with 5% human serum in PBS for 30 minutes. Primary and secondary antibodies were diluted with 5% human serum in PBS. Cells were washed with PBS after primary and after secondary antibody incubations. Primary antibodies used in this paper are mouse anti-CMV IE2 monoclonal antibody (12E2, Santa Cruz Biotechnology), mouse anti-CMV IE1/2 monoclonal antibody (MAB8131, Millipore), and rabbit anti-ZAP polyclonal antibody (PA5-31650, Invitrogen) at 1:500. Alexa-fluor-647 conjugated goat anti-mouse or Alexa-fluor-488 conjugated goat-anti-rabbit IgG secondary antibodies were diluted 1:1000. All images were acquired with Zeiss LSM 710 confocal microscope fitted with 63X/1.4 oil-immersion objective lens.

Real-Time PCR analysis

Total RNA was isolated by using Trizol solution according to the manufacturer's protocol followed by DNase (Turbo DNA-free kit, Ambion) treatment, and then reverse transcribed with poly T primers using High Capacity cDNA Reverse Transcription Kit (Invitrogen). Real-Time PCR was carried out using by Taqman assays with pre-designed gene-specific primer/probe set (Applied Biosystems) on Rotor gene 3000 (Corbet Research). Custom primer/probe set are TCCTCTCTCAGGATCTGTATGT, GGAGAGGAAGGAGTCAAAGATG, and 56-FAM/ ACCATCTAC/ZEN/CCATTGGCTCAAGCA/3IABkFQ for ZAP(L), and AGCATGGTGTGACT GAAAGG, CTTCACAGCTGGAGAAGCTAAA, and 56-FAM/TCTGAAAGG/ZEN/GAAGTCTG AGCGAGTCT/3IABkFQ for ZAP(S).

Fluorescence-activated cell sorting (FACS)

NHDFs were infected with GFP expressing TB40E virus at an MOI of three. 24 hours later, the cells were trypsinized, and resuspended in PBS buffer. The resuspended cells were then sorted into low GFP (parameter set using uninfected cells) or high GFP population by the BD FACSAria IIIu cell sorter. Immediately after sorting, a small portion of the cells from each population were collected and lysed with RIPA buffer, and the remaining low GFP and high GFP cells were re-seeded into a 6-well plate respectively and incubated for another 24 to 72 hours before lysed with RIPA buffer.

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).

8 Simmonds, P., Xia, W., Baillie, J. K. & McKinnon, K. Modelling mutational and selection pressures on dinucleotides in eukaryotic phyla--selection against CpG and UpA in cytoplasmically expressed RNA and in RNA viruses. *BMC Genomics* **14**, 610, doi:10.1186/1471-2164-14-610 (2013).

9 Gaunt, E. *et al.* Elevation of CpG frequencies in influenza A genome attenuates pathogenicity but enhances host response to infection. *Elife* **5**, e12735, doi:10.7554/eLife.12735 (2016).

10 Coulondre, C., Miller, J. H., Farabaugh, P. J. & Gilbert, W. Molecular basis of base substitution hotspots in Escherichia coli. *Nature* **274**, 775-780, doi:10.1038/274775a0 (1978).

11 Takata, M. A. *et al.* CG dinucleotide suppression enables antiviral defence targeting non-self RNA. *Nature* **550**, 124-127, doi:10.1038/nature24039 (2017).

12 Kerns, J. A., Emerman, M. & Malik, H. S. Positive selection and increased antiviral activity associated with the PARP-containing isoform of human zinc-finger antiviral protein. *PLoS Genet* **4**, e21, doi:10.1371/journal.pgen.0040021 (2008).

13 Gao, G., Guo, X. & Goff, S. P. Inhibition of retroviral RNA production by ZAP, a CCCH-type zinc finger protein. *Science* **297**, 1703-1706, doi:10.1126/science.1074276 (2002).

14 Guo, X., Carroll, J. W., Macdonald, M. R., Goff, S. P. & Gao, G. The zinc finger antiviral protein directly binds to specific viral mRNAs through the CCCH zinc finger motifs. *J Virol* **78**, 12781-12787, doi:10.1128/JVI.78.23.12781-12787.2004 (2004).

15 Meagher, J. L. *et al.* Structure of the zinc-finger antiviral protein in complex with RNA reveals a mechanism for selective targeting of CG-rich viral sequences. *Proc Natl Acad Sci U S A* **116**, 24303-24309, doi:10.1073/pnas.1913232116 (2019).

16 Zou, W. & Zhang, D. E. The interferon-inducible ubiquitin-protein isopeptide ligase (E3) EFP also functions as an ISG15 E3 ligase. *J Biol Chem* **281**, 3989-3994, doi:10.1074/jbc.M510787200 (2006).

17 Khan, R., Khan, A., Ali, A. & Idrees, M. The interplay between viruses and TRIM family proteins. *Rev Med Virol* **29**, e2028, doi:10.1002/rmv.2028 (2019).

18 Li, M. M. *et al.* TRIM25 Enhances the Antiviral Action of Zinc-Finger Antiviral Protein (ZAP). *PLoS Pathog* **13**, e1006145, doi:10.1371/journal.ppat.1006145 (2017).

19 Zheng, X. *et al.* TRIM25 Is Required for the Antiviral Activity of Zinc Finger Antiviral Protein. *J Virol* **91**, doi:10.1128/JVI.00088-17 (2017).

20 Odon, V. *et al.* The role of ZAP and OAS3/RNAseL pathways in the attenuation of an RNA virus with elevated frequencies of CpG and UpA dinucleotides. *Nucleic Acids Res* **47**, 8061-8083, doi:10.1093/nar/gkz581 (2019).

