bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Engineering T cell resistance to HIV-1 infection via knock-in of peptides from the heptad repeat 2 domain of gp41

Alexandra Maslennikova[1*], Natalia Kruglova[2*], Svetlana Kalinichenko[1], Dmitriy Komkov[2], Mikhail Shepelev[2], Dmitriy Golubev[2], Andrei Siniavin[3], Andrei Vzorov[3,4], Alexander Filatov[5] and Dmitriy Mazurov[2#]

[1]Cell and Gene Technology Group, Institute of Gene Biology RAS, Moscow, Russia

[2]Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Institute of Gene Biology RAS, Moscow, Russia

[3]Gamaleya Center for Epidemiology and Microbiology, Moscow, Russia

[4]Faculty of Biology, Lomonosov Moscow State University, Moscow, Russia

[5]Laboratory of Immunochemistry, National Research Center Institute of Immunology of Federal Medical Biological Agency of Russia, Moscow, Russia

[*]equal contribution

[#]Correspondence should be addressed to D.M. (dvmazurov@yandex.ru)

Keywords: CRISPR/Cas9, knock-in, GPI-proteins, CD52, gp41, fusion inhibitors, cell sorting, HIV-1 resistance

ABSTRACT

Previous studies suggest that short peptides from the heptad repeat 2 (HR2) domain of gp41 expressed on the cell surface are more potent inhibitors of HIV-1 entry than soluble analogs. However, their therapeutic potential has only been examined using lentiviral vectors. Here, we aimed to develop CRISPR/Cas9-based fusion inhibitory peptide knock-in (KI) technology for the generation and selection of HIV-1-resistant T cells. First, we cloned a series of HIV-1 fusion inhibitory peptides and embedded them in CD52, the shortest GPI-anchored protein, which efficiently delivers epitope tags to the cell surface and maintains a sufficient level of KI. Among the seven peptides tested, MT-C34, HP-23L, and 2P23 exhibited significant activity against both cell-free and cell-to-cell HIV-1 infection. Unlike membrane-bound peptides, the shed variant of MT-C34 provided insufficient protection against HIV-1 due to its low concentrations in the culture medium. Using Cas9 plasmids or ribonucleoprotein electroporation and cell sorting with antibodies raised against gp41 peptides, we generated CEM/R5 cells with biallelic KI of MT-C34 (embedded in CD52 for expression in lipid rafts) and 2P23 (N-terminally fused to CXCR4). In combination, these peptides provided a higher level of protection than individual KI. By extending homology arms and substituting PCR donor DNA with a plasmid containing signals for nuclear localization, we achieved KI of MT-C34 into *CXCR4* loci and HIV-1 proviral DNA at levels of up to 35% in CEM/R5 cells and increased KI occurrence from undetectable to 4% in CD4 lymphocytes. Regardless of HDR efficiency, sorted cells were completely protected from X4 and R5 strains of HIV-1 and capable of expanding during acute viral replication. Thus, the developed CRISPR/Cas9 platform opens up new possibilities for HIV therapy based on protective peptide KI that, in primary human cells, had previously been considered inefficient.

AUTHOR SUMMARY

HIV is a human lentivirus that infects CD4-positive immune cells and, when left untreated, manifests in the fatal disease known as acquired immunodeficiency syndrome. Even after suppression with antivirals, HIV persists in the organism as a latent provirus. One way to decrease the viral reservoir is to engineer CD4 lymphocytes that are genetically resistant to HIV-1 infection via entry molecule knockout or expression of different antiviral genes. Peptides from the heptad repeat (HR) domain of gp41 are potent inhibitors of HIV-1 fusion, especially when designed to express on the cell surface. Individual gp41 peptides encoded by therapeutic lentiviral vectors have been evaluated and some have entered clinical trials. However, a CRISPR/Cas9-based gp41 peptide delivery platform that operates through concomitant target gene modification has not yet been developed due to low knock-in (KI) rates in primary cells. Here, we systematically evaluated the antiviral activity of different HR2-peptides cloned into the shortest carrier molecule, CD52. The resulting small-size transgene constructs encoding selected peptides, in combination with improvements to enhance donor vector nuclear import, helped to overcome precise editing restrictions in CD4 lymphocytes. Using KI into CXCR4, we demonstrated different options for target gene modification, effectively protecting edited cells against HIV-1.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

INTRODUCTION

Human immunodeficiency virus type I (HIV-1) is a retrovirus, the genome of which stably integrates into the host's DNA during viral infection. As such, HIV-1 eradication is only possible when all infected cells in the host organism are either killed or cured via "DNA surgery." The first instance of HIV being cured was reported after the allogeneic transplantation of CCR5(Δ32) bone marrow stem cells to a patient in Berlin (1, 2). In other patients, however, engraftments were not as successful and HIV-1 ultimately rebounded (3). The Berlin patient's case inspired researchers to develop genetic tools to make cells resistant to HIV. Zinc finger nucleases (ZFNs) specific to CCR5 were the first programmed genomic nucleases to enter HIV clinical trials (4). With the adaptation of a bacterial defense system CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9) for gene editing in mammalian cells (5, 6), diverse genetic manipulations, including HIV proviral excision (7, 8) and/or inactivation (9, 10) and CCR5 and CXCR4 knockout (11-13), have been found to be effective and easily achievable. However, three major problems related to CRISPR-Cas9-based editing—emergence of viable HIV escape mutants during indel formation (9, 14, 15), off-target activity of Cas9 (16, 17), and low efficiency of precise genome modification via homology-directed repair (HDR) (18, 19)—create a serious barrier to the use of CRISPR in clinical practice.

Of the various anti-HIV genetic strategies, the inhibition of viral entry is considered to be the most efficient. Short peptides from the heptad repeat 2 (HR2) domain of gp41 inhibit HIV entry by interacting with the HR1 region of gp41, blocking the formation of gp41 6-helix bundles required for membrane fusion (20). The first generation of HR2 peptides consisted of relatively long T-20, C36, and C46 peptides to which the virus rapidly evolved resistance. Subsequent modifications aimed to 1) enhance the interaction of the peptide with the HR1 deep hydrophobic pocket by adding an M-T hook (21, 22) or introducing mutations that stabilized the α-helices (23, 24), or 2) broaden protection against HIV-2 and SIV (25), resulting in the development of a second and third genus of short peptides with higher inhibitory potency and lower viral resistance. To date, only two fusion inhibitor peptides have been approved for clinical use: enfuvirtide (T-20), approved by the FDA, and sifuvirtide, approved for use in China. While synthetic gp41 peptides and their chemical modifications have been studied extensively, much less is known about the potency of genetically encoded analogs. The first studies from Dorothee von Laer's lab showed that T-20 and C46 peptides expressed on the cell surface in the context of the human low-affinity nerve growth factor receptor (LNGFR) are very potent inhibitors of HIV-1 entry (26, 27). In 2016, Liu et al. reported that a shorter C34 peptide linked to a GPI-attachment signal from the decay-accelerating factor (DAF) and expressed in lipid rafts confers strong cellular resistance to HIV-1 infection (28). Recently, He's lab demonstrated that the third-generation peptide 2P23, when expressed in a similar way, fully protects cells against HIV-1, HIV-2, and SIV infection (29). Besides transmembrane linking or GPI-anchoring, direct fusion of C34 to the N-termini of CXCR4 and CCR5 has been reported to confer cells with a strong resistance to HIV infection *in vitro* and *in vivo* (30). Collectively, these data indicate that gp41 peptides, when expressed on a plasma membrane and localized in lipid rafts or co-receptor sites where HIV fusion occurs, become very active in comparison to the soluble form. However, in all these studies, peptides were expressed from retro- or lentiviral vectors with no aim of adapting them for knock-in into endogenous genes. Since HDR-based gene editing is sensitive to the size of the DNA insert, the shortest construct comprised of a peptide and a carrier molecule is desirable in order to achieve a reasonable level of precision in the genetic modification.

We have previously described a novel method for gene-edited cell selection called SORTS (surface oligonucleotide knock-in for rapid target selection), which is based on the targeted biallelic knock-in of epitope tags embedded in CD52 (31). CD52 is the shortest human GPI-anchored protein that can effectively deliver peptides to the plasma membrane, where they serve as makers for fluorescence-activated cell sorting (FACS) isolation of edited cells. In the present study, we modified SORTS by substituting the epitope tag in CD52 with a peptide from the HR2 domain of gp41. We selected the shortest and most potent soluble peptides of the last generation described to date and engineered lentiviral vectors to stably express these peptides on the cell surface. Of the seven tested peptides, MT-C34, HP23L, and 2P23 demonstrated the highest activity against both cell-free and cell-to-cell infection by HIV-1. In contrast, modifications that aimed to secrete, shed, or noncovalently link the MT-C34 peptide to HIV chemokine co-receptors did not improve modified or bystander cell protection. Using CRISPR-Cas9 co-knock-in of MT-C34 and 2P23 (or HP23L) into the endogenous *CXCR4* gene followed by the sorting of double-positive cells labeled with anti-MT-C34 and anti-2P23(HP23L) antibodies (Abs), we isolated a pure population of CEM $CD4^+$ T cells, where the indicated peptides were N-terminally fused to CXCR4 or expressed separately in a GPI-linked form. Both genetic modifications preserved CXCR4 function and completely protected cells from different clades of HIV-1 with various tropisms, as well as from gp41 mutant viruses. Through the optimization of donor DNA, i.e., the extension of homology arms and substitution of the PCR donor with a plasmid, we achieved ~30% HDR efficiency at the *CXCR4* locus in CEM cells. We also achieved a reasonable level of peptide KI into human primary cells using modifications aimed to enhance the nuclear import of donor DNA. Thus, unlike vector-based HIV therapies, this approach relies on precise

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

genetic modification with CRISPR-Cas9 that enables the expression of protective peptides, as well as the concurrent modification of cellular receptors and isolation of modified cells.

RESULTS

Construction and antiviral activity of CD52-embedded gp41 peptides

We have previously demonstrated that SORTS can be adapted for the isolation of T cells that have been effectively cured of HIV-1 via the KI of a CD5HA2 construct, in which the peptide sequence SQTSSP, located in the middle of CD52, is replaced with an HA tag (31). We hypothesized that providing the selected cells with a protective peptide would render "cured" cells resistant to repeat infection with HIV. We selected several peptides from HR2 of gp41 with high and broad anti-HIV activity as reported in the literature and substituted the HA sequence in CD5HA2 constructs with one of the peptides specified in Figure 1A. We then jointed it to the upstream green fluorescent protein mClover via a P2A sequence for separate translation. Three peptides—T-20, C34, and C24 (an inactive form used as a negative control that represents the overlapped sequence of the first two peptides without a pocket-binding domain (32))—were the natural sequences from the HxB2 strain of HIV-1. The other five peptides had the following modifications: MT-C34 had M-T hook residues (22, 33) while P52 (34), MT-WQ-IDL (35), HP23L (36, 37), and 2P23 (25, 29) contained α-helix stabilizing mutations with or without an additional hook-like structure at the N- or C-terminus. To test the inhibitory activity of the peptides, the 293T/CD4/R5 cells—permissive to CCR5- and CXCR4-tropic HIV—were first transduced with one of the generated lentiviral constructs at MOI <0.3 and sorted based on mClover fluorescence. As the T-20 peptide was not properly expressed in our system (likely due to its high C-terminal hydrophobicity), it was excluded from all subsequent experiments. 293T cells with or without receptor/co-receptor/peptide expression were infected with an equal quantity of viral particles pseudotyped with R5-tropic JRFL, ZM135, or X4-tropic NL4-3 Env (referred to below as pseudoviruses (PVs)). Levels of HIV-1 cell-free infection were quantified using replication-dependent inLuc vectors (38, 39) (see Methods for details). P52 and MT-WQ-IDL demonstrated less potent inhibition (60–79% and 79–93% depending on Env, respectively) than other peptides, which blocked HIV-1 infection at levels of 98% or higher (Fig. 1B). The most protective peptides—C34, MT-C34, HP23L, and 2P23—were further tested in cell-to-cell infection. For this purpose, we transduced Raji/CD4/R5 cells with lentiviruses expressing these peptides and sorted them as outlined above. We then measured the levels of infection from non-permissive Raji producer cells to the Raji/CD4/R5 cells with or without peptides through inLuc transduction, as described above (39). Similar to cell-free infection, C34 and MT-C34 also actively inhibited HIV-1 cell-to-cell infection, while HP23L and 2P23 displayed less potent inhibition (Fig. 1C). Ultimately, we selected four peptides—C34, MT-C34, HP23L, and 2P23, embedded in a CD52 carrier molecule—with the highest potencies to inhibit HIV-1 entry.

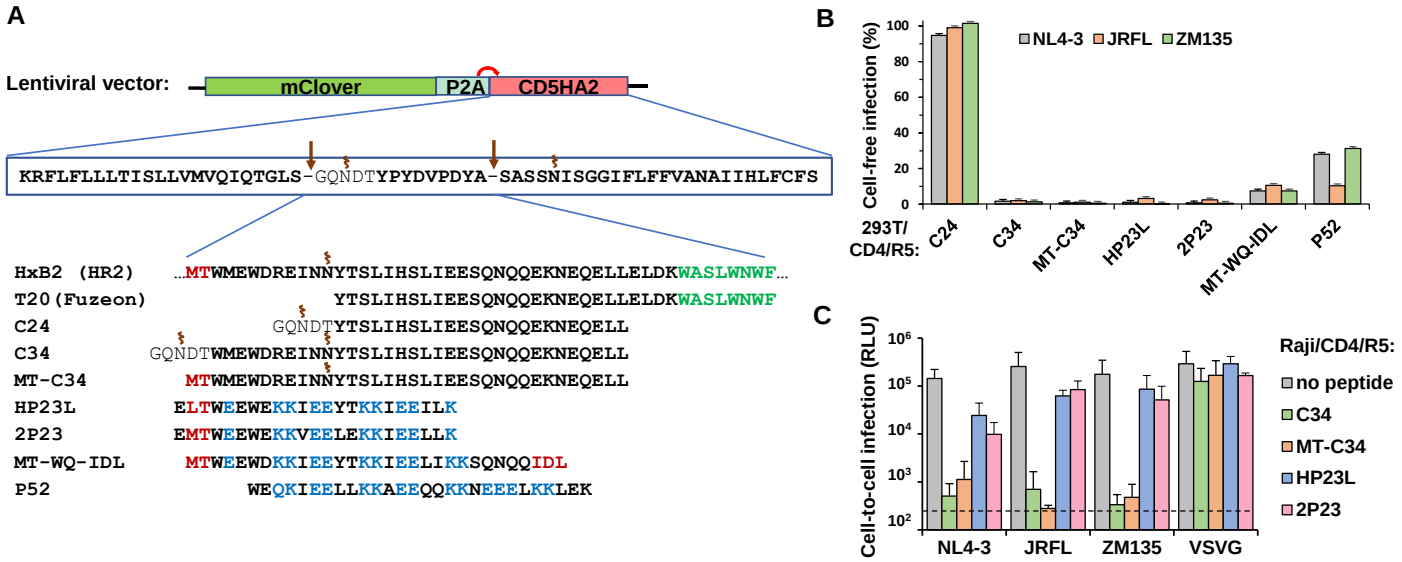

Figure 1. Construction and anti-HIV-1 activity of peptides from HR2-domain of gp41 expressed in the context of CD52. **A.** Schematic of generated lentiviral constructs for stable expression of gp41 peptides. Amino acids forming hook-like structure, α-helix-stabilizing mutations, and hydrophobic parts of peptides are highlighted in red, blue and green, respectively. Sites of proteolytic cleavage and N-glycosylation are indicated by arrows and spirals, respectively. **B.** Effects of GPI-anchored peptides on HIV-1 cell-free infection. 293T/CD4/R5 cells with the stable expression of indicated peptides were infected by the equal doses of HIV-1 VLPs pseudotyped with Envs from different strains of HIV-1 or VSV. Forty eight hours later, the levels of inLuc reporter transduction were measured. The infectivity values obtained for HIV-1 Envs were first adjusted to the respective infectivity data detected with VSV-G; the resulting value estimated in 293T/CD4/R5 cells without peptides was set at 100%, and the levels of infection for cells carrying indicated peptides were recalculated relative to that value. **C.** Effects of selected peptides on cell coculture infection. Raji cells were transfected to

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

produce VLPs pseudotyped with indicated Envs and then directly mixed to the indicated Raji/CD4/R5 target cells at 1:1 ratio. The levels of infectivity were estimated by measuring inLuc transduction in 54 h of coculture and presented as raw data. The background level of Luc activity is shown by dashed line. The data in B and C are representative of at least three independent measurements and shown as average values ± standard deviations.

