Rational attenuation of RNA viruses with zinc finger antiviral protein

Daniel Gonçalves-Carneiro, Emily Mastrocola, Xiao Lei, Justin DaSilva, Yoke Fun Chan and Paul D. Bieniasz

Attenuation of a virulent virus is a proven approach for generating vaccines but can be unpredictable. For example, synonymous recoding of viral genomes can attenuate replication but sometimes results in pleiotropic effects that confound rational vaccine design. To enable specific, conditional attenuation of viruses, we examined target RNA features that enable zinc finger antiviral protein (ZAP) function. ZAP recognized CpG dinucleotides and targeted CpG-rich RNAs for depletion, but RNA features such as CpG numbers, spacing and surrounding nucleotide composition that enable specific modulation by ZAP were undefined. Using synonymously mutated HIV-1 genomes, we defined several sequence features that govern ZAP sensitivity and enable stable attenuation. We applied rules derived from experiments with HIV-1 to engineer a mutant enterovirus A71 genome whose attenuation was stable and strictly ZAP-dependent, both in cell culture and in mice. The conditionally attenuated enterovirus A71 mutant elicited neutralizing antibodies that were protective against wild-type enterovirus A71 infection and disease in mice. ZAP sensitivity can thus be readily applied for the rational design of conditionally attenuated viral vaccines.

The zinc finger antiviral protein (ZAP) inhibits the replication of a broad range of RNA and DNA viruses through the recognition of viral CpG-rich RNA and presents opportunities for the design of attenuated viral vaccines. Live-attenuated viral vaccines offer advantages over other vaccine approaches because they express a complete repertoire of viral proteins and so present the widest range of antigenic determinants to induce durable cellular and humoral responses without adjuvants. However, rational methods for the generation of attenuated viruses are few, and most attenuated vaccines have been empirically produced. One approach used for virus attenuation is recoding of nucleic acid sequences by synonymous mutagenesis. Initial reports using this method replaced codons or codon pairs with counterparts found only rarely in the human genome, a process termed ‘deoptimization’. However, deoptimization of a viral RNA genome can have pleiotropic effects on structure, stability and translation efficiency, conferring virus attenuation through multifactorial mechanisms that are not straightforward to predict.

Attenuation by codon-pair deoptimization incidentally increases the frequency of two dinucleotides, CpG and UpA (TpA in DNA). CpG dinucleotides are severely underrepresented in vertebrate genomes, while TpA/UpA dinucleotides are underrepresented in organisms across the tree of life. The underrepresentation of CpG in vertebrate genomes has created an opportunity for non-self RNA recognition that is exploited by ZAP proteins. The ZAP N-terminal domain employs a highly selective binding pocket that can only accommodate a CpG dinucleotide in a single-stranded configuration. While one ZAP molecule binds to one CpG dinucleotide, individual CpG dinucleotides have negligible effects on viral replication. Rather, it is the cumulative effect of multiple CpG dinucleotides that enables ZAP antiviral activity. However, it is unknown how CpG dinucleotide numbers, juxtaposition and underlying sequence context affect ZAP recognition of viral RNA. Moreover, even though CpG dinucleotides confer ZAP sensitivity, introduction of CpG dinucleotides in an unguided manner can have pleiotropic effects on viral replication through ZAP-independent mechanisms.

Viral genome recoding without consideration of the mechanism(s) of attenuation may produce viruses with reduced immunogenicity, which is clearly an undesirable property for any vaccine. Since ZAP–RNA interactions may be immunostimulatory, optimal recoding strategies would maximize ZAP binding and specify ZAP recognition as the attenuating mechanism. However, so far, delineation of sequence features that could be employed to achieve this goal has not been reported.

Using HIV-1 as a model system, we define how CpG dinucleotide number, spacing and surrounding sequence affect ZAP sensitivity. We then apply these parameters to design a mutant picornavirus genome with precise and stable modifications that function as an effective live-attenuated vaccine whose replication is specifically inhibited by ZAP in cell culture and in vivo.

**Results**

CpG dinucleotides and HIV-1 replication. HIV-1 is naturally CpG-poor and largely ZAP-resistant, while mutant derivatives with elevated CpG content are ZAP-sensitive. We first introduced two unique restriction sites (BstEII-ClaI), each of which contains a single CpG dinucleotide, into the HIV-1 env gene (Fig. 1a and Extended Data Fig. 1a). We synonymously recoded the intervening sequence, which lacks any known proximal cis-acting RNA regulatory elements, to contain zero (CG-0) or 43 (CG-43) CpG dinucleotides (Extended Data Fig. 1b). The CG-0 virus replicated indistinguishably from wild-type (WT) HIV-1 in both unmanipulated and in a CRISPR/Cas9-edited human MT4 T-cell line that expressed only exon 1-edited non-functional ZAP proteins. Conversely, the CG-43 virus replicated like WT virus in ZAP-deficient MT4 cells but was defective in ZAP-expressing cells (Fig. 1b and Extended Data Fig. 1c–f). Next, we generated a collection of HIV-1 mutants that differed from each other by a single CpG dinucleotide, with 1 to 23 CpG dinucleotides positioned as close to the BstEII as allowed by}
synonymous substitution in the BstEII-ClaI bounded region (CG-1 to CG-23). While all mutants replicated indistinguishably from WT HIV-1 in ZAP-deficient cells, virus replication was progressively diminished in ZAP-expressing cells as the number of CpG dinucleotides was increased (Fig. 1c,d). Overall, CG-1 to CG-13 replicated well, CG-15 to CG-23 replicated poorly and CG-14 had an intermediate phenotype (Fig. 1c). The percentage of infected cells at 4 d post infection showed an obvious correlation between the number of introduced CpG dinucleotides and the extent of replication (Fig. 1d). Thus, individual CpG dinucleotides had an incremental impact, and approximately 15 CpG dinucleotides were required to profoundly inhibit HIV-1 replication.

Spacing and base composition between CpG dinucleotides and ZAP activity. We generated a second collection of HIV-1 mutants that each contained 15 CpG dinucleotides but differed in the spacing between each CpG dinucleotide. In ZAP-expressing cells, viruses that contained 15 CpG dinucleotides separated by a mean of 6 or 11 nucleotides replicated with near-WT kinetics. Conversely, viruses with CpG dinucleotides separated by 14 or 32 nucleotides were defective (Fig. 2a,b), while the effect of the 15 CpG dinucleotides was diminished if the spacing between them was further increased to a mean of 40 nucleotides.