21 Fields, B. N. K. D. M. H. P. M. *Fields virology*. (Wolters Kluwer Health/Lippincott Williams & Wilkins, 2013).

22 Kotton, C. N. CMV: Prevention, Diagnosis and Therapy. *Am J Transplant* **13 Suppl 3**, 24-40; quiz 40, doi:10.1111/ajt.12006 (2013).

23 Kowalsky, S., Arnon, R. & Posada, R. Prevention of cytomegalovirus following solid organ transplantation: a literature review. *Pediatr Transplant* **17**, 499-509, doi:10.1111/petr.12118 (2013).

24 Bruminhent, J. & Razonable, R. R. Management of cytomegalovirus infection and disease in liver transplant recipients. *World journal of hepatology* **6**, 370-383, doi:10.4254/wjh.v6.i6.370 (2014).

25 Ballard, R. A., Drew, W. L., Hufnagle, K. G. & Riedel, P. A. Acquired cytomegalovirus infection in preterm infants. *Am J Dis Child* **133**, 482-485 (1979).

26 Griffith, D. M. *et al.* Systemic inflammation after critical illness: relationship with physical recovery and exploration of potential mechanisms. *Thorax* **71**, 820-829, doi:10.1136/thoraxjnl-2015-208114 (2016).

27 Jergovic, M., Contreras, N. A. & Nikolich-Zugich, J. Impact of CMV upon immune aging: facts and fiction. *Med Microbiol Immunol* **208**, 263-269, doi:10.1007/s00430-019-00605-w (2019).

28 Juckem, L. K., Boehme, K. W., Feire, A. L. & Compton, T. Differential initiation of innate immune responses induced by human cytomegalovirus entry into fibroblast cells. *J Immunol* **180**, 4965-4977, doi:10.4049/jimmunol.180.7.4965 (2008).

29 Paijo, J. *et al.* cGAS Senses Human Cytomegalovirus and Induces Type I Interferon Responses in Human Monocyte-Derived Cells. *PLoS Pathog* **12**, e1005546, doi:10.1371/journal.ppat.1005546 (2016).

30 Lio, C. W. *et al.* cGAS-STING Signaling Regulates Initial Innate Control of Cytomegalovirus Infection. *J Virol* **90**, 7789-7797, doi:10.1128/JVI.01040-16 (2016).

31 Diner, B. A., Lum, K. K., Toettcher, J. E. & Cristea, I. M. Viral DNA Sensors IFI16 and Cyclic GMP-AMP Synthase Possess Distinct Functions in Regulating Viral Gene Expression, Immune Defenses, and Apoptotic Responses during Herpesvirus Infection. *MBio* **7**, doi:10.1128/mBio.01553-16 (2016).

32 DeFilippis, V. R., Alvarado, D., Sali, T., Rothenburg, S. & Fruh, K. Human cytomegalovirus induces the interferon response via the DNA sensor ZBP1. *J Virol* **84**, 585-598, doi:10.1128/JVI.01748-09 (2010).

33 Browne, E. P., Wing, B., Coleman, D. & Shenk, T. Altered cellular mRNA levels in human cytomegalovirus-infected fibroblasts: viral block to the accumulation of antiviral mRNAs. *J Virol* **75**, 12319-12330, doi:10.1128/JVI.75.24.12319-12330.2001 (2001).

34 DeFilippis, V. R. *et al.* Interferon regulatory factor 3 is necessary for induction of antiviral genes during human cytomegalovirus infection. *J Virol* **80**, 1032-1037, doi:10.1128/JVI.80.2.1032-1037.2006 (2006).

35 Yurochko, A. D., Kowalik, T. F., Huong, S. M. & Huang, E. S. Human cytomegalovirus upregulates NF-kappa B activity by transactivating the NF-kappa B p105/p50 and p65 promoters. *J Virol* **69**, 5391-5400 (1995).

36 Taylor, R. T. & Bresnahan, W. A. Human cytomegalovirus immediate-early 2 gene expression blocks virus-induced beta interferon production. *J Virol* **79**, 3873-3877, doi:10.1128/JVI.79.6.3873-3877.2005 (2005).

37 Paulus, C., Krauss, S. & Nevels, M. A human cytomegalovirus antagonist of type I IFN-dependent signal transducer and activator of transcription signaling. *Proc Natl Acad Sci U S A* **103**, 3840-3845, doi:10.1073/pnas.0600007103 (2006).

38 Feng, J. *et al.* Interferon-Stimulated Gene (ISG)-Expression Screening Reveals the Specific Antibunyaviral Activity of ISG20. *J Virol* **92**, doi:10.1128/JVI.02140-17 (2018).

39 Rihn, S. J. *et al.* TRIM69 Inhibits Vesicular Stomatitis Indiana Virus. *J Virol* **93**, doi:10.1128/JVI.00951-19 (2019).

40 Dittmann, M. *et al.* A serpin shapes the extracellular environment to prevent influenza A virus maturation. *Cell* **160**, 631-643, doi:10.1016/j.cell.2015.01.040 (2015).

41 Kane, M. *et al.* Identification of Interferon-Stimulated Genes with Antiretroviral Activity. *Cell host & microbe* **20**, 392-405, doi:10.1016/j.chom.2016.08.005 (2016).

42 Schoggins, J. W. *et al.* A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* **472**, 481-485, doi:10.1038/nature09907 (2011).

43 Umashankar, M. *et al.* A novel human cytomegalovirus locus modulates cell type-specific outcomes of infection. *PLoS Pathog* **7**, e1002444, doi:10.1371/journal.ppat.1002444 (2011).

44 Sali, T. M. *et al.* Characterization of a Novel Human-Specific STING Agonist that Elicits Antiviral Activity Against Emerging Alphaviruses. *PLoS Pathog* **11**, e1005324, doi:10.1371/journal.ppat.1005324 (2015).

