



Article

Validation of Promoters and Codon Optimization on CRISPR/Cas9-Engineered Jurkat Cells Stably Expressing aRep4E3 for Interfering with HIV-1 Replication

Koollawat Chupradit, Kanokporn Sornsuwan, Kritayaporn Saiprayong, Methichit Wattanapanitch and Chatchai Tayapiwatana

Special Issue Molecular and Genetic Analysis of HIV-1 Edited by

Dr. Larance Ronsard and Dr. Ashraf S. Yousif





https://doi.org/10.3390/ijms232315049





Article Validation of Promoters and Codon Optimization on CRISPR/Cas9-Engineered Jurkat Cells Stably Expressing αRep4E3 for Interfering with HIV-1 Replication

Koollawat Chupradit ^{1,2}, Kanokporn Sornsuwan ^{2,3}, Kritayaporn Saiprayong ¹, Methichit Wattanapanitch ^{1,*} and Chatchai Tayapiwatana ^{2,3,4,*}

- ¹ Siriraj Center for Regenerative Medicine, Research Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand
- ² Center of Biomolecular Therapy and Diagnostic, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai 50200, Thailand
- ³ Division of Clinical Immunology, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai 50200, Thailand
- ⁴ Center of Innovative Immunodiagnostic Development, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai 50200, Thailand
- * Correspondence: methichit.wat@mahidol.ac.th (M.W.); chatchai.t@cmu.ac.th (C.T.)

Abstract: Persistent and efficient therapeutic protein expression in the specific target cell is a significant concern in gene therapy. The controllable integration site, suitable promoter, and proper codon usage influence the effectiveness of the therapeutic outcome. Previously, we developed a non-immunoglobulin scaffold, alpha repeat protein (αRep4E3), as an HIV-1 RNA packaging interference system in SupT1 cells using the lentiviral gene transfer. Although the success of anti-HIV-1 activity was evidenced, the integration site is uncontrollable and may not be practical for clinical translation. In this study, we use the CRISPR/Cas9 gene editing technology to precisely knock-in aRep4E3 genes into the adeno-associated virus integration site 1 (AAVS1) safe harbor locus of the target cells. We compare the α Rep4E3 expression under the regulation of three different promoters, including cytomegalovirus (CMV), human elongation factor-1 alpha (EF1 α), and ubiquitin C (UbC) promoters with and without codon optimization in HEK293T cells. The results demonstrated that the EF1 α promoter with codon-optimized α Rep4E3mCherry showed higher protein expression than other promoters with non-optimized codons. We then performed a proof-of-concept study by knocking in the αRep4E3mCherry gene at the AAVS1 locus of the Jurkat cells. The results showed that the α Rep4E3mCherry-expressing Jurkat cells exhibited anti-HIV-1 activities against HIV-1NL4-3 strain as evidenced by decreased capsid (p24) protein levels and viral genome copies as compared to the untransfected Jurkat control cells. Altogether, our study demonstrates that the α Rep4E3 could interfere with the viral RNA packaging and suggests that the α Rep4E3 scaffold protein could be a promising anti-viral molecule that offers a functional cure for people living with HIV-1.

Keywords: alpha repeat protein; CRISPR/Cas9; promoter; codon optimization; HIV-1

1. Introduction

Gene therapy involves correcting defective genes, replacing malfunctioning genes, and adding therapeutic genes. Several factors affect the expression levels of transgenes, including promoters, enhancers, polyadenylation signals, and other expression elements. The promoter is one of the critical factors influencing the expression level of the protein of interest. It is known that each gene has its specific promoter. Some promoters can function in specific cell types called tissue-specific promoters. However, several non-specific promoters, such as human cytomegalovirus (CMV) and human elongation factor-1 alpha (EF1 α), are commonly used in gene therapy applications [1,2]. Even though the



Citation: Chupradit, K.; Sornsuwan, K.; Saiprayong, K.; Wattanapanitch, M.; Tayapiwatana, C. Validation of Promoters and Codon Optimization on CRISPR/Cas9-Engineered Jurkat Cells Stably Expressing αRep4E3 for Interfering with HIV-1 Replication. *Int. J. Mol. Sci.* **2022**, *23*, 15049. https://doi.org/10.3390/ijms232315049

Academic Editor: József Tőzsér

Received: 8 November 2022 Accepted: 27 November 2022 Published: 30 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). CMV promoter is a strong promoter, certain studies demonstrated that gene silencing is associated with DNA methylation leading to the downregulation of gene expression [3,4]. The EF1 α promoter has also shown high protein expression levels in primary cells and various cell lines [5,6]. Even though the ubiquitin C (UbC) promoter promotes stable eGFP expression in human embryonic stem cells (hESCs), the level is moderate in comparison to the constitutive CMV enhancer/chicken β -actin promoter (CAG) promoter [7,8]. This indicates that the choice of the promoter is essential for transgene expression in specific cell types. Therefore, selecting appropriate promoters for each cell type is essential for gene therapy applications.

Apart from the promoters, codon usage significantly influences protein expression levels. Even though more than one codon can encode the same amino acid, those codons are not used with the same frequencies and are different in every organism. This phenomenon is called the codon usage bias (CUB) [9–11]. CUB plays an essential role in several cellular processes, for example, transcription, mRNA stability, translation, and protein folding [12,13]. The genome-wide expression profile in the previous studies demonstrated positive correlations between codon biases and mRNA levels in prokaryotes and lower eukaryotes. The results revealed that the effects of CUB are mainly involved in the transcription process [14–17]. Moreover, suboptimal codon usage affected the translation process and mRNA stability [14]. Interestingly, specific cell types, such as immune cells, have distinct codon usage patterns among other immune cell types [18]. Therefore, in the genetic engineering approach, the codon usage of the target gene is usually optimized to increase the protein expression level.

