

Research Article

Open Access

# Stampidine as a Potent Epigenetic Silencer of Host HIV Dependency Factor Genes in HIV-Infected Cells

Sanjive Qazi<sup>1,2,3</sup> and Fatih Uckun<sup>2,3,4\*</sup>

<sup>1</sup>Bioinformatics Program, Department of Biology, Gustavus Adolphus College, St. Peter, MN, USA

<sup>2</sup>Department of Virology and Drug Discovery Program, Parker Hughes Institute, St. Paul, MN 55455, USA

<sup>3</sup>Developmental Therapeutics Program, Children's Hospital Los Angeles, Children's Center for Cancer and Blood Diseases, Los Angeles, CA 90027, USA

<sup>4</sup>Department of Pediatrics, University of Southern California, Keck Medical Center, Los Angeles, CA 90027, USA

## Abstract

Here we report that our lead anti-retroviral (ARV) drug candidate Stampidine [2',3'-didehydro-3'-deoxythymidine-5'-(p-bromophenyl methoxy alaninyl phosphate)] results in methylation of a network of HIV-responsive regulatory genes in T-cells, including several genes for HIV-dependency factors (HDFs). Stampidine epigenetically modulates the host transcriptome in a unique manner, silences expression of a distinct set of genes encoding transcription factors and signal transduction molecules, and prevents HIV infection from distorting and disrupting key cellular transcriptional networks. At nanomolar concentrations that are 4-logs lower than those achieved at its non-toxic dose levels in mice, rats, cats, and dogs, Stampidine switched off genes for several HDFs that are required for HIV replication in T-cells. Notably, Stampidine reversed the effects of HIV exposure on the host transcriptome regardless of NRTI-sensitivity or RT mutations of the HIV isolate used and inhibited the replication of 17 NRTI-resistant HIV-1 strains, including recombinant HIV clones containing common patterns of RT mutations responsible for NRTI resistance, in human peripheral blood mononuclear cells (PBMC) with subnanomolar-nanomolar IC<sub>50</sub> values (Mean ± SEM = 12.0 ± 3.2 nM). Unlike available ARV agents that disrupt a specific step in the life-cycle of HIV, Stampidine has the potential to abrogate all steps in the life cycle of HIV.

## Introduction

Human Immunodeficiency Virus type 1 (HIV-1) infection remains a global health concern affecting millions of individuals worldwide [1-3]. Combination antiretroviral therapy has become the standard of care for patients with HIV infection [1-3]. Antiretroviral (ARV) treatment regimens employing combinations of drugs from at least two of the three classes of ARV therapy, namely Nucleoside Analogue RT Inhibitors (NRTI), Non-Nucleoside Analogue RT Inhibitors (NNRTI), and protease inhibitors, exhibit a potent and sustained antiviral effect and confer consistent long-term viral suppression in patients with HIV infection [1-3]. However, each of these drugs can select for drug-resistant HIV strains and the emergence of ARV drug resistance limits their clinical benefit [4-7]. Acquired resistance to ARV agents hampers the long-term success of contemporary highly active antiretroviral therapy (HAART) regimens [4-7]. Likewise, pre-exposure prophylaxis (PrEP), an evolving new approach to HIV prevention in which ARV agents are used prior to potential HIV exposure in an attempt to reduce the likelihood of HIV infection post exposure in the context of unprotected heterosexual intercourse [8-10], has encountered difficulties in part due to ARV drug resistance [11-13]. Therefore, there is an urgent need for potent anti-HIV agents that are active against ARV drug-resistant HIV strains.

Viruses can only evolve in those host cells that provide the essential microenvironment required for their life cycle [14-18]. Genome-wide small interfering RNA (siRNA)-mediated single gene knockdown studies have shown that HIV exploits several host proteins, also known as HIV-dependency factors (HDFs) during multiple steps of infection, including viral entry, viral integration, and viral transcription [14,15,17,18]. Our work has therefore focused on identifying new agents capable of both preventing and treating HIV infection by leveraging the dependency of HIV on HDFs as well as the viral RT enzyme for infecting and replicating in human cells. Our lead ARV drug candidate Stampidine (5'-(4-bromophenyl methoxy)laninylphosphate]-2',3'-didehydro-3'-deoxythymidine) (CAS 217178-62-6), a rationally designed novel aryl phosphate derivative of the NRTI stavudine (STV)/d4T (CAS 3056-17-5) that is being developed as a promising new drug

candidate against ARV-resistant HIV [3,19], results in methylation of a network of HIV-responsive regulatory genes, including several genes for HDFs. Stampidine epigenetically modulates the host transcriptome in a unique manner, which prevents HIV infection from distorting and disrupting key cellular transcriptional networks. At nanomolar concentrations that are 4-logs lower than those achieved at its non-toxic dose levels in mice, rats, cats, and dogs, Stampidine switched off genes for HDFs that are required for HIV replication. Notably, Stampidine reversed the effects of HIV exposure on the host transcriptome regardless of NRTI-sensitivity or RT mutations of the HIV isolate used and inhibited the replication of 17 NRTI-resistant HIV-1 strains in human peripheral blood mononuclear cells with subnanomolar-nanomolar potency. Stampidine is the first-in-class dual-function ARV agent with both NRTI activity and epigenetic anti-host transcriptome activity for PrEP as well as for treatment of acute or chronic HIV infections. Unlike available ARV agents that disrupt a specific step in the life-cycle of HIV, Stampidine has the potential to abrogate all steps in the life cycle of HIV.

## Materials and Methods

### Anti-retroviral (ARV) potency assays

17 HIV-1 strains were used in Stampidine ARV potency assays, including 11 clinical HIV-1 isolates with NRTI resistance and 6

\*Corresponding author: Fatih M. Uckun, M.D., Ph.D, Children's Hospital Los Angeles, MS#160, Los Angeles, California 90027-0367, USA, Tel: (323)-361-4328; Fax: 323-361-8765; E-mail: [fmuckun@chla.usc.edu](mailto:fmuckun@chla.usc.edu)

Received March 05, 2012; Accepted April 13, 2012; Published April 17, 2012

Citation: Qazi S, Uckun F (2012) Stampidine as a Potent Epigenetic Silencer of Host HIV Dependency Factor Genes in HIV-Infected Cells. J AIDS Clinic Res 3:147. doi:10.4172/2155-6113.1000147

Copyright: © 2012 Qazi S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

prototypical multidrug-resistant recombinant HIV-1 clones containing multiple NRTI resistance mutations (viz.: NIH-7324-1, NIH-7295-1, NIH-52534-2, NIH-1617-1, NIH-4755-5, NIH-56252-1) obtained from the NIH AIDS Research and Reference Reagent program [19,20]. For each HIV strain, RT mutations for up to 6 amino acid positions were queried for NRTI resistance profiles using the web tool on the Stanford University Drug Resistance Database ([http://hivdb.stanford.edu/cgi-bin/RT\\_Phenotype.cgi](http://hivdb.stanford.edu/cgi-bin/RT_Phenotype.cgi)). The database harbors anti-viral profiles for 1131 patterns of major drug resistance mutations from 52,867 clinical isolates obtained from 47,545 persons infected with HIV-1. Fold resistance values for these mutant strains compared to wildtype virus were retrieved from both Antivirogram (Virco, Belgium) and Phenosense (Monogram, South San Francisco) susceptibility tests. The susceptibility of HIV-1 strains to Stampidine was tested according to previously described methods [19]. The  $IC_{50}$  values were calculated using nonlinear regression modeling of the exponential form of the linearized equation.  $IC_{50}$  values were  $\log_{10}$  transformed to homogenize within group variances and normalize data distributions.

