Published in final edited form as:

Top Curr Chem. 2012 ; 317: 181–200. doi:10.1007/128_2011_232.

Fragment Screening and HIV Therapeutics

Joseph D. Bauman, Disha Patel, and Eddy Arnold

Center for Advanced Biotechnology and Medicine, Department of Chemistry and Chemical Biology, Department of Medicinal Chemistry, Rutgers University, Piscataway, NJ

Abstract

Fragment screening has proven to be a powerful alternative to traditional methods for drug discovery. Biophysical methods, such as X-ray crystallography, NMR spectroscopy, and surface plasmon resonance, are used to screen a diverse library of small molecule compounds. Although compounds identified via this approach have relatively weak affinity, they provide a good platform for lead development and are highly efficient binders with respect to their size. Fragment screening has been utilized for a wide-range of targets, including HIV-1 proteins. Here, we review the fragment screening studies targeting HIV-1 proteins using X-ray crystallography or surface plasmon resonance. These studies have successfully detected binding of novel fragments to either previously established or new sites on HIV-1 protease and reverse transcriptase. In addition, fragment screening against HIV-1 reverse transcriptase has been used as a tool to better understand the complex nature of ligand binding to a flexible target.

Keywords

HIV; fragment screening; reverse transcriptase; protease; X-ray crystallography; surface plasmon resonance; drug design

1. Introduction

Human immunodeficiency virus (HIV), the causative agent of acquired immune deficiency syndrome (AIDS), remains a medical challenge with more than 33 million people currently infected worldwide [1]. HIV, like other related retroviruses, relies on the replication of its RNA genome using the host's cellular machinery. Upon infection, reverse transcriptase (RT) copies the viral ssRNA genome to double-stranded pro-viral DNA, which is then transported into the nucleus for integration into the host cell's genome. The provirus then exploits host cellular machinery to produce new infectious viral particles via normal cellular transcription and translation.

The elucidation of the viral replication cycle has identified key viral enzymatic targets--HIV-1 reverse transcriptase (RT), integrase (IN), and protease (PR)--for anti-retroviral drug discovery and design. For the most part, HIV-1 RT and PR have been the focus of extensive drug therapy efforts. It is only recently that drugs targeting viral entry and integrase have been approved. Highly active antiretroviral therapy (HAART), consisting of combination therapy usually with RT and PR inhibitors, has been found to cause a dramatic decrease in HIV-viral load within a few months. Despite such progress, treatment failure stemming from non-compliance, drug-drug interactions, and long-term drug toxicity continues to be a persistent problem [2]. Emergence of multiple drug-resistant strains of the virus due to prolonged chemotherapy continues to press the need for novel, highly efficacious drugs.

2. Fragment Based Drug Discovery

In recent years, the fragment-based approach has greatly facilitated the discovery and optimization for novel leads. Prior to its introduction in 1996, a commonly used drug discovery paradigm involved high throughput screening of several hundred thousand drug-like compounds using *in vivo* assays for the detection of relatively strong inhibiting compounds. Although this approach has been successfully utilized in the development of numerous drugs, drug design efforts were routinely plagued with challenges due to low hit rates, false positives, and substantial labor intensive lead optimization. To circumvent these problems, the fragment-based approach was introduced as an alternative tool for drug discovery and design. Here, chemically diverse libraries of small molecule compounds or fragments are screened against a target protein to find relatively weak binding compounds. The promiscuous nature of fragments allows for higher hit rates while enabling efficient search of diverse chemical space [3–5]. Additionally, the small size of the fragments allows for higher ligand efficiency, which is a measure of the atomic contribution to the overall binding energy of a ligand. Ligand efficiency is typically defined as the free energy of dissociation divided by the number of non-hydrogen atoms [3].

$$LE = \frac{-RT \ln K_D}{n_{HA}}$$

Currently, there are three approaches--fragment evolution, fragment linking/merging, and fragment self-assembly--that can guide lead optimization when utilizing fragment-based drug discovery. Fragment evolution involves the addition of functional groups to the original fragment hit to improve potency and binding. Here, the original hit acts as an “anchor” and often maintains its binding mode during the evolution process [6]. Typically, this process is guided by structural information provided by either X-ray crystallography or NMR spectroscopy.

If the target of interest has multiple fragment binding sites, a fragment linking approach can be utilized. Here, two fragment hits found binding to proximal sites within a target protein are joined using a linking group. This generally results in an improvement in the potency since the expected binding free energy of the linked molecule is greater than the sum of the binding energy of the individual fragments. Alternatively, if two or more fragments bind to overlapping sites, fragment merging can be utilized to join the fragments without the aid of a linker. It is important to note that stereochemical requirements for linking the two fragments can be restrictive since both of the fragments have to retain their original binding mode [6]

Fragment self-assembly also involves the fusion of two fragment hits into one larger molecule. In one popular approach, “click chemistry” is utilized. Chemically reactive fragments are screened such that upon binding to proximal sites within the target protein, the fragments react with each other to produce a larger inhibitor [6].

2.1 Fragment Library Design

One key to the fragment-based approach lies in the design of the fragment library. Typically, fragment libraries are relatively small in size consisting of 500 to 1,000 commercially available molecules. These compounds are selected such that a high degree of chemical diversity and synthetic tractability is achieved [3–5]. In addition, guidelines, such as the Astex “Rule of Three”, are used to ensure that compounds are indeed “fragment-like.” The Rule of Three states that fragments hits have mass < 300 Da, < 3 hydrogen bond acceptors,

3 hydrogen bond donors, a ClogP of 3, rotatable bonds of 3, and a polar surface area $60 \text{ \AA}^2$ [7, 8].

