A Substituted Indole Scaffold as a Class of Allosteric HIV-1 Integrase Inhibitor

Dissertation

Presented in Partial Fulfillment of the Requirements for the Undergraduate Graduation with
Research Distinction at The Ohio State University

By

Ross H. Bockbrader

Bachelors of Science in Pharmaceutical Sciences

College of Pharmacy

The Ohio State University

2019

Research Advisor: James R. Fuchs

Acknowledgements

Abstract

Although a number of drugs have been developed and approved to combat HIV over the past 40 years, mutations of the virus and its protein targets have necessitated the continued discovery and design of even more novel compounds today. In the Fuchs lab, we are approaching this problem by targeting an allosteric binding site on HIV-1 integrase, an integral enzyme that is responsible for integrating the viral genome into the host's DNA. This allosteric site is the same binding site as that used by the cofactor LEDGF/p75, a host cell protein that promotes the activity of integrase. Binding of small molecule drugs at this allosteric binding site along the CCD dimer interface of integrase results in the inactivation of the protein due to the formation of aberrant integrase multimers. This hyper-multimerization occurs as a result of interaction of the CTD subunit of another integrase protein with the CCD dimer interface. We are currently synthesizing new inhibitors based on a well-studied quinoline core scaffold and scaffold hoping to an indole core scaffold. The indole core scaffold was chosen due to the existence of well-established synthetic routes and the ease of functionalization of the indole core. In addition, the tilt of aromatic ring in the allosteric binding pocket has been hypothesized to improve binding of the compounds in the presence of known mutations, namely the A128T mutation. Inhibitors of this type are also structurally unique and may lead to the development of new intellectual property. The ultimate goal of these studies is to develop compounds with greater intregrase inhibitory activity through binding to the LEDGF/p75 site than previously studied compounds by optimizing functional groups around the indole core.

Table of Contents

List of Figures

Introduction to HIV and HIV Therapy

Human immunodeficiency virus (HIV) affects approximately 37.9 million people worldwide. In 2017 alone, HIV led to over 1 million autoimmune deficiency (AIDs) related deaths[1]. Since the rapid emergence of this retrovirus in the late-twentieth century, significant research efforts have led to a better understanding of HIV and the etiology of the symptoms it induces. The study of the biology of HIV and the viral life cycle (Figure 1) has paved the way for the development of therapeutics to prevent or treat the disease by rendering the continuation of the viral life cycle infeasible[2].
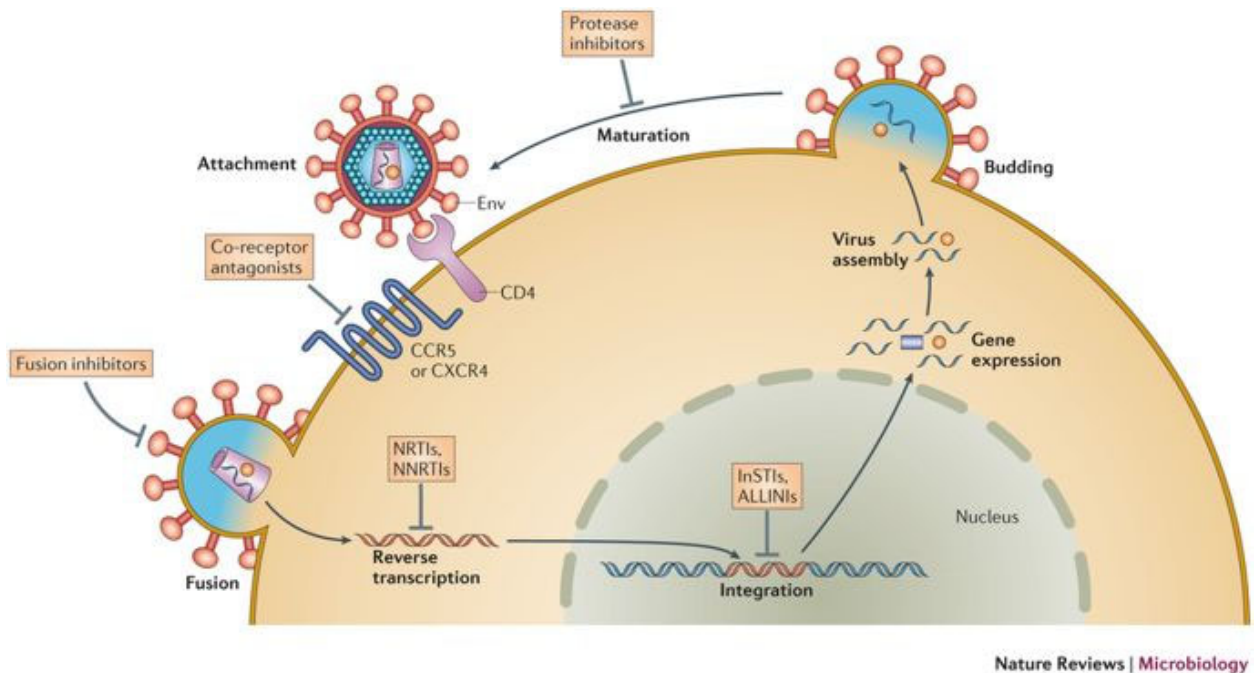
Figure 1. The viral life cycle of HIV and potential drug targets[2]

Interestingly, HIV only encodes 15 proteins, many of which have become important targets in the fight against this disease. By studying the role and structure of these unique protein targets, a variety of therapeutic approaches have been implemented to specifically affect viral fitness and infectivity, and thereby generate effective treatments. Fusion inhibitors, for example,

block the fusion of the virus to the host's cell membrane. Similarly, chemokine receptor type 5 (CCR5) antagonists are used to block the CCR5 co-receptor on the surface of immune cells to prevent the virus from entering (AIDS Info). Nucleotide/nucleoside reverse transcriptase inhibitors (NRTIs) function as chain terminators to block extension of the viral DNA chain during reverse transcription (Immunopaedia) and integrase inhibitors (INISs) impair the integration step by binding to the active site of integrase thus inhibiting the viral DNA's ability to integrate itself into the host DNA. Currently, each of these mechanisms for combating HIV are utilized in currently available FDA-approved therapeutic drugs. A list of the individual HIV drugs "recommended for the treatment of HIV in the United States based on the U.S. Department of Health and Human Services HIV/AIDS medical practice guidelines" is included in Table 1.

Drug Name	Drug Class	Drug Name	Drug Class	Drug Name	Drug Class
Abacavir	NRTI	Etravirine	NNRTI	Saquinavir	PI
Emtricitabine	NRTI	Nevirapine	NNRTI	Tipranavir	PI
Lamivudine	NRTI	Rilpivirine	NNRTI	Enfuvirtide	Fusion inhibitor
Tenofovir DF	NRTI	Atazanavir	PI	Maraviroc	CCR5 antagonist
Zidovudine	NRTI	Darunavir	PI	Raltegravir	Integrase inhibitor
Doravirine	NNRTI	Fosamprenavir	PI	Dolutegravir	Integrase inhibitor
Efavirenz	NNRTI	Ritonavir	PI	Ibalizumab-uiyk	Post-attachment

Table 1. Recommended drugs for the treatment of HIV infections according to the U.S. Department of Health and Human Services. Abbreviations for drug classes: NRTI = Nucleoside Reverse Transcriptase Inhibitor; NNRTI = Non-nucleoside Reverse Transcriptase Inhibitor; PI = Protease Inhibitor

HAART Treatment

Highly active antiretroviral therapy (HAART) is a term that describes treatment of patients using a cocktail of three different drugs that act on different targets of the HIV-1 lifecycle. Prior to HAART, the common treatment of antiretroviral cases were mono or combination therapies using two different antiretroviral compounds. This form of therapy resulted in limited success in treatment of HIV and failed to suppress viral activity. With the introduction of protease inhibitors in the 1990s, it became possible to administer a combination of three separate drug agents to overcome the rapid emergence of resistance to a single agent. This new form of combination therapy for retroviruses showed efficient suppression of viral activity by stopping the lifecycle at different points. This multi-drug therapy has been coined "triple therapy" due to the three active drugs consisting of entry inhibitors, nucleoside inhibitors, non-nucleoside reverse transcriptase inhibitors, integrase inhibitors, and/or protease inhibitors[3].

