Published in final edited form as:

J Med Chem. 2013 April 11; 56(7): 2738–2746. doi:10.1021/jm301271j.

Detecting allosteric sites of HIV-1 reverse transcriptase by X-ray crystallographic fragment screening

Joseph D. Bauman^{1,2,§}, Disha Patel^{1,3,§}, Chhaya Dhar^{1,2}, Marc W. Fromer^{1,2}, Sameer Ahmed^{1,2}, Yulia Frenkel^{1,2}, R. S. K. Vijayan^{1,2}, J. Thomas Eck^{1,2}, William C. Ho^{1,2}, Kalyan Das^{1,2}, Aaron J. Shatkin¹, and Eddy Arnold^{1,2,3,*}

¹Center for Advanced Biotechnology and Medicine, Rutgers University, Piscataway, New Jersey

²Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey

³Department of Medicinal Chemistry, Rutgers University, Piscataway, New Jersey

Abstract

HIV-1 reverse transcriptase (RT) undergoes a series of conformational changes during viral replication and is a central target for antiretroviral therapy. The intrinsic flexibility of RT can provide novel allosteric sites for inhibition. Crystals of RT that diffract X-rays to better than 2 Å resolution facilitated the probing of RT for new druggable sites using fragment screening by X-ray crystallography. A total of 775 fragments were grouped into 143 cocktails, which were soaked into crystals of RT in complex with the non-nucleoside drug rilpivirine (TMC278). Seven new sites were discovered, including the Incoming Nucleotide Binding, Knuckles, NNRTI Adjacent, and 399 sites, located in the polymerase region of RT, and the 428, RNase H Primer Grip Adjacent, and 507 sites, located in the RNase H region. Three of these sites—Knuckles, NNRTI Adjacent, and Incoming Nucleotide Binding—are inhibitory and provide opportunities for discovery of new anti-AIDS drugs.

Introduction

HIV-1 replicates rapidly and develops resistance to antiviral drugs. The most effective treatments of HIV infection consist of combinations of three to four drugs using the approach termed highly active antiretroviral therapy (HAART) that keeps viral load low in infected individuals. Thirteen of the 26 approved anti-AIDS drugs target the viral enzyme reverse transcriptase (RT), which is responsible for generating double-stranded DNA from the single-stranded viral RNA genome.¹ RT is a 117 kDa heterodimer consisting of p66 and p51 subunits. The p66 subunit contains fingers, palm, and thumb polymerase subdomains and a C-terminal RNase H domain linked by a connection subdomain to the polymerase domain (Figure 1A). The p51 subunit is a proteolytic product of p66 lacking the RNase H domain.² The thirteen RT-inhibiting drugs belong to two classes: nucleoside/nucleotide RT inhibitors (NRTIs) and non-nucleoside RT inhibitors (NNRTIs). NRTIs compete with incoming nucleotides and terminate polymerization because they lack a 3'-OH necessary for further nucleotide incorporation.³ NNRTI binding to RT allosterically traps the open binding pocket in the palm subdomain, which is not present in RT structures without NNRTI

*Corresponding Author, Phone: 732-235-5323. arnold@cabm.rutgers.edu.

§Author Contributions

These authors contributed equally.

PDB Codes

4IDK, 4ID4, 4ICL, 4I7G, 4IFV, 4IFY, 4IG0, 4IG3

present. NNRTIs inhibit RT function by locking the conformation of the polymerase domain in an unproductive mode thereby inhibiting polymerization.⁴

Resistance mutations to current drug classes frequently emerge, necessitating the identification of novel drug-binding sites for future drug development. New classes of RT inhibitors under active development are the nucleotide-competing inhibitors, p66/p51 dimerization inhibitors, and RNase H inhibitors.^{5–7} In addition to crystal structures, H/D exchange and nuclear magnetic resonance studies of RT and complexes containing RT suggest that the dynamic and flexible nature of the interdomain hinges between the various subdomains are necessary for a fully functioning RT.^{8–9} These hinge regions can provide novel inhibitory sites to overcome the emerging resistance observed for current classes of drugs. Fragment screening was utilized to chemically interrogate the hinge regions and to investigate their requirement for RT function.

Fragment screening is quickly becoming a valuable technique for identifying new target sites and fragment hit for drug discovery.¹⁰ Traditional drug discovery has often relied on screening hundreds of thousands of drug-like compounds to generate lead inhibitors against a new macromolecular target. Although high-throughput screening (HTS) has been successful in developing many drugs, it suffers from extremely low hit rates, many false positives, and laborious lead optimization as the initial hit is usually far from optimized to its binding pocket. In contrast, fragment screening allows for a far greater fraction of chemical space to be sampled with fewer compounds, ranging from 500–1,500 fragments compared to 100,000+ lead-like compounds for HTS.^{11–13} Due to their small size, fragments normally bind in the micromolar to millimolar range, but this binding is often of high quality as estimated by ligand efficiency (LE). LE is a measure of the binding contribution per non-hydrogen atom of a ligand to its target protein.^{11,14} Numerous techniques are used for fragment screening, including X-ray crystallography, NMR spectroscopy, and surface plasmon resonance.