To facilitate screening efforts, fragment libraries are often grouped into cocktails of four to ten fragments. Compound solubility, toxicity, and potential chemical reactivity of the species within the cocktail itself are taken into consideration during the cocktail design process. Additionally, fragments are selected to minimize the chance of having more than one compound binding within a cocktail to the protein.

2.2 Fragment Screening Techniques

Due to the low binding affinities of fragment hits, the fragment-based approach heavily relies on sensitive biophysical methods, such as nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography, for efficient screening. NMR techniques have been a favorite in the field since it allows for the reliable detection of very weak binders using a ligand-based as well as a target-based approach towards screening of a wide-range of targets. Protein-based NMR methods rely on changes in the protein resonances upon ligand binding. In addition to identifying high affinity fragment hits, protein-based NMR methods can provide information regarding the actual binding pocket. However, these methods require long sample stability and large amounts of protein. Ligand-based NMR methods such as saturation transfer difference (STD) and water LOGSY take advantage of the differences in the ligand resonances between bound and unbound states. Ligand-based detection methods allow for rapid hit identification using relatively little amount of protein. However, tight binding ligands can be false negatives since the rate of dissociation is not large enough to distinguish between the bound and unbound states [9–12].

X-ray crystallography provides a powerful method to perform fragment screening when the drug target can form suitable crystals. For crystals to be amenable for a fragment screening campaign they must meet the following criteria.

1. The crystals must be highly reproducible and diffract X-rays to high resolution (ideally better than $2.5 \text{ \AA}$).
2. The protein must be in a biologically relevant conformation.
3. The druggable sites must not be occluded by protein-protein crystal contacts, by a natural ligand or by chemical used for crystallization or cryoprotection.
4. The crystals must be robust enough to survive the soaking of fragments.
5. The pH and ionic strength of the crystallization solution should optimally be near physiological.

Typically, crystals of the target protein are grown then soaked in solutions of either individual fragments or cocktails. Soaking is conducted at relatively high fragment concentration, ranging from 10–100 mM, since the fragments are weak binders and the protein concentration within a crystal is relatively high. Cocktails are designed such that each of the fragments within a particular cocktail should be diverse with respect to shape to allow for easy detection and deconvolution. At times, the crystal form may not be suitable for soaking, thus other crystal forms may need to be generated. An alternative to a soaking experiment is co-crystallization of either cocktails or individual fragments with the protein itself. However, this approach may require optimization of the crystallization condition for each cocktail or fragment [13–16]. An X-ray crystallographic approach is considered advantageous as it allows for the visualization of multiple binding sites and, specifically, the binding mode to facilitate structure-based lead optimization. The availability of crystallization robotics and advancements in data collection make X-ray crystallography an attractive means for screening. However, it is still an extremely labor-intensive technique

that is limited by the need for highly reproducible crystals that diffract X-rays to a reasonable resolution. The use of X-ray crystallography for fragment screening is also discussed in detail in Davies and Tickle [17].

Surface plasmon resonance (SPR) has recently become a common method of primary screening. Here, ligand binding is detected by changes in the refractive index of the solid support onto which the target protein is immobilized. The analytes (fragments) are injected in a continuous flow and a real time sensogram is recorded. The availability of multiple biosensor channels allows for rapid, high-throughput screening of multiple proteins or protein complexes in parallel. SPR consumes as little as 25–50 μg of protein while retaining a high level of sensitivity to fragments with molecular weight as low as 100 Da [18–20]. In addition to primary hit detection, SPR also provides thermodynamic and kinetic information for ligand binding. Despite such advantages, it is important to note that careful assay design and data analysis are needed, and this is discussed in more detail by Hennig et al. [21].

Mass spectroscopy (MS) and isothermal calorimetry (ITC) also have been utilized as screening tools. MS techniques, such as non-denaturing electrospray ionization MS (ESI-MS), use mass identification as the means for the detection of reversible binding events. MS analysis allows for simultaneous binding of multiple fragments, and hence direct stoichiometric detection of the binding events. Despite such advantages, application is at times limited because the protein of interest may not be stable in the presence of a volatile buffer necessary for analysis [22]. Isothermal calorimetry has widely been used to determine the thermodynamics and stoichiometry of ligand binding in solution. It has not been routinely used in fragment-based drug discovery since it is a low-throughput technique that requires relatively large amounts of protein and time and also lacks the sensitivity for relatively weak binders needed for fragment screening [23, 24]. However, with recent improvements in instrumentation, ITC is slowly gaining ground as a screening tool.

Fragment-based drug discovery has been successfully applied to numerous targets. The success of this method depends heavily dependent upon the design of the fragment library as well as the wide range of techniques available for efficient screening. Thus, it has become an attractive approach for the design of novel HIV therapeutics to circumvent the drug resistance and adherence problems being faced today. Fragment screening against validated viral targets, specifically protease and reverse transcriptase, has been reported thus far.

3 HIV-1 Protease

HIV-1 protease (PR) plays a crucial role in the late stage of viral replication. It is responsible for the formation of viral proteins from the cleavage of the *gag-pol* polypeptide produced from the proviral transcription. Site-directed mutagenesis studies showed that a single point mutation can sufficiently inactivate the enzyme and stop viral infectivity, thus making protease an attractive target for antiretroviral therapy [25].

HIV-1 protease is a symmetrical homodimer consisting of two identical subunits of 99 amino acids. Its active site is formed at the dimer interface and contains two conserved, catalytic aspartic acid residues. A water molecule bound to the enzyme between the two aspartates acts as the nucleophile for catalysis. Each monomer contains a prominent β hairpin loop, known as the “flap,” that projects over the substrate-binding cleft. These flaps are highly flexible and can undergo large localized conformational changes upon substrate and inhibitor binding [25–27].