45 Kapoor, A., Forman, M. & Arav-Boger, R. Activation of nucleotide oligomerization domain 2 (NOD2) by human cytomegalovirus initiates innate immune responses and restricts virus replication. *PLoS One* **9**, e92704, doi:10.1371/journal.pone.0092704 (2014).

46 Zimmermann, A. *et al.* Checks and balances between human cytomegalovirus replication and indoleamine-2,3-dioxygenase. *J Gen Virol* **95**, 659-670, doi:10.1099/vir.0.061994-0 (2014).

47 Bick, M. J. *et al.* Expression of the zinc-finger antiviral protein inhibits alphavirus replication. *J Virol* **77**, 11555-11562, doi:10.1128/jvi.77.21.11555-11562.2003 (2003).

48 Chiu, H. P. *et al.* Inhibition of Japanese encephalitis virus infection by the host zinc-finger antiviral protein. *PLoS Pathog* **14**, e1007166, doi:10.1371/journal.ppat.1007166 (2018).

49 Liu, C. H., Zhou, L., Chen, G. & Krug, R. M. Battle between influenza A virus and a newly identified antiviral activity of the PARP-containing ZAPL protein. *Proc Natl Acad Sci U S A* **112**, 14048-14053, doi:10.1073/pnas.1509745112 (2015).

50 Muller, S. *et al.* Inhibition of filovirus replication by the zinc finger antiviral protein. *J Virol* **81**, 2391-2400, doi:10.1128/JVI.01601-06 (2007).

51 Zhu, Y. *et al.* Zinc-finger antiviral protein inhibits HIV-1 infection by selectively targeting multiply spliced viral mRNAs for degradation. *Proc Natl Acad Sci U S A* **108**, 15834-15839, doi:10.1073/pnas.1101676108 (2011).

52 Honess, R. W. *et al.* Deviations from expected frequencies of CpG dinucleotides in herpesvirus DNAs may be diagnostic of differences in the states of their latent genomes. *J Gen Virol* **70 (Pt 4)**, 837-855, doi:10.1099/0022-1317-70-4-837 (1989).

53 Mao, R. *et al.* Inhibition of hepatitis B virus replication by the host zinc finger antiviral protein. *PLoS Pathog* **9**, e1003494, doi:10.1371/journal.ppat.1003494 (2013).

54 Hayakawa, S. *et al.* ZAPS is a potent stimulator of signaling mediated by the RNA helicase RIG-I during antiviral responses. *Nat Immunol* **12**, 37-44, doi:10.1038/ni.1963 (2011).

55 Li, M. *et al.* Zinc finger antiviral protein inhibits coxsackievirus B3 virus replication and protects against viral myocarditis. *Antiviral Res* **123**, 50-61, doi:10.1016/j.antiviral.2015.09.001 (2015).

56 Guo, X., Ma, J., Sun, J. & Gao, G. The zinc-finger antiviral protein recruits the RNA processing exosome to degrade the target mRNA. *Proc Natl Acad Sci U S A* **104**, 151-156, doi:10.1073/pnas.0607063104 (2007).

57 Meagher, J. L. *et al.* Structure of the zinc-finger antiviral protein in complex with RNA reveals a mechanism for selective targeting of CG-rich viral sequences. *Proc Natl Acad Sci U S A*, doi:10.1073/pnas.1913232116 (2019).

58 Fields, B. N., Knipe, D. M. & Howley, P. M. *Fields' virology*. 5th ed. / editors-in-chief, David M. Knipe, Peter M. Howley ; associate editors, Diane E. Griffin ... [et al.]. edn, (Wolters Kluwer/Lippincott Williams & Wilkins, 2007).

59 Choudhury, N. R. *et al.* RNA-binding activity of TRIM25 is mediated by its PRY/SPRY domain and is required for ubiquitination. *BMC Biol* **15**, 105, doi:10.1186/s12915-017-0444-9 (2017).

60 Meyerson, N. R. *et al.* Nuclear TRIM25 Specifically Targets Influenza Virus Ribonucleoproteins to Block the Onset of RNA Chain Elongation. *Cell Host Microbe* **22**, 627-638 e627, doi:10.1016/j.chom.2017.10.003 (2017).

61 Li, M. M. H. *et al.* Characterization of Novel Splice Variants of Zinc Finger Antiviral Protein (ZAP). *J Virol* **93**, doi:10.1128/JVI.00715-19 (2019).

62 Choudhury, N. R. *et al.* Trim25 Is an RNA-Specific Activator of Lin28a/TuT4-Mediated Uridylation. *Cell Rep* **9**, 1265-1272, doi:10.1016/j.celrep.2014.10.017 (2014).