We aim to compare the effects of promoter and codon usage on protein expression of a scaffold protein, i.e., alpha repeat protein (α Rep). Our previous studies demonstrated that the α Rep4E3 inhibited the RNA packaging process of HIV-1 in the SupT1 cell line using lentiviral transduction [19]. Since the α Rep4E3 is proposed for HIV gene therapy, lentiviral random integration of the transgene into the host chromosome may lead to the risks of insertional mutagenesis and oncogenesis [20]. Recently, the CRISPR/Cas9 system became popular in genetic engineering for precisely editing the chromosome at specific sites [21,22]. Several studies have shown that introducing the transgene into the adeno-associated virus integration site 1 (AAVS1) is safe, and the cells could maintain the transgene expression [23–25]. We formerly demonstrated that the mCherry gene could be delivered into the *AAVS1* site of human induced pluripotent stem cells (hiPSCs). The transduced iPSCs expressed high levels of the transgene and exhibited a normal karyotype after the gene editing [26].

In this study, we introduced an anti-HIV gene, α Rep4E3mCherry, using the CRISPR/ Cas9 technology targeting the *AAVS1* locus. The levels of α Rep4E3 expression under the control of three constitutive promoters, including CMV, human EF1 α , and UbC promoters, along with different codon optimizations, were first validated in HEK293T cells by plasmid vector transfection. Then, a suitable promoter with modified codons of the α Rep4E3mCherry was validated for the effectiveness of α Rep4E3 in interfering with HIV-1 propagation in the Jurkat cell line. The coherence of promoter usage and codon optimization shall be considered when sustainable gene therapy in a specific target cell is applied.

2. Results

2.1. Expression of the α Rep4E3mCherry Protein in HEK293T Cells under Various Constitutive Promoters

We utilized the CRISPR/Cas9 technique to introduce the anti-HIV-1 scaffold gene into HEK293T cells to generate stable cell lines expressing α Rep4E3mCherry under various constitutive promoters. Six SH200 plasmid vectors were constructed to use as donor templates for knock-in. The α Rep4E3mCherry gene construct was placed under the control of different promoters, including CMV, EF1 α , and UbC with or without codon optimization. The mCherry fusion partner was used as a reporter for tracking α Rep4E3 gene expression. The downstream genes expressing the green fluorescent protein (EGFP) and puromycin



resistance gene were under the control of the EF1 α promoter for verifying the transduction efficacy (Figure 1A).

Figure 1. Validation of the α Rep4E3mCherry gene constructs for stable gene expression in the HEK293T cells. (**A**) Schematic diagram of CRISPR/Cas9 plasmid vectors, including SH100 encoding the AAVS1 single-guide RNA and Cas9, and SH200 plasmids encoding the α Rep4E3mCherry gene with or without codon optimization under the control of CMV, EF1 α , and UbC promoters. The downstream of the plasmid vector harbors EGFP and puromycin-resistant genes, LHA = left homology arm, and RHA = right homology arm. (**B**) Representative images of the α Rep4E3mCherry expression in the HEK293T cells after puromycin selection for 13 days, PC = phase contrast, scale bar = 200 µm. (**C**) Flow cytometric analysis of mCherry and EGFP expression. The HEK293T cells stably expressing α Rep4E3mCherry were harvested on day 19 after puromycin selection. (**D**) The percentages of mCherry and EGFP expression in the HEK293T cells were analyzed by flow cytometry.

On day 13 after selection, the red and green fluorescent proteins were observed in six versions of the α Rep4E3mCherry, as shown in B. Notably, the expression levels determined by the flow cytometer on day 19 demonstrated that the codon-optimized α Rep4E3mCherry under the control of CMV and EF1 α promoters displayed a significantly higher number of double positive subpopulation compared to the non-optimized codon (Figure 1C, D). Interestingly, the EF1 α promoter with codon optimization showed the highest double-positive cells compared to other promoters. These results indicated that the α Rep4E3mCherry gene was successfully introduced into the HEK293T cells using the CRISPR/Cas9 technique, and the EF1 α promoter with the α Rep4E3 optimized codon resulted in the highest expression of α Rep4E3mCherry.

2.2. CRISPR/Cas9-Mediated & Rep4E3mCherry Gene Knock-in in the Jurkat Cell Line

We selected the SH200 plasmid DNA containing the EF1 α promoter with the α Rep4E3 optimized codon for further experiments in the Jurkat cell line due to its highest expression in the HEK293T cells (Figure 2A). On day 40 after puromycin selection, 32.5% of the Jurkat cells were double-positive for mCherry and EGFP, and 51.3% of the Jurkat cells were single positive for EGFP. We then FAC sorted the double-positive cells and cultured them for more than 15 days. The percentage of the double-positive cells increased slightly to 45.2% (Figure 2A,B). We performed limiting dilution to select an mCherry-positive clone. After clonal selection, the percentage of the double-positive population (mCherry⁺ EGFP⁺) increased to 76.7% (Figure 2C).



Figure 2. Stable α Rep4E3mCherry expression and clonal validation in the Jurkat cells. (**A**) Schematic diagram of the CRISPR/Cas9-based gene delivery targeting the AAVS1 locus using the SH200 donor plasmid and the RNP complex in the Jurkat cells. (**B**) Representative images of the Jurkat cells stably expressing α Rep4E3mCherry (optimized codon) under the control of the EF1 α promoter after puromycin selection for 40 days and after sorting for 16 days, BF = bright field, scale bar = 200 µm (top panel); flow cytometric analysis of the Jurkat cells expressing α Rep4E3mCherry (bottom panel). (**C**) A single clone of the Jurkat cells stably expressing α Rep4E3mCherry, scale bar = 100 µm (top panel), and flow cytometric analysis of mCherry and EGFP expression (bottom panel).

2.3. Analysis of the α Rep4E3mcherry Transgene Integration in the Jurkat Clone by Multiplex PCR

To confirm the transgene integration after knock-in, we performed a multiplex PCR analysis of the genomic DNA from the Jurkat clone using the primers that amplify the upstream region of the AAVS1 LHA, mCherry reporter protein, and downstream region of the AAVS1 RHA as indicated by the black arrow (Figure 3A). The results demonstrated a successful transgene knock-in at the AAVS1 safe harbor locus on human chromosome 19, as indicated by the 2000-bp PCR product (Figure 3B). However, the lower band (600 bp) was also observed. This result indicated that the α Rep4E3mCherry gene was integrated into one allele (monoallelic-edited).