We have successfully used X-ray crystallography-based fragment screening to discover novel RT inhibitory sites that can be pursued for further design efforts. A total of 16 small-molecule binding sites were discovered during the course of this study, of which seven will be discussed here (Supplemental Table 1). Out of the seven sites, fragment binding to three sites were found to be inhibitory in the enzymatic assay. A brief summary of these results has been reported in a review of HIV fragment screening.¹⁵

Results and Discussion

Library Design

We assembled a library of 775 compounds comprising 500 compounds purchased from Maybridge (Cornwall, UK), 175 compounds purchased based on the published recommendations of Christophe Verlinde and Wim Hol¹⁶ and an additional 100 compounds generously gifted by James Williamson (The Scripps Research Institute, La Jolla, unpublished). These fragments were grouped into 143 cocktails each containing four to eight compounds (100 mM each in d₆-DMSO; screening by NMR was also planned). Fragment cocktails were designed to have maximum structural diversity to enable deconvolution in the electron density maps, cocktail solubility, and no predicted chemical reactivity among the constituents. Additional compounds were purchased throughout the screening process based on derivatization of fragment hits and *in silico* docking (using GlideXP from Schrödinger) to newly discovered binding pockets (unpublished results).

Crystal Optimization

Crystal engineering has been successfully applied to create two enzymatically active, non-drug resistant variants of RT, RT52A and RT69A, whose crystals diffracted X-rays to better than 2 Å resolution in two conformational states, with or without NNRTI bound.¹⁷ The NNRTI-binding pocket has been shown to bind a wide range of chemically diverse hydrophobic compounds and binding of an NNRTI causes large conformational changes in RT.^{18–20} Consequently, most unliganded RT69A crystals were destroyed during the soaking of fragment cocktails. Crystals of rilpivirine (TMC278), recently FDA approved as an anti-AIDS drug,²¹ in complex with RT52A were used for fragment screening since they could be routinely reproduced and diffracted X-rays better than 2 Å resolution.^{17,22} The binding of this nanomolar inhibitor stabilized the open conformation of the NNRTI-binding pocket and minimized crystal damage during fragment soaking. These crystals were extremely robust with 93% yielding a high quality X-ray diffraction dataset after cocktail soaking.

Soaking Optimization

We discovered DMSO to be an excellent cryoprotectant for RT-rilpivirine crystals. Crystals were soaked in 20% (v/v) DMSO (containing fragment cocktail) and 5% (v/v) ethylene glycol as an additional cryoprotectant prior to data collection. The high DMSO tolerance of the crystals allowed soaking of fragments at a final concentration of 20 mM each. Our soaking experiments revealed no noticeable difference in the strength of binding (as assessed by the quality of electron density) of a fragment with soaking times ranging from 30 seconds to 10 minutes. Based on these results, soaking times were usually one to two minutes.

Two additives were found to be beneficial for crystal stability and fragment solubility. L-arginine was found to be a potent solubilizer for fragments. Solubilization is most likely due to the guanidinium group of arginine, which can act as a hydrogen-bonding donor or acceptor while π -stacking with aromatic fragments.²³ Our initial experiments showed that several hydrophobic fragments could bind RT only when soaked in the presence of 80 mM L-arginine (Supplemental Figure 1). All fragments were subsequently soaked into the crystals with 80 mM L-arginine present. During the course of screening, trimethylamine N-oxide (TMAO), a potent protein stabilizer,²⁴ was found to improve the robustness of the crystals for fragment soaking, X-ray diffraction resolution, and crystal mosaicity. Significantly, 6% (w/v) TMAO improved X-ray diffraction resolution of RT-rilpivirine crystals from 1.80 Å to 1.51 Å resolution.²⁵ TMAO may be of general utility for enhancing the diffraction quality of many protein crystals.

Data Collection and Processing

To avoid misinterpretation of solvent binding as fragment binding, a “DMSO blank” reference structure was determined to 1.8 Å resolution, which revealed a total of 16 DMSO binding sites (Figure 1A). This structure was obtained by soaking a crystal in 20% (v/v) DMSO solution without any fragments present. The “DMSO blank” coordinates and scattering factors were then used to create $F_{\text{obs}}(\text{fragment soaked}) - F_{\text{obs}}(\text{DMSO blank})$ ($F_o - F_o$) difference maps to identify fragment binding (Figure 1B). In order to maximize data collection and minimize radiation damage, we initially performed a quick initial data collection pass on each crystal. Limiting exposure time and increasing oscillation range reduced the duration of dataset collection to less than five minutes while maintaining diffraction resolution between 2.0 to 2.3 Å (Figure 1B). After data collection was complete, the dataset was immediately processed and an $F_o - F_o$ difference map was generated. Visual inspection of the $F_o - F_o$ difference map was conducted to determine if changes in the electron density were due to either fragment binding and/or the repositioning of solvent and amino acids. Crystals for which fragment binding was detected were subjected to a high-

resolution data collection pass with typical resolution of 1.8 to 2.1 Å. The high-resolution dataset was then merged with the quick-pass dataset and further refinement was carried out to improve the density for the bound fragment. All fragment hits from cocktails were subsequently verified by repeating the experiments with individual fragment soaks.

***In vitro* Dual Activity Monitor (JDAM)**

A dual activity monitor assay, JDAM, was developed to determine whether fragment binding inhibited either the polymerase or RNase H activities of RT (Figure 1C). A 142 nucleotide T7 polymerase runoff transcript was purified and annealed to a DNA primer. In the presence of nucleoside triphosphates, RT extends the DNA primer, complementing the RNA template using the RNA-dependent DNA polymerase activity, and then degrades the RNA using its RNase H activity. A DNA “molecular beacon” type probe, *i.e.*, a single-stranded oligonucleotide with a stem-loop structure, was used to detect activity. The probe consists of a 25 nucleotide loop that is complementary to the newly synthesized DNA strand and a five base-pair stem that holds a fluorophore and quencher proximal to each other. Upon synthesis of the DNA strand and hydrolysis of the template RNA, the stem-loop structure of the probe collapses when it binds to the complementary DNA. This subsequently causes the fluorophore and quencher to separate and fluorescence to occur. A decrease in the amount of fluorescence in the presence of a fragment relative to a non-inhibited reaction is interpreted as inhibition. The *Z'* score for JDAM was determined to be an acceptable 0.74.²⁶ All fragments and derivatives found bound to RT were tested for inhibition using JDAM. Each inhibitory fragment was then co-crystallized with RT in the absence of TMC278 to confirm the hit identity and to determine whether or not the fragment was an NNRTI.