Figure 1

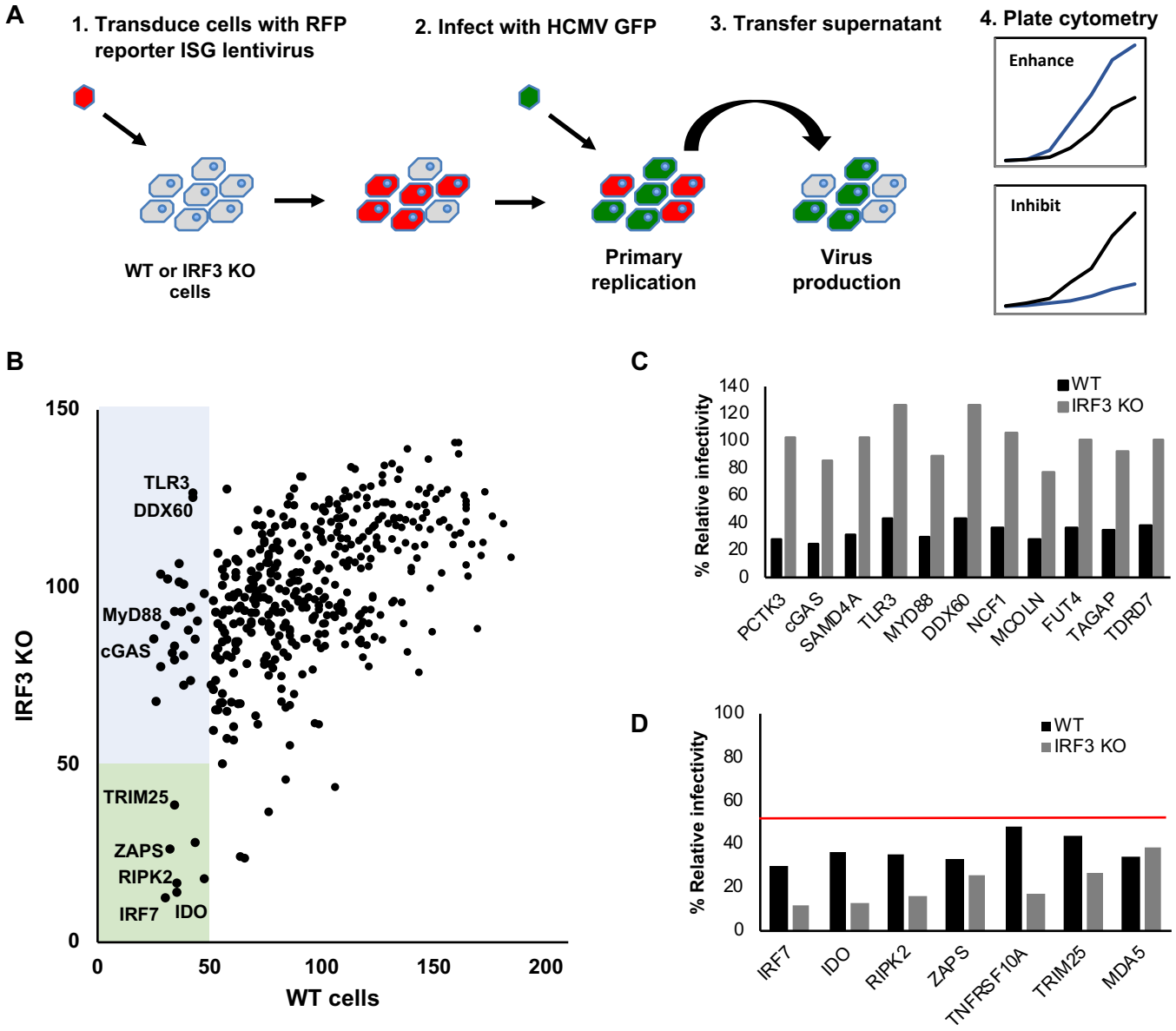

Figure 1. Arrayed ISG expression screening identified ZAPS and TRIM25 as direct inhibitors of HCMV replication. (A) Wild type and IRF3 KO fibroblast cells were seeded in 96-well plates and transduced with arrayed ISG lentivirus library. Cells were infected at 48 hours post-transduction with TB40/E-GFP (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