Analysis of gp41 peptides and HIV receptor/co-receptor expression levels on the surfaces of transduced cells

In order to verify whether the inhibition of HIV-1 entry depends on the level of peptide and/or viral receptor/co-receptor expression on target cells, we generated mouse monoclonal antibodies (mAbs) against C24 (which could also recognize C34 and MT-C34 peptides) and HP23L, as well as a rabbit polyclonal Ab specific to the 2P23 peptide. 293T/CD4/R5 and Raji/CD4/R5 cells with or without peptides were individually labeled with the Abs indicated above or mAbs generated against CD4, CXCR4, and CCR5 and analyzed by flow cytometry. As expected, MT-WQ-IDL, P52, HP23L, and 2P23 were not recognized by anti-C24 mAbs, whereas the peptides C24, C34, and MT-C34 were intensely stained on the surface of 293T cells (Fig. 2A). Similarly, the anti-HP23L mAb and anti-2P23 Ab demonstrated specificity in peptide recognition. Unlike C34-based peptides, which had high levels of expression on both 293T and Raji cells, the HP23L and 2P23 peptides were expressed less intensely, especially on Raji cells (Fig. 2B). These two peptides lack additional sites of N-glycan attachment, which are present in C24, C34, and MT-C34 sequences (Fig. 1A). This suggests that the N-glycosylation of chimeric CD52 proteins at multiple sites may stimulate the export of carried peptides to the plasma membrane, which has been shown in our previous work (31) and similar studies (40, 41). The non-optimal expression of HP23L and 2P23 embedded in CD52 may lower their antiviral potency, as we observed in a cell co-culture infectivity test (Fig. 1C); however, the nature of a peptide can also determine its activity. Monitoring HIV-1 receptor and co-receptor expression on cells transduced with one of the peptides did not reveal any substantial differences between them, suggesting that the inhibitory activity of the peptides was not mediated by the downregulation of viral entry molecules. All the peptides were expressed on transduced cells at comparably high levels, except for CXCR4 on 293T cells, the endogenous level of expression of which is known to be low but sufficient for HIV-1 entry. In summary, gp41 peptides on transduced cells did not alter HIV-1 receptor and co-receptor levels of expression. Meanwhile, the level of expression of these peptides on plasma membranes is an important parameter that may influence their potency to inhibit HIV-1 infection.

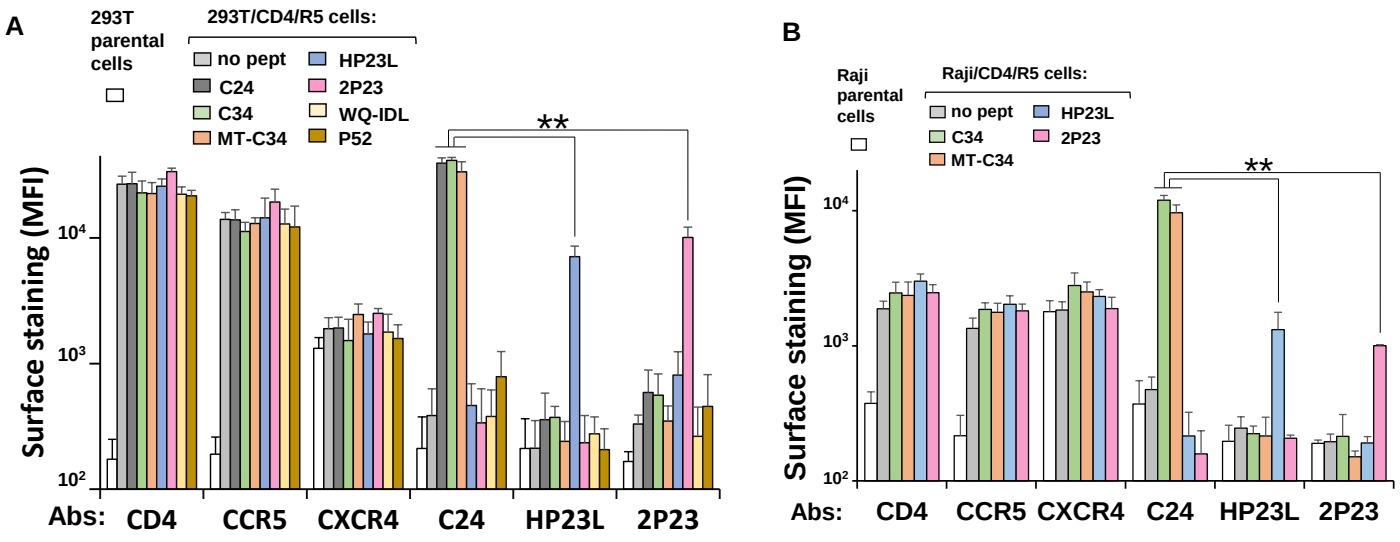

Figure 2. Analysis of gp41 peptide and HIV-1 receptor/coreceptor expression on surfaces of the stably transduced target cells. The 293T (**A**) and Raji **(B)** modified cells were labeled with indicated Abs, washed, and stained with the respective Alexa546-conjugated secondary Ab. Cells were analyzed using flow cytometry, and the data were presented as the mean fluorescence intensity (MFI). The average values of at least three independent measurements with standard deviations are shown. ** The differences are significant at $p<0.01$

Modifying the CD52-based MT-C34 peptide for secretion, shedding, or binding to HIV co-receptors

We reasoned that the anti-viral activity of gp41 peptides can be expanded from gene-modified cells to bystander cells by substituting the GPI-anchored form of a peptide with a soluble one. Matabaro et al. (42) demonstrated that by regulating the fatty acid remodeling of GPI-anchored proteins, it is possible to achieve partial membrane expression of a protein needed for FACS-sorting as well as protein secretion. Another interesting approach for enhancing the activity of an anti-HIV inhibitory peptide is non-covalently linking it to a CXCR4 or CCR5 co-receptor, as demonstrated for the 15D peptide (43).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We modified these approaches to adapt them for the CD52 molecular system. Specifically, to enable MT-C34 peptide secretion, we removed the GPI-attachment signal from the C-terminus of CD52 but left the leader peptide (LD) from CD52 (CD52-MT-C34s) or replaced it with LD from Gaussia luciferase (Glu-MT-C34s). Peptide shedding or binding to HIV-1 co-receptors was engineered by introducing the furin cleavage sequence RIRR (44) (MT-C34-R) or 15D (MT-C34-15D), respectively, between MT-C34 and the CD52 GPI-signal (Fig. 3A). The generated constructs were stably expressed in HIV-permissive 293T cells and tested in cell-free infection as described above. Surprisingly, both variants of secreted MT-C34 peptide failed to efficiently protect cells from HIV-1 infection, and the R and 15D modifications did not enhance the inhibitory activity of MT-C34 (Fig. 3B). We developed a sandwich ELISA system with the newly generated rabbit polyclonal and mouse monoclonal anti-C24 Abs that could detect concentrations of synthetic MT-C34 peptide as low as 5–10 nM. At this level of sensitivity, we were unable to register secreted forms of MT-C34 peptides in the supernatants of transduced 293T cells or detect their antiviral activity in the cell-free infection test (Fig. S1A–D). This suggests that short peptides may have trouble entering the secretory pathway and require elongation or concatemerization for efficient secretion (41). The R modification is characterized by a reduced level of MT-C34 cell surface expression and decreased protection against HIV-1 infection (Fig. 3C, D), while furin inhibitor I partially elevated MT-C34-R expression level (Fig. 3D, right bars). MT-C34 expression was insensitive to Brefeldin A (BrfA) treatment (Fig. 3E, red and orange lines) but MT-C34-R, in the presence of BrfA, largely accumulated inside cells and disappeared from the cell surface (Fig. 3E, dark and light green lines). This suggests that the R modification makes peptide synthesis more dependent on the secretory pathway. Thus, cell analysis data indicate that MT-C34-R is partially cleaved off from the plasma membrane and the remaining amount of the peptide is sufficient for live cell selection and protection against HIV-1, albeit at lower levels of potency. However, the desired level of the cleaved peptide was too low to be assayed by our ELISA system (Fig. S1B) or to inhibit HIV-1 infection (Fig. S1D). In summary, the anti-HIV inhibitory activity of gp41 peptides is highest when they are GPI-anchored to the membrane. Modifications aimed at releasing the peptide from cells or additionally recruiting it to an HIV co-receptor site did not help, or even reduced, the antiviral potency of the peptide.

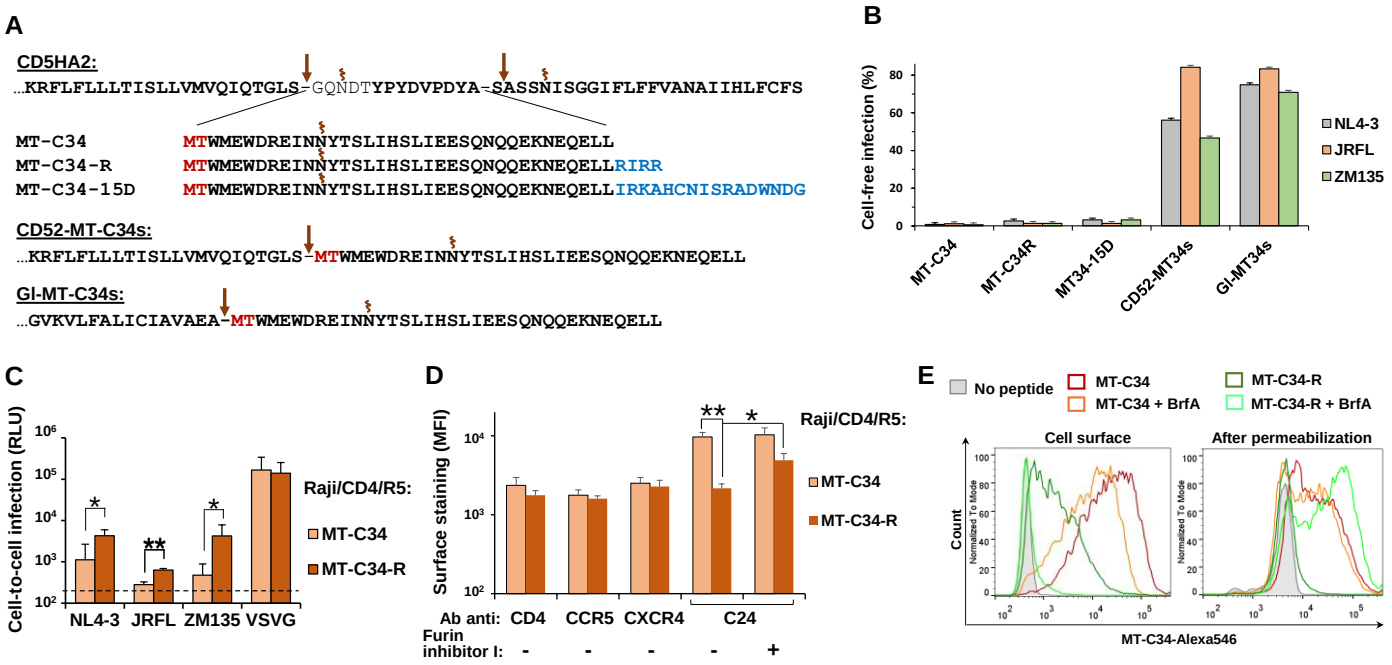

Figure 3. Modifying MT-C34 GPI-anchored peptide for secretion, shedding or coreceptor binding. A. Amino acid sequences of the generated constructs. The furin cleavage site (R) and the peptide from HIV-1 gp120 (15D) are shown in blue; see Fig.1A for the other designations. The effects of MT-C34 modifications on the levels of HIV-1 cell-free (**B**) and cell coculture (**C**) infection. The infectivity assays were set up as described for Fig.1B-C. **D**. The influence of modification R on the levels of HIV-1 receptor/coreceptors and MT-C34 expression on the surface of Raji/CD4/R5 cells. Immunofluorescent staining and analysis were performed as outlined for Fig. 2B. **E**. Comparison of the surface and intracellular expression of MT-C34 and MT-C34-R in Raji/CD4/R5 cells prior or after 16 h-treatment with brefeldin A (BrfA). Live cells were stained with anti-C24 mAb and anti-mouse Alexa546-conjugated secondary Ab (on the left), or cells were fixed and permeabialized with saponin first and then immunolabeled (on the right). Representative histogram overlays are shown. The average values with standard deviations from at least three independent experiments are shown in B-D. The differences calculated by Student *t*-test are significant at $p<0.05$ (*) or $p<0.01$ (**).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Selecting a strategy for HIV-1 fusion inhibitory peptide knock-in

SORTS offers a quick way to test KI efficiencies at multiple sites, as it does not require the generation of a donor plasmid. We have previously applied this method to target the capsid region in HIV-1 proviral DNA (31). Here, we focused on the *CD4* and *CXCR4* HIV entry cellular genes, which have a high T cell expression, and developed strategies to knock in peptides without respective gene knockout. Specifically, we designed guide RNAs (gRNAs) and PCR donors to target the start and stop codons of the *CD4* gene and the start codon of the *CXCR4* gene. To enable transgene expression, an *in-frame* strategy with an upstream or downstream P2A sequence (31) was applied during the construction of homology arms (see Fig. 4A and Table S2 for details). To assess KI rates, we first used CD5HA2 PCR donor DNAs. CEM/R5 cells were co-transfected with Cas9 and gRNA expression plasmids and donor DNA, as previously described (31). To monitor both KI and targeted gene expression rates, cells double stained for HA and for the respective cellular protein were analyzed by flow cytometry at the indicated time post transfection. Despite a rather efficient NHEJ-mediated inactivation of *CD4* and *CXCR4* gene ORFs at the start codons, a clear population of cells with HA KI (~6%) was detected only for the *CXCR4* gene (Fig. 4B). Importantly, HA-positive cells were also CXCR4-positive, indicating that at least one allele was precisely modified.

