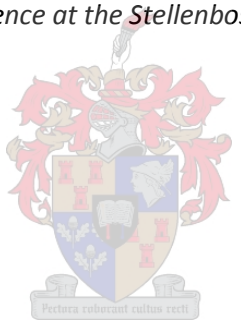
LEAD OPTIMISATION OF AN INDOLE BASED HIV-1 NON-NUCLEOSIDE REVERSE TRANSCRIPTASE INHIBITOR

---

by

Siobhan Ernan Brigg

Thesis presented in fulfilment of the requirements for the degree of Master of Science in the Faculty of Science at the Stellenbosch University



Supervisor: Dr Stephen C. Pelly

Co-supervisor: Dr Margaret A. L. Blackie

December 2017

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the authorship owner thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2017

Abstract

HIV-1 remains the worst pandemic faced by mankind since its discovery as the causative agent of AIDS in the early 1980s. An enormous amount of research has been done to find a cure, but to date there has been no success and resistance is widespread among the available treatment. This project focused on the development of novel non-nucleoside reverse transcriptase inhibitors (NNRTIs) using a rational design approach. The lead compound, ethyl 5-chloro-3-(methoxy(phenyl)methyl)-*1H*-indole-2-carboxylate, was shown to have low nano-molar potency against HIV-1 (IC_{50} = 16 nM), however it had two main shortcomings which needed to be addressed; poor resistance profile and poor acid stability. Previous research had shown the resistance profile could be improved by introducing meta substitution on the phenyl moiety which interacts with Tyr181 of the NNRTI binding pocket (NNIBP). We were successful in synthesising several meta substituted phenyl derivatives of the lead compound and these were shown to be equally as potent as the lead compound. Their activity against resistant strains is yet to be determined as we are awaiting the results from biological testing. The presence of an acid labile methyl ether functionality on the molecule which was susceptible to an acid catalysed indole mediated S_N1 substitution in aqueous acidic medium meant that the lead compound could never be considered as a candidate for an orally available drug. The methyl ether moiety was exchanged for a sulfide moiety and several of these derivatives were successfully synthesised. Acid stability tests showed that we were successful in our endeavour to improve the acid stability, offering an advantage over the lead compound despite a slight reduction in potency. However to completely eliminate the possibility of substitution, we replaced the methyl ether moiety for an ethyl group, successfully synthesising ethyl 5-chloro-3-(1-phenylpropyl)-*1H*-indole-2-carboxylate and 5-chloro-3-(1-phenylpropyl)-*1H*-indole-2-carboxamide and we are currently awaiting the results from biological testing to determine whether this derivative is active against HIV-1. The functionality in the 2-position of the indole was also investigated through the synthesis of 5-chloro-3-(methoxy(phenyl)methyl)-*1H*-indole and 5-chloro-3-((methylthio)(phenyl)methyl)-*1H*-indole. These derivatives lacking a group in the 2-position of the indole showed significant reduction in potency. Replacement of the ethyl ester for an isobutyl ester to give isobutyl 5-chloro-3-((3,5-dimethylphenyl)(methylthio)methyl)-*1H*-indole-2-carboxylate, showed some maintenance of potency, however the larger side chain was not well accommodated in the NNIBP.

The presence of a chiral centre on the lead compound, and all derivatives synthesised in the project, resulted in our final aim; we set out to develop a method for resolving these enantiomers. Unfortunately, although we employed a variety of different strategies, including the use of chiral auxiliaries and the classical resolution method of attempting to make diastereomeric salts, we were not successful in achieving this aim.

Uittreksel

MIV-1 bly die ergste pandemie wat die mensdom sedert die ontdekking van die oorsaak van VIGS in die vroeë 1980s. ‘n Enorme hoeveelheid navorsing is al onderneem om ‘n kuur te ontdek, maar daar was tot op hede geen sukses nie. Hierdie projek het gefokus op die ontwikkeling van nuwe nie-nukleosied trutranskriptase inhibeerders (NNRTIs) met gebruik van ‘n rasioniële ontwerpbenadering. Die loodverbinding, etiel-5-chloor-3- (metoksiel (fenyl) metiel) -1H-indool-2-karboksilaat, het getoon dat dit 'n lae nano-molêre sterkte teen MIV-1 (IC50 = 16 nM) het, maar dit het twee hoof tekortkominge wat aangespreek moes word; swak weerstandsprofiel en swak suur stabiliteit. Vorige navorsing het getoon dat die weerstandsprofiel verbeter kan word deur die meta-substitusie op die fenielgroep in te voer wat met Tyr181 van die NNRTI-bindingsak (NNIBP) in wisselwerking tree. Ons was suksesvol om verskeie meta-gesubstitueerde fenielederivate van die hoofverbinding te sintetiseer en dit is getoon dat dit ewe sterk as die hoofverbinding is. Hul aktiwiteit teen weerstandbiedende stamme moet nog nie bepaal word nie, aangesien ons die resultate van biologiese toetse afwag. Die teenwoordigheid van 'n suur labiele metiel eter funksie op die molekule wat vatbaar was vir 'n suur gekataliseerde indool gemedieerde SN1 substitusie in waterige suur medium beteken dat die lood verbinding nooit beskou kan word as 'n kandidaat vir 'n mondelinge beskikbare geneesmiddel. Die metiel-etergroep is vir 'n sulfieddeel uitgeruil en verskeie van hierdie afgeleides is suksesvol gesintetiseer. Suur stabiliteit toetse het getoon dat ons suksesvol was in ons strewe om die suur stabiliteit te verbeter, wat 'n voorsprong bo die hoofverbindings bied ten spyte van 'n effense vermindering in sterkte. Om die moontlikheid van vervanging egter heeltemal uit te skakel, het ons die metiel-etergroep vervang vir 'n etielgroep, met die suksesvolle sintetisering van etiel 5-chloor-3- (1-fenylpropyl) -*1H*-indool-2-karboksilaat en 5-chloor-3- 1-fenylpropyl) -1H-indool-2-karboksamied en wag tans op die resultate van biologiese toetsing om vas te stel of hierdie afgeleide teen MIV-1 aktief is. Die funksionaliteit in die 2-posisie van die indool is ook ondersoek deur die sintese van 5-chloor-3- (metoksiel (fenyl) metiel) -*1H*-indool en 5-chloor-3 -((metielthio) (feniel) metiel) -1H-indool. Hierdie afgeleides wat 'n groep in die 2-posisie van die indool het, het aansienlike vermindering in sterkte getoon. Vervanging van die etiel ester vir 'n isobutiel ester om isobutiel 5-chloor-3 - ((3,5-dimetylfenyl) (metielthio) metiel) -*1H*-indool-2-karboksilaat te gee, het 'n mate van instandhouding van sterkte getoon, maar die groter kant ketting is nie goed geakkommodeer in die NNIBP nie. Die

teenwoordigheid van 'n chirale sentrum op die loodverband, en al die afgeleides wat in die projek gesintetiseer is, het tot ons finale doel gelei; Ons het 'n metode ontwikkel om hierdie enantiomere op te los. Ongelukkig het ons ongelukkig nie geslaag om hierdie doelwit te bereik nie, alhoewel ons 'n verskeidenheid verskillende strategieë gehad het, insluitende die gebruik van chirale hulpmiddels en die klassieke resolusie metode om diastereomere soute te maak.

Acknowledgements

I have so many people to thank for helping me put this MSc project together. First and foremost, my supervisor, Dr Stephen Pelly, who has been an amazing teacher to me and has made me feel competent and confident in the lab. His passion for this project made it very enjoyable for me throughout, and I am grateful to him for building my interest in medicinal chemistry over the last few years and I know this will shape my decisions for my future in a big way. My co-supervisor, Dr Blackie for keeping an eye on me when Steve left for greener pastures. Prof van Otterlo and Dr Arnott for always having great advice and giving me a different perspective when I was stuck with a problem.

I would also like to thank my lab mates and members of the GOMOC team. You have all been so helpful and I always felt like I had someone to turn to when the synthesis wasn't working out for me. I would like to personally thank Tanya for her endless patience with me and for always making me laugh. Leandi for putting up with my absent mindedness and for never allowing a dull moment in the lab. Nicole for becoming such a close friend and for being supportive of me even when I should have been the support for her! Lebu for keeping me company in the early hours of the morning in the lab and for always being there to listen to my synthesis problems and offering me his expert advice (Thanks Professor). Last but not least, Monica for not only helping me with getting a crystal structure but also always being so organized and making sure I am in the right place at the right time.

Thank you to the NRF for funding, without which I would have had to drop out of university a long time ago. An abundance of gratitude to Adri at the NICD for doing all the biological assays and for maintaining such a close collaboration with our group. CAF for their help with my compound characterization. I would like to say a special thanks to Elsa and Jaco at the NMR facility – always service with a smile. Also, Malcolm and Marietjie at the MS facility.

Out of the lab, I owe my MSc to my wonderful partner Amir. You have been a continuous strength for me and even when I was so exhausted from juggling too much work you pushed me to keep going and made me believe I could achieve anything. To my parents who have always pushed me to be an academic and for putting me through undergraduate studies despite it not being easy financially. For their endless love and emotional support, and never once making me second guess my dreams.

***

I would like to dedicate this thesis to my Grandmother Erna McNeill Maclean who passed away in 2016 at the age of 92. The strongest and most fascinating person I have known, with sharpest mind, memory like an elephant and just generally a beautiful and peaceful lady always with a twinkle in her eye. She has always been a strength and inspiration in my life and will continue to be forever more. I miss you every day Granny.

***

Outputs

The work reported in this thesis has contributed to the following outputs:

Papers:

Brigg, S.; Pribut, N.; Basson, A. E.; Avgenikos, M.; Venter, R.; Blackie, M. A.; van Otterlo, W. A. L.; Pelly, S. C. Novel Indole Sulfides as Potent HIV-1 NNRTIs. *Bioorg. Med. Chem. Lett.* **2016**, *4* (6), 1580-1584.

Oral presentations:

Brigg, S. Indole Sulfides as Potent HIV-1 NNRTIs. Presented at the Frank Warren Conference, Rhodes University, Grahamstown, South Africa, December 2016.

Poster presentations:

Brigg, S. Indole Sulfides as Potent HIV-1 NNRTIs. Poster presented at the Frank Warren Conference, Rhodes University, Grahamstown, South Africa, December 2016.

Contents

Chapter 1: An Introduction to HIV and AIDS

1.1. The Discovery and Impact of HIV/AIDS

In 1970, the enzyme reverse transcriptase (RT) was first isolated from a group of viruses which converted single-stranded RNA into double-stranded DNA.[1] Since HIV belongs to the group *Retroviridae*, a viral family now characterized by the presence of this enzyme, the discovery of RT allowed HIV (HIV-1 and HIV-2) to be recognized as the causative agent of acquired immunodeficiency syndrome (AIDS) in the early 1980s.[2–5]

HIV-1 can be classified into three groups; Group M, N and O, which originate from independent zoonotic transfer events of simian immunodeficiency virus (SIV) from primates.[3] Groups N and O account for only 2% of all HIV-1 cases and are mostly restricted to Cameroon and surrounding countries, whereas type M is widespread and accounts for the other 98% of cases worldwide.[6] Group M is further classified into subsets A to J.[7] HIV-2 also arose from zoonotic transfer of SIV, however, the prevalence is very low and the progression of the disease is often too slow to take affect within a person's lifetime.

HIV is passed on through bodily fluids (blood, breast milk and semen), with 80% of HIV infections are passed on by unprotected heterosexual intercourse with an infected individual.[8] Other common modes of infection include homosexual unprotected intercourse, needle sharing and mother-to-child transmission during birth or breastfeeding. Although there has been a decrease in the number of new HIV infections and AIDS related deaths since 2000, owing to increased HIV/AIDS education and the effectiveness of antiretroviral treatment (ART), HIV/AIDS remains the worst pandemic faced by humankind.[9] UNAIDS reports that in 2014 there were an estimated 40 million people infected with HIV, 2.0 million new HIV infections and 1.2 million AIDS-related deaths.[9] Nearly two-thirds of people infected with HIV reside in Sub-Saharan Africa.[9] South Africa has the highest reported incidence of HIV infection in the world, with over 6 million infected and the Kwazulu-Natal province showing prevalence as high as 26.4% amongst the working age population.[8,10] It has been estimated that more than half the population of South Africa do not know their HIV status and for this reason statistics regarding number of infected persons are often under reported.[10]

1.2. The Structural Biology, Life cycle and Treatment of HIV-1

HIV-1 primarily targets CD4+ T cells and macrophages. The virus hijacks the host cells replication machinery and as a final step in the life cycle, the progeny virions are released into the blood stream to go on to infect other cells.[5] The virus manages to complete this life cycle and sabotage both the innate and adaptive immunity of humans whilst having a genome which codes for only 15 proteins.[5,11] Nine open reading frames (ORFs) code for four Gag proteins (matrix, capsid, nucleocapsid and p6), two Env proteins (gp120 and gp41), three Pol proteins (protease, reverse transcriptase and integrase).[5] The other six proteins are known as accessory proteins and are involved in various stages of the HIV life cycle. The life cycle can be described in thirteen key steps.

1.2.1 Steps One to Three: Attachment, Fusion and Uncoating

For attachment of the viral particle to the CD4+ T cells to occur, the virus has protruding structures which can form an interaction with the exo-proteins of the host cell membrane. Wild type HIV-1 virion has 7-21 spikes protruding, at random distances from each other, from the viral envelope.[12] Each spike is made up of two glycoproteins (gp), namely the trans-membrane gp41 and the surface gp120.[5,12] Gp120 subunits are made up of a variable domain consisting of five loops (V1 to V5) on the surface of the protein, and the inner core containing the conserved regions.[13] The design of this protein is critical to immune evasion; the variable loops protect the conserved inner domain and act as a "moving target" preventing the adaptive immune system from developing resistance.[13] The gp41 consists of a hydrophobic fusion peptide on the surface domain, a trans membrane domain, two heptad repeat (HR) regions and a cytoplasmic tail.[13]

The gp120 protein binds to the surface receptor CD4 on helper T lymphocytes and macrophages with high affinity (K_d approximately 4 nM) leading to a cascade of events (Figure 1).[5] CD4 binding results in movement of V1 and V2, and rearrangement of V3. These rearrangements create the co-receptor binding site, revealing part of gp41 and cause the CD4 receptor to fold in on itself bringing the host and viral membranes in closer proximity to each other.[13] All together this makes binding of gp120 to the chemokine co-receptor possible which in turn facilitates fusion.[5,10,14,15] Once gp120 is bound to both CD4 and co-receptors, gp41 fusion peptide projects itself into the host membrane, tethering the virus to the host membrane.[13] The trans membrane region of gp41 moves closer to the fusion peptide as the HR regions form a six-helix bundle facilitating pore

formation in the host membrane.[13] Finally, the nucleocapsid containing viral RNA and enzymes is released into the host's cytoplasm.[10]

