

## Supporting Information

### Detailed Methods

**SI Figure 1.** Workflow and Bayesian statistics employed in virtual screening.

**SI Figure 2.** Anti-HIV-1 activity of compound **3** in MAGI and PBMC assays.

**SI Figure 3.** Structures of all active compounds from virtual screen.

**SI Figure 4.** Model of compound **3** binding to CXCR4 derived from the IT1t crystal structure.

**SI Figure 5.** Comparison of compound **3** pose derived from CXCR4:CVX15 structure with residues critical for SDF-1 binding and HIV-1 infection.

**SI Figure 6.** Anti-HIV-2 activity of compound **3** in PBM cells.

**SI Table 1.** Table 1 from text including control compounds and activities.

**SI Table 2.** MAGI assay activity of compounds from screen.

**SI Table 3.** Radio-labeled chemokine ligand displacement results.

### SI REFERENCES

## Detailed Methods

### *Virtual Screen Methods*

Bayesian models were constructed using methods as previously outlined by several other groups using SciTegic Pipeline Pilot (v. 7.0.1.100) (Accelrys Software Inc. 2008).<sup>1,2</sup> The GPCR-SARfari database (Version 3.0) was used for this study containing 147,292 compounds with over 1 million bioactivity datapoints against 912 GPCR targets (accessed September, 2012. <https://www.ebi.ac.uk/chembl/sarfari/gpcrsarfari>).<sup>3</sup> This set was first filtered for compounds with molecular weight of less than 650 g/mol and reported activity in “nM” units leaving approximately 90,000 compounds. Active compounds against CCR5 were identified if the reported IC<sub>50</sub>/K<sub>i</sub> value was less than 10,000 nM. This resulted in 1502 actives for CCR5. The number of unique CXCR4 active compounds in this data set was low, and additional CXCR4 active ligands (< 10μM potency) were obtained from the BindingDB<sup>4</sup> and literature review<sup>5</sup> bringing the final set of CXCR4 active compounds to 221. The remaining compounds from the GPCR-SARfari were presumed to be inactive. Training and test sets were constructed for each GPCR by randomly dividing the actives and inactives into two groups. Laplacian-modified Bayesian statistical models were constructed for either CXCR4 or CCR5 training sets using the functional chemical fingerprints to the 6<sup>th</sup> neighbor (FCFP\_6) as a descriptor set, and these models were then scored the test set. Bayesian models were also constructed using the total set of active and inactive compounds and scored using the same total set. Both exercises yielded similar statistics; we used and report the total data set statistics to maximize chemical information. The ROC curves, leave-one-out statistics and virtual screen workflow are provided (**Supplementary Fig. 2**). Prior to scoring and ranking, the database was filtered for good high-throughput compound filters, Lipinski filters and organic filters available in the program.

### *Molecular Modeling Methods*

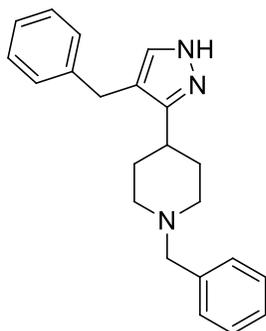
Schrodinger's Maestro Suite (New York, NY, 2012) was employed to generate models of compound **3** interacting with the CXCR4, CCR5, and HIV-RT proteins. Default values were used within the programs unless otherwise stated. The crystal structures used in modeling were the CCR5:Maraviroc crystal structure chain A (pdbid 4MBS),<sup>6</sup> the CXCR4:IT1t chain A (pdbid 3ODU), the CXCR4:CVX15 crystal structure (pdbid 3OE0)<sup>7</sup> and the HIV-RT:5h crystal structure (pdbid 3M8Q).<sup>8</sup> Each structure was subjected to a similar preparation. All waters were removed from the GPCR crystal structures with the exception of the 567 ordered water molecule in the HIV-RT crystal structure. The crystal structures were prepared using the protein preparation wizard workflow, and a Glide Grid was generated centered on the co-crystallized ligand (the Arg2 – Nal3 residues in the case of the CXCR4:CVX15 structure) with a 10x10x10Å boundary box that spanned the entire binding pocket. 410 low-energy conformers and tautomers of compound **3** with the piperidine nitrogen protonated were generated using MacroModel with the OPLS-2005 force field as ligand entries for docking. The low-energy conformers of cationic **3** were flexibly docked into the grids using Glide at standard precision, and the pose with the most favorable glide score for each protein was reported. The HIV-RT:Efavirenz (HIVRT:EFV) and HIV:Etravirine crystal structures (pdbid 1FK9 and 3M8P)<sup>8,9</sup> were also examined in addition to the HIV-RT:5h structure as models for the NNRTI activity of **3**. Alignment of the HIV-RT structures reveals that the NNRTI binding pocket site may be flexible. Namely, the HIVRT:EFV site is more compact, and the RT-5h structure is larger due to different ligand sizes. Docking 410 different conformers/tautomers of **3** into the HIV-RT:EFV site results in poses that are highly strained (e.g. axial orientations in the piperidine ring and/or distortion of the ligand) in

order to fit into the more compact site. This suggests that the HIVRT:5h structure may be a more accurate representation of binding because the 5h ligand contains 4 ring structures similar to our compound.

### Re-Synthesis and Characterization of Hit Compounds

The purity of all purchased compounds were >95% as determined by HPLC.

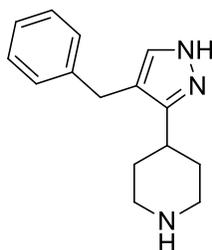
**Precursor 1.** Prepared as described previously.<sup>10</sup> <sup>1</sup>H NMR (600 MHz, Chloroform-d)  $\delta$  7.32 –



7.21 (m, 8H), 7.18 – 7.13 (m, 2H), 3.79 (s, 2H), 3.67 (s, 1H), 3.50 (s, 2H), 2.94 (dt, J = 11.3, 2.9 Hz, 2H), 2.66 – 2.56 (m, 1H), 1.99 (tt, J = 11.7, 1.5 Hz, 2H), 1.89 – 1.80 (m, 2H), 1.75 – 1.67 (m, 2H). <sup>13</sup>C NMR (151 MHz, Chloroform-d)  $\delta$  161.70, 141.37, 138.29, 129.49, 128.58, 128.41, 127.27, 126.18, 116.00, 70.74, 63.60, 54.14, 31.84, 30.03. LCMS

75-95% 3 minutes MeOH:H<sub>2</sub>O gradient >95% pure rt= .620 HRMS calc'd for C<sub>22</sub>H<sub>26</sub>N<sub>3</sub> 332.21212; found [M+H] 332.21145.