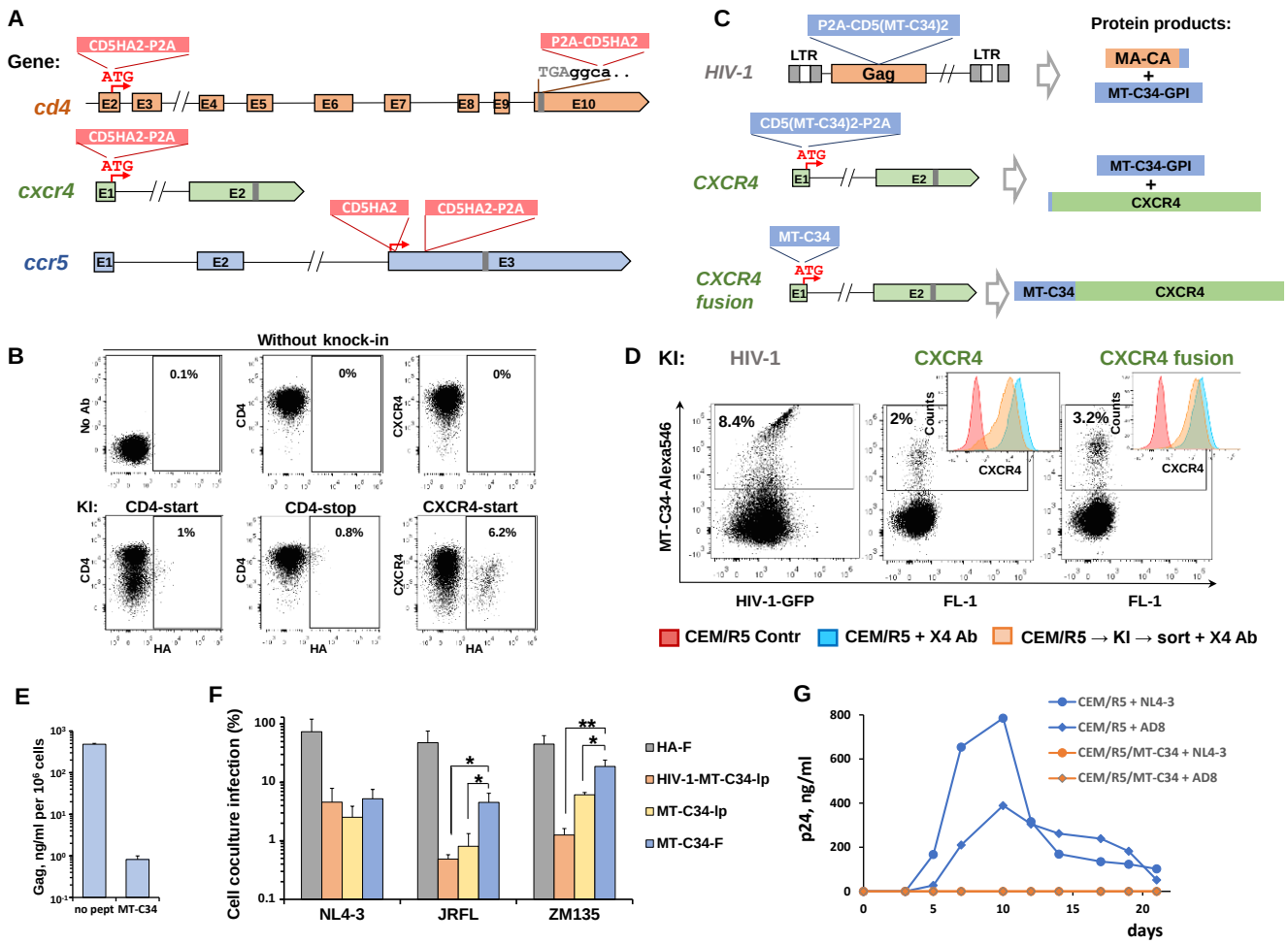

Figure 4. Testing different anti-HIV-1 strategies based on MT-C34 knock-in. A. Schematics showing cellular genes, targeted sites in exons (E) and donor DNA constructs (red boxes) selected for HA KI. **B**. The efficiency of HA surface expression after KI into different cellular genes. CEM/R5 cells were electroporated to initiate KI. At day 3, transfected and control cells were double-stained for surface HA and respective cellular protein and analyzed by flow cytometry. **C**. A schematic picture illustrating working strategy for MT-C34 KI into integrated HIV-1 provirus or *cxcr4* gene (on the left) and resulting protein expression products (on the right). **D**. The levels of MT-C34 and CXCR4 expression after KI in CEM/R5-HIV-1-GFPt (left Plot) and CEM/R5 (middle and right hand graphs). Cells were immunostained for MT-C34 at 3d posttransfection, and analyzed by flow cytometry. MT-C34^{+} cells were sorted and labeled with anti-CXCR4 mAb along with controls (embedded histograms). Typical FACS images are shown in B and D. **E**. The levels of viral Gag produced by CEM/R5-HIV-1-GFPt cells prior (left) and after (right) MT-C34 KI and FASC sorting. **F.** Susceptibility of isolated CEM/R5 cells with indicated KIs to HIV-1 cell coculture infection **G**. HIV-1 spreading assay on permissive CEM/R5-HA-F and resistant CEM/R5-MT-C34-lp cells challenged with NL4-3 or NL4(AD8) viruses. The levels of HIV-1 production in E and G were quantified by p24 ELISA. The data are presented as average values with standard deviations of at least three independent experiments. F – fusion, lp – lipid rafts. * and ** - the differences are significant at $p<0.05$ and $p<0.01$, respectively, by Student *t*-test.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Following this, we replaced HA with the MT-C34 sequence in pUCHR-CD5HA2-P2A-mClover and pUCHR-mClover-P2A-CD5HA2 plasmids and PCR-amplified donor DNAs from these vectors. This allowed us to express MT-C34 from an HIV-1 provirus or a *CXCR4* gene as a GPI-anchored peptide or to fuse MT-C34 directly to the N-terminus of CXCR4 (30) (Fig. 4C). The FACS analysis of CEM/R5 cells, transfected for KI of these donors, showed that all strategies were effective and resulted in the parallel expression of the MT-C34 peptide and CXCR4 (Fig. 4D). The MT-C34-positive CEM/R5 and CEM/R5-HIV-1 cells from the above experiments were sorted and then tested in a cell-to-cell infection assay with three different Envs (as described above), while concomitant inactivation of HIV-1 proviral DNA in CEM/R5-HIV-1 cells was estimated by p24 ELISA. Similar to HA (see Fig. 5C in (31)), integration of MT-C34 into the HIV-1 provirus greatly reduced viral particle production (Fig. 4E). CEM/R5 cells expressing HA-CXCR4 were relatively sensitive to X4- and R5-tropic HIV-1 infection (Fig. 4F). This is in line with the findings of Leslie et al. (30) and suggests that KI-based N-terminal modification preserved the function of CXCR4 at least at the level of viral co-receptors. Remarkably, all variants of MT-C34 peptide knocked into CXCR4 exhibited potent and comparable levels of protection against X4-tropic HIV-1 (Fig. 4F, left bars). In contrast, infection with the R5-tropic Envs JRFL and ZM-135 was more effectively inhibited by the variant MT-C34-lp, expressed either from proviral DNA or the *CXCR4* locus, than by MT-C34-F, which was fused directly to CXCR4 (Fig. 4F, middle and right bars). This indicates that peptide localization in lipid rafts can better facilitate the inhibition of both CXCR4 and CCR5-dependent HIV-1 fusion than its localization at the CXCR4 co-receptor site. Additionally, we performed an HIV-1 spreading assay with the molecular clones NL4-3 and NL4(AD8) and found that CEM/R5 cells bearing MT-C34-lp were completely resistant to these strains of HIV-1 (Fig. 4G). In summary, we adapted SORTS for KI of the gp41 peptide MT-C34 into the *CXCR4* gene, where it can be N-terminally fused to CXCR4 or translated and expressed separately. Both modifications led to the effective and broad inhibition of HIV entry without any CXCR4 knockout.

Biallelic knock-in of two distinct gp41 peptides into the *CXCR4* gene synergistically enhances protection of T cells against HIV-1

As previously demonstrated, the SORTS procedure is capable of isolating a highly pure population of cells, the majority of which carry a precise gene modification in both alleles, when using two different epitope tags (31). We hypothesized that the biallelic KI of two distinct fusion inhibitory peptides would increase the cell's level of protection against HIV-1. To explore this, we performed all possible combinations of single or dual peptide KI into the *CXCR4* gene, excluding peptide constructs that were poorly expressed (i.e., 2P23 and HP23L embedded in CD52; Fig. 2) or could not be stained simultaneously by antibodies. Following KI, CEM/R5 cells were immunostained, analyzed by FACS (see two typical examples and schematics in Fig. 5A), sorted, and tested in the cell-to-cell infection assay (Fig. 5B). As expected, the two control CEM/R5 sublines, expressing HA alone or in combination with C24 fused to CXCR4, were not substantially protected from HIV-1. The levels of inhibition mediated by MT-C34, HP23L, and 2P23 peptides fused to CXCR4 were relatively comparable and protected cells both from X4- and R5-tropic HIV-1. However, all four paired combinations of peptides that we tested exerted higher levels of HIV-1 inhibition compared to the individual peptides. From these results, we concluded that biallelic KI conveyed a protective advantage to cells.

For further analysis, we selected MT-C34-lp/2P23-F cells with two relatively distant HIV-1 fusion inhibitory peptides expressed at distinct locations on the plasma membrane. Genetic analysis of these cells showed that PCR-amplified target loci contained two DNA integration products of expected size (Fig. 5C); these were then purified from the gel and verified by sequencing (Fig. S2). Using Triton X-100 protein extraction followed by the immediate fixation and flow cytometry analysis of cell residues (45), we confirmed that the peptide MT-C34 was localized in a detergent-resistant membrane microdomain, whereas 2P23 fused to CXCR4 as well as CD4, characterized by intermediate levels of TX-100 solubilization (Fig. S3). Next, we tested whether our combination of cell surface peptides was active against gp41 mutant HIV-1. To this end, four single mutations—V38A and E49K in the HR1 domain and N126K and E136G in the HR2 domain of gp41—that have been shown to confer HIV-1 resistance to the soluble peptides T-20, C34, and MT-C34 (46, 47) were introduced in an NL4-3 genome (Table S1). All generated mutants replicated well in CEM/R5 cells, whereas MT-C34-lp$^+$/2P23-F$^+$ cells were completely protected from infection with any of the tested HIV-1 mutants (Fig. 5D). To test whether our gene-edited HIV-resistant cells would be able to grow during acute infection with HIV-1, we first optimized transfection with a Cas9 ribonucleoprotein (RNP) complex (Fig. S4A) and generated MT-C34-lp$^+$/2P23-F$^+$ CEM/R5 cells using either a CRISPR/Cas9 plasmid or RNP electroporation. We then set up a culture mixture that included ~20% MT-C34-lp$^+$/2P23-F$^+$ cells and ~80% parental CEM/R5 cells and challenged them with high-dose X4- or R5-tropic HIV-1. Parental cells died following an infection peak at day eight, whereas edited cells—generated using the CRISPR/Cas9 plasmid or RNP—survived and quickly expanded over the next six days (Fig. 5E and S4B). In summary, using a CEM/R5 cell line model, we showed that it was possible to precisely modify endogenous CXCR4 and select T cells with two fusion inhibitory peptides to render cells fully resistant to wild-type or gp41 mutant HIV-1.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

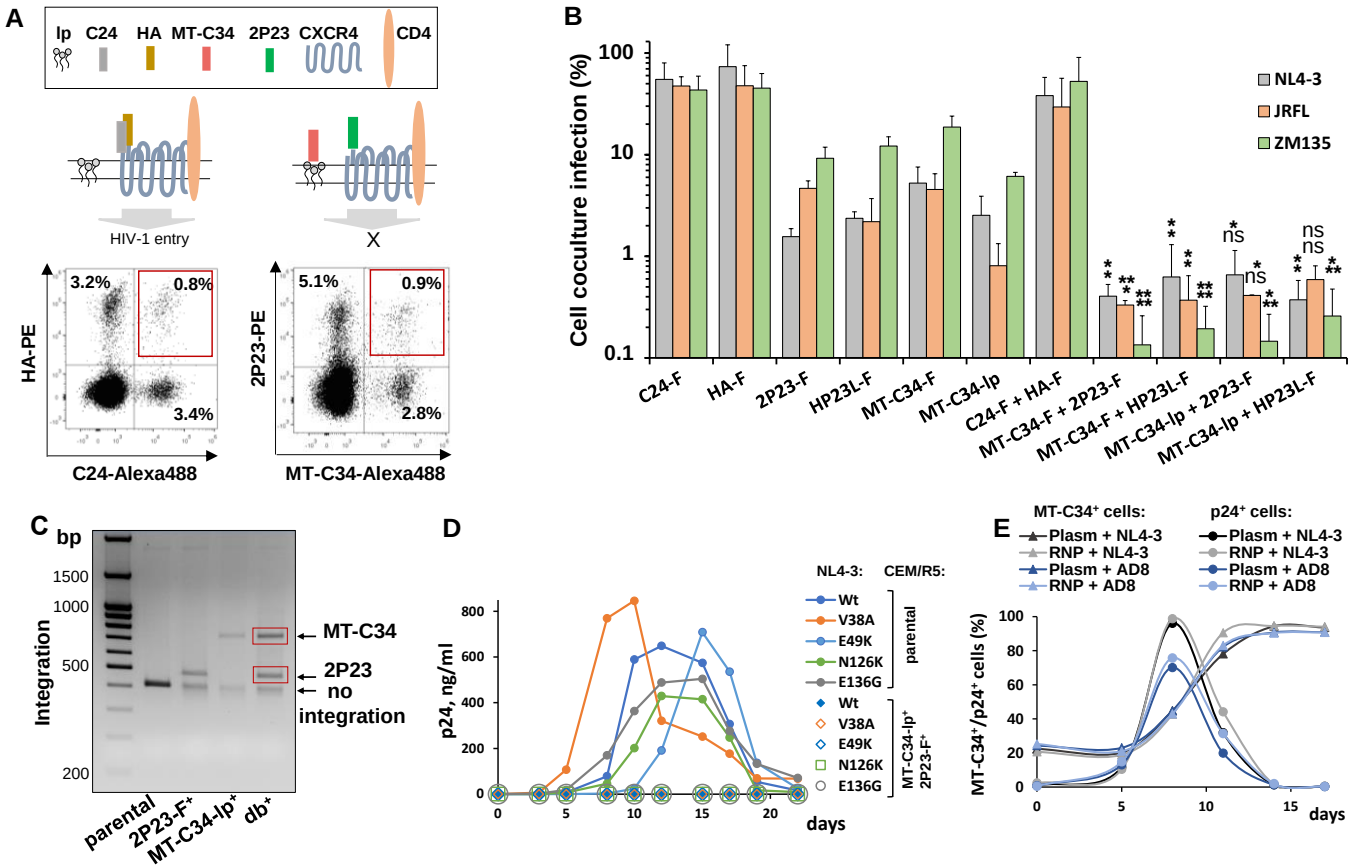

Figure 5. Exploring biallelic knock-in of two different peptides to straighten cell protection against HIV-1. A. A schematic illustrating localization of indicated peptides or epitope tags on plasma membrane after biallelic KI into *cxcr4* gene (top) and representative FACS DotPlots (bottom) recorded at day 3 after KI initiation in CEM/R5 cells. **B**. A combination of two fusion inhibitory peptide KI enhances the level of cell protection from HIV-1 infection. CEM/R5 cells with mono- or biallelic KI of indicated peptides were isolated using respective Ab staining and FACS sorting, and added as targets to transfected Raji cells to measure cell coculture infectivity levels relative to unmodified CEM/R5 control. *, **, ns indicate how significant are the differences between double and corresponding single peptide KIs at $p<0.05$, $p<0.01$, and $p>0.05$, respectively, by Student *t*-test; the results for left peptides were placed above the results for right peptides indicated in sample names. **C**. CXCR4 target locus was PCR-amplified from parental or indicated sorted CEM/R5 cells, and the resulting DNA fragments were resolved on agarose gel; DNA bands indicated in red boxes were excised and analyzed by Sanger sequencing. **D**. CEM/R5 cells with double KI of MT-C34 and 2P23 peptides are fully resistant to gp41 mutant HIV-1. Parental or modified cells (10^6) were challenged with 25 ng of wt or gp41 mutant NL4-3. The kinetics of HIV-1 replication was monitored by p24 ELISA. **E**. Survival and growth advantage of MT-C34-lp+/2P23-F+ cells during HIV-1 infection. CEM/R5 cells with biallelic KI generated using CRISPR/Cas9 plasmid (Plasm) or ribonucleoprotein (RNP) were mixed to permissive CEM/R5 cells at 1:5 ratio and challenged with a high dose of NL4-3 or AD8 HIV-1 (100 ng per 10^6 cells). Cells were double-stained for surface MT-C34 peptide and intracellular Gag to simultaneously detect infection and resistant/sensitive cell proportion using flow cytometry. Representative mages and graphs are shown in A and C-E.

Optimization of MT-C34 peptide knock-in into human primary cells

Precise genetic modification of primary cells is challenging. To accomplish this, we first optimized stimulation and transfection protocols for peripheral blood mononuclear cells (PBMC) by testing different methods of activation with PHA or anti-CD2, CD3, and CD28 Abs and Neon transfection protocols. The best results (60–70% GFP-positive cells) were obtained for cells activated with a combination of plastic-adsorbed anti-CD3 mAb and soluble anti-CD28 mAb or with CD2/CD3/CD28 beads for 2–3 days and electroporated using three pulses of 10 ms at 1,600 V. Thus, following optimization, transfection efficiencies of PBMC and CEM cells became comparable (Fig. 6A and B). A well-known way to improve HDR rate is to extend homology arms. By exploring KI of MT-C34 and HA into the *CXCR4* and *CD4* genes and the HIV-1 provirus in CEM and Raji cells, we demonstrated that extension of homology arms in a PCR donor from ~100 bp to ~500 bp increased the level of KI from 2- to 4-fold depending on the target gene. Meanwhile, the substitution of the PCR donor with a circular plasmid donor encoding 500 bp homology arms further enhanced HDR, which reached 35% at the *CXCR4* exon-2 locus (Fig. S5). After the donors were improved, we started to see MT-C34 KI into exon-2 of *CXCR4* in CD4 lymphocytes, albeit at a rate nearly 50-fold lower (~0.6%) than in CEM cells (Fig. 6C, upper row). Using cell staining with anti-CXCR4 mAbs (Fig.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

6C, low row) and a T7 endonuclease assay (Fig. 6D), we detected only a mild decrease in *CXCR4* knockout (KO) or indel formation in CD4 lymphocytes relative to those in CEM cells. Likewise, MT-C34 KI into HIV-1 proviruses with a plasmid donor was substantially lower in CD4 lymphocytes than in CEM cells (Fig. 6E and S5). These results suggest that human primary cells are prone to genetic modifications through the NHEJ mechanism of DNA repair but are resistant to HDR. It has long been known that naked transfected plasmid DNA in the cytoplasm binds to various transcription factors that transport it to the nucleus (48). We hypothesized that without a promoter our donor plasmids are poorly localized in the cell nucleus, thereby decreasing HDR. To enhance this mechanism, we added one or two X4ex2 gRNA-recognition sites (sg) (49, 50), as well as a 72-bp DNA transport signal (DTS) from the SV40 promoter (51), at the flanks of the donor sequence in the pJet-X4ex2 plasmid. These sites should have to recruit Cas9 or various transcription factors, respectively, to the DNA. Introducing sg or DTS at the 3'-position did not have a substantial effect on HDR in CD4 lymphocytes (Fig. 6F). In contrast, 5'-localization of sg, especially in combination with 3'-sg or 3'-DTS, increased the occurrence of MT-C34 KI up to 6-fold relative to KI with an unmodified plasmid donor. We then selected the best construct, pJet-sg-X4-DTS, and compared the abilities of different types of CRISPR/Cas9 to initiate KI at the *CXCR4* exon-2 locus (Fig. 6G). Cas9 RNP was slightly less efficient than transfection with gRNA and Cas9 expression plasmids. The highest level of MT-C34 KI was achieved with a single gRNA-Cas9 pX330 plasmid that contained two nuclear localization signals (NLS) instead of one NLS in the others two. Using the optimized protocol for MT-C34 KI at *CXCR4* exon-2, we generated and sorted MT-C34-negative and -positive CD4 lymphocytes. Equal numbers of lymphocytes from each group were mixed with transfected Raji cells in a 1:3 ratio to measure ZM135 HIV-1 cell coculture infectivity at day five. Contrary to MT-C34$^-$ cells, both Raji cells added as targets, and MT-C34$^+$ CD4 lymphocytes were completely resistant to R5-tropic HIV-1 coculture infection (Fig. 6H). We also showed that KI of MT-C34 into proviral DNA greatly reduced the production of HIV-1 particles by CD4 lymphocytes infected with HIV-1-GFPt (Fig. 6I).