Fragment Screening Overview

Over the course of this study, 606 datasets were collected with an average resolution of 2.1 Å. A total of 742 fragments were successfully screened with 34 fragments found to bind throughout RT at 16 sites, giving a hit rate of 4.4% (Figure 1D). Out of 606 datasets, 43% of the datasets collected were from cocktail soaks, 34% for individual compounds, 14% fragment derivatives, and 9% were co-crystals with individual fragments. The high percentage of individual soaks stemmed from difficulty with identifying weakly bound fragments.

Halogenated fragments gave substantially higher hit rates than the non-halogenated fragments as a whole. A hit rate of 24.1% was observed for fluorinated fragments with seven of the 29 fluorine-containing fragments identified as hits. Similarly, a hit rate of 23.5% was observed with four out of the 17 brominated fragments screened binding to RT. One of the brominated fragments, 4-bromopyrazole, was observed at 16 different sites throughout RT. Our fragment library also contained two chlorinated compounds with one of them found to bind. Biasing the fragment library with halogenated compounds may be beneficial based on these observations.

Brominated and fluorinated compounds can also be advantageous due to their use for anomalous X-ray diffraction²⁷ and isotope enrichment for NMR spectroscopy²⁸, respectively. The use of halogens can also allow for additional interactions in the form of halogen binding. Although these interactions are not normally as strong as the typical hydrogen bonding or hydrophobic interactions, halogens can contribute to the overall binding interactions and make excellent handles for further expansion.²⁹

NNRTI Adjacent Binding Site

The “NNRTI Adjacent” binding site (Figure 2A) is located at the p66/p51 interface adjacent to Glu138 of p51, and separated from the NNRTI-binding pocket by the β 9 strand. The binding site is a tunnel-like pocket with Thr139 (p51), Pro140 (p51), Thr165, Leu168, Lys172, and Ile180 forming an opening of the tunnel, which then continues through the protein towards another cavity near the polymerase active site at Gln91. The fragment binding to the site with the best electron density, **1** (ethyl pyrazolo[1,5-a]pyridine-3-carboxylate), has hydrogen-bonding interactions with the backbone carbonyl of Ile180 and the backbone amide nitrogen of Gln182. Hydrophobic interactions were also observed with Pro140 and Ile180. The side chains of Gln161 and Gln182 are repositioned to accommodate the fragment. This, in turn, allows for hydrogen-bonding interactions between Gln161 to the backbone amide of Gln91 and Gln91 to the backbone carbonyl of Gln182 (Figure 2B). This new hydrogen-bond network causes a 2.1 Å C α backbone movement of Gln91 that propagates to surrounding residues. Significantly, a 1.1 Å C α backbone movement at Gln182 amplifies into a 1.5 Å C α movement of the catalytic YMDD motif. Of the fragments bound to this site, the best inhibition and electron density was observed for **1**, which had an IC₅₀ value of 350 μ M with a ligand efficiency of 0.34 kcal/mol●NHA (non-hydrogen atom). The formation of the new hydrogen-bond network in combination with the distortion of the YMDD motif is apparently responsible for the inhibition by **1**.

Intriguingly, an ordered water molecule is located approximately 4.1 Å away from **1** and 3.7 Å from the NNRTI rilpivirine. Thus, it may be possible to extend an NNRTI to bridge to the NNRTI Adjacent pocket. Alternatively, fragment evolution along the NNRTI Adjacent tunnel may be a promising avenue for optimization against the wildtype and current drug-resistant forms of RT.

Avoidance of drug resistance is a crucial goal for lead development against RT. To ascertain the conservation of the binding pocket residues, 1,809 sequences of RT from clinical isolates³⁰ were analyzed. Highly conserved residues were defined as having < 1% mutations compared with the consensus at a given position. Mutational analysis of the pocket for **1** reveals a great degree of conservation for the residues not involved in resistance of NNRTIs with the exception of Lys172, which is more commonly Arg. The residues involved in key interactions (Pro140, Ile180, and Gln182) are highly conserved, thereby suggesting this to be an attractive area for lead development. However, development will require consideration of NNRTI-resistance mutations, specifically Glu138Lys, Thr139Ala, Thr165Ile, and Tyr181Cys.