The first series of HIV-1 PR inhibitors, referred to as peptidomimetic inhibitors, are transition state mimics that resemble peptide substrates. They are relatively flexible, linear molecules with a well-defined backbone from which hydrophobic groups are projected into

four or more of the subsites of the HIV-1 protease active site. The inhibitors function by creating a hydrogen bond network with a tetra-coordinated structural water molecule that is tightly bound between the inhibitor and the flaps. Inhibition is also depended upon critical interactions between the catalytic aspartates and the inhibitor [27].

Poor pharmacokinetic profiles and complex syntheses of peptidomimetic inhibitors led to a second class of HIV-1 PR inhibitors, loosely termed non-peptidic protease inhibitors. These inhibitors typically consist of a rigid, cyclic core with groups projected into the central subsites of the enzyme. Interestingly, the structural water was found to be absent in the crystal structures of non-peptidic inhibitors bound to HIV-1 PR. Hydrogen bond accepting groups were found capable of retaining the hydrogen bonding interactions with the flap amide nitrogen directly without the presence of the water [27, 28]. The ability of non-peptidic inhibitors to retain the necessary interactions for inhibition has made them attractive for drug design efforts. The availability of structural data has greatly facilitated the design of novel, highly efficacious peptidic and non-peptidic HIV-1 PR inhibitors and the search for possible allosteric HIV-1 PR inhibitors.

3.1 Fragment Screening

Examination of several different ligand bound complexes of HIV-1 PR suggests that some of the drug-resistance mutations observed may alter the equilibrium between the closed and open states of the protein, thereby possibly decreasing drug binding affinity. Based on molecular dynamic simulations comparing the wild-type HIV-1 protease to the V82F/I84V drug resistant mutant, the mutant was found capable of opening the flaps much farther with a greater degree of flexibility than the wild-type HIV-1 PR. In addition, the simulations revealed the presence of a solvent-exposed cleft, referred to as the “exo site”, for both the wild-type and mutant HIV-1 PR in the closed conformation. This suggested a potential allosteric pocket that could inhibit protease by suppressing the motions of the flaps [29–30].

An X-ray crystallography-based fragment screening was undertaken by Perryman *et al.*, [30] to identify potential molecules targeting the newly discovered site. The Active Sight fragment library (San Diego, USA), consisting of 384 compounds dissolved in DMSO was screened against HIV-1 PR with and without an active site inhibitor, TL-3. The library itself consisted predominantly of compounds with a single rigid core with one to three small substituents. It was subdivided into a total of 96 cocktails with each cocktail consisting of four highly shape diverse fragments. A combination of soaking and co-crystallization approach was utilized against five different crystal forms. Altogether, 808 crystals were screened and 378 datasets were collected and refined.

Individual fragment soaks using the apo C222₁ crystal form at a 10 mM fragment concentration yielded no hits. Similar results were also observed for soaking experiments for both the apo and TL-3 bound P2₁2₁2 crystal form. Co-crystallization of the P6₁22 crystal form with an active site inhibitor, TL-3, and fragment cocktails or individual fragments at concentrations of 2.5 mM and 10 mM, respectively, was undertaken. For the large part, no fragment binding was observed. However, three cocktails---D9, F1, and F4---produced two new crystal forms.

As with any X-ray crystallography based experiment, solvent was found to play a crucial role in the screening process. This was also true in the case of HIV-1 protease. Initial soaking experiments revealed a sensitivity of the P4₁ crystal form of the unliganded protein towards the DMSO concentration. To circumvent this problem, similar concentration of DMSO was used for both drop preparation and soaking to retain diffraction quality. A new crystallization condition was sought to produce crystals that could handle the DMSO concentration and avoid the original precipitant, PEG 8000, which was found bound to HIV-1 PR.

Despite optimizing the crystal form, a large percentage of soaking experiments using the apo protease crystal form, P2₁2₁2, revealed that the exo site was occupied with acetate and water molecules from the buffer. Similarly, the co-crystallization of HIV-1 PR with TL-3 and fragments revealed that DMSO and water molecules occupied the exo site. This is a common occurrence with an X-ray crystallographic approach. Both the solvent and fragment compete for the same binding sites; however, a high solvent to fragment ratio in the soaking solution can permit solvent binding to occur despite the relatively weak interactions observed.

Co-crystallization experiments revealed interesting electron density at the exo site for the D9 cocktail, which changed the expected crystal form of the HIV-1 PR-TL-3 complex from P6₁22 to P2₁2₁2. Subsequent cocktail deconvolution led to the identification of 4D9, 2-methylcyclohexanol, as the bound fragment (Figure 1). This was further confirmed by soaking the P2₁2₁2 crystal form of the HIV-1 PR-TL-3 complex in a 20 mM solution of 4D9. Interestingly, the complementary experiment using the P6₁22 form did not reveal fragment binding. This result was attributed to the packing interactions observed between the two different crystal forms. Comparison of the structures revealed that in the P2₁2₁2 form I72 from subunit A faces the solvent to allow the side chain of L63 to flip up to accommodate binding of 4D9. In the P6₁22 form, I72 interacts with itself through a crystallographic two-fold axis, thereby preventing the rearrangement of L63 needed for fragment binding.