To examine ZAP binding to these mutant viral sequences, we performed crosslinking immunoprecipitation assays coupled with RNA sequencing (CLIP-seq) using viruses with 15 CpG dinucleotides, separated by a mean of 11 or 32 nucleotides. We measured the frequency of CLIP-seq reads mapping to each nucleotide position in the BstEII-ClaI interval of the HIV-1 genome. While viral RNA containing no CpG dinucleotides in the BstEII-ClaI interval showed little ZAP binding, there was abundant ZAP binding when the CpG dinucleotides were positioned at a mean of 32 nucleotides apart (Fig. 2c).
Fig. 2 | Spacing and composition between CpG dinucleotides affect HIV-1 replication. a, Viral replication in cells infected with mutants of HIV-1 GFP reporter viruses containing 15 additional CpG dinucleotides with a mean of 6, 11, 14, 32 or 40 nucleotides between each CpG dinucleotide. Each day, the percentage of infected cells was measured by flow cytometry. b, Summary of the percentage of GFP-positive cells at day 4 post initial infection. Mean ± s.d. from 3 independent experiments; two-way ANOVA for the presence of ZAP (column factor) P < 0.0001, spacing (row factor) P = 0.002. Šídák’s multiple comparisons test was used to calculate adjusted P values between ZAP+/+ and ZAP−/− groups, comparisons of virus mutants with spacing between CpGs ≥14 nt display P(adj) < 0.0001. c, CLIP-seq analysis of ZAP binding to HIV-1 RNA: ZAP-deficient and TRIM25-deficient 293T cells were transfected with a plasmid encoding ZAP-L as well as HIV-1 proviral plasmids with 0 CpG dinucleotides (black line), or 15 CpG dinucleotides with average spacing of 11 nucleotides (red line) or 32 nucleotides (blue line) in the BstEII–ClaI interval. CLIP reads that mapped to the BstEII–ClaI interval are plotted as the normalized fraction of total reads. Circles indicate the positions of CpG dinucleotides. d,e, Replication of CG-14/A+/U+/C+/G+ mutant viruses in ZAP-expressing and ZAP-deficient MT4 cells. The percentage of GFP-positive cells was measured daily (d) and the mean ± s.d. of 3 independent experiments were measured at day 4 post infection (e).

However, ZAP binding to the modified sequence was minimal when CpG dinucleotides were positioned at a mean of 11 nucleotides apart (Fig. 2c). Thus, these data suggest that adequate spacing between CpG dinucleotides is important for ZAP recognition.

Based on previous CLIP-seq experiments13,15 and ZAP–RNA crystal structures14,15, ZAP specificity is determined solely by the target CpG dinucleotide and not flanking nucleotides. However, whether the overall sequence context in which CpG dinucleotides are present contributes to ZAP antiviral activity is unknown. We generated HIV-1 mutants that contained either 0 or 14 CpG dinucleotides (CG-0 and CG-14) and synonymously mutated the surrounding sequence in the BstEII–ClaI interval to contain the maximum possible number of adenine (A+), cytidine (C+), guanine (G+) or uridine (U+) nucleotides (Extended Data Fig. 2a). The CG-0 viruses with elevated A, U or G content (CG-0/A+, CG-0/U+ and CG-0/G+) replicated with close to wildtype kinetics, while the cytidine-enriched virus (CG-0/C+) showed severe replication defects, independent of the presence of ZAP (Extended Data Fig. 2b,c). The CG-14/A+, CG-14/U+ and CG-14/G+ viruses replicated similarly to wild type in ZAP-deficient cells while the CG-14/C+ virus exhibited a ZAP-independent defect similar to the CG-0/C+ virus (Fig. 2d). Notably, while elevating G content (CG-14/G+) had little impact on virus replication, the CG-14/A+ and CG-14/U+ viruses were severely attenuated, specifically in ZAP-expressing
cells (Fig. 2d,e). Thus, increasing A or U content apparently increased the ability of CpG dinucleotides to impart ZAP sensitivity.

We next generated 7 HIV-1 mutants, each containing 15 additional CpG dinucleotides, with the CpG-enriched sequences positioned at different locations across the env gene (Extended Data Fig. 3a). All these viruses replicated similarly to WT HIV-1 in ZAP-deficient cells, and 5/7 exhibited ZAP-dependent attenuation. The exceptions were two viruses with CpG-enriched segments located 3’ to env nucleotide positions 110 or 889 (CG-15(110) and CG-15(889), Extended Data Fig. 3b,c). Notably, the A and U frequencies in these two regions were reduced compared with other HIV-1 genome regions, and the mean spacing between CpG dinucleotides was the lowest among the mutants (Extended Data Fig. 3d). Strikingly, increasing the adenine frequency of CG-15(889) in the CpG-enriched interval to generate CG-15(889)/A increased ZAP-dependent attenuation such that CG-15(889)/A+ was specifically defective in ZAP-expressing cells (Extended Data Fig. 3d,e). We conclude that apparent position-dependent effects on ZAP sensitivity are probably mediated by surrounding nucleotide composition.

Selective pressure by ZAP can deplete CpG dinucleotides from viral genomes. The paucity of CpG dinucleotides in mammalian virus genomes may have been driven by ZAP selection10,11. The potential utility of reversing this property to generate ZAP-sensitive, live-attenuated vaccines depends on the stability of attenuating mutations. While codon-pair deoptimization has been reported to be stable during virus passage20,21, the stability of introduced CpG mutations is enterovirus A71 (EV-A71) that causes hand, foot and mouth disease in young children, with occasional severe complications including acute flaccid paralysis, brainstem encephalitis and meningitis23. EV-A71 has a low frequency of CpG dinucleotides (Fig. 4a) and is therefore a good candidate for genetic recoding to confer ZAP sensitivity and generate a potential live-attenuated vaccine.

4a) and is therefore a good candidate for genetic recoding to confer ZAP sensitivity and generate a potential live-attenuated vaccine.

Engineering enterovirus A71 to confer ZAP sensitivity. Members of the Picornaviridae are important human pathogens that cause morbidity27 and most lack efficacious vaccines. One such picornavirus is enterovirus A71 (EV-A71) that causes hand, foot and mouth disease in young children, with occasional severe complications including acute flaccid paralysis, brainstem encephalitis and meningitis21. EV-A71 has a low frequency of CpG dinucleotides (Fig. 4a) and is therefore a good candidate for genetic recoding to confer ZAP sensitivity and generate a potential live-attenuated vaccine.