**Precursor 2.** To a solution of **precursor 1** (.30 g, .9 mmol) in *t*-BuOH (9 mL, .1M) and AcOH



(.9 mL, .01 M) was added wet Pd/C (.06 g, 10% Pd by weight). The reaction was then hydrogenated under a 50 psi atmosphere of hydrogen gas (Parr-Shaker overnight). Upon completion the hydrogen was removed *in vacuo* and then flushed with argon. The crude reaction mixture was then filtered through

two fluted pieces of filter paper and concentrated *in vacuo*. The oily residue was then diluted with brine and DCM followed by basification with 10% NaOH. The layers were separated and

the aqueous layer extracted with DCM (3 times). The organic layers were combined, dried over anhydrous sodium sulfate, filtered and concentrated to afford **Precursor 2**, which was used without further purification. <sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 7.30 (s, 1H), 7.27 – 7.20 (m, 2H), 7.18 – 7.08 (m, 3H), 3.77 (s, 2H), 3.41 (dt, *J* = 12.8, 3.0 Hz, 2H), 2.91 (td, *J* = 13.1, 3.2 Hz, 2H), 2.81 (tt, *J* = 11.7, 3.8 Hz, 1H), 2.02 – 1.91 (m, 2H), 1.85 – 1.74 (m, 2H). <sup>13</sup>C NMR (101 MHz, Chloroform-*d*) δ 177.14, 149.03, 140.68, 131.58, 128.74, 128.52, 126.51, 116.92, 43.99, 31.80, 31.23, 29.91, 28.64. LCMS 75-95% 3 minutes MeOH:H<sub>2</sub>O gradient >95% pure *rt* = .581 HRMS calc'd for C<sub>15</sub>H<sub>20</sub>N<sub>3</sub> 242.16517; found [M+H] 242.16551

**General Reductive Amination:** To a solution of the amine in DCM (.1M) is added the aldehyde (1.1 eq) and stirred at room temperature for 30 minutes. Then sodium triacetoxyborohydride (1.5 eq) is added as one portion and the reaction is tracked by LCMS. The reaction is usually complete within 5 hours. Upon completion the mixture is diluted with brine and basified with 10% NaOH. The layers are separated and the aqueous layer extracted with DCM (3 times). The organic layers were combined, dried over anhydrous sodium sulfate, filtered and concentrated to afford the crude product which is purified by column chromatography.

**Compound 1.** Prepared by the general reductive amination procedure. Purified on a 12 gram combiflash column with a gradient from 0-70% DCM:MeOH:NH<sub>4</sub>OH (9:1:5) in DCM to afford 2-((4-(4-benzyl-1H-pyrazol-3-yl)piperidin-1-yl)methyl)pyridine (275 mg, 62% yield over two steps). <sup>1</sup>H NMR (400 MHz, Chloroform-*d*) δ 8.53 (ddd, *J* = 5.0, 1.9, 0.9 Hz, 1H), 7.63 (td, *J* = 7.6, 1.8 Hz, 1H), 7.40 (d, *J* = 7.8 Hz, 1H), 7.30 – 7.21 (m, 3H), 7.19 – 7.11 (m, 3H), 3.79 (s, 2H), 3.65 (s, 2H), 2.95 (d, *J* = 11.4 Hz, 2H), 2.62 (tt, *J* = 12.1, 4.0 Hz, 1H), 2.10 (td, *J* = 11.8, 2.7 Hz,

2H), 1.84 (qd,  $J = 12.3, 3.8$  Hz, 2H), 1.72 (d,  $J = 12.3$  Hz, 2H).  $^{13}\text{C}$  NMR (101 MHz, Chloroform-d)  $\delta$  158.23, 149.12, 141.20, 136.54, 128.38, 125.95, 123.53, 122.33, 115.71, 64.79, 54.19, 33.18, 31.51, 29.80. LCMS 75-95% 3 minutes MeOH:H<sub>2</sub>O gradient >95% pure  $rt = .555$ . HRMS calculated for C<sub>21</sub>H<sub>25</sub>N<sub>4</sub> 333.20737; found [M+H] 333.20695.

**Compound 3.** Prepared by a general reductive amination procedure. Purified on a 12 gram combiflash column with a gradient of 0-70% DCM:MeOH:NH<sub>4</sub>OH 9:1:5 in DCM to afford 4-((4-(4-benzyl-1H-pyrazol-3-yl)piperidin-1-yl)methyl)pyridine (225 mg, 51% yield over two steps).  $^1\text{H}$  NMR (400 MHz, Chloroform-d)  $\delta$  8.71 – 8.36 (m, 2H), 7.35 (s, 1H), 7.27 – 7.21 (m, 4H), 7.19 – 7.12 (m, 3H), 3.82 (s, 2H), 3.48 (s, 2H), 2.88 (dt,  $J = 11.5, 3.1$  Hz, 2H), 2.62 (tt,  $J = 12.0, 3.9$  Hz, 1H), 2.03 (td,  $J = 11.5, 2.3$  Hz, 2H), 1.89 (qd,  $J = 12.4, 3.6$  Hz, 2H), 1.79 – 1.65 (m, 2H).  $^{13}\text{C}$  NMR (101 MHz, Chloroform-d)  $\delta$  149.62, 148.08, 141.13, 128.34, 125.98, 123.80, 115.74, 61.98, 54.13, 33.32, 31.69, 29.82. LCMS 25-95% 8 minutes MeOH:H<sub>2</sub>O gradient >95% pure  $rt = 3.410$ . LCMS 75-95% 3 minutes MeOH:H<sub>2</sub>O gradient >95% pure  $rt = .589$ . HRMS calc'd for C<sub>21</sub>H<sub>25</sub>N<sub>4</sub> 333.20737; found [M+H] 333.20729.

*Viruses and Cells.* The K103N/Y181C NL4-3-based replication competent virus was created through site-directed mutagenesis of a proviral clone and contains a Renilla luciferase gene in place of the Nef gene.<sup>11</sup> A pseudotype virus HIV-1 LAI  $\Delta env-luc$  (LAI-Luc)<sup>12</sup> was prepared by co-transfecting HEK 293T cells with both a plasmid containing proviral DNA of HIV-1 LAI  $\Delta env-luc$  (Firefly luciferase) and pVSVenv expressing the Vesicular Stomatitis Virus G envelope protein. The CXCR4-independent HIV-8X envelope clone was a gift of Dr. Robert Doms (University of Pennsylvania). HeLa-C14 cells were constructed and express CD4 and both HIV-

1 co-receptors, CCR5 and CXCR4.<sup>13</sup> HeLa-C14 cells are grown and maintained in DMEM + 10% FCS, PenStrep, 400 µg/ml zeocin, 100 µg/ml hygromycin, 200 µg/ml G418. HeLa-CD4/CCR5/CXCR4 cells express the three proteins plus it contains an integrated copy of the pTRE2Hyg-Luc plasmid. HeLa-CD4/CCR5/CXCR4 are maintained in 10% FBS-DMEM + 200 µg/ml G418, 200 hygromycin, 1.5 µg/ml puromycin, 400 µg/ml zeocin). ACTOne-CXCR4 are modified HEK293 cells that express CXCR4, but not CD4 or CCR5 (purchased from Codex BioSolutions, Inc). ACTOne cells are maintained in DMEM + 10% FBS selection, 200 µg/ml mg/ml G418, 2 µg/ml Puromycin.