### Gene expression profiling

Three HIV-1 strains from NIH AIDS Research and Reference Reagent Program, namely RT-MDR (Catalog No. 252), HTLV-IIIB (Catalog No. 398), and BR92019 (Catalog No. 1778) were used in gene expression profiling experiments. MT-2 T-cells cultured in RPMI 1640 supplemented with 15% fetal bovine serum were infected with one of these 3 HIV-1 strains at a multiplicity of infection (MOI) of 0.01. RNA was extracted from HIV-infected MT-2 cells at 24h, 48h, or 7 days after HIV-infection using previously reported standard procedures [21]. In brief, total RNA was prepared from cells infected with virus using TriPure isolation reagent (Boehringer Mannheim). DNA was removed from this RNA preparation by DNase treatment (RQ1 RNase-Free DNase, Promega) and phenol chloroform extraction. Further purification of RNA was achieved using an RNeasy column prior to *in-vitro* transcription reaction. Total RNA was quantified using standard spectrophotometric methods. First strand cDNA was synthesized with a T7-(dT)<sub>24</sub> primer and second strand cDNA was synthesized with *E coli* DNA polymerase I and ligase. The cDNA product was *in-vitro* transcribed and labeled with the Enzo BioArray High Yield RNA transcript labeling kit. The *in vitro* transcription products were purified on an RNeasy spin column prior to fragmentation in buffer containing 40 mM Tris-Acetate (pH 8.1), 100 mM KOAc and 30 mM MgOAc for 35 minutes at 94°C. We measured expression levels of gene transcripts using an U95 GeneChip microarray system from Affymetrix, which interrogates the expression levels of 12,625 genes, using our previously reported procedures [21]. Biotinylated fragmented cDNA was hybridized to an Affymetrix U95 array for 15 h prior to wash and staining, as per manufacturer's recommendations. Arrays were scanned with a confocal scanner manufactured for Affymetrix by Agilent. All HIV-infected samples were compared to the uninfected control samples (time zero) at each time point using comparisons for pairs of gene chips performed on the same day of experimentation. We compared gene expression levels using a T-test (Unequal variance correction) on  $\log_{10}$  transformed signal values as reported [22-24]. The Signal Difference metric was obtained as a weighted mean (Tukey's Biweight) by determining the difference in the Signal of the HIV-infected sample relative to the uninfected control sample. Analysis of variance (ANOVA) models, as implemented in the Genesight Software (Biodiscovery Inc, El Segundo), were used to screen for genes differentially affected by Stampidine treatment in the absence or presence of HIV infection across the 3 time points. This analysis identified 3484 transcripts that were most significantly affected using

the signal difference metric ( $p < 0.05$ ; False Discovery Rate = 18%). We also examined the effects of Stampidine alone without HIV infection. MT-2 cells were treated with 0 nM (N = 8) or 1-1000 nM (N = 17) Stampidine for 24 hrs, 48 hrs, or 7 days.

In order to evaluate the dynamic changes in expression levels of transcriptional factor genes as well as genes for signal transduction molecules, we used the NetAffx.com web site and searched for oligos that contained the string 'transcription' (1220 gene transcripts) as well as 'receptor signal transduction' (853 gene transcripts) on the Affymetrix U95 chipset. The transcript levels for transcription factor genes were statistically filtered using the following criteria to identify the statistically significant HIV effects: 1. Sum of the absolute values at 24-h, 48-h and 7-day for HIV- induced changes > 30 (556 genes); 2. ANOVA p-value for the virus effect on signal difference values < 0.05 (229 genes, False discovery rate = 12.2%, this set of genes was used for the correlation analysis). Significantly affected gene transcripts across the three time points were used in the correlation-cluster analysis to characterize network relationships. Samples were pooled from the HIV-infected no drug control group (treatment with all viral strains at all time points (N = 11) and correlations were performed between all pairs of genes. The correlation coefficient (r) matrix was loaded into the Genesight software for analysis using the Hierarchical Clustering method performed on both rows and columns (Complete Linkage using Pearson Correlation as the Distance metric). The network structure was deduced from examination of the clustered correlation matrix. Data from the Stampidine treatment groups were pooled at all time points, (N = 17) and a correlation analysis was performed between all gene pairs and correlations were ranked according to the Pearson correlation coefficient (1 = perfect positive correlation, 0 = no correlation, -1 = negative correlation). Most correlated gene pairs were ranked for each of the two functional classes with p-values of < 0.001, 0.1, and 0.05. The number of connections per gene was determined for correlations with p-values less than 0.001 for each functional class. To identify gene pairs that showed the greatest difference in correlation coefficient values between no drug treatment controls and Stampidine treatment samples, t-tests were performed following Fishers r to z-transformation ( $z = 0.5 * \ln((1+r)/(1-r))$ ; standard error of mean =  $1/(N-3)$ , where N = numbers of data points). Correlation coefficients (r) were determined between all gene pairs and Hierarchical Cluster analysis (Complete Linkage, Pearson correlation as Distance metric) was applied to the matrix of correlation coefficients (229 rows, 229 columns with 'r' values) both rows and columns of gene identifications. Similar analyses were performed for HIV-responsive genes encoding signal transduction molecules (232 genes out of 853 genes). We evaluated the dynamic effect of virus at early time points between 24 and 48 hours. All virus treated samples were compared to the no-virus control at each time point using the Signal Difference metric in the absence and presence of Stampidine performed on the same day. This ensured specific contrasts of virus effects controlling for baseline changes in control levels of expression that changed from 24 to 48 hours or the effects of Stampidine alone at each time point. Dynamic changes in gene expression induced by virus was calculated by comparing the Signal Difference metric for 4 virus samples (2 RT-MDR, 1 for BR92019 and 1 for HTLV-IIIB subtypes) at 24 hours with 4 virus treated samples at 48 hours. T-test (unequal variances) identified 1992 differentially affected transcripts of which 837 were up regulated at 48 hours compared to 24 hour treatment with virus. Signal Difference values for virus treated samples in the presence of 1 nM Stampidine was compared at 48 hours (7 samples pooled from 2 samples for BR92019, 2 samples for HTLV- IIIB and 3 samples for RT-MDR) versus 24 hours (5 samples pooled from 1 sample for BR92019, 1

sample for HTLV-IIIB and 3 samples for RT-MDR). Fold difference and P-values (T-test Unequal variances) for the virus effect in the presence of Stampidine was reported for each transcript that showed significant up regulation in the absence of Stampidine. The differentially expressed genes that were downregulated by Stampidine in HIV-infected cells were cross-referenced to the HDF genes identified in 3 separate siRNA-based genomic screens [14,17,18]. The effects of Stampidine on DNA methylation levels in T-cells was determined using 3 different T-cell lines (LOUCY, CEM, MOLT3). We utilized the Infinium HumanMethylation450 (HM450) BeadChip (Illumina) [25,26] to interrogate methylation levels for more than 485,000 methylation sites at the single nucleotide levels covering 99% of RefSeq genes with an average of 17 CpG sites per gene region distributed across the promoter, 5'UTR, first exon, gene body, and 3'UTR, and 96% of CpG islands, with additional coverage in island shores and the regions flanking them. The data were generated at the USC Epigenome Center. Data points were reported as beta values, which is a ratio of the signal intensity of the methylated probe type (M) over the sum of the methylated and unmethylated probe signal intensities (M+U) and a corresponding detection p-value calculated by comparing the signals from each

analytical probe to a panel of negative control probes on the array. Data points with corresponding detection p-value >0.05 were masked as "NA" in the data set. Beta-values exhibited marked heteroscedasticity in the low and high methylation range necessitating the transformation of beta-value, which corresponds to % methylation, to M-value [27] using logistic function  $\text{Log}_2(\text{beta}/(1-\text{beta}))$ . M-values exhibited a bimodal distribution that correspond to a positive mode (methylated mode) and a negative mode (unmethylated mode) with statistical properties to allow T-test comparisons of no treatment control versus Stampidine treatment for each cell line. We performed a paired T-test on the difference of M-values between control and Stampidine-treated samples for the 3 cell lines and filtered methylation sites exhibiting an increase in methylation in the presence of Stampidine (>5% increase in beta values) and p-value <0.05 (Paired t-test for 3 cell lines).