In summary, by modifying donor DNAs and CRISPR/Cas9 we improved KI of the CD5(MT-C34)2-bglpA construct into the *CXCR4* locus in human primary CD4 lymphocytes from an undetectable level up to 3.5–4.5% and demonstrated the possibility of generating HIV-1-resistant and "HIV-1-cured" primary cells via KI of an MT-C34 fusion inhibitory peptide.

DISCUSSION

In this study, we adopted SORTS to engineer T cells with resistance to HIV-1 infection. First, we tested several short fusion inhibitory peptides from the HR2-domain of gp41, expressed on the cell surface using the shortest GPI-protein, CD52, and showed that some peptides (C34, MT-C34, and 2P23) can potently inhibit entry of HIV-1 independent of co-receptor usage. It is important to emphasize that these GPI-linked peptides were effective at protecting cells against not only cell-free viruses but also cell-to-cell infection, a dominant mechanism of HIV transmission *in vivo* (52). To demonstrate this, we employed a previously developed reporter system, which is based on the inactivation of a reporter gene with an intron and the reverse orientation of the expression cassette relative to the viral genome (38). This feature allows the reporter gene to remain silent in a viral producer (transfected) cell but become activated in a target cell after completing one cycle of replication including reverse transcription and integration steps. This type of vector allows for the precise measurement of infectivity in mixed producer–target cell cultures; background activity of the reporter gene is zero and there is no effect of active reporter protein carry-over during the measurement of cell-free infection, which is often observed with conventional vectors.

Using antibodies developed against selected gp41 peptides, we demonstrated that the degree of HIV-1 inhibition depends on the level of peptide expression on the cell surface (Fig. 2). We also noticed that if a hybrid CD52-peptide molecule has two or three sites of N-glycan attachment, its surface expression is elevated compared to a hybrid with only one N-glycosylation site. We (31) and other groups (40, 41) have previously reported similar findings that suggest that N-glycosylation of GPI-anchored proteins may stabilize their expression. However, other factors such as amino acids at N- and C-terminal junctions of a peptide with the leader and GPI-anchoring signals of CD52 can also be important and require optimization for efficient peptide expression on the cell surface.

The idea of expanding protection against HIV-1 to bystander cells by genetically modifying T cells to secrete fusion inhibitory peptides has been explored previously (41, 53, 54). In our study, we aimed to achieve the partial release of an active peptide into the medium, as a portion of the peptide should remain on the plasma membrane for SORTS-mediated cell isolation. We achieved this by introducing a furin cleavage site between the C-terminus of MT-C34 and the CD52 GPI-anchoring signal (Fig. 3). This was accompanied by a decrease in MT-C34-R cell surface expression and protection against HIV-1 that could be partially restored by furin inhibitor I. In addition, the conditioned medium from these cells did not inhibit HIV-1 infection. Based on the MT-C34 ELISA results, we concluded that this inefficient protection resulted from a low level of MT-C34 peptide in the supernatant (Fig. S1). The problem of low anti-HIV peptide secretion by genetically modified T

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

cells has been previously discussed (41); one way to overcome difficulties related to inefficient entry of a short peptide into the secretory pathway is to elongate or concatemerize these peptides.

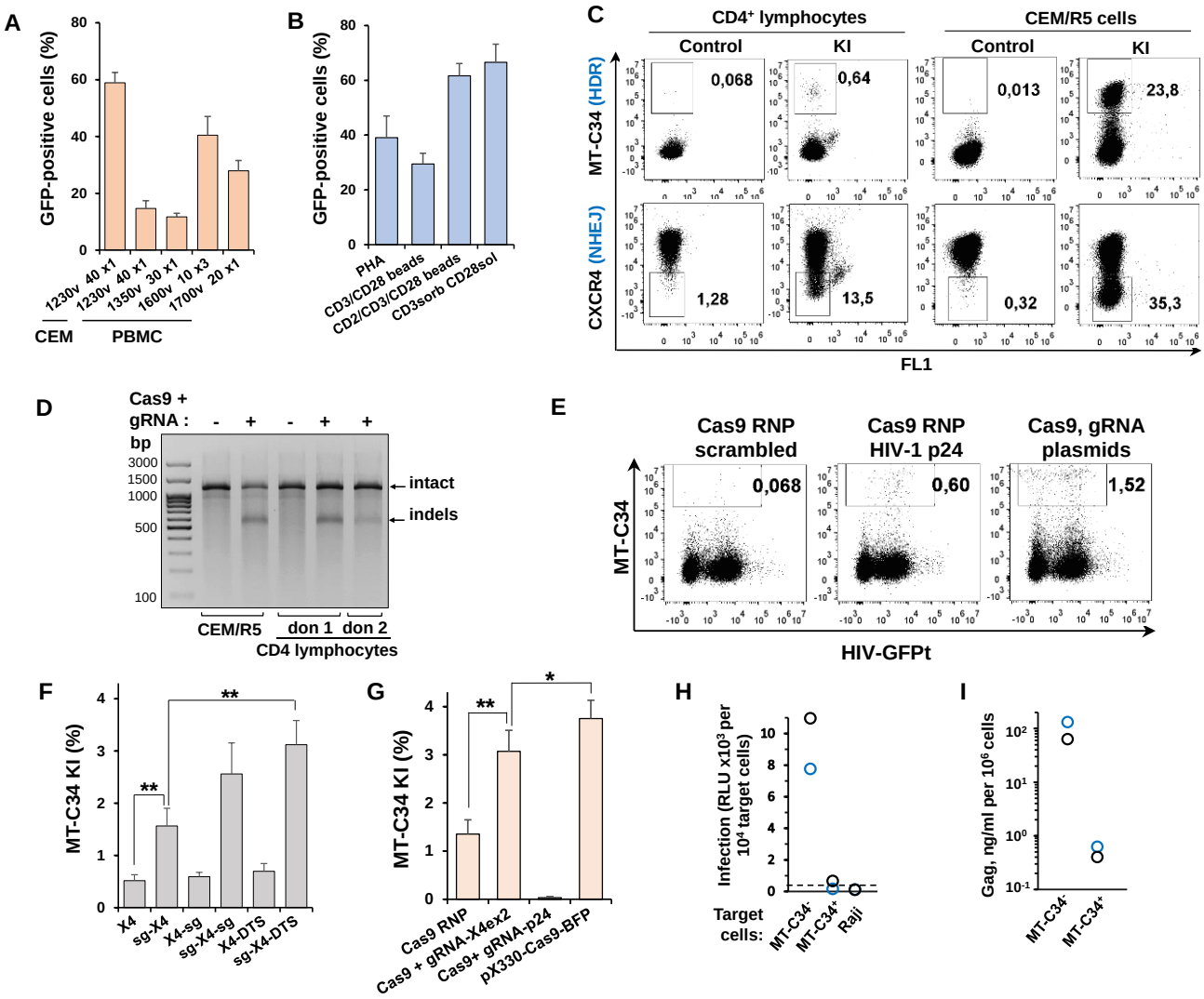

Figure 6. Evaluating efficiency of MT-C34 peptide KI into human primary cells and protection from HIV-1 infection. A. Optimization of electroporation conditions. PBMC were stimulated with 5 µg/ml PHA for 48 h and electroporated with 5 µg of GFP-expressing plasmid DNA. Transfection efficiency was measured 24 h later. **B**. Optimization of PBMC stimulation. PBMC were activated with 5 µg/ml PHA, or magnetic beads preloaded with indicated mAbs, or anti-CD3 (OKT3) mAb sorb on plastic surface with soluble 2 µg/ml anti-CD28 mAb for 48 h. Then GFP transfection efficiencies were estimated as in A. **C.** Comparison of the levels of MT-C34 KI (upper row) and CXCR4 KO (low row) in primary CD4 lymphocytes and CEM/R5 cell line. Indicated cells were co-transfected with pJet-X4 donor vector, Cas9 and gRNA expression plasmids to initiate KI into exon-2 of *CXCR4* locus. Four days later, cells were stained for MT-C34 or CXCR4 and analyzed by flow cytometry. **D.** Genetic analysis of CXCR4 exon-2 locus editing efficiency. The indicated cells were cotranscfected as described in C, but without donor. At day 3, the genomic DNAs were purified, the target locus was PCR-amplified, and the formation of indels was determined using T7 endonuclease assay. **E.** MT-C34 KI into HIV-1 provirus. CD4 lymphocytes activated for 2 d were infected with HIV-1-GFPt. Next day, cells were cotransfected with indicated CRISPR/Cas9 and pJet-HIV-1 p24 donor plasmid, and analyzed by flow cytometry at day 4 posttransfection. **F**. Donor modification with protospacer and DTS sequences enhances HDR efficiency. CD4 lymphocytes were cotransfected with Cas9 and X4ex2 gRNA expression plasmids along with one of the indicated pJet donor plasmids. **G**. Type of CRISPR/Cas9 influence the level of MT34 KI. CD4 lymphocytes were electroporated with pJet-sg-X4-DTS donor plasmid and CRISPR/Cas9 in a form of RNP, two individual plasmids or a single gRNA-Cas9 expression plasmid with two NLS. Cells in F and G were analyzed at day 3 posttransfection. The average values with standards deviations are shown in A-B, and E-F. The representative graphs are shown. **H**. MT-C34^{+} CD4 lymphocytes are resistant to HIV-1 cell coculture infection. The equal numbers of MT-C34 positive or negative sorted CD4 lymphocytes from a healthy donor or Raji cells were mixed at 1:3 ratio with Raji effector cells co-transfected with packaging, reporter and ZM135 Env expression plasmids. Cell coculture infection was measured in 5 days and normalized to cell numbers. Dashed line shows background level of luciferase activity. **I.** The levels of viral particle production by CD4 lymphocytes infected with HIV-1-GFPt followed by MT-C34 peptide KI into provirus and cell sorting. The equal number of indicated cells were incubated for 3 d before HIV-1 p24 ELISA assesement. The data obtained using two different donor cells are shown in H and I. * and ** - the differences are significant at $p<0.05$ and $p<0.01$, respectively, by Student t-test.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Third-generation lentiviral vectors have become safer and have been used to protect T cells from HIV-1 infection by transducing different antiviral genes such as modified restriction factors APOBEC3G-D128K (55) and TRIM5α-CypA (56), CCR5-interfering RNAs (57), fusion inhibitory peptides (C34-CXCR4 clinical trial), or a combination of sh5 and C46 (58). Although they can efficiently deliver relatively large genes into primary cells without producing double-strand breaks in the host's genomic DNA, several problems remain. These issues include the silencing of an integrated transgene over time, an immune response if selecting markers are included (fluorescent proteins, antibiotic resistance), and no possibility of knocking out or modifying endogenous genes, which can interfere with transgenes and reduce their therapeutic effects. In our study, we attempted CRISPR/Cas9 KI of preselected gp41 peptides into the HIV-1 entry genes *CD4* and *CXCR4*. We succeeded in targeting the *CXCR4* gene at the start codon (exon-1) or the beginning of exon-2. The first strategy allowed us to save CXCR4 expression and concomitantly obtain biallelic KI of two different fusion inhibitory peptides, one in the context of the CD52 molecule (i.e., GPI-anchored) and the other fused in frame with the N-terminus of CXCR4, as previously explored by Leslie et al. (30). We showed that the combination of these two peptides provided CEM/R5 cells with stronger protection against HIV than individual peptides. These cells were fully protected from some gp41 mutants of HIV-1 and could survive and grow during acute HIV-1 infection.

Knowing that primary cells are resistant to HDR-based gene editing, we first optimized the PBMC activation/transfection protocol to obtain 60–70% efficiency. Second, through the extension of homology arms from 100 to 500 bp and substitution of the PCR donor with a circular plasmid, we reached unusually high levels of HDR in CEM cells (~25–35%) when targeting exon-2 of CXCR4 or HIV-1 capsid (Fig. S5A and C). This can logically be explained by the more efficient nuclear entry of plasmid supercoiled DNA than of linear PCR fragments during electroporation, though some other possible explanations may exist. Søndergaard et al. reported that small plasmid cotransfection increased levels of CRISPR/Cas9 delivery and cell survival (59). Hornstein et al. showed that increasing vector length to 1.9 kb enhanced transfection, which then proceeded to plateau as plasmids further increased in size (up to 4.5 kb) (60). Nevertheless, despite fairly efficient locus editing, primary human lymphocytes remained resistant to HDR, displaying only minor enhancements of 0.4–1%. This inherited resistance may be mediated by unknown mechanisms of donor DNA sensing, interferon response, and foreign DNA nuclear efflux/degradation that are preserved in primary cells but damaged in cell lines. Our data showing that HDR in primary cells can be greatly improved via the modification of a donor plasmid to actively transport it into the nucleus using Cas9 guide sequences and SV40 DTS partially supports the above assumption.

In conclusion, we developed a gp41 fusion inhibitory peptide CRISPR/Cas9 HDR-based platform to generate T cells with high and broad resistance to HIV-1 infection. For the first time, we demonstrated the possibility of using this platform to target the *CXCR4* locus and HIV provirus in human primary CD4 lymphocytes. We believe that further improvements to HDR efficiency in primary cells are possible and, together with system adaptation to the *CCR5* locus, the described CRISPR/Cas9 peptide KI platform will become a widespread and effective instrument in future preclinical and clinical HIV trials.

METHODS

Cell cultures. The human embryonic kidney 293T cells and HeLa-TZM-blue reporter cells were obtained through NIH AIDS Research and Reference Reagent Program. The human CD4 T cell line CCRF-CEM and the B cell line Raji were purchased from ATCC. The peripheral blood mononuclear cells (PBMC) from fresh donor blood or buffy coat were isolated on the density gradient of Ficoll-Paque (Paneco, Russia). The $CD4^+$ lymphocytes were isolated using Dynabeads Untouched Kit (Invitrogen, USA). The purified PBMC and CD4 lymphocytes were activated with 5 μg/ml of phytohemaglutinin (PHA) (Sigma-Aldrich, USA), CD2/CD3/CD28 activation beads (Miltenyi Biotec, Germany) or plastic-adhered anti-CD3 mAb plus soluble anti-CD28 mAb for 2-3 days, and grown in the presence of 100 U/ml of recombinant human interleukin-2 (Ronkoleikin, Biotech, Russia). The 293T cell line was cultured in high glucose Dulbecco's modified Eagle's medium (DMEM) (Sigma-Aldrich, USA) with sodium pyruvate, sodium bicarbonate, 10% fetal calf serum (FCS), 2 mM glutamine and 40 μg/ml gentamicin. PBMCs, CEM and Raji cells were maintained in RPMI 1640 medium containing 10% fetal calf serum, 2 mM glutamine, 50 μM 2-beta-mercaptoethanol, and 40 μg/ml gentamicin. All transfections were performed on low passage cells; cultured cell lines were checked periodically for mycoplasma contamination.

Plasmid construction. To create vectors for stable expression of HIV-1 fusion inhibitory peptides, the HA epitope with or without upstream GQNDT from the CD52 coding sequence in pUCHR-mClover-smAID-P2A-CD5HA2 or pUCHR-CD5HA2-P2A-smAID-mClover lentiviral plasmid (31) was substituted with the respective peptide coding sequence by using overlapping extension PCR followed by cloning at Afe I/PspX I or Afe I/Age I restriction sites, respectively. The pUCHR-CCR5 plasmid was constructed by PCR amplification of CCR5 coding sequence from U937 cell-derived cDNA and

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

subsequent cloning the resulting DNA at EcoR I and Xma I restriction sites into pUCHR vector. The HIV-1 molecular clones pNL4-3 and pNL(AD8) were kindly provided by Dr. Eric Freed (NCI-Frederick, USA); CH077 clade B and K3016 clade C transmitted-founder viruses were obtained through NIH AIDS Research and Reference Reagent Program. For generation of NL4-3 mutants in gp41 region of Env, the BsaBI/BamH I DNA fragment from pNL4-3 plasmid was amplified by fusion PCR using overlapping primers with a mutation. It was, first, cloned into pJet plasmid (Thermo Scientific, USA), and a mutation was verified by DNA sequencing; then the fragment was cloned back into pNL4-3 using aforementioned restriction sites. The detailed information about primers and generated constructs is presented in Table S1. All PCR DNA fragments prepared for cloning were generated using Pfu polymerase (Sibenzyme, Russia) and verified by sequencing. The gRNA expressing vector pKS gRNA BB was described earlier (61). Plasmids for the expression of wild type spCas9 alone or in combination with gRNA were pcDNA3.3-hCas9 (Addgene #41815) and pX330-T2A-BFP (Addgene # 64216), respectively.

