Crystal Structures of HIV-1 Reverse Transcriptase with Picomolar Inhibitors Reveal Key Interactions for Drug Design

Kathleen M. Frey,[‡] Mariela Bollini,[†] Andrea C. Mislak,[‡] José A. Cisneros,[†] Ricardo Gallardo-Macias,[†] William L. Jorgensen,[†,*] and Karen S. Anderson[‡,*]

[†] *Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107*

[‡] *Department of Pharmacology, Yale University School of Medicine, New Haven, CT 06520-8066*

Supporting Information

Table of Contents

Materials and Methods for Crystallography

Recombinant RT52A enzyme was expressed in *E. coli* BL21(DE3) Gold cells upon induction with 1 mM IPTG at an OD_{600} ~ 0.8-1.0. Cells pellets were lysed with 1 x Bugbuster (EMD-Millipore) and 1500 units of DNase diluted in a buffer of 50 mM Tris, 0.5 M NaCl (pH 8.0). The enzyme was purified using Cobalt Affinity chromatography. The *n*-terminal histidine tag located on the p51 subunit was cleaved overnight with HRV 3C protease at 4°C. RT52A was further purified with Ion exchange chromatography (Q-sepharose resin) then desalted into a final buffer of 50 mM Tris, 50 mM NaCl, 10% (v/v) glycerol, and 2 mM DTT. Purified enzyme was concentrated to ~ 3 mg/ml and stored in -80°C until crystal tray setup. ESI-MS was used to characterize each enzyme purification experiment by approximating the ratio of p66:p51 subunit and confirming the histidine tag cleavage.

Samples for crystallization were prepared by incubating the RT52A enzyme (~ 3 mg/ml) with the inhibitor (0.5 mM, dissolved in DMSO) for 1 hour on ice. Following incubation, the RT52A:inhibitor complex was clarified of any inhibitor precipitate by high speed centrifugation at 4°C. Clarified RT52A:inhibitor sample was transferred to a 10K MWCO microcon and concentrated to a final protein concentration of ~ 18-20 mg/ml. Hanging drop vaporization using a protein:well solution ratio of 1:1 was used to screen crystallization conditions. Optimized conditions for crystal growth included 20% (w/v) PEG 8000, 100 mM ammonium sulfate, 15 mM magnesium sulfate, 5 mM spermine-HCl, and 50 mM citric acid pH 5.5 or 50 mM HEPES pH 7.0. Crystals grew in approximately 3-5 days at 4°C.

Several crystals of RT52A:**1** and RT52A:**2** were screened for diffraction at Brookhaven NSLS on beam line X29A. High-resolution data sets for the best diffracting, single crystals were collected, indexed, integrated, and scaled into the appropriate space group using HKL2000[1]. In order to obtain phases, molecular replacement (MR) was performed with Phaser[2,3] using a previously determined RT:NNRTI structure (PDB code: 1S9E) as the probe model[4]. The final probe model used for MR had both the inhibitor and water molecule coordinates removed from the structure, while the atoms for all residues were restored to an occupancy=1.0. The program COOT[5] was used for model building into the electron density. Maximum-likelihood restrained refinement in Phenix[6] was used to refine the structure after each cycle of model building and modification until acceptable *R*-factors, geometry statistics (rmsds for ideal bond lengths and angles), and Ramachandran statistics were achieved for the respective structures. Unbiased omit F_o-F_c electron density maps were generated using simulated annealing (using 28 boxes/omit regions) within the "create omit map" feature in Phenix Autobuild.[7] PyMOL[8] was used to visualize and analyze the structures, in addition to C^{α} backbone structure alignment and rmsd calculations.

Table S1. *Statistics for data collection and refinement of crystal structures*

Complex	***RT52A:1***	***RT52A:2***
PDB Code	**4H4M**	**4H4O**
X-ray Source	NSLS X29A	NSLS X29A
Wavelength, Å	1.075	1.075
Space group	C2	C2
No. Molecules in the asymmetric unit	1	1
Unit cell, a,b,c in Å (α,β,γ in °)	a=225.133, b=69.208, c=104.138 (α=γ=90°, β=103.12)	a=225.451, b=69.575, c=104.401 (α=γ=90°, β=106.19)
Resolution Range, Å	50 – 2.85	50.0 – 2.90
Last shell, Å	2.90 – 2.85	2.95 – 2.90
R_{sym}, (last shell)	0.097 (0.546)	0.073 (0.596)

Completeness, % (last shell %)	98.5 (98.4)	99.8 (98.8)
No. of Reflections	35716	34776
Redundancy (last shell)	3.8 (3.7)	3.8 (3.7)
Avg. I/σ (last shell)	16.6 (2.7)	16.2 (2.6)
Refinement Statistics		
Total Number of Atoms (Protein/Inhibitor/Solvent)	8075 (7996/30/45)	8092 (8003/30/49)
R_{free}, R_{factor}	0.2733, 0.2334	0.2650, 0.2324
Rms deviation bond lengths (Å), angles (°)	0.003, 0.91	0.003, 0.88
Avg. B-factor (Protein/Inhibitor/Solvent)	60.30 (60.40/45.96/34.20)	164.10 (164.30/143.29/135.10)
Ramachandran Favored, Allowed, Outliers (%)	91.8, 7.0, 0.6	92.8, 6.0, 1.2

Table S2. *Comparison of RT-Catechol Diether Structures with other RT-NNRTI structures. RMSD calculations are based on atom alignment using rms_cur in Pymol. All atoms for each structure were used to calculate the RMSD.*