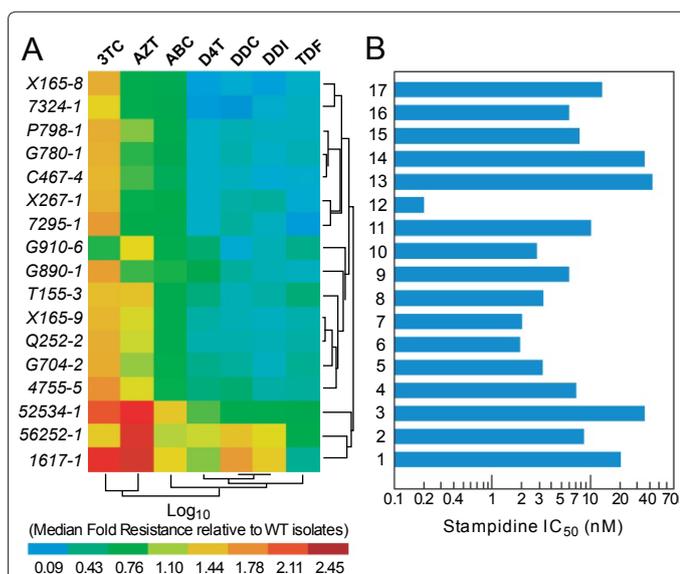
## Results and Discussion

### Broad-spectrum anti-HIV activity of Stampidine against HIV isolates/clones with complex drug resistance profiles

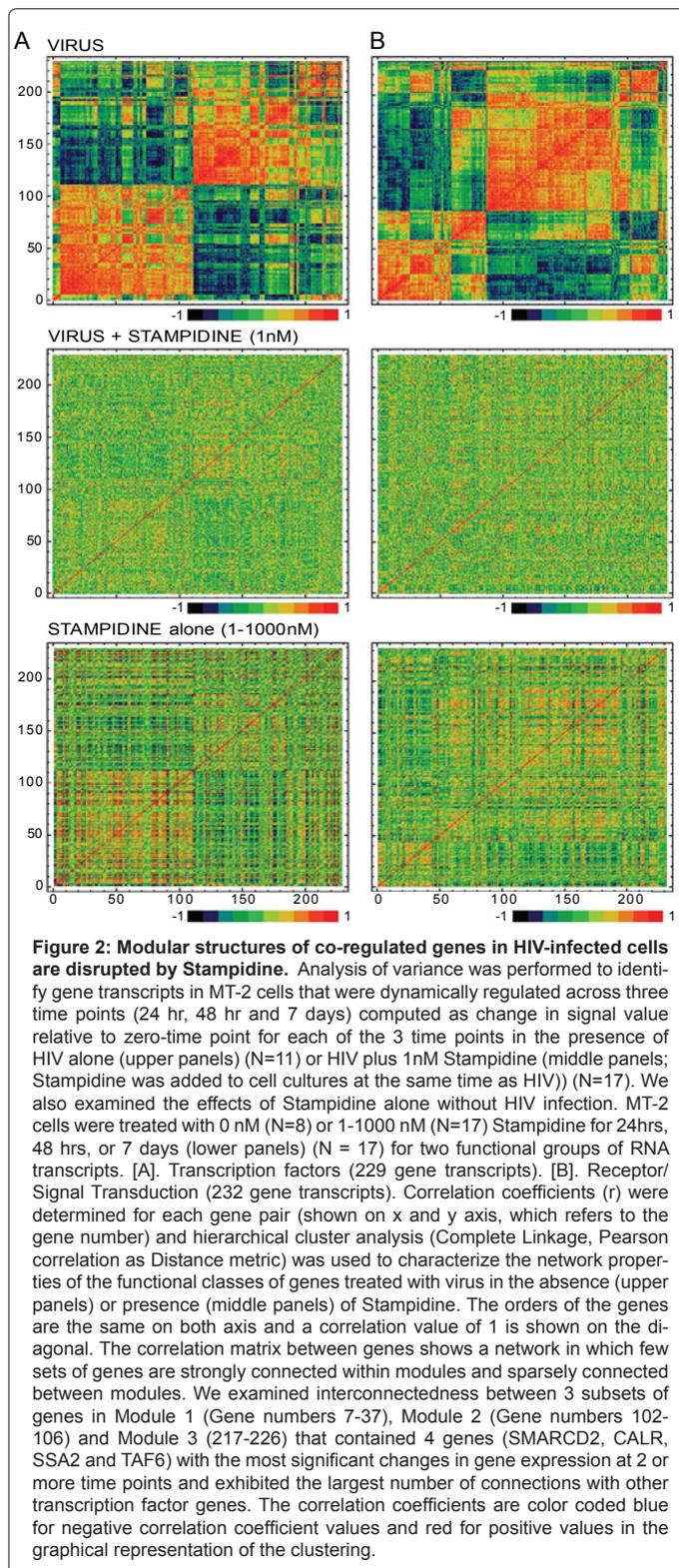
Stampidine exhibited nanomolar-subnanomolar potency against ARV drug-resistant HIV-1 strains (Figure 1). Two-way cluster analysis of the Fold Resistance data retrieved from the Stanford Database for HIV isolates showed two clusters of NRTI compounds according to the NRTI-response profiles of 17 ARV drug-resistant HIV strains, one for a group of two drugs: AZT and 3TC, and a second group of 4 drugs: D4T, DDC, DDI and TDF (Figure 1A). Stampidine exhibited nanomolar potency against each of the 17 HIV-1 strains ( $\text{IC}_{50} = 0.2\text{-}42$  nM; Mean = 12.0 nM, SEM = 3.2 nM) even though there were marked differences in fold resistance values between the 7 NRTIs for the 17 HIV-1 strains (Figure 1B, Table S1). Fourteen out of the 17 strains of HIV demonstrated greater than 30 median fold resistance values for 3TC, 10 strains showed greater than 10 median fold resistance values for AZT. Two strains of HIV-1 (1617-1 and 52534-1) showed median fold resistance values greater than 100 for 3TC and AZT and greater than 10 fold resistance for 6 NRTIs (Table S1). Stampidine maintained potency of 20 and 35 nM for 1617-1 and 52534-1 respectively, both harboring mutations in NRTI-resistance positions K70, V75, F77 and Q151. A cluster of 7 HIV-1 strains (G190-6, G890-1, T156-3, X165-9, Q252-2, G704-2, 4755-5) containing common mutations in positions M41, D67, L210 and T215 in at least 5 of the HIV-1 strains exhibited  $\text{IC}_{50}$  values for Stampidine less than 10 nM. Stampidine was most effective against X267-1 strain ( $\text{IC}_{50} = 0.2$  nM) that exhibited greater than 2 median fold resistance values against 5 NRTIs (Table S1).

