Supporting Information

Material and Methods

General. All chemicals and DNA oligomers were obtained from commercial sources and used without further purification. Unnatural amino acids (UAA*s), 4-azidophenylalanine (AzF) and 4-iodophenylalanine (IodoF), are purchased from Bachem. The synthesis of 4-acetylphenylalanine (AcF) follows the procedure that was previously reported.[1] HIV-1 clone, pSUMA, was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH, from Dr. John Kappes and Dr. Christina Ochsenbauer. Standard molecular biology techniques (Sambrook et al., 2001) were used throughout. *E. coli* DH10B and GeneHogs were used for routine cloning and DNA propagation. *E coli* Stbl-2 (Life Technologies) was use for pSUMA manipulations.

Plasmid construction

pcDNA3.1-tRNA^{Tyr}. The amber suppressor tRNA^{Tyr} was derived from *E. coli*. The G35 of the tRNA^{Tyr} was mutated to C to generate CUA anticodon. The tRNA^{Tyr} gene lacking 3'-CCA but with a 3'-TTTTTCT sequence was PCR amplified. A human U6 promoter was PCR amplified from pCMV-pylRSNBK-1[2] and added in front of the tRNA^{Tyr} gene by overlapping PCR. The U6-tRNA^{Tyr} gene cassette was PCR amplified and inserted into pcDNA3.1/hygro(+, Life Technologies) in front of the CMV promoter to afford pcDNA3.1-tRNA^{Tyr}.

pAzFRS. The AzFRS-encoding gene was PCR amplified from plasmid pSWAN-pAzpaRS[3] and inserted into pcDNA3.1-tRNA^{Tyr} behind the non-regulated CMV promoter to afford plasmid pAzFRS.

pAcFRS. The AcFRS-encoding gene was PCR amplified from plasmid pSWAN-pApaRS[3] and inserted into pcDNA3.1-tRNA^{Tyr} behind the non-regulated CMV promoter to afford plasmid pAcFRS.

pIodoFRS. The IodoFRS-encoding gene was PCR amplified from plasmid pSWAN-pIpaRS[3] and inserted into pcDNA3.1-tRNA^{Tyr} behind the non-regulated CMV promoter to afford plasmid pIodoFRS.

pEGFP-TAG40. The EGFP-encoding gene containing an amber stop codon (UAG) at a permissive site (Tyr40) was inserted into pcDNA3.1-tRNA^{Tyr} behind the non-regulated CMV promoter to afford plasmid pEGFP-TAG40.

pSUMA mutants. Overlapping PCR was used to introduce amber mutations onto the HIV-1 genome that is encoded by plasmid pSUMA. The following primers are used to introduce amber mutations:

pSUMA-Tyr132
5'-CTAGAACGATTCGCAGTTAACCCTG-3'
5'-GTTCTGCACTATAGGCTAATTTTGGCTGACCTGGCTG-3'
5'-TAGCCTATAGTGCAGAACCTCCAG-3'
5'-GTGCCTATAGCTTTGTGTCCACAG-3'

pSUMA-Ala119
5'-CTAGAACGATTCGCAGTTAACCCTG-3'
5'-TTTCCTGCGTCAGCctaTGCTTGCTGTGCTTTTTTCTTAC-3'
5'-CTGACGCAGGAAACAACAGCCAG-3'
5'-GTGCCTATAGCTTTGTGTCCACAG-3'

pSUMA-Leu365
5'-CAAAGTAAGACAATATGATCAGGTAAC-3'
5'-CACCTCTACAGATGTTCTCTCAGTTCCTC-3'
5'-CATCTGTAGAGGTGGGGATTTACCACAC-3'
5'-GCTTCCCATGTTTCTCTTTGTATG-3'

pSUMA-Tyr59
5'-CTAGAACGATTCGCAGTTAACCCTG-3'
5'-ATTTCTATGGTTACCTGATCCTATTGTCTTACTTTGATAAAAC-3'

pSUMA-Trp36Gln127
5'-GTAGAGGATCCACTAGTAAC-3'
5'-TAGCTCCCTGCTTGCCTATACTATATGTTTTAATTG-3'
5'-GGCAAGCAGGGAGCTAGAAC-3'

5’-GTGCCTATAGCTTTGTGTCCACAG-3’

pNL-GI-Tyr40. Overlapping PCR was used to introduce amber mutations onto the GFP gene that is encoded on plasmid pNL-GI.[4] The following primers are used to introduce the amber mutation:
5’-GCTATAAGACGCGTCCACCATG-3’
5’-CCTAGGTGGCATCGCCCTC-3’
5’-CGATGCCACCTAGGGAAAGCTGACCCTGAAGTTC-3’
5’-CGTCTAGATTACTTGTACAGCTCATC-3’