Structure	**Ref.**	**PDB Code**	**Resolution (Å)**	**RMSD C^{α} RT52A:1** (Green)	**RMSD C^{α} RT52A:2** (Cyan)	**Ribbon Color** (See Figure S4)
RT52A:TMC278 (rilpivirine)	9	2ZD1	1.80	2.885	2.835	Brown
RT:R129385	4	1S9E	2.60	1.916	1.874	Pink
RT:R221239	10	2BE2	2.43	2.318	2.340	Blue
RT:UC781	11	1RT4	2.90	3.016	3.024	Purple
RT:TMC125 (etravirine)	12	3M8P	2.67	2.500	2.536	Orange
RT:nevirapine	13	3HVT	2.90	2.974	2.832	Yellow

Table S3. *Hydrogen bonding interaction distances between Lys102/Lys103 and the uracil group of* ***1*** *and* ***2****.*

Complex	C=O (K103) – NH (uracil)	NH (K103) – C=O (uracil)	NH_3^+ (K102) – C=O4 (uracil)
RT52A:**1**	3.26 Å	3.17 Å	3.47 Å
RT52A:**2**	2.83 Å	2.91 Å	3.18 Å

***Figure S1**: Alignment of RT52A:**1** and RT52A:**2** structures. (A) C^{α} trace for structures of RT52A bound to catechol diether compounds **1** (green) and **2** (pink). (B) View of the NNBP for the aligned structures with interacting residues in cartoon and stick representation. Structures differ by an rmsd of 1.205 Å (with no atoms rejected during alignment in Pymol).*

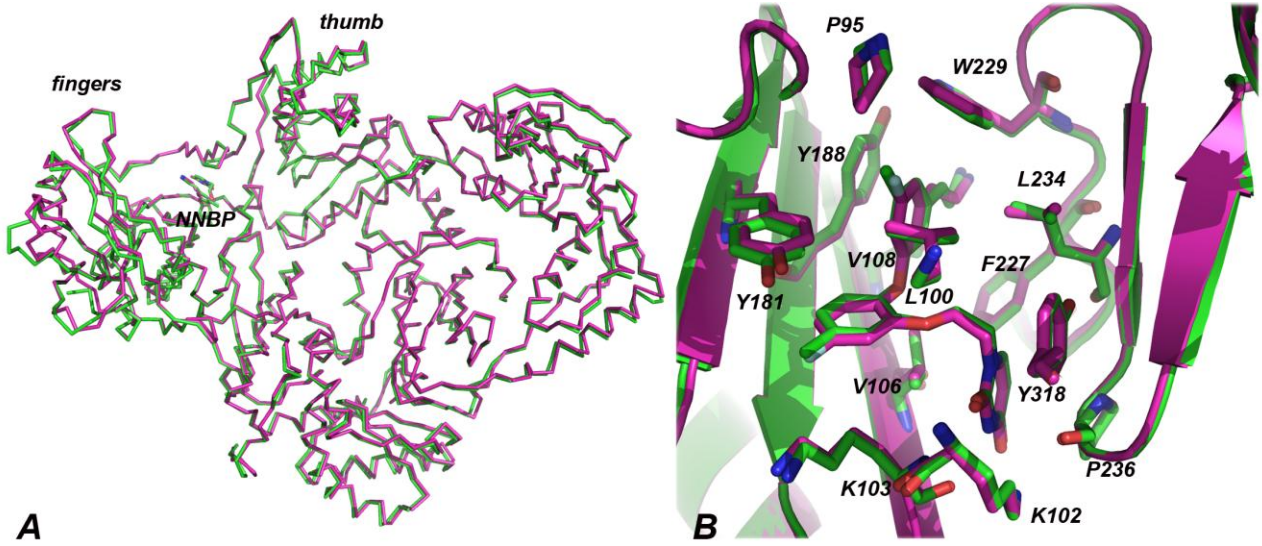
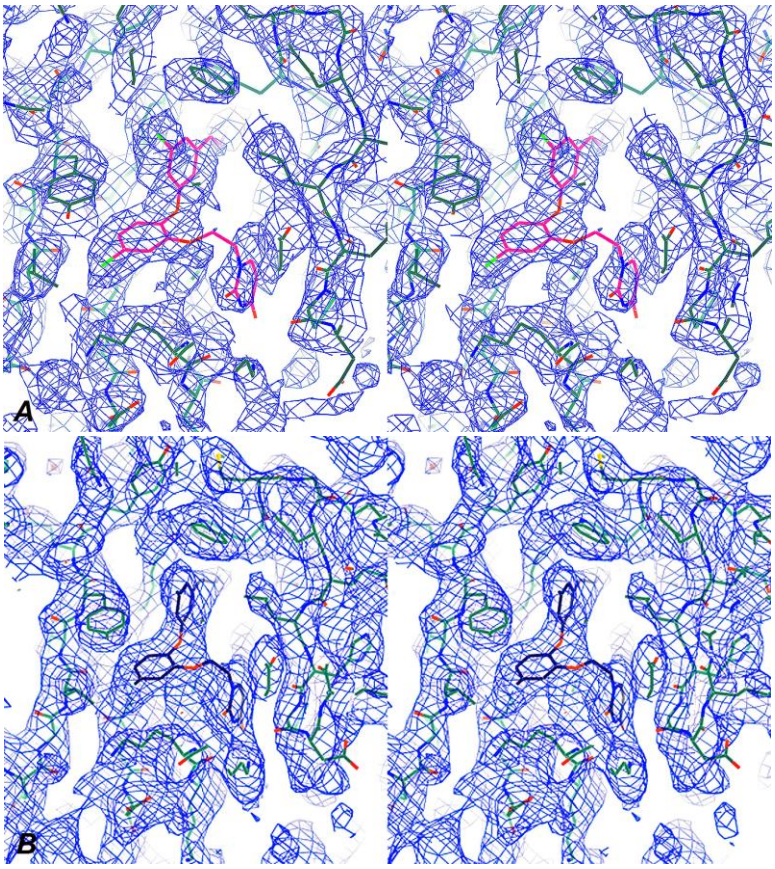