*HIV-Nef-M1 Induced Mitochondrial Membrane Depolarization Assay.* 11-mer peptides Nef Motif-1 (Nef M1, TNAACAWLEAQ) and Nef sMotif-1 (Nef sM1, ALAETCQNAWA) were obtained from Sigma Genosys (Houston, TX). For the depolarization assay, 5,5',6,6'-tetrachloro-1,1',3,3'-tetraethyl-benzimidazolylcarbocyanine iodide (JC-1; CBIC) was obtained from Molecular Probes Inc. (Eugene, Oregon). Jurkat cells were treated with 10 ng/ml of NefM1 or sM1 and different doses of AMD3100 (50, 200, 400, 600, 800, and 1000 nM), maraviroc (0.1, 1, 10, 100, 1000, and 10,000) or AZT (0.1, 1, 10, 100, and 1000 nM) for 24 hours. The cultures were then washed in 1 x PBS and 100 µl of fresh JC-1 stain solution (a stock solution of JC-1 at a concentration of 2.5 mg/ml in DMSO was diluted into culture media to a final concentration of 10 µg/ml) was added to the cell pellet and incubated at 37 °C for 10 min. The cells were washed in 1 x PBS, and rinsed in imaging buffer (High K<sup>+</sup> buffer, 137 mM KCl, 3.6 mM NaCl, 0.5 mM MgCl<sub>2</sub>, 1.8 mM CaCl<sub>2</sub>, 1.6 mM NaH<sub>2</sub>PO<sub>4</sub> and 4.3 mM NaHCO<sub>3</sub>, pH 7.4). Stained cultures were viewed immediately by epifluorescence on a computer-controlled microscope system (Carl Zeiss, Thornwood, NY) and images were captured via a charge-coupled device camera (MC 100

SPOT 60910; Photonic Science, East Sussex, United Kingdom). Further image processing was conducted with Image-Pro 2.0 software (Media Cybernetics, Silver Spring, MD). SigmaPlot 10 was used for numerical and graphical analyses of all data obtained.

*MAGI Assay.* MAGI assays were conducted by Southern Research Institute. MAGI-CCR5 cells (obtained from the NIH AIDS Research and Reference reagent Program) were passaged in T-75 flasks prior to use in the antiviral assay. MAGI-CCR5 cells were derived from HeLa-CD4-LTR- $\beta$ -gal cells, and were used for either CXCR4- or CCR5-tropic assays depending on the specific virus employed in the study (see below). The cells naturally express CXCR4 and have been engineered to express high levels of CD4 and CCR5 and contain one copy of the HIV-1 LTR promoter, which drives expression of the  $\beta$ -galactosidase gene upon HIV-1 Tat transactivation. On the day preceding the assay, the cells were plated at  $1.0 \times 10^4$  per well and incubated at 37°C overnight. Total cell and viability quantification was performed using a hemacytometer and trypan blue exclusion. Cell viability was greater than 95% for the cells utilized in the assay. The viruses used for these tests were the CXCR4-tropic virus strain HIV-1<sub>IIIIB</sub> or the CCR5-tropic virus strain HIV-1<sub>Ba-L</sub>. Both viruses were obtained from the NIH AIDS Research and Reference Reagent Program. HIV-1<sub>IIIIB</sub> was grown in CEM-SS cells, and HIV-1<sub>Ba-L</sub> was grown in Ghost Hi5/MAGI-CCR5 co-cultures for the production of stock virus pools. For each assay, a pre-titered aliquot of virus was removed from the freezer (-80°C) and allowed to thaw slowly to room temperature in a biological safety cabinet. The virus was resuspended and diluted into tissue culture medium such that the amount of virus added to each well in a volume of 50  $\mu$ L was approximately ten TCID<sub>50</sub>/well (~0.001 TCID<sub>50</sub>/cell). Compounds were evaluated at six concentrations (triplicate wells/concentration). On the day of assay setup, compound dilutions

were prepared at two-times (2X) the final required concentrations. Media used for plating the cells the day before assay setup was aspirated from the plates and replaced with 50  $\mu$ L of the 2X compounds, followed by the addition of 50  $\mu$ L of virus, which diluted the compounds to the final 1X concentrations. Cell control wells (cells only) and virus control wells (cells plus virus) were included on each assay plate. Identical uninfected assay plates (virus replaced with media) were prepared for parallel cytotoxicity testing. The cultures were incubated for 48 hours after which antiviral efficacy was measured as the inhibition of  $\beta$ -galactosidase reporter expression. Cytotoxicity was monitored by MTS staining. A chemiluminescent endpoint was used to determine the extent of  $\beta$ -galactosidase expression as a measure of HIV-1 infection of the cells. Once HIV-1 attaches to and enters the MAGI-CCR5 cells, HIV-1 Tat transactivates the LTR dependent –  $\beta$ -galactosidase enzyme to express  $\beta$ -galactosidase. Thus, there is a direct relationship between the level of HIV-1 infection and the level of  $\beta$ -galactosidase detected in the cells. At 48 hours post infection, plates were aspirated and PBS was added to each well. Gal-screens reagent (Tropix, Bedford, MA) was then added per the manufacturer's instructions for chemiluminescent detection of  $\beta$ -galactosidase activity and incubated at room temperature for 90 minutes. The resulting chemiluminescence signal was then read using a Microbeta Trilux luminescence reader (PerkinElmer/Wallac). At assay termination, the cytotoxicity assay plates were stained with the soluble tetrazolium-based dye MTS (CellTiter Reagent, Promega) to determine cell viability and quantify compound toxicity. MTS was metabolized by the mitochondrial enzymes of metabolically active cells to yield a soluble formazan product, allowing the rapid quantitative analysis of cell viability and compound cytotoxicity. The MTS was a stable solution that does not require preparation before use. At termination of the assay, 15  $\mu$ L of MTS reagent was added per well. The microtiter plates were then incubated 1.5-2 hrs at

37°C; the incubation interval was chosen based on empirically determined times for optimal dye reduction. The plates were read spectrophotometrically at 490/650nm with a Molecular Devices Vmax plate reader. Using an in-house computer program, IC<sub>50</sub> (50% inhibition of virus replication), IC<sub>90</sub> (90% inhibition of virus replication), TC<sub>50</sub> (50% cytotoxicity) and therapeutic index values (TI = TC<sub>50</sub>/IC<sub>50</sub>; also referred to as Antiviral Index or AI) were calculated. Raw data for both antiviral activity and cytotoxicity with a graphical representation of the data were provided in a printout summarizing individual compound activity.