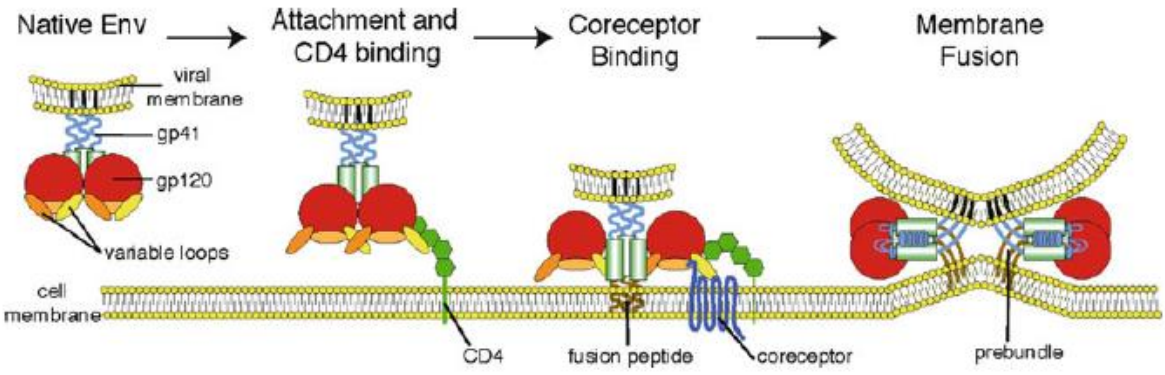


Figure 1: The fusion process. Diagram extracted from Wilen, Tilton and Doms, 2012[13]

Once inside the host cell, a process known as uncoating occurs whereby the viral RNA is released into the host cell's cytoplasm. Although it is still not known exactly when and how this occurs, it has been shown that the translocation of the virus by host cytoplasmic protein, Dynein, along cytoplasmic microtubules facilitates uncoating.[16]

There are currently only two FDA approved agents for this part of the HIV life cycle (Figure 2). The first is the fusion inhibitor, enfuvirtide, which is a homolog of the HR regions of gp41.[4] Enfuvirtide must be taken intravenously since it is a 36 amino acid polypeptide hence is degraded by digestive enzymes and stomach acid when taken orally. The second, maraviroc, is a co-receptor inhibitor (CRI) approved in 2007.[4] Maraviroc acts as an antagonist of chemokine CCR5, preventing binding of gp120 to the co-receptor.[4]

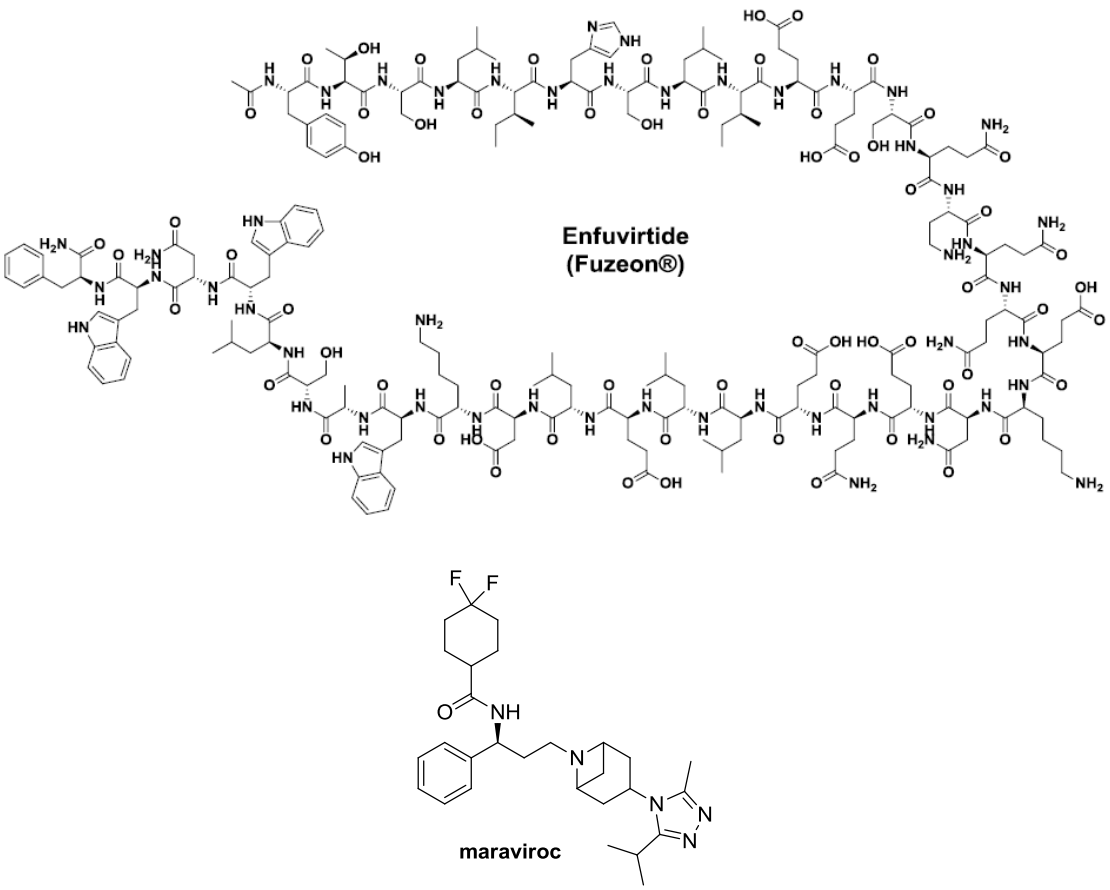


Figure 2: Structures of maraviroc and enfuvirtide, currently the only two entry inhibitors

1.2.2. Reverse Transcription

HIV-1 RT catalyzes the conversion of the single-stranded RNA genome of HIV-1 to double-stranded linear DNA. Each virion contains up to 50 copies of the enzyme, a heterodimer composed of two subunits, p66 and p51, derived from the Pol poly-protein (Figure 3).[11] p66 contains the catalytic sites for both the polymerase activity, as well as RNase H activity.[17] The polymerase domain resembles a hand, complete with fingers, thumbs and palm, and is connected to the RNase H domain.

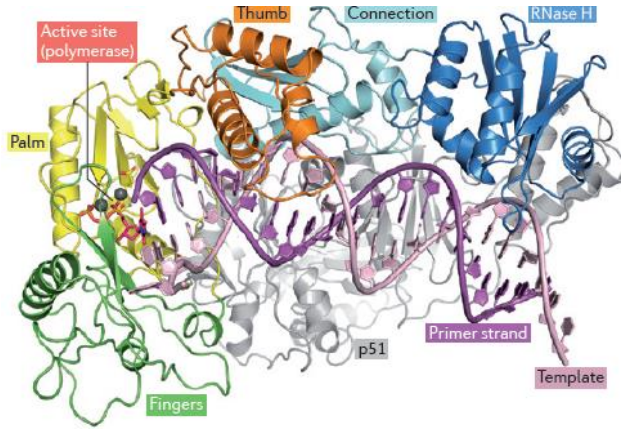


Figure 3: Structure of Reverse transcriptase containing a bound molecule of dNTP from Engelman and Cherepanov, 2012[11]

To get the polymerase activity started, both a template and a primer are needed. The template is the (+)-stranded viral RNA and the primer is the hosts (-)-stranded t-RNA, Lys3.[1] The primer's 3' end is base paired to a sequence on the 5' end of the template known as the primer binding site (PBS) and DNA synthesis can begin creating an RNA-DNA complex.[1] The RNA-DNA complex acts as a substrate for RNase H, which degrades the 5' end of the template to expose the newly synthesized (-)-strand DNA.[1]

Catalytic residues on the palm subdomain, aided by coordination to Mg^{2+} cations (shown as dark grey spheres in figure 3), activate the 3'-hydroxyl group of the DNA primer.[11] The 2;-deoxyribonucleoside 5'-triphosphate (dNTP) acts as a substrate and is incorporated into the growing DNA strand. Overall, the catalytic process results in single-stranded viral RNA being converted to double-stranded pro-viral DNA. There are three classes of ARVs which specifically target the enzyme reverse transcriptase; nucleoside RT inhibitors (NRTIs), nucleotide RT inhibitors (NtRTIs) and non-nucleoside RT inhibitors (NNRTIs).

3'-Azido-2',3'-dideoxythymidine (AZT) was the first HIV-1 ARV to get FDA approval.[4] AZT falls into the category NRTIs since it is a nucleoside analog, specifically a thymidine analog, competing with dNTP as a substrate in the active site.[18] There are currently seven FDA approved NRTIs, with several more undergoing clinical trials.[18] The key difference between NRTIs and the natural substrate of RT is the lack of a 3'-hydroxyl group which is essential for elongating the growing chain of DNA, resulting in termination of this process.[4] Like the natural substrate, NRTIs are phosphorylated by a series of enzymes, deoxycytidine kinase, deoxycytidine monophosphate kinase, and nucleoside diphosphate kinase, to give the triphosphate derivative of the drug; it is

only in this form that the NRTI can bind to RT.[4] The formation of the monophosphorylated species is the rate determining step for the formation of the active drug molecule.[4] There are several ways in which drug resistance has developed against NRTIs. Unfortunately, HIV-1 has developed the ability to incorporate the drug molecule into the viral DNA, and continue as normal with the rest of the life cycle through a mechanism not yet known. Mutations in the active sites, for example Met184 exchanged for Val or Ile, eliminate the ability for oxathiolane based NRTIs binding, whilst maintaining the ability to bind deoxyribose based natural substrate.[11]

Since phosphorylation of the drug molecule is the rate determining step in the binding of the drug to the catalytic site, NtRTIs were developed which already contain the phosphonate group, eliminating this delay in action.[18] The phosphonate group on NtRTIs are resistant to esterases, thus cannot be cleaved - a property which greatly increases the K_D of the drug molecule.[18] A major downfall of current NtRTIs is the presence of two negative charges on the molecules, this creates problems with transport across cell membranes.[18] This problem is overcome by formulating these drugs as the phosphate ester pro-drugs which are able to cross cell membranes and are metabolized to the active compound *in vivo*. There are currently two NtRTIs on the market, namely Tenofovir and Adefovir, both of which are also used to treat Hepatitis B.[19]

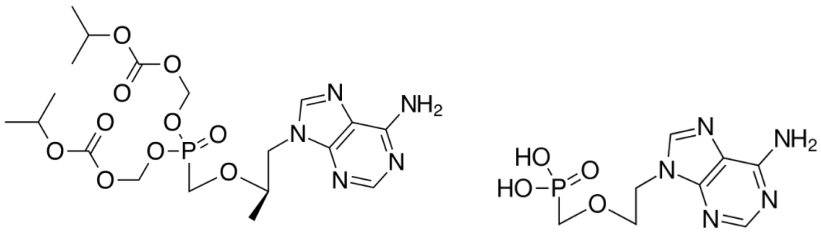


Figure 4: Structures of Tenofovir and Adefovir, the two NtRTIs currently approved for clinical use

NNRTIs are non-competitive allosteric inhibitors of HIV-1 RT, they target a small cleft, known as the NNRTI binding pocket (NNIBP), approximately 10 Å away from the active site whereby binding initiates a conformational change in the enzyme rendering it inactive.[15] The pocket is elastic and primarily hydrophobic in nature due to several aromatic amino acid residues (Trp and Tyr); however a number of hydrophilic residues are also present.[15] The overall structure and composition of the pocket lends to NNRTIs being highly variable in structure, though they are generally small, hydrophobic molecules.[20] Their hydrophobic nature of these compounds gives them the advantage of being able to cross the blood brain barrier (BBB), the only category of ARVs which are able to do so.[21]

First generation NNRTIs, such as nevirapine and delavirdine, are rigid, butterfly shaped compounds which allowed very little leeway for mutations to the NNIBP, making these drugs highly susceptible to resistance (Figure 5). Second generation NNRTIs are more conformationally flexible, allowing them to adapt to mutations in the binding site, giving them improved resistance profiles. Efavirenz was put on the market in 1998. It is considered a second generation NNRTI owing to the improved resistance profile, despite the lack of conformational flexibility.[15] Interestingly, efavirenz will only bind in the pocket of the enzyme-substrate complex and not when the enzyme is unbound.[15] Efavirenz has some interference with cytochrome P450 and has some interactions with protease inhibitors (PIs). Despite several problems, to this day, efavirenz is used in highly active ARV treatment (HAART). Etravirine and rilpivirine, also second generation NNRTIs, have much improved conformational flexibility increasing the potential number of binding modes in the NNIBP, greatly improving their resistance profiles.[18] Up to three mutations are required within the NNIBP to render these compounds ineffective against HIV-1 RT.[10] A downside of having flexibility is decreased selectivity, a number of the second generation NNRTIs have the ability to bind to other enzymes present in the human body, resulting in lower selectivity, thus increased toxicity.[10]

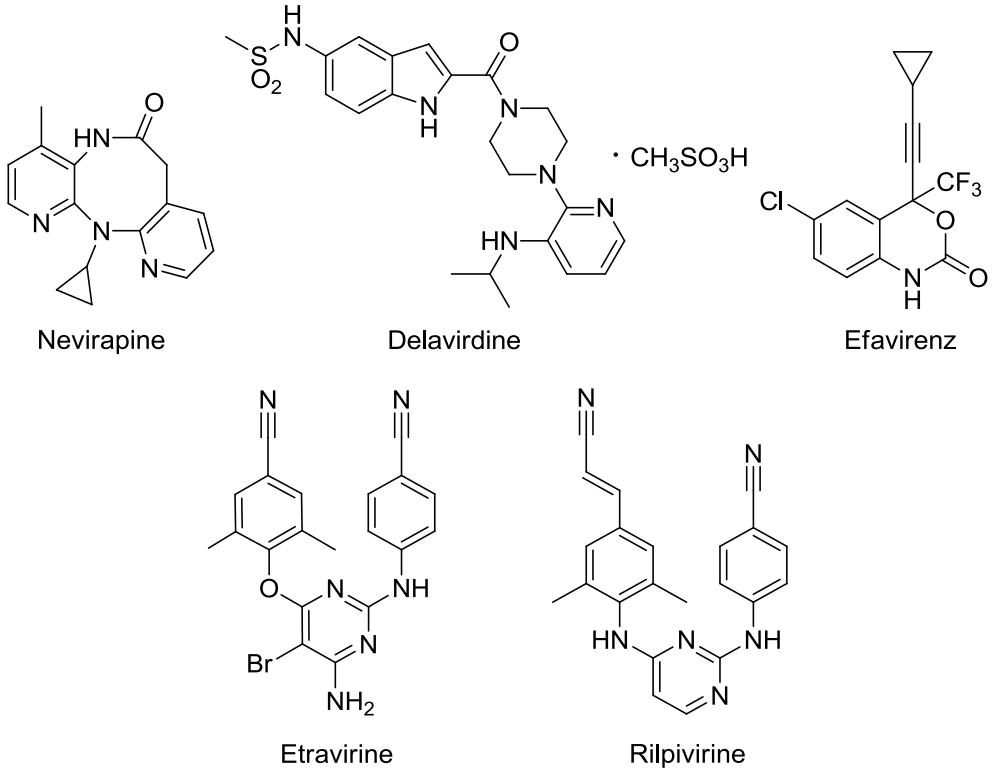


Figure 5: The structures of the five FDA approved NNRTIs

Another potential target in this stage of the HIV-1 life cycle is the Vif-APOBEC3G complex. RT is inhibited by a host cell cytidine deaminase restriction factor, namely APOBEC3G, which is part of human's innate immune system. This enzyme catalyses the conversion of cytidine to uridine in single stranded DNA, disrupting the process of reverse transcription by preventing binding of the DNA primer. To counteract the effect of the APOBEC3G, the virus uses a protein called Viral infectivity factor (Vif) to flag the APOBEC3G for degradation by proteolytic enzymes present in the host cell cytoplasm.[11] Targeting this complex could be an interesting way of indirectly preventing reverse transcription of the viral RNA.