***Figure S2**: Stereo view of electron density maps for the NNBP of RT52A-catechol diether complexes. Final $2F_o$-F_c electron density maps (blue) for RT52A:**1** (A) and RT52A:**2** (B) complexes at 1.0 σ contour level. Omit F_o-F_c electron density (gray) generated by simulated annealing for RT52A:**1** (C) and RT52A:**2** (D) at 3.0 σ contour level.*



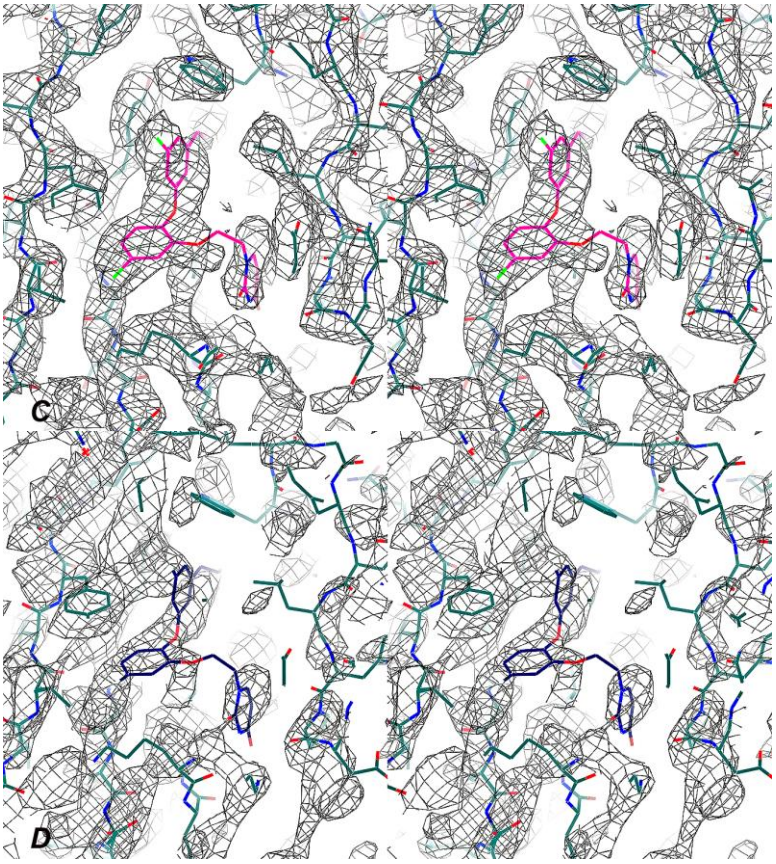


Figure S3. *Mono view of electron density maps for the NNBP of RT52A-catechol diether complexes. $2F_o$-F_c electron density maps for RT52A:**1** (A) and RT52A:**2** (B) complexes at 1.0 σ contour level. Omit F_o-F_c electron density generated by simulated annealing for RT52A:**1** (C) and RT52A:**2** (D) at 3.0 σ contour level.*

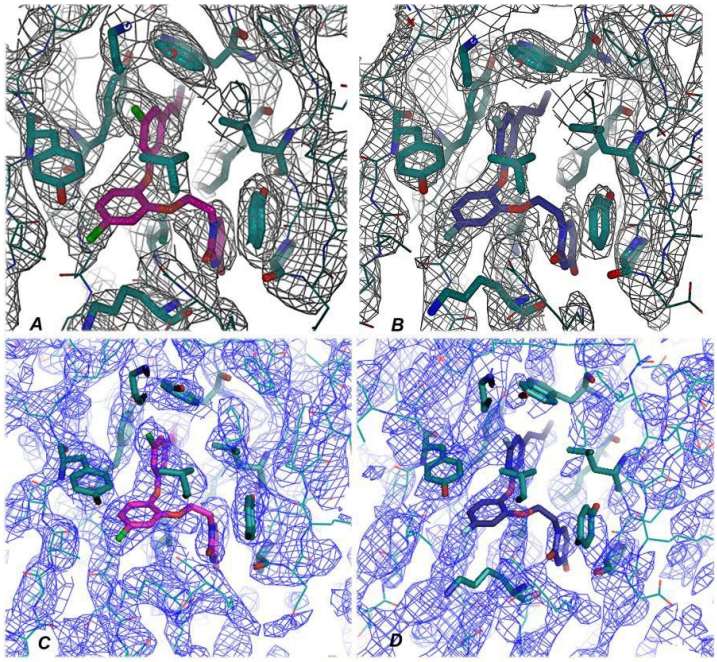


Figure S4. *Stereo view of the crystal structure for* ***2*** *complexed with HIV-RT. Multiple contacts with residues in the NNBP are apparent; the dashed lines highlight the hydrogen bonds with K102 and K103.*