Lentiviral transduction. To generate lentiviral particles, 8×10^5 293 T cells in a 6 cm dish were cotransfected with 2 µg of HIV-1 packaging plasmid pCMVΔ8.2 R (Addgene # 12263); 0.5 µg of the pCMV-VSV-G plasmid (Addgene # 8454) expressing protein G from vesicular stomatitis virus (VSV-G); and 3 µg of one of the pUCHR transfer vector constructed as described in previous paragraph. Transfection was performed for 6 h using Lipofectamine 2000 reagent (Thermo Scientific, USA) according to the manufacturer's instruction, then the culture medium was replaced. Supernatants containing PVs were harvested and cleared through 0.45 µm pore size filters at 54 h posttransfection. The 293T/CD4 or Raji/CD4 cells that we described previously (62) were, first, infected with PVs to express CCR5. After immunostaining with respective mAb and positive sorting, cells were transduced to express one of gp41 peptides. Peptide-expressing cells were then sorted based on mClover fluorescence. Transductions were set up using different doses of PVs. Two-three days postinfection the levels of transduction were quantified by flow cytometry, and samples with less than 30% of positive events (low MOI) were selected for sorting.

Cas9, guide RNAs and donor DNAs. Cas9 protein was produced in *E.coli* BL21 (DE3) strain after transformation with SP-Cas9 plasmid (a gift from Niels Geijsen, Addgene plasmid # 62731) followed by purification on Ni-NTA agarose (Qiagen, USA), and stored at -80°C at a concentration of 6.27 µg/µl (38.93 pM/µl). Protospacer sequences for guide RNAs (gRNAs) were selected using two web-based resources https://www.benchling.com/crispr/ and http://chopchop.cbu.uib.no/ and cloned into pKS gRNA BB or pX330-T2A-BFP plasmid using BbsI restriction site. Guide RNAs were transcribed *in vitro* from PCR templates using MEGAshortscript High Yield Transcription Kit (Ambion, Thermo Fisher Scientific, USA) in accordance to manufacturer's protocol and purified from the reaction mix with miRNeasy kit (Qiagen, Germany). To prepare PCR template for an *in vitro* transcription reaction, the gRNA sequence was PCR-amplified from the respective pKS-gRNA plasmid using 5'- primer bearing T7 promoter sequence TGTAATACGACTCACTATAGGG and a specific protospacer sequence and 3'-U6 terminator primer AAAAGCACCGACTCGGTGCC. PCR products were purified and used in an *in vitro* transcription reaction at 150 nM concentration. Donor DNAs for SORTS were prepared by PCR, as described earlier (31), i.e., with primers designed to include ~100 nt homology arms in a close proximity to DNA cut site and with respective peptide expression plasmid, used as a template. PCR products were run on 1% agarose gel and purified using gel extraction Kit (Thermo Scientific). To generate donor plasmids, ~400-500 bp homology arms were amplified from genomic DNA isolated from PBMC and fused with a respective short PCR-donor DNA to get a long PCR donor. The latter was cloned into pJet plasmid and verified by sequencing. The modifications of pJet donor plasmids with protospacers and DTS were performed by cloning annealed oligos at the respective restriction sites that flank a donor sequence. The detailed information about gRNAs, primers and templates used for donor generation is presented in Table S2.

Knock-ins and SORTS. For SORTS principle and experimental protocol, please, refer to our previous paper (31). Briefly, to initiate knock-in (KI) of gp41 peptide into proviral DNA or cellular gene, $1\text{-}2 \times 10^6$ CEM, CEM-HIV-1-GFPt cells, activated human PBMC or $CD4^+$ lymphocytes were cotransfected with 3 µg of pcDNA3.3-hCas9, 1 µg of pKS-gRNA plasmid DNAs, and 2 pM of respective PCR or plasmid donor/s. Alternatively, Cas9 and gRNA expression plasmids were substituted with 4 µg of pX330-Cas9-BFP plasmid or 200 pM Cas9 RNP. For RNP preparation, Cas9 protein and gRNA were mixed at equimolar concentration in Buffer R and incubated 20-30 min before electroporation. All cell types were transfected using Neon electroporation system with 100 µl pipet tips (Invitrogen, USA) at the following settings 1,230 V, 40 ms × 1 for CEM cells, and 1,600 V, 10 ms x 3 for human primary cells. At day 3 posttransfection, cells were immunostained with respective Abs, and the level of KI as a percentage of positively labeled cells was assessed by flow cytometry. Finally, cells expressing peptide/s were isolated using several rounds of FACS sorting procedure until high level of purity was achieved.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Genetic analysis of targeted loci. Genomic DNA from ~10^6 parental or sorted CEM/R5 cells or CD4 lymphocytes was purified using Quick-gDNA MiniPrep Kit (Zymo Research, USA). The *CXCR4* target regions were PCR-amplified using 200 ng of genomic DNA, Taq DNA polymerase (SibEnzyme, Russia) and primers listed at the end of Table S2. PCR reaction was set up using the following parameters: (95 °C-2′) × 1, (95 °C-30″, 65 °C-30″, 72 °C-30″) × 30, (72 °C-5′) × 1 The DNA fragments amplified from double positive cells and containing integrated transgene were resolved on agarose gel, purified and cloned into pJet 1.2 PCR cloning vector (Thermo Scientific, USA). The DNAs from single bacterial clones were sequenced using Sanger method. The level of indels formation in PCR-amplified target locus was analyzed by T7 endonuclease assay using standard protocol.

Antibody generation and purification. Synthetic peptides C24, MT-C34, HP23L, and 2P23 (see sequences in Fig.1A) with or without an additional C-terminal cysteine residue were ordered from Syneuro company (Moscow, Russia). Each cysteine-modified peptide was conjugated to two carrier proteins, bovine serum albumin (BSA) and keyhole limpet hemocyanin (KLH) (both from Sigma-Aldrich, USA) using sulfosuccinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate reagent (ThermoFisher Scientific, USA) in accordance to manufacturer's protocol. Total four 8-week-old female BALB/c mice and two New Zealand rabbits (per antigen) were immunized three times every 2 weeks with a BSA-conjugated peptide injected intraperitoneally (mice) or subcutaneously (rabbits) in complete (first shot) and incomplete (shots 2 and 3) Freund's adjuvant. At day 35, test bleeds were performed, and sera from individual animals were tested by ELISA with KLH-conjugated peptide immobilized on plastic surface and anti-mouse or anti-rabbit horse radish peroxidase (HRP)-conjugated detector Ab (Cell Signaling Technology, USA). The animals with the highest Ab titer were selected for booster immunization. Four days after the final immunization, mice were sacrificed, and spleen lymphocytes were hybridized with the Sp2/0 myeloma cell line, and cultivated on a monolayer of feeder cells (mouse peritoneal macrophages) for about 2 weeks. The first screening of hybridoma supernatants was performed by using ELISA with KLH-conjugated peptide. The positive hybridomas were then screened on peptide-expressing 293T cells using immunofluorescence (IF) and flow cytometry. In a week after booster shot, ~ 50 ml of sera from rabbits were collected and analyzed as outlined above for mice. The selected mouse and rabbit immunoglobulins were affinity purified on protein G SepFast resins (BioToolmics, UK) or on peptide used for immunization and conjugated to a cyanogen bromide-activated-Sepharose® 4B (Sigma-Aldrich, USA), respectively, and in accordance to manufacturer's protocols.

Triton X-100 solubilization assay. The sensitivity of cellular antigens and peptides to TX-100 solubilization using flow cytometry was performed as essentially described by Filatov A. et al. (45). Briefly, $1x10^7$ cells were washed in PBS, lysed in 1 ml ice-cold 1% TX-100 in PBS supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) for 30 min at 4 °C, and fixed by diluting sample with 10 ml of ice-cold 1% paraformaldehyde (Sigma-Aldrich, USA) in PBS and incubating for additional 30 min . TX-100 extracted cell residues were collected by centrifugation at 1700 x g for 10 min, stained with primary and secondary Abs and analyzed by flow cytometry as described below. To avoid losing cellular remains, the smaller volume of reaction mixture and the higher centrifuge forces (see above) were applied than those used for live cell labeling. For better visualization of live and TX-100 extracted cells on flow cytometer, the axis scales on forward/side scatter plot were turned from linear to logarithmic.

Immunofluorescence, flow cytometry and cell sorting. Cell surface antigen immunolabeling was performed on live cells by incubating them with a primary Ab for 20 min., followed by two washes with PBS, incubation with an appropriate secondary fluorophore-conjugated Ab for 20 min and final washes with PBS. Prior addition, all Ab were diluted in PBS supplemented with 2% FCS at a final concentration of 5 µg/ml. For intracellular staining, cells were fixed with 4% paraformaldehyde (Sigma-Aldrich, USA) in PBS for 15 min, washed with PBS, and incubated for 30 min in permeabilization solution containing 0.1% saponin (Sigma-Aldrich, USA) and 2% normal mouse serum in PBS. HIV-1 Gag was stained by addition of rhodamine-labeled anti-p24 mAb at 1:200 dilution for 30-40 min. Cells were washed 3 time with permeabilization solution and resuspended in PBS prior analysis. The list of primary commercial Ab used in this study included: the rabbit anti-HA mAb clone C29F4 (Cell Signaling Technology, USA), the mouse anti-human mAbs CD3 clone OKT3, CD4 clone EM4, CD45 clone LT45, CD71 clone LT-4E3, CD98 clone MC7E7 (all from Sorbent, Moscow, Russia), CD59 clone MEM43 (Exbio, Czech Republic), CXCR4 clone 12G5 (Santa Cruz Biotechnology, Dallas, TX, USA), anti-human CD195 clone 2D7/CCR5 (BD Pharmingen™, USA), and anti-HIV-1 p24 clone KC57-RD1 (Beckman Coulter, CA, USA). The secondary goat anti-mouse and anti-rabbit Abs conjugated to Alexa488, Alexa546 or phycoerythrin (PE) were purchased from Thermo Scientific, USA. Immunolabeled samples were analyzed on CytoFLEX S (Beckman Coulter, USA) flow cytometry instrument equipped with 488-nm and 561-nm lasers, which were used to detect Alexa488 (or GFPt) and PE (or Alexa546) fluorescent signals, respectively. Cells were sorted using a FACSAria II (Becton Dickinson Biosciences, San Jose, CA, USA)

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

or Sony MA900 (Sony Biotechnology, San Jose, CA, USA) Instrument. The collected data were analyzed by CytExpert and presented using FlowJo LLC software.

HIV-1 cell-free and cell coculture infections. For HIV-1 single-round cell-coculture and cell-free infection tests with intron-regulated reporter vector inLuc, please, refer to our previous papers (38, 39, 62). Briefly, to generated viral particles pseudotyped with different Envs, 293T cells in 10-cm dish were cotransfected with 4 µg of pCMVΔ8.2 R plasmid, 6 µg of pUCHR-inLuc-mR reporter vector, and 1 µg of one of the HIV-1 Env expressing vectors pIIINL4env (a gift from Dr. Eric Freed, NCI-Frederick, USA), pJRFLenv (63) and pZM135 (a gift from Drs. E. Hunter and J. Blackwell, Emory University, GA USA). At 48 h posttransfection, supernatants containing PVs were harvested, cleared through filters with 0.45 µm pores, and the levels of HIV-1 Gag in supernatants were quantified by p24 ELISA. To initiate cell-free infection, 7.5 x 10^4 293T/CD4/R5 cells with or without peptide in a 24-well plate were infected with an equal amount of PVs (~2 µg of p24 per well) for 48 h. To set up cell-coculture infection, 10^6 Raji cells were Neon electroporated (1,350 V, 30 ms × 1) with the same viral vectors as above in the amount of 2 µg, 3 µg and 1 µg, respectively. Next day, transfected Raji cells were directly mixed with Raji/CD4/R5 targets cells with or without peptide expression at a 1:1 ratio, and cultured in 5 ml of complete growth medium for 48 h. At 16 h prior harvesting, 20 nM of phorbol 12-myristate 13-acetate (PMA) was added into co-cultures to enhance reporter expression in infected cells. The luciferase activity as a measure of infectivity was estimated using the GLO lysis buffer, Bright-GLO luciferase assay system, and GloMax®-Multi Jr Single-Tube Multimode Reader (all from Promega, Madison, WI, USA).

HIV-1 spreading assay. A panel of infectious HIV-1 (NL4-3 wt and mutants, NL(AD8) was generated by transfecting HEK293T cells with a respective molecular clone as described above. Virus input quantities were normalized for inoculation by measuring p24 levels and in some cases by determining 50% tissue culture infectious dose ($TCID_{50}$) on TZM-bl cells, as described (64). Total 10^6 CEM/R5 cells were challenged with a high dose of virus (25-100 ng of p24 or ~ 100-400 $TCID_{50}$) in 200 µl of growth medium for 5 h, then washed of the virus three times and placed in 1 ml of complete growth medium. Every 2-3 days, the 1/2-2/3 of the resuspended cell culture was removed and replaced with a fresh medium. The cells in collecFted probed were pelleted by centrifugation and where appropriate double-stained for surface MT-C34 peptide and intracellular HIV-1 p24, while supernatants were stored in a freezer for subsequent p24 ELISA measurements. Dynamics of HIV-1 replication was monitored at least for 21 days.

ELISA. The levels of HIV-1 core protein Gag in cell supernatants were quantified using HIV-1 p24 ELISA Kit (Zeptometrix, Bufalo NY, USA) in accordance to the manufacturer's instruction. For hybridoma screening, the 96-well ELISA plates (SPL, Life Sciences, Korea) were coated by a conjugate of immunogenic peptide with KLH diluted in PBS at 5 µg/ml concentration. All subsequent steps, i.e. washing, blocking, adding hybridoma supernatants, and detecting with the secondary HRP-conjugated Ab were performed using ELISA commercial solutions and diluents purchased from Xema-Medica Co (Russia). The levels of peptides in supernatants were assessed by a sandwich ELISA with a mouse monoclonal Ab against peptide acting as a capture reagent, and a rabbit polyclonal Ab raised against the same peptide together with a secondary anti-rabbit HRP-conjugated Ab used for detection purpose. Prior addition, all Abs were diluted to a final concentration of 1 µg/ml.
Data availability statement. The data that support the findings of this study, some reagents generated in this study (antibodies and plasmids) are available from the corresponding author upon reasonable request.

Ethic statement.
All experiments that involved human blood samples and laboratory animals were approved by the Human and Animal Ethics Committees of the Institute of Gene Biology (Moscow). The animals used in this study were handled in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health (USA). Blood donors gave informed consent for the use of their samples in the described experiments. All methods were performed in accordance with relevant the guidelines and regulations.

ACKNOWLEGEMENTS
We thank Alexey Deykin and Evgeniy Korshunov from Core Facility Center of the Institute of Gene Biology (Moscow) for laboratory animal handling and immunization, Olga Arkova and Artem Bonchuk from the Center for Precision Genome Editing and Genetic Technologies for Biomedicine IBG RAS for Cas9 protein purification. We are grateful to Eric Freed from NCI-Frederick for the critical reading of the manuscript. This work was mostly supported by the grant 18-29-07052 from Russian Foundation for Basic Research and partially (studies on primary cells) by the grant 075-15-2019-1661 from the Ministry of Science and Higher Education of the Russian Federation.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

AUTHOR CONTRIBUTIONS

D.M. – concept development, construct design, data analysis, antibody generation, writing manuscript; A.M. – lentiviral plasmid construction, cell transduction, single-round infection tests, immunofluorescent analysis; N.K. – primer design, donor DNA preparation, primary cell activation and transfections, knock-in, FACS analysis; S.K. – hybridoma generation and screening, antibody purification, genetic analysis; D.K. – cell sorting, ELISA; M.S. – guide RNA and Cas9 RNP preparation, KI analysis, editing manuscript; D.G. – TX-100 antigen solubilization, immunofluorescence, flow cytometry analysis; A.S and A.V. – HIV spreading assay, virus titration, data analysis, editing manuscript; A.F. – monoclonal hybridoma generation, data analysis.

COMPETING INTERESTS

The authors declare no competing financial and non-financial interests.