Protein expression and purification. 293T cells were grown in media containing DMEM, 10% FBS (v/v), and 2 mM L-glutamine at 37 ºC in a humidified atmosphere of 5% CO_2 (v/v). When cells reached 60-70% confluency, they were transfected with plasmids pEGFP-40TAG and pAzFRS using Lipofectamine™ 2000 (Life Technologies) according to the manufacturer’s protocol (36 μL Lipofectamine + 12 μg of pAzFRS + 12 μg of pEGFP-TAG40 for 12 mL cell culture in 75 cm^2 cell-culture flask). Six hours posttransfection, the culture medium was carefully removed and replaced with 12ml of fresh medium containing 1 mM AzF. Cells were grown for an additional 36 h before being washed with DPBS, lysed with RIPA buffer (Thermo Scientific), and partially purified using Ni-NTA resin (GE Healthcare) according to the manufacturer’s protocol.

Transfection and generation of live HIV-1. 293T cells were grown in a medium containing DMEM, 10% FBS, and 2 mM L-glutamine at 37 ºC in a humidified atmosphere of 5% CO_2. When cells reached 60-70% confluency, they were transfected with appropriate plasmid(s) using lipofectamine2000 (10 μl lipofectamine2000 + 3 μg of plasmid(s) for 2 ml cell culture in a 6-well plate) by following standard procedures provided by the manufacture (Life Technologies). Six hours posttransfection, the culture medium was carefully removed and replaced with 2ml of fresh medium. When it was needed, UAA* was added to a final concentration of 1 mM. After 44-48 hours of incubation, virus-containing culture supernatant was harvested by collecting the medium from the well using a pipette. The FBS concentration in the virus-containing culture medium was then adjusted to 20% (i.e. for each 1ml of harvested virus, add 0.125 ml of FBS). The virus-containing culture medium was then filtered through a 0.45-micron filter. The virus was used directly or aliquoted into sterile screw-cap vials and stored at − 80°C.

p24 assay. P24 assay was measured with Retrotek HIV-1 p24 Antigen ELISA 2.0 by following standard procedures provided by the manufacture (ZeptoMetrix Corporation). Briefly, 200 ml cell culture samples, the standard, and the control were added to the monoclonal antibody-coated micro-plate wells and incubated at 37°C for 1.5 hour. After washing and the addition of antibody-HRP conjugate, the micro-plate was incubated at 37°C for one hour. Plates were then washed and the HRP-substrate was added for color development. The reactions were stopped by the addition of 1 M H_2SO_4 and absorbance values were determined at 450 nm. The amount of p24 is determined by interpolation from a point-to-point plot or from a linear regression analysis of the standard curve.

Infection assay. HIV-1 infection was quantified with X-gal staining based assay with TZM-bl cells. Briefly, virus (generated in the presence of UAA*), 50 ml 10% DMEM growth medium, and DEAE-dextran at a final concentration of 40 mg/ml were added to each well (96-well flat bottom plate) that contains TZM-bl cells (1×10^4 per well in 100 ml volume) in triplicate. Assay controls included: (1) replicate wells of TZM-bl cells with virus that are generated in the absence of UAA*; (2) growth medium only control; and (3) TZM-bl cell only control. After 48-hour incubation at 37°C, 200 ml of assay medium was removed from each well. Plates were then washed, fixed, and stained using X-gal solution followed by examination using a light microscope.

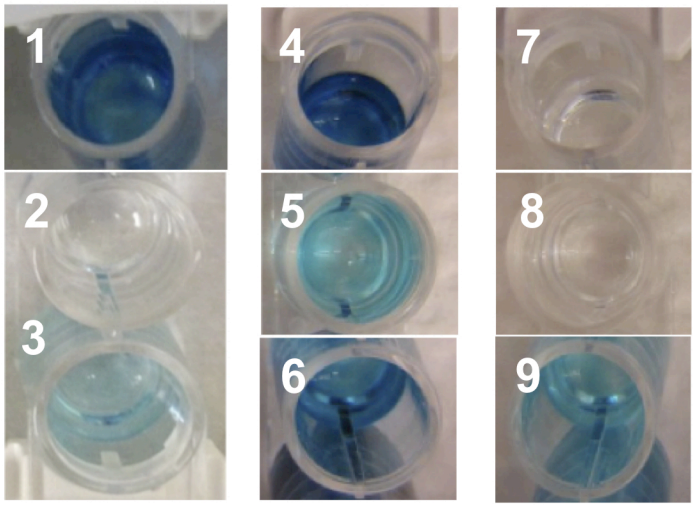
Tissue Culture Infectious Dose 50 (TCID50). Infectious titers of all viruses were determined by standard Tissue Culture Infectious Dose 50 (TCID50) method with X-gal staining assay (see above) in TZM-bl cells. Briefly, four-fold serial dilutions of virus were performed in quadruplicate (96-well flat bottom plate) in the X-gal staining assay. The TCID50 of the virus was calculated by Spearman-Karber formula according to the negative end-point and dilution folds.