HIV-1 Integrase as a Target for Drug Development

HIV-1 integrase is one of "newest" targets of the approved HIV drugs in terms of FDA approval dates. This enzyme is responsible for the critical integration of the viral DNA into the host DNA in a two-step reaction. In the first step, IN removes the two terminal nucleotides (G and T) from each 3' end of the double-stranded viral DNA. "Joining" or "strand transfer" is the second step, which consists of a S_n2 nucleophilic attack by the free 3' hydroxyl group on the viral DNA to the host cell's chromosomal DNA producing two joined molecules[4,5]. The strand transfer process is then completed by cellular DNA repair mechanisms that fills any gaps between the host DNA and the newly integrated provirus[6,7].
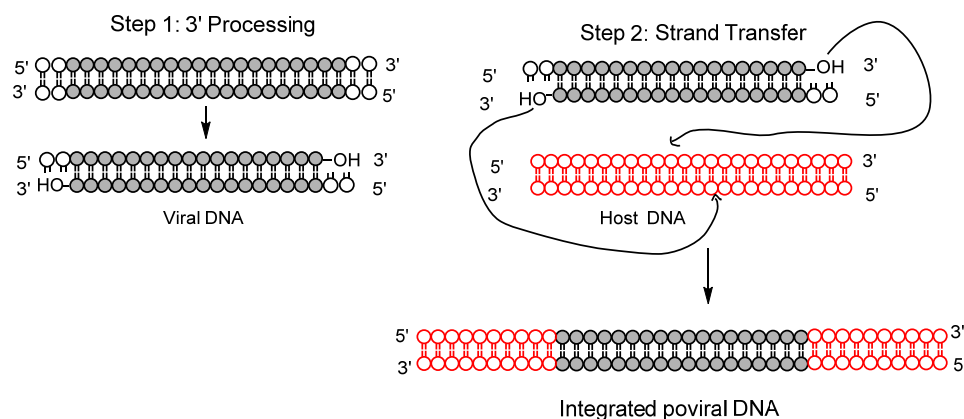
Figure 2. Mechanism of HIV-1 integrase[8]

HIV-1 IN is made up of 288 amino acid residues, which are the building blocks for the three domains: the N-terminal domain (NTD), the catalytic core domain (CCD), and the C-terminal domain (CTD). Together these three domains make up a single IN protein or monomer. In the NTD, two histidine and cysteine residues form a HHCC zinc-finger motif which chelates a zinc atom in each monomer. This motif ultimately facilitates the formation of a tetramer (or higher order oligomer) through multimerization of the IN monomers. The tetramer is believed to be required for IN activity, since it is necessary for DNA binding and the subsequent strand transfer reactions. Through this integration process, the viral DNA is successfully incorporated into the host genome, enabling the formation and maturation of additional viral particles[8].

Of particular interest for the treatment of HIV is the targeting of the viral life cycle at the integration stage. INISs are one class of HIV drugs that accomplish this through binding to the active site of IN. Raltegravir was the first member of this class to make it to the market, receiving approval in 2007. Dolutegravir was introduced six years later, but has a much higher barrier to resistance than raltegravir. Elvitegravir and Bictavir are also INIS drugs prescribed in conjunction with other treatments for the management of HIV. All of these INISs contain a

similar central pharmacophore that interacts with the two magnesium (or manganese) ions in the integrase active site. However, frequent aberrant mutations of the viral genome have begun to render the efficacy of these current treatment options sub-therapeutic, suggesting the need to develop new mechanisms of action for the inhibition of the integration state of the viral life cycle[8].
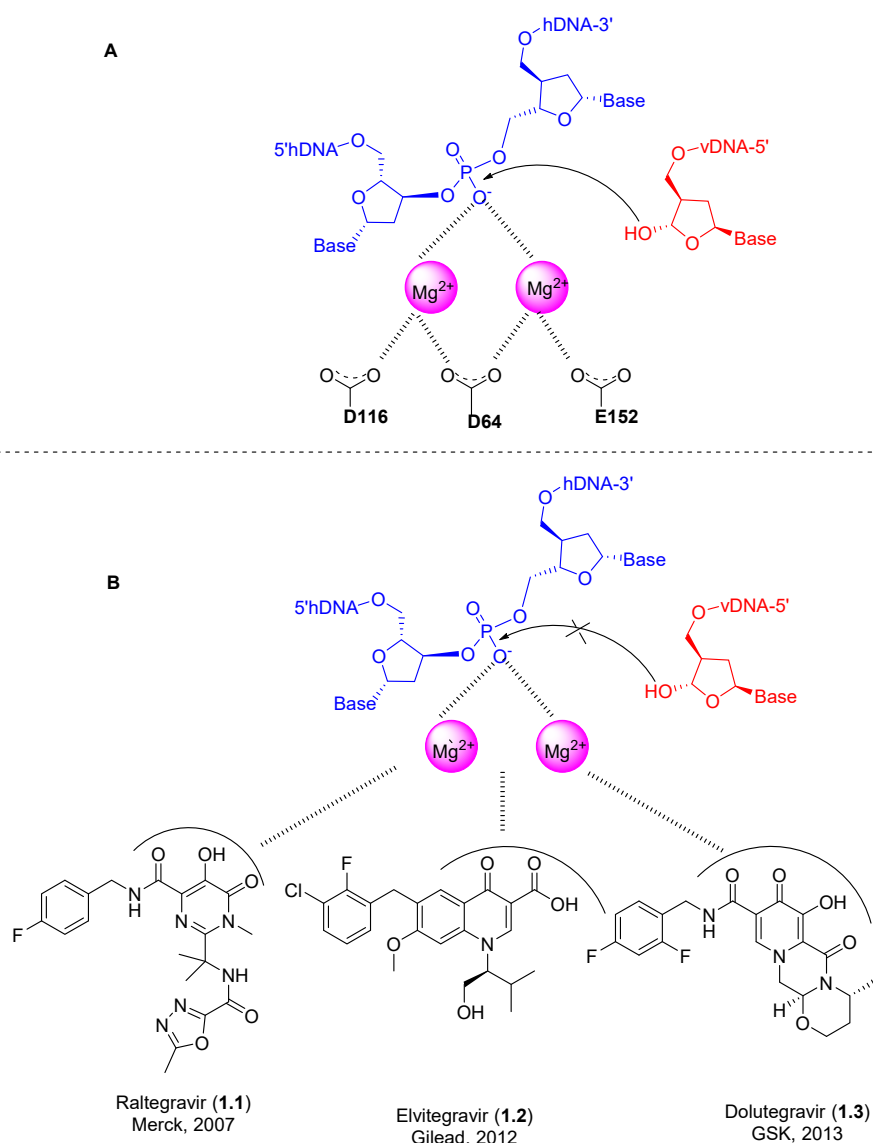

Figure 3. FDA approved IN inhibitors to the divalent Mg^{2+} in the active site and inhibition of the IN enzyme.

One form of INIS that has been studied more recently is the allosteric integrase inhibitors (ALLINIs). These compounds have been shown to bind outside of the active site of integrase in the same pocket as the natural cofactor LEDGF/p75 at the CCD dimer interface. Binding of small molecules at this allosteric site results in the inactivation of integrase due to the formation of aberrant integrase multimers. This hyper-multimerization occurs as a result of interaction of the CTD subunit of another integrase protein with the CDD dimer interface. It is hypothesized that while these small molecules do not directly compete with the natural cofactor, the induction of multimers of integrase through binding of the ALLINI after the natural ligand is no longer bound results in a malformation of the packaging of the viral load, thereby rendering the new virus noninfectious[9]. One such small molecule, BI224436 by Gilead Science, advanced to phase one clinical trials, but failed to provide the necessary safety profile for continued development. Currently, no FDA-approved small molecules for the treatment of HIV that target the allosteric site of integrase are on the market[8].

Quinoline scaffold

Previous studies had discovered that small molecules can target the LEDGF/p75 binding site and disrupt IN binding, thus halting the integration process. This led Debyser and coworkers to rationally design a high throughput screening that ultimately resulted in the discovery of a series of 2-(quinoline-3-yl) acetic acid derivatives that act as inhibitors of the LEDGF/p75-integrase interaction[10]. These screenings started with a limited set of 20,000 commercially available compounds to establish a pharmacophore. In silico screening of these compounds produced 2,000 hits after two passes. From these 2000 hits, the top 25 compounds which scored the best in this simulation were selected for biological testing resulting in one hit compound (LEDGIN-1). Multiple rounds of structure-activity refinement produced a highly potent 2-

(quinoline-3-yl) acetic acid analog (LEDGIN-6). This analog was reported to be ten times more potent (IC_{50}= 1.37 µM) against IN-LEDGF/p75 interactions in vivo while exhibiting a reported twenty-fold increase in antiviral activity (EC_{50}= 2.35 µM) over other studied quinoline acidic acids[11].