1.2.3. Integration

Integrase (IN) is the key mediator of the integration process whereby viral DNA is integrated into the host DNA. The process begins directly after reverse transcription, where a pre-integration complex is formed. The pre-integration complex is made up of several proteins including viral matrix proteins, reverse transcriptase, integrase and host cell proteins including LEDGF/p75, as well as the viral DNA strand.[22] The pre-integration complex enters the nucleus of the cell through a nuclear pore. Integration is initiated by the removal of two nucleotides from the 3' end of the viral DNA by nucleophilic attack by a water molecule and is aided by the presence of a magnesium ion in the active site of integrase which deprotonates water, increasing its nucleophilicity.[22] This process, known as end processing, exposes a conserved CA dinucleotide sequence with a 3'-hydroxyl group initiating a single step transesterification reaction, inverting the chirality of the host DNA 5'- phosphate.[22] The choice of the insertion site on the host DNA is not well understood, however, there has been shown to be a trend for choosing sections of the DNA with sharp bend, the theory behind this is that the slight distortion of the DNA due to its strained shape may help lower the activation energy needed for the transesterification reaction.[23] Another observation is that integration seems to occur favorably at active transcription units.[23] This is thought to be due to the short half-life of infected T cells (typically 1 day) before the cell is destroyed by the immune system, so the virus has a limited time period in which to produce progeny.[23] The host factor LEDGF/p75 plays a big role in several steps of the integration process and has been found to be involved in tethering the pre-integration complex to the host DNA.[23] Once the viral DNA is inserted into the host genome, host enzymes complete the process by filling in the single strand gaps between the 3' end of the host DNA, and 5' end of the viral DNA.[11] Once fully inserted into the host cell DNA, the virus is now known as a provirus.

There are currently three FDA approved drugs which inhibit integration, raltegravir, elvitegravir, and dolutegravir (Figure 6), all of which are known as integrase strand transfer inhibitors (INSTIs).[24] Unfortunately, first generation INSTIs (raltegravir and elvitegravir) were high susceptible to resistance due to genetic mutations of integrase.[10] A second generation INSTI, dolutegravir, overcomes these problems and has a much better resistance profile.[10] INSTIs are currently used as part of first line and salvage therapy (in combination with two NRTIs) and a protease-based regimen (used in combination with a protease inhibitor).[24]

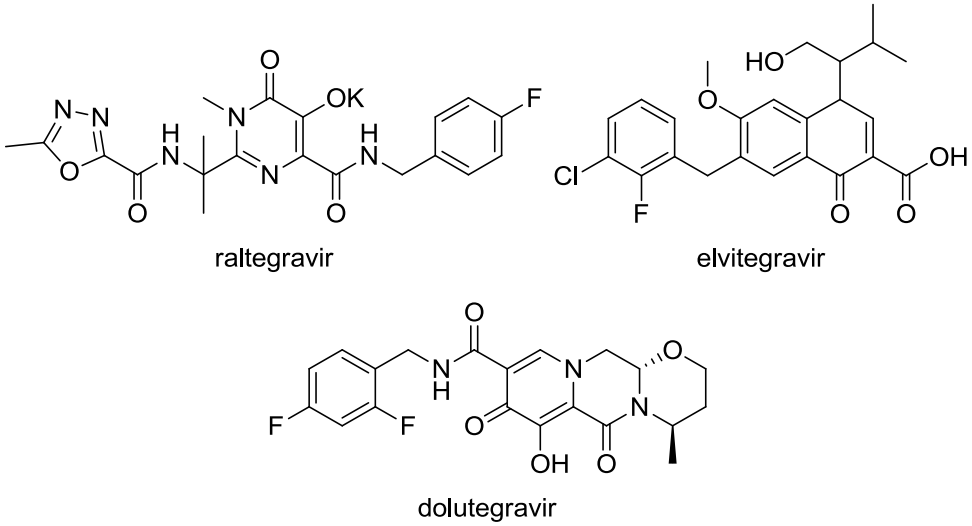


Figure 6: The three FDA approved INSTIs

1.2.4. Transcription

Once host cell transcription enzymes have initiated transcription of the viral DNA, viral protein Tat forms a complex with several host cell proteins which recognizes a sequence of transcribed RNA known as the transactivation-responsive region (TAR) found downstream to the initiation site for transcription.[11] TAR is characterized by a U-rich sequence at its 3' end which is most likely the recognition site for the Tat complex which binds here and promotes transcription elongation.[25] Several host cell proteins form a complex with certain regions on the viral RNA called a spliceosome, the structure which performs splicing under tightly regulated conditions.[25] The transcribed viral RNA is spliced into a total of forty different mRNA species, including incompletely spliced mRNA encoding Env, Vpu, Vif and Vpr, unspliced RNA used later as the virion genome or mRNA encoding the Gag-Pol polyprotein, and completely spliced mRNA encoding Tat and Rev.[25] Any incompletely spliced host cell mRNA would usually be destroyed in the nucleus at this point,

however, HIV-1 expresses the Rev protein which facilitates the transport of the incompletely spliced viral mRNA out of the nucleus (Figure 7).[25] Rev binds to a highly conserved region on the Env gene, known as the Rev-responsive element, inducing the formation of two purine-purine base pairs and is followed by the complexation of several other monomers to Rev as well as Crm1, which is part of the nuclear export complex.[25] Rev and the nuclear export complex interact with nuclear pore proteins and the whole assembly is transported across the nuclear membrane into the cytoplasm.[25] Crm1 is transported back to the nucleus immediately and Rev is released from the mRNA and forms a new interaction with importin-b which facilitates Rev's transport back into the nucleus. Completely spliced viral mRNA utilizes the same transport systems used for host cell mRNA.[25]

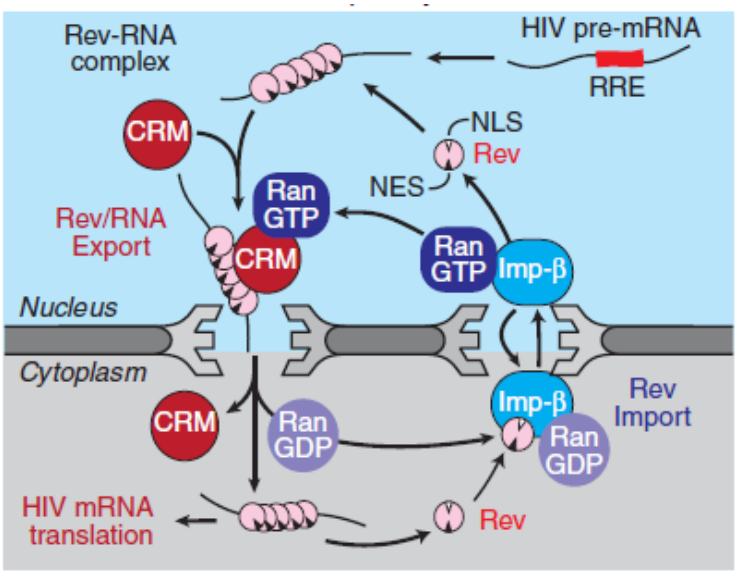


Figure 7: Rev:RRE interaction and the Rev nuclear export/import cycle from Karn and Stolzfus, 2012[25]

1.2.5. Viral Assembly, Budding and Maturation

Following transport across the nuclear membrane into the cytoplasm of the host cell, viral mRNA is translated to several precursor proteins by the host cell translation machinery.[11] From here assembly can begin starting with the 55kDa Gag and 160 kDa GagPol precursor polyproteins which move to the plasma membrane of the host cell, which is now used as the new virion cell membrane (Figure 8).[23] The Env glycoprotein is synthesized by the Golgi apparatus of the host cell and move into the plasma membrane where the gp120 subunit protrudes outside the host cell, later making the characteristic spikes on the viral cell, and the gp41 subunit anchors the

glycoprotein to the plasma membrane.[26] The Gag precursor includes the matrix, nucleocapsid, capsid and p6 proteins. The matrix protein tethers Gag to the plasma membrane whilst ensuring incorporation of the Env protein into the membrane of the newly formed virions.[27] The capsid helps with polymerization of the Gag (with intermittent inclusion of GagPol) to form the Gag lattice, whilst the nucleocapsid ensures the incorporation of the viral genome into the virions.[27] p6 interacts with host transport proteins which facilitate the budding process. Following budding, the enzyme protease, part of the GagPol polyprotein, cleaves the Gag and GagPol polyproteins into their corresponding proteins, resulting in the formation of mature virions which can now go on to infect more host cells.[26]

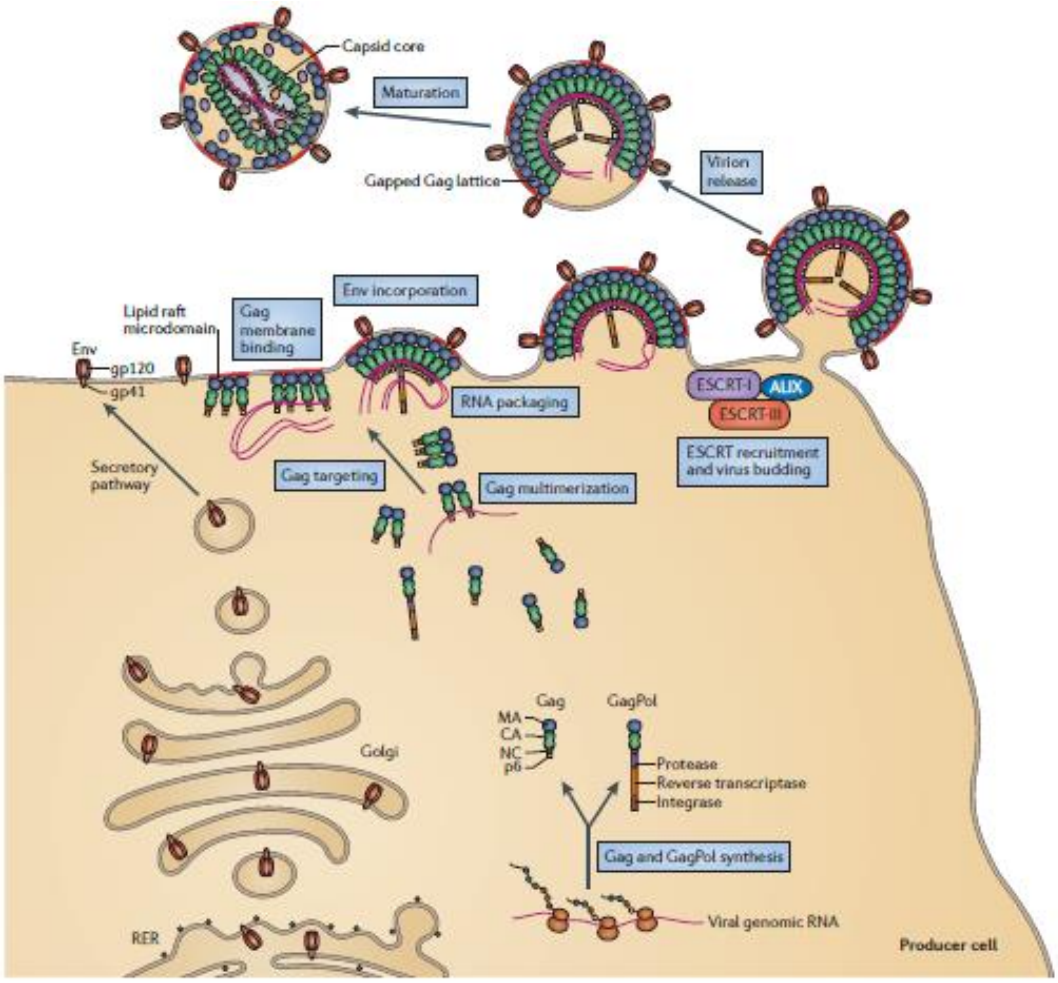


Figure 8: Viral assembly, budding and maturation from Freed, 2015 [26]

Investigation into why Group M HIV-1 had been so successful in the widespread infection of humans pointed to an explanation which involved the way the virus had evolved whilst adapting

from one species to another. The simian immunodeficiency virus (SIV) isolated from the chimpanzee subspecies (SIV_{cpz}) has been shown to originally have come about through recombination of two strains of monkey SIV, SIV_{agm} and SIV_{rcm}, neither of which were pathogenic.[3] The recombined genome of SIV_{cpz}, however, is pathogenic and this is due to two genes acquired during the recombination event. The first gene gives rise to the *Nef* protein whereas the second gene gives rise to the *Vpu* protein, both proteins have several functions one of which is the anti-tetherin activity.[3] Tetherin is a mammalian protein which binds between the viral envelope and the host cell membrane, preventing release of progeny virions.[28] The viral *Nef* protein binds to the cytoplasmic tail (CT) of tetherin, whereas *Vpu* binds to the transmembrane region.[3,28] Whereas, in chimpanzees only the SIV *Nef* protein has anti-tetherin activity due to the chimpanzee tetherin CT exhibiting less diversity, in humans, the CT of tetherin has a deleted pentamer and this has resulted in complete loss of anti-tetherin activity for the viral *Nef* protein, thus HIV-1 uses only *Cpu* protein for anti-tetherin activity.[3] Group M HIV-1 has better been able to adapt its *Cpu* protein for anti-tetherin activity without losing other important functions, such as degradation of CD4 cells as has occurred in Group N, or in the case of Group O, where loss of anti-tetherin function has occurred altogether.[3] This could be a major reason for the increased "replicative and transmission fitness" of the Group M HIV-1, compared to Group N and O.[3,29]

The only FDA approved treatment for these stages, assembly, budding and maturation, are the protease inhibitors (PIs) and there are currently ten of them.[18] PIs mimic the natural substrate of HIV-1 protease and can be divided into two categories; peptidomimetic inhibitors (saquinavir, ritonavir, darunavir, indinavir, fosamprenavir, nelfinavir, atazanavir, lopinavir, and amprenavir) and non-peptidomimetic inhibitors (tipranavir).[18] Due to their peptide-based structure, peptidomimetic PIs suffer from poor bioavailability and specificity issues. There is also a high prevalence of resistance.[18] The non-peptidomimetic PIs circumvent most of these problems and are also less susceptible to cross resistance with peptidomimetic PIs.[18]

1.3. The Stages of HIV Infection

According to the World Health Organisation (WHO) HIV infection goes through three main phases:[7]

1) Acute phase. The first two to four weeks after infection are asymptomatic but thereafter the patient may suffer from acute retroviral syndrome which results in flu-like symptoms as the viral

load (HIV RNA copies per mL of plasma) spikes, corresponding to a drop in CD4+ T cell count. It is during this stage that risk of transmission is greatest. Once the viral load has reached a stable level, the patient may again be asymptomatic, though swollen lymph glands are common.

2) Chronic phase. This phase is characterized by a steady decrease in the CD4+ T cell count and a steady increase in the viral load. Thus, the patient will get progressively sicker and experience a wide range of opportunistic infections, as well as becoming more prone to certain types of cancer. This phase usually lasts ten to twelve years.

4) AIDS. An HIV positive person is diagnosed with AIDS when their CD4+ T cell count drops below 200 cells/mm^3 of blood and/or has more than one opportunistic infection listed in the WHO manual for HIV Infection. With no treatment, AIDS can result in death within three years.