Fluorescence spectroscopy and cell imaging. The fluorescent images and bright-filed images were taken by a Nikon ECLIPSE TE3000 microscope and an EVOS FL Auto Imaging System with DIC. The cells were excited at 488 nm to acquire EGFP fluorescence images at 530/25 nm.

LC/MS/MS. The corresponding protein band of EGFP-40TAG from SDS-PAGE was cut and in-gel digested with trypsin (in 50 mM ammonium bicarbonate, pH 8.0) overnight at 37 °C. The resulting peptide fragments were extracted with 0.1% formic acid/75% acetonitrile, and then

subjected to LC/MS/MS analysis using a Waters Q-TOF Ultima. Database searches were performed on an in-house Mascot server (Matrix Science Ltd., London, UK). For AzF substitution site mapping on EGFP, AzF substitution for tyrosine was included as a variable modification.

Supplemental Figures.



Normalized OD values of above p24 assays

well 1	well 4	well 7
6.181	7.262	0.003
well 2	well 5	well 8
0.000	1.364	-0.009
well 3	well 6	well 9
0.785	3.201	1.912

Figure S1. **p24 assay after transfection of 293T with pSUMA variants.**
1, wild-type pSUMA control; **2**, negative ELISA control, no p24; **3**, positive ELISA control, 125 pg/ml p24; **4**, pSUMA-Tyr132 + $tRNA^{Tyr}$-AcFRS pair, with 1mM AcF; **5**, pSUMA-Ala119 + $tRNA^{Tyr}$-AcFRS pair, with 1mM AcF; **6**, pSUMA-Leu365 + $tRNA^{Tyr}$-AcFRS pair, with 1mM AcF; **7**, pSUMA-Tyr132 + $tRNA^{Tyr}$-IodoFRS pair, with 1mM IodoF; **8**, pSUMA-Ala119 + $tRNA^{Tyr}$-IodoFRS pair, with 1mM IodoF; **9**, pSUMA-Leu365 + $tRNA^{Tyr}$-IodoFRS pair, with 1mM IodoF. Well-2 was used as blank for p24 assays.