This highlights one key disadvantage of using X-ray crystallography as a screening tool. Compared to solution-based screening tools, protein flexibility within a crystal is greatly limited. Additionally, the crystal packing can hinder fragment binding or create artificial fragment binding sites through non-physiological crystal subunit interactions. Tight packing may require longer soaking times, whereas, in the case with HIV-1 PR, crystal packing can also affect the availability of pockets for binding. It is also important to note that fragment binding reported in this study was observed for only one monomer despite the fact that protease is a symmetrical homodimer. This can be explained by the fact that the crystallographic environment around each monomer is different. Thus, effective screening may require the use of multiple crystal forms of the protein to take into consideration protein flexibility and accessibility.

Co-crystallization of cocktails F1 and F4 with protease and TL-3, respectively, also revealed two additional fragments---1F1 (indole-6-carboxylic acid) and 2F4 (2-acetylbenzothiophene)---binding in a similar manner to a pocket on the surface of the flap in Monomer B (Figure 1). Fragment binding induces significant conformational changes with a lateral shift in the anti-parallel β strand segments at residues 45–47, 53–56, and 78–81 and a rearrangement at residues 35–41. Structural comparison of HIV-1 PR monomers with the flaps open, closed, and with and without 1F1, specifically at the base of the flap, revealed that 1F1 binds in a part open form despite having TL-3 occupying the active site and the flaps being closed. This suggests that the allosteric pocket may be functionally relevant for protease activity however further experiments need to be conducted for validation.

X-ray crystallography based fragment screening was successfully utilized for the validation of the exo site suggested by molecular dynamics studies. Additionally, a novel binding pocket was discovered during the process. Although further studies are needed to demonstrate the inhibitory potential of these sites, the availability of structural data can facilitate further design. The screening process highlighted several key advantages and disadvantages to consider when utilizing an X-ray crystallography based approach.

4 HIV-1 Reverse Transcriptase

As mentioned previously, HIV-1 reverse transcriptase (RT) is a critical enzyme for viral replication. It is responsible for the conversion of the (+) single-stranded RNA viral genome into double-stranded DNA. At the same time, it is also responsible for the degradation of the RNA genome after it is transcribed into DNA. X-ray crystal structures have revealed that HIV-1 RT is a heterodimer consisting of p66 and p51 subunits. Both p66 and p51 have the same sequence except the p51 subunits lacks the C-terminal RNase H domain. Despite sequence commonality, both subunits vary greatly with respect to conformation. The p66 subunit is arranged such that its N-terminal region resembles an open right hand containing three subdomains, aptly referred to as fingers, palm, and thumb. Following the thumb subdomain, there is a connection domain, which leads to the C-terminal RNase H domain. In the p66 subunit, three catalytic residues are exposed in the nucleic acid binding cleft. However, these three residues are buried in the p51 subunit, which lacks this cleft [31, 32].

Currently, there are two broad classes of drugs targeting reverse transcriptase activity-- nucleoside/nucleotide RT inhibitors (NRTIs) and non-nucleoside RT inhibitors (NNRTIs). NRTIs are analogs of endogenous 2'-deoxy-nucleosides that lack the 3'-hydroxyl needed for polymerization. Upon incorporation by RT, they act as chain-terminators of the viral reverse transcripts. NNRTIs bind to a normally closed allosteric binding pocket in the palm subdomain. This stabilizes a single conformation of the palm/thumb subdomains, which is not sufficient for polymerization to occur. Additional classes of inhibitors in active development are nucleotide-competing RT inhibitors (NcRTIs) [33], p66/p51 dimerization inhibitors [34], and RNase H inhibitors [35].

4.1 Fragment Screening Against HIV-1 Reverse Transcriptase by X-ray Crystallography

The flexible nature of RT is a critical feature both from a biological and a drug design perspective [36, 37]. Biologically, the flexibility of the enzyme plays a crucial role in the catalytic activity of the protein by allowing for the exact coordination of several domains in the protein to occur. From a drug design perspective, protein flexibility combined with various interdomain hinges present throughout the protein suggests the possible existence of novel druggable sites.

We utilized fragment screening by X-ray crystallography to investigate potential new inhibitory sites in HIV-1 RT. Initially, the apo crystal form of HIV-1 RT was utilized for screening. However, it was found to be unsuitable for soaking experiments due to sensitivity towards the soaking condition and methodology despite extensive optimization. As a result, an alternate crystal form, RT in complex with TMC278, a potent NNRTI, was used for screening. An engineered form of HIV-1 RT was used; it had been developed because crystallization trials of wild-type HIV-1 RT complexed with TMC278 had failed to give crystals of X-ray diffraction quality.

Taking advantage of the tremendous amount of available structural data, crystal engineering was used to improve the crystal quality and subsequent X-ray diffraction. The termini of p66 and p51 were truncated to remove residues found to be disordered in previous crystal structures. In addition, common crystal contacts were mutated to increase the likelihood of

obtaining new crystal forms and high-entropy residues (lysines and glutamic acids) were mutated to alanine to lessen the entropic penalty of forming crystal contacts. After several rounds of mutagenesis and testing, an HIV-1 RT variant (RT52A) was made to diffract X-rays to 1.8 Å resolution when crystallized in complex with TMC278 and other NNRTIs [38]. These crystals were found to be highly reproducible and robust for fragment screening but could not be used to find new NNRTIs since the NNRTI binding pocket is occupied.