The EV-A71 genome encodes a single polyprotein (Fig. 4b) that is cleaved to generate structural (VP1, VP2, VP3 and VP4) and non-structural proteins (2A, 2B, 2C, 3A, 3C, VPg and the RNA-dependent RNA polymerase). To monitor EV-A71 replication, we generated a reporter virus encoding NanoLuc luciferase followed by a 2A cleavage site at the N terminus of the viral polyprotein, as previously described44. For recoding, we applied the CpG number, spacing and intervening mononucleotide content criteria determined above using HIV-1. We recoded a ~1 kb target region that spanned 2C, 3A, VPg and 3C coding sequences and was selected arbitrarily, other than the fact that it lacks any known proximal cis-acting RNA regulatory elements (Fig. 4b). In WT EV-A71, this region contains 32 CpG dinucleotides and 261 adenines (Fig. 4c). We changed the number and distribution of CpG dinucleotides, as well as the adenine content, generating three mutants: (1) EV-A71/A+ that contained a high frequency of A nucleotides without increasing the number of CpG dinucleotides; (2) EV-A71/C+G-48 with 16 additional CpG dinucleotides that combined with the existing 32 CpG dinucleotides, generating a segment with 48 CpG dinucleotides at a mean of 19 nucleotides apart but retaining WT mononucleotide composition; and (3) EV-A71/C+G-48/A+ that had the 16 additional CpG dinucleotides positioned as above but in the A-rich sequence context (Fig. 4c). All viruses replicated well in ZAP-deficient cells (Fig. 4d–f). While no replication defects were observed in the WT EV-71, EV-A71/A+ or EV-A71/C+G-48 mutants, the EV-A71/C+G-48/A+ mutant was specifically attenuated in ZAP-expressing cells (Fig. 4d–f).

When the EV-A71/C+G-48/A+ mutant was repeatedly passaged in ZAP-expressing cells, luciferase activity progressively decreased with each passage and was ultimately below the detection limit (Fig. 4g). Conversely, no replication defect was evident in ZAP-deficient cells, and sequence analyses revealed no reversion mutations acquired during passage. While viral RNA was abundant in all EV-A71/C+G-48/A+ replicates in ZAP-deficient cells, RNA levels were below the detection limit in ZAP/+ cells (Extended Data Fig. 5a,b). Moreover, when we followed virus replication for 4 days using NanoLuc assays, determination of infectious virus yield (TCID50) and viral RNA quantification, all three measurements of EV-A71/C+G-48/A+ mutant replication were dramatically reduced in ZAP/+ cells compared with ZAP-deficient cells (Extended Data Fig. 5c–e). Thus, the observed replication deficits reflected bona fide effects of ZAP and not reporter gene instability. Moreover, the CpG and A enrichment was stable and EV-A71/C+G-48/A+ could not escape ZAP under these conditions. Thus, the principles governing ZAP sensitivity identified using HIV-1 could be applied to an unrelated RNA virus, leading to stable, ZAP-dependent attenuation.

ZAP-dependent attenuation of recoded EV-A71 in mice. To determine whether ZAP could inhibit EV-A71/C+G-48/A+ replication in vivo, we generated a ZAP−/− C57BL/6 mouse line using CRISPR guide RNAs targeting ZC3HAV1 exon 1. A germline-transmissible edited ZAP allele contained a 2 nt insertion, introducing a frame-shift mutation that abrogated ZAP expression (Extended Data Fig. 6a,b). We introduced substitutions in EV-A71, these substitutions being required for symptomatic infection of mice29, thus generating ‘mouse-adapted’ viruses, hereafter referred to as mEV-A71 and mEV-A71/C+G-48/A+. These viruses lacked the NanoLuc reporter, but EV-A71/C+G-48/A+ exhibited the same ZAP-dependent replication deficits as the aforementioned NanoLuc encoding EV-71 reporter constructs (Extended Data Fig. 6c). Because we observed more consistent mEV-A71 pathogenesis in IFNAR−/− neonatal mice, we crossed the ZAP−/− C57BL/6 mouse line to an IFNAR−/− C57BL/6 mouse line. We then infected ZAP−/−/IFNAR−/− or ZAP−/−/IFNAR−/− neonates with mEV-A71 or mEV-A71/C+G-48/A+ and scored disease progression according to a previously described 0–4 scale; from asymptomatic (0) to dead or moribund (4) (Fig. 5a). Limb paralysis events (Fig. 5b) characteristic of EV-A71 infection in similar mouse models26,27 were observed upon mEV-A71 infection of either ZAP−/− or ZAP−/− mice, with disease progression that was ultimately fatal in most cases (Fig. 5a–c). Conversely, while ZAP−/− mice succumbed to the mEV-A71/C+G-48/A+ virus,
nearly all ZAP+/+ mice survived infection and presented with low clinical scores (Fig. 5a–c). Specifically, 20/26 mEV-A71/CG-48/A+ infected ZAP+/+ mice showed no symptoms, while 4 mice exhibited enterovirus-specific symptoms (that is, limb paralysis) that quickly resolved (Extended Data Fig. 7a). Two ZAP+/+ mice from a single litter died suddenly without limb paralysis, suggesting that their death was not due to enteroviral disease. In contrast, all mEV-A71/CG-48/A− infected ZAP−/− mice developed enterovirus-specific symptoms before death (Extended Data Fig. 7b). Viral RNA levels in muscle were equivalent in mEV-A71-infected ZAP+/+ and ZAP−/− mice but 60-fold lower in mEV-A71/CG-48/A+ infected ZAP+/+ mice compared with ZAP−/− mice (Fig. 5d). Sequences of PCR amplicons encompassing the engineered regions of the mEV-A71/CG-48/A+ genome from all mice in which viral RNA was detectable (Fig. 5d) revealed no mutations. Together, these data indicate that mEV-A71/CG-48/A+ was strongly and stably attenuated in vivo, and its attenuation was strictly dependent on ZAP.

**Recoded EV-A71 elicits protective immunity in mice.** We collected plasma from ZAP+/+ mice 5 weeks after previous inoculation with mEV-A71/CG-48/A+ at 1 d or 5 d of age (Fig. 5e). While plasma from mock-inoculated mice did not neutralize EV-A71, plasma from mice inoculated with mEV-A71/CG-48/A+ neutralized EV-A71 infection, with 50% neutralizing titres (NT50) ranging from 895 to 10,602 in mice infected at 1 d of age (median NT50 = 2,957) and titres ranging from 627 to 1,514 in mice infected...
at 5d of age (median NT₅₀ = 1,162) (Fig. 5f and Extended Data Fig. 8a). Next, we aimed to determine whether these neutralizing antibodies were protective in vivo. Since productive infection of mEV-A71 is age-sensitive in mice, we performed passive protection experiments in which the neonatal offspring of females that were previously mock-inoculated or inoculated with mEV-A71/CG-48/A+ were challenged with WT mEV-A71 (Fig. 5e). In this type of experiment, suckling pups acquire antibodies via maternal milk. Pups from females previously inoculated with mEV-A71/CG-48/A+ at 1 d or 5 d of age (median clinical score at day 20 of 0.52 and 0.24 in 1-day-old and 5-day-old infected mice, respectively) and increased survival compared with pups from mock-treated females (median clinical score at day 20 of 3.14 and 2.86 in 1-day-old and 5-day-old infected mice, respectively) (Fig. 5g and Extended Data Fig. 8b). We conclude that ZAP-attenuated mEV-A71/CG-48/A+ replication in mice elicits antibodies that are passively transferred and protective against mEV-A71 disease in the offspring of inoculated females.