1.4. Highly active retroviral therapy (HAART)

Despite the amount of research that has gone into finding a cure for HIV, to date the disease can only be managed with chronic medication. The main problem facing drug design for HIV-1 is the high mutation rate of the virus, as well as high tolerance to these mutations. This makes the viral proteins moving targets for inhibition. Despite the difficulties, there has been considerable headway in developing agents which suppress viral loads and allow individuals to have quality of life, with major delay in onset of AIDS-related infections. The current treatment for individuals living with HIV/AIDS is HAART, which generally makes use of two NRTIs or NtRTIs, in combination with one of either a NNRTI or PI.[8] The WHO recommended first line treatment in 2015 as a combination of tenofivir, lamivudine and efavirenz and second line treatment as a combination of two NRTIs with either atazanavir with a low dose of ritonavir, or lopinavir, with a higher dose of ritonavir.[30] Unfortunately, due to high mutation rates and issues with drug adherence, it is still necessary to continue to design new therapeutic agents for the treatment of HIV.

Chapter 2: Introduction to our strategy

2.1 A closer look at the NNRTI binding pocket (NNIBP)

The NNIBP is an induced-fit allosteric pocket on the enzyme RT, which is not present in the absence of an NNRTI, but rather exists as cavity on the p66 subunit of the RT enzyme lined with hydrophobic residues.[8,17,18,31] In the unbound enzyme, the small solvent-free cavity is occupied by three tyrosine residues, namely Tyr181, Tyr183 and Tyr188. Tyr181 and Tyr188 mostly exist in the 'down' position in the free enzyme and during binding of an NNRTI they rotate to an 'up' position, though newer NNRTIs bind slightly differently (Figure 9).[8,14,15,21,31–33] The NNIBP is formed during this rotation, exposing the interactive residues within the pocket and allowing the inhibitor to be accommodated.[8] Once bound, the NNRTIs cause a change in the tertiary structure of the RT enzyme, causing shifting of both the p66 and p51 subunits.[8] More specifically the shift in the aspartic acid residues of the catalytic triad is thought to directly result in a dramatic decline in catalytic activity, or even halt it completely.[8]

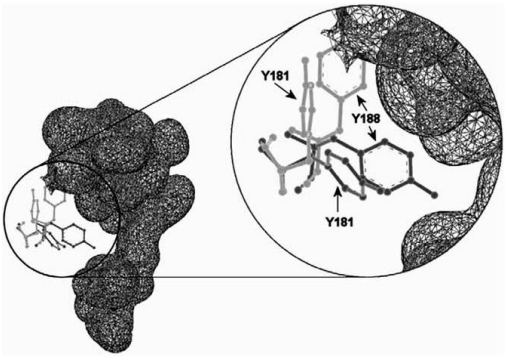


Figure 9: The 'down' conformation of the Tyr181 and Tyr181 residues of the NNRTI free enzyme (black) and the 'up' conformation of the same residues on the NNRTI bound enzymes (grey) from Martins et al, 2008. [15]

2.2 Interactions within the NNIBP

The vast structural diversity of the NNRTIs due to the flexibility of the NNIBP itself is quite remarkable.[8,34] However, when analyzing the interactions of these compounds within the pocket, it becomes apparent that they bind in similar ways and from similar interactions.[34] Due to the mostly hydrophobic nature of the pocket, there are only a handful of electrostatic interactions observed.[8,21] The interactions within the pocket start with simple Van der Waals interactions for amino acids such as Val179, Tyr181, Tyr188 and Phe227.[15] Hydrogen bonds can form between the

amino acid backbones, or can be more targeted to amino acid side chains, for example Val189 and Tyr318, but the latter is much less commonly observed.[15] Generally, the backbone amide carbonyls contribute most to the binding affinity of the inhibitors through hydrogen bonding, however a very dominant interaction is also seen with Pi-Pi stacking interactions between either Tyr181 or Tyr188 with the inhibitor, as well as a Pi-Pi interaction with Trp229.[15]

The interactions of different NNRTIs within the NNIBP is shown in Figure 10. Nevirapine shows almost no direct interactions. There is only Pi-Pi stacking interaction between its methyl pyridine ring with Tyr181 and a hydrogen bonding interaction with Leu234 *via* a bridging water molecule.[21] Efavirenz, etravirine and rilpivirine all show interactions with the backbone of Lys101 and this corresponds with a 50 to 100 times increase in potency when compared to neviripine.[8,21] Etravirine and rilpivirine also show Pi-Pi stacking with Tyr181 which appears to be an dominant interaction for all NNRTIs.[35] Studying these interactions of the successful NNRTIs gave great insight into which functionalities and scaffolds could be utilised in our drug design project.

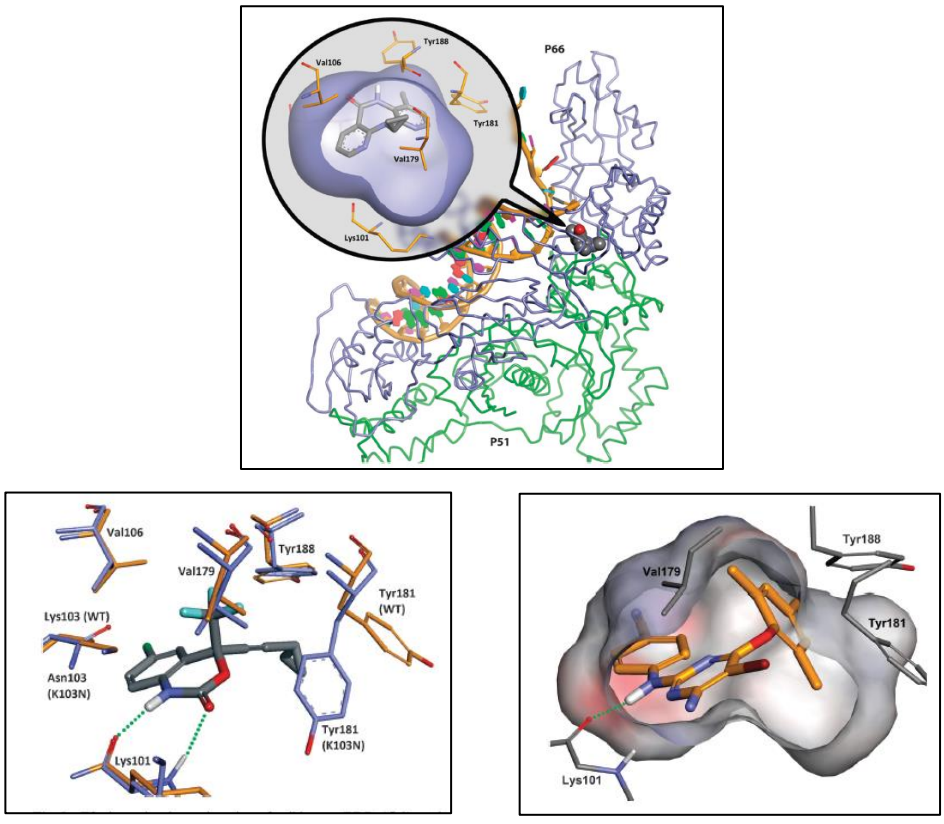


Figure 10: Interactions of different NNRTIs within the NNIBP. Nevirapine (top), efavirenz (bottom left), etravirine (bottom right) from Reynolds et al, 2012[8]

2.2 Resistance

Resistance is almost entirely responsible for the failure of HAART, and it has been seen for every NNRTI put on the market so far.[15] It is commonly attained by failure to comply with the drug regime, cross-infection (from one HIV positive person to another), monotherapy and pharmacological issues.[8,15] Since HIV-1 has a mutation rate on average of about one mutation per replication cycle, it makes the likelihood of development of resistance all that more probable.[15] Of course, many mutations are disruptive to the virus' life cycle so are not passed on easily but the incorrect use of NNRTIs can quickly result in the selection of drug-resistant mutants, since they are at an advantage when compared to the wild-type virus.[15] Resistance to NNRTIs arises through mutations in the non-critical residues in the NNIBP which reduces the binding affinity of the inhibitor either by inducing steric clashes, by removing favourable inhibitor-enzyme interactions which contributed to binding or changing charges within the pocket.[15] Since first generation NNRTIs generally have high specificity to the NNIBP due to their rigid structure, a single point mutation can render the drug ineffective and cross-resistance is common.[31] Second generation NNRTIs have a much higher barrier to resistance. Most mutations are found in or around the NNIBP which is what would be expected for allosteric inhibitors.[8] These directly affect inhibitor binding, however there are a number of mutations which are found outside this area that still successfully confer resistance.[15] Since there are such a large number of mutations which result in the loss of efficacy of NNRTIs, only a few key mutants are discussed here which are most relevant to this project.

The Y181C mutant, where tyrosine is replaced by a cysteine residue, has been shown to have around 20% prevalence in HIV-positive patients undertaking a treatment which includes an NNRTI with or without an NRTI.[36] This mutant results in loss of a key aromatic residue which is important for interactions with aromatic rings present in NNRTIs. This is a common mutant in Nevirapine treatment, since it results in loss of the only direct interaction that the drug has within the pocket.[31] Efavirenz is especially tolerant to this mutation since it does not rely on Pi-Pi stacking to this residue.

Unfortunately, efavirenz is not tolerant to all mutations, the K103N mutant is often selected for by treatment with Efavirenz.[36] This mutant occurs when Lys103 is replaced by an asparagine residue. The mutant appears to stabilize the NNRTI-free enzyme by a hydrogen bond forming between the new Asn103 residue and Tyr188, keeping Tyr188 in a permanent 'down' position

thereby blocking entry of NNRTIs into the binding pocket.[15] This is a unique way of developing resistance and has been very successful for the virus. The K103N mutant is currently by far the most problematic of the HIV-1 mutants. A study by Rhee *et* al showed this strain to have a prevalence of just 1.15% in treatment naïve patients sharply rising to 36.38% in patients treated with an NNRTI with or without an NRTI.[36] It has been known to decrease the binding of nevirapine for the NNIBP by 40 fold, and similarly, the binding of efavirenz by 6-fold.[37] An interesting observation of the ability of several new NNRTIs to bind to the NNIBP whether Tyr188 is in the 'up' or 'down' position offers insight as to how the drug is still able to maintain a fair amount of potency in spite of the mutation.[8]

The K101E mutant is generally a much less prevalent mutant (less than 8% in HIV-positive individuals undergoing treatment with an NNRTI with or without an NRTI) however it emphasizes a more complicated relationship with inhibitors and their binding pocket then previously observed.[36] Although nevirapine does not have a direct interaction with Lys101, this mutation leads to an 8-fold loss of potency.[8] The development of resistance is possibly explained by the fact that in the wild-type virus Lys101 interacts with Glu138 through the formation of a salt bridge. When the lysine residue is exchanged for a glutamic acid not only does no salt bridge form but the Glu138 is actually repulsed by the like charges, this repulsion changes the shape of the binding pocket.[8]

Due to the large prevalence of resistance of HIV-1 against NNRTIs, it is necessary to continue to develop new NNRTIs to maintain the drug development pipeline in order to keep pace with the exceptionally high mutation rate of the virus. There are a number of highly conserved residues within the NNIBP, which are required for the replicative fitness of the virus, including Trp229, Tyr318 and Trp401.[15] Since the virus cannot survive if there are changes to these residues, targeting them is an excellent strategy for developing new NNRTIs with a higher barrier to resistance.

2.3 Our strategy

2.3.1 Why NNRTIs?

Acknowledging the need for a continuous supply of new ARVs to enter the drug development pipeline, our group set out to develop new strategic therapeutic targets for HIV-1, specifically NNRTIs. There are several reasons for our interest in NNRTIs. Firstly, they are the only category of

ARVs that can cross the blood brain barrier (BBB). This is of importance since it has been shown that HIV infection stimulates the secretion of several substances from host cells, including neurotoxins, leading to a condition known as AIDS dementia complex.[38] Secondly, NNRTIs have shown great structural diversity which creates a unique opportunity for probing creative drug design strategies.[21] Previous studies in our research group involved reviewing the crystal structures of the ligand bound to the NNIBP of RT to gain a better understanding of interactions in this pocket.[21]

2.3.2 Why Indoles?

Using efavirenz as a starting point, it was envisaged that exchanging the main scaffold for an indole could result in a compound with similar interactions within the NNIBP as efavirenz which could then be further optimized.[14,21] A study of the literature showed that this was not the first time that indoles had been considered as a scaffold for NNRTIs.[21] In fact, in 1993 a phenyl sulfinyl indole (Figure 11) was shown to have an IC_{50} value on 63 nM.[39] This lead to a number of optimized structures which evolved from having a sulfinyl moiety such as for **1**[40], to a sulfonyl (IC_{50} = 3 nM)[41] in **2**, to a series of sulfonamide derivatives including **3** and finally a series of phosphorous based derivatives such as IDX-899 (compound **4**) were investigated.[8] IDX-899 made it to FDA clinical trials but this has been subsequently put on hold pending further investigation after there were reports of seizures in several patients.[42]

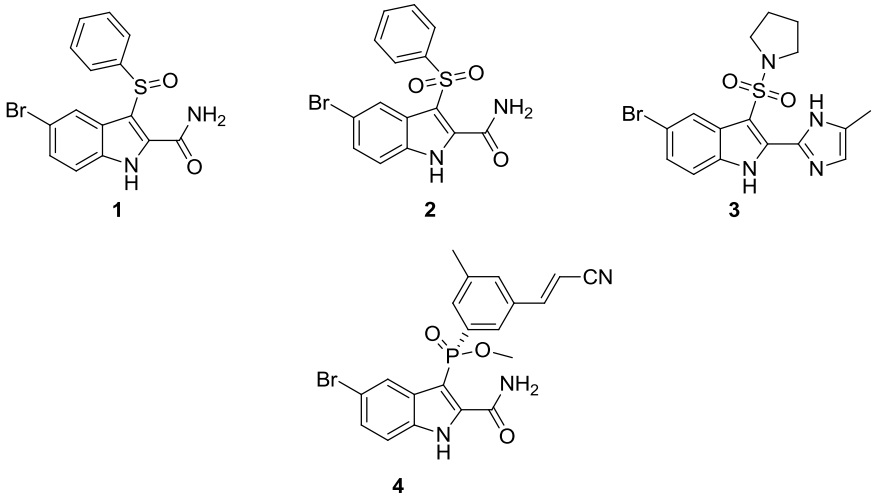
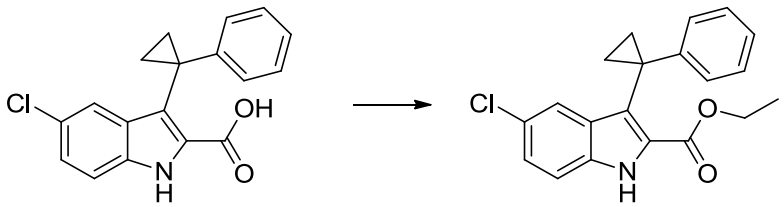


Figure 11

Since these indole-based compounds have been shown to be excellent inhibitors of RT we decided to base our design strategy on this scaffold. To get an understanding of what gave these compounds such good activity a thorough study of the binding of these inhibitors to the NNIBP was warranted and this revealed several important interactions within the pocket. First and most important was the double hydrogen bond interaction with Lys101 *via* the indole NH and the amide in the 2-position of the indole.[21] The amide shows further interactions with water molecules at the entrance of the NNIBP.[21] The phenyl ring showed Pi-Pi stacking with Tyr 181.[21] An apparent mismatched interaction in many of these inhibitors, however, was the placement of the polar sulfonamide oxygen in the small hydrophobic Val179 pocket. With this in mind, the group set out to design an indole based compound which maintained all these interactions, whilst improving potency and barriers to resistance.