A fragment library consisting of 775 commercially available compounds was assembled in-house. This library consisted of 500 compounds purchased from Maybridge (Cornwall, UK), 175 individual compounds purchased based on the recommendations of Christophe Verlinde and Wim Hol [39], and an additional 100 compounds were generously gifted by James Williamson (The Scripps Research Institute, La Jolla, unpublished). The fragments were divided into cocktails containing an average of five compounds. The cocktails were designed such that structural diversity and cocktail solubility was maximized and the chemical reactivity between fragments of the same cocktail was not an issue. To ensure maximum structural diversity within the cocktail, a program, FROCIVANTO, was developed (Eck and Arnold, unpublished results from this laboratory). FROCIVANTO produces shape fingerprints for all fragments in the library and then generates the distance matrix for the fingerprints, utilizing the Euclidean-style Ultrafast Shape Recognition distance metric [40]. The fragments are then ordered based on shape similarity and a simple partitioning technique is used whereby cocktails are generated sequentially, with every N^{th} fragment being selected from the shape-based ordering. Typically, cocktails with an average of five compounds were designed and prepared such that the final concentration of each fragment was 100 mM in d_6 -DMSO.

Initial screening experiments pressed the need for the optimization of the soaking and freezing conditions with respect to fragment concentration and solubility, soaking time, and crystal stability. The RT52A-TMC278 crystals were found to survive soaking conditions with the final fragment concentration of 20 mM. Importantly, 20% (v/v) d_6 -DMSO was also found to serve a dual purpose, both as a solvent for the fragment as well as a cryo-protectant during crystal freezing. This facilitated the soaking process since only one solution was needed for both soaking and freezing. Fragment solubility in the soaking solution proved to be a major obstacle to overcome. The crystal soaking time was increased from several seconds to approximately one to two minutes to counter the lower fragment concentration in solution and allow for the crystal to equilibrate in the soaking solution. Also, the addition of 80 mM L-arginine was found to improve fragment solubility in the soaking solution. This was found to have a tremendous impact on the screening results since many of the fragment hits were found to bind RT-TMC278 complex only in the presence of L-arginine. It is important to note that the L-arginine served only as an additive to improve solubility and was not found present in the electron density.

In addition to optimizing experimental conditions, a protocol was designed to efficiently screen and process the wealth of data collected during the course of the screening. Due to the high concentration of d_6 -DMSO present in the soaking solution, a reference structure for a crystal soaked in 20% (v/v) d_6 -DMSO soaking condition without any fragment present was determined to identify the d_6 -DMSO binding sites. Multiple binding sites for d_6 -DMSO were observed (Figure 2A). To avoid misinterpretation of the solvent molecules as fragment binding, this structure served as a 'blank' for subsequent analysis. As shown in Figure 2b, a high-speed pass was initially performed for a crystal to collect X-ray diffraction to no better than 2.1 Å. The high-speed pass maximizes the quantity of fragments screened during a time-limited X-ray data collection trip (often five to ten datasets per hour are possible).

The diffraction calculated in the CNS program system was then immediately processed at the beamline using HKL2000 and a map using F_o-F_o coefficients comparing the fragment dataset to the 'blank' reference dataset. The maps were then evaluated for changes in electron density for the presence of any strong positive density that may be indicative of fragment binding or the movement of the protein backbone, residue side chains. A high-resolution pass was conducted when the F_o-F_o maps suggested fragment binding. These datasets were further refined using CNS to clearly delineate fragment binding. Binding detected from cocktail soaks was subsequently verified by soaking of individual fragments. These compounds were then tested for polymerization and RNase H inhibition using activity assays. In addition, fragments were co-crystallized with the apo protein to see whether or not binding was maintained without the NNRTI-pocket being occupied and to verify that inhibition was not due to fragment binding to the NNRTI pocket. Chemical derivatives of the most promising fragments were purchased and tested for binding and inhibition using X-ray crystallography and activity assays, respectively.

A total of 705 datasets with an average resolution of 2.1 Å were collected for 742 compounds out of the 775 screened. From these datasets, 34 compounds were found to bind to the protein complex giving a hit rate of 4.4%. Fifty binding events were observed with many of these compounds binding to multiple locations throughout the protein complex, including the hinge regions as well as the polymerase and RNase H active sites (shown in Figure 3). In addition, two new allosteric pockets--the knuckles and the NNRTI adjacent--were discovered from the screening efforts (Figure 4). The knuckles pocket had not been observed in any of the RT structures thus far and was found to be present only in conjunction with fragment binding.

Prior to fragment binding, the knuckles pocket is a non-solvent exposed cavity near the incoming dNTP substrate-binding site (Figure 4a). Upon fragment binding to this cavity, the polypeptide backbone rearranges (Ser117 C α is displaced 2.8 Å) to create a solvent-exposed pocket. The formation of the pocket results in a 3.2 Å movement in the backbone for active site residues Tyr115 and Phe116, which are involved in incoming nucleotide binding during polymerization. The pocket opens also at residue Ser163 creating a potential direction for fragment growth. The original fragment hits and derivatives found binding to this site were found to have inhibitory activity with the top derivative having an IC₅₀ of 600 μM with a ligand efficiency of 0.37 kcal/mol NHA.

In a typical NNRTI-bound structure without fragment binding, the NNRTI adjacent binding site is a solvent-accessible pocket at the p66/p51 interface that is separated from the NNRTI-binding pocket by Val179 and Ile180. Bordered by Thr139, Pro140, Thr165, Leu168, Lys172, and Ile180, this pocket expands upon fragment binding (Figure 4b). The fragments binding to this pocket were found to be inhibitory in the absence of an NNRTI. The best primary hit was found to have an IC₅₀ value of 450 μM and a ligand efficiency of 0.33 kcal/mol•NHA. The close proximity to the NNRTI-binding pocket may allow for expansion of known NNRTIs by fragment linking. Alternatively, compounds targeting solely the NNRTI adjacent pocket can be potentially promising inhibitors against both wild-type and NNRTI drug-resistant forms of the protein.