### Stampidine prevents HIV infection from distorting and disrupting key cellular transcriptional networks to enable HIV replication in T-cells

By using U95 GeneChip microarrays from Affymetrix, gene expression profiling of MT-2 T-cells was performed before and at 24 h, 48 h, and 7 days after infection with one of 3 different HIV strains (viz.: 92BR019, HTLV-IIIB, RT-MDR). ANOVA identified HIV-induced changes across the three time-points in expression levels of 229 gene transcripts representing 209 genes annotated for transcription factors. Correlation-Cluster analysis was used to characterize the connectivity between these significantly affected transcript levels (Figure 2, Table S2). Among 26,106 transcription factor gene pairs, 649 showed highly significant HIV-induced correlations ( $p < 0.001$ , R-squared > 0.71; False Discovery Rate = 4.0%) and 190 transcription factor gene transcripts intercorrelated with  $p < 0.001$  (Figure 2A). Eleven members in this network of transcription factor genes were downregulated when



**Figure 1: Stampidine Inhibits NRTI-Resistant HIV with Subnanomolar/Nanomolar Potency.** **A:** Dendrograms were drawn to illustrate similar Fold Resistance profiles from joining pairs of closely related median Fold Resistance values profiles, whereby drug and HIV isolates joined by short branch lengths showed most similarity. Mutation combination for each of the HIV-1 isolates/clones was cross-referenced with drug susceptibility results of sequences matching input query for known NRTI Drug Resistance Mutations on the Stanford University HIV Drug Resistance Database. Fold resistance values were compiled for all HIV-1 isolates/ clones against 7 nucleoside/nucleotide analog drugs (Abacavir (ABC); Didanosine (DDI); Zalcitabine (DDC); Lamivudine (3TC); Stavudine (D4T); Tenofovir (TDF); Zidovudine (ZDV)), and the median value was depicted using a heat map for each isolate (rows) and each drug (column) following a 2-way agglomerative hierarchical clustering technique to organize expression patterns using the Average distance linkage method such that HIV isolates having similar Fold Resistance values across drug profiles and across HIV isolates were grouped together. Dendrograms were drawn to illustrate similar Fold Resistance profiles from joining pairs of closely related median Fold Resistance values profiles, whereby drug and HIV isolates joined by short branch lengths showed most similarity. Stampidine exhibited potent anti-HIV activity against each isolate with subnanomolar-low nanomolar  $\text{IC}_{50}$  values. **B:** Anti-HIV potency of Stampidine for each of the HIV-1 isolates (virus isolates are ordered vertically as shown in panel A) is illustrated using a bar chart of the  $\text{IC}_{50}$  values (log scale).



uninfected cells were treated with Stampidine (Table S2A). Likewise, ANOVA identified 232 gene transcripts representing 196 genes annotated for signal transduction molecules changed across the three time-points in HIV-infected cells. Of 26,796 gene pairs for signal transduction molecules, 596 showed highly significant HIV-induced

correlations ( $p < 0.001$ , False Discovery Rate = 4.5 %) and 189 gene transcripts for signaling molecules exhibited inter-correlations with  $p < 0.001$  (Figure 2B). 22 members in this network of signal transduction related genes were downregulated when uninfected cells were treated with Stampidine (Table S2B). The correlation coefficient (r) matrix was loaded into Genesight software for analysis using Hierarchical Clustering method (Complete Linkage using Pearson Correlation as the Distance metric) performed on both rows and columns. Figure 2 shows the graphical representation of the clustered correlation matrix and highlights the clear modular organization of the connectivity with 3 large modules that showed positive correlations. The network structure could be deduced from examination of the clustered correlation matrix. Notably, the modular connectivity structure of HIV-responsive genes for transcription factors or signal transduction proteins was abolished in the presence of 1 nM Stampidine (Figure 2A and 2B, middle panels). Stampidine treatment up to 1000 nM did not exhibit modular correlation clustering of the HIV induced genes (Figure 2, lower panels) indicating that the widespread effect of Stampidine in the presence of virus was due to disruption of a few key hubs of the gene interaction network formed by HIV and this disruption was only in the presence of HIV.

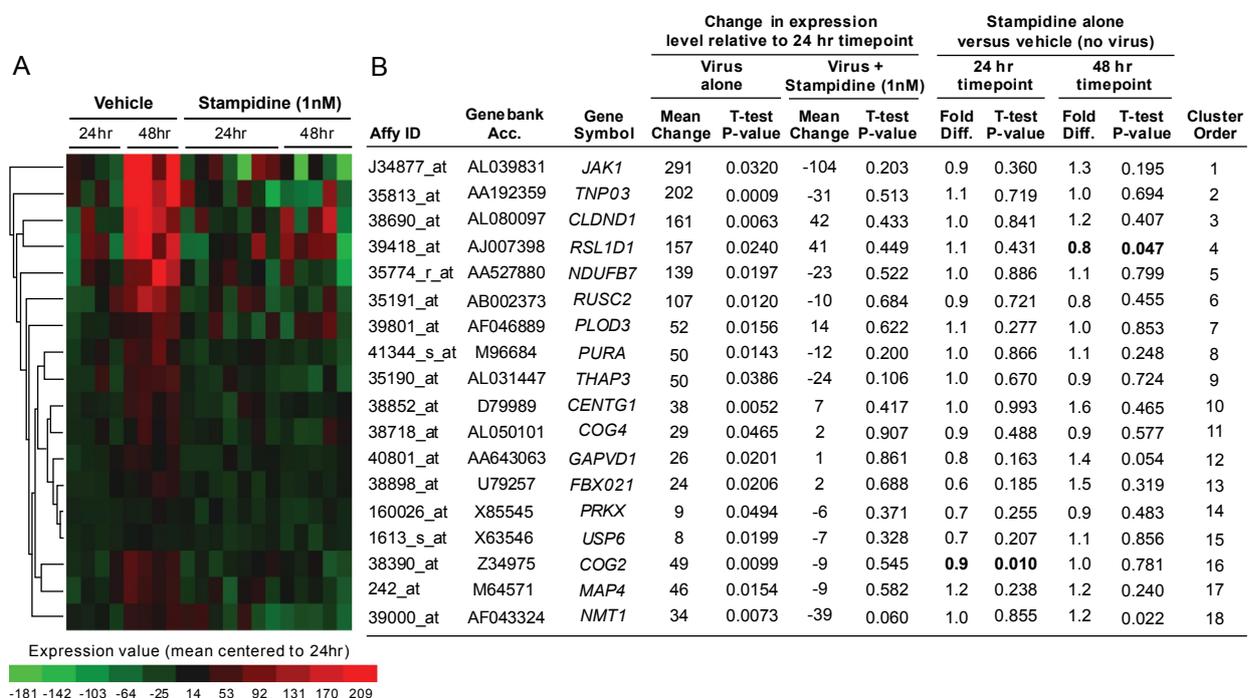
Within the modular network of transcription factor genes, members of each module showed a high degree of strong positive correlations with each other (Table S2C). Four genes that were significantly affected by virus treatment at multiple time points were found to be members of three modules (*TAF6* and *CALR* (Module 1), *SSA2* (Module 2) and *SMARCD2* (Module 3). *TAF6* was significantly affected by viral treatment at 24 hours ( $p=0.039$ ), 48 hrs ( $p=0.0027$ ) and 7 days ( $p=0.0092$ ); *CALR* was significantly affected at 24 hours ( $p=0.039$ ) and 7 days ( $p=0.019$ ); *SSA2* (*35295\_g\_at*) was significantly affected at 24 hr ( $p=0.0081$ ) and 48 hours ( $p=0.004$ ); and, *SMARCD2* (*34695\_at*) was affected at 24 hours ( $p=0.0039$ ) and 7 days ( $p=0.009$ ). These 4 genes were among the most connected (correlation coefficients with  $p < 0.001$ ) of the 190 inter-correlated genes (24<sup>th</sup> most connected), 16 (7<sup>th</sup> most connected), and 12 (12<sup>th</sup> most connected) significant correlations were observed with *SMARCD2*, *TAF6*, *CALR* and *SSA2* respectively). *ILF3* was the most connected gene exhibiting 35 significant correlations with other transcription factor transcripts. Members of each module showed a high degree of strong positive correlations with each other. Besides *TAF6* and *CALR*, there were 29 members of Module 1 with 465 connections (390 correlations showed R-squared greater than 0.36 ( $P < 0.05$ ; False Discovery rate = 5.9%), and 317 of these connections were significantly reduced by Stampidine treatment ( $p < 0.05$ ; False Discovery rate = 5%). Two connections showed r-values of 0.95: *NSEP1* with *ILF3*, and *HTATSF1* with *KLF5*. For Module 2 correlations, 9 out of 10 connections showed R-squared greater than 0.36 ( $p < 0.05$ ) and 8 of these connections were affected by Stampidine treatment. The strongest correlation occurred between *ENPP2* and *SURB7* ( $r=0.91$ ). In Module 3, there was a smaller number of significant correlations (17 out of 45 correlation coefficients were significant with  $p < 0.05$ ); but a high proportion of the significant correlations that were detected in the presence of HIV alone were affected by Stampidine treatment (15 out of 17). *SMARCD2* had significant connections with all the members of Module 3. *E2F4* was most strongly correlated with *PDEF* ( $r=0.87$ ). Detailed examination of the inter-module connections revealed three candidate genes as major communication points between modules: *SURB7*, *PFDN5* and *ENPP2*. The strongest negative correlation for change in expression induced by HIV was observed between *PFDN5* and *HTATSF1* (-0.92) and the strongest positive correlation was observed with *SURB7* and *RXR8* (0.83). The effect of Stampidine