Discussion

The delineation of sequence features that affect sensitivity to ZAP in HIV-1 (numbers of CpG, spacing and context) enabled us to develop design rules that we applied to engineer a picornavirus mutant that is strongly attenuated in a strictly conditional manner. The close spacing between each CpG dinucleotide affecting ZAP sensitivity and ZAP binding in CLIP-seq suggests that ZAP molecules binding to adjacent CpG dinucleotides may compete with each other, consistent with modelling studies indicating that ZAP competes for RNA secondary structure with ~13 nucleotides. Closely spaced CpG dinucleotides may also promote RNA secondary structure that might inhibit ZAP access. Conversely, wide CpG spacing conferred reduced sensitivity. Interactions between ZAP and TRIM25 and between ZAP and KHNYN, two known co-factors of ZAP, appear to be mediated by protein-protein contacts; thus, it is possible that ZAP molecules and co-factors bound to adjacent CpG dinucleotides may coalesce, with close spacing facilitating assembly of active ZAP:TRIM25:KHNYN complexes.
Fig. 5 | Recoded EV-A71 has ZAP-dependent attenuation and elicits protective antibodies in mice.  

**a**, Clinical score following infection of 1-day-old ZAP\(^+/+\) and ZAP\(^{-/-}\) mice with mEV-A71 WT or mEV-A71/CG-48/A+ (mean ± s.d.; \(n = 11–26\) mice per group; \(P\) values calculated using two-way ANOVA; NS, non-significant). **b**, Examples of characteristic pathology developed in infected mice following mEV-A71 infection, including one-limb and two-limb paralysis. **c**, Probability of ZAP\(^+/+\) and ZAP\(^{-/-}\) mouse survival (%) following infection with mEV-A71 WT or mEV-A71/CG-48/A+ (\(n = 11–36\) mice per group; \(P\) value calculated using Mantel-Cox test). **d**, One-day-old mice (\(n = 4–6\) per group) were infected with indicated virus; 6-days-post-infection mice were sacrificed and muscles from both posterior limbs were collected, homogenized and total RNA was extracted. EV-A71-specific RNA was quantified by qPCR. Dashed line indicates limit of qPCR detection. **e**, Schematic representation of experimental design. ZAP\(^+/+\)/IFNAR\(^{-/-}\) female mice previously inoculated with mEV-A71/CG-48/A+ (or mock-infected females) were mated with ZAP\(^+/+\)/IFNAR\(^{-/-}\) male mice. Resulting offspring was challenged at 1 d of age with mEV-A71 WT. **f**, Neutralizing activity in plasma from mice after mEV-A71/cG-48/A+ infection (infection at 1 d of age), blood collection at 5 weeks, or mock-infected mice evaluated using EV-A71 NanoLuc luciferase reporter virus. 293T cells were incubated with the antibody:virus mixture for 48 h and luciferase activity was measured. **g**, Clinical score and survival probability following mEV-A71 (WT) infection of the offspring of ZAP\(^+/+\)/IFNAR\(^{-/-}\) females previously inoculated with mEV-A71/CG-48/A+ (\(n = 25\) pups from 4 different females) or previously mock-infected (\(n = 22\) pups from 4 different females) at 1 d of age. Clinical score and survival were assessed daily until weaning. Statistical significance was inferred by two-way ANOVA and Mantel-Cox tests.
Studies of RNA-binding protein specificity typically focus on the recognition of particular RNA sequences. Although many RNA-binding proteins recognize specific sequence motifs, contextual features, such as flanking nucleotide composition, can be crucial for determining target specificity. Increasing adenine or uridine content in viral genomes may increase sensitivity to RNAes, such as RNAse L that cleaves UpA and UpU dinucleotides in viral RNA, and we note that our A- or U-enriched viral genomes contain greater numbers of UpA dinucleotides. Nevertheless, A/U-enriched viruses replicated identically to wildtype HIV-1 in ZAP-deficient cells, indicating that the defects imposed by A/U enrichment are ZAP specific. An obvious effect of A or U enrichment would be to reduce stable RNA secondary structure, perhaps increasing CpG dinucleotide accessibility to ZAP.

All approaches to viral attenuation must balance reduced pathogenesis versus reduced antigen levels that accompany impaired viral genome expression and replication. In principle, programmed attenuation of viruses based on ZAP sensitivity might be adjustable through variation in CpG number and accessibility. Because RNA features conferring ZAP sensitivity in HIV-1 were readily transferrable to a very different virus (EV-A71), these approaches may be generally applicable, and it is possible that nearly any virus that exhibits CpG depletion could be a suitable target for this recoding approach. Importantly, no loss-of-function mutations in ZAP have been identified so far in humans, but further investigation would be required to evaluate whether existing genetic variation in ZAP might render some individuals more susceptible to CpG-enriched viruses and whether the stability of the introduced mutations that we observed for HIV-1 and EV-A71 in cell culture and in mice would be generalizable to CpG-enriched virus-vaccinated humans. Attenuation through engineered ZAP sensitivity might be ineffective for viruses with alternative mechanisms of ZAP evasion, such as targeting ZAP or co-factors for depletion, although no such viruses are currently known. Notably, strictly ZAP-dependent attenuation allows for the cultivation of high-titre, live-vaccine stocks in ZAP-deficient cells. Additional advantages may stem from the reported observation that RNA recognition by ZAP is immunostimulatory. Indeed, infection with alternative mechanisms of ZAP evasion, such as targeting ZAP through engineered ZAP sensitivity might be ineffective for viruses with modified mononucleotide composition were designed by maintaining the position of the 14 CpG dinucleotides in the CG-14 virus and replacing all other codons with those that contained the desired nucleotide. For example, in A+ mutants, leucine codons (CUU) were replaced with adenine-containing leucine codons (CUA). In all cases, no additional CpG dinucleotides were introduced. Viruses were designed with in-house built scripts and checked for inadvertent introduction of splice sites using MaxEntScan. Synthetic DNA sequences encoding the modified sequences were purchased (Twist Bioscience) inserted into the HIV-1 NHG BstEII and ClaI modified proviral plasmid using standard cloning procedures.

Viruses. HIV-1 wildtype and mutants viruses were generated as described previously. Briefly, viral plasmids were linearized with MluI restriction enzyme and column purified. Linearized DNA was then used to generate viral RNA using the T7 RiboMAX Express large-scale RNA production system according to the manufacturer’s guidelines. Viral RNA was then transcribed in ZAP-deficient RD cells using the TransIT-mRNA transfection kit. After overnight incubation, media were replaced for cytopathic effect. When the cytopathic effect was observed in ~80% of cells, supernatants were collected, clarified by centrifugation (10 min, 2,000 × g) and filtered through a 0.22 µm filter. Collected virions were concentrated using Lenti-X concentrator (Clontech) according to the manufacturer’s guidelines and resuspended in serum-free DMEM.