Knuckles Site

One promising pocket, the “Knuckles,” is formed at the junction between the p66 fingers and palm subdomains (Figure 1D). Prior to fragment binding, the Knuckles pocket is a buried cavity near the polymerase active site. Small molecule binding stabilizes an open conformation of the Knuckles where Ser117 is repositioned 2.8 Å away from its position in the unliganded form (Figure 2C–D). Critical active site residues Tyr115 and Phe116, which are involved in binding the incoming dNTP during polymerization³¹, have a polypeptide backbone shift of 3.2 Å. The fingers subdomain is shifted 1.2 Å, even within the constraints of the crystal lattice. Several fragments were found to bind to this allosteric site. Among these, **2** (4-(trifluoromethoxy)benzylamine) had the clearest density but no detectable inhibition at 1 mM (Figure 2C). The trifluoro group of **2** makes halogen-bonding interactions with Met164 and Ser117. Additionally, hydrophobic interactions are seen between the aromatic ring of **2** and residues Ile5, Ala114, Leu214, and Val118 of p66. Of the fragments found to bind in the Knuckles pocket, **3** (4-(trifluoromethoxy)phenol) was found to have an IC₅₀ of 600 μ M corresponding to a ligand efficiency of 0.37 kcal/

mol●NHA. Residues comprising the Knuckles pocket are highly conserved with the exception of Leu214, which is predominantly phenylalanine. Upon fragment binding, the Knuckles pocket expands and becomes solvent exposed at Ser117 and Ile167. This provides room for fragment expansion in two directions. Potential expansion at Ile167 is further supported by the discovery of the compound, 4 ((S)-6,6-dimethyl-5-((R)-8-oxo-6,8-dihydro-[1,3]dioxolo[4,5-e]isobenzofuran-6-yl)-5,6,7,8-tetrahydro-[1,3]dioxolo[4,5-g]isoquinolin-6-ium), found bound approximately 5 Å from 2 (Figure 2E). Lys166 repositions to allow binding of 4 to occur. Key interactions observed include hydrogen bonding with Ser163 and hydrophobic interactions with Lys166, Gly51, and Glu6.

Incoming Nucleotide Binding Site

Fragment binding was observed to a pocket flanked by Tyr115, Phe116, and Gln151 nearby the highly conserved polymerase active site. Upon binding, Gln151 reorients allowing the pocket to expand (Figure 2F). Key interactions with 5 (4-(4-methylpiperazino)benzoic acid) include electrostatic interactions between the piperazine ring of 5 and the carboxylate of Asp185, an essential catalytic residue for polymerization. Hydrogen-bonding interactions are seen between the carboxylic acid moiety of 5 and the amine of Lys73, hydroxyl of Tyr146, and backbone amide of Gln151. Edge-to-face π -interactions are observed for the aromatic ring of 5 with Tyr115 and Phe116. The site is attractive for drug design efforts because a compound bound here would directly interfere with dNTP binding. All residues surrounding the fragment are highly conserved making resistance evolution at this site difficult. Compound 5 has an IC₅₀ value of 200 μ M and LE of 0.31 kcal/mol●NHA, making it an excellent candidate for development as an inhibitor that will block access to the polymerase active site and nucleotide-binding site.

399 Site

The p66/p51 dimer is essential for RT activities and disruption of dimerization inhibits DNA polymerization and viral replication. Peptides and small molecules have been reported to inhibit RT as dimerization inhibitors; however, no structure is available of RT complexed with any of these molecules.^{32,33} We observed fragment binding in a pocket located at the p66/p51 interface. The pocket is comprised of residues Trp24, Pro25, Glu399, Thr400, Trp402, and Thr403 from p51 as well as Thr377 and Ile380 from the p66 subunit (Figure 2G). A fragment, 6 (1-(4-fluorophenyl)-5-methyl-1H-pyrazol-4-yl), was observed at this site making hydrogen-bonding interactions with the backbone carbonyl of Glu399 as well as hydrophobic interactions with Trp24 and Pro25. No inhibition was seen from fragments binding at this site as a possible consequence of the small size of the compounds and the apparent rigidity of the site in numerous RT structures. Further exploration using larger compounds may validate whether or not binding to this site is inhibitory.

428 Site

Fragment binding was also observed at the 428 site in the connection subdomain. Residues Glu396, Thr397, Thr400, Leu425, and Gln428 from the p66 subunit form the binding pocket of the 428 site (Figure 3A). Unambiguous electron density was observed for fragment 7 binding to this site. To accommodate 7 (2-amino-N-(2-oxo-2,3-dihydro-1H-benzo[d]imidazol-5-yl)acetamide) binding, Gln428 rotates by 70 degrees. Hydrogen-bonding interactions were observed between the hydroxyl group of 7 and Thr400 as well as the amine of 7 and the carboxylate side chain of Glu430, respectively. Hydrophobic interactions of 7 with Gln428 and Leu425 are also observed. Inhibition by 7 was determined to be 30% at 1 mM. Although 7 is a weak inhibitor, further development is supported by the presence of a DMSO molecule approximately 2.6 Å away from the bound fragment that can provide a handle for potential fragment growth.

507 Site

Located approximately 9 Å away from the 428 site, the 507 site is a large solvent-exposed pocket at the base of the RNase H domain. Strong binding for 8 (2-(((2-(3,4-dihydroquinolin-1(2H)-yl)-2-oxoethyl)(methyl)amino)methyl)quinazolin-4(1H)-one), a compound purchased based on *in silico* docking results to the Knuckles pocket, was unexpectedly observed at the 507 site. Hydrogen bonding was observed between 8 and the hydroxyl of Tyr532 as well as the backbone carbonyl of Leu533 (Figure 3B). Hydrophobic interactions with Lys259 from p51 and Leu429, Leu533, Ala534, and Trp535 from the p66 subunit were also observed (Figure 4B). Interestingly, an independent analysis of the RT/rilpivirine structure using Schrödinger SiteMap revealed both 428 and 507 sites as potential ligand-binding sites.³⁴ The IC₅₀ was found to be 150 μM. However, the ligand efficiency for the molecule is only 0.19 kcal/mol●NHA due to its modest inhibition and relatively large size (362 Da).