on the major inter-modal connections is shown in Table S2D and Figure S1. Correlation coefficients were calculated in the presence or absence of Stampidine (significant correlations for HIV in the presence of Stampidine (n = 31) were R-squared >0.32 (r = ± 0.57), p<0.001; R-squared (r = ±0.46) > 0.21, p<0.01; R-squared >0.126 (r = ±0.35), p<0.05). Correlation coefficients for all inter-module connections between Module 1 and Module 2 were significantly reduced in the presence of Stampidine, except for the *ILF3-SURB7* connection for which a significant r-value (r = 0.41) was maintained in the presence of Stampidine (p<0.05). For Module 1 versus 3 connections, 8 out the 10 negative correlations were significantly affected by Stampidine treatment. The *KLF5-PFDN5* connection in HIV-infected samples was reversed from a significantly negative correlation in the absence of Stampidine (r = -0.89, p<0.001) to a significantly positive correlation in the presence of Stampidine (r = 0.44, p<0.05). Likewise, 5 out of 7 connections between Module 2 and 3 were significantly affected by the addition of Stampidine.

Notably, there was a significant over-representation of 18 genes (Fisher's Exact, 2-tailed, p = 0.020) identified in the Brass siRNA screen of HDF genes critical for HIV replication [14] among the 540 HIV-responsive genes in our study that were upregulated after HIV infection without Stampidine (p<0.05) but silenced in the presence of Stampidine (p>0.2) (Figure 3A and 3B). Three (*MAP4*, *JAK1*, *TNPO3*) of these 18 genes were also identified as HDF genes in both the Konig et al. siRNA screen of 293T cells [17] and Zhou et al. siRNA screen of T cells [18] while the remaining 15 were identified as HDF genes in at least one of these independent HIV siRNAi screens. Further, 2 of 18 genes that were

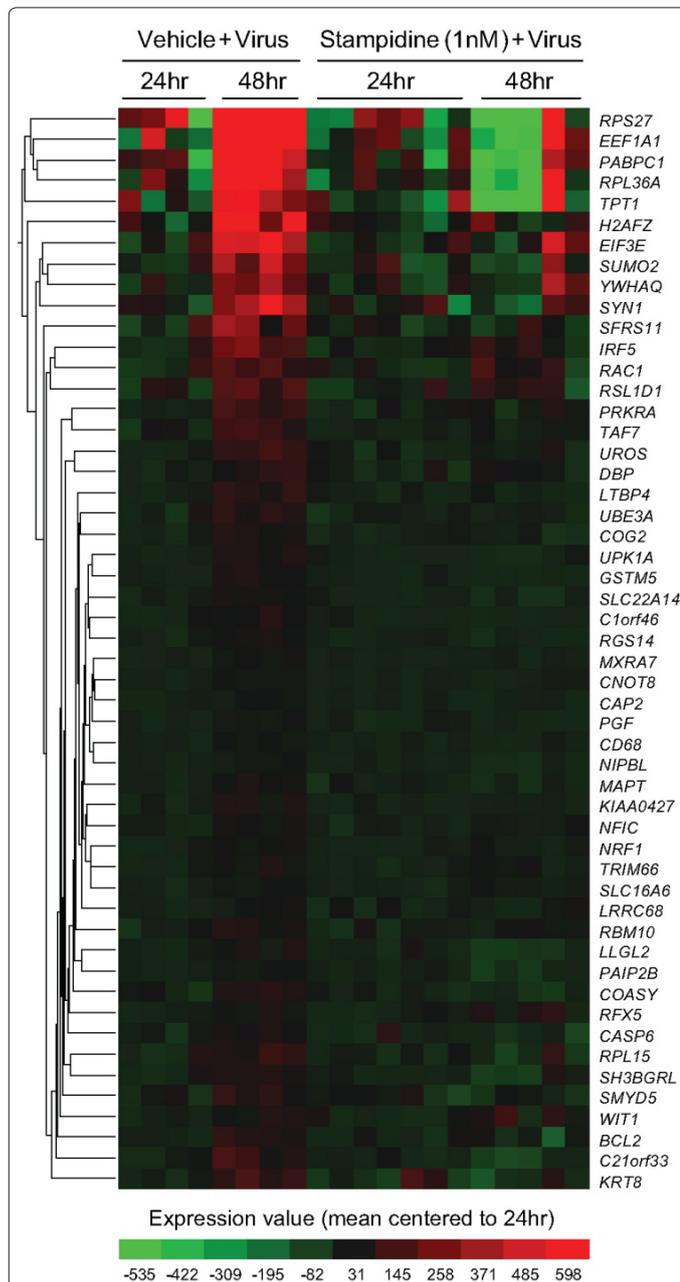
cross-referenced with the Brass siRNA screen [14] were down regulated by Stampidine alone within 48 hours (*RSL1D1* at 48 hours and *COG2* at 24 hours) (Figure 3B). These genes affected by Stampidine alone treatment appeared in two sub-clusters of co-regulated expression sets highlighting 4 genes with the greatest signal difference values (*JAK1*, *TNPO3*, *CLDND1*, *RSL1D1*) and 3 genes (*COG2*, *MAP4*, *NMT1*) with HIV-induced increases that were only slightly reduced by Stampidine (Figure 3). Hence, in HIV infected cells, Stampidine silences HDF genes previously shown to facilitate multiple steps of HIV infection, including viral entry, viral integration, and viral transcription. Stampidine is the first pharmaceutical ARV drug candidate capable of down-regulating critical genes required for the organization of transcriptional networks enabling HIV replication.