2.3.3 The road to our lead compound

The first proof of concept compound (**Error! Reference source not found.**) was a cyclopropyl derivative **5**, which was made with a carboxylate in the two position of the indole with the idea of improving solubility.[21] The choice of the cyclopropyl group was not a random one, efavirenz and nevirapine both bear these moieties and modelling showed that superposition of the proof of concept compound with nevirapine ended up with both the cyclopropyl groups over lapping almost perfectly in the hydrophobic Val179 pocket.[21] Upon synthesizing this compound, an *in vitro*, non-replicative phenotypic assay which employed a HIV-1 retroviral vector system to produce virus-like particles (VLP)s was utilized and revealed a disappointing IC_{50} value of only 28.5 µM.[21] However, the subsequent synthesis of an ester derivative **6** gave efficacy results equaling nevirapine, with an IC_{50} of 0.1 µM. The proposed explanation was that binding was in fact occurring with the carboxylate derivative, however the hydrophilicity of the carboxylate was causing problems with the compound **5**'s ability to cross the cell membranes of host cells.[21]



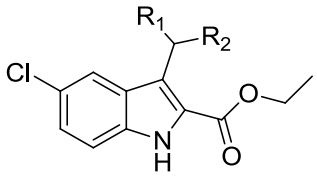
IC_{50} = 28.5 µM **IC_{50} = 0.085 µM**

5 **6**

Scheme 1

With a proof of concept compound now looking promising, the group set out to prepare further derivatives for a small SAR study. Firstly, the bromo *versus* chloro in the five position of the indole was investigated and both derivatives showed similar potency.[21] To confirm the necessity of a phenyl group in the 2-position of the indole, thiophene derivative **7** and methyl derivative **8** were also synthesized (Table 1) and as expected both showed low IC_{50} values, also adding weight to the modelling predictions.[21] The next step was to introduce modifications to the moiety occupying the Val179 pocket, currently a cyclopropyl derivative in compound **6** which had given difficulties in the synthesis, with the hope to get a more easily attainable bioisostere in this position. Firstly, complete omission of this group as seen for compound **9** decreased the efficacy more than ten-fold. This showed the importance of the van der Waals hydrophobic interactions in this pocket. Introducing a methyl ether to this position to give compound **10**, as predicted by modelling, gave a successful candidate for this Val179 pocket which was easy to obtain and showed a five times improvement in potency. The slightly larger ethyl ether derivative **11** showed similar potency. The hydroxy derivative **12** was also tested for activity but showed significantly less potency when compared to the ether derivatives, most likely owing to the hydrophilic nature of the group which was a mismatch for the hydrophobic Val179 pocket.

Table 1: Efficacy values for derivatives of the lead compound from Müller et al, 2014[20]

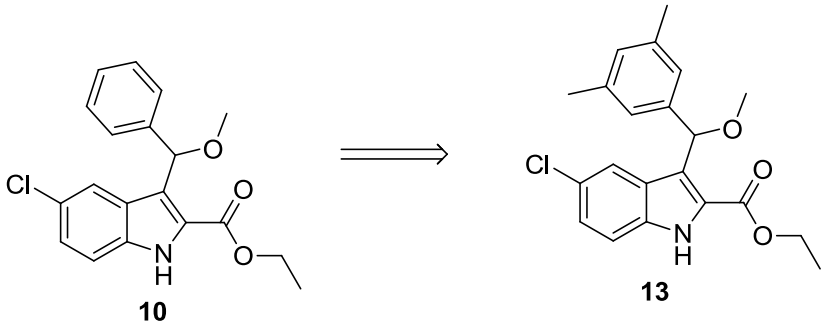


Compound	R_1	R_2	IC_{50} (µM)
6	phenyl	cyclopropyl	0.085
7	thiophene	cyclopropyl	0.065
8	methyl	cyclopropyl	2.25
9	H	phenyl	0.24
10	OMe	phenyl	0.02
11	OEt	phenyl	0.03
12	OH	phenyl	1.15

To show that these compounds were inhibiting HIV-1 by reducing RT enzymatic activity, the most potent compound in the series, derivative **10**, was assayed against a series of resistant RT strains containing mutations in the NNIBP. Mutations of Lys103, Tyr188 and Ty181 all caused significant changes to the potencies of these compounds, showing that indeed they are inhibitors of HIV-1 RT.[20,21]

2.3.4 Strategies for optimization of the lead compound and aims for this project

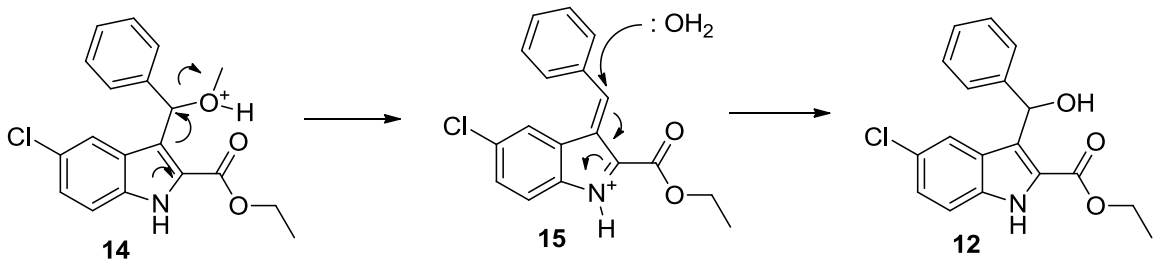
With lead compound **10** established as a potent inhibitor of wild-type HIV-1, a thorough investigation into similar analogues of this compound was necessary. This lead to the development of several aims for this project. Previously in our research group, a 3,5-dimethylphenyl derivative **13** (Scheme 2) was synthesised and tested again wild type HIV-1 and also was tested for efficacy against a mutant panel. Compound **13** showed excellent potency towards not only the wild-type HIV-1 but also several resistant strains, most importantly the K103N strain which as discussed previously is hugely prevalent.[20] The first aim of this project was to investigate the effect of changing substituents in the meta positions of the phenyl group on the compounds ability to inhibit common resistant strains.



Scheme 2

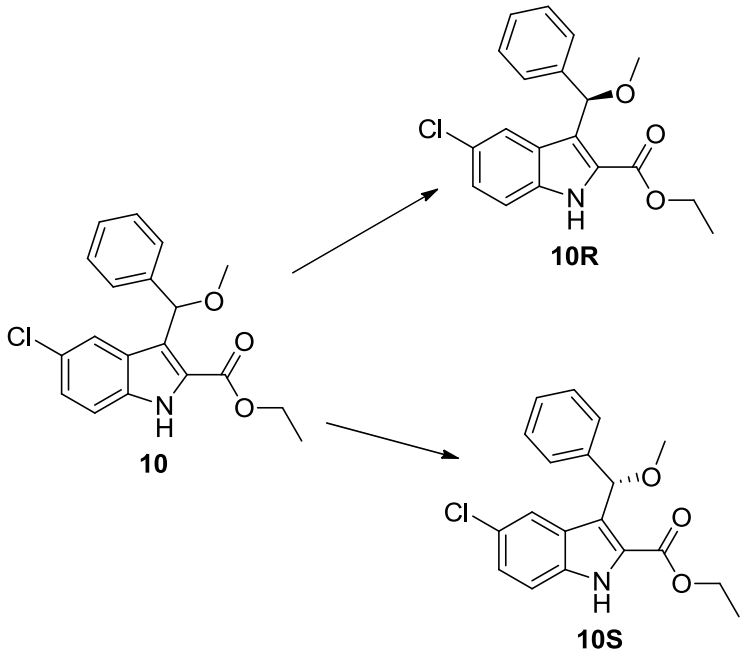
It was acknowledged early on that this compound had a major flaw when it came to acid stability in aqueous environments (Scheme 3). During an acidic work-up the methyl ether group is protonated to give intermediate **14**. An indole mediated S_N1 reaction, first leads to the elimination intermediate **15**, which then undergoes nucleophilic attack of water to afford the alcohol **12**.[43] This property of the compound was in fact utilised during the synthetic pathway, however from a drug design perspective this meant that the compound would not be viable as an orally administered drug since it would be converted to the significantly less potent hydroxy derivative **12** once in the acidic environment of the stomach. This gave us our second aim which was to

improve on the acid stability of the group occupying this Val179 pocket by using suitable bioisosteres.



Scheme 3: The mechanism for substitution of the methyl ether for a hydroxy group in an aqueous acidic environment

A final problem associated with lead compound **10** is the presence of a chiral centre, resulting in enantiomers **10R** and **10S**. To date only racemic mixtures have been tested for efficacy against HIV-1. Modelling shows that the R and the S enantiomer are not hugely different in terms of their binding scores, with the R enantiomer only scoring marginally better that the S enantiomer. To this end, our third aim was to design a strategy for separating these enantiomers so that their efficacies can be attained separately.



Scheme 4

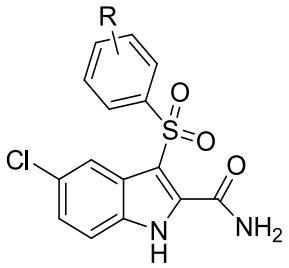
Chapter 3: Improving the resistance profile and acid stability of the lead compound

3.1 Inspiration from literature – Improving the resistance profile

Despite the excellent potency of lead compound **10** against wild type HIV-1, it was still necessary to further investigate the compound's efficacy against mutant strains of the virus, in order for it to be advantageous over other drugs already on the market. Subsequent research showed that introducing substituents on the phenyl ring, specifically a 3,5-dimethyl substitution, as seen in compound **13**, led to improved resistance properties.[20]

The concept of using 3,5-substitution on the phenyl ring was not novel. Silvestri and Artico (2005) had performed a thorough study on the effects of substitution on the phenyl ring of a sulfonyl indole derivative (Table 2) on the resistance properties of this compound.[40] The group synthesised a whole range of derivatives with different substituents on the phenyl ring including F, Cl, NH_2, *i*-Pr, *t*-Bu and Me. After the derivatives were assessed for their biological activity against HIV-1, mutant studies were performed on those which showed the greatest potency, paying special attention to results for the Y181C mutant and the double mutant Y181C/K103N.[40] The results for the study showed an interesting correlation between having substituents on the 3 and 5 positions of the phenyl group and maintaining efficacy against mutant strains of the virus (Table 2). No substitution in these positions, seen in compound **16**, showed hundred-fold loss of potency from wild-type to single mutant strain, and a thousand-fold loss of potency from single to double mutant strain. The presence of a methyl group in the three-position **17** maintained potency very well against the Y181C mutant, dropping only from a 0.001 µM to 0.006 µM EC_{50}, for the double mutant Y181C-K103N, however, this derivative lost most of its activity showing an EC_{50} of just 7 µM. An intriguing result came from the 3,5-dimethyl substituted phenyl derivative **18** where despite a ten times reduction in efficacy against the Y181C mutant compared to the wild-type HIV-1, there was not a significant loss of potency against the double mutant, again only a ten times reduction and even better it still showed activity against the triple mutant K103R-V179D-P225H.[40,44]

Table 2: Investigating the effect of substitution on the resistance profiles of different sulfonyl indole NNRTIs against HIV-1 mutants. Data was obtained from Silvestri and Artico, 2005.[40]



Compound number	R	EC_{50} wild-type IIB (µM)	EC_{50} Y181C mutant (µM)	EC_{50} K103N mutant (µM)
16	H	0.001	0.02	8
17	3-Me	0.001	0.006	7
18	3,5-dimethyl	0.004	0.03	0.65

This study was further backed up when Romines *et al* (2006) performed an extensive SAR study on benzophenone derivatives (Figure 12) with potent anti-HIV activity.[45] These compounds bind to the NNIBP in such a way that the terminal phenyl ring interacts with Tyr181 and Tyr188 in a similar way to our lead compound. Once again it was observed that substitution on the three and five position of the phenyl ring resulted in maintaining efficacy against mutant strains, including K103N and Y181C. In this study, CN, Me and Cl substituents had been utilised. Certain groups were not well accommodated such as a methoxy group in both positions.

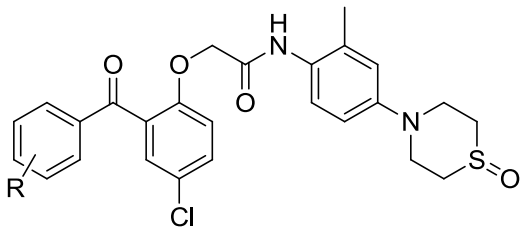


Figure 12: Benzophenone derivatives used in an SAR study by Romines et al, 2006[45]

Having established certain trends from literature our group set out to perform an SAR study on our indole based lead compound. The group was able to synthesise a 3,5-dimethyl phenyl derivative **13** and the results for the mutant studies thoroughly support the importance of 3,5 substitution in maintaining efficacy against mutant strains of HIV-1 (Figure 13). What stands out

from these results is that activity of compound **13** is maintained for K103N strain, the most problematic mutant strain of HIV-1 seen to date. Other mutants which did not become resistant to compound **13** include Y188C, V106M and G190A. Unfortunately, the Y181C and Y181I both displayed resistance to compound **13** and we hoped that investigating the substitution on the phenyl ring would give rise to a derivative which better maintained efficacy against these strains.

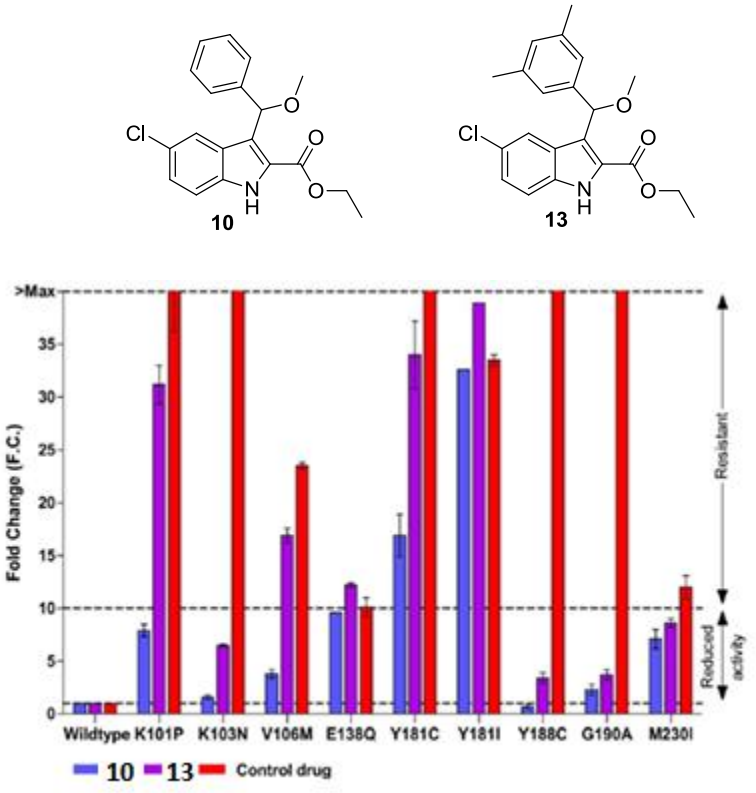


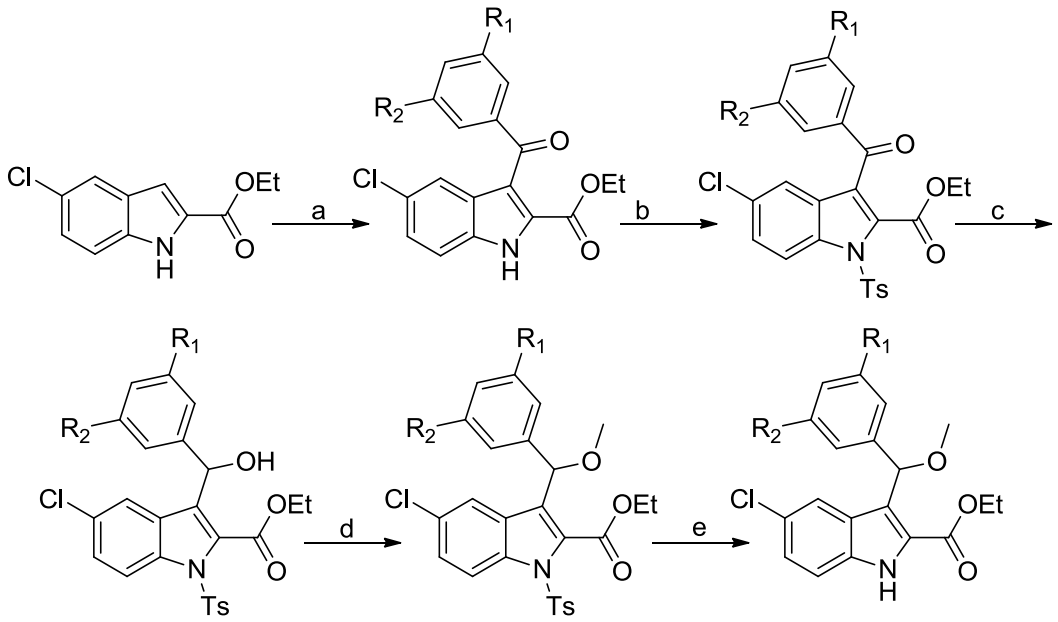
*Figure 13: Mutant panel for compounds **10** and **13** from Müller et al, 2014[20]*

With the success of this previous research we set out to develop several 3, or 3,5-substituted phenyl derivatives which we envisaged would have an impact on the resistance profiles of these derivatives.