Co-crystallization of 8 in complex with RT without rilpivirine present revealed that 8 binds to the NNRTI-binding pocket (Figure 3C). Key interactions within the NNRTI pocket includes hydrogen bonding with the backbone amide of Lys101 and hydrophobic interactions with Lys103, Val106, Val179, Ile180, Tyr181, Tyr188, Phe227, Trp229, Leu234, and Tyr318. Based on the quality of electron density, 8 is believed to be inhibiting RT predominantly as an NNRTI. Exploration of the inhibitory potential of the 507 site requires further investigation by screening for compounds that would exclusively bind the 507 site and not the NNRTI-binding pocket.

RNase H Primer Grip Adjacent

The RNase H Primer Grip Adjacent site consists of Leu469, Lys476, Gln480, Tyr483, Glu516, Leu517, and Gln520 (Figure 3D). Pocket formation occurs when the side chains of Gln480 and Gln520 rotate to allow fragment binding. A small network of hydrogen-bonding interactions is formed between the carboxylate moiety of 9 (1-methyl-5-phenyl-1H-pyrazole-4-carboxylic acid) and Gln480 and Lys476, which is part of the RNA primer grip³⁵ with an additional hydrogen bond between the pyrazole of 9 and Gln520. Additional hydrophobic interactions with Lys476, Gln480, Tyr483, Leu517, and Gln520 were also observed. No inhibition was detected for fragments binding to this site. Fragment expansion could be used to extend this fragment to the RNA primer grip or active site.

Similarity Analysis

A Tanimoto similarity analysis was conducted to evaluate whether or not the fragment hits identified during the screening were potentially new chemical scaffolds compared to published RT inhibitors.³⁶ Fragment hits were compared to a database of 2,961 published RT inhibitors extracted from the BindingDB (<http://www.bindingdb.org>).³⁷ Of the 2,961 RT inhibitors, no compound has a Tanimoto similarity score greater than 60% when compared with the nine fragment hits (see Table 2 in Supplemental). Given that most published RT inhibitors are considerably larger than fragments, higher Tanimoto coefficients are not expected. A simple substructure search of the fragment hits within the 2,961 RT inhibitors in BindingDB revealed no compounds containing the chemical scaffolds identified in this study.

Conclusion

Prior to this study, Geitman et al. identified one novel scaffold that targeted the NNRTI pocket from a fragment screening campaign by surface plasmon resonance (SPR) against both wild-type and drug-resistant variants of RT using a 1,040 compound library.¹⁸ Fragment screening by SPR has also been used as a tool for understanding pocket flexibility,

specifically the flexibility of NNRTI pocket. A deconstruction analysis, which consisted of screening a fragment library based on known NNRTIs, revealed that the compound size was an important factor in binding to the NNRTI pocket.³⁸

Unlike previous studies that focused primarily on the NNRTI site of RT, we took advantage of the inherent flexibility of RT observed in crystal structures to chemically interrogate potential novel druggable binding sites that can be used for new drug design strategies. Fragment screening by X-ray crystallography revealed seven new binding sites within HIV-1 RT, of which three sites (the Knuckles, the NNRTI Adjacent, and the Incoming Nucleotide Binding sites) were found to be inhibitory in an enzymatic assay. Tanimoto and substructure analyses revealed that fragments binding to these sites were novel scaffolds when compared with previously published RT inhibitors. This combined with the structural information gained from this approach can allow for rapid structure-based drug design of novel lead candidates.

Experimental Section

Expression, Purification, and Crystallization

RT constructs RT52A (crystallization optimized mutant) and RT35A (wild-type) were constructed, expressed, and purified as described previously.¹⁷ Crystallization was performed using the hanging-drop method with EasyXtal DG-Tools (Qiagen, Valencia, CA) crystallization trays. Prior to crystallization, RT52A (20 mg/ml) was incubated with rilpivirine (TMC278/Edurant) at 1:1.5 molar protein to drug ratio at room temperature (~23°C) for 30 minutes. RT52A-rilpivirine crystals were produced in hanging drops at 4°C with a 1:1 ratio of protein and well solution containing 11% PEG 8000, 4% PEG 400, 50 mM imidazole pH 6.6, 10 mM spermine, 15 mM MgSO₄, 100 mM ammonium sulfate, and 5 mM tris(2-carboxyethyl)phosphine and an experimentally optimized concentration of microseeds from previously generated crystals (preseeding).

Fragment Library

A library of 775 compounds purchased from Maybridge, Sigma-Aldrich, or Acros was assembled for screening. The fragments were purchased with nominal purity of > 99% and used without further purification. A new lot of highly pure fragment (> 99% purity) from Sigma-Aldrich, Acros, or Tocris was purchased to validate fragment hits from the Williamson library.

Fragment Soaking, Data Collection, and Processing

The fragment/cryo soaking solutions were prepared with crystallization well solution with the addition of 80 mM L-arginine, 5% (v/v) ethylene glycol, and 20% (v/v) d₆-DMSO (containing 20 mM final concentration each fragment). Crystals of RT52A-rilpivirine were harvested two weeks to four months after crystals formed. The crystals were placed in fragment/cryo soaking drops for one to two minutes before flash cooling in liquid nitrogen. Data collection was performed at the Cornell High Energy Synchrotron Source (CHESS) F1 beamline and the National Synchrotron Light Source (NSLS) X25 and X29 beamlines. The diffraction data were indexed, processed, scaled and merged using *HKL2000I*.³⁹ Initial datasets from crystals were collected to minimize the time of collection by increasing the oscillation range per image and decreasing exposure time. F_o-F_o maps (as described previously) were immediately calculated using CNS and visualized with Coot.^{41,42} Datasets for crystals containing bound fragments were then recollected to improve maximum X-ray diffraction resolution. High-resolution datasets containing bound fragments were further refined using PHENIX and Coot.^{40,41} Crystal structure figures were made with MacPyMol (Schrödinger, New York, NY).