We next sought to determine the effects of Stampidine on the temporal order of dynamic changes in gene expression in MT-2 cells between 24 and 48 hours after exposure to one of 3 different HIV strains. We compared HIV-induced changes relative to control levels of expression at 24 and 48 hours and filtered genes that exhibited increases in HIV-induced expression levels at 48 hours relative to 24 hours. By comparing 48 hour with 24 hour time points and ordering the genes according to the T-test P-value for mean change in signal difference metric, we identified 83 HIV-responsive genes that were significantly upregulated after HIV infection. Notably, Stampidine completely abrogated the HIV-induced increased expression of these 83 genes (Table S3). For 52 genes selected for cluster analysis, the signal difference values were greater than 20 with P-values <0.01. *UROS*, *UPK1A*, *EIF3E*, *PAIP2B*, *MXRA7* and *YWHAQ* were the most significantly



**Figure 3: HIV-induced increases at 48 h abrogated by Stampidine showed cross hits with Host Dependent Factors identified in RNAi study.** Cross-referencing of HIV-induced increases in expression between 24 and 48 hours that were abrogated by Stampidine with HDF genes validated in the Brass siRNA screen [1] resulted in identification of 18 genes depicted in the heat map. Gene expression values (signal difference metric) were mean centered to HIV-infected 24 hour time point in the presence and absence of Stampidine to assess genes dynamically induced in expression by virus at 48 hours. One-way agglomerative hierarchical clustering technique was used to organize gene expression patterns across treatments to identify genes with similar profiles. Dendrograms depicted in A illustrated similar expression from joining pairs of expression profiles using the average distance metric, whereby profiles joined by short branch lengths showed most similarity. The heat map depicts genes up regulated (red) and down regulated (green) at 48 hours relative to 24 hours. The table in B shows the mean centered increases in gene expression relative to the 24 hour time point in the presence of the virus (N=5) and the abrogation in the presence of virus and Stampidine (N=7). The fold change for the effect of Stampidine alone at 24hr (N=6) and 48 hr (N=6).

affected 6 genes ( $P < 0.01$ ). Cluster analysis (Figure 4) exhibited coregulation of 5 genes (*RPS27*, *EEF1A1*, *PABPC1*, *RPL36A* and *TPT1*) with expression changes greater than 500 using the signal difference metric following infection with HIV and abrogation of the HIV-



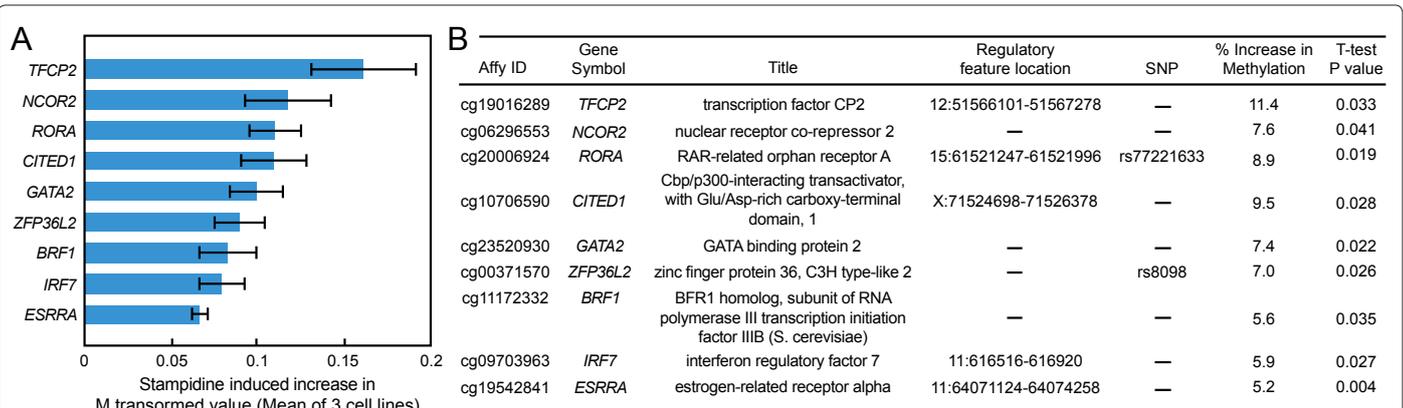
**Figure 4: Effects of Stampidine on HIV-induced Dynamic Gene Expression Changes.** Gene expression values (signal difference metric) were mean centered to 24 h time point post-HIV exposure in the presence and absence of Stampidine to assess genes dynamically induced in expression by the virus at 48 h. One-way agglomerative hierarchical clustering technique was used to organize gene expression patterns across treatments to identify genes with similar profiles. The heat map depicts genes upregulated (red) vs. downregulated (green) at 48 h relative to 24 h. Fifty-two genes exhibited significant upregulation after HIV infection in the absence of Stampidine and signal difference values greater than 20. Comparison of control versus Stampidine treatment showed that all of these HIV-induced genes were downregulated with Stampidine treatment within 48 hours (24 hours or 48 hours or combined 24 and 48 hour time points).

induced increased expression by Stampidine. A subcluster of 4 genes represented transcripts that were significantly upregulated with HIV infection ( $P$ -value  $< 0.01$ ) but showed no HIV-induced upregulation in the presence of Stampidine (*EIF3E*, *SUMO2*, *YWHAQ*, and *SYN1*).

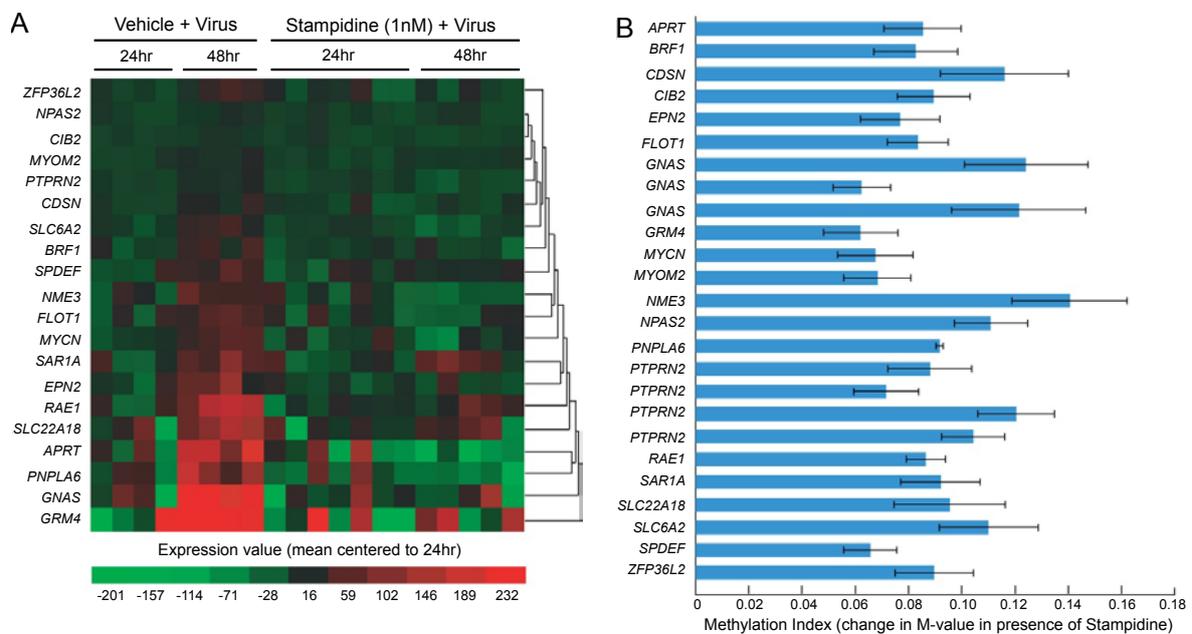
### Stampidine causes methylation of a subset of HIV-responsive regulatory genes in T-cells