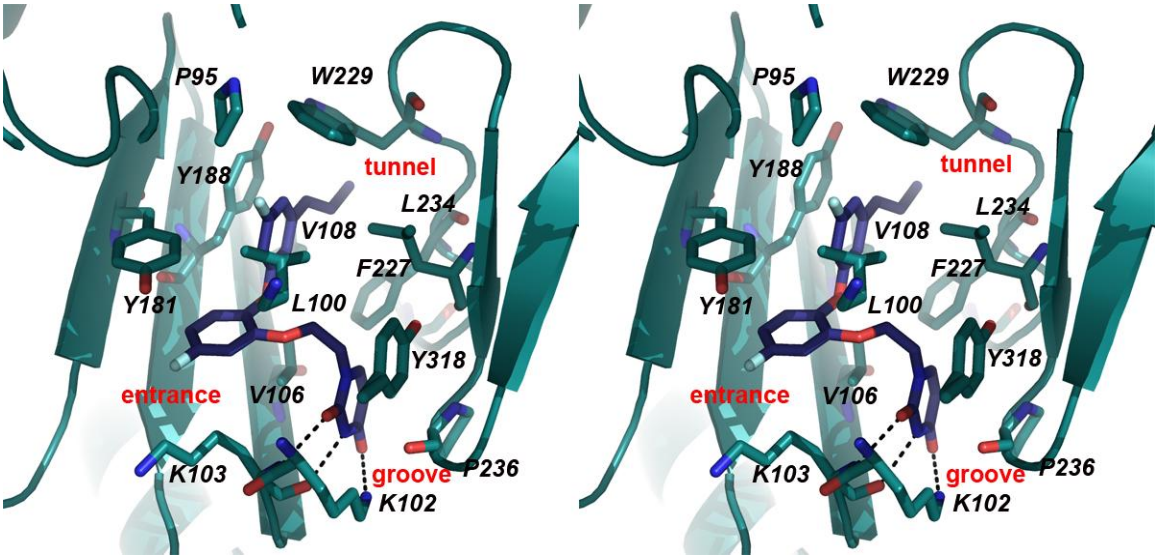
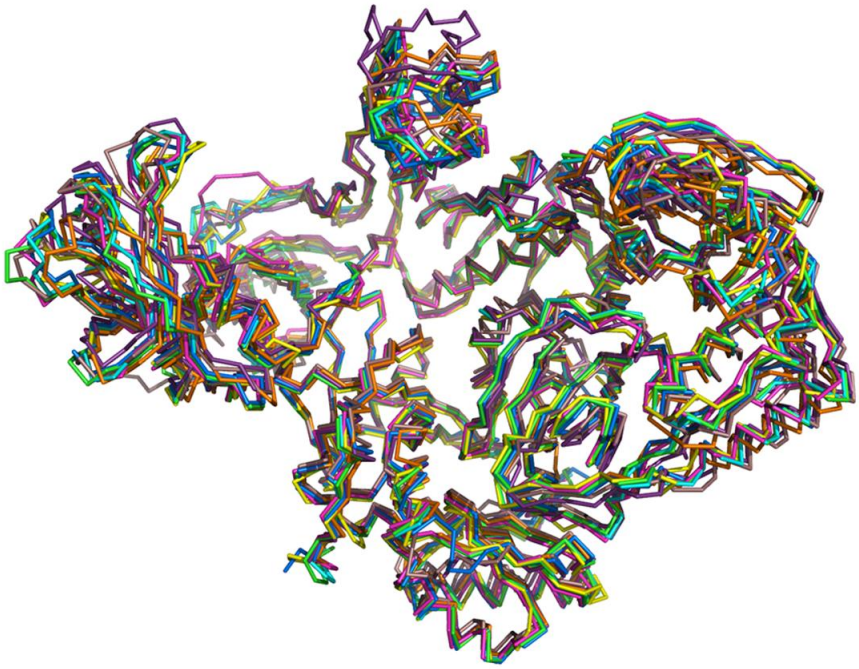


Figure S5: *C^{α} Trace for the RT-NNRTI complexes used for the structural comparison in Table S2.*



Crystallography References

(1) *Otwinowski, Z., Minor, W. Methods in Enzymology* ***1997****, 276, 307-326.*

(2) *McCoy, A. J. Acta Crystallogr D Biol Crystallogr* ***2007****, 63, 32-41.*

(3) *McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. J Appl Crystallogr* ***2007****, 40, 658-674.*

(4) *Das, K.; Clark, A. D., Jr.; Lewi, P. J.; Heeres, J.; De Jonge, M. R.; Koymans, L. M.; Vinkers, H. M.; Daeyaert, F.; Ludovici, D. W.; Kukla, M. J.; De Corte, B.; Kavash, R. W.; Ho, C. Y.; Ye, H.; Lichtenstein, M. A.; Andries, K.; Pauwels, R.; De Bethune, M. P.; Boyer, P. L.; Clark, P.; Hughes, S. H.; Janssen, P. A.; Arnold, E. J Med Chem* ***2004****, 47, 2550-60.*

(5) *Emsley, P.; Cowtan, K. Acta Crystallogr D Biol Crystallogr* ***2004****, 60, 2126-32.*

(6) *Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; measure the inhibition of the enzymatic activity of HIV-RT Terwilliger, T. C.; Zwart, P. H. Acta Crystallogr D Biol Crystallogr* ***2010****, 66, 213-21.*

(7) *Terwilliger, T. C.; Grosse-Kunstleve, R. W.; Afonine, P. V.; Moriarty, N. W.; Zwart, P. H.; Hung, L. W.; Read, R. J.; Adams, P. D. Acta Crystallogr D Biol Crystallogr* ***2008****, 64, 61-9.*

(8) *The Pymol Molecular Graphics System, Version 1.5.0.4; Schrödinger, LLC.*

(9) *Das, K.; Bauman, J. D.; Clark, A. D., Jr.; Frenkel, Y. V.; Lewi, P. J.; Shatkin, A. J.; Hughes, S. H.; Arnold, E. Proc Natl Acad Sci U S A* ***2008****, 105, 1466-71.*

(10) *Himmel, D. M.; Das, K.; Clark, A. D., Jr.; Hughes, S. H.; Benjahad, A.; Oumouch, S.; Guillemont, J.; Coupa, S.; Poncelet, A.; Csoka, I.; Meyer, C.; Andries, K.; Nguyen, C. H.; Grierson, D. S.; Arnold, E. J Med Chem* ***2005****, 48, 7582-91.*

(11) *Ren, J.; Esnouf, R. M.; Hopkins, A. L.; Warren, J.; Balzarini, J.; Stuart, D. I.; Stammers, D. K. Biochemistry* ***1998****, 37, 14394-403.*