MOA of Quinoline ALLINIs

Previous members of the Fuchs lab initially synthesized several of the early quinoline-based ALLINIs to study their biological profiles. Among the compounds synthesized were LEDGIN-6 and BI-1001, both of which were previously reported by the Debyser lab and Boehringer Ingelheim, respectively[11,12]. LEDGN-6 and BI-1001 were evaluated by the Kvaratskhelia lab using homogeneous time resolved fluorescence (HTRF)-based IN-LEDGF/p75 binding assays. These studies revealed that both compounds have identical modes of actions and similar IC_{50} values when inhibiting IN-LEDGF/p75. Determination of IC_{50} values and mechanism of action used a 3' processing and strand transfer activity assay to measure LEDGF/p75 independent inhibition. The 3' processing assay showed that LEDGIN-6 and BI-1001 inhibited integrase activity with IC_{50} values of 3.9 and 2.3 µM, respectively. Stand transfer assays reported that LEDGN-6 inhibited stand transfer with an IC_{50} value of 4.2 µM and 1.7µM, respectively. The k_d values obtained from these experiments for LEDGN-6 and BI-1001 were 10.0 and 1.0 µM, respectively. These quinoline structures were also observed to promote aberrant multimerization of IN. The term "aberrant IN multimerization" refers to the formation of inactive IN oligomers that vary from the correctly assembled tetramer. LEDGIN-6 and BI-1001 induced this aberrant multimerization with IC_{50} values of 11.3 and 4.9 µM, respectively. Lastly, the anti-viral activity for LEDGN-6 and BI-1001 were 12.2 and 5.8 µM respectively[12]. In

the head-to-head comparison of these compounds, BI-1001 showed very slightly better potency

than LEDGN-6 in the various assays[8].

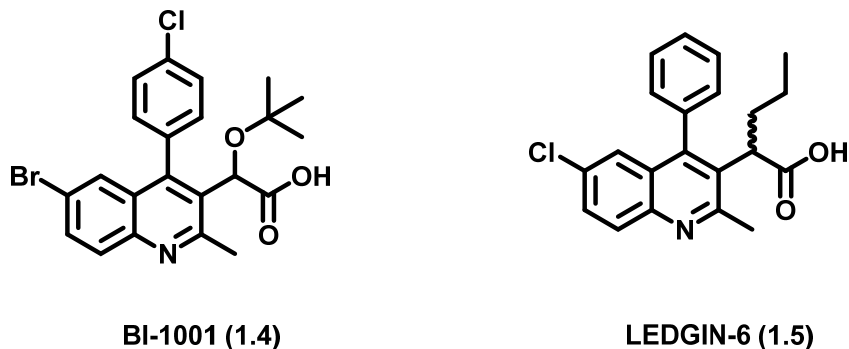**BI-1001 (1.4)**   **LEDGIN-6 (1.5)**

Figure 4. Chemical structures of BI-1001 and LEDGIN-6.

Crystal structure analysis of BI-1001 bound to the CCD dimer interface of IN indicated

that the carboxylic acid moiety of the compound hydrogen bonds with histidine-171 (H171) and

glutamate-170 (E170). Additionally, the α-methoxy group on BI-1001 shows an additional

interaction through hydrogen bonding with threonine-174 (T174) which is hypothesized to be the

reason that it shows a better potency than LEDGN-6. Both LEDGN-6 and BI-1001 express

multimodal mechanisms of action, each leading to the inhibition of LEDGF-IN binding in

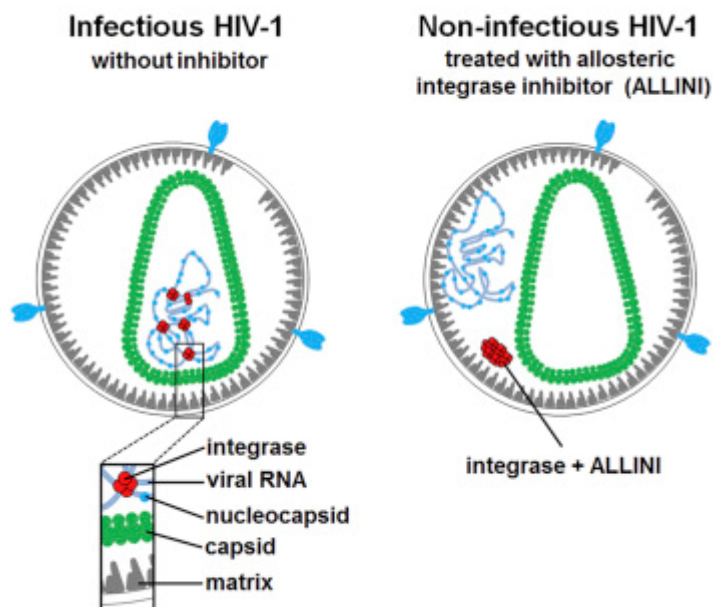LEDGF/p75 independent binding assays as well as induction of aberrant multimerization of IN[12].

Figure 5. The formation of aberrant mutlimerization of IN in the presence of ALLINIs

Resistance Mechanism

One of the main drivers of the necessity of novel targets and compounds to treat HIV is the frequent and aberrant mutations leading to resistance to many of the current compounds being used or tested as therapies. During the course of HIV drug development, resistance mutations to drugs are deliberately developed in order to study the escape mechanisms employed by the virus. Dr. Kvaratskhelia's group has studied one mutation in particular, A128T, which was developed upon treatment with these prototypical ALLINIs. A128T is located in the IN-dimer interface where LEDGF/p75 or ALLINIs would occupy. It is hypothesized that this resistance is due to a shift of the quinolone core downwards in the binding pocket due to the mutation of the alanine residue to the bulkier, more polar, threonine[9].

Scaffold hopping to indole

The A128T resistance mutation pushed the necessity to discover novel allosteric integrase inhibitors. Scaffold hopping techniques were employed to assist in developing new compounds. The term "Scaffold Hopping" refers to the method of functionalizing groups around a central core. Scaffold hopping techniques have been used for many reasons including i) replacement of a lipophilic scaffold by a more polar one, ii) substitution of a metabolically labile scaffold with a more stable or less toxic one, and iii) replacement of a very flexible scaffold with a rigid central core[8].

Scaffold hoping was employed as a part of this project to discover novel allosteric IN inhibitors using the already established quinoline-based compounds as a reference core. Numerous heterocyclic compounds have been employed or considered by the Fuchs lab as potential core structures including pyridine, thiophene, isoquinoline, and indole cores. As a part of these studies, the indole scaffold was employed as a candidate for further ALLINI development for three reasons. These reasons being: 1) a wide variety of established synthetic routes were available to ease of development of these biologically active compounds, 2) the innate reactivity of the indole core which allows ease of functional group additions, and 3) the limited patent coverage of the indole core being used a ALLINIs leaving room for the development of novel intellectual property[8, 13]. The goal using this approach was to explore the differences in binding, geometry, electronics, and hydrogen bonding effects of the electron rich indole core against the electron deficient quinoline core.

Before synthesis on the indole scaffold began, computational models developed by Guqin Shi in Dr. Chenglong Li's lab using AutoDock 4.0 were used to predict possible binding modes of the indole core. This prediction was necessary due to the difference in size between the highly

substituted five membered pyrrole-like ring of the indole as compared to the six membered

pyridine-like ring of the quinolone system. Although the indole ring system has a different shape

then the quinoline scaffold, the binding modes for these scaffolds were found to be maintained.

According to docking studies, the pharmacophoric acetic acid and aryl groups overlay in the

same areas of space as seen in quinoline binding. The major difference between the two

molecules is the scaffold positioning itself. As the acid side chain anchors the molecule in the

pocket, the smaller size of the five membered ring of the indole necessitates a slight tilt in the

opposite direction of the A128T residue which previously was the means of resistance to the

quinoline-based ALLINIs. With these results, it was established that the indole scaffold should

be pursued in the synthesis of ALLINI analogs[8].
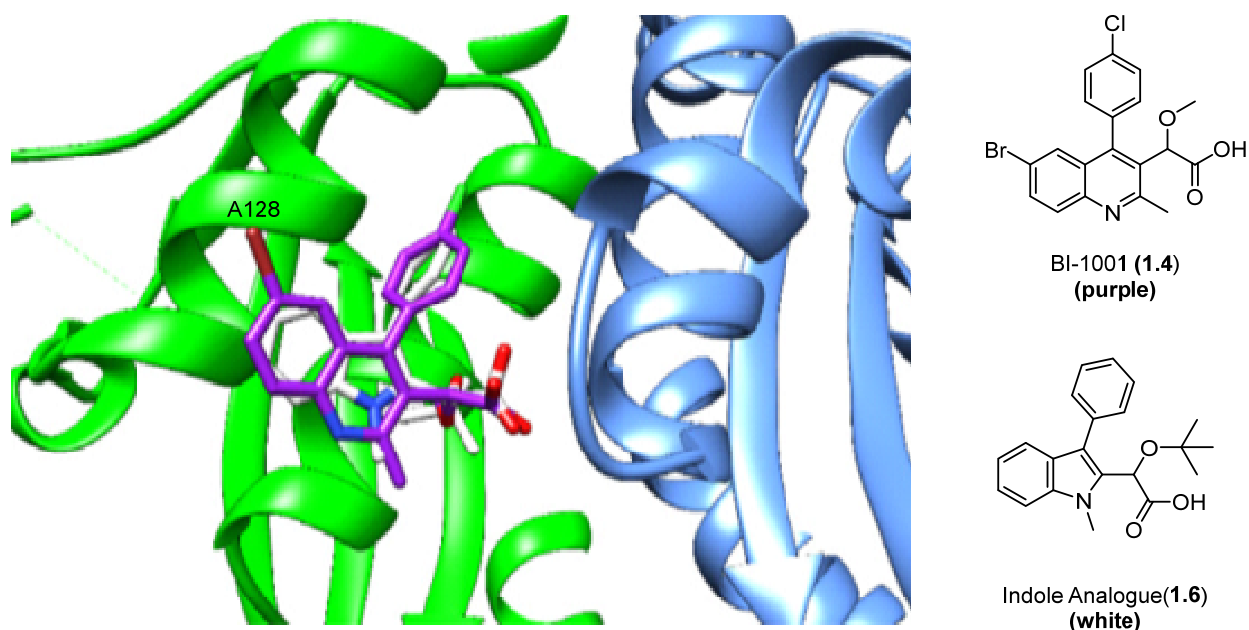
Figure 6. A docking model BI-1001 (purple) and a first-generation indole analog (white) in the

LEDGF/p75 binding pocket to the HIV-1 IN dimer interface (subunits shown in green and blue).