We utilized the Infinium HumanMethylation450 BeadChip (Illumina) to interrogate in 3 different human T-cell lines methylation levels for more than 485,000 methylations sites at the single nucleotide levels covering 99% of RefSeq genes with an average of 17 CpG sites per gene region distributed across the promoter, 5'UTR, first exon, gene body, and 3'UTR, and 96% of CpG islands, with additional coverage in island shores and the regions flanking them. 821 probes revealed increased methylation of 568 genes in the presence of Stampidine (Table S4). *TFAP4*, *LEFTY1*, and *NCOR2*, 3 HDF genes identified by Brass siRNA screen as being critical for HIV replication, were methylated by Stampidine. By cross-referencing these genes with the list of HIV-responsive transcription factor genes from gene expression profiling experiments and correlation cluster analysis performed using MT-2 cells, we identified 9 genes (*TFCP2*, *NCOR2*, *RORA*, *CITED1*, *GATA2*, *ZFP36L2*, *BRF1*, *IRF7*, *ESRRA*) methylated by Stampidine that were components of the network of correlated transcription factors formed in the presence of HIV and silenced in the presence Stampidine (Figure 5). These 9 genes formed 73 significant correlations with a total of 66 genes (9 methylated genes plus 57 other genes) in the network in which positive correlation coefficient values were reduced in the presence of Stampidine (Table S5). *ZFP36L2*, *NCOR2* and *RORA* were identified as potential hubs as they formed 16, 14 and 14 connections respectively. Regulatory regions of 5 genes (*TFCP2*, *RORA*, *CITED1*, *IRF7*, *ESRRA*) showed an increase in methylation, 2 genes exhibited methylation close to SNP sites (*RORA* (rs77221633) and *ZFP36L2* (rs8098)). Of the 66 genes, 17 (2 methylated genes plus 15 other highly correlated genes) showed dynamic changes in their expression level between 24-h and 48-h after HIV infection. These dynamic changes were abrogated by Stampidine (Table S6). Notably, a total of 20 genes showing dynamic increases in expression after HIV infection were methylated by Stampidine (Figure 6). Stampidine prevented HIV-induced upregulation of 19 of these 20 genes (Figure 6A, Table S7A). Of these 19 genes, 15 showed methylation of their regulatory regions and/or CPG islands. Regulatory regions of 4 genes (*CIB2*, *EPN2*, *FLOT1*, *PNPLA6*) showed an increase in methylation and 9 genes exhibited methylation close to SNP sites (Table S7B: *APRT* (rs35132384), *CDSN* (rs2302399), *FLOT1* (rs74956736), *GNAS* (rs77997868), *NME3* (rs11543214), *PTPRN2* (rs75100955), *SLC22A18* (rs451443), *SPDEF* (rs56262607), *ZFP36L2* (rs8098)) (Figure 6B, Table S7B). Five genes exhibited an increase in expression in the presence of virus between 24 and 48 hours with p-value less than 0.01 (*PTPRN2*, *MYCN*, *RAE1*, *APRT* and *GNAS*), and two of these genes showed an increase in methylation at multiple CPG motifs (*GNAS* (3 sites) and *PTPRN2* (4 sites)) (TableS7 A and B).

Taken together, these studies establish our lead ARV drug candidate Stampidine as a novel epigenetic modulator of HIV infection-associated gene expression. We propose that Stampidine disrupts critical steps in HIV replication via downregulation of host genes that configure gene networks to enable viral replication. Stampidine is the first dual-function ARV drug candidate with both NRTI activity and epigenetic anti-host transcriptome activity for PrEP as well as for therapy against HIV/AIDS. Unlike available ARV agents that disrupt a specific step in the life-cycle of HIV, Stampidine has the potential to abrogate all steps in the life cycle of HIV.



**Figure 5: Stampidine methylates transcription factors in the HIV network.** The effects of Stampidine on gene methylation levels was determined in 3 T-cell lines (LOUCY, CEM, MOLT3) by comparing the methylation levels in samples treated with Stampidine to those in vehicle-treated control samples. Nine genes exhibited significant increases (Mean and SEM for M-values shown) in methylation that formed interconnections in the HIV network for transcription factors.



**Figure 6: Effects of Stampidine on Methylation Levels of HIV-Responsive Transcription Factor Genes. HIV-Induced Dynamic Increases in Expression of Methylated Transcription Factor Genes.** A total of 20 genes showing dynamic increases in expression after HIV infection were methylated by Stampidine. Gene expression values (signal difference metric) were mean centered to 24 h time point post-HIV exposure in the presence and absence of Stampidine to assess genes dynamically induced in expression by the virus at 48 h. One-way agglomerative hierarchical clustering technique was used to organize gene expression patterns across treatments to identify genes with similar profiles. The heat map depicts genes upregulated (red) vs. downregulated (green) at 48 h relative to 24 h (A). Gene expression values for GNAS were scales by a factor of 0.2 to a comparable dynamic range observed with other gene transcripts. Stampidine prevented HIV-induced upregulation of 19 of these 20 genes. The effects of Stampidine on gene methylation levels was determined in 3 T-cell lines (LOUCY, CEM, MOLT3) by comparing the methylation levels in samples treated with Stampidine to those in vehicle-treated control samples. Methylation profile depicts the increase in methylation in the presence of Stampidine for methylation probes corresponding to these 20 genes (B). A total of 20 genes showing dynamic increases in expression after HIV infection were methylated by Stampidine. See Table S7 for details.

Stampidine is a promising ARV drug candidate against HIV because of (a) its remarkable subnanomolar to low nanomolar *in vitro* anti-retroviral potency against genotypically and phenotypically NRTI-resistant primary clinical HIV isolates, non-nucleoside reverse transcriptase inhibitor (NNRTI)-resistant HIV-1 isolates, as well as clinical non-B subtype HIV-1 isolates (subtypes A, C, F, and G) originating from South America, Asia, and sub-Saharan Africa with resistance to stavudine, adefovir and tenofovir [28], (b) favorable pharmacokinetics profile in rodents, dogs, and cats with 25 mg/kg or 50

mg/kg tolerable dose levels yielding micromolar plasma concentrations of Stampidine in rodents, cats, and dogs, which are 1,000-fold higher than its *in vitro* IC<sub>50</sub> value against HIV [29,30], (c) favorable, safety profile in rodents, dogs, and cats [29,30], and (d) *in vivo* anti-retroviral activity in Hu-PBL-SCID mice as well as FIV-infected domestic cats [31]. In a recently completed placebo-controlled Phase I study involving 30 therapy-naïve adult HIV-infected adult patients, formulated cGMP-grade Stampidine [32] administered as a single oral bolus dose did not cause dose-limiting toxicity at single dose levels ranging from 5 to 25

mg/kg, but it resulted in modest decreases in HIV load in 9 patients [33]. The favorable preclinical and early clinical safety/activity profile of Stampidine warrants its further development as a new anti-HIV drug candidate against ARV drug-resistant HIV strains.

The epigenetic effects of Stampidine are very different from those of azacytidine (Aza-C), another nucleoside analog with marked epigenetic modulatory effects on gene methylation and expression [34]. In contrast to Aza-C, which acts as a demethylating agent that induces gene expression, Stampidine is a methylating agent that represses gene expression. Furthermore, its epigenetic effects are not only in the opposite direction (i.e., methylation by stampidine vs. demethylation by AZA-C), but also much more selective: Stampidine methylates a very small proportion of CpG sites in the genome (821 sites (0.2%)) interrogated on the Infinium HumanMethylation450 BeadChip. By comparison, in a recently published genome-wide analysis examining the methylation levels of more than 27,000 CG dinucleotides representing more than 14,000 genes, Aza-C resulted in strong demethylation ( $DB \leq -0.2$ ) of 16% (1,839 of 11,406) of the CGs methylated in control cells [34]. Cross referencing of the transcription factor genes that were methylated by Stampidine with the transcription factor binding sites that were enriched in 121 genes that became demethylated after Aza-C treatment revealed that only 3 transcription factor binding sites (GATA2, EN1 and HNF1B) were shared. Aza-C would demethylate and thereby increase gene expression of the genes harboring binding sites for the 3 transcription factors (very much like HIV exposure), whereas Stampidine would methylate the genes for these transcription factors that would result in the down regulation of genes activated by these transcription factors. Our results show that Stampidine selectively methylates a few transcription factors found at hubs of correlated gene expression changes to cause widespread disruption of HIV induced gene networks.