Interestingly, during the course of the screening it became apparent that halogenated compounds were frequently found to be hits. In fact, four out of the 17 brominated compounds screened and seven out of the 29 fluorinated compounds screened were found to bind throughout the protein giving a hit rate of 23.5% and 24.1%, respectively. Although this preference for halogen containing compounds is not fully understood, the higher hit rates suggest that use of halogenated compounds may be advantageous. One of the compounds, 4-bromopyrazole, was found bound to 11 sites through out the protein,

including the two new sites described (Figure 5). Burley and coworkers at SGX Pharmaceuticals (San Diego, CA), described using a fragment library containing a large fraction of brominated compounds, in part to permit Br anomalous scattering measurements for rapid assessment of fragment binding [41].

X-ray crystallography was successfully used for primary fragment screening despite the challenges faced. Fragment screening was surprisingly effective in discovering new inhibitory sites on HIV-1 RT when hundreds of drug discovery projects using other methods had failed to detect them. This demonstrates the utility of fragment screening by X-ray crystallography to probe a protein for new allosteric sites.

4.1.1 Fragment Screening and Surface Plasmon Resonance—Recently, Geitmann *et al.* [42], has reported the use of SPR as a screening tool for identifying novel NNRTI scaffolds from a fragment library consisting of 1,040 compounds screened against HIV-1 RT. Prior to screening, the fragment library was first evaluated for chemical diversity using a Tversky similarity analysis to compare the fragments to each of 826 published NNRTIs extracted from the BindingDB. The analysis revealed that the majority of the compounds in the library were significantly different from the published NNRTIs. The results also indicated that 28 out of 1,040 fragments were substructures of NNRTIs with respect to simple atom connectivity, with many of them having different functionality and polarity compared to the established NNRTI. Overall, the use of the Tversky similarity analysis provided a means for ensuring the exploration of novel chemical scaffolds.

A primary screen was conducted against the wild-type HIV-1 RT with the fragment concentration ranging from 50 to 400 μM and nevirapine, a potent NNRTI, as the positive control. The use of multiple concentrations enabled the identification of false positives, a common problem with SPR-based assays. A total of 165 compounds were selected with apparent K_D values less than 1 mM and a stoichiometry of 0.75–5 times the value obtained with nevirapine. The sensograms of the 165 compounds were then subjectively evaluated for basic interaction characteristics, such as rate of dissociation, to eliminate any false positives. In addition, sensograms were evaluated for secondary effects stemming from poor compound solubility, conformational changes within the bound protein, and clearance of the biosensor surface for false positives. A total of 69 compounds were eliminated due to strong secondary effects or slow dissociation thereby leaving only 96 hits from the primary screen.

The 96 compounds identified as primary hits were then subjected to two independent experiments in parallel. First, an-SPR based competition assay was used to screen each of the 96 fragments and nevirapine at a concentration of 200 μM and 20 μM , respectively. The sensograms were also evaluated to remove any false positives based on the previously mentioned criteria. Out of the 96 compounds originally identified from the primary screen, only 20 were found to compete with nevirapine. At the same time, the hits from the primary screen were also evaluated in an enzyme inhibition assay. Of the 96 compounds assayed, 27 compounds were considered hits with IC_{50} values lower than 1 mM.

Ten compounds from the set of 27 were found to compete with nevirapine as well as inhibit wild-type RT activity with submillimolar IC_{50} values. The compounds were then re-evaluated under more fully optimized conditions against the wild-type RT for further validation. Fragments **9** and **10** failed to reproducibly show inhibition and were dropped from further analysis. The remaining 8 fragments were then screened against drug-resistant mutants of HIV-1 RT, K103N, L100I, and Y181C, respectively. Only fragment **1**, bromoindanone, was found to have an IC_{50} value lower than 25 μ M against all four enzyme variants. This gives bromoindanone a very high ligand efficiency of greater than 0.57 kcal/mol $\cdot$ NHA, making it a potentially promising starting point for lead development for a novel NNRTI.

4.1.2 Fragment Screening and Pocket Flexibility—Interestingly, only two of the 28 fragments initially identified as substructures of published NNRTIs from the similarity analysis were found to be hits among the 20 compounds found through the competition screen but failed to show enzymatic inhibition. Although the lack of inhibition of the two compounds could be explained by the change in polarity and conformation preference compared to the parent NNRTI, the poor hit rate from the pool of NNRTI-based substructures suggests a lack of an efficient binding hot spot. Brandt *et al.* [43] utilized fragment screening and SPR as a tool to understand the nature of fragment binding to flexible pockets, specifically the NNRTI pocket in HIV-1 RT.

Twenty-one fragments stemming from three NNRTIs were purchased and assayed in an SPR-based biosensor assay. Based on the ligand efficiency of the parent compound, the deconstructed fragments were expected to have high ligand efficiency and IC_{50} values below 3 mM. However, only nine of the 21 compounds tested were found to be binders. Three of the 21 fragments were found to have reproducible K_D values in the micromolar range. The six remaining compounds were also found to bind but K_D values fell above the highest concentration tested and were estimated despite a large linear component in the equation or by a less good fit to data. A closer look at the binders and the non-binders from the deconstruction experiment revealed that the substructures found to bind were relatively large compounds. In fact, only three (**11**, **14**, **15**) out of the nine that bound could be considered as fragments based on a molecular weight less than 300 Da and the average number of heavy atoms for the binders and non-binders was 21 and 14, respectively.