16

Strategies for the Optimization of the Indole Scaffold

When trying to develop a strategy for the design and synthesis of novel indole scaffolds of ALLINIs, the positions of the key pharmacophoric groups being transferred from the quinolone needed to be considered. Since the primary ring of the indole is smaller than that of the quinolone, the positioning of these substituents as well as the consideration for further synthetic manipulation was confirmed through comparison of docking models with the quinoline crystal structures. Perhaps the most important group on the indole scaffold is the carboxylic acid side chain that is attached at the C2 position of the indole nucleus (which is adjacent to the N atom on the C2 carbon of the pyrrole-like ring). This group is necessary for the direct binding to the IN protein and has been stated to be "critical for antiviral potency and that there was no tolerated isosteric replacement for the acid" (Fader et al.)[14]. With the acid in place, the only possible areas for modifications off the indole ring are the N1, C5, C6, and C3 positions (based on what is used for starting material).

3-methyl indole with necessary pharmacophore
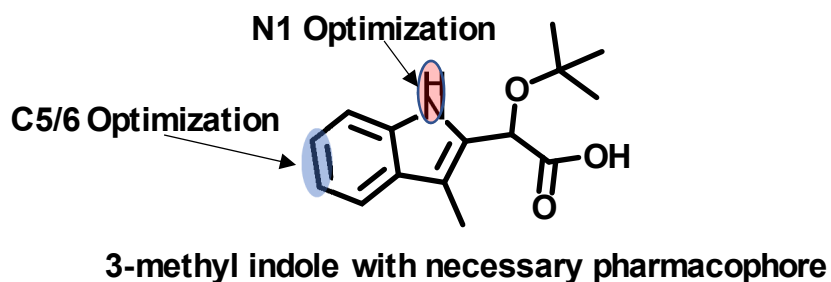
Figure 7. A structural map of sites for possible modifications on the indole scaffold

The N1 and C3 positions would initially be held mostly constant with a methyl at the C3 position and an anisole or chromane ring at N1. With this is mind, the C5 and C6 positions were targeted first for optimization. This strategy served two primary pursposes: 1) no previous findings of indole-based ALLINIs had been optimized at the C5 and C6 positions and 2) the

introduction a solvent facing group would be expected to interact with subunit 1 in the IN-CCD

dimer interface causing a disruption in protein formation. With regard to the C3 or N1 position,

depending on which starting material is used, the position would be largely unchanged from the

general hydrophobic group of an anisole or chromane group occupying the position for

hydrophobic interactions.

The synthetic route that was chosen to be used was the addition of the hydrophobic

chromane or anisole at the N1 position and optimization at the C5 position. This was chosen in

analogy to the previous studies on the novel ALLINI KF116 that showed that having the solvent

exposed groups being positioned in the "lower left quadrant" of the molecule presented a greater

binding affinity. With this, the synthetic route was optimized using the N1 group as the anchor

for the hydrophobic group for binding to the CCD dimer interface.
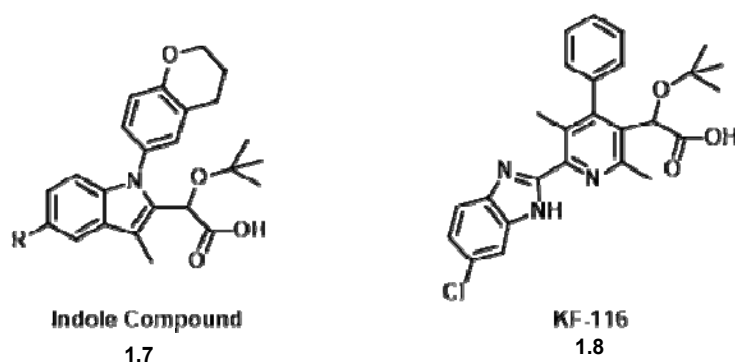
Figure 8. Comparison of the proposed indole compounds (1.7) to KF-116 (1.8).

Synthesis of C5 substituted indoles

The synthesis of the substituted indole series began with the commercially available 3-

methyl indole, which was then subjected to an Ullmann coupling[15] reaction with either an anisole

or chromane moiety. These rings served as the hydrophilic/aromatic groups that would project from the N1 position of the indole of the molecule into a narrow, capped hydrophobic channel in the dimer interface. Next the keto-ester side chain was added via the conditions shown in Figure 8. The keto-ester group would require subsequent functionalization to produce the biologically active acetic acid side chain. Once the keto-ester was established, however, the C5 position of the indole was brominated, via the conditions in Figure 8. To add the sterically bulky *t*-butyl group to the keto-ester side chain, the ketone was reduced to an alcohol using sodium borohydride, yielding an alcohol. After purification of the resulting alcohol, the t-butyl group was attached using *t*-buOAc/perchloric acid to yield the desired "Key Intermediate" product. From this intermediate, substitutions of the previously established C5 bromine using Suzuki coupling methods were employed to yield derivatives of the original indole scaffold. Upon successful coupling at the C5 position, the ester group on the C2 position could be saponified to produce the carboxylic acid that would be subjected to biological testing.
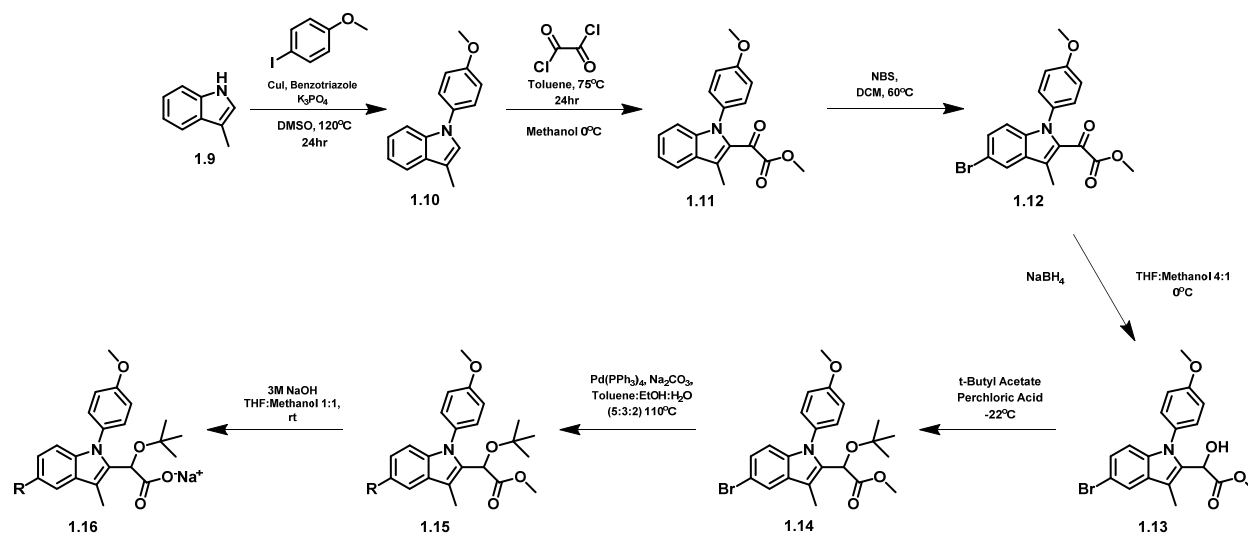
Figure 9. Synthetic route of N1-anisole-containing indole compounds devised by Dr. Janet Antwi.

The *t*-butyl ether represents a point of divergence in the synthetic route, potentially

facilitating the production of a number of diverse substrates. At this point, the brominated C5

position allows for the customization and optimization of the compound through the addition of

different functional groups via a Suzuki reaction using various boronic acid reagents. The main

goal of these newly added functional groups is to point away from the IN-CCD binding area in

hopes of picking up interactions with the incoming CTD subunit to efficiently promote

multimerization and thus effective IN inhibition.