MATERIALS & CORRESPONDENCE

All correspondence and material requests should be addressed to dvmazurov@yandex.ru

REFERENCES

1. Jessen H, Allen TM, Streeck H. 2014. How a single patient influenced HIV research--15-year follow-up. N Engl J Med 370:682-3.
2. Allers K, Hutter G, Hofmann J, Loddenkemper C, Rieger K, Thiel E, Schneider T. 2011. Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation. Blood 117:2791-9.
3. Henrich TJ, Hanhauser E, Marty FM, Sirignano MN, Keating S, Lee TH, Robles YP, Davis BT, Li JZ, Heisey A, Hill AL, Busch MP, Armand P, Soiffer RJ, Altfeld M, Kuritzkes DR. 2014. Antiretroviral-free HIV-1 remission and viral rebound after allogeneic stem cell transplantation: report of 2 cases. Ann Intern Med 161:319-27.
4. Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, Spratt SK, Surosky RT, Giedlin MA, Nichol G, Holmes MC, Gregory PD, Ando DG, Kalos M, Collman RG, Binder-Scholl G, Plesa G, Hwang WT, Levine BL, June CH. 2014. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med 370:901-10.
5. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. Science 339:823-6.
6. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339:819-23.
7. Hu W, Kaminski R, Yang F, Zhang Y, Cosentino L, Li F, Luo B, Alvarez-Carbonell D, Garcia-Mesa Y, Karn J, Mo X, Khalili K. 2014. RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. Proc Natl Acad Sci U S A 111:11461-6.
8. Kaminski R, Chen Y, Fischer T, Tedaldi E, Napoli A, Zhang Y, Karn J, Hu W, Khalili K. 2016. Elimination of HIV-1 Genomes from Human T-lymphoid Cells by CRISPR/Cas9 Gene Editing. Sci Rep 6:22555.
9. Wang G, Zhao N, Berkhout B, Das AT. 2016. CRISPR-Cas9 Can Inhibit HIV-1 Replication but NHEJ Repair Facilitates Virus Escape. Mol Ther 24:522-6.
10. Liao HK, Gu Y, Diaz A, Marlett J, Takahashi Y, Li M, Suzuki K, Xu R, Hishida T, Chang CJ, Esteban CR, Young J, Izpisua Belmonte JC. 2015. Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. Nat Commun 6:6413.
11. Ye L, Wang J, Beyer AI, Teque F, Cradick TJ, Qi Z, Chang JC, Bao G, Muench MO, Yu J, Levy JA, Kan YW. 2014. Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5Δ32 mutation confers resistance to HIV infection. Proc Natl Acad Sci U S A 111:9591-6.
12. Liu S, Wang Q, Yu X, Li Y, Guo Y, Liu Z, Sun F, Hou W, Li C, Wu L, Guo D, Chen S. 2018. HIV-1 inhibition in cells with CXCR4 mutant genome created by CRISPR-Cas9 and piggyBac recombinant technologies. Sci Rep 8:8573.
13. Yu S, Yao Y, Xiao H, Li J, Liu Q, Yang Y, Adah D, Lu J, Zhao S, Qin L, Chen X. 2018. Simultaneous Knockout of CXCR4 and CCR5 Genes in CD4+ T Cells via CRISPR/Cas9 Confers Resistance to Both X4- and R5-Tropic Human Immunodeficiency Virus Type 1 Infection. Hum Gene Ther 29:51-67.
14. Wang Z, Pan Q, Gendron P, Zhu W, Guo F, Cen S, Wainberg MA, Liang C. 2016. CRISPR/Cas9-Derived Mutations Both Inhibit HIV-1 Replication and Accelerate Viral Escape. Cell Rep 15:481-489.
15. Yoder KE, Bundschuh R. 2016. Host Double Strand Break Repair Generates HIV-1 Strains Resistant to CRISPR/Cas9. Sci Rep 6:29530.
16. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD. 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 31:822-6.
17. Kosicki M, Tomberg K, Bradley A. 2018. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. Nat Biotechnol 36:765-771.
18. Mao Z, Bozzella M, Seluanov A, Gorbunova V. 2008. Comparison of nonhomologous end joining and homologous recombination in human cells. DNA Repair (Amst) 7:1765-71.
19. Miyaoka Y, Berman JR, Cooper SB, Mayerl SJ, Chan AH, Zhang B, Karlin-Neumann GA, Conklin BR. 2016. Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. Sci Rep 6:23549.
20. De Feo CJ, Weiss CD. 2012. Escape from human immunodeficiency virus type 1 (HIV-1) entry inhibitors. Viruses 4:3859-911.
21. Chong H, Yao X, Qiu Z, Qin B, Han R, Waltersperger S, Wang M, Cui S, He Y. 2012. Discovery of critical residues for viral entry and inhibition through structural Insight of HIV-1 fusion inhibitor CP621-652. J Biol Chem 287:20281-9.
22. Chong H, Yao X, Sun J, Qiu Z, Zhang M, Waltersperger S, Wang M, Cui S, He Y. 2012. The M-T hook structure is critical for design of HIV-1 fusion inhibitors. J Biol Chem 287:34558-68.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

23. Dwyer JJ, Wilson KL, Davison DK, Freel SA, Seedorff JE, Wring SA, Tvermoes NA, Matthews TJ, Greenberg ML, Delmedico MK. 2007. Design of helical, oligomeric HIV-1 fusion inhibitor peptides with potent activity against enfuvirtide-resistant virus. Proc Natl Acad Sci U S A 104:12772-7.
24. Zhang X, Zhu Y, Hu H, Zhang S, Wang P, Chong H, He J, Wang X, He Y. 2018. Structural Insights into the Mechanisms of Action of Short-Peptide HIV-1 Fusion Inhibitors Targeting the Gp41 Pocket. Front Cell Infect Microbiol 8:51.
25. Xiong S, Borrego P, Ding X, Zhu Y, Martins A, Chong H, Taveira N, He Y. 2017. A Helical Short-Peptide Fusion Inhibitor with Highly Potent Activity against Human Immunodeficiency Virus Type 1 (HIV-1), HIV-2, and Simian Immunodeficiency Virus. J Virol 91.
26. Egelhofer M, Brandenburg G, Martinius H, Schult-Dietrich P, Melikyan G, Kunert R, Baum C, Choi I, Alexandrov A, von Laer D. 2004. Inhibition of human immunodeficiency virus type 1 entry in cells expressing gp41-derived peptides. J Virol 78:568-75.
27. Hildinger M, Dittmar MT, Schult-Dietrich P, Fehse B, Schnierle BS, Thaler S, Stiegler G, Welker R, von Laer D. 2001. Membrane-anchored peptide inhibits human immunodeficiency virus entry. J Virol 75:3038-42.
28. Liu L, Wen M, Zhu Q, Kimata JT, Zhou P. 2016. Glycosyl Phosphatidylinositol-Anchored C34 Peptide Derived From Human Immunodeficiency Virus Type 1 Gp41 Is a Potent Entry Inhibitor. J Neuroimmune Pharmacol 11:601-10.
29. Tang X, Jin H, Chen Y, Li L, Zhu Y, Chong H, He Y. 2019. A Membrane-Anchored Short-Peptide Fusion Inhibitor Fully Protects Target Cells from Infections of Human Immunodeficiency Virus Type 1 (HIV-1), HIV-2, and Simian Immunodeficiency Virus. J Virol 93.
30. Leslie GJ, Wang J, Richardson MW, Haggarty BS, Hua KL, Duong J, Secreto AJ, Jordon AP, Romano J, Kumar KE, DeClercq JJ, Gregory PD, June CH, Root MJ, Riley JL, Holmes MC, Hoxie JA. 2016. Potent and Broad Inhibition of HIV-1 by a Peptide from the gp41 Heptad Repeat-2 Domain Conjugated to the CXCR4 Amino Terminus. PLoS Pathog 12:e1005983.
31. Zotova A, Pichugin A, Atemasova A, Knyazhanskaya E, Lopatukhina E, Mitkin N, Holmuhamedov E, Gottikh M, Kuprash D, Filatov A, Mazurov D. 2019. Isolation of gene-edited cells via knock-in of short glycophosphatidylinositol-anchored epitope tags. Sci Rep 9:3132.
32. Zhu Y, Ding X, Yu D, Chong H, He Y. 2019. The Tryptophan-Rich Motif of HIV-1 gp41 Can Interact with the N-Terminal Deep Pocket Site: New Insights into the Structure and Function of gp41 and Its Inhibitors. J Virol 94.
33. Chong H, Yao X, Qiu Z, Sun J, Qiao Y, Zhang M, Wang M, Cui S, He Y. 2014. The M-T hook structure increases the potency of HIV-1 fusion inhibitor sifuvirtide and overcomes drug resistance. J Antimicrob Chemother 69:2759-69.
34. Chong H, Zhu Y, Yu D, He Y. 2018. Structural and Functional Characterization of Membrane Fusion Inhibitors with Extremely Potent Activity against Human Immunodeficiency Virus Type 1 (HIV-1), HIV-2, and Simian Immunodeficiency Virus. J Virol 92.
35. Su S, Zhu Y, Ye S, Qi Q, Xia S, Ma Z, Yu F, Wang Q, Zhang R, Jiang S, Lu L. 2017. Creating an Artificial Tail Anchor as a Novel Strategy To Enhance the Potency of Peptide-Based HIV Fusion Inhibitors. J Virol 91.
36. Chong H, Qiu Z, Su Y, Yang L, He Y. 2015. Design of a highly potent HIV-1 fusion inhibitor targeting the gp41 pocket. Aids 29:13-21.
37. Chong H, Wu X, Su Y, He Y. 2016. Development of potent and long-acting HIV-1 fusion inhibitors. Aids 30:1187-96.
38. Mazurov D, Ilinskaya A, Heidecker G, Lloyd P, Derse D. 2010. Quantitative comparison of HTLV-1 and HIV-1 cell-to-cell infection with new replication dependent vectors. PLoS Pathog 6:e1000788.
39. Shunaeva A, Potashnikova D, Pichugin A, Mishina A, Filatov A, Nikolaitchik O, Hu WS, Mazurov D. 2015. Improvement of HIV-1 and Human T Cell Lymphotropic Virus Type 1 Replication-Dependent Vectors via Optimization of Reporter Gene Reconstitution and Modification with Intronic Short Hairpin RNA. J Virol 89:10591-601.
40. Liu Y, Nguyen A, Wolfert RL, Zhuo S. 2009. Enhancing the secretion of recombinant proteins by engineering N-glycosylation sites. Biotechnol Prog 25:1468-75.
41. Egerer L, Volk A, Kahle J, Kimpel J, Brauer F, Hermann FG, von Laer D. 2011. Secreted antiviral entry inhibitory (SAVE) peptides for gene therapy of HIV infection. Mol Ther 19:1236-44.
42. Matabaro E, He Z, Liu YS, Zhang HJ, Gao XD, Fujita M. 2017. Molecular switching system using glycosylphosphatidylinositol to select cells highly expressing recombinant proteins. Sci Rep 7:4033.
43. Chertov O, Zhang N, Chen X, Oppenheim JJ, Lubkowski J, McGrath C, Sowder RC, 2nd, Crise BJ, Malyguine A, Kutzler MA, Steele AD, Henderson EE, Rogers TJ. 2011. Novel peptides based on HIV-1 gp120 sequence with homology to chemokines inhibit HIV infection in cell culture. PLoS One 6:e14474.
44. Zhou Y, Wang J, Zhou I, Lou H, Li CZ, Chen ZR, Zhang ZH, Liu S, Wu S, Tan W, Jiang S, Zhou C. 2013. Simultaneous expression of displayed and secreted antibodies for antibody screen. PLoS One 8:e80005.
45. Filatov AV, Shmigol IB, Kuzin, II, Sharonov GV, Feofanov AV. 2003. Resistance of cellular membrane antigens to solubilization with Triton X-100 as a marker of their association with lipid rafts--analysis by flow cytometry. J Immunol Methods 278:211-9.
46. Su Y, Chong H, Xiong S, Qiao Y, Qiu Z, He Y. 2015. Genetic Pathway of HIV-1 Resistance to Novel Fusion Inhibitors Targeting the Gp41 Pocket. J Virol 89:12467-79.
47. Yu D, Ding X, Liu Z, Wu X, Zhu Y, Wei H, Chong H, Cui S, He Y. 2018. Molecular mechanism of HIV-1 resistance to sifuvirtide, a clinical trial-approved membrane fusion inhibitor. J Biol Chem 293:12703-12718.
48. Lam AP, Dean DA. 2010. Progress and prospects: nuclear import of nonviral vectors. Gene Ther 17:439-47.
49. Zhang JP, Li XL, Li GH, Chen W, Arakaki C, Botimer GD, Baylink D, Zhang L, Wen W, Fu YW, Xu J, Chun N, Yuan W, Cheng T, Zhang XB. 2017. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. Genome Biol 18:35.
50. Nguyen DN, Roth TL, Li PJ, Chen PA, Apathy R, Mamedov MR, Vo LT, Tobin VR, Goodman D, Shifrut E, Bluestone JA, Puck JM, Szoka FC, Marson A. 2020. Polymer-stabilized Cas9 nanoparticles and modified repair templates increase genome editing efficiency. Nat Biotechnol 38:44-49.
51. Dean DA, Dean BS, Muller S, Smith LC. 1999. Sequence requirements for plasmid nuclear import. Exp Cell Res 253:713-22.
52. Murooka TT, Deruaz M, Marangoni F, Vrbanac VD, Seung E, von Andrian UH, Tager AM, Luster AD, Mempel TR. 2012. HIV-infected T cells are migratory vehicles for viral dissemination. Nature 490:283-7.
53. Falkenhagen A, Singh J, Asad S, Leontyev D, Read S, Zúñiga-Pflücker JC, Joshi S. 2017. Control of HIV Infection In Vivo Using Gene Therapy with a Secreted Entry Inhibitor. Mol Ther Nucleic Acids 9:132-144.
54. Kuhlmann AS, Haworth KG, Barber-Axthelm IM, Ironside C, Giese MA, Peterson CW, Kiem HP. 2019. Long-Term Persistence of Anti-HIV Broadly Neutralizing Antibody-Secreting Hematopoietic Cells in Humanized Mice. Mol Ther 27:164-177.
55. Delviks-Frankenberry KA, Ackerman D, Timberlake ND, Hamscher M, Nikolaitchik OA, Hu WS, Torbett BE, Pathak VK. 2019. Development of Lentiviral Vectors for HIV-1 Gene Therapy with Vif-Resistant APOBEC3G. Mol Ther Nucleic Acids 18:1023-1038.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

56. Neagu MR, Ziegler P, Pertel T, Strambio-De-Castillia C, Grutter C, Martinetti G, Mazzucchelli L, Grutter M, Manz MG, Luban J. 2009. Potent inhibition of HIV-1 by TRIM5-cyclophilin fusion proteins engineered from human components. J Clin Invest 119:3035-47.
57. Urusov F, Glazkova D, Omelchenko D, Bogoslovskaya E, Tsyganova G, Kersting K, Shipulin G, Pokrovsky V. 2018. Optimization of Polycistronic Anti-CCR5 Artificial microRNA Leads to Improved Accuracy of Its Lentiviral Vector Transfer and More Potent Inhibition of HIV-1 in CD4(+) T-Cells. Cells 7.
58. Wolstein O, Boyd M, Millington M, Impey H, Boyer J, Howe A, Delebecque F, Cornetta K, Rothe M, Baum C, Nicolson T, Koldej R, Zhang J, Keech N, Camba Colón J, Breton L, Bartlett J, An DS, Chen IS, Burke B, Symonds GP. 2014. Preclinical safety and efficacy of an anti-HIV-1 lentiviral vector containing a short hairpin RNA to CCR5 and the C46 fusion inhibitor. Mol Ther Methods Clin Dev 1:11.
59. Søndergaard JN, Geng K, Sommerauer C, Atanasoai I, Yin X, Kutter C. 2020. Successful delivery of large-size CRISPR/Cas9 vectors in hard-to-transfect human cells using small plasmids. Commun Biol 3:319.
60. Hornstein BD, Roman D, Arévalo-Soliz LM, Engevik MA, Zechiedrich L. 2016. Effects of Circular DNA Length on Transfection Efficiency by Electroporation into HeLa Cells. PLoS One 11:e0167537.
61. Tarasevich A, Filatov A, Pichugin A, Mazurov D. 2015. Monoclonal antibody profiling of cell surface proteins associated with the viral biofilms on HTLV-1 transformed cells. Acta Virol 59:247-56.
62. Zotova A, Atemasova A, Pichugin A, Filatov A, Mazurov D. 2019. Distinct Requirements for HIV-1 Accessory Proteins during Cell Coculture and Cell-Free Infection. Viruses 11.
63. Sato K, Aoki J, Misawa N, Daikoku E, Sano K, Tanaka Y, Koyanagi Y. 2008. Modulation of human immunodeficiency virus type 1 infectivity through incorporation of tetraspanin proteins. J Virol 82:1021-33.
64. Kimpton J, Emerman M. 1992. Detection of replication-competent and pseudotyped human immunodeficiency virus with a sensitive cell line on the basis of activation of an integrated beta-galactosidase gene. J Virol 66:2232-9.