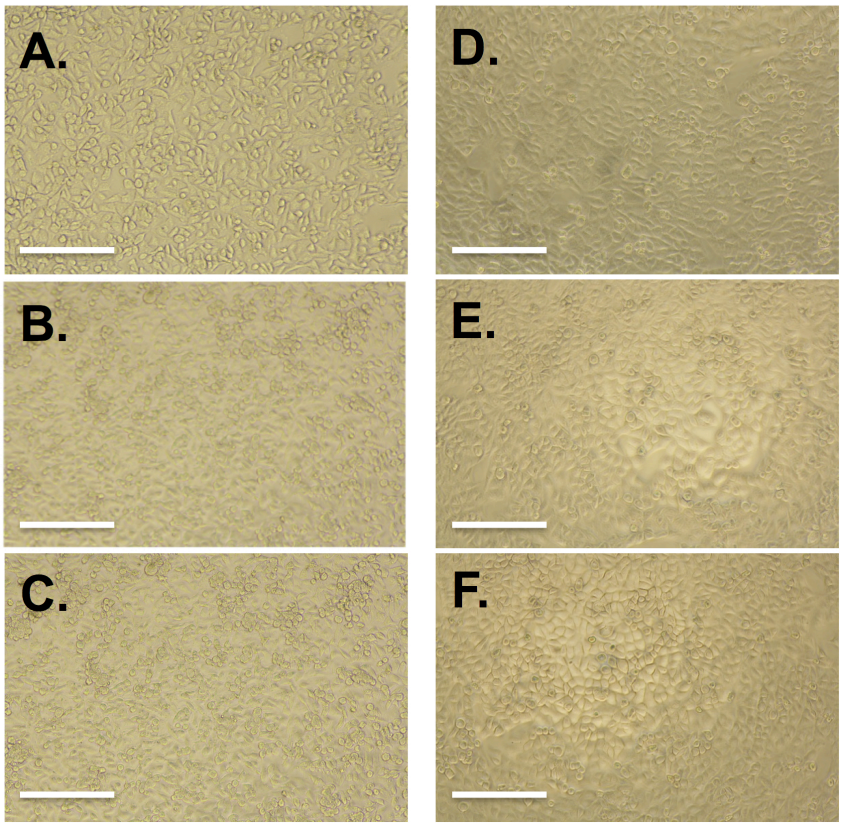


Figure S2. **Infection assays with TZM-bl cells using tRNA^{Tyr}-AcFRS pair and tRNA^{Tyr}-IodoFRS pair.**
(A) infected with virus collected from pSUMA-Tyr132 + tRNA^{Tyr}-AcFRS pair + 1mM AcF experiment; (B) infected with virus collected from pSUMA-Ala119 + tRNA^{Tyr}-AcFRS pair + 1mM AcF experiment; (C) infected with virus collected from pSUMA-Leu365 + tRNA^{Tyr}-AcFRS pair + 1mM AcF experiment; (D) infected with virus collected from pSUMA-Tyr132 + tRNA^{Tyr}-IodoFRS pair + 1mM IodoF; (E) infected with virus collected from pSUMA-Ala119 + tRNA^{Tyr}-IodoFRS pair + 1mM IodoF experiment; (F) infected with virus collected from pSUMA-Leu365 + tRNA^{Tyr}-IodoFRS pair + 1mM IodoF experiment. Scale bars, 200 μm.

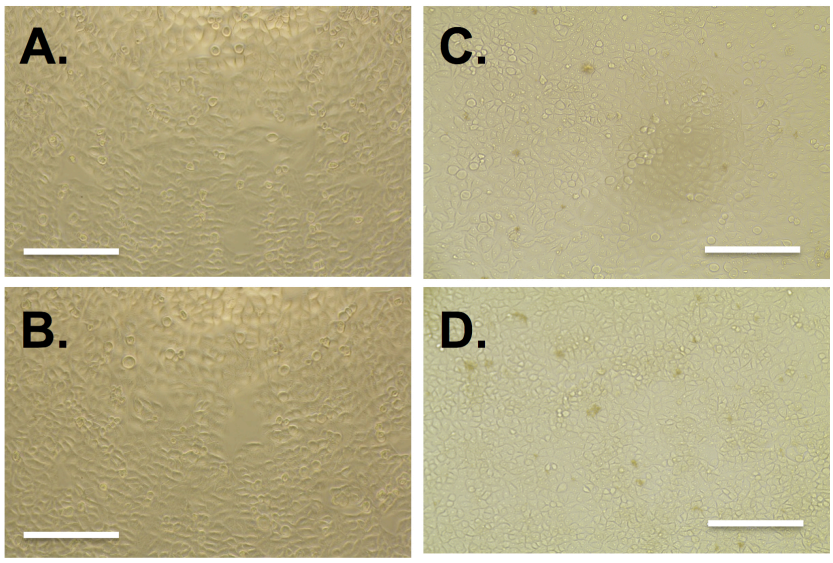


Figure S3. **Infection assays with TZM-bl cells using tRNA^{Tyr}-AzFRS pair.**
(A) infected with virus collected from pSUMA-Ala119 + tRNA^{Tyr}-AzFRS pair, without 1 mM AzF experiment; (B) infected with virus collected from pSUMA-Leu365 + tRNA^{Tyr}-AzFRS pair, without 1mM AzF experiment; (C) Infected with pSUMA-Tyr59 + tRNA^{Tyr}-AzFRS pair, without 1 mM AzF experiment (the harvested viruses were concentrated 15 folds); (D) Infected with pSUMA-Trp36Gln127 + tRNA^{Tyr}-AzFRS pair, without 1 mM AzF experiment (the harvested viruses were concentrated 15 folds). Scale bars, 200 μm.