*Anti-HIV Potency in PBM Cells.* Anti-viral activity in PBMC assays were conducted at Southern Research Institute. Fresh human PBMCs, seronegative for HIV and HBV, were isolated from screened donors (Biological Specialty Corporation, Colmar, PA). Cells were pelleted/washed 2-3 times by low speed centrifugation and re-suspension in PBS to remove contaminating platelets. The Leukophoresed blood was then diluted 1:1 with Dulbecco's Phosphate Buffered Saline (DPBS) and layered over 14 mL of Lymphocyte Separation Medium (LSM; Cellgro® by Mediatech, Inc.; density 1.078+/-0.002 g/ml; Cat.# 85-072-CL) in a 50 mL centrifuge tube and then centrifuged for 30 minutes at 600 X g. Banded PBMCs were gently aspirated from the resulting interface and subsequently washed 2X with PBS by low speed centrifugation. After the final wash, cells were enumerated by trypan blue exclusion and re-suspended at 1 x 10<sup>6</sup> cells/mL in RPMI 1640 supplemented with 15 % Fetal Bovine Serum (FBS), and 2 mM L-glutamine, 4 µg/mL Phytohemagglutinin (PHA, Sigma). The cells were allowed to incubate for 48-72 hours at 37°C. After incubation, PBMCs were centrifuged and re-suspended in RPMI 1640 with 15% FBS, 2 mM L-glutamine, 100 U/mL penicillin, 100 µg/mL streptomycin, and 20 U/mL recombinant human IL-2 (R&D Systems, Inc). IL-2 was included in the culture medium to

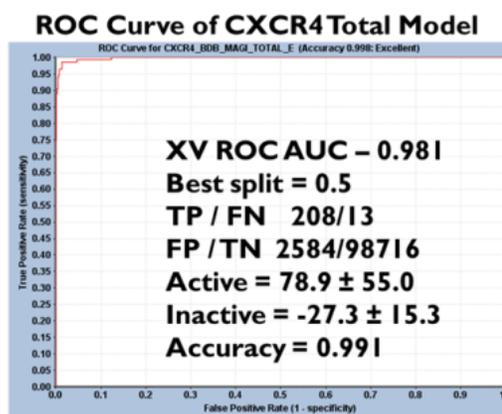
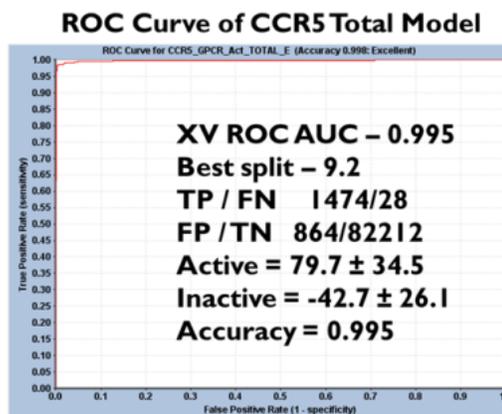
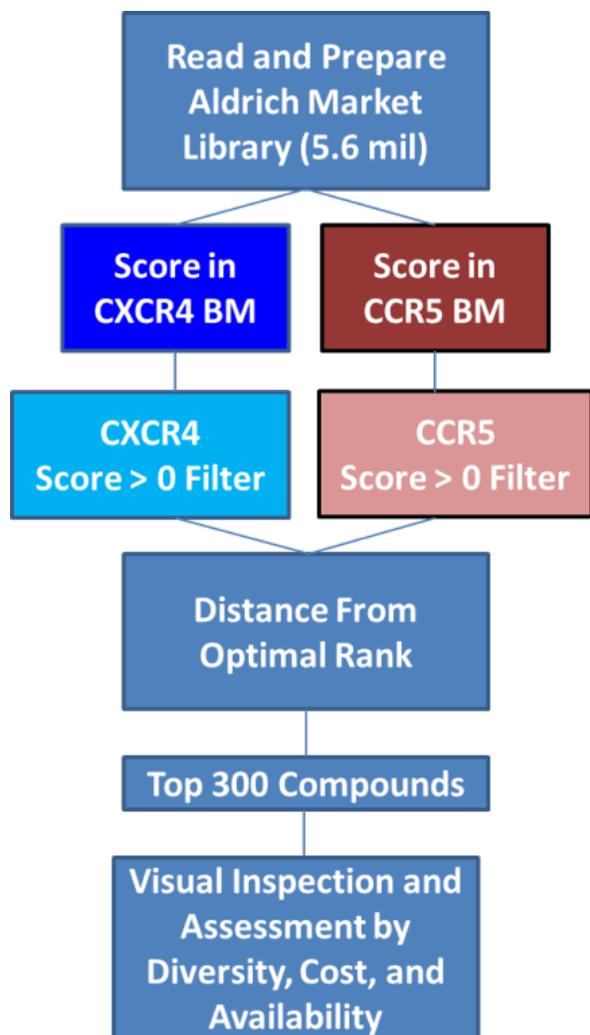
maintain the cell division initiated by the PHA mitogenic stimulation. PBMCs were maintained in this medium at a concentration of  $1-2 \times 10^6$  cells/mL with biweekly medium changes until used in the assay protocol. Cells were kept in culture for a maximum of two weeks before being deemed too old for use in assays and discarded. MDMs were depleted from the culture as the result of adherence to the tissue culture flask. For the standard PBMC assay, PHA stimulated cells from at least two normal donors were pooled (mixed together), diluted in fresh medium to a final concentration of  $1 \times 10^6$  cells/mL, and plated in the interior wells of a 96 well round bottom microplate at  $50 \mu\text{L}/\text{well}$  ( $5 \times 10^4$  cells/well) in a standard format developed by the Infectious Disease Research department of Southern Research Institute. Pooling (mixing) of mononuclear cells from more than one donor was used to minimize the variability observed between individual donors, which results from quantitative and qualitative differences in HIV infection and overall response to the PHA and IL-2 of primary lymphocyte populations. Each plate contained virus/cell control wells (cells plus virus), experimental wells (drug plus cells plus virus) and compound control wells (drug plus media without cells necessary for MTS monitoring of cytotoxicity). In this in vitro assay, PBMC viability remained high throughout the duration of the incubation period. Therefore, infected wells were used in the assessment of both antiviral activity and cytotoxicity. Test drug dilutions were prepared at a 2X concentration in microtiter tubes and  $100 \mu\text{L}$  of each concentration (nine total concentrations) were placed in appropriate wells using the standard format.  $50 \mu\text{L}$  of a predetermined dilution of virus stock was placed in each test well (final MOI  $\sim 0.1$ ). The PBMC cultures were maintained for seven days following infection at  $37^\circ\text{C}$ ,  $5\% \text{CO}_2$ . After this period, cell-free supernatant samples were collected for analysis of reverse transcriptase activity and/or p24 antigen content. Following removal of supernatant samples, compound cytotoxicity was measured by addition of MTS to the plates for