(12) *Kertesz, D. J.; Brotherton-Pleiss, C.; Yang, M.; Wang, Z.; Lin, X.; Qiu, Z.; Hirschfeld, D. R.; Gleason, S.; Mirzadegan, T.; Dunten, P. W.; Harris, S. F.; Villasenor, A. G.; Hang, J. Q.; Heilek, G. M.; Klumpp, K. Bioorg Med Chem Lett, 20, 4215-8.*

(13) *Wang, J.; Smerdon, S. J.; Jager, J.; Kohlstaedt, L. A.; Rice, P. A.; Friedman, J. M.; Steitz, T. A. Proc Natl Acad Sci U S A* ***1994****, 91, 7242-6.*

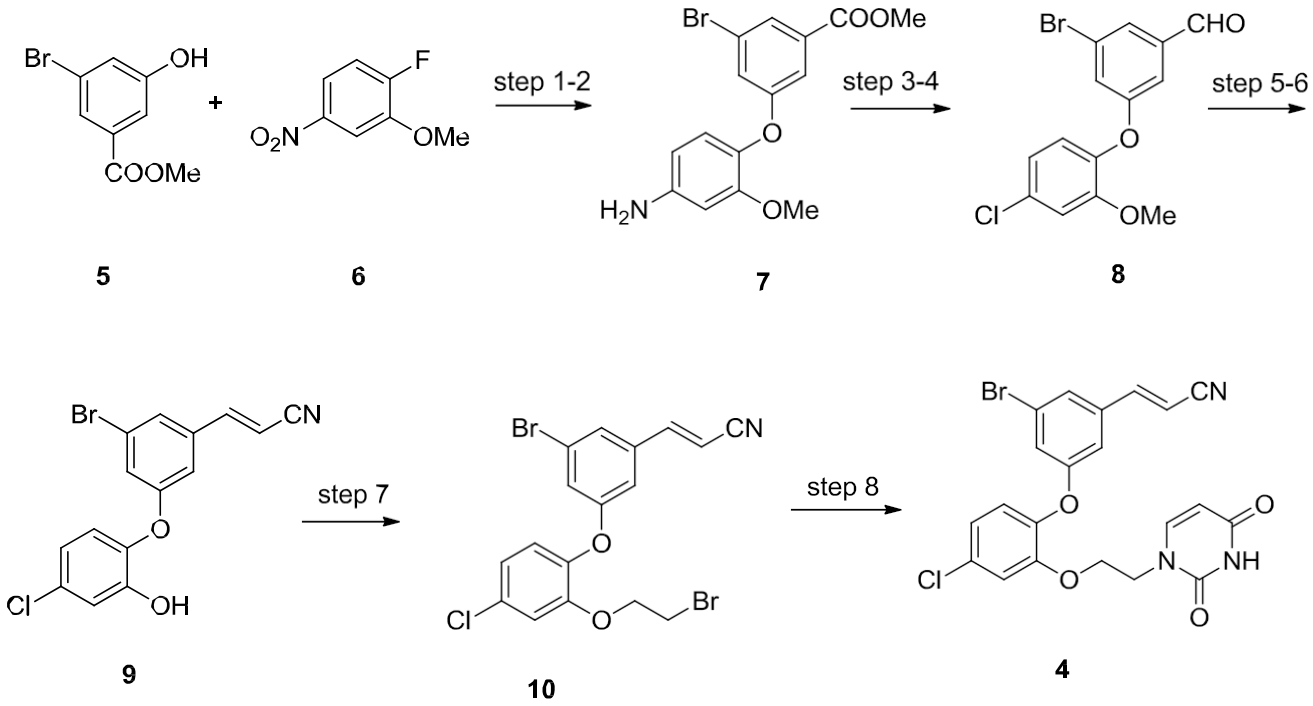
Synthesis of Compound 4 and Enzymatic Assay

General Information

NMR spectra were recorded on a Bruker Avance DRX-500 (500 MHz) and DRX-400 (400 MHz) instruments. Column chromatography was carried out using CombiFlash over redisep column cartridges employing Merck silica gel (Kieselgel 60, 63-200 μm). Precoated silica gel plates F-254 were used for thin-layer analytical chromatography. Mass determination was performed using Waters Xevo QTOF equipped with Z-spray electrospray ionization source. The purity (≥ 95%) of final synthesized compound was determined by reverse phase HPLC, using a Waters 2487 dual λ absorbance detector with a Waters 1525 binary pump and a Phenomenex Luna 5μ C18(2) 250 x 4.6 mm column. Sample was run at 1 mL/min using gradient mixtures of 5-100% of water with 0.1% trifluoroacetic acid (TFA) (A) and 10:1 acetonitrile:water with 0.1% TFA (B) for 22 min followed by 3 min at 100% B.

Compounds **1** – **3** were previously reported [Bollini, M.; Domaoal, R. A.; Thakur, V. V.; Gallardo-Macias, R.; Spasov, K. A.; Anderson, K. S.; Jorgensen, W. L. *J. Med. Chem.* ***2011***, *54*, 8582].