EV-A71 wildtype and mutant viruses were generated as described previously. Briefly, viral plasmids were linearized with MluI restriction enzyme and column purified. Linearized DNA was then used to generate viral RNA using the T7 RiboMAX Express large-scale RNA production system according to the manufacturer’s guidelines. Viral RNA was then transcribed in ZAP-deficient RD cells using the TransIT-mRNA transfection kit. After overnight incubation, media were replaced for cytopathic effect. When the cytopathic effect was observed in ~80% of cells, supernatants were collected and filtered through a 0.1 µm filter. Virus stocks were passaged once in ZAP-deficient RD cells. All virus stocks were stored at −80°C before use.

HIV-1 replication assays. For spreading infections, 1.5 × 10^4 MT4 cells were infected with 400 infectious units of YSV-G pseudotyped HIV-1 NHG in a total of 2 ml of complete RPMI. Each day after infection, cells were resuspended, 100 µl of cell suspension was collected and fixed in 4% paraformaldehyde, and cultures were replenished with 100 µl of fresh RPMI. The percentage of GFP-positive cells was determined using flow cytometry and calculated using Flowjo. For the long-term virus passage experiments, 7.5 × 10^4 MT4 cells were infected with 2,000 infectious units of the CG-43 or CG-15 HIV-1 NHG mutants. Every 2 d, cells were resuspended, 100 µl of the cell suspension was fixed in 4% paraformaldehyde and the percentage of GFP-positive cells was measured using a flow cytometer. When the percentage of GFP-positive cells was greater than 85%, supernatants were collected, filtered through a 0.22 µm filter and used as inoculum to infect fresh MT4 cells. The RNA was isolated from an aliquot of the passaged supernatant using TRIzol, reverse transcribed using the SuperScript III first-strand synthesis system, and an env fragment was amplified using the following primers: 5′-ACAGAAAAATTTGGCTGACCCGCTTTATTGGG-3′ and 5′-GGTTTGTTATCTAGTGGTATATCGAAG-3′. Mutations identified in revertant viruses were then introduced into the HIV-1 NHG CG-15 construct by site-directed mutagenesis.
EVA-71 replication assays. Hela cells were infected at a multiplicity of infection (MOI) of 0.02 for 1 h at 37 °C. Cells were then washed twice in PBS and incubated in complete DMEM at 37 °C. At the indicated timepoints, 100 µl of the supernatant was collected by centrifugation (10,000 x g, 5 min) and filtered through a 0.45 µm pore size filter. Collected virus was diluted and used to re-infect cells at MOI = 0.02. Median tissue culture infectious doses (TCID₅₀) were determined using a cytopathic effect readout and calculated as previously described²⁸ using RD ZAP⁺ target cells.

Immunoblotting. Cells were lysed in NuPAGE LDS sample buffer (Invitrogen) supplemented with β-mercaptoethanol. Samples were then heated at 72 °C for 20 min and sonicated for 15 s. Protein samples were resolved onto NuPAGE 4–12%, Bis-Tris protein gels (Invitrogen), transferred to nitrocellulose membranes and blocked with Intercept blocking buffer (Li-Cor) and incubated with the following antibodies: anti-ZC3H1AV1 (rabbit polyclonal antibody, 16820-1-AP, Proteintech) was used at 1:5,000 dilution in PBS supplemented with Tween20 in human MT4 cell line samples, anti-ZC3H1AV1 (rabbit polyclonal antibody, ab124715, Abrex) was used at 1:300 dilution in 5% milk in PBS-Tween20 in mouse peripheral blood mononuclear cell samples and anti-α-tubulin (mouse monoclonal antibody, T5168, Sigma). After overnight incubation at 4 °C, membranes were washed in PBS-Tween20, blotted with horseradish peroxidase-conjugated secondary antibodies and developed using a C-Digiti chemiluminescent western blot scanner.

CLIP-seq. All CLIP-seq experiments were performed as described previously³⁸ with the following modifications. HEK293T ZAP⁺/− TRIM25⁺/− cells were transfected with a proviral plasmid and a plasmid encoding the long isoform of Cas9 complexed with a guide RNA targeting exon 1 of mouse ZAP. Two DNA lesions were identified, ZAP-SA-R1' and ZAP-SA-R1''. The results indicated that after 4 days post infection, plasma was heat-inactivated as described previously⁷, serially diluted and incubated with EV-A71 NanoLac reporter virus for 1 h at 37 °C. Viruses were then used to infect HEK293T cells and incubated for a further 48 h at 37 °C. Infected cells were then lysed with passive lysis buffer (Promega) and nano luciferase activity was measured as described above. NT₅₀ titres were calculated in GraphPad Prism using non-linear regression (least squares regression without weighting).

Statistical analyses. All statistical analyses were performed using GraphPad Prism 9. Spreading infection data were plotted as mean ± s.d. For comparison between virus mutants and presence of ZAP, two-way analysis of variance (ANOVA) was used with Sidák's multiple comparisons test. Statistical significance in the clinical scores of infected mice was performed two-way ANOVA, while differences in probability of survival were assessed using the Mantel-Cox test. Details of statistical tests, sample sizes as well as P values are described for each experiment in the respective figure legends.

Received: 21 January 2022; Accepted: 3 August 2022; Published online: 8 September 2022

References

1. Gao, G., Gao, X. & Goff, S. P. Inhibition of retroviral RNA production by ZAP; a CCCH-type zinc finger protein. Science 297, 1703–1706 (2002).
2. Müller, S. et al. Inhibition of filovirus replication by the zinc finger antiviral protein. J. Virol. 81, 2391–2400 (2007).
3. Lin, Y.-T. et al. Human cytomegalovirus evades ZAP detection by suppressing CpG dinucleotides in the major immediate early 1 gene. PLoS Pathog. 16, e1008844 (2020).
4. Takata, M. A. et al. CG dinucleotide suppression enables antiviral defence targeting non-self RNA. Nature 550, 124–127 (2017).
5. Pollard, A. J. & Bijker, E. M. A guide to vaccinology: from basic principles to high-fidelity polymerase as next-generation vaccine candidate. J. Virol. 93, e02308–e02318 (2019).
6. Le Nouën, C. et al. Attenuation of human respiratory syncytial virus by genome-scale codon-pair deoptimization. Proc. Natl Acad. Sci. USA 111, 13169–13174 (2014).
7. Tsai, Y.-H. et al. Enterovirus A71 containing codon-deoptimized VP1 and high-fidelity polymerase as next-generation vaccine candidate. J. Virol. 93, e02328–e02338 (2019).
8. Li, P. et al. Zika virus attenuation by codon pair deoptimization induces sterilizing immunity in mouse models. J. Virol. 92, e00701–e00718 (2018).
9. Mueller, S. et al. Live attenuated influenza virus vaccines by computer-aided rational design. Nat. Biotechnol. 28, 723–726 (2010).
10. Schick, K., Trzcinski, S. & Wellin, D. Attenuation of a very virulent Marek's disease herpesvirus (MDV) by codon pair bias deoptimization. PLoS Pathog. 14, e1006857–e1006857 (2018).
11. González-Carneiro, D., Bieniasz, P. D., Berkhour, B. & Prasad, V. R. Mechanisms of attenuation by genetic recoding of viruses. mBio 12, e02328–e02320 (2021).
12. Talloch, E., Atkinson, N. J., Evans, D. J., Ryan, M. D. & Simmonds, P. RNA virus attenuation by codon pair deoptimisation is an artefact of increases in CpG/UpA dinucleotide frequencies. eLife 3, e04531 (2014).
13. Gonçalves-Carneiro, D., Takata, M. A., Ong, H., Shilton, A. & Bieniasz, P. D. Origin and evolution of the zinc finger antiviral protein. *PloS Pathog.* **17**, e1009545 (2021).