determination of cell viability. Wells were also examined microscopically and any abnormalities were noted. At assay termination, assay plates were stained with the soluble tetrazolium-based dye MTS (CellTiter 96 Reagent, Promega) to determine cell viability and quantify compound toxicity. The mitochondrial enzymes of metabolically active cells metabolized MTS to yield a soluble formazan product. This allows the rapid quantitative analysis of cell viability and compound cytotoxicity. The MTS was a stable solution that does not require preparation before use. At termination of the assay, 20  $\mu$ L of MTS reagent was added per well. The microtiter plates were then incubated 4-6 hrs at 37°C. The incubation intervals were chosen based on empirically determined times for optimal dye reduction. Adhesive plate sealers were used in place of the lids. The sealed plate was inverted several times to mix the soluble formazan product, and the plate was read spectrophotometrically at 490/650 nm with a Molecular Devices Vmax or SpectraMaxPlus plate reader.

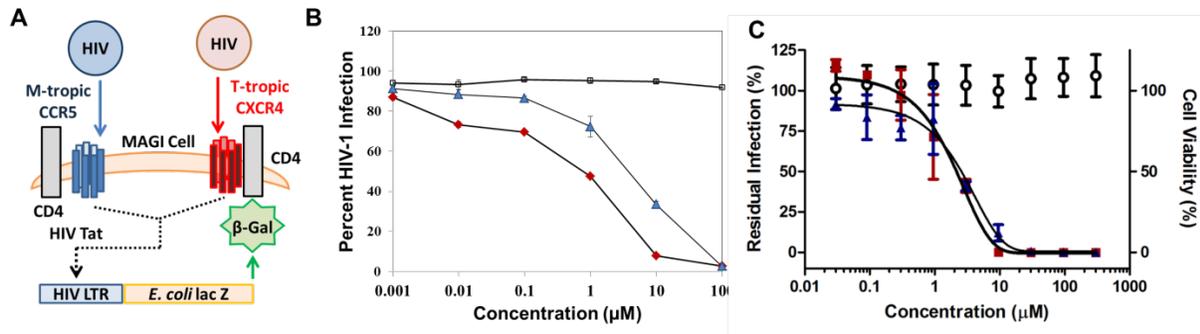
*Radio-Ligand Displacement Assay.* Displacement assays for  $^{125}$ I-MIP-1 $\beta$ :CCR5 and  $^{125}$ I-SDF-1:CXCR4 were conducted by Eurofins Cerep Panlabs at two concentrations using previously reported methods.<sup>14,15</sup>

*HIV-RT Inhibition Assay.* This assay was conducted by Southern Research Institute. Purified recombinant HIV (pNL4-3) heterodimeric (p66/p51) Reverse Transcriptase (RT) was purchased from a commercially available source. The assay was performed in 96-well filter plate, where RT activity was determined by the incorporation of radiolabeled deoxyribonucleotides into the newly synthesized DNA strand. The standard RT reaction mixture contained in vitro transcribed viral RNA derived from the HIV-1NL4-3 5'-LTR region (position 454 to 652) and primer that is

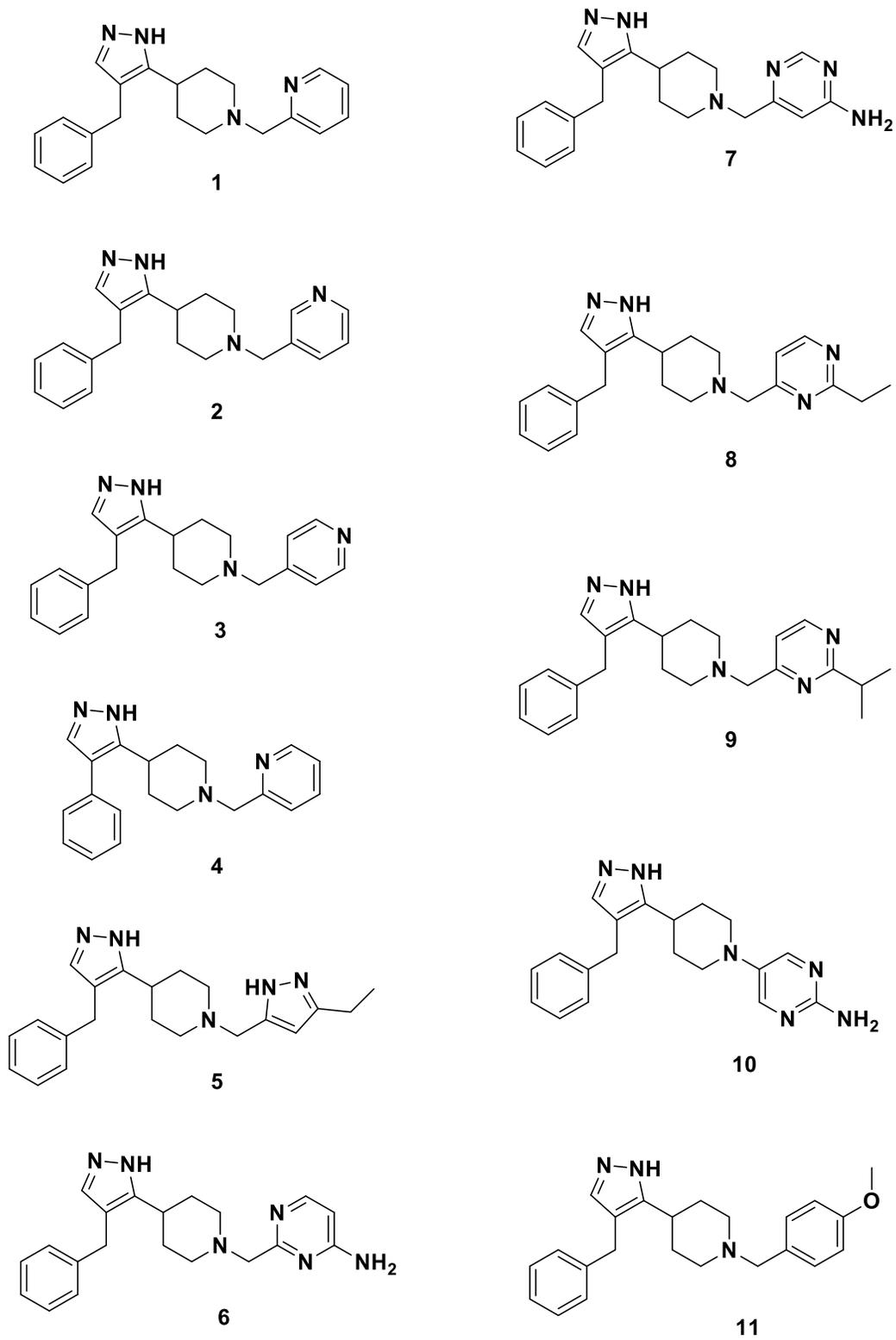
complementary to the primer binding site (PBS, nucleotide residues nucleotides 636 to 652), radiolabeled deoxyribonucleotide, dNTPs and reverse transcriptase. Briefly, the reaction was carried out in a volume of 50  $\mu$ l containing 50 mM Tris HCl, pH 7.8, 50 mM KCl, 5mM MgCl<sub>2</sub>, 1mM DTT, 50  $\mu$ M each of dATP, dCTP, dGTP, 50 nM dTTP, 1 $\mu$ C<sub>i</sub> of [<sup>3</sup>H] dTTP (70-90C<sub>i</sub>/mM) and 5 nM template/primer. The reaction was initiated by the addition of 10 nM RT. For compound screening, serially diluted test articles were added to the reaction followed by the addition of RT. The reaction mixture is incubated at 37°C for 1h, and then quenched by the addition of ice-cold trichloroacetic acid (TCA) to the final concentration of 10%. The plate was incubated at 4°C for 1h to precipitate the synthesized DNA, then rinsed 3-times with 10% TCA and 1 time with 70% ethanol. After addition of 25  $\mu$ l scintillation fluid to completely dried wells, radioactivity was counted by MicroBeta scintillation counter. The reduction of radioactivity represents the potency of compound inhibition.