Though the ALLINIs based on the indole scaffold are still a relatively novel potential

drug class, several compounds have been previously synthesized in the Fuchs lab by Dr. Janet

Antwi. Many of the tested compounds that were produced consisted of an anisole residing on the

N1 position of the scaffold while the carboxylic acid chain remained on the C2 position to ensure

biological activity. Different functional groups were added via Suzuki coupling reactions and

biological data of some her most potent compounds can been seen in Table (2) along with

structures of each respective compound.

Dr. Antwi's compounds were tested in the Kvaratshelia lab using a LEDGF/p75

independent assay. For clarification, this particular assay was run without LEDGF/p75, while

dependent assays are run in the presence of LEDGF/p75. The assay itself was designed to mimic

the stage in the HIV-1 where LEDGF/p75 is no longer present after the integration phase has

already finished. Instead of measuring the direct inhibition of the ALLINI against LEGDF/p75,

the assay determines the direct effect of the drug on IN, ultimately leading to aberrant multimer

formation of the IN enzyme. These aberrant multimers of IN are the inactive form of the IN

enzyme and can no longer continue the life cycle of the virus.

My initial work in the Fuchs lab focused on the synthesis of the bromide in an effort to prepare novel ALLINIs. Utilizing the methods established by Dr. Antwi, I was able to repeat and troubleshoot these synthetic steps to produce the desired product. With this intermediate in hand, it was coupled to a thiphene to produce a new ALLINI compound. Upon purification and characterization of the compound, it was submitted to Dr. Ross Larue for biological evaluation.
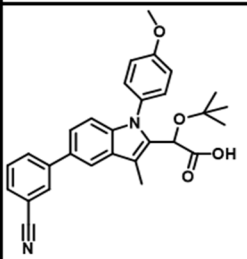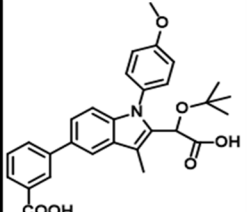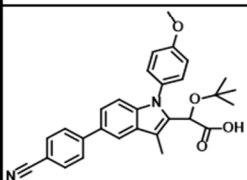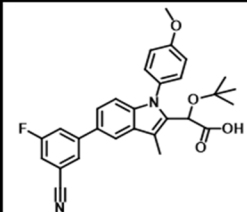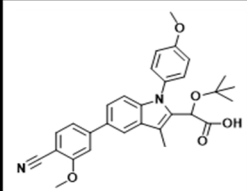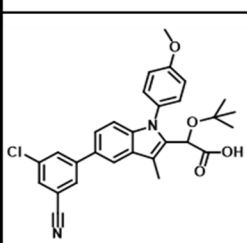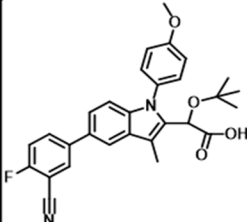
Compound (phenyl indole)	$IC_{50}(\mu M)$	Structure
3-Cyano-	1.17	
3-carboxylate-	1.10	
4-cyano-	1.15	
3-cyano-5-fluoro-	0.83	
4-cyano-3-methoxy-	1.16	
3-chloro-5-cyano-	1.80	
3-cyano-4-fluoro-	2.04	

Table 2. Dr. Antwi's biological compounds

Chromane-indole Synthesis

In addition to the introduction of the anisole ring on the N1 position, other ring systems, including the chromane system, have been shown to possess even greater potency against IN in other series of ALLINI analogues. The attachment of the anisole itself was initially chosen for the known hydrophobic characteristics that it possesses, but also for the ease of synthesis of the anisole-indole intermediate due to iodo-anisole (a reagent for Ullman coupling of the anisole to the 3-methyl indole) is commercially available. It is hypothesized that substituting the anisole ring with a more hydrophobic functional group would increase the binding affinity into the allosteric site of IN. The chromane derivatives, therefore, were expected to provide better potency, but the chemistry necessary for their introduction was not as well established.
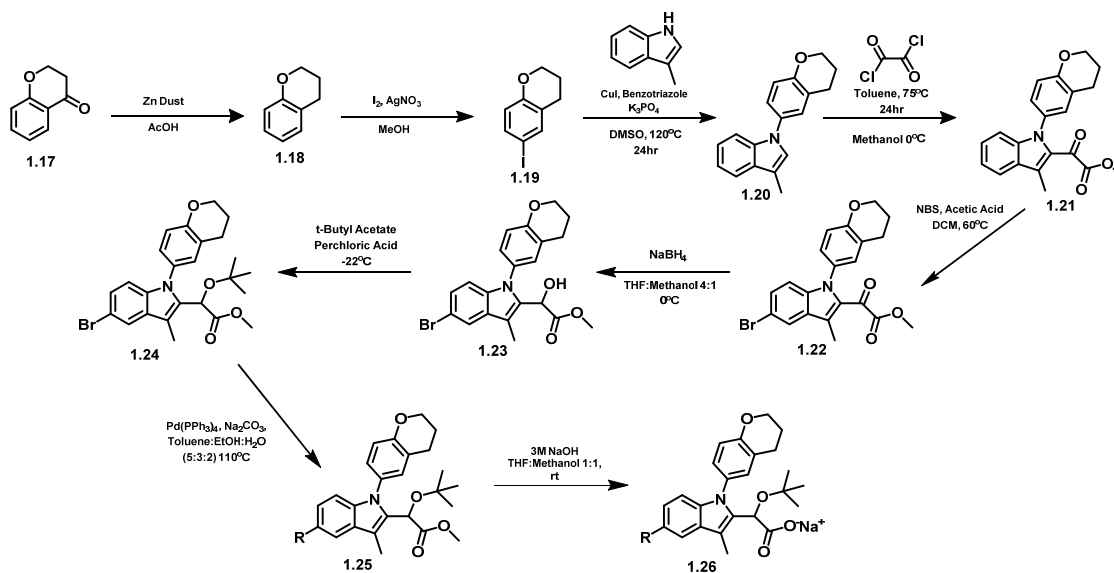
Figure 10. Synthetic scheme of chromane substituted indole based on anisole indole synthesis

As previously mentioned, the reagent, 4-iodo-chromane, to synthesize the chromane indole intermediate is not commercially available. This inconvenience required the synthesis of the necessary starting material in a two-step reaction. 4-chromanone was reduced using Zn dust

in acetic acid yielding chromane. The chromane was then subjected to a halogenation reaction in the presence of I_2 and $AgNO_3$ yielding 4-iodo-chromane. 4-iodo-chromane was subjected to an Ullmann coupling reaction in the presence of 3-methyl-indole to yield the chromane substituted indole on the N1 position. To establish the keto-ester on the C2 of the scaffold, the conditions represented in figure (9). To establish the *t*-butyl ester, the keto-ester was reduced using sodium borohydride yielding the alcohol. Finally, treatment with *t*-BuOAc/perchloric acid provided the desired "Key Intermediate" product.

Having successfully completed the synthesis of an anisole derivative, I was also able to carry out the synthesis of the more synthetically challenging chromane. This synthesis was important not only to generate more of the key intermediate for analogue synthesis, but also to produce all of the synthetic intermediates in sufficient quantity and purity for chemical characterization. With this synthesis in hand, therefore, all of the intermediates in this sequence were characterized using ^{1}H NMR, [13]NMR, and HRMS. The bromide intermediate was also saponified to generate a new analogue for testing. This compound has also been submitted to Dr. Larue for testing. Subsequent efforts will be focused on the derivatization of this bromide in an effort to introduce substituted aromatic rings onto this chromane-containing indole system. It is expected that the introduction of the most potent aromatic systems in Dr. Antwi's studies with the anisole system into the chromane series will produce compounds with higher potency for integrase inhibition. These studies are currently ongoing.

Chapter 2.