3.2 An extension on previous research

Previous studies in our lab to develop compounds **10** and **13** meant that the synthetic route for development of these derivatives was already well precedented. The synthesis started with ethyl 5-chloro-2-indole carboxylate and comprised of only five steps (Scheme 5). The first was a Friedel-Crafts acylation (reaction **a)** using the corresponding acid chloride to produce the ketone. This

was followed by a tosyl protection of the indole nitrogen (reaction **b)** and reduction of the ketone (reaction **c)** to the corresponding alcohol using $NaBH_4$. The reversible substitution reaction (reaction **d**) that followed allowed the methyl ether derivative to be obtained by stirring up the alcohol derivative with MeOH and *p*-toluenesulfonic acid, albeit in low yield. Finally, a deprotection of the tosyl group (reaction **e**) gave the desired compounds.

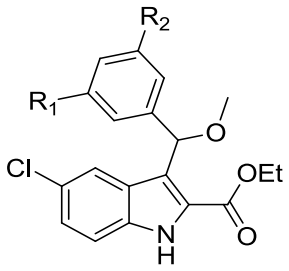


Scheme 5: Reagents and conditions: (a) aryl acid chloride, AlCl3, DCE, 2 h, 0–85 °C (b) p-TsCl, NaH, DMF, 18 h, 0 °C—rt (c) NaBH4, EtOH, THF, 3 h, 0 °C—rt (d) p-TsOH, MeOH, 18 h, rt (f) KOH, EtOH, THF, 4 h, 100 °C.

With a well precedented synthetic route at hand, we only needed to decide which substituents we wanted on the 3,5-position of the phenyl ring. This depended on the acid chlorides which were available to be synthesised or bought, so we used this as a starting point, additionally looking at common substituents used in other drug like molecules (Table 3). We specifically chose to synthesise the 3-methyl substituted phenyl derivative **19** since there was some evidence that methyl groups contributed to an improved resistance profile of the compounds.[40] The chloro, bromo and nitrile derivatives, compound **20, 21, 22**, were of interest since we envisaged these could potentially form interactions with Pi systems on the amino acid residues, such as Trp229. The nitrile derivative **22** we envisaged could be made by substitution of the bromine using a cyanide ion source such as CuCN. Research by Romines *et al* (2006), showed that methoxy groups substituted on the phenyl ring contributed negatively to the activity of these benzophenone derivatives, therefore we did not consider using this substituent.[45] However, nitroaromatic groups

are commonly encountered in drug molecules, including antibiotics (i.e. chloramphenicol) and antiparasitic drugs.[46] They unfortunately have been linked to bone marrow suppression and can have carcinogenic properties.[46] Dinitro derivative **24** could also be converted to the corresponding diamine **25** and then to the corresponding isonitrile **26**. We were interested in the isonitrile derivative **26** since it also had the potential to form an interaction with the Pi system of Trp229.

Table 3



Compound	R_1	R_2
10	H	H
13	Me	Me
19	Me	H
20	Cl	H
21	Br	H
22	CN	H
23	CF_3	CF_3
24	NO_2	NO_2
25	NH_2	NH_2
26	NC	NC

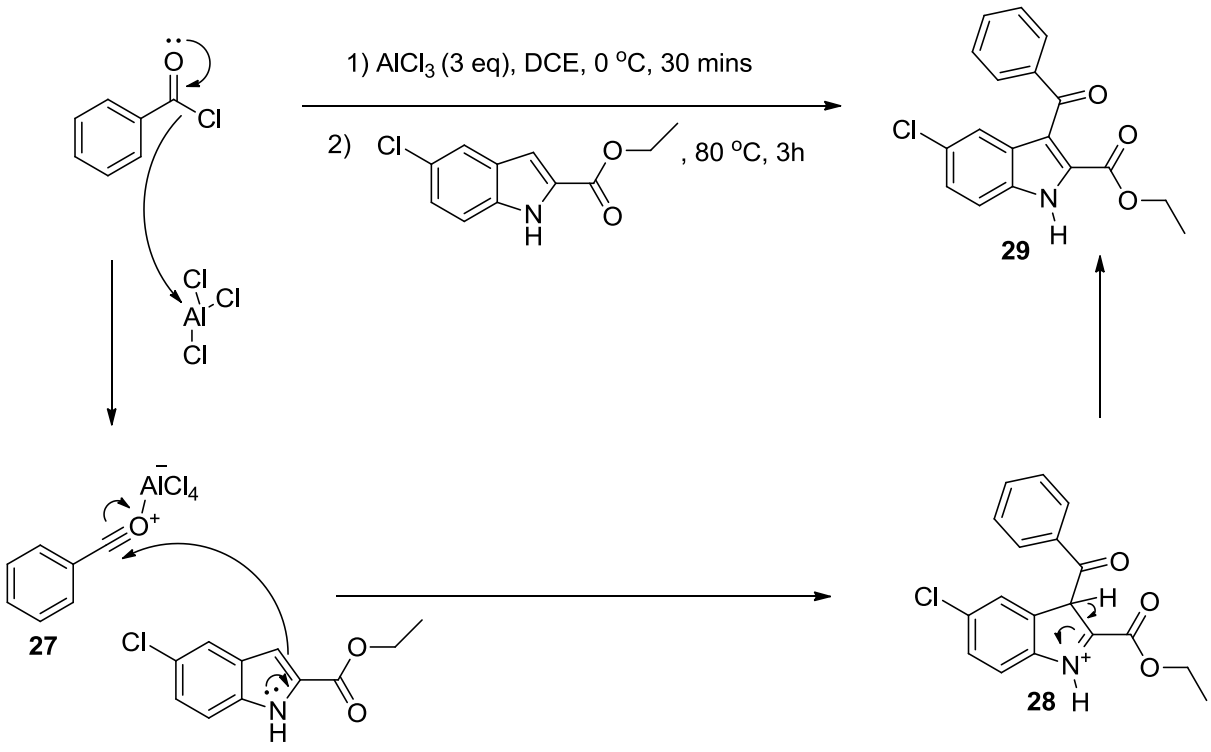
3.3 Using Friedel-Crafts acylation to acylate the 3-position of the indole

3.3.1 Introduction to Friedel-Crafts acylations

In 1887 two chemists, Charles Friedel and James Mason Crafts, found that in the presence of $AlCl_3$, benzene and amyl chloride, amyl benzene was formed.[47] This was the first Friedel-Crafts alkylation, and over the years a number of Lewis acids have been employed in this reaction

including BF_3, $TiCl_4$ and $SnCl_4$.[47] The major drawback with Friedel-Crafts alkylations is the lack of control when it comes to keeping the reaction stoichiometric. Due to the activating effect of alkyl groups on an aromatic ring, once one alkyl group has been added, it is very likely that more will follow. For this reason, Friedel-Crafts acylations became popular since these reactions circumvented this problem and the ketone that is formed can easily be reduced to an alkyl group.

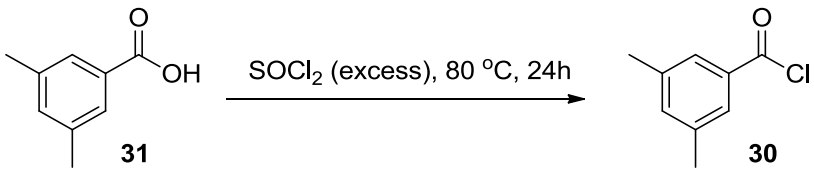
The mechanism for Friedel-Crafts acylation on an indole system is shown in Scheme 6. Elimination of the chlorine atom from benzoyl chloride forms an acylium ion, with $AlCl_4^-$ as a counter ion. The acylium salt **27** which is formed is a powerful electrophile, which allows the weakly nucleophilic double bond between the 2 and 3-position of the ethyl 5-chloroindolecarboxylate to perform the nucleophilic attack on the acylium ion carbon. This attack is aided by the lone pairs on the nitrogen of the indole which form a new double bond between the nitrogen and the 2-position, thereby facilitating the nucleophilic attack and stabilising the intermediate **28** by repositioning the positive charge onto the nitrogen atom. Elimination of the hydrogen from the 3-position of intermediate **28** followed by return of the nitrogen lone pairs, allows the indole to regain aromaticity and reveal the product, compound **29**.



Scheme 6: The proposed mechanism for Friedel-Crafts acylation

3.3.2 Synthesis of 3,5-dimethylbenzoyl chloride 30

To introduce variation in our target compounds, for the Friedel-Crafts acylation we needed to use a variety of substituted benzoyl chlorides in the Friedel-Crafts acylation. Most of the substituted benzoyl chlorides were available commercially so these did not have to be synthesised. However, 3,5-dimethyl benzoyl chloride **30** (Scheme 7) was not available and had to be synthesised from the corresponding carboxylic acid **31**. The 3,5-dimethyl benzoic acid **31** was refluxed in thionyl chloride for 24h, whereupon it was clear to see by TLC analysis that a product with a much higher R_f had formed, which is expected when converting a carboxylic acid to an acid chloride. The thionyl chloride was removed *in vacuo* and the crude sample of compound **30** was used directly in the Friedel-Crafts acylation without further purification.

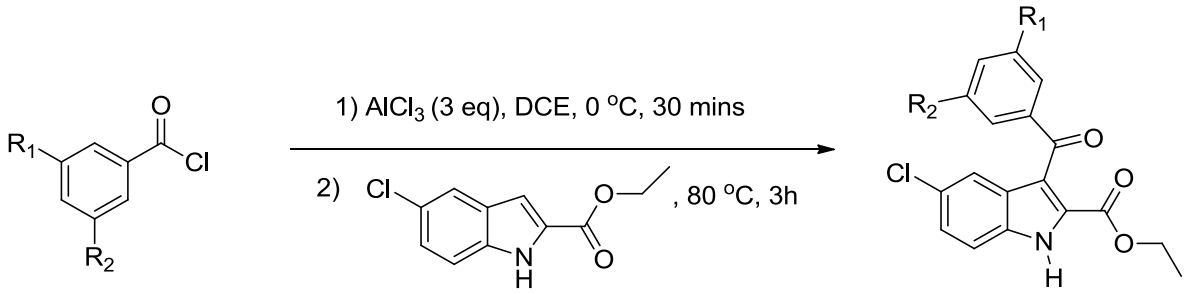


Scheme 7

3.3.3 Synthesis of the 3-acyl indoles

For the first step in our synthetic pathway, the synthesis of the 3-acyl indoles, we made use of a Friedel-Crafts acylation reaction (Scheme 8). The reaction was performed in 1,2-dichlorethane (DCE) as opposed to the more commonly used solvent for these reactions, 1,2-dichloromethane (DCM), in order for the reaction to be performed at higher temperatures. It was necessary to use elevated temperatures during this reaction since the indole system we were working with was not very reactive as a nucleophile due to the presence of the ester which draws electron density away from the system. A 3-necked round bottomed flask (RBF) fitted with a condenser was charged with DCE, followed by the corresponding benzoyl chloride. The flask was placed on ice before aluminium chloride was added and this solution was stirred at 0 °C for 30 minutes to allow the acylium ion to be generated before the indole was added and the reaction mixture was refluxed at 85 °C for 3h. For these reactions, a colour change from clear to green-black was observed after refluxing for 3h. The colour was indicative of the presence of aluminium salts generated as degradation products due to the high temperatures. Once cooled and added over a sodium bicarbonate/ice mixture a milky yellow emulsion formed which first had to be filtered

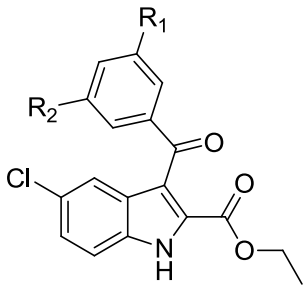
through celite for the work-up to be carried out successfully. This was followed by purification by column chromatography 10-40% EtOAc/Hex to yield the desired products.



Scheme 8

In general yields for this reaction were consistent (Table 4), and the substituents on the phenyl ring did not appear to have much of an effect on these reactions. The low yield for the nitro substituted derivative, compound **36**, was due to difficulty in purification as a result of insolubility of the product. The work-up itself did not cause any problems but after the organic layer (ethyl acetate) was dried over magnesium sulphate and filtered then concentrated under vacuum, the product started precipitating out of solution. Due to this observation, it was thought that filtering off this precipitate would provide us with a purified product. Unfortunately, after several recrystallization attempts, we were left with only 340 mg of a yellow solid and TLC analysis showed this was still not pure. Unfortunately, the compound isolated was not soluble in $CDCl_3$ or d_6-DMSO. ^1H and ^{13}C NMR spectra only revealed towering solvent peaks and no discernible product peaks. MS was also inconclusive giving a 100% intensity peak at 274.2730, which did not correspond to any product which we could predict from fragmentation or any common impurities seen for mass spectrometry. For the *bis*-trifluoromethyl derivative **37** the reaction and work-up all went per what would be expected, the only odd observation was when the crude product was dry loaded the silica gel went pitch black. Around 800 mg of a white solid was obtained from 1 g of starting material, which corresponds to only a 30% yield. However, ^1H and ^{13}C NMR analysis showed many impurities so the sample was recolumned but still showed impurities in the ^1H and ^{13}C NMR spectra despite only showing one spot on TLC. The MS of this compound intriguingly gave us a near identical value to the dinitro derivative (m/z = 274.2743) at 100% intensity, though we were still not able to identify this fragment.