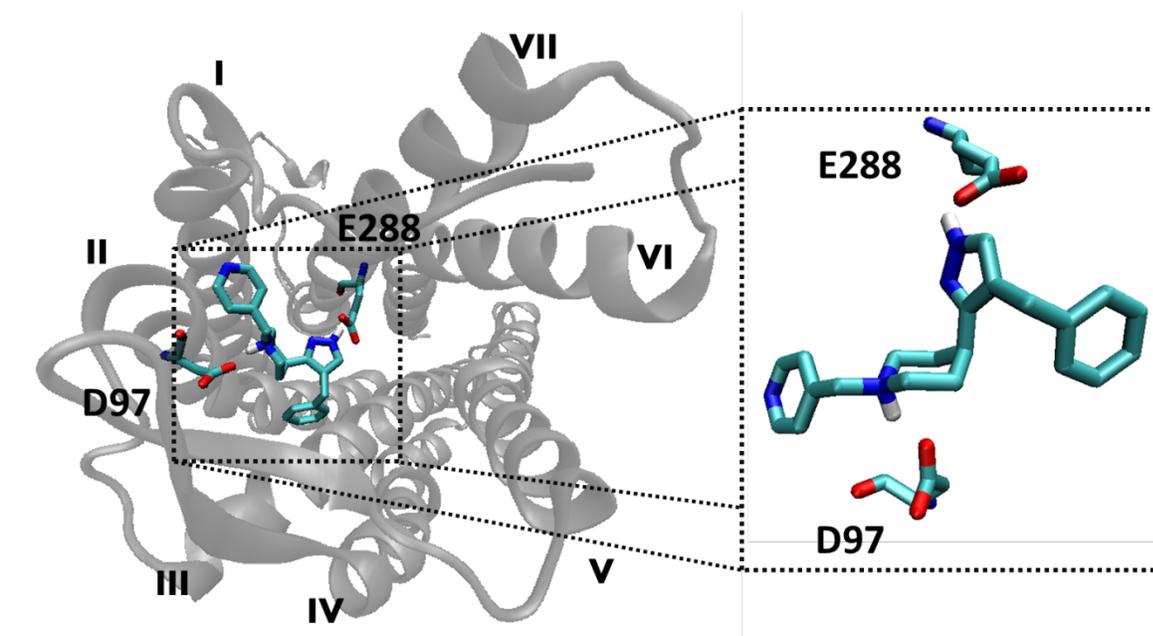
**SI Figure 1.** Workflow and Bayesian statistics employed in virtual screening.



**SI Figure 2.** Anti-HIV-1 activity of compound **3** in (B) MAGI and (C) PBMC assays. A) Representation of the MAGI assay. B) Anti-HIV activity of compound **3** against M-tropic HIV-1<sub>Ba-L</sub> virus (CCR5, solid blue triangles) or T-tropic HIV-1<sub>IIIB</sub> virus (CXCR4, solid red diamonds) and cytotoxicity measurements (open black squares). C) Anti-HIV activity of compound **3** against M-tropic HIV-1<sub>Ba-L</sub> virus (CCR5, solid blue triangles) or T-tropic HIV-1<sub>IIIB</sub> virus (CXCR4, solid red diamonds) in PBM cells and cytotoxicity measurements (open black squares).