Experimentals
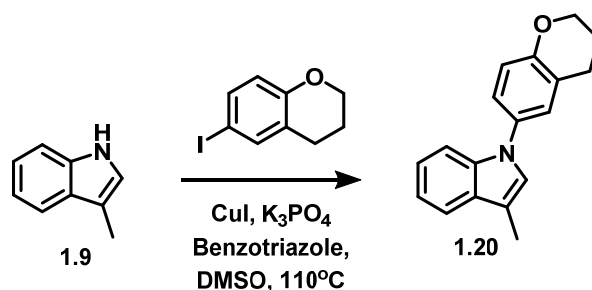
1.20. To a solution of 3-methyl indole (306.5 mg, 2.34 mmol) in DMSO (0.25 M), was added

iodochromane (650 mg, 2.5 mmol), CuI (44mg, 0.231 mmol), K_3PO_4 (989.18 mg, 4.66 mmol),

and benzotriazle (111.38 mg, 0.935 mmol). The mixture was heated to 120 °C and stirred

overnight. Upon completion, the reaction mixture was quenched with water and extracted with

ethyl acetate (3x) and evaporated under reduced pressure. The crude mixture was washed with

water, brine, dried over sodium sulfate and concentrated under educed pressure. Flash

chromatography (silica gel, 1-2% ethyl acetate in hexanes) yielded the coupled product as a

bright pink oil. ^{1}H NMR (400 MHz, CDCl$_3$). δ 7.67 (d, J = 7.5 Hz, 1H), 7.51 (d, J = 7.5 Hz, 1H),

7.32-7.17 (m, 4H), 7.10 (s,1H), 6.95 (d, J = 8.7 Hz, 1H), 4.32-4.28 (m, 2H), 2.88 (t, J = 6.5 Hz,

2H), 2.44 (s, 3H), 2.06-2.12 (m, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 153.34, 136.47, 132.49,

129.36, 125.96, 125.77, 123.64, 123.14, 122.08, 119.41, 119.09, 117.54, 111.94, 110.34, 66.63,

31.67, 25.06, 22.25, 9.63.
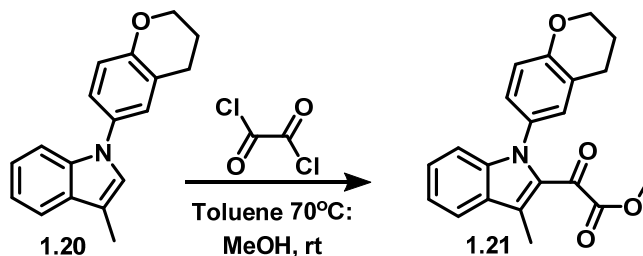
1.21. To a solution of indole-chromane (598 mg, 2.27 mmol) in toluene (11.4 ml, 0.20 M) was

added oxalyl chloride (1.8ml, 20.4mmol). The resulting mixture was heated to 75 °C and stirred

overnight. Upon completion, the mixture was cooled to 0 °C and methanol (34ml, 60% M) was slowly added while stirring. The solution was then concentrated under reduced pressure and taken back up in EtOAc. The organic layer was washed with sat. NaHCO$_3$ solution (2x). The combined aqueous layers were washed with EtOAc. The combined organic layers were vigorously washed with water (5x), brine (1x), dried over sodium sulfate, and concentrated under reduced pressure. Flash chromatography (silica gel, 10% EtOAc in hexanes) yielded product as yellow oil. ^{1}H NMR (400 MHz, CDCl$_3$). δ 7.77 (d, J = 8.1 Hz, 1H), 7.37 (t, J = 7.7 Hz, 1H), 7.21 (d, J = 7.3 Hz, 1H), 7.18 (d, J = 8.5 Hz, 1H), 7.10 (dd, J = 8.6 Hz, 1H), 6.99 (d, J = 2.5 Hz, 1H), 6.92 (d, J = 8.6, 1H), 4.30-4.21 (m, 2H), 3.56 (s, 3H), 2.81 (t, J = 6.5 Hz, 2H), 2.64 (s, 3H), 2.04 (qd, J = 6.3, 4.4 Hz, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 179.72, 164.38, 154.63, 140.46,130.58, 129.58, 129.35, 127.61, 127.38, 126.91, 126.29, 123.03, 121.41, 121.11, 117.46, 111.53, 66.66, 52.42, 24.83, 22.07, 10.37.
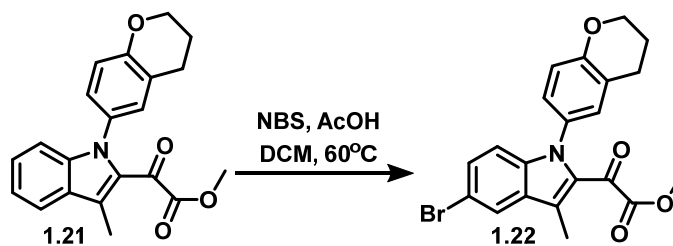
1.22. To a solution of indole-chromane keto-ester (156 mg, 0.45 mmol) in DCM (4.5ml, 0.10 M) was added freshly recrystallized NBS (87.4 mg, 0.5 mmol) and acetic acid (28 μL, 0.5 mmol). The mixture was stirred at 60 °C for 30 minutes. Upon completion, the reaction was quenched with water and extracted with DCM (3x). The combined layers were washed with water, brine, dried over sodium sulfate, and concentrated under reduced pressure. Flash chromatography (silica gel, 10% EtOAc, 15% DCM, in hexanes) afforded the brominated product as a yellow oil. ^{1}H NMR (400 MHz, CDCl$_3$). δ 7.87 (dd, J = 1.9, 0.7 Hz, 1H), 7.40 (dd, J = 8.9, 1.9 Hz, 1H),

7.06 – 7.02 (m, 2H), 6.94 (d, J = 2.7 Hz, 1H), 6.89 (d, J = 8.6 Hz, 1H), 4.28 – 4.22 (m, 2H), 3.55 (s, 3H), 2.81 (t, J = 6.4 Hz, 2H), 2.58 (s, 3H), 2.04 (p, J = 6.2 Hz, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 179.64, 164.00, 154.91, 138.81, 131.30, 130.37, 129.19, 129.02, 128.87, 126.76, 124.86, 123.76, 123.16, 117.58, 114.24, 113.13.
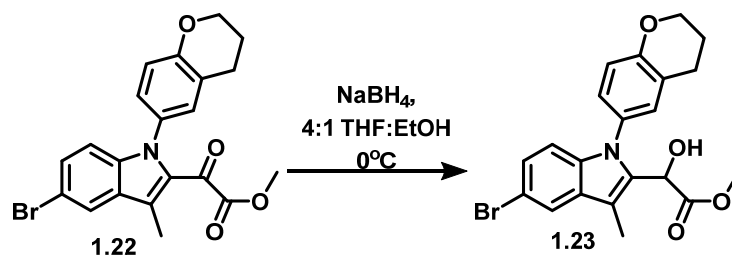
1.23. Brominated chromane-indole (122 mg, 0.29 mmol) was taken up in a 4:1 mixture of THF:EtOH (0.1 M total) and cooled to 0 °C. NaBH$_4$ (11 mg, 0.29 mmol) was added slowly added and stirred at 0 °C for 1h. The reaction was quenched with a slow addition of water. The aqueous solution was extracted with EtOAc (3x). The combined organic layers were washed with brine, dried over sodium sulfate, and concentrated under reduced pressure. Flash chromatography (silica gel, 20% EtOAc, in hexanes) afforded the reduced product. ^{1}H NMR (400 MHz, CDCl$_3$). δ 7.72 (d, J = 1.9 Hz, 1H), 7.22 (dd, J = 8.7, 1.9 Hz, 1H), 7.14 (dt, J = 8.3, 2.1 Hz, 1H), 6.97 (d, J = 2.4 Hz, 1H), 6.90 (dd, J = 9.5, 5.7 Hz, 2H), 5.25 (dd, J = 13.9, 3.4 Hz, 1H), 4.26 (dd, J = 5.9, 4.3 Hz, 2H), 3.69 (d, J = 12.2 Hz, 3H), 3.31 (t, J = 5.0 Hz, 1H), 2.82 (q, J = 6.8 Hz, 2H), 2.31 (d, J = 6.0 Hz, 3H), 2.05 (dq, J = 9.3, 5.9, 4.3 Hz, 2H).
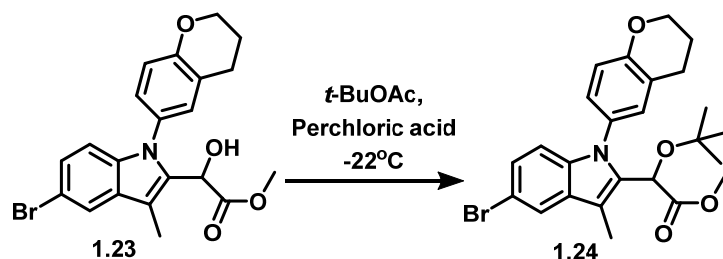
1.24. Hydroxyester chromane-indole (70.6 mg, 0.164 mmol) was dissolved in *t*-BuOAc (3.28 ml, 0.05 M) and cooled to -22 °C. Perchloric acid (0.38 ml, 0.437 M) was added to the solution and was stirred. Every 20 minutes the solution was monitored by TLC for the presence of an undesired byproduct. Upon the formation of the undesired product, the solution was quenched with Na_2CO_3. The aqueous layer was extracted with EtOAc (3x), washed with brine, dried over sodium sulfate, and concentrated under reduced pressure. The crude material was purified by flash column chromatography (silica gel, 10% EtOAc, in hexanes) to afford *t*-butyl ether. ^{1}H NMR (400 MHz, CDCl3). δ 7.70 (d, *J* = 1.9 Hz, 1H), 7.18 (dt, *J* = 9.4, 2.7 Hz, 2H), 7.00 – 6.84 (m, 3H), 5.08 (s, 1H), 4.27 (t, *J* = 5.2 Hz, 2H), 3.64 (d, *J* = 1.8 Hz, 3H), 2.84 (q, *J* = 6.6 Hz, 2H), 2.38 (d, *J* = 2.4 Hz, 3H), 2.07 (dt, *J* = 8.5, 3.7 Hz, 2H), 1.08 (d, *J* = 4.9 Hz, 9H). ^{13}C NMR (101 MHz, CDCl3) δ 171.52, 154.95, 136.69, 134.78, 131.16, 130.09, 129.78, 129.01, 128.55, 127.36, 124.99, 122.88, 121.39, 117.32, 112.77112.11, 110.44, 75.79, 66.68, 66.51, 52.21, 27.90, 24.84, 22.08, 8.72.
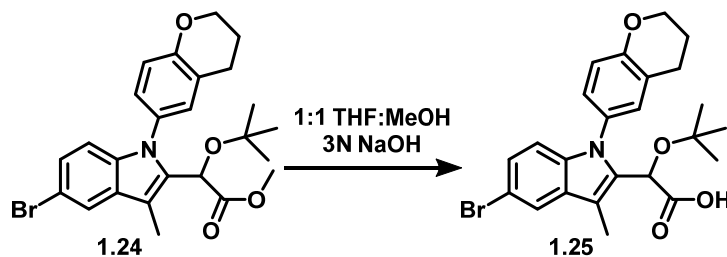
1.25. To a solution of chromane-indole (1 equiv) in a 1:1 mixture of THF (0.1 M) and MeOH (0.1 M) was added 3N NaOH solution (5 equiv). The resulting solution was then stirred at rt overnight until completion. The reaction was quenched with water and 2N HCl was slowly added to reach a pH of ≈ 4. The aqueous layer was then extracted with EtOAc (3x). The combined organic layers were washed with water, brine, dried over sodium sulfate, and concentrated under reduced pressure. The resulting acid was triturated with hexanes to afford the