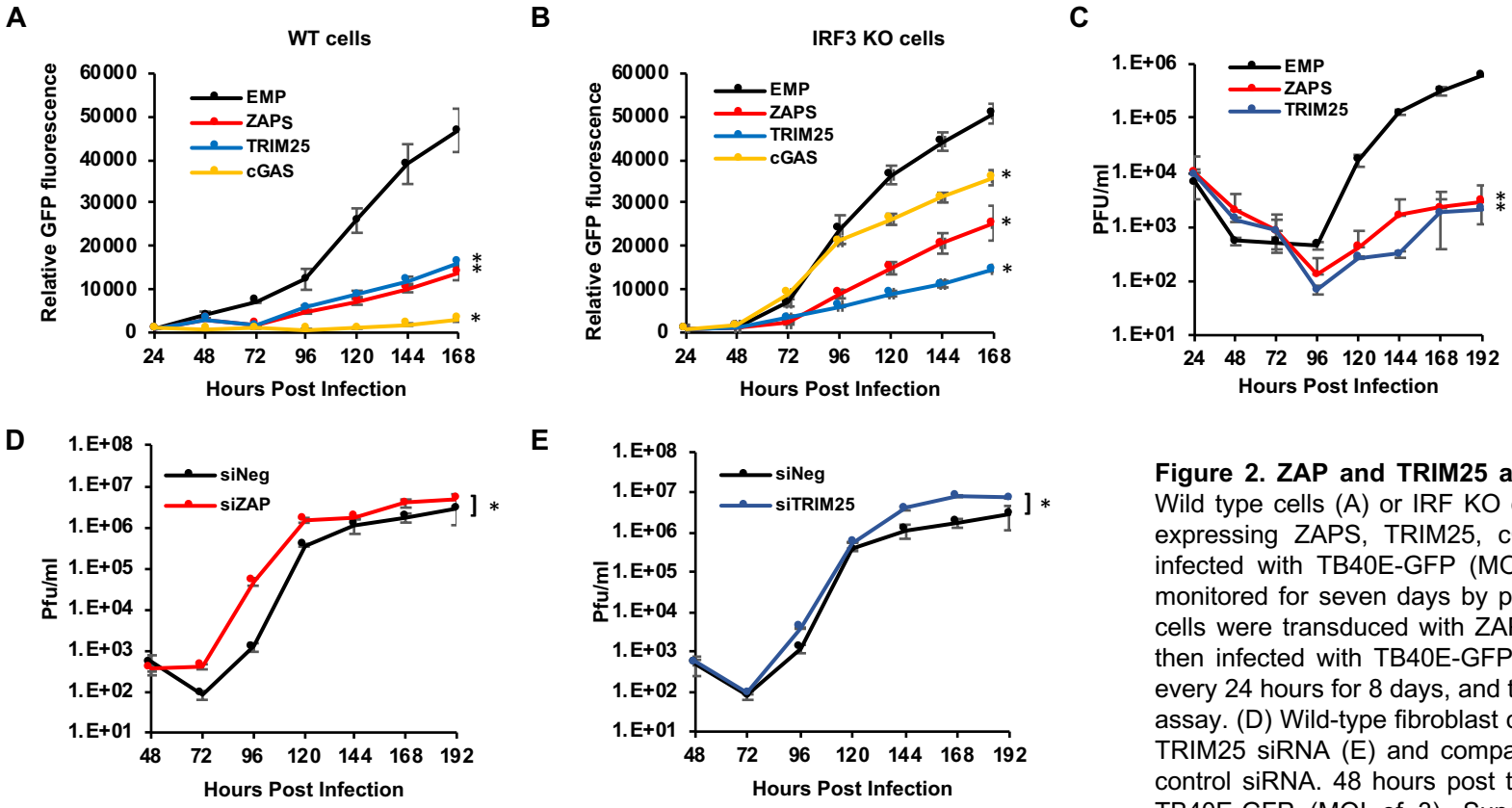

Figure 2. ZAP and TRIM25 are host restriction factors for HCMV. Wild type cells (A) or IRF KO cells (B) were transduced with lentivirus expressing ZAPS, TRIM25, cGAS or an empty control vector, and infected with TB40E-GFP (MOI of 3). GFP fluorescence levels were monitored for seven days by plate fluorometry. (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

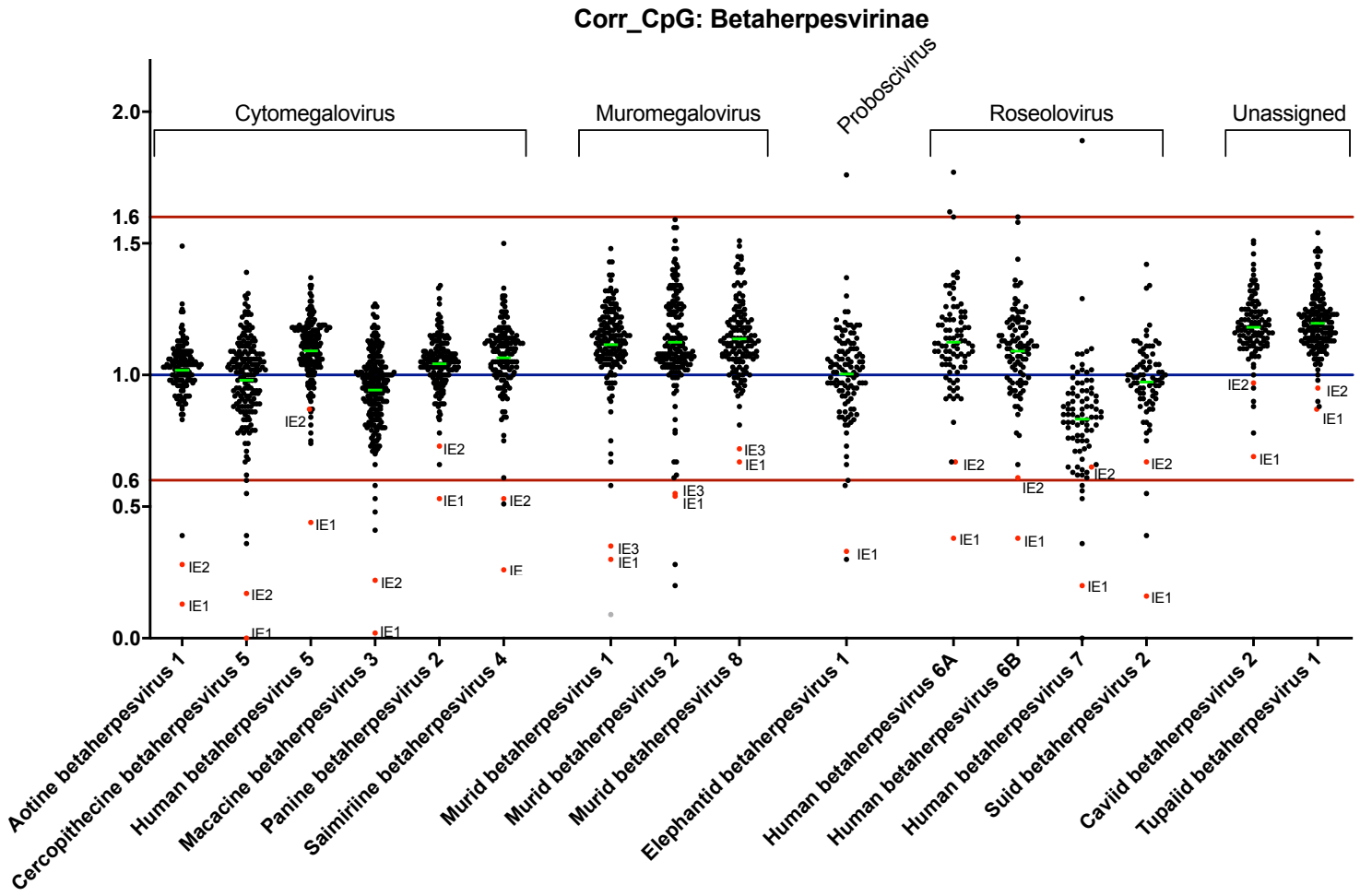

Figure 3. Specific suppression of CpG nucleotides within the immediate early genes of beta-herpesviruses. The CpG content of annotated open reading frames from 16 beta-herpesvirus genomes are shown, following normalization for length and GC content. A corrected CpG ratio of one reflects the expected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