Table 4



acid chloride used	Compound	R_1	R_2	% yield
benzoyl chloride	**29**	H	H	61
compound **30**	**32**	Me	Me	56
m-toloyl chloride	**33**	Me	H	65
3-chlorobenzoyl chloride	**34**	Cl	H	62
3-bromobenzoyl chloride	**35**	Br	H	44
3,5-dinitrobenzoyl chloride	**36**	NO_2	NO_2	Could not purify
3,5-*bis*-(trifluoromethyl)benzoyl chloride	**37**	CF_3	CF_3	-

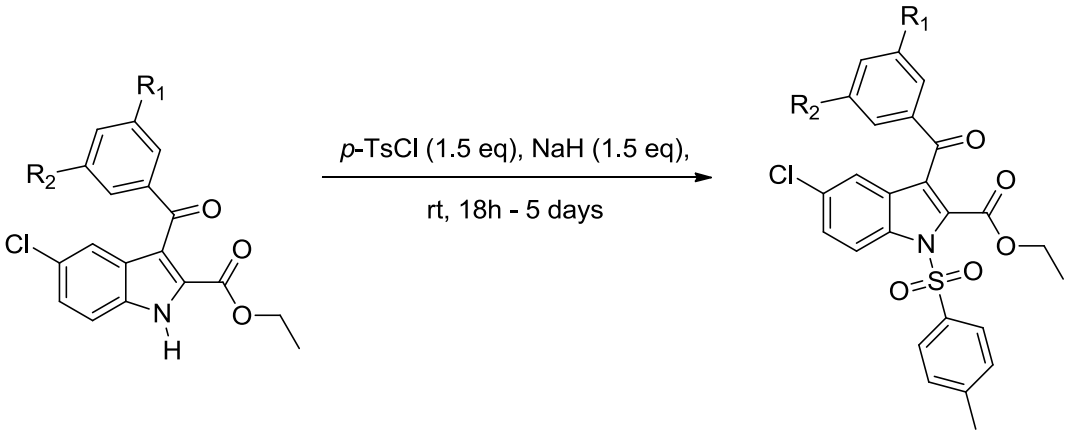
For derivatives **29** and **32** - **35** we could confirm that we had successfully synthesised the desired 3-acyl indole products. For the ^{1}H NMR spectra, an N – H peak was present between 9 and 10 ppm and the characteristic quartet at around 4.2 ppm and triplet at around 0.9 ppm corresponded to the ethyl of the ester in the 2-position of the indole. The presence of the aromatic acyl group could be confirmed since all aromatic peaks can be accounted for in the ^{1}H NMR spectra, as well as the C=O peak seen at around 196 ppm in the carbon spectra. The methyl substituents on the phenyl ring for compounds **32** and **33** were also accounted for in the ^{1}H and ^{13}C NMR spectra. Supplementary to this we also performed IR, MS and melting point analyses on all the samples and everything corresponded to the expected products.

3.4 Tosyl protection of the acylated indoles

Previous studies in our group showed that without a protecting group on the nitrogen, subsequent reactions gave poorer yields. We chose to use a tosyl protecting group instead of Boc due to the different conditions needed for the removal of these groups in our final step of the

synthesis. We knew that the methyl ether group is acid labile and since Boc groups often requires acidic conditions to be removed, the tosyl group seemed an obvious choice since they are typically removed under basic conditions and would not interfere with the methyl ether.

Having chosen our protecting group, we proceeded with the reaction (Scheme 9). The reaction for the different derivatives took between 18h to 5 days. Once all starting material was consumed, the reaction mixture was neutralised with ammonium chloride, extracted with ethyl acetate and column chromatography was performed to isolate the products.

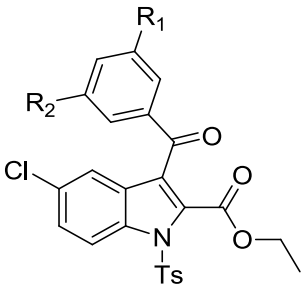


Scheme 9

Yields for the tosyl protection were consistent for derivatives **38** - **41**, however we saw a considerable reduction in yield for the bromo derivative **42** (

Table 5). The 3,5-dinitro derivative 43 was synthesised from an impure sample of compound **36**, with the hope that the compound might be easier to handle once the tosyl group was added. However, even in DMF compound **36** suffered from solubility issues. For this reason, the reaction had to be heated and left for several days, even so we only obtained a small amount of a solid which we were unable to characterise since the sample was not soluble in solvents needed to do ^{1}H and ^{13}C NMR spectral analysis and mass spectroscopy. It was decided that the dinitro derivative was too difficult to work with and due to the solubility issues would probably not fare well in a bioassay in any case. This meant we also had to adapt the original strategy for synthesis of the amine derivative **25** and isonitrile derivative **26**, since we had planned to make these directly from the dinitro derivative **24**. However, due to time constraints we never revisited these syntheses.

Table 5



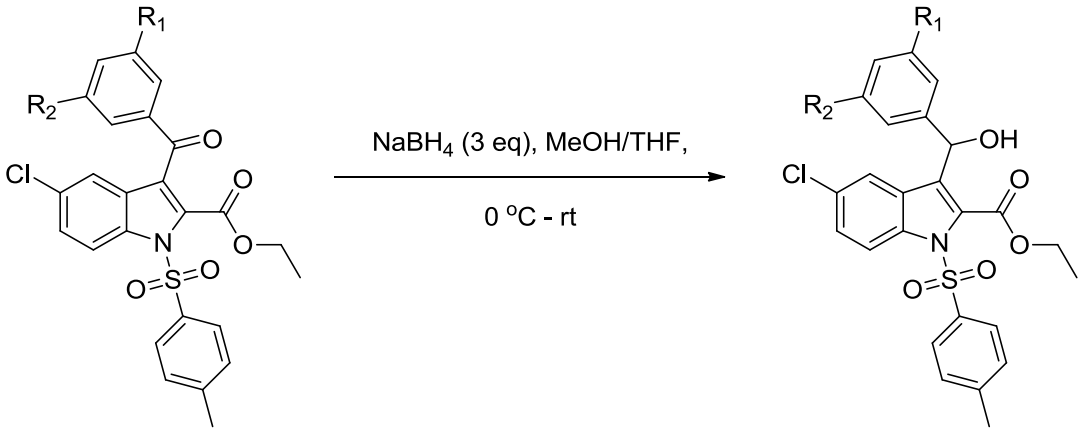
Compound	R_1	R_2	% yield
38	H	H	90
39	Me	Me	88
40	Me	H	87
41	Cl	H	89
42	Br	H	30
43	NO_2	NO_2	-

To obtain the bromo derivative **42**, the reaction had to be carried out for five days for all the starting material to be consumed. Product 42 was isolated in only 30% yield, and a considerable amount of a side product was also formed which could not be purified. Unfortunately, this side product was not characterised.

Derivatives **38 - 42** were fully characterised. ^{1}H NMR spectroscopy confirmed the presence of the tosyl protecting group for the derivatives firstly by the absence of the N – H peak, but also the presence of the methyl group of the tosyl at around 4.1 ppm, as well as the extra aromatic peaks in the aromatic region. IR spectroscopy showed the loss of the primary amine N – H stretch at around 3300 cm^{-1} as well as the presence of an S=O stretch peaks at around 1300 and 1200 cm^{-1}

3.5 Synthesis of the alcohol derivatives

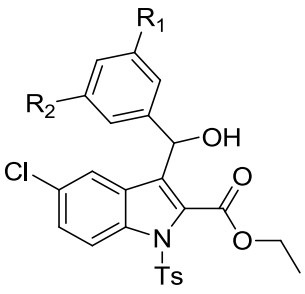
Having protected the nitrogen on the indole we were now able to carry out the reduction of the ketone to the alcohol by utilising $NaBH_4$ (Scheme 10).



Scheme 10

The indole was dissolved in a small amount of THF (2 mL) to aid with solubility, before 30 mL of MeOH was added. $NaBH_4$ was introduced at 0 °C and the reaction mixture was allowed to reach rt. The reaction was complete after 3-18h depending on the derivative, thereafter it was neutralised with concentrated ammonium chloride solution and extracted with ethyl acetate. Column chromatography was performed at 5-30% EtOAc/Hexane giving clean products in decent yields (Table 6). We saw decreased reactivity and yields for the halogen derivatives and these had to be left overnight for all starting material to be consumed.

Table 6



Compound	R_1	R_2	% yield	Reaction time
12	H	H	77	3h
44	Me	Me	77	4h
45	Me	H	84	4h
46	Cl	H	64	18h

47	Br	H	44	24h

For all the derivatives **12Error! Reference source not found.** and **44** - **47**, a side product with a slightly lower R_f to the product was seen on TLC, which we assume was most likely the product of the same side reaction which was occurring for all derivatives. For the synthesis of compound **45**, we analysed the side product which was formed. ^{1}H NMR analysis indicated this was a deprotected product, as indicated by the presence of an N – H peak of the indole nitrogen. Unfortunately, further characterisation was not possible due to the rapid degradation of this product, as observed by TLC analysis. It is possible that, since we were not using freshly distilled methanol for this reaction, the small amount of water in the reaction mixture reacted with the excess $NaBH_4$ and made a very weakly basic solution necessary for the tosyl group to be removed. This theory is supported by the observation that more of the side product was formed during the synthesis of the halogen substituted phenyl derivatives **46** and **47** which had longer reaction times. Due to the halogen substituents on the phenyl ring making the ketone less reactive, hence the longer reaction time needed, it makes sense that the longer exposure to the very mild basic conditions led to the increased formation of the side product. The reason for the halogen substituted derivatives **46** and **47** being less reactive is discussed further in Section 3.19

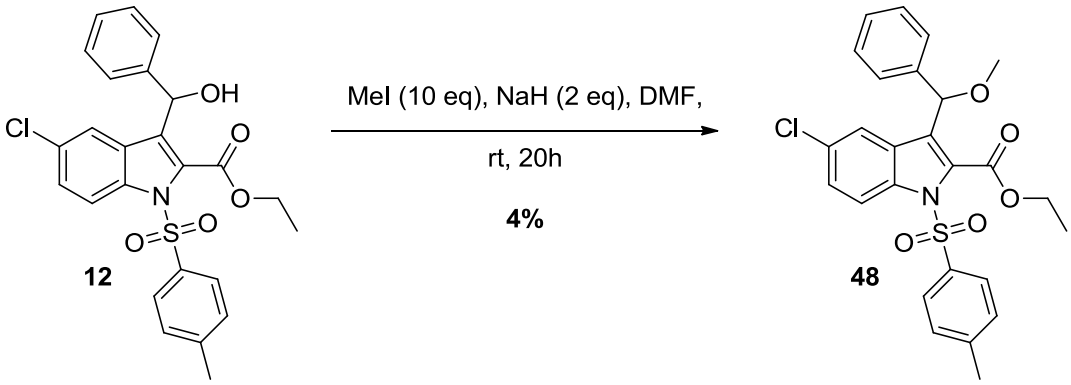
Derivatives **12** and **44** - **47** were fully characterised. The appearance of a singlet corresponding to the benzylic C – H at around 6.1 ppm and a broad O – H peak between 2 and 3 ppm confirmed that the ketone had been reduced, along with the disappearance of the C=O peak in the ^{13}C NMR spectra. In the IR spectra, loss of the characteristic conjugated C=O stretch peak at around 1700 cm^{-1} and appearance of the O – H stretch between 3200 and 3000 cm^{-1} further confirmed the formation of the alcohol.

3.6 Synthesis of the methyl ether derivatives

3.6.1 Synthesis of ethyl 5-chloro-3-(methoxy(phenyl)methyl)-1-tosyl-1H-indole-2-carboxylate 48

Having successfully synthesised the alcohol derivatives **12** and **44** - **47**, we were ready to introduce the methyl ether group. Since previous studies used a low yielding substitution reaction to obtain the methyl ether, we wanted to see if it was possible to optimise the formation of the methyl

ether. We attempted to make the methyl ether derivative **48** from alcohol derivative **12** by means of a methylation with MeI (Scheme 11). The reaction was performed in DMF, using NaH to deprotonate the alcohol, which was followed by addition of the MeI. After 20h, the reaction mixture showed multiple spots on TLC but not all starting material had been consumed. Nevertheless, we continued to work-up the reaction starting by acidifying with concentrated ammonium chloride and then extracting with ethyl acetate. Column chromatography resulted in the isolation of several fractions, with only 4 mg of the desired product **48** (4% yield) isolated from 100 mg of starting material. One of the side products we believe was the deprotected indole, and in the presence of excess NaH and MeI it is possible that the unprotected nitrogen of the indole could be both deprotonated and methylated attributing to yet another side product, ironically the desired final compound 10. Although we did not specifically isolate these products, a reaction done later in the project showed that NaH in DMF can remove the tosyl group easily at rt. With this hindsight, we could conclude that using a different base, such as NEt_3 or pyridine may have been successful for this reaction along with the use of DCM as a solvent.

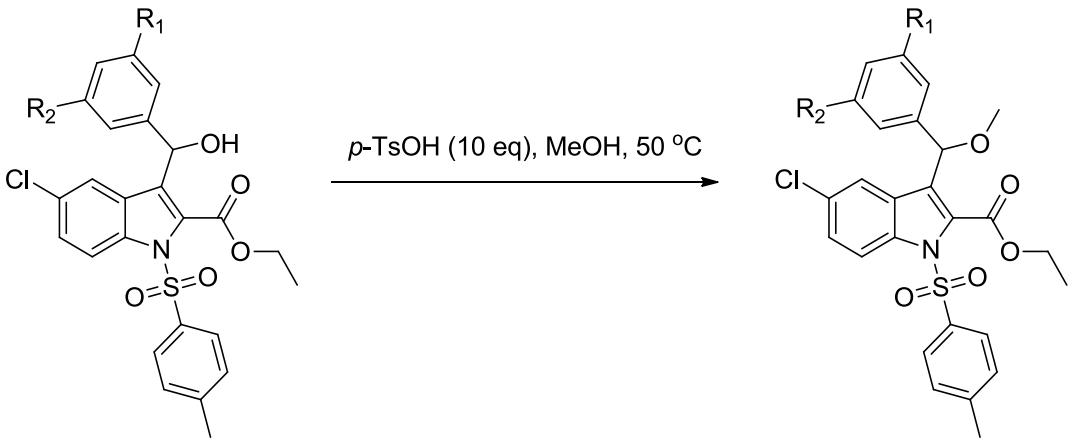


Scheme 11

3.6.2 Synthesis of the methyl ether derivatives using as substitution reaction

With the failure of the methylation attempt, we returned to the original synthetic route whereby the alcohol was substituted for the methyl ether by an acid catalysed indole promoted S_N1 reaction (Scheme 12). The indole was heated in MeOH, with ten equivalents of *p*-toluenesulfonic acid at 50 °C for 18h or longer (up to five days). Once we no longer saw any further conversion of starting material to product on TLC, the reaction was basified with saturated sodium bicarbonate and extracted with EtOAc. Purification was done by column chromatography (5 – 30% EtOAc/Hex) to yield the desired compounds. Unfortunately, we were unable to remove an impurity from the

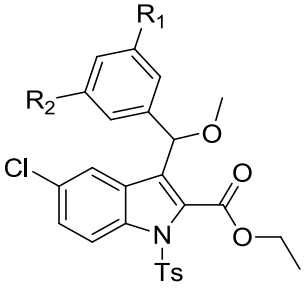
sample of compound **49** and continued to the next step without any characterisation for this compound.