RT Activity Assay

The 142-nucleotide RNA strand with the following sequence was produced with the MEGAshortscript kit and purified with the MEGAclear kit (Ambion, Austin, TX): GTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAA GTAGATAAATTAGTCAGTGTCTGGATAAAGTCTGAGTCTGGTAAAGAAACCGCTG CTGCGAAATTTGAACGCCAGCACATGGACTCGTCTACTA. The DNA primer has the sequence TAGTAGACGAGTCCATGTGC. The probe was (FAM)CCGGGGAATTGGAGGAAATGAACAAAGTAGCCGG(Iowa-black). All oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). RT (construct RT35A – wild type) was preincubated with inhibitor for 30 minutes at 37°C. The preincubation mixture contained 100 nM RT35A, 10 μM dNTPs, 1 mg/ml BSA, 1 mM DTT, 10 mM CHAPS, 1% (v/v) d₆-DMSO (with inhibitor), and 50 nM probe. Reactions were started with the addition of premixed primer (225 nM) and RNA (200 nM). Reactions were incubated at 37°C and then polymerization was stopped with the addition of 500 nM RNase A (Qiagen, Valencia, CA) after 2 minutes. Fluorescence was immediately measured in a Varioskan microplate reader (Thermo Scientific, Rockford, IL). Dose-response curves were calculated by fitting data to a sigmoidal four-parameter logistics equation using Prism (GraphPad). Ligand efficiency was calculated using $LE \text{ (in kcal/mol)} = -RT \ln(IC_{50})/NHA$.

Supporting Information

Figure 1 demonstrates the effect of L-arginine on fragment binding. Table 1 summarizes the X-ray data and refinement statistics for each structure discussed. Table 2 describes the Tanimoto coefficient analysis. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Jamie Williamson from The Scripps Research Institute for contributing to our fragment library. We also gratefully acknowledge Tom Blundell and members of his laboratory, Wim Hol, Christophe Verlinde, and Jim Wells for valuable discussions about fragment screening. We thank the laboratories of Ann Stock and Gaetano Montelione for access to equipment used in this study. We greatly appreciate Matthew Miller for his assistance in data collection and Stephen Hughes for valuable discussions as well as Rajiv Bandwar and Arthur D. Clark, Jr. for their assistance in this study. Data collection was largely conducted at the Cornell High Energy Synchrotron Source (CHESS) and the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory. CHESS is supported by the NSF & NIH/NIGMS via NSF award DMR-0225180, and the MacCHESS resource is supported by NIH/NCRR award RR-01646. Use of NSLS is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-98CH10886. We are extremely grateful to NIH MERIT Award AI27690 (to EA) for generous support of these efforts.

Abbreviations Used

RT	HIV-1 reverse transcriptase
HAART	highly active antiretroviral therapy
HTS	high throughput screening
NNRTI	non-nucleoside reverse transcriptase inhibitor
LE	ligand efficiency

H/D exchange	hydrogen/deuterium exchange
TMAO	trimethylamine N-oxide
JDAM	<i>in vitro</i> dual activity monitor
NHA	non-hydrogen atom
dNTP	deoxyribonucleoside triphosphate

References

- Coffey, S. Antiretroviral Drug Profiles, HIVInSite. 2012. University of California San Francisco [Online] Available: <http://hivinsite.ucsf.edu/InSite?page=ar-drugs>[accessed July 2012]
- Sarafianos SG, Marchand B, Das K, Himmel DM, Parniak MA, Hughes SH, Arnold E. Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *J Mol Biol.* 2009; 385:93–713.
- De Clercq E. Non-nucleoside reverse transcriptase inhibitors (NNRTIs): past, present, and future. *Chem Biodivers.* 2004; 1:4–64.
- Das K, Martinez SE, Bauman JD, Arnold E. HIV-1 reverse transcriptase complex with DNA and nevirapine reveals non-nucleoside inhibition mechanism. *Nat Struct Mol Biol.* 2012; 19:53–259.
- Jochmans D, Deval J, Kesteleyn B, Van Marck H, Bettens E, De Baere I, Dehertogh P, Ivens T, Van Ginderen M, Van Schoubroeck B, Ehteshami M, Wigerinck P, Götte M, Hertogs K. Indolopyridones inhibit human immunodeficiency virus reverse transcriptase with a novel mechanism of action. *J Virol.* 2006; 80:2283–12292.
- Divita G, Restle T, Goody RS, Chermann JC, Baillon JG. Inhibition of human immunodeficiency virus type 1 reverse transcriptase dimerization using synthetic peptides derived from the connection domain. *J Biol Chem.* 1994; 269:3080–13083.
- Tramontano E, Di Santo R. HIV-1 RT-associated RNase H function inhibitors: recent advances in drug development. *Curr Med Chem.* 2010; 17:837–2853.
- Seckler JM, Howard KJ, Barkley MD, Wintrode PL. Solution structural dynamics of HIV-1 reverse transcriptase heterodimer. *Biochemistry.* 2009; 48:646–7655.
- Zheng X, Mueller GA, DeRose EF, London RE. Solution characterization of [methyl-¹³C]methionine HIV-1 reverse transcriptase by NMR spectroscopy. *Antiviral Res.* 2009; 84:205–214. [PubMed: 19665484]
- Kuo, LC., editor. Fragment-Based Drug Design Tools, Practical Approaches, and Examples; *Methods in Enzymology* Vol 493. Elsevier, Inc; San Diego, CA: 2011.
- Kuntz ID, Chen K, Sharp KA, Kollman PA. The maximal affinity of ligands. *Proc Natl Acad Sci USA.* 1999; 96:997–10002.
- Rees DC, Congreve M, Murray CW, Carr R. Fragment-based lead discovery. 2004; 3:60–672.
- Hesterkamp T, Whittaker M. Fragment-based activity space: smaller is better. *Curr Opin Chem Biol.* 2008; 12:60–268. [PubMed: 18282482]
- Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. *Drug Disc Today.* 2004; 9:30–431.
- Bauman JD, Patel D, Arnold E. Fragment screening and HIV therapeutics. *Top Curr Chem.* 2012; 317:81–200.
- Verlinde CLMJ, Fan E, Shibata S, Zhang Z, Sun Z, Deng W, Ross J, Kim J, Xiao L, Arakaki T, Bosch J, Caruthers JM, Larson ET, LeTrong I, Napuli A, Kelley A, Mueller N, Zucker F, Van Voorhis WC, Buckner FS, Merritt EA, Hol WGJ. Fragment-based cocktail crystallography by the medical structural genomics of pathogenic protozoa consortium. *Curr Top Med Chem.* 2009; 9:678–1687. [PubMed: 19689374]
- Bauman JD, Das K, Ho WC, Baweja M, Himmel DM, Clark AD, Oren DA, Boyer PL, Hughes SH, Shatkin AJ, Arnold E. Crystal engineering of HIV-1 reverse transcriptase for structure-based drug design. *Nucl Acids Res.* 2008; 36:083–5092.