Synthesis of compound 4



- Step 1

A mixture of 1-fluoro-2-methoxy-4-nitrobenzene **5** (1.57 g, 9.7 mmol), methyl 3-bromo-5-hydroxybenzoate (**6**) (2.12 g, 9.7 mmol) in DMSO (3.0 mL) and anhydrous K_2CO_3 (0.41 g, 3.0 mmol) was heated at 90 °C for 5 h. The mixture was poured into ice water and extracted with EtOAc (3 x 50 mL). The organic layer was sequentially washed with brine (2 x 75 mL), dried over anhydrous Na_2SO_4, and concentrated in vacuo. The residue was purified by column chromatography to give methyl 3-bromo-5-(2-methoxy-4-nitrophenoxy)benzoate **6a** (3.5 g, 100%). LC-MS (ES) for $C_{15}H_{12}Br\,NO_6$ $[M+1]^+$ 382.88

- Step 2

Compound **6a** (3.5 g, 9.16 mmol), Fe (5.0 g, 91.8 mmol) and a solution of NH_4Cl (1.9 g in 19.0 mL H_2O, 36.64 mmol) were suspended in 60 mL of EtOH and heated at 75 °C for 3 h. The mixture was allowed to cool to room temperature; the suspended solid was filtered over celite and the filtrate was concentrated in vacuo. The residue was partitioned between EtOAc and water; combined organic extracts were dried over anhydrous Na_2SO_4 and concentrated in vacuo, to give methyl 3-(4-amino-2-methoxyphenoxy)-5-bromobenzoate (**7**) (2.1 g, 66 %). The crude of **7** (1.3 g, 3.7 mmol) was suspended in concentrated HCl (6.5 mL) at 0 °C and stirred for 30 min. After this period, a solution of $NaNO_2$ in H_2O (0.38 g, 5.56 mmol) was added dropwise and stirred for 1 h at 0 ^{o}C. This solution was added dropwise to a solution of CuCl (1.47 g, 14.84 mmol) in concentrated HCl (7.0 mL) at 60 °C. After addition, the mixture was heated at 80 °C for 30 minutes. The mixture was allowed to cool to room temperature and was extracted with ethyl acetate, dried over anhydrous $MgSO_4$ and evaporated in vacuo. The crude product was purified by column chromatography (SiO_2, hexane/EtOAc 100: 0 to 80:20) to give methyl 3-bromo-5-(4-chloro-2-methoxyphenoxy)benzoate **7a** (1.0 g, 76%) . LC-MS (ES) for $C_{15}H_{12}BrCl\,O_4$ $[M+1]^+$ 372.19

- Step 3

To a solution of **7a** (1.0 g, 2.8 mmol) in dry MeOH (10. 0 mL), $NaBH_4$ (0.63 g, 16.2 mmol) was added in portions over period of 20 min at 0 ^{o}C. After addition, the reaction mixture was stirred for 30 min at the same temerature. The solution was poured into cold aqueous HCl and extracted with EtOAc (3 x 50 mL). The organic layer was sequentially washed with brine (2 x 75 mL), dried over anhydrous Na_2SO_4, and concentrated in vacuo to give 3-bromo-5-(4-chloro-2-methoxyphenoxy)phenyl)methanol **7b** (0.8 g, 88%) which was used without further purification to the next step.

- Step 4

Compound **7b** (0.88 g, 2.5 mmol) was added to a 0.16 M solution of the Des-Martin (DM) reagent (4.2 g, 10 mmol) in methylene chloride and the mixture was stirred at room temperature for 3 h.

The reaction mixture was then diluted with 20 mL of ether and poured into a solution of sodium thiosulfate. The organic layer was sequentially washed with brine (2 x 75 mL), dried over anhydrous

Na_2SO_4, and concentrated in vacuo to give **8** (0.34 g, 40 %). LC-MS (ES) for $C_{14}H_{10}BrClO_3$ $[M+1]^+$ 342.78

- Step 5

To a solution of diethylcyanomethyl phosphonate (0.23 mL, 1.4 mmol) in THF (15 mL) was added *t*-BuOK (0.16 g, 1.4 mmol) at ice-water bath temperature with stirring for 30 min. After that, **8** (0.34 g, 0.9 mmol) in THF (15 mL) was added dropwise into the above mixture at room temperature and was stirred overnight. The reaction mixture was quenched with water and extracted with EtOAc, washed with brine, dried over anhydrous Na2SO4, filtered and concentrated in vacuo. The residue was purified by column chromatography (SiO2, Hexane/CH2Cl2 80:20) to give (*E*)-3-(3-bromo-5-(4-chloro-2-methoxyphenoxy)phenyl)acrylonitrile **8a** (0.17 g, 34 %) LC-MS (ES) for $C_{16}H_{11}BrClNO_2$ $[M+1]^+$ 365.19

- Step 6

A solution of BBr_3 (3mL, 2.7 mmol, 1M in CH_2Cl_2) was added dropwise to a solution of **8a** (0.23 g, 0.76 mmol) in dry CH_2Cl_2 (4.0 mL) under N_2 at -78 °C. The reaction mixture was stirred at this temperature for 1 h. After this period, the reaction was allowed to warm to room temperature and stirred for 12 h. After completion, the solution was quenched with water, the solvent was removed in vacuo and the residue was extracted with CH_2Cl_2 and washed with a solution of $NaHCO_3$. The combined organic layer was washed with brine, dried over anhydrous Na_2SO_4 and concentrated in vacuo. The residue was purified by column chromatography (SiO_2, Hexane/EtOAc 80:20) to give (*E*)-3-(3-bromo-5-(4-chloro-2-hydroxyphenoxy)phenyl)acrylonitrile **9** (0.25 g, 100 %). LC-MS (ES) for $C_{15}H_9BrClNO_2$ $[M+1]^+$ 352.01