desired product. ^{1}H NMR (400 MHz, CDCl$_3$). δ 7.70 (d, J = 1.9 Hz, 1H), 7.28 – 7.18 (m, 2H), 7.01 (dd, J = 8.5, 2.6 Hz, 1H), 6.96 – 6.85 (m, 2H), 5.16 (d, J = 22.6 Hz, 1H), 4.26 (q, J = 4.2, 3.7 Hz, 2H), 2.81 (dt, J = 15.5, 6.5 Hz, 2H), 2.36 (d, J = 6.0 Hz, 3H), 2.06 (d, J = 12.3 Hz, 2H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 171.09, 155.07, 136.96, 136.79, 133.68, 131.22, 129.83, 129.73, 129.59, 128.77, 128.72, 128.60, 127.29, 125.48, 125.40, 123.30, 123.03, 121.50, 121.46, 117.46, 117.36, 112.99, 112.92, 112.15, 112.12, 110.77, 110.50, 66.73, 66.21,65.99, 28.02. 24.94, 24.83, 22.02, 8.89, 8.88.

3. References

1. Content Source: HIV.govDate last updated: July 31, 2019. (2019, October 2). Global Statistics. Retrieved from https://www.hiv.gov/hiv-basics/overview/data-and-trends/global-statistics.

2. Laskey, S. B., & Siliciano, R. F. (2014, September 29). A mechanistic theory to explain the efficacy of antiretroviral therapy. Retrieved from https://www.nature.com/articles/nrmicro3351.

3. Cichocki, M. (2019, November 22). How HAART (Highly Active Antiretroviral Therapy) Works. Retrieved from https://www.verywellhealth.com/haart-highly-active-antiretroviral-therapy-48967.

4. Savarino, A. In-Silico Docking of HIV-1 Integrase Inhibitors Reveals a Novel Drug Type Acting on an Enzyme/DNA Reaction Intermediate. *Retrovirology* **2007**, *4* (1), 21.

5. Grawenhoff, J.; Engelman, A. N. Retroviral Integrase Protein and Intasome Nucleoprotein Complex Structures. *World J. Biol. Chem.* **2017**, *8* (1), 32–44.

6. Krishnan, L.; Engelman, A. Retroviral Integrase Proteins and HIV-1 DNA Integration. *J. Biol. Chem.* **2012**, *287* (49), 40858–40866.

7. Karn, J.; Stoltzfus, C. M. Transcriptional and Posttranscriptional Regulation of HIV-1 Gene Expression. *Cold Spring Harb. Perspect. Med.* **2012**, *2* (2), a006916–a006916.

8. Antwi, J. (2017). *THE DESIGN, SYNTHESIS AND OPTIMIZATION OF ALLOSTERIC HIV-1 INTEGRASE INHIBITORS* (Doctoral dissertation). The Ohio State University, Columbus

9. Feng, L., Sharma, A., Slaughter, A., Jena, N., Koh, Y., Shkriabai, N., … Kvaratskhelia, M. (2013, May 31). The A128T resistance mutation reveals aberrant protein

multimerization as the primary mechanism of action of allosteric HIV-1 integrase inhibitors. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3668738/

10. De Luca, L.; Barreca, M. L.; Ferro, S.; Christ, F.; Iraci, N.; Gitto, R.; Monforte, A. M.; Debyser, Z.; Chimirri, A. Pharmacophore-Based Discovery of Small-Molecule Inhibitors of Protein-Protein Interactions between HIV-1 Integrase and Cellular Cofactor LEDGF/p75. *ChemMedChem* **2009**, *4* (8), 1311–1316.

11. Christ, F.; Voet, A.; Marchand, A.; Nicolet, S.; Desimmie, B. A.; Marchand, D.; Bardiot, D.; Van der Veken, N. J.; Van Remoortel, B.; Strelkov, S. V.; et al. Rational Design of Small-Molecule Inhibitors of the LEDGF/p75-Integrase Interaction and HIV

12. Kessl, J. J.; Jena, N.; Koh, Y.; Taskent-Sezgin, H.; Slaughter, A.; Feng, L.; de Silva, S.; Wu, L.; Le Grice, S. F. J.; Engelman, A.; et al. Multimode, Cooperative Mechanism of Action of Allosteric HIV-1 Integrase Inhibitors. *J. Biol. Chem.* **2012**, *287* (20), 16801–16811.

13. Taber, D. F.; Tirunahari, P. K. Indole Synthesis: A Review and Proposed Classification. *Tetrahedron* **2011**, *67* (38), 7195–7210.

14. Fader, L. D.; Carson, R.; Morin, S.; Bilodeau, F.; Chabot, C.; Halmos, T.; Bailey, M. D.; Kawai, S. H.; Coulombe, R.; Laplante, S.; et al. Minimizing the Contribution of Enterohepatic Recirculation to Clearance in Rat for the NCINI Class of Inhibitors of HIV. *ACS Med. Chem. Lett.* **2014**, *5* (6), 711–716.

15. Joshi, M.; Tiwari, R.; Verma, A. K. Regioselective Preferential Nucleophilic Addition of *N* -Heterocycles onto Haloarylalkynes over *N* -Arylation of Aryl Halides. *Org. Lett.* **2012**, *14* (4), 1106–1109.