Figure 4

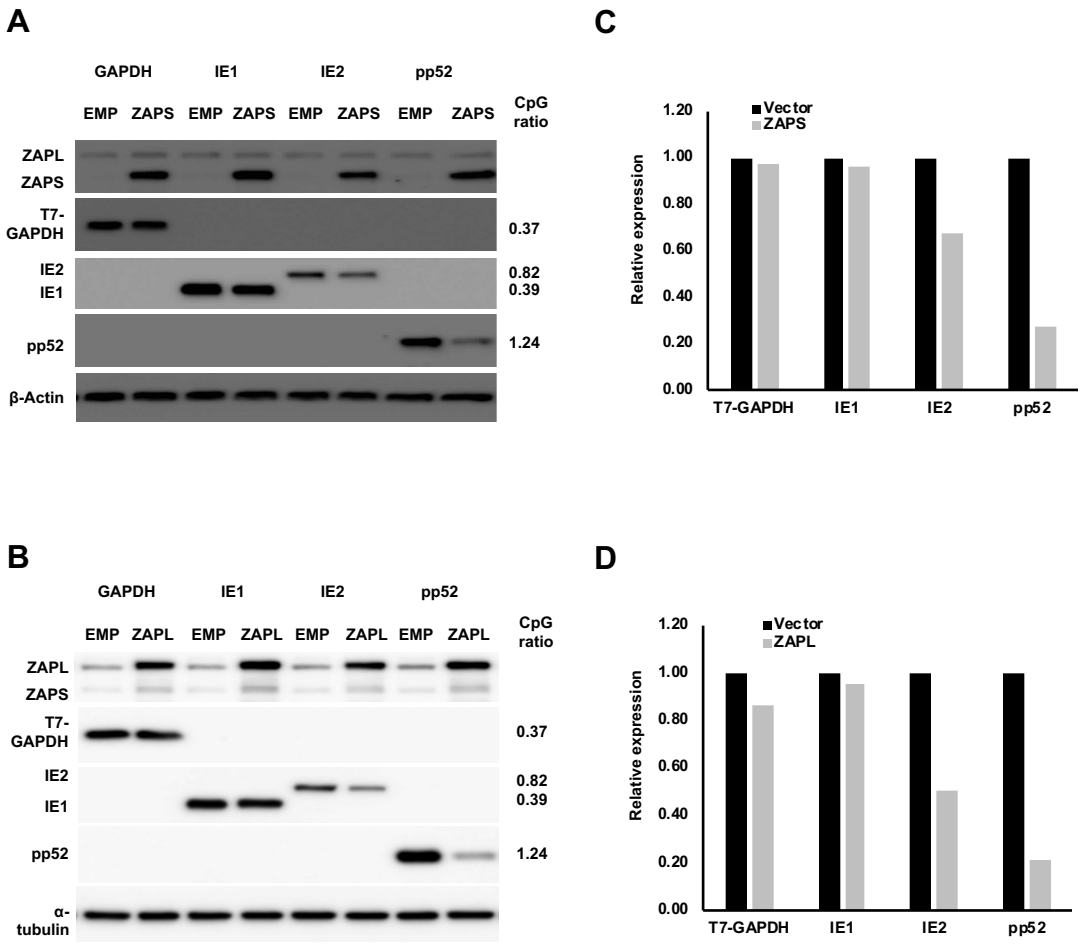

Figure 4. ZAP targets HCMV genes with high CpG dinucleotide frequencies. (A) 293T cells were co-transfected with plasmids expressing ZAPS (A), ZAPL (B) or a control empty vector with plasmids expressing HCMV genes with varied CpG dinucleotide frequencies (IE1, IE2 and pp52). A plasmid expressing T7-tagged 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

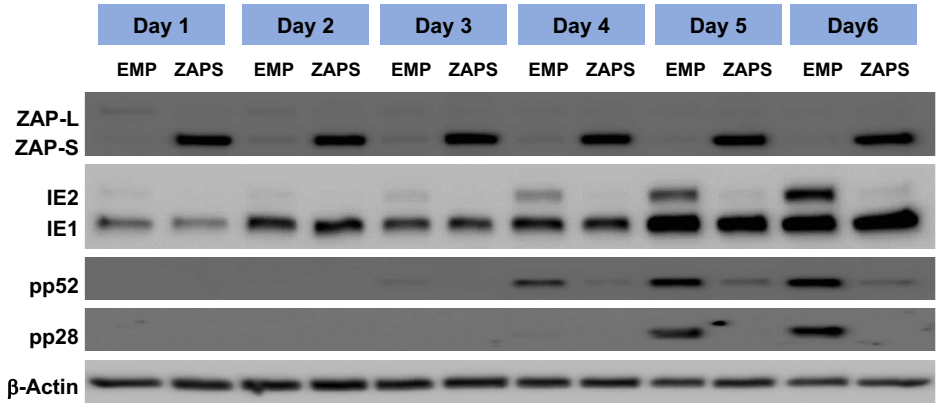

Figure 5. IE1 expression is not affected by ZAP. Wild type fibroblast cells were transduced with ZAPS or empty vector control lentiviruses and then infected with TB40E-GFP 48 hours post transduction. Total protein lysates were harvested every 24 hours and the expression levels of the viral proteins of 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