- Step 7

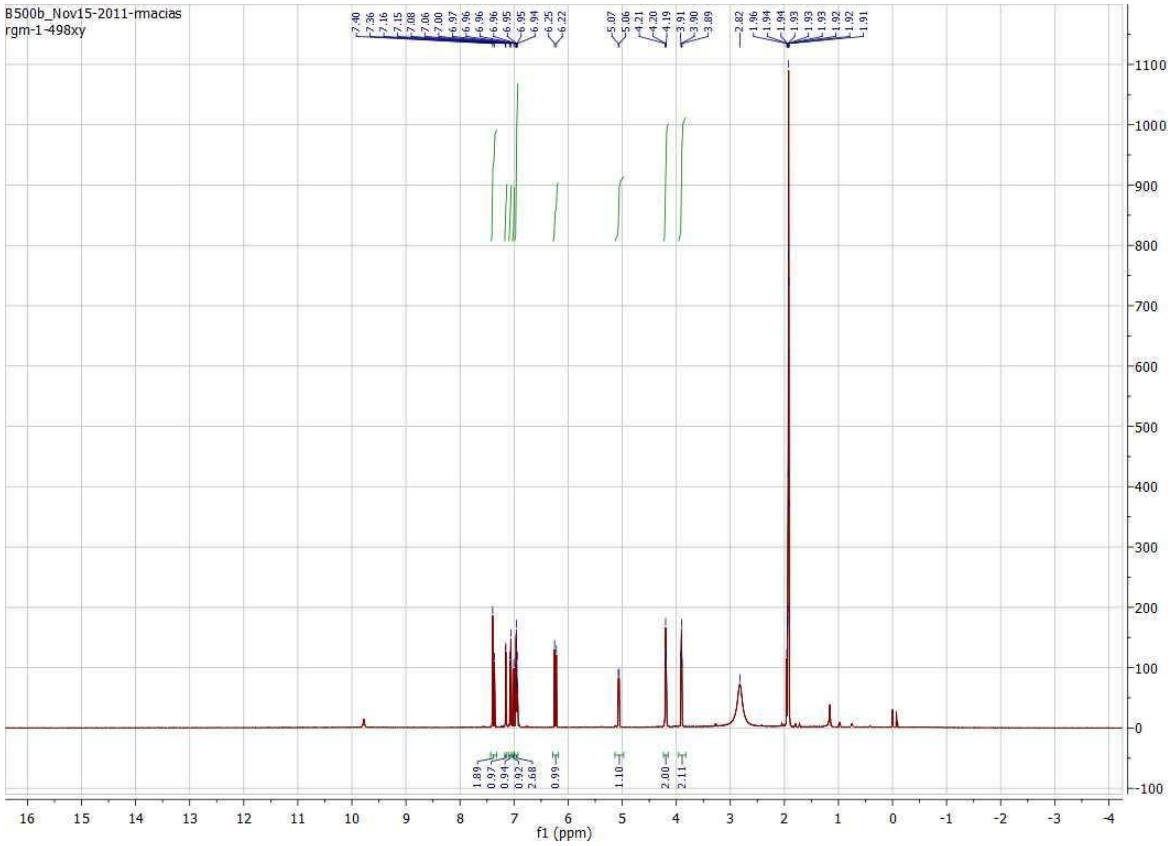
Compound **9** (0.326 g, 0.93 mmol), Cs_2CO_3 (0.071g, 0.22 mmol), 1,2 dibromoethane (0.32 mL, 3.72 mmol) in DMF was stirred at 60 ^oC for 15 h. The solution was poured into ice water and extracted with EtOAc (3 x 50 mL). The organic layer was sequentially washed with brine (2 x 75 mL), dried over anhydrous Na_2SO_4, and concentrated in vacuo. The crude product was purified by column chromatography (SiO_2, hexane/EtOAc 70:30) to give (*E*)-3-(3-bromo-5-(2-(2-bromoethoxy)-4-chlorophenoxy)phenyl)acrylonitrile **10** (0.16 g, 40 %). LC-MS (ES) for $C_{17}H_{12}Br_2ClNO_2$ $[M+1]^+$ 458.13.