Appendix: Characterization Data of Selected Compounds
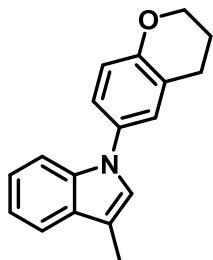
¹H NMR (400 MHz, CDCl₃)
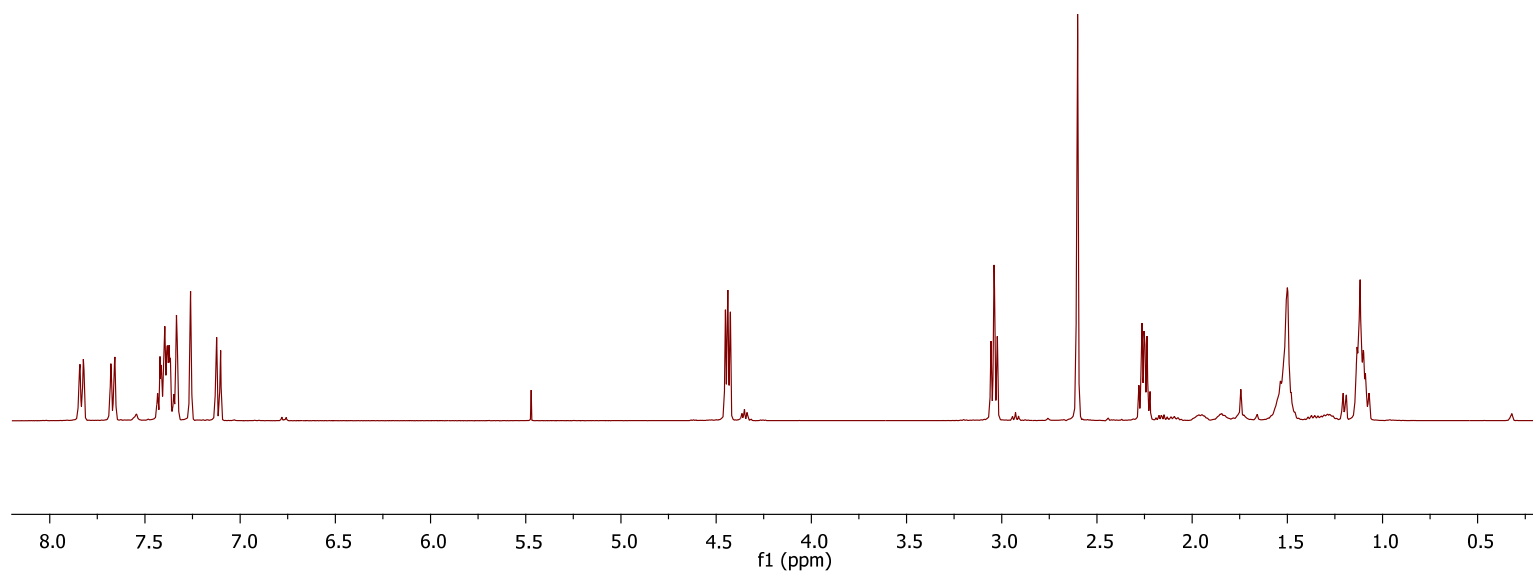
Sample is contaminated with residual 4-iodochromane, DCM, ethyl acetate, and hexanes that will be removed next step
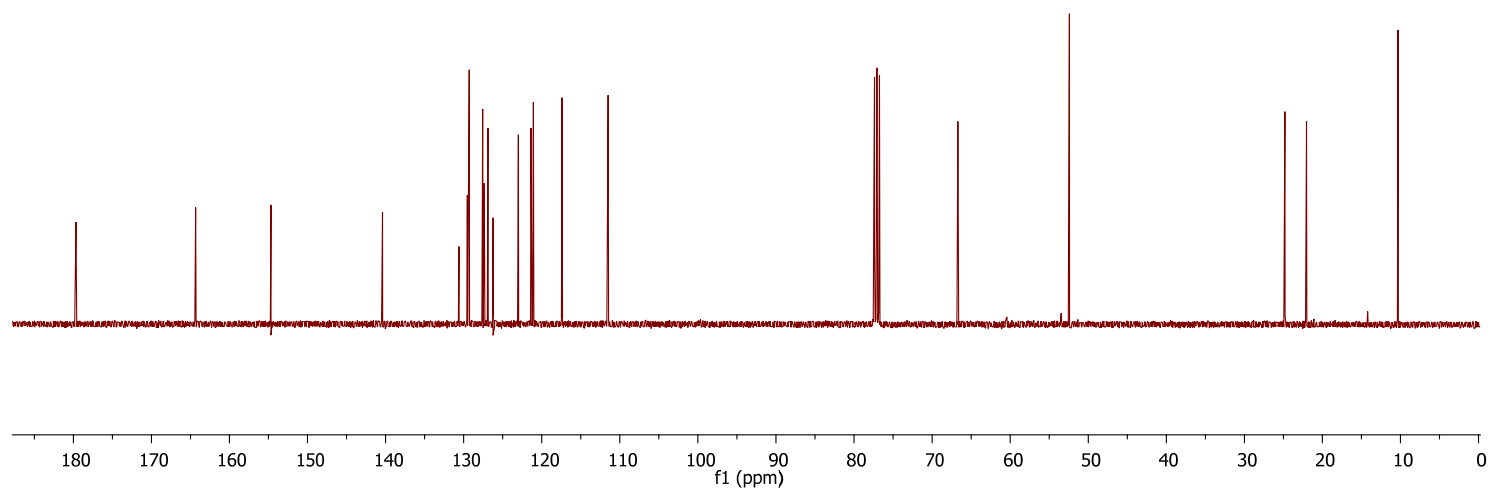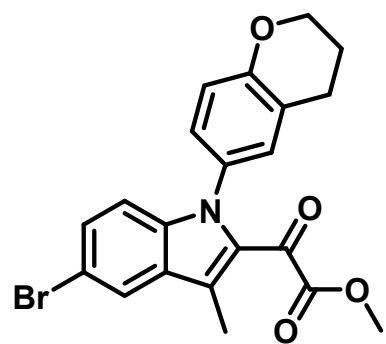
¹³C NMR (400 MHz, CDCl₃)

^{13}C NMR (400 MHz, CDCl$_3$)

150   140   130   120   110   100   90   80   70   60   50   40   30   20   10   0
f1 (ppm)

36

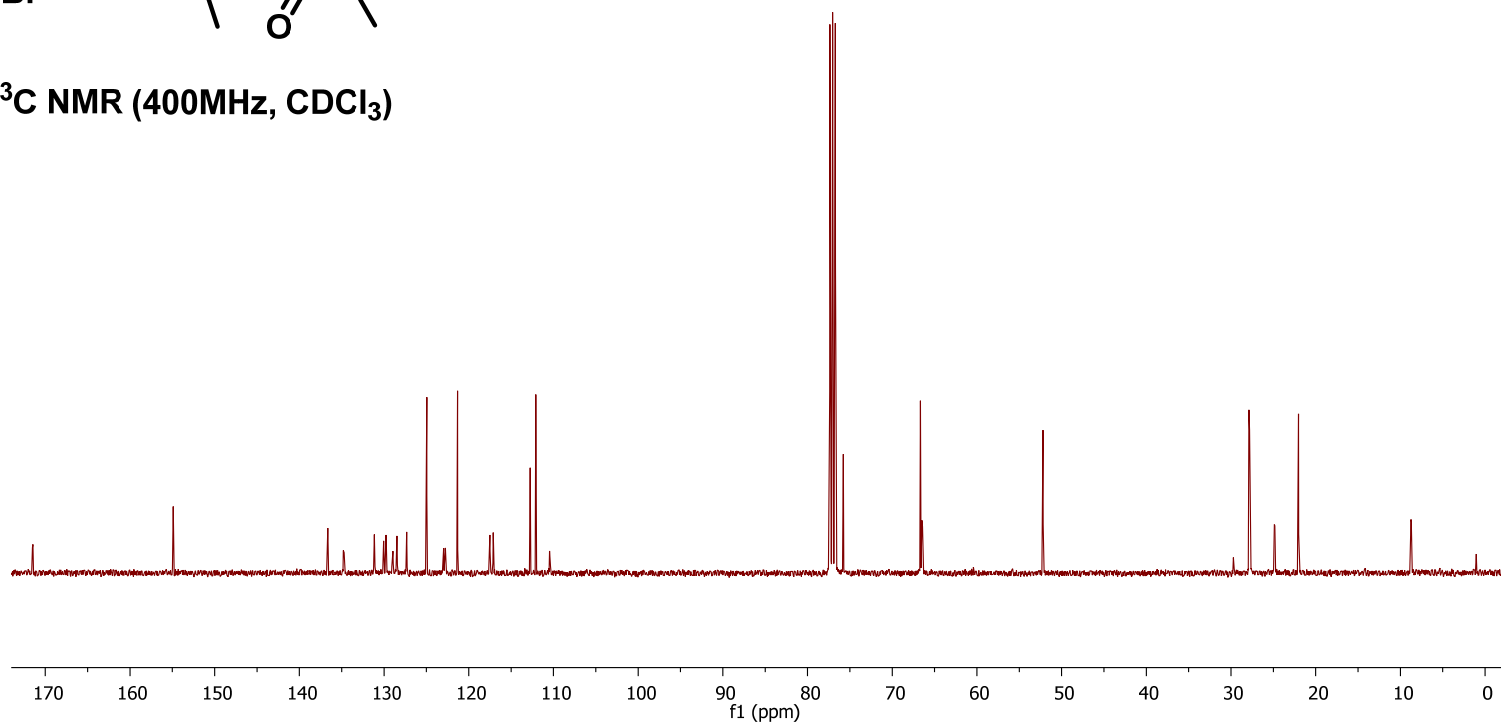**¹H NMR (400MHz, CDCl₃)**

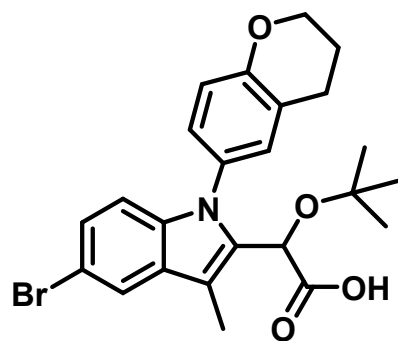^{13}C NMR (400MHz, CDCl$_3$)

^{1}H NMR (400MHz, CDCl$_3$)

¹³C NMR (400MHz, CDCl₃)

¹H NMR (400MHz, CDCl₃)

¹H NMR (400MHz, CDCl₃)

^{13}C NMR (400MHz, CDCl$_3$)

^{1}H NMR (400MHz, CDCl$_3$)

^{13}C NMR (400MHz, CDCl$_3$)

45