14. 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. USA* **116**, 24303–24309 (2019).

15. Luo, X. et al. Molecular mechanism of RNA recognition by zinc-finger antiviral protein. *Cell Rep.* **30**, 46–52.e6 (2020).

16. Ficarelli, M. et al. CpG dinucleotides inhibit HIV-1 replication through zinc finger antiviral protein (ZAP)-dependent and -independent mechanisms. *J. Virol.* **94**, e0337–19 (2022).

17. le Nouët, C. et al. Optimization of the codon pair usage of human respiratory syncytial virus paradoxically resulted in reduced viral replication in vivo and reduced immunogenesis. *J. Virol.* **94**, e01296–19 (2020).

18. 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 (2011).

19. Greenbaum, B. D., Levine, A. J., Bhanot, G. & Rabdan, R. Patterns of evolution and host gene mimicry in influenza and other RNA viruses. *PloS Pathog.* **4**, e100079 (2008).

20. le Nouët, C. et al. Genetic stability of genome-scale deoptimized RNA virus vaccine candidates under selective pressure. *Proc. Natl. Acad. Sci. USA* **114**, E386–E395 (2017).

21. Burns, C. C. et al. Modulation of poliovirus replicative fitness in HeLa cells by deoptimization of synonymous codon usage in the capsid region. *J. Virol.* **80**, 3225–3227 (2006).

22. Zell, R. Picornaviridae—the ever-growing virus family. *Arch. Virol.* **163**, 299–318 (2018).

23. Tee, H. K., Zainol, M. I., Sam, I.-C. & Chan, Y. F. Recent advances in the understanding of enterovirus A71 infection: a focus on neuropathogenesis. *Expert Rev. Anti Infect. Ther.* **19**, 733–747 (2021).

24. Tan, C. W., Tee, H. K., Lee, M. H. P., Sam, I.-C. & Chan, Y. F. Enterovirus A71 DNA-launched infectious clone as a robust reverse genetic tool. *PloS ONE* **11**, e0162771 (2016).

25. Tee, H. K. et al. Electrostatic interactions at the five-fold axis alter heparin-binding phenotype and drive enterovirus A71 virulence in mice. *PloS Pathog.* **15**, e1007863 (2019).

26. Lee, M. H. P. et al. Vaccine candidates generated by codon and codon pair deoptimization of enterovirus A71 protect against lethal challenge in mice. *Vaccine* **39**, 1708–1720 (2021).

27. Caine, E. A., Moncla, L. H., Ronderos, M. D., Friedrich, T. C. & Osorio, J. E. A single mutation in the VP1 of Enterovirus 71 is responsible for increased virulence and neurotropism in adult interferon-deficient mice. *J. Virol.* **90**, 8592–8604 (2016).

28. Shih, C. et al. Immunocompetent and immunodeficient mouse models for Enterovirus 71 pathogenesis and therapy. *Viruses* **10**, 674 (2018).

29. Chen, S. et al. Structure of N-terminal domain of ZAP indicates how a zinc-finger protein recognizes complex RNA. *Nat. Struct. Mol. Biol.* **19**, 430–435 (2012).

30. Wang, Y. et al. Stable stem enabled Shannon entropies distinguish non-coding RNAs from random backgrounds. *BMC Bioinformatics* **13**, S1 (2012).

31. Ficarelli, M. et al. KHNYN is essential for the zinc finger antiviral protein (ZAP) to restrict HIV-1 containing clustered CpG dinucleotides. *elife* **8**, e46767 (2019).

32. Dominguez, D. et al. Sequence, structure, and context preferences of human RNA binding proteins. *Mol. Cell* **70**, 854–867.e9 (2018).

33. Floyd-Smith, G., Slattery, E. & Lengyel, P. Interferon action: RNA cleavage pattern of a 2′-5′-oligoadenylate-dependent endonuclease. *Science* **212**, 1030–1032 (1981).

34. Wreschner, D. H., McCauley, J. W., Skehel, J. J. & Kerr, I. M. Interferon action—sequence specificity of the ppp(A2 p)nA-dependent ribonuclease. *Nature* **299**, 414–417 (1981).

35. Gaunt, E. et al. Elevation of CpG frequencies in influenza A genome attenuates pathogenicity but enhances host response to infection. *elife* **5**, e12735 (2016).

36. Yew, N. S. et al. CpG-depleted plasmid DNA vectors with enhanced safety and long-term gene expression in vivo. *Mol. Ther.* **5**, 731–738 (2002).

37. Faust, S. M. et al. CpG-depleted adenov-associated virus vectors evade immune detection. *J. Clin. Investig.* **123**, 2994–3001 (2013).

38. Yeo, G. & Burge, C. B. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.* **11**, 377–394 (2004).

39. Gonçalves-Carneiro, D., McKeating, J. A. & Bailey, D. The measles virus receptor SLAMF1 can mediate particle endocytosis. *J. Virol.* **91**, e02255–16 (2017).

40. Ulrike, M. et al. Functional role of type I and type II interferons in antiviral defense. *Science* **264**, 1918–1921 (1994).

41. Liberatori, R. A. et al. Rhabdo-immunodeficiency virus, a murine model of acute HIV-1 infection. *elife* **8**, e49875 (2019).

Acknowledgements

We thank the Rockefeller University CRISPR and Genome Editing Resource Center for assistance with knockout mouse generation and the Rockefeller University Genome Resource Center for sequencing of CLIP libraries. This work was supported by grants from the NIAID R01AI50111 and the Center for HIV RNA studies U54AI150470 (to P.D.B.). This article is subject to HHMI’s Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the author-accepted manuscript of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

Author contributions

D.G.-C. and P.D.B. designed the experiments. D.G.-C. performed the experiments. D.G.-C., E.M. and J.D.S. developed and maintained mouse lines. X.L. developed the CLIP methodology. Y.F.C. provided critical reagents. D.G.-C. and P.D.B. wrote the paper with input from Y.F.C.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41564-022-01223-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41564-022-01223-8.