SUPPLEMENTARY INFORMATION

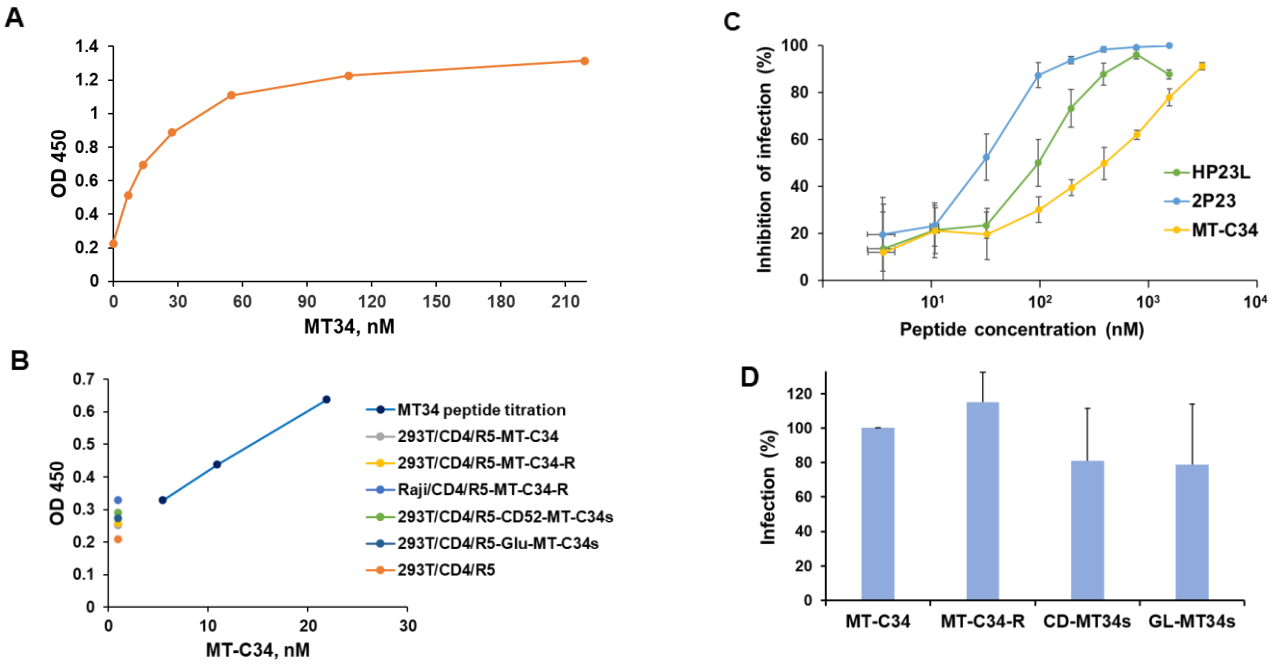

Figure S1. Secreted or shed forms of MT-C34 peptide are not detected by sandwich ELISA and do not inhibit HIV-1 infection. **A**. Determining sensitivity of developed MT-C34 ELISA with a pair of rabbit and mouse anti-MT-C34 Abs. The synthetic peptide MT-C34 was titrated and assayed by ELISA. **B**. The levels of MT-C34 peptide in cell supernatants are below the detection limit of the established sandwich ELISA. The indicated cells were passaged and cultured for 48 h, then the supernatants were harvested and assayed by ELISA without dilution. In parallel, the control MT-C34 peptide at several dilutions was evaluated in the same plate. **C**. Inhibitory activities of synthetic peptides against HIV-1_{JRFL}. VLPs psedotyped with JRFL Env were preincubated with one of the indicated peptides for 1h and the levels of infectivity were measured on 293T/CD4/R5 cells as described in Methods. The results are presented as an inhibitory activity relative to that with no peptide. **D**. Supernatants from 293T cells engineered to secrete or shed MT-C34 peptide do not inhibit HIV-1 infection. VLPs with JRFL Env were preincubated for 1 h with 0.25 ml of the supernatant derived from one of the indicated modified 293T cells and added to 293T/CD4/R5 cells cultured in 24-well plate in 0.25 ml of growth medium. The levels of infectivity were quantified as in C and presented relative to MT-C34 sample that does not contain released peptide. The graphs in A and B are representative of two separate experiments. The average values of at least three independent experiments with standard deviations are shown in C and D

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

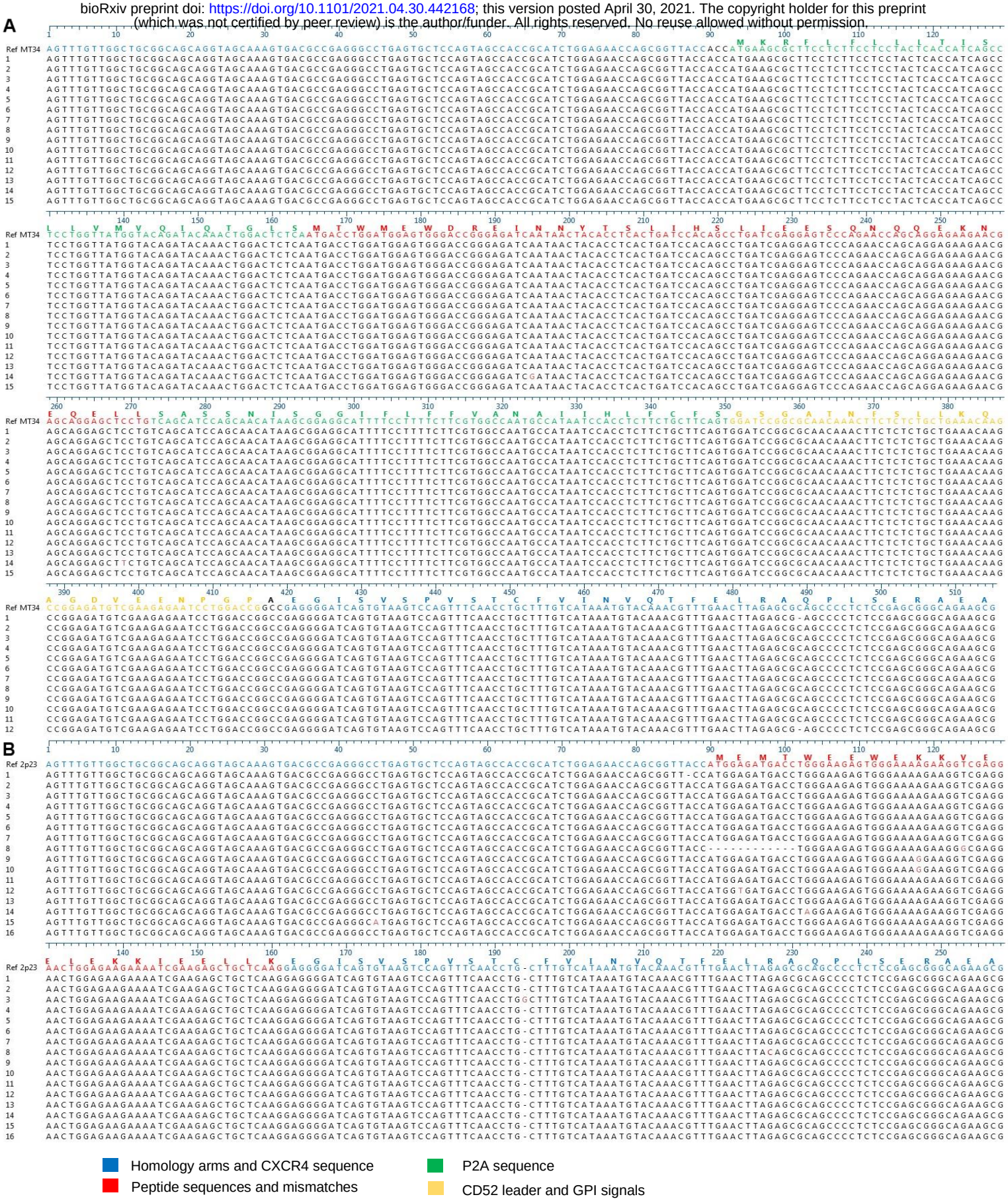

Figure S2. Sequence analysis of CXCR4-integrated DNA products PCR-amplified from MT-C34-lp^{+}/2P23-F^{+} CEM/R5 cells. PCR-products highlighted by red boxes (Fig.5C) were gel excised, purified, cloned into pJet1.2 vector and analyzed by Sanger sequencing after E.coli transformation and selection of individual bacterial clones.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

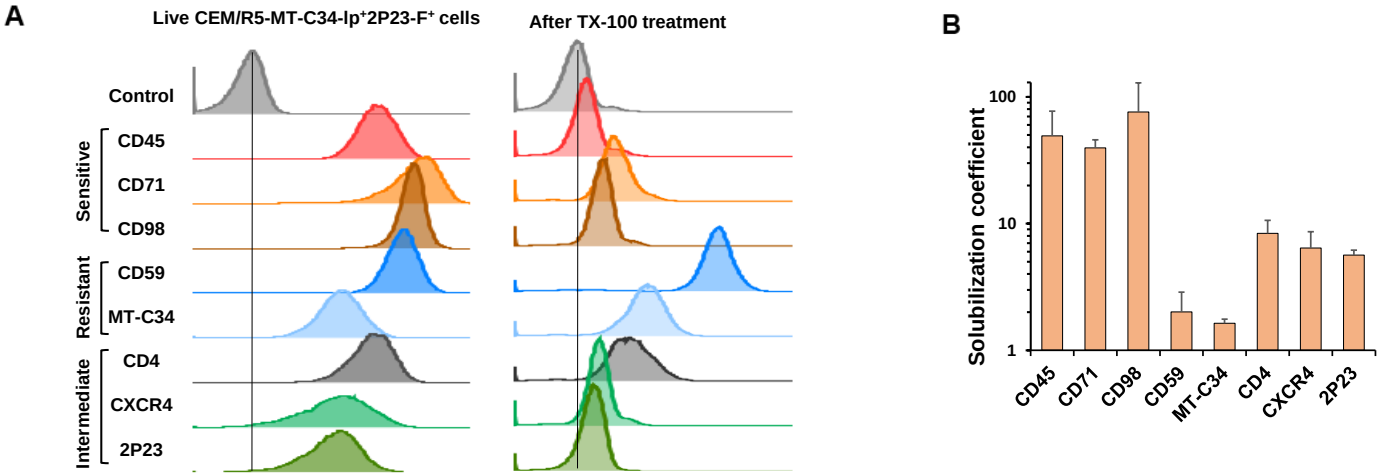

Figure S3. Different sensitivity of lipid raft-associated MT-C34 and 2P23 fused with CXCR4 to Triton X-100 solubilization. A. CEM/R5-MT-C34-lp$^+$2P23-F$^+$ cells were either surface stained with Abs against indicated antigens and the secondary Alexa-488-conjugated Ab (on the left) or treated, first, with TX-100/paraformaldehyde as outlined in Method section, and then the resulting cellular remains were immunostained likewise (on the right). Samples were analyzed by flow cytometry and presented as overlayed histograms with partial off-set. **B.** Mean fluorescence intensities (MFI) measured for each antigen were first adjusted to respective control (without primary Ab). Solubilization coefficients were calculated by dividing adjusted MFIs prior TX-100 treatment to those obtained after TX-100 extraction. Representative histograms and average values with standard deviations obtained from three independent experiments are shown. Non-raft antigens CD45, CD71, and CD98 were easily solubilized by TX-100. The GPI-anchored CD59 and MT-C34 were resistant to detergent treatment, while 2P23 covalently linked to CXCR4 and CD4 displayed intermediate sensitivity to TX-100 solubilization.

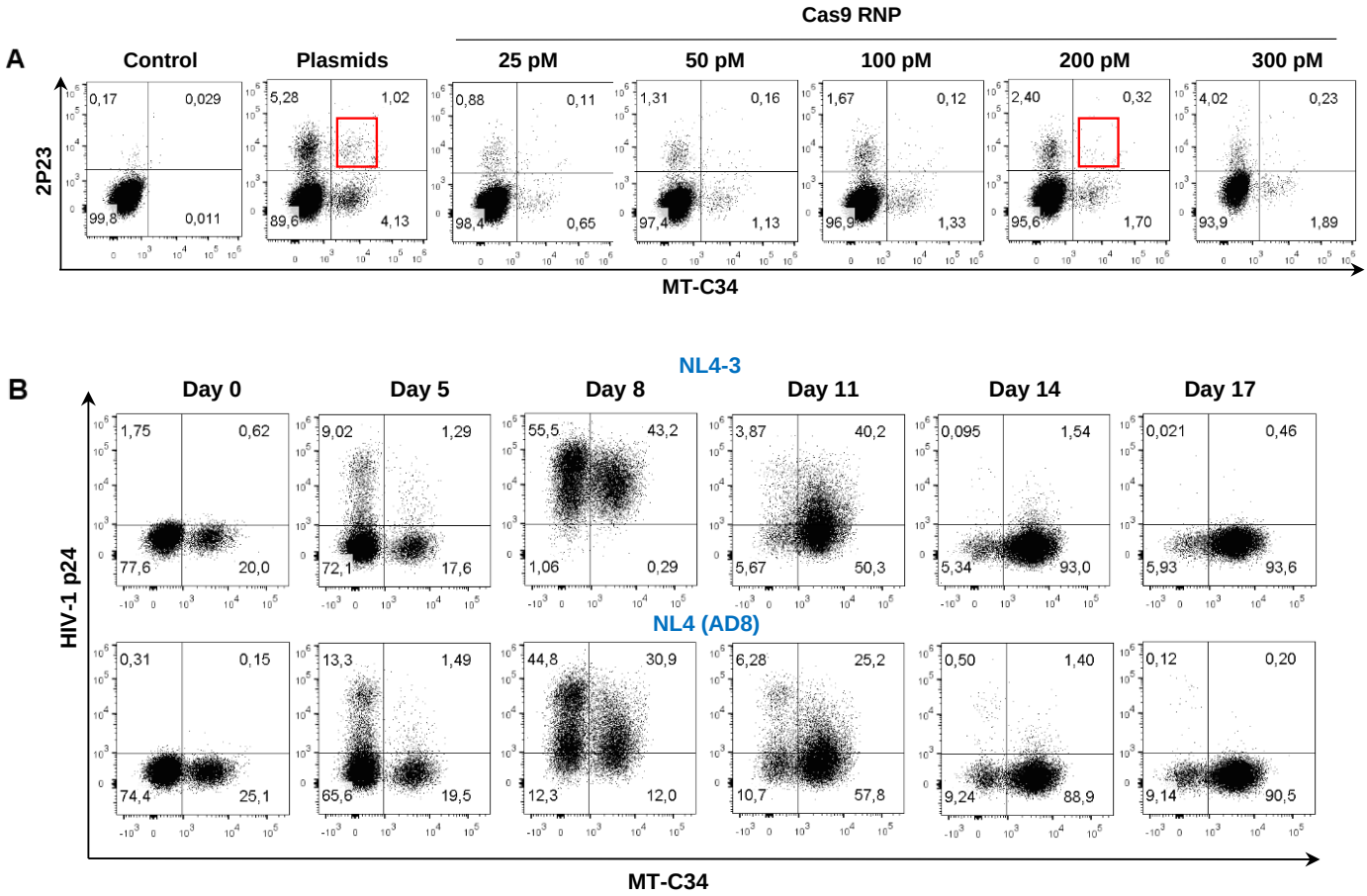

Figure S4. Survival and selective expansion of cells with MT-C34 and 2P23 peptide KI during HIV-1 infection. A. Knock-in optimization with Cas9 ribonucleoprotein (RNP) complex. CEM/R5 cells were co-transfected using 1 µg of 2P23-F, 1 µg of MT-C34-lp PCR-donors, with either gRNA (1 µg) and Cas9 (3 µg) expression plasmids or Cas9 RNP. Double-positive cells (in red boxes) were sorted and used for further analysis. **B.** A culture mixture consisting of 2x10^5 CEM/R5-MT-C34-lp$^+$/2P23-F$^+$ cells generated using Cas9 RNP and 8x10^5 CEM/R5 cells was challenged with 100 ng of X4-tropic NL4-3 or R5-tropic AD8 strain of HIV-1. The levels of infection spread and the ratios of two cell populations were monitored by flow cytometry after p24 and MT-C34 double-staining. DotPlots from a typical experiment are shown.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