Figure 3. Analysis of α Rep4E3mcherry transgene integration in the Jurkat clone by multiplex PCR. (**A**) Schematic diagram of the multiplex PCR amplification. (**B**) The PCR products amplified from the genomic DNA extracted from the untransfected Jurkat cells and the Jurkat cells expressing α Rep4E3mCherry, and non-template control (NTC). The 2000-bp band indicates the integrated α Rep4E3mCherry gene at the AAVS1 locus, while the 600-bp band represents the wild-type AAVS1 locus.

2.4. The *cRep4E3mCherry Jurkat Clone Conferred Resistance to HIV-1 Infection*

To evaluate the protection levels of the Jurkat clone stably expressing α Rep4E3mCherry against HIV-1, the untransfected and α Rep4E3mCherry Jurkat cells were challenged with

HIV-1NL4-3 at an MOI of 10. From day 15 post-infection (pi) onwards, the untransfected Jurkat cells displayed multinucleated giant cells, a sign of viral infection, slightly more than the α Rep4E3mcherry Jurkat cells. However, the α Rep4E3mCherry Jurkat cells still expressed mCherry reporter protein until day 23 pi (Figure 4A). The Alu-gag PCR on day 15 pi showed a delayed Cq of 22.95 in the α Rep4E3mcherry Jurkat cells compared to 13.36 in the control Jurkat cells, indicating that the Jurkat cells stably expressing α Rep4E3mCherry exhibited lower HIV-1 integration compared to the control Jurkat cells (Figure 4B). The levels of anti-viral protection were evaluated in the culture supernatants collected on days 5, 9, 15, 20, and 23 pi using an ELISA. Following HIV-1 infection, high levels of p24 protein were observed in the culture supernatant from the control Jurkat cells on day 15 pi, while the levels of p24 protein released from the α Rep4E3mcherry Jurkat cells were suppressed until day 20 pi (Figure 4C). The result from the Alu-gag PCR was concordant with the ELISA results, indicating that the α Rep4E3 protein initially delayed HIV-1 infection in the Jurkat cells. We further investigated the role of the α Rep4E3 protein in RNA packaging interference following our published protocol [19]. The incorporated viral RNA was quantified using the viral load assay. Our results revealed that at the equal concentration of 20 ng/mL p24 protein on day 23 pi, the viral genomic RNA in the culture supernatant harvested from the α Rep4E3mCherry Jurkat cells was significantly lower than those of the control Jurkat cells (Figure 4D). This finding reflects an anti-HIV-1 activity of the α Rep4E3 protein at the RNA packaging process, although a slight increase in the level of p24 protein was observed on day 23 pi.



Figure 4. Inhibition of HIV-1 replication in the α Rep4E3mCherry-expressing Jurkat cells. (A) Morphology and mCherry expression were monitored by light and fluorescence microscopy using an

inverted fluorescence microscope, BF = bright field, scale bar = 100 µm. (**B**) The *Alu-gag* quantitative real-time polymerase chain reaction (PCR) (*Alu-gag* qPCR) was performed in the untransfected (black line) and α Rep4E3mCherry (red line) expressing Jurkat cells on day 15 pi to determine the HIV-1 integration. (**C**) The culture supernatants were collected at 3-day intervals, from day 5 to day 23 pi. The culture supernatants were assayed for p24 protein. (**D**) HIV RNA copies were determined per 20 ng of p24 protein on day 23 pi using the HIV viral load assay. The bar graph represents the mean \pm SD from triplicate wells. ** $p \leq 0.01$ using unpaired *t*-test.

3. Discussion

Gene therapy is an approach to treat patients with both inherited and acquired diseases by genetically modifying patients' cells. Viral vectors are widely used since they rapidly invade target cells using their natural ability. However, random integration can lead to the risk of genotoxicity [27,28]. Previously, we delivered the gene encoding α Rep4E3 controlled by the MND promoter into the SupT1 cells using lentiviral transduction. Even though the results showed stable protein expression and anti-HIV-1 activity, the transgene copy number and integration site were not verified [19]. These factors could probably bias the data interpretation. CRISPR/Cas9 has become a versatile tool for gene editing due to low off-target activity, ease of use, and the ability to perform multiplex gene editing [29]. Furthermore, the advantages of CRISPR/Cas9 are high controllability and the ability to identify the gene editing sites correctly. The previous study showed that CRISPR/Cas9 targeting CXCR4 or CCR5 chemokine receptors, the co-receptors for HIV-1, protected the primary CD4⁺ T cells from viral infection [30]. Moreover, using a single-guide RNA (sgRNA) targeting the long terminal repeat (LTR) sequence or other essential genes for viral replication could inactivate the HIV-1 provirus by introducing a mutation into the target region [31]. However, the leakage of the virus from using one sgRNA still occurred. Thus, molecules that bind to the conserved regions could overcome this limitation.

Several attempts have been made to inhibit HIV-1 infection. Our previous studies demonstrated the use of scaffold proteins, including ankyrin repeat protein, zinc finger protein, and α Rep4E3 protein, as anti-HIV-1 molecules. The results showed that these designed scaffold proteins could inhibit HIV-1 infection at various steps of the viral life cycle [19,32,33]. The present study used the α Rep4E3 targeting the downstream conserved region of the zinc finger domain in the nucleocapsid (NC). The binding of the α Rep4E3 to the NC results in the inhibition of the HIV-1 RNA packaging process, in which the viral RNA molecules are recognized by the nucleocapsid (NC) domain during the viral assembly [19]. Previous studies showed that the CHCC residues of the NC zinc fingers are highly conserved, and the mutation of this region resulted in the production of non-infectious viruses [34–37]. Hence, using the α Rep4E3 protein for HIV-1 treatment may alleviate the probability of viral escape.