Correspondence and requests for materials should be addressed to Paul D. Bieniasz.

Peer review information *Nature Microbiology* thanks Jeremy Luban, Yi-Ling Lin and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Extended Data Fig. 1 | See next page for caption.
Extended Data Fig. 1 | HIV-1 platform for analysis of CpG effects and ZAP-deficient MT4 cells. (a) Details of the insertion of BstEII and ClaI restriction sites at nucleotide positions 6325 and 6771, respectively, in the HIV-1 genome. (b) Summary of the number of CpG dinucleotides present in the HIV-1 reporter virus genome and in the BstEII-ClaI interval in env. (c) Western blot analysis of 293T cells cotransfected with HIV-1 (WT) or HIV-1 (CG-43) and plasmids expressing WT ZAP or the ZAP deletion mutants (Δ3, Δ24 and Δ191) recovered from the CRISPR/Cas9 edited MT4 cells. Blots are representative of 1 experiment. (d) Infectious virus yield from 293T cells cotransfected with HIV-1 (WT) or HIV-1 (CG-43) and plasmids expressing WT ZAP or the ZAP deletion mutants (Δ3, Δ24 and Δ191) recovered from the CRISPR/Cas9 edited MT4 cells. (e) Viral replication in ZAP-expressing and ZAP-knockout MT4 cells infected with wildtype (WT) or CG-43 mutant or a mutant containing the BstEII and ClaI restriction sites with 0 CpGs in the intervening region (CG-0). Each day, the percentage of infected cells was measured by flow cytometry. (f) Summary of the percentage of GFP-positive cells at day 4 post initial infection. Mean and SD from 3 independent experiments is plotted.
Extended Data Fig. 2 | Mononucleotide composition of HIV-1 mutants. (a) Summary of mononucleotide composition changes relative to the wildtype HIV-1 in the BstEII-ClaI interval for CG-0/A+/U+/C+/G+ and CG-14/A+/U+/C+/G+ mutants. (b, c) Replication of CG-0/A+/U+/C+/G+ mutant viruses in ZAP-positive and ZAP-negative MT4 cells. The percentage of GFP positive cells was measured daily (b) and the mean and SD of 3 independent experiments measured at day 4 post infection (c).
Extended Data Fig. 3 | Effect of genome position in ZAP antiviral activity. (a) Diagram of the location of the mutated sequence in the env gene. Rev response element (RRE) is located between nucleotides 1560 and 1800. Position coordinates for each mutant are indicated by the location of the first and last CpG dinucleotides. (b) Replication of a collection of HIV-1 mutants (CG-15(110) through CG-15(1330)) each containing 15 CpG dinucleotides at different positions in the env gene as indicated in (a). Numbers in parentheses indicate the nucleotide position within env of the first introduced CpG. Percentage of infected cells was monitored by flow cytometry. (c) Summary of percentage of infected cells 4 days post infection, mean and s.d. of three experiments is plotted. (d) Summary of overall A and U content in the mutated regions of the mutants depicted in (a) (plus 10 nucleotides 5’ and 3’ of the first and last CpG, respectively) and mean number of nucleotides between each CpG. nt, nucleotide. (e) Replication of CG-15(889) recoded to contain high levels of adenine (CG-15(889)/A+). Cells were infected and infection was monitored as in (a).
Extended Data Fig. 4 | Replication and monitoring of CpG reversion during passage of CpG enriched HIV-1. (a) Replication of the cG-43 HIV-1 mutant in ZAP-deficient MT4 cells in three replicate cultures (Rep1-3). Percentage of infected cells was monitored by flow cytometry. (b, c) Sequence analysis of CG-15 HIV-1 mutant during replication. Unique point mutations in each replicate (b) and the total number of CpGs in each replicate at each passage (c) are indicated.
Extended Data Fig. 5 | Replication and monitoring of CpG reversion during passage of CpG enriched EV-A71. (a) EV-A71 RNA levels measured at the end of passage #5, 60h post-infection in three replicate cultures (mean and s.d. of 3 replicates is plotted). (b) PCR amplification of the modified region of EV-A71/CG-48/A+ in three replicates at the end of a long-term passage in ZAP+/+ and ZAP−/− cells (amplified region is approximately 1kb). (c–e) Replication of EV-A71 wildtype and mutant reporter viruses in ZAP+/+ and ZAP−/− HeLa cells. NanoLuc luciferase activity was measured every 12h (c), or quantification of viral RNA at day 4 (d), mean and s.d. plotted; data from 3 replicates. Alternatively replication of EV-A71 wildtype and mutant viruses monitored by TCID50 measurement at each timepoint (e); 3 replicates shown.
Extended Data Fig. 6 | Generation of a ZAP knockout mouse line and ZAP-dependent attenuation of EV-A71 lacking the NanoLuc reporter.

(a) Schematic diagram of the ZC3HAV1 locus in the Mus musculus genome. Sequence chromatogram of CRISPR target site PCR product amplified from homozygous ZAP−/− mice. A frameshift mutation in the essential N-terminal RNA binding domain via the insertion of two nucleotides is indicated. (b) Western blot analysis of mouse peripheral blood mononuclear cells (PBMCs) isolated from ZAP+/+IFNAR−/− and ZAP−/−IFNAR−/− mouse lines. (c) Replication of EV71 WT and EV71/CG48/A+ in ZAP+/+ and ZAP−/− HeLa cells, monitored by TCID50 measurement at each timepoint; 3 replicates are shown.
Extended Data Fig. 7 | Clinical score in individual mEV71/CG48/A+ infected ZAP+/+ and in ZAP−/− mice. (a, b) Clinical scores of individual mice following infection of one-day old ZAP+/+ (a) and ZAP−/− (b) mice with mEV-A71/CG-48/A+.
Extended Data Fig. 8 | mEV-A71/CG-48/A+ elicits protective antibodies in ZAP+/+ mice. (a) Neutralizing activity in plasma from mice after mEV-A71/CG-48/A+ infection (n = 4, at day 5 of age, blood collection at 6-weeks) or mock-infected mice (n = 3) evaluated using EV-A71 NanoLuc luciferase reporter virus. 293T cells were infected with the antibody:virus mixture, incubated for 48 h and luciferase activity was measured. Mean and SD plotted.