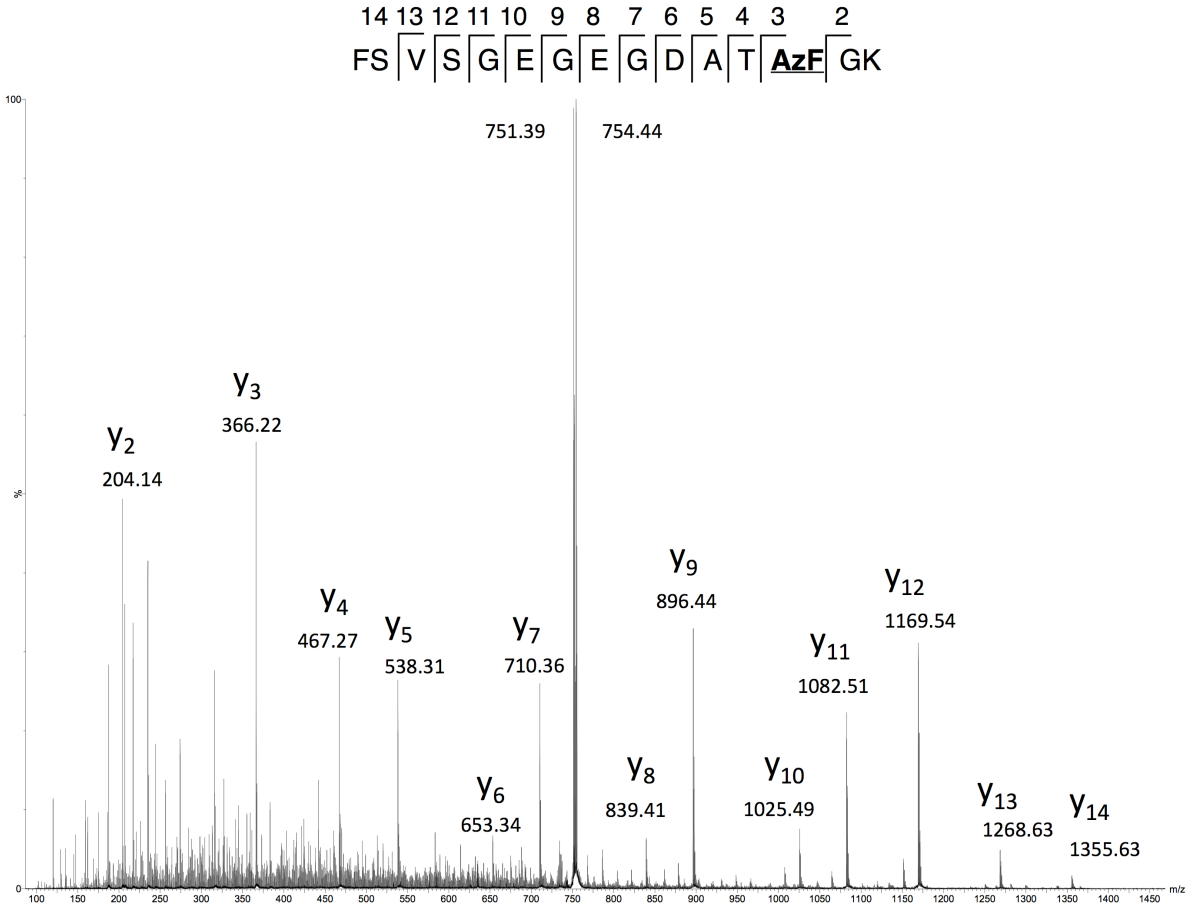


Figure S4. **Mass spectrometry analysis of EGFP containing AzF at position 40.** The y ions are marked in the spectrum. The amino acid sequence of the peptide fragment, FSVSGEGEGDAT**AzF**GK, from mutant EGFP containing AzF is shown on top. The amber mutation site contained exclusively 4-aminophenylalanine (aminoF), which is the reduction product of AzF. This observation is consistent with previous reports on the mass spectrometry analyses of AzF-containing proteins.[3,5]

References.

[1] L. Wang, Z. Zhang, A. Brock, P. G. Schultz, *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 56-61.

[2] P. R. Chen, D. Groff, J. Guo, W. Ou, S. Cellitti, B. H. Geierstanger, P. G. Schultz, *Angew. Chem. Int. Ed.* **2009**, *48*, 4052-4055; *Angew. Chem.* **2009**, *121*, 4112–4115.

[3] W. Liu, A. Brock, S. Chen, S. Chen, P. G. Schultz, *Nat. Methods* **2007**, *4*, 239-244.

[4] K. L. Collins, B. K. Chen, S. A. Kalams, B. D. Walker, D. Baltimore, *Nature* **1998**, *391*, 397-401.

[5] J. W. Chin, T. A. Cropp, J. C. Anderson, M. Mukherji, Z. Zhang, P. G. Schultz, *Science* **2003**, *301*, 964-967.