18. Geitmann M, Elinder M, Seeger C, Brandt P, de Esch IJP, Danielson UH. Identification of a novel scaffold for allosteric inhibition of wild type and drug resistant HIV-1 reverse transcriptase by fragment library screening. *J Med Chem.* 2011; 54:99–708.
19. Rodgers DW, Gamblin SJ, Harris BA, Ray S, Culp JS, Hellmig B, Woolf DJ, Debouck C, Harrison SC. The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. *Proc Natl Acad Sci USA.* 1995; 92:222–1226.
20. Hsiou Y, Ding J, Das K, Clark AD, Hughes SH, Arnold E. Structure of unliganded HIV-1 reverse transcriptase at 2.7 Å resolution: implications of conformational changes for polymerization and inhibition mechanisms. *Structure.* 1996; 4:53–860.
21. Janssen PAJ, Lewi PJ, Arnold E, Daeyaert F, de Jonge M, Heeres J, Koymans L, Vinkers M, Guillemont J, Pasquier E, Kukla M, Ludovici D, Andries K, de Béthune MP, Pauwels R, Das K, Clark AD, Frenkel YV, Hughes SH, Medaer B, De Knaep F, Bohets H, De Clerck F, Lampo A, Williams P, Stoffels P. In search of a novel anti-HIV drug: multidisciplinary coordination in the discovery of a 4-[[4-[(1E)-2-cyanoethenyl]-2,6-dimethylphenyl]amino]-2-pyrimidinyl]amino]benzonitriles (R278474, Rilpivirine). *J Med Chem.* 2005; 48:901–1909. [PubMed: 15715458]
22. Das K, Bauman J, Clark A, Frenkel Y, Lewi P, Shatkin A, Hughes S, Arnold E. High-resolution structures of HIV-1 reverse transcriptase/TMC278 complexes: strategic flexibility explains potency against resistance mutations. *Proc Natl Acad Sci USA.* 2008; 105:466.
23. Flocco MM, Mowbray SL. Planar stacking interactions of arginine and aromatic-side chains in proteins. *J Mol Biol.* 1994; 235:9–717.
24. Zou Q, Bennion BJ, Daggett V, Murphy KP. The molecular mechanism of stabilization of proteins by TMAO and its ability to counteract the effects of urea. *J Am Chem Soc.* 2002; 124:192–1202. [PubMed: 11782169]
25. Zhang J, Chung T, Oldenburg K. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J Biomol Screen.* 1999; 4:8–73.
26. Kuroda DG, Bauman JD, Challa JR, Patel D, Troxler T, Das K, Arnold E, Hochstrasser RM. Snapshot of the equilibrium dynamics of a drug bound to HIV-1 reverse transcriptase. *Nature Chem.* in press.
27. Blaney, J.; Nienaber, V.; Burley, SK. Fragment-based lead discovery and optimisation using X-ray crystallography, computational chemistry, and high-throughput organic synthesis. In: Jahnke, W.; Erlanson, DA., editors. *Fragment-based Approaches in Drug Discovery Methods and Principles in Medicinal Chemistry.* Vol. 34. Wiley–VCH; Weinheim, Germany: 2006. p. 15-248.
28. Dalvit C, Flocco M, Veronesi M, Stockman BJ. Fluorine-NMR competition binding experiments for high-throughput screening of large compound mixtures. *Comb Chem High Throughput Screen.* 2002; 5:05–611.
29. Hernandez MZ, Cavalcanti SMT, Moreira DRM, de Azevedo Junior WF, Leite ACL. Halogen atoms in the modern medicinal chemistry: hints for drug design. *Curr Drug Targets.* 2010; 11:03–314.
30. HIV sequence database. 2009. Los Alamos National Laboratory [Online]. Available: www.hiv.lanl.gov (accessed April 2011)
31. Martín-Hernández AM, Domingo E, Menéndez-Arias L. Human immunodeficiency virus type 1 reverse transcriptase: role of Tyr115 in deoxynucleotide binding and misinsertion fidelity of DNA synthesis. *EMBO J.* 1996; 15:434–4442.
32. Andreola ML. Therapeutic potential of peptide motifs against HIV-1 reverse transcriptase and integrase. *Curr Pharm Design.* 2009; 15:508–2519.
33. Grohmann D, Corradi V, Elbasyouny M, Baude A, Horenkamp F, Laufer SD, Manetti F, Botta M, Restle T. Small molecule inhibitors targeting HIV-1 reverse transcriptase dimerization. *Chembiochem.* 2008; 9:16–922.
34. Felts AK, Labarge K, Bauman JD, Patel DV, Himmel DM, Arnold E, Parniak MA, Levy RM. Identification of alternative binding sites for inhibitors of HIV-1 ribonuclease H through comparative analysis of virtual enrichment studies. *J Chem Inf Model.* 2011; 51:986–1998. [PubMed: 21495663]