Since the promoter choice affects the protein expression levels, we first screened different constitutive promoters, including CMV, EF1 α , and UbC, with and without codon optimization in the HEK293T cells. All the transfected HEK293T cells expressed mCherry protein after antibiotic selection. Interestingly, the EF1 α promoter resulted in the highest percentages of double-positive cells, followed by the UbC and CMV promoters, indicating that the transgene promoter significantly influences the gene expression. Our result contradicts the previous study that HEK293T cells expressed the highest level of EGFP integrated at the *AAVS1* locus under the CMV promoter, followed by EF1 α , SV40, and TK promoters [38]. Apart from the promoter, the secondary structure of the transcribed RNA influences individual protein expression levels [39]. Fine-tuning the coding sequence to optimize mRNA structure should be further considered [40].

In addition to the promoter choice, a significant difference in protein expression was observed after codon optimization in the HEK293T cells. This result implied that CUB also affects protein expression. Several studies showed that codon changes could result in changes in the conformation and stability of the protein [41,42]. Codon optimization is an

algorithm designed to increase the expression of the protein of interest by changing the base for synonymous codons. Some studies demonstrated that optimizing codon usage bias could enhance the recombinant protein expression in different host organisms [43–45]. The crucial codon bias is evident in specific cells, such as monocytes, B and T lymphocytes. Previously the mutational bias influenced the codon bias in monocytes, while the translational selection could affect codon usage in B and T cells [18]. Thus, optimization for T cell codon usage should benefit the expression of the Rep4E3mCherry in Jurkat cells.

To determine whether codon usage affects protein expression, we compared the noncodon and codon-optimized α Rep4E3 protein expression. The results showed that the codon-optimized a Rep4E3mCherry HEK293T cells had higher mCherry⁺EGFP⁺ expression than the non-codon optimization in the CMV and EF1 α promoters. However, there was no difference in the double-positive cells in the non-codon and codon-optimized α Rep4E3 under the UbC promoter. Hence, the EF1 α promoter with the codon-optimized α Rep4E3mCherry was chosen for further experiments in the Jurkat cells. The Jurkat cells expressing α Rep4E3mCherry showed anti-HIV activity when infected with HIV-1_{NL4-3}. The increase in the p24 level in the α Rep4E3mCherry Jurkat cells was observed on day 23 pi, indicating that the virus could assimilate viral particles in the engineered cells after a certain period. However, the infectivity of these particles is lower since most of them are HIV genome free suggesting that the α Rep4E3 functions in interfering with the RNA packaging process. This result is consistent with our previous study, in which lentiviral transduction was applied to generate the SupT1 cells expressing the α Rep4E3 protein. In addition, monoallelic expression of the aRep4E3mCherry gene was adequate to interfere with HIV-1 replication. After day 20, the p24 level was saturated and beyond the interference efficacy of the α Rep4E3mCherry protein. The load of HIV-1 in the culture supernatant is much higher than the median viral loads in HIV-1-infected individuals [46]. Thus, the HIV-1 protection by α Rep4E3mCherry should be further validated to ensure its clinical significance.

In conclusion, we successfully delivered the α Rep4E3mCherry gene into the *AAVS1* locus and explored the effects of the promoter and codon usage of α Rep4E3mCherry gene in HEK293T cells. The results demonstrated that the EF1 α promoter with the codon-optimized α Rep4E3 showed the highest protein expression among other promoters. In addition, the Jurkat stably expressing the α Rep4E3mCherry protein exhibited anti-HIV-1 activity by inhibiting the viral RNA packaging process resulting in the reduction of viral genome copies. These findings will be a foundation for establishing engineered stem cells for HIV gene therapy.

4. Materials and Methods

4.1. Cells

Human embryonic kidney (HEK293T) cells and Jurkat cells, Clone E6-1 (a human T lymphoblast cell line susceptible to HIV-1 infection) were obtained from the ATCC (Manassas, VA, USA). HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Gibco, Invitrogen, Grand Island, NY, USA), and Jurkat cells were maintained in RPMI-1640 medium, both supplemented with 10% fetal bovine serum (FBS) (Gibco, Invitrogen), 2 mM GlutaMax (Gibco, Invitrogen), and 100 units/mL of penicillin-streptomycin (Gibco, Invitrogen). All cells were maintained in a 37 °C humidified incubator containing 5% CO₂. The cells were passaged every 2–3 days at a subcultivation ratio of 1:5–1:6.

4.2. Optimization of Codon Usages

We optimized the codon of the α Rep4E3 gene using GenSmart Codon Optimization (GenScript) by inserting the α Rep4E3 gene into the website and selecting a human-T cell as a host organism. The optimized α Rep4E3 DNA sequences were synthesized by Integrated DNA Technologies. Both the non-optimized and optimized codons were used for the SH200 donor plasmid (GeneCopoeia, Rockville, MD, USA) construction.

4.3. Construction of the Donor Plasmid Vector Harboring *cRep4E3mCherry* Gene

The CRISPR/Cas9 system used in this study is a plasmid-based method consisting of the SH100 plasmid (GeneCopoeia) carrying sgRNA targeting the *AAVS1* locus and Cas9 gene, and the SH200 plasmid carrying *AAVS1* left homology arm (LHA), different promoters including CMV, UbC or EF1 α with either codon-optimized or non-codon-optimized aRep4E3 sequence, mCherry, EGFP/puro, and *AAVS1* right homology arm (RHA) (Figure 1A). The non-codon-optimized aRep4E3 sequence was obtained from our previous work [19]. The optimized α Rep4E3 DNA sequences were synthesized by Integrated DNA Technologies. The α Rep4E3mCherry gene was then amplified and ligated into the SH200 plasmid vector using *Xba*I and *BstB*I restriction enzyme digestions to obtain the SH200 α Rep4E3mCherry donor plasmid vector. The DNA sequences of promoter and α Rep4E3mCherry genes were confirmed by DNA sequencing.