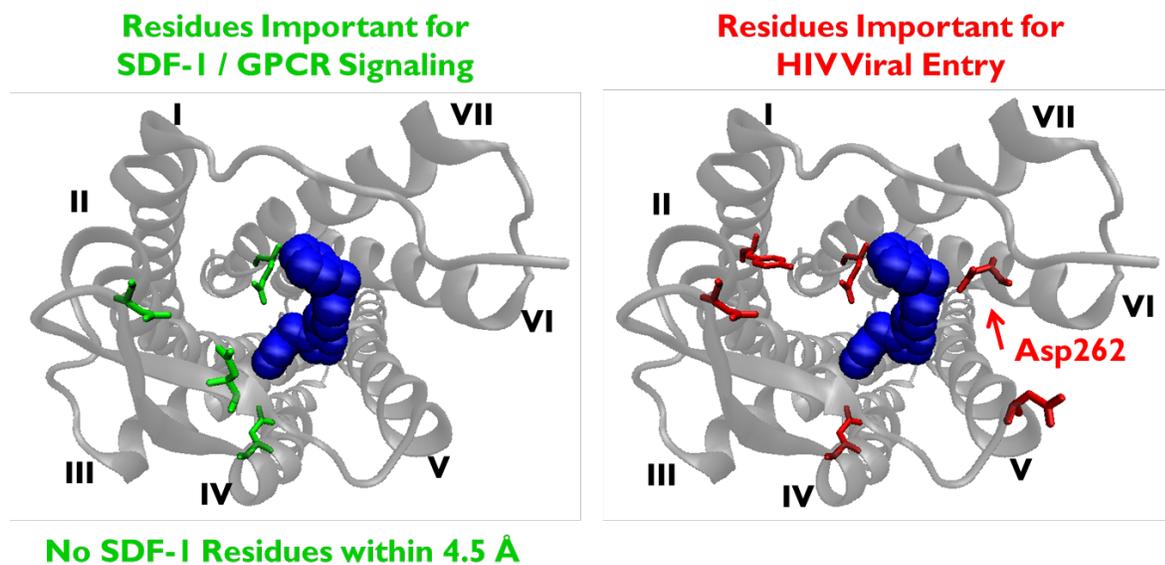


SI Figure 3. Structures of active compounds from virtual screen.

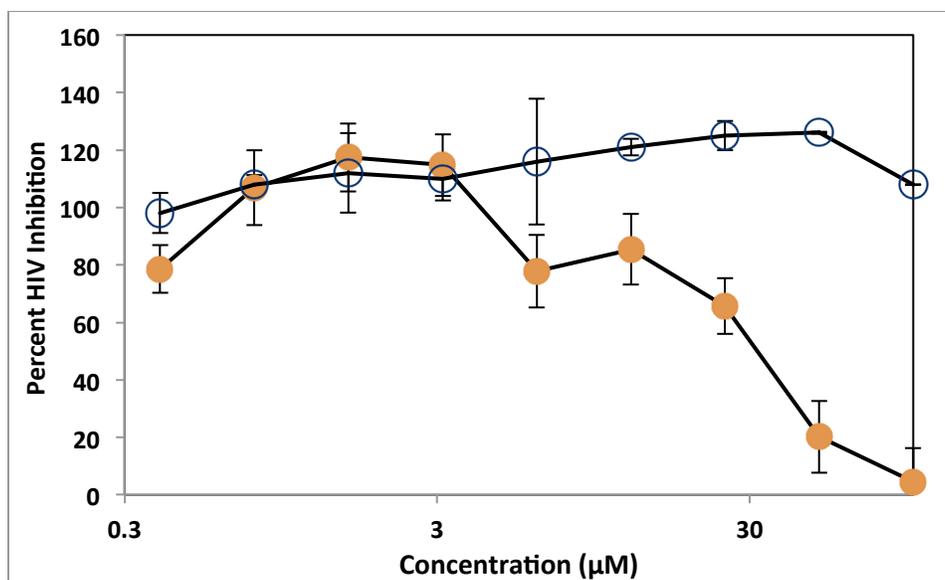


**SI Figure 4.** Model of compound **3** binding to CXCR4 derived from the IT1t crystal structure.

*Inset:* **3** forms a high-energy conformation when modeled into the minor sub-pocket.



**SI Figure 5.** Comparison of compound **3** pose derived from CXCR4:CVX15 structure with residues critical for SDF-1 binding and HIV-1 infection. Residues identified by mutagenesis obtained from previous studies.<sup>16</sup>



**Figure 6.** Anti-HIV-2 activity of compound **3** in PBM cells. The anti-HIV-2 inhibition is given as orange circles, and the cytotoxicity is provided as open circles.

<b>SI Table 1. Bioactivities of Compound 3</b>		
<b>Anti-HIV-1 Assay</b>	<b>Compound 3</b>	<b>Controls</b>
MAGI – HIV-1 <sub>Ba-L</sub> (CCR5)	IC <sub>50</sub> = 3.8 μM	[TAK-779] <sup>a</sup> IC <sub>50</sub> = 5 nM
MAGI – HIV-1 <sub>IIIIB</sub> (CXCR4)	IC <sub>50</sub> = 0.8 μM	[AMD3100] <sup>b</sup> IC <sub>50</sub> = 2 nM
MAGI – Toxicity	TC <sub>50</sub> > 300 μM	
PBMC – HIV-1 <sub>Ba-L</sub> (CCR5)	IC <sub>50</sub> = 2.3 μM	[AZT] <sup>d</sup> IC <sub>50</sub> = 4.1 nM
PBMC – HIV-1 <sub>IIIIB</sub> (CXCR4)	IC <sub>50</sub> = 2.3 μM	
PBMC - Toxicity	TC <sub>50</sub> > 100 μM	
<b>CCR5 Activity Assay</b>	<b>Compound 3</b>	<b>Controls</b>
HeLa-C14:JRFL CD4-Dependent Fusion	IC <sub>50</sub> = 36 μM	[TAK-779] <sup>a</sup> IC <sub>50</sub> = 1 nM [AMD3100] <sup>b</sup> EC <sub>50</sub> > 1.25 μM
<sup>125</sup> I-MIP-1β Displacement	38% at 100 μM/ 88% at 1mM	[MVC] <sup>c</sup> 100% at 100 μM
<b>CXCR4 Activity Assay</b>	<b>Compound 3</b>	<b>Controls</b>
Hela/67:Lai CD4-Dependent Fusion	EC <sub>50</sub> = 70 μM	[AMD3100] <sup>b</sup> EC <sub>50</sub> = 12 nM
ACTOneX4:HIV-8x CD4-Independent Fusion	EC <sub>50</sub> = 52 μM	[AMD3100] <sup>b</sup> EC <sub>50</sub> = 10 nM
PBMC – HIV-2	IC <sub>50</sub> = 31 μM	[AZT] <sup>d</sup> IC <sub>50</sub> = 12 nM
Nef-M1:CXCR4 MMD Inhibition	IC <sub>50</sub> = 1.2 μM	[AMD3100] <sup>b</sup> IC <sub>50</sub> = 0.5 μM [MVC, AZT] <sup>c,d</sup> > 10 μM
<sup>125</sup> I-SDF-1 Displacement	0% at 1mM	[AMD3100] <sup>b</sup> IC <sub>50</sub> = 0.45 μM
<b>NNRTI Activity Assay</b>	<b>Compound 3</b>	<b>Controls</b>
VSV-G Pseudovirus	EC <sub>50</sub> = 1.3 μM	[AMD3100] <sup>b</sup> EC <sub>50</sub> > 6.25 [EFV] <sup>e</sup> EC <sub>50</sub> = 0.38 nM
Isolated HIV-RT Inhibition	IC <sub>50</sub> = 9.0 μM	[NVP] <sup>f</sup> IC <sub>50</sub> = 0.26 μM
Multi-cycle with Mutant HIV-RT <sup>g</sup>	EC <sub>50</sub> (WT) = 0.9 μM EC <sub>50</sub> (Mut <sup>g</sup> ) = 19 μM (21x)	[EFV] <sup>e</sup> EC <sub>50</sub> (WT) = 0.046 nM [EFV] <sup>e</sup> EC <sub>50</sub> (Mut <sup>g</sup> ) = 4 nM (71x)
<sup>a</sup> Tak-779 is a Dual CCR5 – CCR2 antagonist (Takeda Chemical Industries, Ltd.) <sup>b</sup> AMD3100 is a CXCR4 antagonist FDA-approved for stem cell mobilization (Mozobil, Genzyme) <sup>c</sup> MVC is the abbreviation for FDA-approved CCR5 antagonist Maraviroc (Selzentry, ViiV Healthcare) <sup>d</sup> AZT is the abbreviation for the FDA-approved NRTI azidothymine (Retrovir, ViiV Healthcare) <sup>e</sup> EFV is the abbreviation for the FDA-approved NNRTI Efavirenz (Sustiva, Bristol-Myers-Squibb) <sup>f</sup> NVP is the abbreviation for the FDA-approved NNRTI Nevirapine (Boehringer Ingelheim) <sup>g</sup> Mutant HIV-RT are NNRTI resistant Lys103Gln and Tyr181Cys variants		