## References

1. Palella FJ, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, et al. (1998) Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *N Engl J Med* 338: 853-860.
2. Richman DD, Margolis DM, Delaney M, Greene WC, Hazuda D, et al. (2009) The challenge of finding a cure for HIV infection. *Science* 323: 1304-1307.
3. Uckun FM (2006) Unmet challenges in HIV therapy and potential of stampidine. *Arzneimittelforschung* 56: 117-120.
4. Johnson VA, Brun-Vezinet F, Clotet B, Gunthard HF, Kuritzkes DR, et al. (2010) Update of the Drug Resistance Mutations in HIV-1: December 2010. *Top HIV Med* 18: 156-163.
5. Kantor R, Katzenstein DA, Efron B, Carvalho AP, Wynhoven B, et al. (2005) Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: results of a global collaboration. *PLoS Med* 2: e112.
6. Rhee SY, Fessel WJ, Zolopa AR, Hurley L, Liu T, et al. (2005) HIV protease and reverse transcriptase mutations: correlations with antiretroviral therapy in subtype B isolates and implications for drug resistance surveillance. *J Infect Dis* 192: 456-465.
7. Shafer RW, Kantor R, Gonzales MJ (2000) The genetic basis of HIV-1 resistance to reverse transcriptase and protease inhibitors. *AIDS Rev* 2: 211-228.
8. Kelesidis T, Landovitz RJ (2011) Preexposure prophylaxis for HIV prevention. *Curr HIV/AIDS Rep* 8: 94-103.
9. Lampe MA, Smith DK, Anderson GJ, Edwards AE, Nesheim SR (2011) Achieving safe conception in HIV-discordant couples: the potential role of oral preexposure prophylaxis (PrEP) in the United States. *Am J Obstet Gynecol* 204: 488.e1-8.
10. Mayer KH, Venkatesh KK (2010) Chemoprophylaxis for HIV prevention: new opportunities and new questions. *J Acquir Immune Defic Syndr* 2: S122-127.
11. Das K, Bandwar RP, White KL, Feng JY, Sarafianos SG, et al. (2009) Structural basis for the role of the K65R mutation in HIV-1 reverse transcriptase polymerization, excision antagonism, and tenofovir resistance. *J Biol Chem* 284: 35092-35100.
12. Hurt CB, Eron JJ Jr, Cohen MS (2011) Pre-Exposure Prophylaxis and Antiretroviral Resistance: HIV Prevention at a Cost? *Clin Infect Dis* 53: 1265-1270.
13. White KL, Chen JM, Margot NA, Wrin T, Petropoulos CJ, et al. (2004) Molecular mechanisms of tenofovir resistance conferred by human immunodeficiency virus type 1 reverse transcriptase containing a diserine insertion after residue 69 and multiple thymidine analog-associated mutations. *Antimicrob Agents Chemother* 48: 992-1003.
14. Brass AL, Dykxhoorn DM, Benita Y, Yan N, Engelman A, et al. (2008) Identification of host proteins required for HIV infection through a functional genomic screen. *Science* 319: 921-926.
15. Bushman FD, Malani N, Fernandes J, D'Orso I, Cagney G, et al. (2009) Host cell factors in HIV replication: meta-analysis of genome-wide studies. *PLoS Pathog* 5: e1000437.
16. König R, Stertz S, Zhou Y, Inoue A, Hoffmann HH, et al. (2010) Human host factors required for influenza virus replication. *Nature* 463: 813-817.
17. König R, Zhou Y, Elleder D, Diamond TL, Bonamy GM, et al. (2008) Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* 135: 49-60.
18. Zhou H, Xu M, Huang Q, Gates AT, Zhang XD, et al. (2008) Genome-scale RNAi screen for host factors required for HIV replication. *Cell Host Microbe* 4: 495-504.
19. Uckun FM (2006) Aryl phosphate derivatives of d4T having activity against resistant HIV strains. US Patent No. 7144874.
20. Johnston E, Dupnik KM, Gonzales MJ, Winters MA, Rhee SY, et al. (2005) Panel of prototypical infectious molecular HIV-1 clones containing multiple nucleoside reverse transcriptase inhibitor resistance mutations. *AIDS* 19: 731-733.
21. Qazi S, Soiseth NL, Uckun FM (2010) Early detection of subtype-specific HIV-1 infection using transcriptomes from host cell responses. *EJBS* 2: 65-79.
22. Qazi S, Uckun FM (2010) Gene expression profiles of infant acute lymphoblastic leukaemia and its prognostically distinct subsets. *Br J Haematol* 149: 865-873.
23. Uckun FM, Goodman P, Ma H, Dibirdik I, Qazi S (2010) CD22 Exon 12 Deletion as a Novel Pathogenic Mechanism of Human B-Precursor Leukemia. *Proc Natl Acad Sci U S A* 107: 16852-16857.
24. Uckun FM, Qazi S, Ma H, Tuel-Ahlgren L, Ozer Z (2010) STAT3 is a substrate of SYK tyrosine kinase in B-lineage leukemia/lymphoma cells exposed to oxidative stress. *Proc Natl Acad Sci U S A* 107: 2902-2907.
25. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, et al. (2010) Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res* 20: 440-446.
26. Weisenberger DJ, Den Berg DV, Pan F, Berman BP, Laird PW (2008) Comprehensive DNA methylation analysis on the Illumina Infinium assay platform. Technical report. Illumina, Inc; San Diego.
27. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, et al. (2010) Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 11: 587.
28. Uckun FM, Pendergrass S, Venkatachalam TK, Qazi S, Richman D (2002) Stampidine is a potent inhibitor of zidovudine and nucleoside analog reverse transcriptase inhibitor-resistant primary clinical HIV-1 isolates with thymidine analog mutations. *Antimicrob Agents Chemother* 46: 3613-3616.
29. Uckun FM, Qazi S, Pendergrass S, Lisowski E, Waurzyniak B, et al. (2002) *In vivo* toxicity, pharmacokinetics, and anti-HIV activity of stavudine-5'-(p-bromo phenyl) methoxyalaninyl phosphate) (stampidine) in mice. *Antimicrob Agents Chemother* 46: 3428-3436.
30. Uckun FM, Waurzyniak B, Tibbles H, Venkatachalam TK, Erbeck D (2006) *In vivo* pharmacokinetics and toxicity profile of the anti-HIV agent stampidine in dogs and feline immunodeficiency virus-infected cats. *Arzneimittelforschung* 56: 176-192.

31. Uckun FM, Chen CL, Samuel P, Pendergrass S, Venkatachalam TK, et al. (2003) *In vivo* antiretroviral activity of stampidine in chronically feline immunodeficiency virus-infected cats. *Antimicrob Agents Chemother* 47: 1233-1240.
32. DuMez D, Venkatachalam TK, Uckun FM (2006) Large-scale synthesis and formulation of GMP-grade stampidine, a new anti-HIV agent. *Arzneimittelforschung* 56: 136-151.
33. Cahn P, Rolon MJ, Gun AM, Ferrari I, Dibirdik I, et al. (2012) Preclinical and First-In-Human Phase I Clinical Evaluation of Stampidine, A Potent Anti-HIV Pharmaceutical Drug Candidate. *J AIDS Clinic Res* 3: 1-8.
34. Hagemann S, Heil O, Lyko F, Brueckner B (2011) Azacytidine and Decitabine Induce Gene-Specific and Non-Random DNA Demethylation in Human Cancer Cell Lines. *PLoS One* 6: e17388.