4.4. CRISPR/Cas9 αRep4E3mCherry Gene Delivery and Expression in Cell Lines

We first compare the expression of the α Rep4E3mCherry gene in HEK293T cells using a donor plasmid with different types of promoters and codon usage. The HEK293T cells were transfected with a 1:1 ratio of the SH100 plasmid (GeneCopoeia) encoding the AAVS1 sgRNA and Cas9 protein, and the SH200 plasmid DNA harboring α Rep4E3mCherry gene. Briefly, 1 µg of SH100 and 1 µg of SH200 plasmid DNA were co-transfected into 2 × 10⁵ HEK293T cells using Lipofectamine 3000 Reagent (Thermo Fisher, Waltham, MA, USA). The transfected cells were cultured at 37 °C, 5% CO₂. Two days post-transfection, the transfected HEK293T cells were treated with 200 ng/mL of puromycin. The puromycin concentration was gradually increased to 500 ng/mL for enrichment after 4 days. The puromycin selection was carried out until all the untransfected control cells died.

The SH200 plasmid DNA that gave the highest αRep4E3mCherry expression level was selected for further experimentation in the Jurkat cells. Gene knock-in was carried out using the ribonucleoprotein (RNP) complex. Briefly, 150 µg/mL of Alt-R[®] S.p. Cas9 Nuclease V3 (Integrated DNA Technologies, Coralville, IA, USA) was mixed with 90 µg/mL of sgRNA targeting the *AAVS1* site generated by GeneArtTM Precision gRNA Synthesis Kit (Invitrogen), and incubated at room temperature for 10 min.

The Jurkat cells were transfected with the RNP complex using Lonza P3 Primary Cell 4D-Nucleofector X Kit S (Lonza, Basel, Switzerland). Briefly, 5×10^5 Jurkat cells were resuspended with 20 µL of P3 4D-Nucleofector solution containing preincubated RNP and 3.6 µg of SH200 plasmid DNA using CL120 protocol by a 4D-Nucleofector X unit. The transfected cells were incubated in a complete RPMI-1640 medium at 37 °C, 5% CO₂ for 16–18 h. After incubation, 100 ng/mL of puromycin was added to the culture medium to select the transfected cells.

4.5. Fluorescence-Activated Cell Sorting (FACS)

After puromycin selection, the Jurkat cells stably expressing α Rep4E3mCherry were expanded for cell sorting— 1×10^7 cells were resuspended in 2%FBS/PBS. The cell sorting was performed by a flow cytometer (FACSAria III, BD Biosciences, Franklin Lakes, NJ, USA). The sorted pool was then cultured in a complete RPMI-1640 medium supplemented with 200 units/mL of penicillin-streptomycin.

4.6. Clonal Isolation by Limiting Dilution

The α Rep4E3mCherry-sorted pool was resuspended in a complete RPMI-1640 medium to obtain a seeding density of 4 cells/100 µL/well. Then 100 µL of the cell suspension was transferred to each well of the 96-well flat-bottom plate and incubated at 37 °C, 5% CO₂. A single colony was selected and expanded for further experiments.

4.7. Determination of the α Rep4E3mCherry Integration Site in the Knock-in Cell Lines

The genomic DNA of the untransfected Jurkat cells and the Jurkat cells stably expressing α Rep4E3mCherry was extracted using the High Pure PCR Template Kit (Roche, Basel, Switzerland). Twenty-five nanograms of genomic DNA was used as a template for PCR using primers specific to the left homology arm of the *AAVS1* locus and *mCherry* gene. The primer sequences were: Fwd AAVS1 primer, 5'-TCTTGTAGGCCTGCATCATCACC-3', Rev AAVS1 primer, 5'-AGGATCCTCTCTGGCTCCAT-3', and mCherry primer, 5'-TCGTAGGGGGGGGCCCTCGC-3'. The reactions were performed in the final volume of 25 μ L using DreamTaq Green PCR Master Mix (2X) (Thermo Fisher) using the following program: 20 s at 95 °C, followed by 35 cycles of denaturation at 95 °C for 3 s and annealing at 70 °C for 30 s and extension at 72 °C for 2 min and 10 s.

4.8. Viral Stocks

Replication-competent HIV-1_{NL4-3} virus (X4-tropic) was produced by transfecting the pNL4-3 plasmid into HEK293T cells. Briefly, 3.5×10^6 HEK293T cells were seeded onto 10-cm dishes and transfected with 5 µg of the pNL4-3 plasmid, using Lipofectamine 3000 reagent (Thermo Fisher), as previously described [33]. After incubation for 5 h, the transfection mixture was removed, replaced with 10 mL of complete medium, and incubated at 37 °C, 5% CO₂. The HIV-1 virus was harvested from the culture supernatants by filtration through sterile syringe filters with a Millex-HA syringe filter unit, 0.45-µm pore size (Merck, Rahway, NJ, USA). HIV-1 samples were aliquoted and kept frozen at -80 °C. The culture supernatants were subjected to HIV viral load assay. The level of HIV virion was evaluated using reverse transcription quantitative polymerase chain reaction (RT-qPCR) by COBAS Ampliprep/COBAS Taqman HIV-1 test (Roche).

4.9. HIV-1 Infection

To determine the anti-viral protection of α Rep4E3mCherry, the Jurkat cells expressing α Rep4E3mCherry were challenged with the X4-tropic HIV-1NL4-3. The untransfected Jurkat cells and the Jurkat cells expressing α Rep4E3mCherry were maintained in a growth medium for at least 4 weeks prior to the HIV-1 challenge. The cells were incubated with HIV-1 at a multiplicity of infection (MOI) of 10 for 16 h. The cells were then washed three times with a serum-free medium and resuspended in a fresh growth medium. They were split (1:2) at 3-day intervals to maintain a cell density of approximately 10⁶ cells/mL. HIV-1 replication was monitored in the culture supernatant by enzyme-linked immunosorbent assay (ELISA) using Genscreen ULTRA HIV Ag-Ab (Biorad, Hercules, CA, USA) by capsid (p24) protein detection and viral load assay described above. The cell pellets were kept for determining the HIV-1 integration using the *Alu-gag* PCR. To further determine the HIV-1 viral RNA, the culture supernatants on day 23 post-infection (pi) were subjected to HIV viral load assay. The levels of HIV virion in the culture supernatant were normalized to 20 ng/mL of p24 protein. The viral load was determined by RT-qPCR by COBAS Ampliprep/COBAS Taqman HIV-1 test (Roche).