**SI Table 2.** MAGI assay activity of compounds from screen.

Compound Number <sup>a</sup>	MAGI <sub>IIIb</sub> (CXCR4)		MAGI <sub>Ba-L</sub> (CCR5)		TC <sub>50</sub> (μM)
	IC <sub>50</sub> (μM)	Bayesian Score	IC <sub>50</sub> (μM)	Bayesian Score	
1 <sup>b</sup>	25	15	17	12	>300
2	5.8	-3	16	16	>300
3 <sup>b</sup>	0.8	0	3.8	12	>300
4	44	-7	64	-1	>300
5	68	-5	60	20	>300
6	81	-3	83	8	>300
7	13	1	17	9	>100
8	57	0	66	16	>100
9	6.4	0	13	22	>100
10	105	-6	95	10	>300
11	>200	-12	72	15	>300

a. All active compounds were >95% purity by HPLC.

b. Results from a re-synthesized non-commercial sample.

**SI Table 3.** Radio-labeled chemokine ligand displacement results.

Compound Number	<sup>125</sup> I-SDF-1 CXCR4 <sup>d</sup>		<sup>125</sup> I-MIP-1β CCR5 <sup>e</sup>	
	% Displacement		% Displacement	
	100 μM	1 mM	100 μM	1 mM
1 <sup>a</sup>	0	9	23	59
3 <sup>a</sup>	0	0	38	85
9 <sup>b</sup>	0	19	14	70
11 <sup>b</sup>	0	0	65	86

- a. Resynthesized sample at >95% purity by HPLC and <sup>1</sup>H NMR.  
b. Results from commercial sample >95% purity by HPLC.  
c. Absolute experimental values were less than -50%.  
d. Control IC<sub>50</sub> of 0.13 nM for non-radiolabeled SDF-1.  
e. Control IC<sub>50</sub> of 0.21 - 0.26 nM for non-radiolabeled MIP-1β.

## REFERENCES

- (1) Navratilova, I.; Besnard, J.; Hopkins, A. L. *ACS Med. Chem. Lett.* **2011**, *2*, 549.
- (2) Besnard, J.; Ruda, G. F.; Setola, V.; Abecassis, K.; Rodriguiz, R. M.; Huang, X. P.; Norval, S.; Sassano, M. F.; Shin, A. I.; Webster, L. A.; Simeons, F. R.; Stojanovski, L.; Prat, A.; Seidah, N. G.; Constam, D. B.; Bickerton, G. R.; Read, K. D.; Wetsel, W. C.; Gilbert, I. H.; Roth, B. L.; Hopkins, A. L. *Nature* **2012**, *492*, 215.
- (3) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B. *Nucleic Acids Research* **2012**, *40*, D1100.
- (4) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. *Nucleic Acids Research* **2007**, *35*, D198.
- (5) Debnath, B.; Xu, S.; Grande, F.; Garofalo, A.; Neamati, N. *Theranostics* **2013**, *3*, 47.
- (6) Tan, Q.; Zhu, Y.; Li, J.; Chen, Z.; Han, G. W.; Kufareva, I.; Li, T.; Ma, L.; Fenalti, G.; Li, J. *Science* **2013**, *341*, 1387.
- (7) Wu, B.; Chien, E. Y.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. *Science* **2010**, *330*, 1066.
- (8) Kertesz, D. J.; Brotherton-Pleiss, C.; Yang, M.; Wang, Z.; Lin, X.; Qiu, Z.; Hirschfeld, D. R.; Gleason, S.; Mirzadegan, T.; Dunten, P. W. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4215.
- (9) Ren, J.; Milton, J.; Weaver, K. L.; Short, S. A.; Stuart, D. I.; Stammers, D. K. *Structure* **2000**, *8*, 1089.
- (10) Prosser, A. R.; Liotta, D. C. *Tet. Lett.* **2014**.
- (11) Li, Z.; Terry, B.; Olds, W.; Protack, T.; Deminie, C.; Minassian, B.; Nowicka-Sans, B.; Sun, Y.; Dicker, I.; Hwang, C. *Antimicrob. Agents Chemother.* **2013**, *57*, 5500.
- (12) Lin, P.-F.; Nowicka-Sans, B.; Terry, B.; Zhang, S.; Wang, C.; Fan, L.; Dicker, I.; Gali, V.; Higley, H.; Parkin, N. *Antimicrob. Agents Chemother.* **2008**, *52*, 1759.
- (13) Li, Z.; Zhou, N.; Sun, Y.; Ray, N.; Lataillade, M.; Hanna, G. J.; Krystal, M. *Antimicrob. Agents Chemother.* **2013**, *57*, 4172.
- (14) Raport, C. J.; Gosling, J.; Schweickart, V. L.; Gray, P. W.; Charo, I. F. *J. Biol. Chem.* **1996**, *271*, 17161.
- (15) Valenzuela-Fernández, A. n.; Planchenault, T.; Baleux, F.; Staropoli, I.; Le-Barillec, K.; Leduc, D.; Delaunay, T.; Lazarini, F.; Virelizier, J.-L.; Chignard, M. *J. Biol. Chem.* **2002**, *277*, 15677.
- (16) Brelot, A.; Heveker, N.; Montes, M.; Alizon, M. *J. Biol. Chem.* **2000**, *275*, 23736.