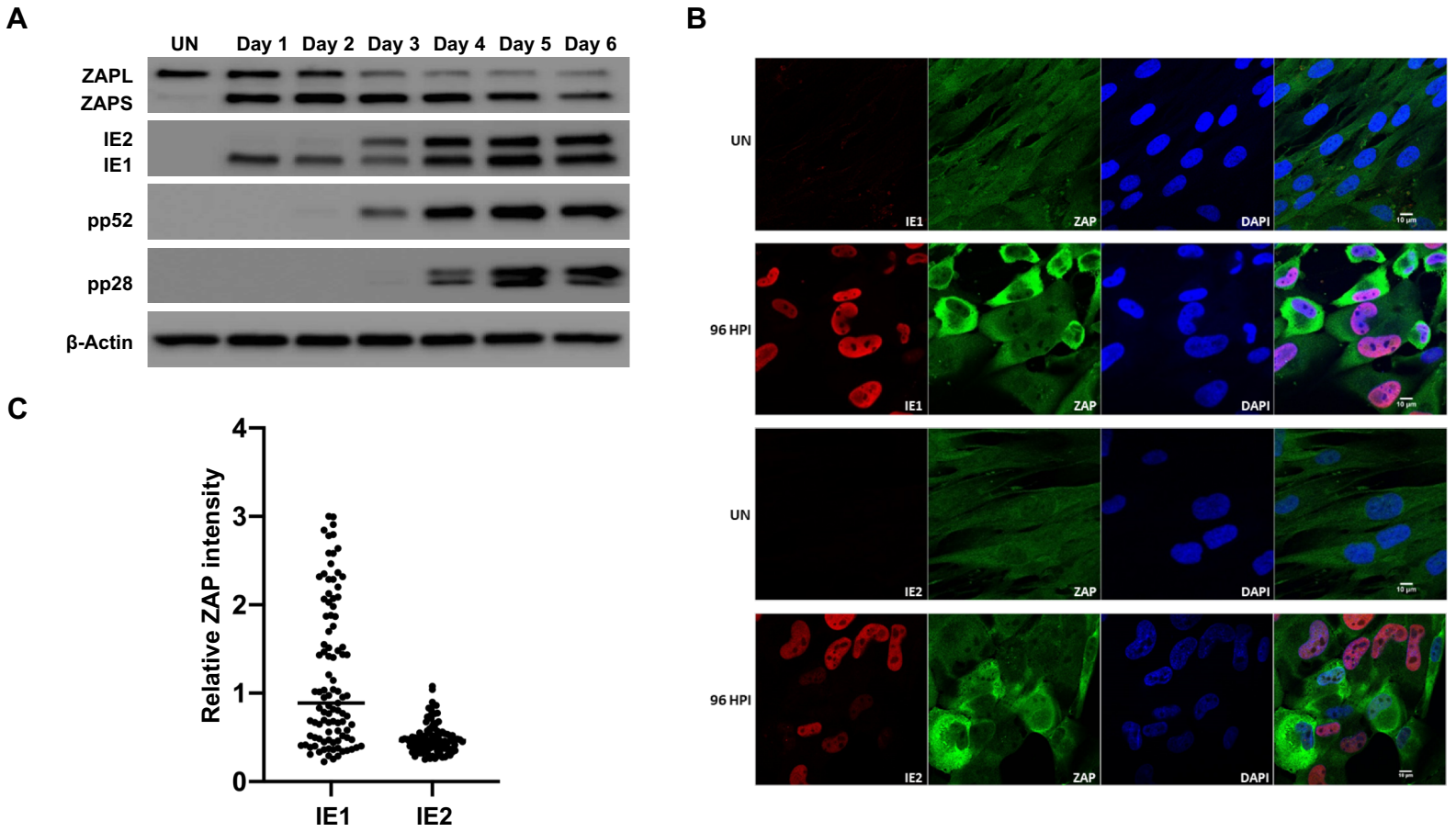

Figure 6. ZAP expression is reduced in IE2 positive cells. (A) Wild type fibroblast cells were infected with TB40E-GFP at an MOI of 3. Total protein lysates were harvested every 24 hours and the expression levels of ZAP and the viral proteins of each major kinetic class of HCMV were determined by western 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