4.10. HIV-1 Integration Assay

The HIV-1 integration was determined by a conventional *Alu-gag* qPCR assay [47,48]. Briefly, the genomic DNA of the HIV-1 infected cells was extracted using the High Pure PCR Template Kit. The first round of PCR was performed using a pair of specific primers against *Alu* (human) and *gag* (HIV-1) sequences. The primer sequences were: *Alu* forward, 5'-GCC TCC CAA AGT GCT GGG ATT ACA G-3', and HIV-1 *gag* reverse, 5'-GTT CCT GCT ATG TCA CTT CC-3'. The reactions were performed in the final volume of 25 µL using DreamTaq Green PCR Master Mix (2X) (Thermo Fisher). The second round of RU5 kinetic PCR was performed using 10 µL of undiluted first-round PCR product. The primer sequences were U3_FWD_SS_RIHES01, 5'- AGT GGA GGT TTG ACA GCC GCC TA -3' and deltaU3_REV_SS_RIHES, 5'-CCA GTA CAG GCA AAA AGC AGC TGC-3'. The deltaU3 molecular probe, labeling at the 5' terminus with the 6-carboxyfluorescein (FAM) reporter and at the 3' terminus with the Black Hole Quencher (BHQ), contains the following sequence: 5'-6FAM- CCT CCC TGG AAA GTC CCC AGC GGA AAG TCC-BHQ-3'. The 25-µL reaction volume contained THUNDERBIRD[®] Probe qPCR Mix (Toyobo, Osaka,

Japan), 400 nM of U3_FWD_SS_RIHES01 primer, 400 nM of deltaU3_REV_SS_RIHES primer, and 140 nM deltaU3 molecular beacon probe. The PCR reactions were performed using the MiniOpticon real-time PCR system using the following program: 20 s at 95 °C, followed by 50 cycles of denaturation at 95 °C for 3 s and annealing and extension at 63 °C for 30 s. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as an internal control for the DNA quantification in each qPCR assay, using the following GAPDH primer sequences: GAPDH_FWD, 5'-GAA GGT GAA GGT CGG AGT C-3' and GAPDH_REV, 5'-GAA GAT GGT GAT GGG ATT TC-3'. The GAPDHTM molecular beacon probe contained the following sequence: 5'-FAM-CAA GCT TCC CGT TCT CAG CCT-BBQ-3'. The reactions were carried out as previously described.

4.11. Statistical Analysis

Statistical analyses were performed using GraphPad Prism software. Statistical significance was determined by the unpaired *t*-test analysis. We indicated the significance levels as follows: p < 0.01 **.

Author Contributions: Conceptualization, C.T.; methodology, K.C.; validation, K.C.; formal analysis, K.C.; investigation, K.C. and K.S. (Kanokporn Sornsuwan); resources, M.W. and C.T.; data curation, K.C. and K.S. (Kritayaporn Saiprayong); writing—original draft preparation, K.C.; writing—review and editing, K.C., M.W. and C.T.; supervision, M.W. and C.T.; funding acquisition, M.W. and C.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Siriraj Research Fund, Faculty of Medicine Siriraj Hospital, Mahidol University (grant number (IO) R016433029), the Distinguished Research Professor Grant (NRCT 808/2563), the Program Management Unit for Human Resources & Institutional Development, Research and Innovation (grant number B05F640041 and B05F640056). The Fundamental Fund, Ministry of Higher Education, Science, Research and Innovation (grant number R000030097). KSa is supported by the Siriraj Graduate Scholarship, Faculty of Medicine Siriraj Hospital, Mahidol University. MW is supported by the Chalermphrakiat Grant, Faculty of Medicine Siriraj Hospital, Mahidol University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: We thank the Core Facilities and Research Support Service, Faculty of Medicine Siriraj Hospital, Mahidol University.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Maguire, C.A.; Ramirez, S.H.; Merkel, S.F.; Sena-Esteves, M.; Breakefield, X.O. Gene therapy for the nervous system: Challenges and new strategies. *Neurotherapeutics* 2014, 11, 817–839. [CrossRef] [PubMed]
- McIntyre, C.; Derrick-Roberts, A.L.; Byers, S.; Anson, D.S. Correction of murine mucopolysaccharidosis type IIIA central nervous system pathology by intracerebroventricular lentiviral-mediated gene delivery. J. Gene Med. 2014, 16, 374-87. [CrossRef] [PubMed]
- 3. Hsu, C.C.; Li, H.P.; Hung, Y.H.; Leu, Y.W.; Wu, W.H.; Wang, F.S.; Lee, K.D.; Chang, P.J.; Wu, C.S.; Lu, Y.J.; et al. Targeted methylation of CMV and E1A viral promoters. *Biochem. Biophys. Res. Commun.* **2010**, 402, 228–234. [CrossRef] [PubMed]
- 4. Osterlehner, A.; Simmeth, S.; Gopfert, U. Promoter methylation and transgene copy numbers predict unstable protein production in recombinant Chinese hamster ovary cell lines. *Biotechnol. Bioeng.* **2011**, *108*, 2670-81. [CrossRef] [PubMed]
- 5. Wang, X.; Xu, Z.; Tian, Z.; Zhang, X.; Xu, D.; Li, Q.; Zhang, J.; Wang, T. The EF-1alpha promoter maintains high-level transgene expression from episomal vectors in transfected CHO-K1 cells. *J. Cell. Mol. Med.* **2017**, *21*, 3044–3054. [CrossRef] [PubMed]
- Kim, D.W.; Uetsuki, T.; Kaziro, Y.; Yamaguchi, N.; Sugano, S. Use of the human elongation factor 1 alpha promoter as a versatile and efficient expression system. *Gene* 1990, 91, 217–223. [CrossRef]
- Liew, C.G.; Draper, J.S.; Walsh, J.; Moore, H.; Andrews, P.W. Transient and stable transgene expression in human embryonic stem cells. *Stem Cells* 2007, 25, 1521–1528. [CrossRef]
- 8. Norrman, K.; Fischer, Y.; Bonnamy, B.; Wolfhagen Sand, F.; Ravassard, P.; Semb, H. Quantitative comparison of constitutive promoters in human ES cells. *PLoS ONE* **2010**, *5*, e12413. [CrossRef]
- 9. Ikemura, T. Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. 1985, 2, 13–34.
- 10. Sharp, P.M.; Tuohy, T.M.; Mosurski, K.R. Codon usage in yeast: Cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res.* **1986**, *14*, 5125–5143. [CrossRef]