(b) Clinical score and survival probability following mEV-A71 (wildtype) infection of the offspring of ZAP+/+IFNAR−/− females previously inoculated with mEV-A71/CG-48/A+ (n = 17 pups from 3 different females) or mock-infected (n = 7 pups from 1 female) at day 5 of age. Clinical score and survival were assessed daily until weaning age. Mean and SEM plotted.
Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

CLIP-Seq data were processed and analyzed using python scripts as part of the FASTX-Toolkit version-0.0.14 (available here: Code used to map and calculate read counts from CLIP-Seq experiments is available from http://hannonlab.cshl.edu/fastx_toolkit/ and can be cloned from GitHub under https://github.com/agordon/fastx_toolkit).

Data analysis

Flow cytometry data were analyzed using FlowJo version 10.8.0. All other data were plotted and analyzed using GraphPad Prism version 9.2.0. Statistical test were performed using GraphPad Prism version 9.2.0. Analysis of potential splicing sites in recoded sequences was performed using MaxEntScan (available here: http://hollywood.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. Github). See the Nature Portfolio guidelines for submitting code & software for further information.
Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data that support the findings of this study are available in the Source Data files. Sequencing data resulting from CLIP-Seq experiments have been deposited in the NCBI GEO database and can be accessed using the accession code GSE208611. The NHG HIV-1 genome we used in this study can be accessed through the NCBI Nucleotide database using the accession code MF944225.1.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

| Reporting on sex and gender | N/A |
|----------------------------|-----|
| Population characteristics  | N/A |
| Recruitment                | N/A |
| Ethics oversight           | N/A |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- [x] Life sciences  
- [ ] Behavioural & social sciences  
- [ ] Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size | Sample size was calculated based on previous experiments and the authors' experience. For animal experiments, animal group size was determined based on Tee et al. 2019 Plos Pathogens (PMID: 31730673), who used similar infection conditions with EV-A71. |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Data exclusions | No data was excluded from this study with the exception of some CLIP experiments with read counts that were too low to be reliable. |
| Replication | All experiments presented in this manuscript were reliably replicated, at least 3 times, and presented equivalent results. CLIP experiments were performed twice, since these are inherently complex experiments with frequent reagent deterioration. Both experiments yielded remarkably similar results with very high read counts. |
| Randomization | All mice were infected at the same age and with the same infectious dose. Littermates of both sexes were included in the same experimental groups. Mice litters were randomly chosen to be infected with wildtype or mutant viruses as they became available from ongoing matings. Female mice that survived infection were then paired with non-exposed males; their offspring was challenged with wildtype viruses as mice litters become available without deliberate selection. All other experiments that did not involve mice were not subject to randomization since experimental units measurements were performed by machines and not subject to operator's bias. |
| Blinding | No blinding was used in this study since most of the quantitative data presented were measured by a machine. In infection experiments in mice that required the attribution of a previously published clinical score, we based scores on the observation of clear and obvious symptoms including death of the animal, the complete paralysis of 1 or 2 limbs, and hunched position adopted by some mice. All infections in mice resulted in a binary survival score (dead or alive) that could not be subject to interpretation by the operator. |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.
Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| ☒   | Antibodies            |
| ☒   | Eukaryotic cell lines |
| ☒   | Palaeontology and archaeology |
| ☒   | Animals and other organisms |
| ☒   | Clinical data         |
| ☒   | Dual use research of concern |

Methods

| n/a | Involved in the study |
|-----|-----------------------|
| ☒   | ChIP-seq              |
| ☒   | Flow cytometry        |
| ☒   | MRI-based neuroimaging |

Antibodies

**Antibodies used**

- Anti-ZC3HAV1 Rabbit Polyclonal Antibody (Anti-ZAP), ProteinTech, 16820-1-AP, Lot# 00046982. Used in western blotting at 1:5000 dilution.
- Anti-α-Tubulin Mouse Monoclonal Antibody, Millipore-Sigma, TS168, Batch# 0000089494. Used in western blotting at 1:10,000 dilution.
- Anti-ZC3HAV1 Rabbit Polyclonal Antibody, Abbexa, abx124715, Lot#A2002958K. Used in western blotting at 1:300 dilution.
- Anti-HA epitope tag Rabbit Monoclonal Antibody, clone 600-401-384, Rockland. Used in western blotting at 1:5000.
- Anti-HIV-1-Env (Gp160/Gp120), Goat Polyclonal Antibody, clone 12-6205-1, American Research Products. Used in western blotting at 1:1000 dilution.

**Validation**

- Anti-ZC3HAV1 (ProteinTech) antibody was validated for western blotting in siRNA and gRNA experiments targeting ZAP in several human cells (PMID: 28953888).
- Anti-α-Tubulin antibody was validated by manufacturer (Sigma’s Enhanced Validation antibodies, https://www.sigmaaldrich.com/US/en/product/sigma/TS168) and by other research groups against protein extracts from human cells (PMID: 3143934) and multiple mouse organs (PMID: 23748901).
- Anti-ZC3HAV1 (Abbexa) antibody was validated by manufacturer for western blotting by using mouse lung and thymus protein extracts (https://www.abbexa.com/zc3hav1-antibody) and in this study (Extended Figure 6) using mouse PBMCs from ZAP+/+ and ZAP-/- mouselines.
- Anti-HA (Rockland) antibody was validated for western blotting by using mouse lung and thymus protein extracts from mouse PBMCs containing HA-tagged proteins.
- Anti-HIV-1-Env was validated for western blotting previously (PMID: 33901262) and in this study by analyzing protein extracts from cells infected with HIV-1.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

**Cell line source(s)**

- Human Embryonic Kidney 293T (HEK293T) and HeLa cells were obtained from ATCC. MT4 human lymphocyte cell line was obtained from the NIH AIDS Reagent repository.

**Authentication**

- Cells were authenticated by microscopic inspection.

**Mycoplasma contamination**

- Cell lines were not tested for the presence of mycoplasma.

**Commonly misidentified lines**

(See ICLAC register)

- No commonly misidentified lines were used in this study.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

**Laboratory animals**

- House mouse (Mus musculus), C57BL/6J, used in experiments and/or colony maintenance between 1-day old and 9-months of age. A transgenic ZC3HAV1 knockout line (homozygous ZAP-knockout) was derived from C57BL/6J. We also used a C57BL/6J Ifnar1-/- knockout line (MMRRC #32045). All mice were housed in standard housing conditions (12h light-dark cycles, room temperature of 65-75°F and 40-60% humidity). Mice had unrestricted access to the food and water.

**Wild animals**

- We did not use any wild animals.

**Reporting on sex**

- Sex was not considered in this study. We used mice of both sexes in disease progression experiments and females for maternal antibody protection experiments.

**Field-collected samples**

- This study did not involve field-collected samples.

**Ethics oversight**

- Mouse handling and experimental protocol (18047-H) was approved by institutional animal care and use committee (IACUC) protocol.
Ethics oversight of the Rockefeller University.

Note that full information on the approval of the study protocol must also be provided in the manuscript.