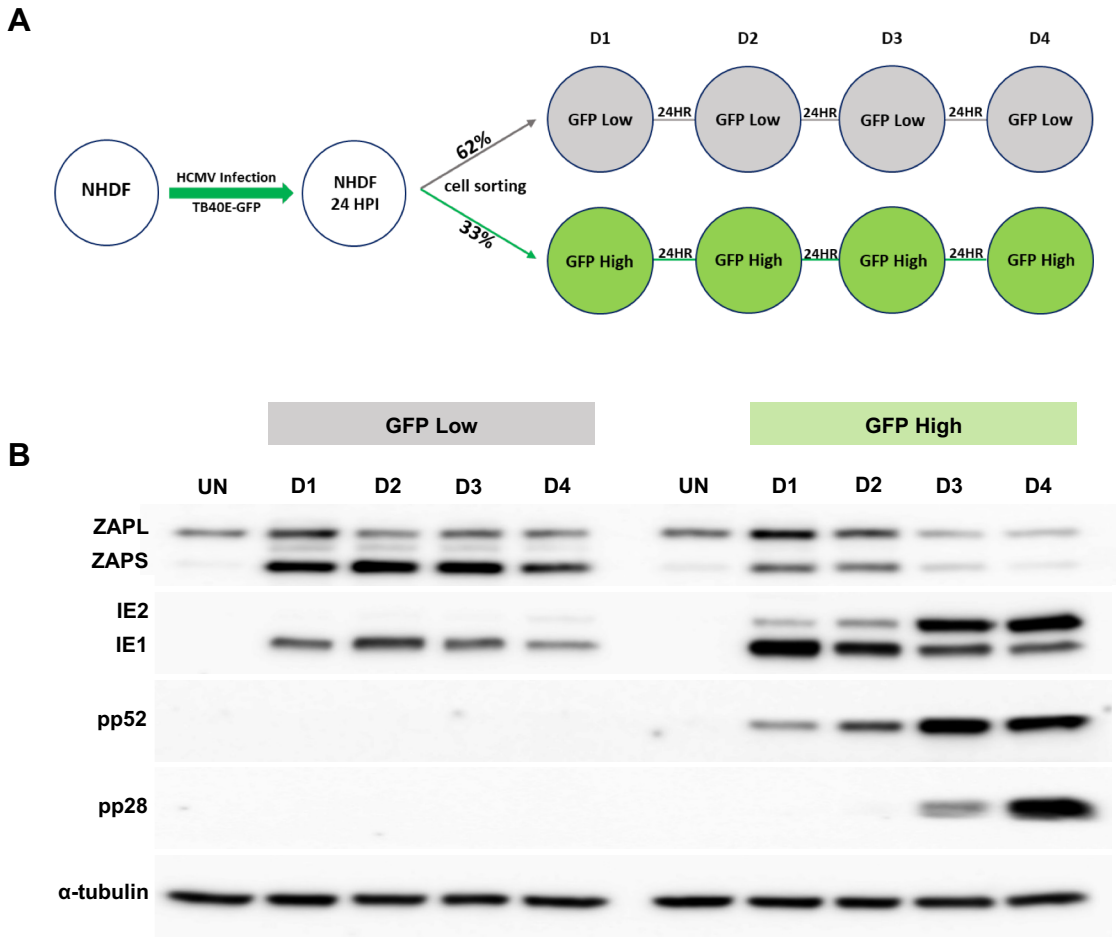

Figure 7. High ZAP expression correlates with failure in HCMV acute replication. (A) Diagram of cell sorting following TB40E-GFP infection. 24 hours post infection, fibroblast cells were sorted into low and high GFP expressing populations, and then re-seeded. Total protein lysates were harvested at the 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

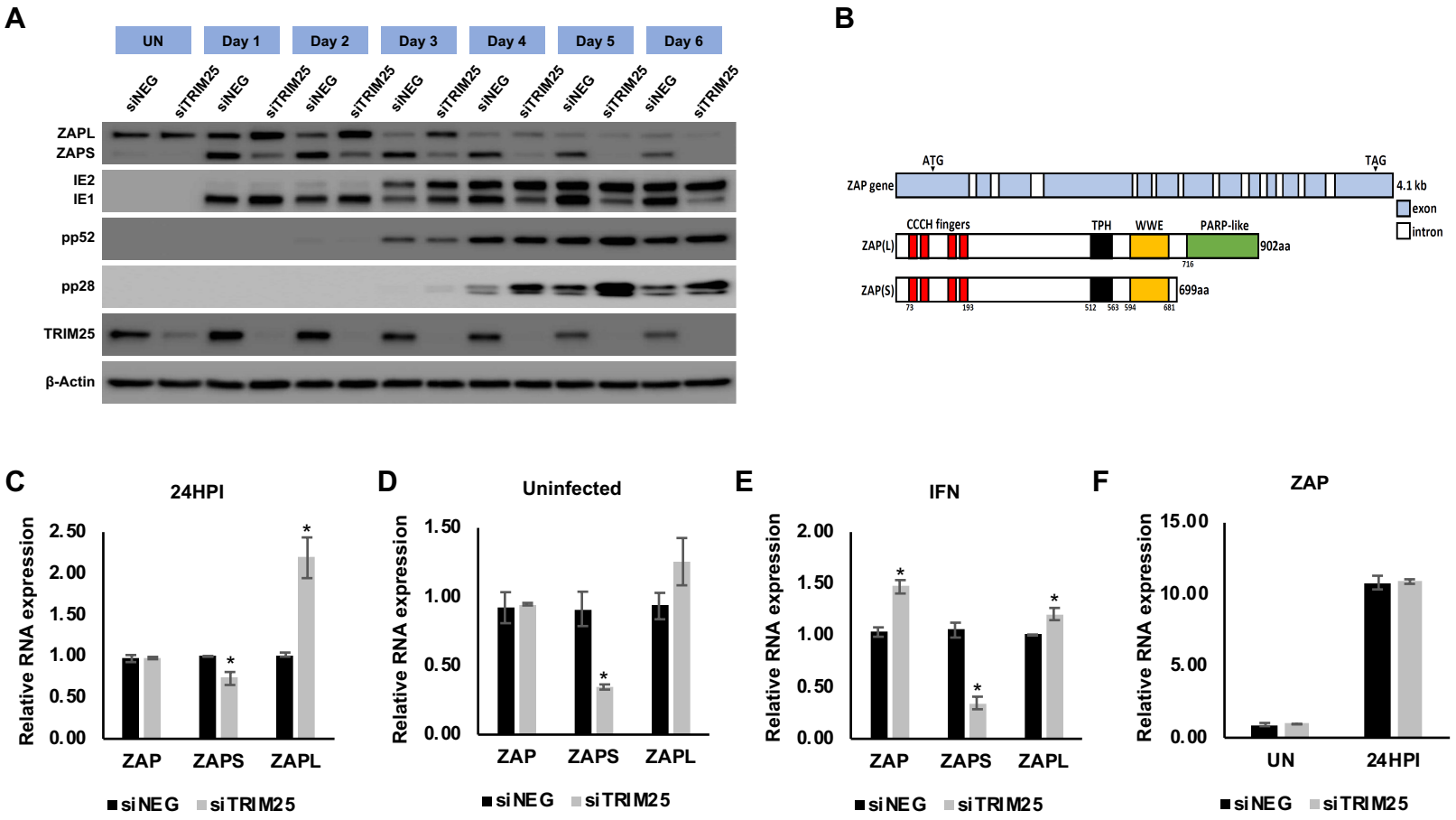

Figure 8. TRIM25 regulates differential splicing of ZAP. (A) Wildtype fibroblast cells were transfected with TRIM25 siRNA or negative control siRNA and infected with TB40/E-GFP 48 hours post transfection (MOI of 3). Total protein lysates were harvested every 24 hours and expression levels of ZAPS and 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.