- 11. Plotkin, J.B.; Kudla, G. Synonymous but not the same: The causes and consequences of codon bias. *Nat. Rev. Genet.* **2011**, *12*, 32–42. [CrossRef] [PubMed]
- 12. Quax, T.E.; Claassens, N.J.; Soll, D.; van der Oost, J. Codon Bias as a Means to Fine-Tune Gene Expression. *Mol. Cell.* **2015**, *59*, 149-61. [CrossRef] [PubMed]
- 13. Liu, Y. A code within the genetic code: Codon usage regulates co-translational protein folding. *Cell Commun. Signal.* **2020**, *18*, 145. [CrossRef] [PubMed]
- 14. Zhou, Z.; Dang, Y.; Zhou, M.; Li, L.; Yu, C.H.; Fu, J.; Chen, S.; Liu, Y. Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E6117-E25. [CrossRef] [PubMed]
- 15. Jeacock, L.; Faria, J.; Horn, D. Codon usage bias controls mRNA and protein abundance in trypanosomatids. *eLife* **2018**, *7*, e32496. [CrossRef]
- 16. Coghlan, A.; Wolfe, K.H. Relationship of codon bias to mRNA concentration and protein length in *Saccharomyces cerevisiae*. *Yeast* **2000**, *16*, 1131-45. [CrossRef]
- 17. dos Reis, M.; Wernisch, L.; Savva, R. Unexpected correlations between gene expression and codon usage bias from microarray data for the whole *Escherichia coli* K-12 genome. *Nucleic Acids Res.* **2003**, *31*, 6976–6985. [CrossRef]
- 18. Huo, X.; Liu, S.; Li, Y.; Wei, H.; Gao, J.; Yan, Y.; Zhang, G.; Liu, M. Analysis of synonymous codon usage of transcriptome database in Rheum palmatum. *PeerJ* 2021, *9*, e10450. [CrossRef]
- Hadpech, S.; Nangola, S.; Chupradit, K.; Fanhchaksai, K.; Furnon, W.; Urvoas, A.; Valerio-Lepiniec, M.; Minard, P.; Boulanger, P.; Hong, S.S.; et al. Alpha-helicoidal HEAT-like Repeat Proteins (alphaRep) Selected as Interactors of HIV-1 Nucleocapsid Negatively Interfere with Viral Genome Packaging and Virus Maturation. *Sci. Rep.* 2017, *7*, 16335. [CrossRef]
- Schlimgen, R.; Howard, J.; Wooley, D.; Thompson, M.; Baden, L.R.; Yang, O.O.; Christiani, D.C.; Mostoslavsky, G.; Diamond, D.V.; Duane, E.G.; et al. Risks Associated With Lentiviral Vector Exposures and Prevention Strategies. *J. Occup. Environ. Med.* 2016, 58, 1159–1166. [CrossRef]
- 21. Ran, F.A.; Hsu, P.D.; Wright, J.; Agarwala, V.; Scott, D.A.; Zhang, F. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **2013**, *8*, 2281-308. [CrossRef] [PubMed]
- 22. Shalem, O.; Sanjana, N.E.; Hartenian, E.; Shi, X.; Scott, D.A.; Mikkelson, T.; Heckl, D.; Ebert, B.L.; Root, D.E.; Doench, J.G.; et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **2014**, *343*, 84–87. [CrossRef] [PubMed]
- Smith, J.R.; Maguire, S.; Davis, L.A.; Alexander, M.; Yang, F.; Chandran, S.; Ffrench-Constant, C.; Pedersen, R.A. Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. *Stem Cells* 2008, 26, 496–504. [CrossRef] [PubMed]
- Inderbitzin, A.; Loosli, T.; Kouyos, R.D.; Metzner, K.J. Quantification of transgene expression in GSH AAVS1 with a novel CRISPR/Cas9-based approach reveals high transcriptional variation. *Mol. Ther. Methods Clin. Dev.* 2022, 26, 107–118. [CrossRef] [PubMed]
- Papapetrou, E.P.; Schambach, A. Gene Insertion Into Genomic Safe Harbors for Human Gene Therapy. *Mol. Ther.* 2016, 24, 678-84. [CrossRef]
- 26. Chupradit, K.; Thongsin, N.; Tayapiwatana, C.; Wattanapanitch, M. A precise gene delivery approach for human induced pluripotent stem cells using Cas9 RNP complex and recombinant AAV6 donor vectors. *PLoS ONE* **2022**, *17*, e0270963. [CrossRef]
- 27. Hacein-Bey-Abina, S.; de Saint Basile, G.; Cavazzana-Calvo, M. Gene therapy of X-linked severe combined immunodeficiency. *Methods Mol. Biol.* **2003**, 215, 247–259.
- Hacein-Bey-Abina, S.; von Kalle, C.; Schmidt, M.; Le Deist, F.; Wulffraat, N.; McIntyre, E.; Radford, I.; Villeval, J.L.; Fraser, C.C.; Cavazzana-Calvo, M.; et al. A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* 2003, 348, 255–256. [CrossRef]
- 29. Chen, K.; Gao, C. Targeted genome modification technologies and their applications in crop improvements. *Plant Cell Rep.* **2014**, 33, 575–583. [CrossRef]
- 30. Liu, Z.; Chen, S.; Jin, X.; Wang, Q.; Yang, K.; Li, C.; Xiao, Q.; Hou, P.; Liu, S.; Wu, S.; et al. Genome editing of the HIV co-receptors CCR5 and CXCR4 by CRISPR-Cas9 protects CD4(+) T cells from HIV-1 infection. *Cell Biosci.* **2017**, *7*, 47. [CrossRef]
- Wang, G.; Zhao, N.; Berkhout, B.; Das, A.T. CRISPR-Cas based antiviral strategies against HIV-1. Virus Res. 2018, 244, 321–332.
 [CrossRef] [PubMed]
- Nangola, S.; Urvoas, A.; Valerio-Lepiniec, M.; Khamaikawin, W.; Sakkhachornphop, S.; Hong, S.S.; Boulanger, P.; Minard, P.; Tayapiwatana, C. Antiviral activity of recombinant ankyrin targeted to the capsid domain of HIV-1 Gag polyprotein. *Retrovirology* 2012, 9, 17. [CrossRef] [PubMed]
- Sakkhachornphop, S.; Barbas, C.F., 3rd; Keawvichit, R.; Wongworapat, K.; Tayapiwatana, C. Zinc finger protein designed to target 2-long terminal repeat junctions interferes with human immunodeficiency virus integration. *Hum. Gene. Ther.* 2012, 23, 932–942. [CrossRef]
- 34. Darlix, J.L.; Lapadat-Tapolsky, M.; de Rocquigny, H.; Roques, B.P. First glimpses at structure-function relationships of the nucleocapsid protein of retroviruses. *J. Mol. Biol.* **1995**, 254, 523–537. [CrossRef] [PubMed]
- 35. Levin, J.G.; Guo, J.; Rouzina, I.; Musier-Forsyth, K. Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: Critical role in reverse transcription and molecular mechanism. *Prog. Nucleic Acid Res. Mol. Biol.* **2005**, *80*, 217–286.
- 36. Muriaux, D.; Darlix, J.L. Properties and functions of the nucleocapsid protein in virus assembly. *RNA Biol.* **2010**, *7*, 744–753. [CrossRef]