35. Sarafianos SG, Das K, Tantillo C, Clark AD, Ding J, Whitcomb JM, Boyer PL, Hughes SH, Arnold E. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J*. 2001; 20:449–1461.
36. Harper G, Bradshaw J, Gittins JC, Green DV, Leach AR. Prediction of biological activity for high-throughput screening using binary kernel discrimination. *J Chem Inf Comput Sci*. 2001; 1:295–1300.
37. Liu T, Lin Y, Wen X, Jorissen RN, Gilson MK. BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucl Acids Res*. 2007; 35:D198–201. [PubMed: 17145705]
38. Brandt P, Geitmann M, Danielson UH. Deconstruction of non-nucleoside reverse transcriptase inhibitors of human immunodeficiency virus type 1 for exploration of the optimization landscape of fragments. *J Med Chem*. 2011; 54:709–718. [PubMed: 21207958]
39. Otwinowski, Z.; Minor, W.; Borek, D.; Cymborowski, M.; DENZO and SCALEPACK. *International Tables for Crystallography Volume F Crystallography of Biological Macromolecules*. Second Edition. Arnold, E.; Himmel, DM.; Rossmann, MG., editors. John Wiley & Sons; West Sussex: 2011. p. 282-295.
40. Brunger AT, Adams PD, Clore GM, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges N, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system (CNS): a new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*. 1998; 4:905–921. [PubMed: 9757107]
41. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 2010; 66:486–501. [PubMed: 20383002]
42. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. 2010; 66:213–221. [PubMed: 20124702]

Figure 1.

Overview of fragment screening. (a) Color-coded cartoon of the RT-rilpivirine complex with 20% d_6 -DMSO present. Rilpivirine (brown space filling) is bound at the NNRTI-binding pocket. The p66 subdomains are color-coded fingers (blue), palm (red), thumb (green), connection (yellow), RNase H (red), and the p51 subunit (cyan). Ordered waters are shown as blue dots and d_6 -DMSO molecules are shown as green, yellow, and red spheres. (b) Flowchart for fragment screening by X-ray crystallography with approximate completion times indicated. Hit to lead development is the current focus of this project. (c) Schematic drawing depicting the JDAM activity assay. (d) Drawing representing p66 (dark green), p51 (cyan), and rilpivirine (yellow spheres) with all fragments discovered in this study (orange spheres). The sites of fragment binding described in this report are circled.

Figure 2.

The NNRTI Adjacent, Knuckles, Incoming Nucleotide Binding, and 399 sites. Residues that shift due to ligand binding are shown in magenta in their non-bound positions and dark green in their bound position except p51 residues shown in cyan. (a) **4** (gray) bound at the NNRTI Adjacent site (dark green). Rilpivirine is shown on the right side of the panel (yellow). Electron density is shown for **1** from a $F_o - F_c$ omit map (3.5 σ contours). (b) Top-down view of panel (A) showing the new hydrogen-bonding network as green dashes. (c) Structure of RT-rilpivirine-**1**. (d) Structure turned 90 degrees. (e) Structure of RT-rilpivirine-**3** with the position of **2** shown at the bottom of the panel. Electron density is

shown for **2** from an F_o-F_c omit map (2σ contours). (**f**) Incoming nucleotide binding site with **5** bound. Q151 shifts from the unbound position (magenta) to the bound position (green). Electron density is shown for **5** from an F_o-F_c omit map (3.0σ contours). (**g**) **6** bound at the 399 site. An ordered d_6 -DMSO molecule is shown at the bottom of the panel. Electron density is shown for **6** from an F_o-F_c omit map (1.5σ contours).

Figure 3.

428, 507, and RNase H Primer Grip Adjacent sites. Residues that shift due to ligand binding are shown in magenta in their non-bound positions and dark green in their bound position. (a) **7** (gray) bound at the 428 site. A bound d_6 -DMSO molecule is in the bottom right of the panel. Electron density is shown for **7** from an F_o-F_c omit map (2.5σ contours). (b) **8** (gray) bound at the 507 site. Electron density is shown for **8** from an F_o-F_c omit map (2.5σ contours). (c) **8** (gray) at the NNRTI-binding pocket without rilpivirine present. The normal position of rilpivirine at the NNRTI-binding pocket is shown in magenta. Electron density is shown for **8** from an F_o-F_c omit map (4.0σ contours). (d) **9** (gray) bound at the RNase H primer grip adjacent site. Electron density is shown for **9** from an F_o-F_c omit map (3.0σ contours).