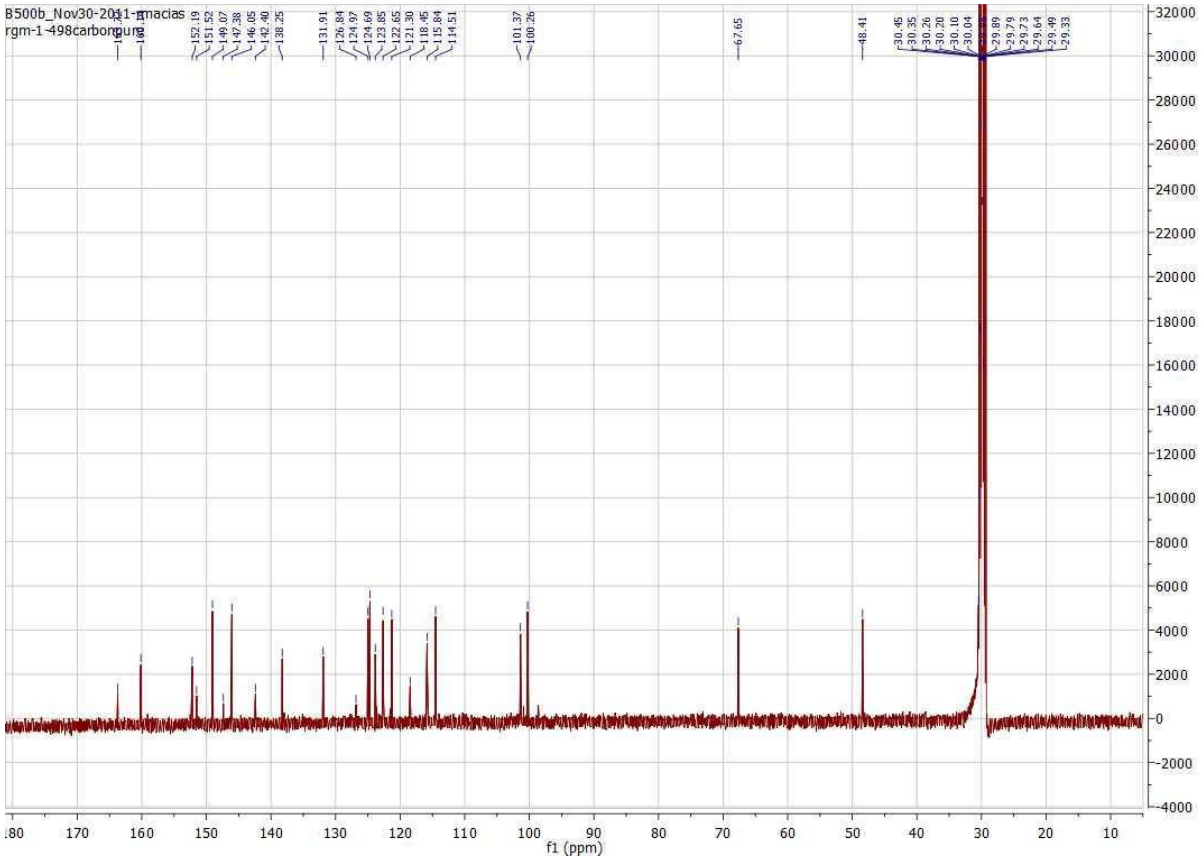
- Step 8

10 (0.15 g, 0.32 mmol), 3-benzoylpyrimidine-2,4(1*H*,3*H*)-dione (0.078 g, 0.36 mmol), anhydrous K_2CO_3 (0.05 g, 0.36 g) in DMF (4 mL) was stirred at rt overnight and 1 h at 60 ^oC to complete the reaction. The reaction mixture was poured into a solution of NH_4Cl and extracted with EtOAc (3 x 20 mL). The organic layer was sequentially washed with brine (2 x 15 mL), dried over anhydrous Na_2SO_4, and concentrated in vacuo. The crude product was dissolved in MeOH (4.0 mL) and NH_4OH (3.0 mL) was added. The reaction mixture was stirred at rt for 4 h. After completion, the reaction mixture was concentrated under reduced pressure and purified by column chromatography (SiO_2, hexane/EtOAc

20:80) to give **4** (0.07g, 43 %). HR-MS (ES) calcd for $C_{21}H_{15}NBrCl_3O_4$ $[M+1]^+$ 489.9833, found 489.9821. ^{1}H NMR: 500 MHz ($CDCl_3$) δ 4.03(t, *J*= 5 Hz, 2H), 4.32 (t, *J* = 5 Hz, 2H), 5.19 (d, J = 8 Hz, 1H), 6.36 (d, *J* = 16.5 Hz, 1H), 7.06-7.09 (m, 3H), 7.13 (s, 1H), 7.19 (d, J = 8.5 Hz, 1H), 7.28 (d, J = 2.5 Hz, 1H), 7.51 (d, *J* = 16.5 Hz, 1H), 7.52 (s, 1H). ^{13}C NMR: 125 MHz ($CDCl_3$) δ 48.34, 67.6, 100.19, 101.37, 114.51, 115.41, 118.41, 121.29, 122.64, 123.82, 124.75, 124.97, 131.90, 138.19, 142.33, 145.92, 149.21, 152.03, 159.09, 163.73.

^{1}H and ^{13}C NMR Spectral Charts for 4





Assay for Inhibition of HIV-1 Reverse Transcriptase [1]

The EnzChek® Reverse Transcriptase assay kit was purchased without enzyme from Molecular Probes (E-22064). The enzyme, HIV-1 RT, was obtain form Applied Biosystems. The assay was performed by preparing a master mixture of different concentrations of the compounds. Stock solutions were serially diluted 2-fold in 10 nM Tris-HCl at pH 7.4, 50 % DMSO in a microcentrifuge tube (5-0.001 µM). The reaction was performed by adding 20 µL of annealing mixture[2], 2 µL solution of each compound, and 0.5 µL of enzyme RT and 2.5 µL of enzyme dilution buffer (50 mM Tris-HCl, 20 % glycerol, 2 mM DTT, pH 7.6) into individual wells of the microplate, and then mixed well. After incubation at 25 ^{o}C for 30 min, 2 µL of 200 mM EDTA was added to stop the reaction, followed by addition of 173 µL of picoGreen reagent diluted 345-fold with 1X TE (tris-EDTA) buffer for each well, making the final volume 200 µL. Incubation at room temperature for 3 min, during which PicoGreen binds to double-stranded DNA and RNA-DNA hybrids, was followed by measurement of fluorescence with a fluorometer (excitation 485 nm; emission 520 nm). The assay was performed with triplicate samples at each concentration. The inhibitory effect of HIV-1RT activity was compared by using the percent inhibition, which was calculated via the following eq:

$$\% \, inhibition = /1 - \frac{RTs - RTb}{RTFc - RTFb} x100$$

where

RTs= fluorescence with sample

RTc= fluorescence without sample (normal RT reaction)

RTb= fluorescence without enzyme (negative control, background)

IC_{50} values were determined using GradPhpad Prism 5 with a non-linear regression model.

To check precision, the assay was repeated three times for nevirapine yielding IC_{50} results of 1.00, 0.89, and 1.30 µM. It was also repeated three times for rilpivirine yielding results of 40, 33, and 42 nM. Thus, the results are consistent within ±30%.

References for HIV-RT Assay

1. Silprasit, K.; Thammaporn, R.; Tecchasakul, S.; Hannongbua, S.; Choowongkomon, K. *J. Virol. Methods* **2011,** 171**,** 381.

2. 1mg/mL of poly(A) ribonucleotide template was mixed with 50 µg/mL of oligo(dT)16 primer at the same volume ratio, followed by incubation at room temperature for 1h for annealing. The mixture was diluted 200-fold with polymerizarion buffer (60 mM Tris-HCl, 60 mM KCl, 8 mM $MgCl_2$, 13 mM DTT, 100 µM, dTTp, pH 8.1).