- 37. Mougel, M.; Cimarelli, A.; Darlix, J.L. Implications of the nucleocapsid and the microenvironment in retroviral reverse transcription. *Viruses* **2010**, *2*, 939–960. [CrossRef]
- Shin, S.; Kim, S.H.; Shin, S.W.; Grav, L.M.; Pedersen, L.E.; Lee, J.S.; Lee, G.M. Comprehensive Analysis of Genomic Safe Harbors as Target Sites for Stable Expression of the Heterologous Gene in HEK293 Cells. ACS Synth. Biol. 2020, 9, 1263–1269. [CrossRef]
- Mauger, D.M.; Cabral, B.J.; Presnyak, V.; Su, S.V.; Reid, D.W.; Goodman, B.; Link, K.; Khatwani, N.; Reynders, J.; Moore, M.J.; et al. mRNA structure regulates protein expression through changes in functional half-life. *Proc. Natl. Acad. Sci. USA* 2019, 116, 24075–24083. [CrossRef]
- Leppek, K.; Byeon, G.W.; Kladwang, W.; Wayment-Steele, H.K.; Kerr, C.H.; Xu, A.F.; Kim, D.S.; Topkar, V.V.; Choe, C.; Rothschild, D.; et al. Combinatorial optimization of mRNA structure, stability, and translation for RNA-based therapeutics. *Nat. Commun.* 2022, 13, 1356. [CrossRef]
- Agashe, D.; Martinez-Gomez, N.C.; Drummond, D.A.; Marx, C.J. Good codons, bad transcript: Large reductions in gene expression and fitness arising from synonymous mutations in a key enzyme. *Mol. Biol. Evol.* 2013, 30, 549–560. [CrossRef] [PubMed]
- Tsai, C.J.; Sauna, Z.E.; Kimchi-Sarfaty, C.; Ambudkar, S.V.; Gottesman, M.M.; Nussinov, R. Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. *J. Mol. Biol.* 2008, 383, 281–291. [CrossRef] [PubMed]
- 43. Gaspar, P.; Oliveira, J.L.; Frommlet, J.; Santos, M.A.; Moura, G. EuGene: Maximizing synthetic gene design for heterologous expression. *Bioinformatics* **2012**, *28*, 2683–2684. [CrossRef] [PubMed]
- 44. Komar, A.A.; Guillemet, E.; Reiss, C.; Cullin, C. Enhanced expression of the yeast Ure2 protein in Escherichia coli: The effect of synonymous codon substitutions at a selected place in the gene. *Biol. Chem.* **1998**, *379*, 1295–1300.
- 45. Inouye, S.; Sahara-Miura, Y.; Sato, J.; Suzuki, T. Codon optimization of genes for efficient protein expression in mammalian cells by selection of only preferred human codons. *Protein Expr. Purif.* **2015**, *109*, 47–54. [CrossRef]
- 46. Selik, R.M.; Linley, L. Viral Loads Within 6 Weeks After Diagnosis of HIV Infection in Early and Later Stages: Observational Study Using National Surveillance Data. *JMIR Public Health Surveill.* **2018**, *4*, e10770. [CrossRef]
- O'Doherty, U.; Swiggard, W.J.; Jeyakumar, D.; McGain, D.; Malim, M.H. A sensitive, quantitative assay for human immunodeficiency virus type 1 integration. J. Virol. 2002, 76, 10942–10950. [CrossRef]
- Agosto, L.M.; Yu, J.J.; Dai, J.; Kaletsky, R.; Monie, D.; O'Doherty, U. HIV-1 integrates into resting CD4+ T cells even at low inoculums as demonstrated with an improved assay for HIV-1 integration. *Virology* 2007, 368, 60–72. [CrossRef]