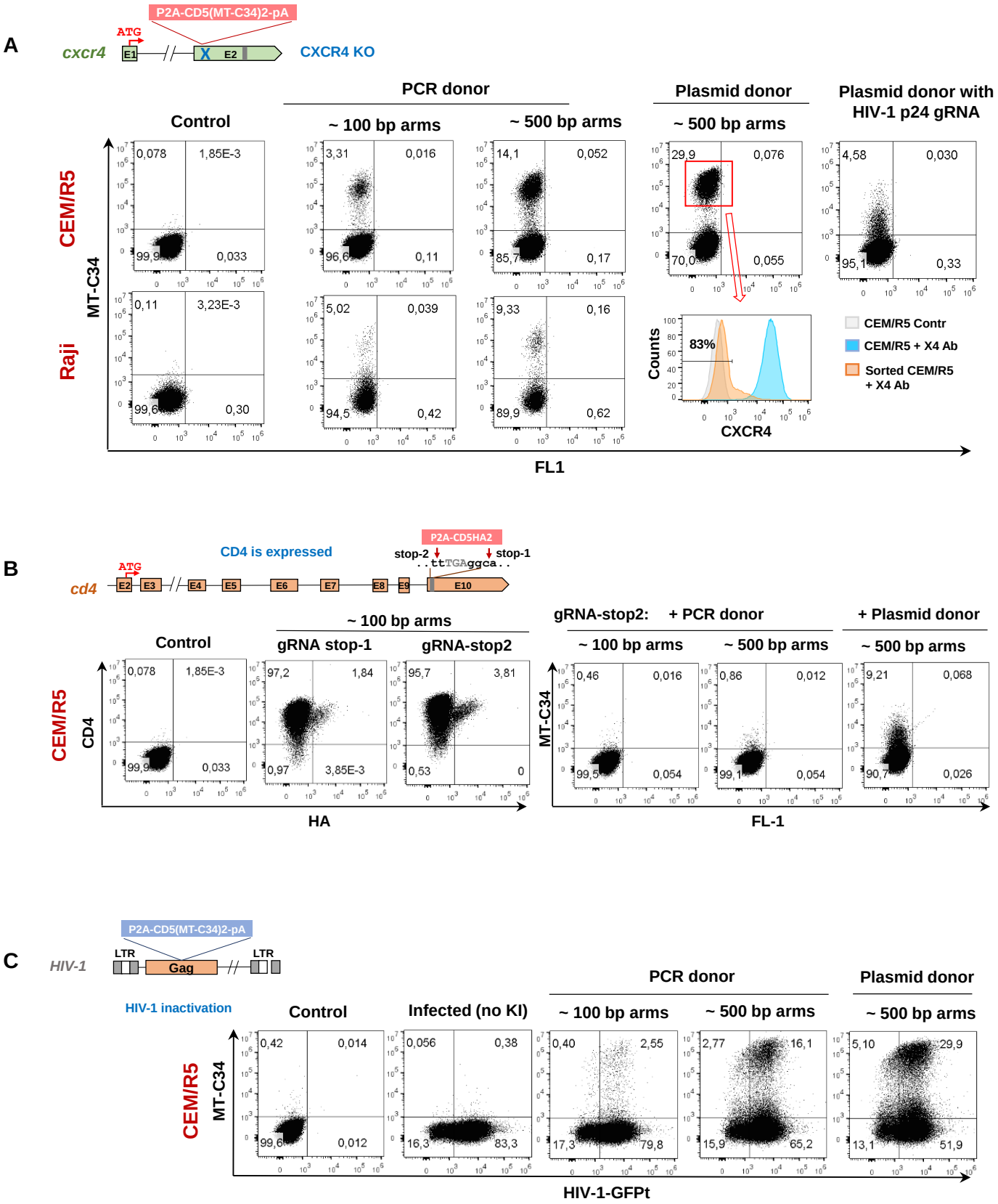

Figure S5. Knock-in efficiency can be improved by homology arm extension, type of DNA donor and gRNA selection. Cells were transfected with 2 pM of PCR or plasmid DNA donor, 1 µg of respective gRNA and 3 µg of Cas9 expression plasmids to initiate KI at loci indicated schematically next to the plotted results (**A** – CXCR4, **B** – CD4 and **C** – HIV-1 integrated DNA). At day 3 posttransfection, cells were immunostained and analyzed by flow cytometry. The levels of targeted gene expression were estimated either at the same time of KI evaluation or followed by positive cell sorting selection and expansion (see histogram). Images are representative of at least two independent experiments.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table S1. Primers used for plasmid construction.

Construct name	Primer name	Primer sequence	Templates
Generation of *pUCHR-mClover-CD5-peptide-2* constructs using fusion PCR and Afe I/PspX I restriction sites			
Flanking primers	5'-mClover	CTGTACAAGCCAGCCAAACCTCCGG	PCR products 1 + 2
	3'-PspXI	CATTGGTCTTAAAGGTACTCGAG	
C24	5'-C24	GATCGAGGAGTCCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTCCTGTCAGCATCCAGCAACATAAG	pUCHR-mClover-CD5HA2
	3'-C24		
C34	5'-C34	TACACCTCACTGATCCACAG	pUCHR-mClover-C24
	3'-C34	CTGTGGATCAGTGAGGTGTAGTTATTGATCTCCCGGTCCCACTCCATCCAGGTGTCGTTTTGTCCTGAG	
MT-C34	5'-MT34	GGACTCTCAATGACCTGGATGGAGTGGGACCG	pUCHR-mClover-C34
	3'-MT34	CCAGGTCATTGAGAGTCCAGTTTGTATCTG	
HP23L	5'-HP23L	GGAAAAGAAGATCGAGGAATACACCAAGAAAATCGAAGAGATCCTCAAGTCAGCATCCAGCAACATAAG	pUCHR-mClover-CD5HA2
	3'-HP23L	GTATTCCTCGATCTTCTTTTCCCACTCTTCCCAGGTCAGCTCTGAGAGTCCAGTTTGTATCTG	
2P23	5'-2P23	GAAAAGAAGGTCGAGGAACTGGAGAAGAAAATCGAAGAGCTGCTCAAGTCAGCATCCAGCAACATAAG	pUCHR-mClover-CD5HA2
	3'-2P23	CAGTTCCTCGACCTTCTTTTCCCACTCTTCCCAGGTCATCTCTGAGAGTCCAGTTTGTATCTG	
MT-WQ-IDL	5'-WQ-IDL	CACCAAAAAGATCGAAGAGCTCATCAAGAAATCCCAGAACCAGCAGATCGACCTGTCAGCATCCAGCAACATAAG	pUCHR-mClover-CD5HA2
	3'-WQ-IDL	GCTCTTCGATCTTTTTGGTGTATTCCTCGATCTTCTTGTCCCACTCTTCCCAGGTCATTGAGAGTCCAGTTTGTATCTG	
P52	5'-P52	CCGAAGAGCAGCAGAAAAAGAACGAGGAAGAGCTGAAGAAGCTGGAGAAGTCAGCATCCAGCAACATAAG	pUCHR-mClover-CD5HA2
	3'-P52	CTTTTTCTGCTGCTCTTCGGCTTTCTTGAGCAGTTCCTCGATCTTCTGCTCCCATGAGAGTCCAGTTTGTATCTG	
MT-C34-R	5'-34furin	CTGCGGATCAGGCGCTCAGCATCCAGCAACATAAG	pUCHR-mClover-MT-C34
	3'-34furin	GAGCGCCTGATCCGCAGGAGCTCCTGCTCG	
MT-C34-15D	5'-34-15D	CAACATCAGCAGGGCCGACTGGAACGACGGCTCAGCATCCAGCAACATAAG	pUCHR-mClover-MT-C34
	3'-34-15D	CGGCCCTGCTGATGTTGCAGTGGGCCTTCCTGATCAGGAGCTCCTGCTCG	
CD52-MT-C34s*	5'-mClover	CTGTACAAGCCAGCCAAACCTCCGG	pUCHR-mClover-MT-C34
	3'-MT34-tr	GTACTCGAGCTCACAGGAGCTCCTGCTCG	
Gl-MT-C34s*	5'-Gl-LD	AAGCGCTTCGTTCTGTTTGCCCTGATCTGCATCGCTGTGGCCGAGGCTATGACCTGGATGGAGTGG	pUCHR-mClover-MT-C34
	3'-MT34-tr	GTACTCGAGCTCACAGGAGCTCCTGCTCG	
* constructs were generated by a standard PCR with two primers			
Generation of *pUCHR-MT-C34-mClover* construct using fusion PCR and Afe I/Age I restriction sites			
Flanking primers	5'-Afe I	GAAGCGCTTCCTCTTCCTC	PCR products 1 + 2
	3'-Age I	CTCACCGGTGGCCATCC	
N-MT-C34	5'-52-P2A	CCACCTCTTCTGCTTCAG	pUCHR-CD5HA2-sAID-mClover
	3'-52-P2A	CTGAAGCAGAAGAGGTGG	pUCHR-mClover-MT-C34
Construction of *pUCHR-CCR5* plasmid using EcoR I/Xma I restriction sites			
CCR5	5'-EcoR-R5	CGAATTCGCCACCATGGATTATCAAGTGTCAAGTCC	cDNA from U937 cells
	3'-Xma-R5	GCCCGGGTCACAAGCCCACAGATATTTCC	
Generation of *pNL4-3* with mutations in gp41 using overlapping primers and Bsa BI/Bam HI sites of restriction			
Flanking primers	5'-NB-BsaBI	CCATCAGTGGACAAATTAGATG	PCR products 1 + 2
	3'-BamH-NB	GATCGTCCCAGATAAGTGC	
V38A	5'-V38A	CTGGTATAgcgCAGCAGCAGAACAATTTGC*	pNL4-3
	3'-V38A	GCTGCTGcgcTATACCAGACAATAATTGTCTGGCC	
E49K	5'-E49K	GAGGGCTATTaagGCGCAACAGCATCTGTTG	
	3'-E49K	GTTGCGCcttAATAGCCCTCAGCAAATTGTTC	
N126K	5'-N126K	GAGAAATTAACaagTACACAAGCTTAATACACTCCTT	
	3'-N126K	GCTTGTGTActtGTTAATTTCTCTGTCCCACTC	
E136G	5'-E136G	CACTCCTTAATTggaGAATCGCAAAACCAGCAAGA	
	3'-E136G	CGATTCtccAATTAAGGAGTGTATTAAGCTTG	
* The triplets coding mutated amino acids are shown by lowercase, and mutation is highlighted in red			

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442168; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table S2. Guide RNA and donor DNA generation.

Gene/ gRNA	Primers for donor generation	Templates
CXCR4 start AGAACCAG CGGTTACC ATGG***AGG***	*5'-MT-C34(HA)-lp:* AGTTTGTTGGCTGCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCAC CGCATCTGGAGAACCAGCGGTT**ACCACCATGAAGCGCTTCC** *3'-MT-C34(HA)-lp:* CGCTTCTGCCCGCTCGGAGAGGGGCTGCGCTCTAAGTTCAAACGTTTGTACATTTATGACAAAGC AGGTTGAAACTGGACTTACACTGATCCCCTC**GGCCGGTCCAGGATTCTC**	pUCHR-MT-C34-mClover or pUCHR-CD5HA2-mClover
	5'-MT-C34-F: AGTTTGTTGGCTGCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGC ATCTGGAGAACCAGCGGTTACC**ATGACCTGGATGGAGTG** *3'-MT-C34-F:* CGCTTCTGCCCGCTCGGAGAGGGGCTGCGCTCTAAGTTCAAACGTTTGTACATTTATGACAAAGCAGG TTGAAACTGGACTTACACTGATCCCCTC**CAGGAGCTCCTGCTCG**	pUCHR-mClover-MT-C34
	5'-C24-F: AGTTTGTTGGCTGCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATC TGGAGAACCAGCGGTTACCATG**TACACCTCACTGATCCAC** + *3'-MT-C34-F*	pUCHR-mClover-MT-C34
	5'-HA-F: AGTTTGTTGGCTGCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCT GGAGAACCAGCGGTTACCATG**TACCCTTACGACGTGCC** *3'-HA-F:* CGCTTCTGCCCGCTCGGAGAGGGGCTGCGCTCTAAGTTCAAACGTTTGTACATTTATGACAAAGCAGGTTG AAACTGGACTTACACTGATCCCCTC**GGCGTAGTCGGGCAC**	pUCHR-mClover-CD5HA2
	5'-2P23-F: AGTTTGTTGGCTGCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCAT CTGGAGAACCAGCGGTTACCATG**GAGATGACCTGGGAAGAG** *3'-2P23-F:* CGCTTCTGCCCGCTCGGAGAGGGGCTGCGCTCTAAGTTCAAACGTTTGTACATTTATGACAAAGCAGGTT GAAACTGGACTTACACTGATCCCCTC**CTTGAGCAGCTCTTCGATT**	pUCHR-mClover-2P23
	5'-HP23L-F: AGTTTGTTGGCTGCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCA TCTGGAGAACCAGCGGTTACCATG**GAGCTGACCTGGGAAGAG** *3'-HP23L-F:* CGCTTCTGCCCGCTCGGAGAGGGGCTGCGCTCTAAGTTCAAACGTTTGTACATTTATGACAAAGCAGGT TGAAACTGGACTTACACTGATCCCCTC**CTTGAGGATCTCTTCGATTT**	pUCHR-mClover-HP23L
CXCR4 exon-2 CACTTCAG ATAACTAC ACCG***AGG***	*5'-MT-C34-lp:* GCTTTAAAAATTTTTTTTAACTGGGTTAATGCTTGCTGAATTGGAAGTGAATGTCCATTCCTTTGCCTC TTTTGCAGATATACACTTCAGATAAC**GGATCCGGCGCAACAAAC** *3'-MT-C34-lp:* GTAGATGGTGGGCAGGAAGATTTTATTGAAATTAGCATTTTCTTCACGGAAACAGGGTTCCTTCATGG AGTCATAGTCCCCTGAGCCCATTTCCT**CACACAAAAAACCAACACAC**	pUCHR-mClover-MT-C34
	5'-X4-arm499bp-fwd: TTTGCGCTTTAGGAGAATGAGT + *5'-X4-arm499bp-rev:* CCAATTCAGCAAGCATTAACCCA	Genomic DNA from PBMC
	3'-X4-arm440bp-fwd: ACCCTGTTTCCGTGAAGAAAATG + *3'-X4-arm440bp-rev:* AGTCGGGAATAGTCAGCAGG	
CD4 start AGGGACTC CCCGGTTC ATTG***TGG***	*5'-HA-lp:* GTTTGCTGACTAATGATTGGCATTTCCCTCAGGCCCTGCCATTTCTGTGGGCTCAGGTCCCTACTGGCTCAG GCCCCTGCCTCCCTCGGCAAGG**CCACCATGAAGCGCTTCC** *3'-HA-lp:* GCCTTTCCTCCCACGTCATCTGCATTGAGACCCCAGGTCTGAGAACTTACCCAGTTGCAGCACCAGAAGCA AGTGCCTAAAAGGGACTCCCCGG**TTGGCCGGTCCAGGATTC**	pUCHR-CD5HA2-mClover
CD4 stop-1 GTAGCCCC ATTTGAGG CACG***AGG*** **CD4 stop-2** CTGGCCTC GTGCCTCA AATG***GGG***	*5'-MT-C34(HA)-lp:* GGAGGTGCTAGAACGCAAAGGGGTTGCAGTGGGGACAGACCTGCTCCCCTTCTTCTTTGTTCCTG CAGCCGGTTTCAGAAGACATGTAGCCCCATT**GGATCCGGCGCAACAAAC** *3'-MT-C34(HA)-lp:* GAGGCTGGGGATCTGCTACATTCATCTGGTCCGCAGGCAGGAAACGCGGGGCAGACACCTGGGG AGGCTGCAAGTGGGATCTGCCTGGCCTCGTGCC**TCAACTGAAGCAGAAGAGG**	pUCHR-mClover-MT-C34 or pUCHR-mClover-CD5HA2
	5'-CD4-arm526bp-fwd: CCTGGGCCATGTAACTGCTT + *5'-CD4-arm526bp-rev:* TGCAACCCCTTTGCGTTCTA	Genomic DNA from PBMC
	3'-CD4-arm536bp-fwd: CAGATGAATGTAGCAGATCCCCA + *3'-CD4-arm536bp-rev:* ATGTGAGACATCCCTTGAGGC	
HIV-1 capsid GTTAAAAG AGACCATC AATG***AGG***	*5'-MT-C34-lp:* CCCATGTTTTCAGCATTATCAGAAGGAGCCACCCCACAAGATTTAAACACCATGCTAAACACAGTGG GGGGACATCAAGCAGCCATGCAAATGTTAAAA**GGATCCGGCGCAACAAAC** *3'-MT-C34-lp:* CTAGTAGTTCCTGCTATGTCACTTCCCCTTGGTTCTCTCATCTGGCCTGGTGCAATAGGCCCTGCATGC ACTGGATGCACTCTATCCCATTCTGCAGC**CACACAAAAAACCAACACAC**	pUCHR-mClover-MT-C34
	5'-p24-arm530bp-fwd: GTGCGAGAGCGTCAGTATTAA + *5'-p24-arm530bp-rev:* TTCTGATAATGCTGAAAACATGGG	pHIV-1-GFPt
	5'-p24-arm515bp-fwd: GTGACATAGCAGGAACTACTAG + *5'-p24-arm515bp-rev:* GCTATGTGCCCTTCTTTGCC	
Modification of donor plasmids with protospacers and DTS sequences		
CXCR4 exon-2	CACTTCAGATAACTACACCG***AGG*** cloned upstream of the donor at Not I/PspX I sites after annealing	pJet-X4ex2-don
	CCTCGGTGTAGTTATCTGAAGT cloned downstream of the donor at Xba I/Nco I sites after annealing	
HIV-1 p24	***CCT***CATTGATGGTCTCTTTTAAC cloned upstream of the donor at Xba I/Nco I sites after annealing	pJet-p24-don
	GTTAAAAGAGACCATCAATG***AGG*** cloned downstream of the donor at Not I/PspX I sites after annealing	
DTS	ATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACCCTAACTGACACACATTCCACAG cloned upstream of the p24 donor or downstream of X4ex2 donor at Xba I/Nco I sites after annealing	pJet-p24-don, pJet-X4ex2-don
Primers for targeted locus amplification		
CXCR4 start	5'-gDNA: CGGACTCACTACCGACCAC 3'-gDNA: GTACGGGTACCTCCAATGTCC	Genomic DNA from CEM cells
CXCR4 exon-2	5'-gDNA: GTATTTAGGCAGGCGTGGGA 3'-gDNA: AGTCATTGGGGTAGAAGCGG	Genomic DNA from CEM or CD4 lymphocytes

Protospacer adjacent motifs are italicized and in bold. Primer sequences that have complementarity to the plasmid DNA template are shown in bold.