To account for these observations, a corrected ligand efficiency, which incorporates a ligand independent free energy fee, ΔG_{ind} , was proposed. Here, the ΔG_{ind} is defined as the energy required to create the NNRTI-binding site ($\Delta G_{\text{opening}}$) as well as the change in free energy caused by the loss in translational and rotational entropy ($T\Delta S_{\text{tr}}$).

$$LE^* = \frac{\Delta G_{\text{opening}} - T\Delta S_{\text{tr}} - RT \ln K_D}{n_{\text{HA}}}$$

A strong correlation between experimental and predicted results was observed when using a value of 7.0 kcal/mol for ΔG_{ind} . However, this model assumes that the protein-ligand interaction energy is uniformly distributed over the ligand-protein interface thereby ignoring the existence of hot spots which in turn underestimates the potential maximum affinity of fragments.

The modified ligand efficiency, LE^* , will have a greater preference for smaller fragments, thereby rendering it inefficient as a measure for prioritizing compounds for lead development. To circumvent the observed size dependency on LE^* , Brandt and coworkers proposed the use of a modified fit quality (FQ^*), which is a function of the number of heavy atoms.

$$FQ^* = \frac{LE^*}{LE_Scale^*}$$

where

$$\text{LE}_{\text{Scale}} = 0.0975 + \frac{17.3}{n_{\text{HA}}} + \frac{35.1}{n_{\text{HA}}^2} - \frac{493}{n_{\text{HA}}^3}$$

FQ* analyses of the experimental library as well as the 946 NNRTIs published in the database were found to be suboptimal. Additionally, a dramatic drop in the FQ* was observed for compounds with less than 20 heavy atoms; thereby, suggesting that either larger concentrations may be needed for fragment screening or alternatively small fragments may not be ideal for screening. Although FQ*, like LE*, cannot help prioritize hits for lead development, it can provide information about the binding site landscape which in turn can facilitate the course of lead design.

Fragment screening has been used as a tool to identify novel scaffolds targeting the NNRTI-binding pocket of HIV-1 RT. However, only one compound out of the 1,040 screened was found to be active against both the wild-type and mutant variants of HIV-1 RT. Additionally, a poor hit rate for NNRTI-like fragments in primary screens was observed. A deconstruction analysis, which involves screening of fragments based on known NNRTIs, was used to assess the amenability of the NNRTI pocket towards fragment screening.

5 Closing Remarks

Fragment screening is breathing new life into drug discovery against HIV-1 targets. The relatively low cost of purchasing and curating a collection of 500–1000 compounds makes it an attractive method of drug discovery for academic as well as industrial groups. Fragment screening has been used to discover new druggable sites as well as novel scaffolds against well-established binding pockets. The novel nature of the compounds discovered by fragment screening will feed drug design efforts against HIV-1 for years to come.

Abbreviations

AIDS	acquired immune deficiency syndrome
HIV	human immunodeficiency virus
PR	protease
RT	reverse transcriptase
HAART	highly active antiretroviral therapy
NMR	nuclear magnetic resonance
SPR	surface plasmon resonance
MS	mass spectroscopy
ITC	isothermal calorimetry
NRTI	nucleoside/nucleotide reverse transcriptase inhibitor
NNRTI	non-nucleoside reverse transcriptase inhibitor
ClogP	calculated logarithm of octanol-water partition coefficient
IC₅₀	concentration of a compound leading to 50% enzyme inhibition
K_D	dissociation constant
LE	ligand efficiency

References

1. World Health Organization. World Health Statistics. 2008. <http://www.who.int/whosis/whostat/2008/en/index.html>
2. Mehellou Y, De Clercq E. *J Med Chem*. 2010; 53:521–538. [PubMed: 19785437]
3. Hadjuk PJ, Greer J. *Nat Rev Drug Discov*. 2007; 6:211–219. [PubMed: 17290284]
4. Hestekamp T, Whittaker M. *Curr Opin Chem Biol*. 2008; 12:260–268. [PubMed: 18316043]
5. Erlanson DA, Wells JA, Braisted AC. *Annu Rev Biophys Biomol Struct*. 2004; 33:199–223. [PubMed: 15139811]
6. Rees DC, Congreve M, Murray CW, Carr R. *Nat Rev Drug Discov*. 2004; 3:660–672. [PubMed: 15286733]
7. Congreve M, Chessari G, Tisi D, Woodhead AJ. *J Med Chem*. 2008; 51:3661–3680. [PubMed: 18457385]
8. Congreve M, Carr R, Murray C, Jhoti H. *Drug Discov Today*. 2003; 8:876–877. [PubMed: 14554012]
9. Zartler ER, Shapiro MJ. *Curr Opin Chem Biol*. 2005; 9:366–370. [PubMed: 15925537]
10. Siegal G, Ab E, Schultz J. *Drug Discov Today*. 2007; 12:1032–1039. [PubMed: 18061882]
11. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. *Science*. 1996; 274:1531–1534. [PubMed: 8929414]
12. Hadjuk PJ, Meadows RP, Fesik SW. *Quart Rev Biophys*. 1999; 32:211–240.
13. Jhoti H, Cleasby A, Vedonk M, Williams G. *Curr Opin Chem Biol*. 2007; 11:485–493. [PubMed: 17851109]
14. Carr R, Jhoti H. *Drug Discov Today*. 2002; 7:522–527. [PubMed: 11983569]
15. Hartshorn MJ, Murray CW, Cleasby A, Frederickson M, Tickle IJ, Jhoti H. *J Med Chem*. 2005; 48:403–413. [PubMed: 15658854]
16. Murray CW, Blundell TL. *Curr Opin Struct Biol*. 2010; 20:497–507. [PubMed: 20471246]
17. Davies TG, Tickle IJ. Fragment screening using X-ray crystallography. *Top Curr Chem*. 2011; 10.1007/128_179
18. Navratilova I, Hopkins A. *ACS Med Chem Lett*. 2010; 1:44–48.
19. Neumann T, Junker HD, Schmidt K, Sekul R. *Curr Top Med Chem*. 2007; 7:1630–42. [PubMed: 17979772]
20. Huber W, Mueller F. *Curr Pharm Des*. 2006; 12:3999–4021. [PubMed: 17100609]
21. Hennig M, Ruf A, Huber W. Combining biophysical screening and X-ray crystallography for fragment-based drug discovery. *Top Curr Chem*. 2011; 10.1007/128_225
22. Zartler, ER.; Shapiro, MJ. Fragment-based drug discovery: a practical approach. Zartler, ER.; Shapiro, MJ., editors. Wiley; United Kingdom: 2008.
23. Torres FE, Recht MI, Coyle JE, Bruce RH, Williams G. *Curr Opin Struct Biol*. 2010; 20:598–605. [PubMed: 20888754]
24. Ladbury JE, Klebe G, Freire E. *Nat Rev Drug Discov*. 2010; 9:23–27. [PubMed: 19960014]
25. Robins T, Plattner J. *J Acq Immun Def Syn*. 1993; 6:162–170.
26. Kempf DJ, Sham HL. *Curr Pharm Des*. 1996; 2:225–246.
27. Abdel-Rahman HM, Al-karamany GS, El-Koussi NA, Youssef AF, Kiso Y. *Curr Med Chem*. 2002; 9:1905–1922. [PubMed: 12369875]
28. Huff JR. *J Med Chem*. 1991; 34:2305–2314. [PubMed: 1875332]
29. Perryman AL, Lin JH, McCammon JA. *Prot Sci*. 2004; 13:1108–1123.
30. Perryman A, Zhang Q, Soutter HH, Rosenfeld R, McRee DE, Olson AJ, Elder JE, Stout CD. *Chem Biol Drug Des*. 2010; 75:257–268. [PubMed: 20659109]
31. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. *Science*. 1992; 256:1783–1790. [PubMed: 1377403]
32. Jacobo-Molina A, Ding J, Nanni RG, Clark AD, Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL, Clark P. *Proc Natl Acad Sci USA*. 1993; 90:6320–6324. [PubMed: 7687065]

33. 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, Gotte M, Hertogs K. *J Virol*. 2006; 80:12283–12292. [PubMed: 17020946]
34. Camarasa M-J, Velázquez S, San-Félix A, Pérez-Pérez MJ. *Antivir Chem Chemother*. 2005; 16:147–153. [PubMed: 16004078]
35. Tramontano E, Di Santo R. *Curr Med Chem*. 2010; 17:2837–2853. [PubMed: 20858167]
36. Gotte M, Rausch JW, Marchand B, Sarafianos S, Le Grice SFJ. *Biochim Biophys Acta*. 2010; 1804:1202–1212. [PubMed: 19665597]
37. Sarafianos SG, Marchand B, Das K, Himmel DM, Parniak MA, Hughes SH, Arnold E. *J Mol Biol*. 2009; 385:693–713. [PubMed: 19022262]
38. Bauman JD, Das K, Ho WC, Baweja M, Himmel DM, Clark AD, Oren DA, Boyer PL, Hughes SH, Shatkin AJ, Arnold E. *Nucleic Acids Res*. 2008; 36:5083–5092. [PubMed: 18676450]
39. 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. *Curr Top Med Chem*. 2009; 9:1678–1687. [PubMed: 19929835]
40. Ballester PJ, Richards WG. *J Comput Chem*. 2007; 28:1711–1723. [PubMed: 17342716]
41. Blaney, J.; Nienaber, V.; Burley, SK. *Fragment-based Approaches in Drug Discovery*. Jahnke, W.; Erlanson, DA., editors. Wiley-VCH Verlag GmbH & Co. KGaA; 2006.
42. Gietmann M, Elinder M, Seeger C, Brandt P, de Esch IJP, Danielson UH. *J Med Chem*. 2011; 54:699–708. [PubMed: 21207961]
43. Brandt P, Gietmann M, Danielson UH. *J Med Chem*. 2011; 54:709–718. [PubMed: 21207958]

Figure 1.

A cartoon representation of HIV-1 protease with TL-3 bound in the active site. 4D9 and 1F1 represent fragment binding to the novel sites identified through screening efforts.

Figure 2.

(A) X-ray crystal structure of HIV-1 RT without any fragments soaked but frozen with all other solution components present including 20% d₆-DMSO. The fingers, palm, thumb, connection and RNase H subdomains are color coded blue, red, green, yellow, and orange, respectively. The NNRTI-binding pocket is shown in brown and the polymerase active site is colored dark green. The positions of the bound d₆-DMSO molecules are shown as cyan spheres. (B) Scheme of data collection and initial processing.

Figure 3.

Cartoon representation of HIV-1 RT with p66 colored green and p51 colored blue. Each fragment found bound is shown as a space-filling model. The polymerase active site, knuckles, NNRTI adjacent, 428, 507, RNase H backside, and RNase H active site are colored coded.

Figure 4.

Cartoon representation of novel inhibitory binding sites. (A) The knuckles binding site at the fingers/palm of p66 junction. Unbound is colored orange and bound is colored blue. The transparent representation shows the molecular surface of the open pocket. The 4-bromopyrazole fragment is shown in the pocket. (B) The NNRTI adjacent binding pocket is shown with a transparent surface representation. At the top is the NNRTI-binding pocket with TMC278 shown in green and blue. A conserved water molecule is shown between the two pockets as a blue sphere. 4-bromopyrazole is shown in bound in the lower NNRTI adjacent pocket.

Figure 5.
Cartoon representation of HIV-1 RT with p66 colored green and p51 colored blue. Each 4-bromopyrazole molecule found are shown as colored spheres.