Scheme 12

Although these reactions did not give good yields (Table 7) since they are reversible, there were generally no side products observed so starting material could be recovered and the reaction set up multiple times until enough product was obtained. This was especially utilised to obtain chloro derivative 51 where the yields were so low that even after 3-5 days there was a maximum of 11% yield. The bromo derivative **47** was left to react for 3 days and had very low yields and since we repeatedly had low yields for this derivative throughout the synthesis there was not enough starting material left at this stage to set up multiple reactions. We decided that we would revisit this derivative at a later stage. Nonetheless, the low reactivity of these halogen substituted derivatives **47** gave us some idea of how we could improve the acid stability of the lead compound. This was further explored and the findings are discussed later in this chapter.

Table 7

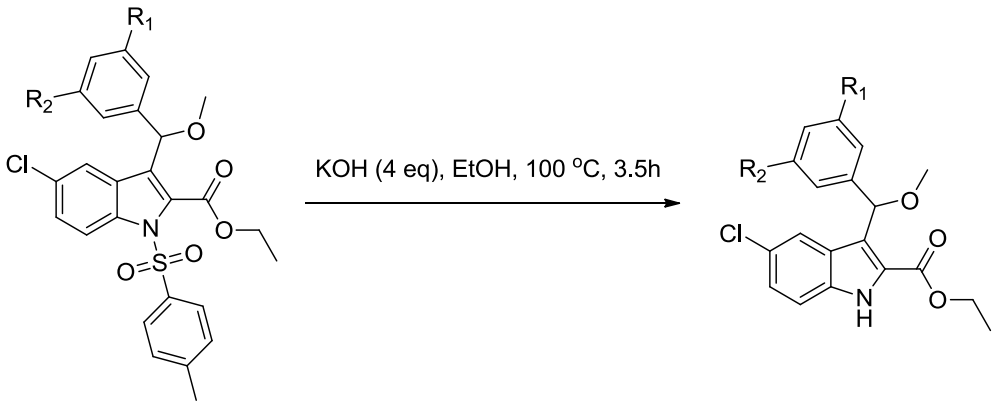


Compound	R_1	R_2	% yield
48	H	H	62
50	Me	Me	45
49	Me	H	Unable to purify
51	Cl	H	11
52	Br	H	9

Derivatives **48**, **50 - 52** were characterised by MS, NMR and melting point analysis. IR analysis was performed on compound **52** only. The presence of a new methyl peak in the ^{1}H NMR spectra at 3.24 – 3.31 ppm and in the ^{13}C NMR spectra at 76.1 – 77.8 ppm confirmed that the alcohol had been substituted for the methyl ether.

3.7 Synthesis of the unprotected methyl ether derivatives

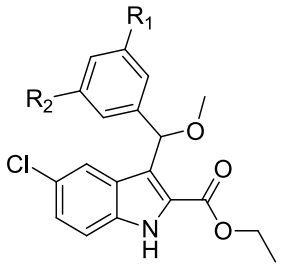
Now that we had successfully obtained the desired methyl ether derivatives, our final step was to remove the tosyl protecting group (Scheme 13). We chose to use KOH in EtOH to remove the tosyl, knowing that using any other alcohol would most likely result in transesterification under these conditions. After dissolving the indole in EtOH and adding KOH, the reaction mixture was heated to 100 °C for 3.5h. Once cooled, the reaction mixture was acidified with saturated ammonium chloride and extracted with ethyl acetate, before purification was done by column chromatography.



Scheme 13

The reaction worked very well for all remaining derivatives with yields ranging from quantitative to 72% (Table 8). The monomethyl substituted phenyl derivative 19 gave a 15% yield over two steps. All derivatives were characterised by IR, MS, NMR and melting point analysis. The removal if the tosyl group was confirmed by the presence of the unprotected N – H peak in the 1H NMR, and the loss of all signals corresponding to the tosyl group. Having successfully synthesised four methyl ether derivatives **10**, **13**, **19** and **20** we could now send these compounds for biological evaluation by our collaborators at the National Institute for Communicable Diseases (NICD) in Johannesburg, South Africa where they could be tested for their efficacy against both the wild type and several mutant strains of HIV-1.

Table 8



Compound	R_1	R_2	% yield
10	H	H	96%
13	Me	Me	quantitative
19	Me	H	15% over two steps
20	Cl	H	quantitative

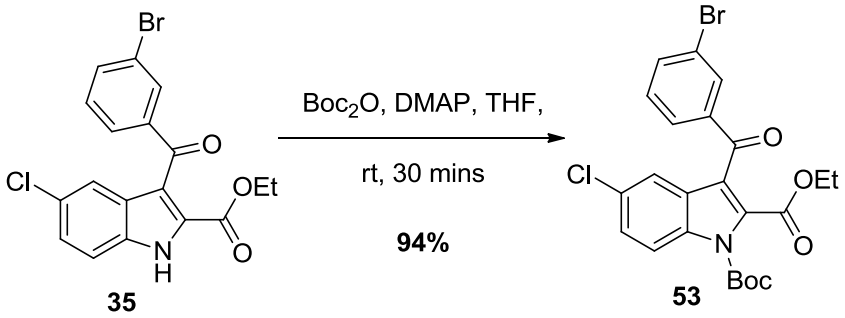
3.8 Revisiting the bromo derivative

The bromo derivative **21** had offered several challenges during the attempted synthesis with a tosyl protecting group. Firstly, the Friedel-Crafts acylation had been low yielding, but this being the first step in the synthesis meant that we could just repeat the reaction until we had a workable amount of the acylated product to continue. The problem came from the low yielding tosyl protection where a major side product and had resulted in a low yield of the desired product 42. To improve the yield of the protecting step, we decided to change the protecting group to a Boc group. Our original decision to use a tosyl group had been due to the Boc group requiring acidic

conditions to be removed, and we were worried this would interfere with the acid sensitive methyl ether moiety. A procedure used previously in our group using K_3PO_4 in an alcohol, offered us a milder method for removing the Boc and since the K_3PO_4 is not acidic we did not envisage that it would influence the methyl ether. Furthermore, since the Boc group was slightly less electronegative than the tosyl, we considered the reduction and substitution reactions may also have an improved yield. Having established a new approach for the synthesis of bromo derivative 21 we continued to test this strategy.

3.8.1 Synthesis of 1-*tert*-butyl 2-ethyl 3-(3-bromobenzoyl)-5-chloro-1H-indole-1,2-dicarboxylate 53

Since we already had the acylated product **35** at hand, we could proceed with the synthesis immediately, starting with a Boc protection (Scheme 14). The Boc protection of compound **35** carried out in THF. The indole was added followed by Boc_2O and a few crystals of DMAP. Compound **53** was synthesised in good yield (94%), which was already a major improvement from the 30% yield for the tosyl protection. The product was fully characterised by IR, NMR, MS and melting point analysis which confirmed we had been successful in making compound **53**. The key features which made this confirmation were the large singlet at 1.63 ppm integrating for nine protons corresponding to the *t*-butyl group on the Boc, as well as the absence of the N – H peak in the ^{1}H NMR confirming that the indole is protected with a Boc group.

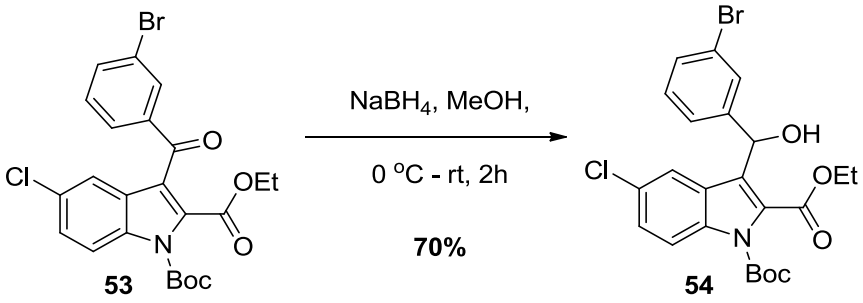


Scheme 14

3.8.2 Synthesis of 1-*tert*-butyl 2-ethyl 3-((3-bromophenyl)(hydroxy)methyl)-5-chloro-1H-indole-1,2-dicarboxylate 54

We now turned to the next step in the synthesis, reducing the ketone of compound **53** to an alcohol to give compound **54 (**Scheme 15**)**. This was carried out in the same way we had done for

the tosyl protected derivative **42**, using $NaBH_4$ as the reducing agent and performing the reaction in MeOH. The reaction proceeded with a 70% yield of compound **54** and gave no side products. Comparing the yields for the reduction of compound 42, which gave only a 44% yield, shows that the different protecting groups have a considerable effect on the reactivities of these derivatives. The possible reasons for this are discussed further in Section 3.19

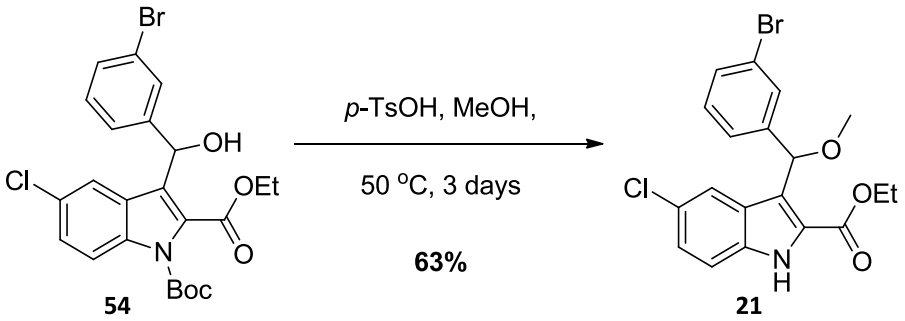


Scheme 15

To confirm we had reduced the ketone, compound **54** was fully characterised. The absence of the ketone signal at around 196 ppm in the ^{13}C NMR spectrum confirmed that this moiety had been modified. Two sets of doublets with matching coupling constants, corresponding to the benzylic C – H signal and the O – H signal, were present in the ^{1}H NMR spectrum confirming that the ketone had been reduced to the alcohol. The coupling of these two signals was expected, but it was interesting that we had not seen this for the tosyl derivative **47**, which had only shown a singlet for the benzylic C – H and a broad signal for the O – H.

3.8.3 Synthesis of ethyl 3-((3-bromophenyl)(methoxy)methyl)-5-chloro-1H-indole-2-carboxylate 21

The next step in the synthesis (Scheme 16) was carried out with much scepticism. This substitution reaction had been very low yielding for the tosyl derivative **47**, and we were not expecting much improvement on this yield. The reaction was carried out in the same manner, with MeOH and *p*-TsOH resulting in substitution of the alcohol for the methyl ether. However, to our delight not only did the substitution proceed but the Boc group was also removed during this reaction, giving us a 63% yield for these two steps and gave the final compound **21**.

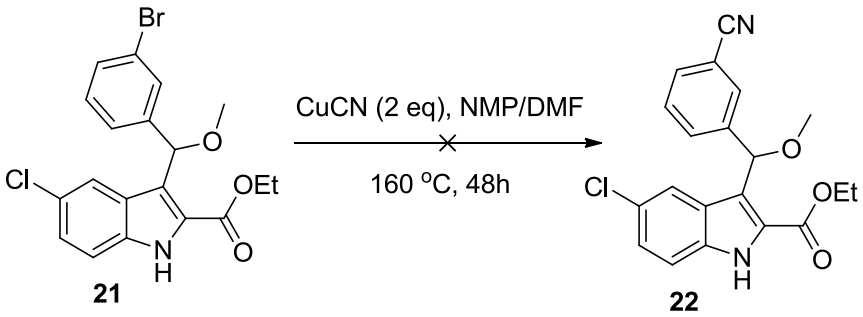


Scheme 16

With regards to the substitution, it is understandable that the Boc group would be removed under the acidic conditions presented by the *p*-TsOH, what was surprising was the extent to which the alcohol had exchanged for the methyl ether. When compared to the tosyl protected derivative **47** which had very poor yields for the substitution reaction, the Boc protected bromo derivative **54** had very successfully undergone this conversion. There are several possibilities as to why this indole mediated S_N1 reaction was so much improved for the Boc protected derivative. It may be due to the Boc group being less electron withdrawing than the tosyl group allowing the reaction may proceed more easily. Alternatively, if we assume that the Boc group was eliminated and without the deactivating effect of the electron withdrawing carbamate the indole mediated substitution may proceed with relative ease to yield the final product 21. Compound **21** was fully characterised. The absence of a signal integrating for 9 protons in the ^{1}H NMR spectrum at 1.62 ppm and the presence of a signal for the N – H proton at 8.90 ppm confirmed that the indole was deprotected. The presence of a singlet at 3.43 ppm integrating for 3 protons confirmed that the methyl ether group had been installed. Overall, this synthetic strategy had been much cleaner with the Boc protecting group in place, resulting in less side products and higher yields. The two-in-one reaction where the Boc group was removed along with the substitution reaction also made this synthesis one step shorter. In hindsight, this would have been a better protecting group to use for all derivatives. With the bromo derivative **21** successfully synthesised, we could send it for biological testing.

3.9 Attempted synthesis of ethyl 5-chloro-3-((3-cyanophenyl)(methoxy)methyl)-1H-indole-2-carboxylate **22**

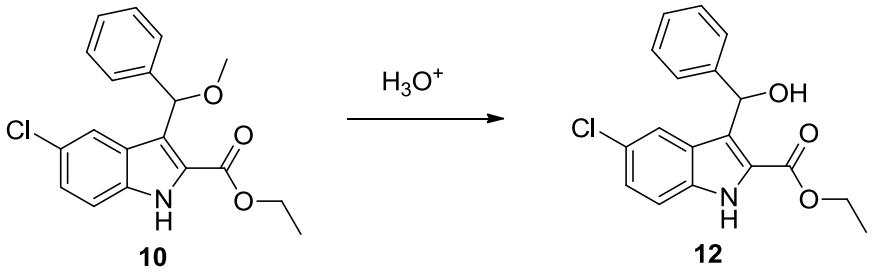
To expand our library of derivatives, we envisaged performing an electrophilic substitution of the bromo in compound **21** for a nitrile group to give compound **22** (Scheme 17). Compound **21** was dissolved in DMF and NMP to give a clear solution. CuCN was added and the reaction mixture was heated to 160 °C for 48 h. Although colour changes were observed, from clear to orange, TLC analysis only showed starting material and a spot on the baseline, presumably the CuCN salt. Following work-up there was no longer a spot on the baseline when a TLC was run. Purification allowed starting material to be recovered but no significant side products were isolated. Since this reaction was unsuccessful it was decided that there would not be any further attempt to make the nitrile until after the current library had been tested for efficacy.



Scheme 17

3.10 The problem with acid stability and our strategy

During the acid workup of compound **10** when it was first synthesised, a troubling observation came to light.[14] The compound was unstable in an aqueous acidic environment.[14] This meant that the compound could never be a viable candidate for an orally administered drug and continued studies on this lead compound would be of no use. The problem comes from inherent resonance of the indole which results in the lone pair on the nitrogen of compound **10** initiating an indole-mediated S_N1 reaction which eliminates the methyl ether and allows substitution for an alcohol group to give compound **12** if there is water present (Scheme 18).[43] This mechanism was in fact utilised to install the methyl ether group in previously synthesised derivatives, however to move forward with this project it was necessary to improve on the acid stability of the lead compound.



Scheme 18

In a previous study in our group, a tertiary amine derivative **55** (Figure 14) was synthesised and tested for efficacy against HIV-1.[43] Docking studies showed that it was well accommodated in the NNIBP, and had similar binding scores to the original compound.[43] Unfortunately, efficacy results did not agree with the modelling results, an uncommon occurrence for this project, and the compound was found to be not overly active (1.2 μM).[43]

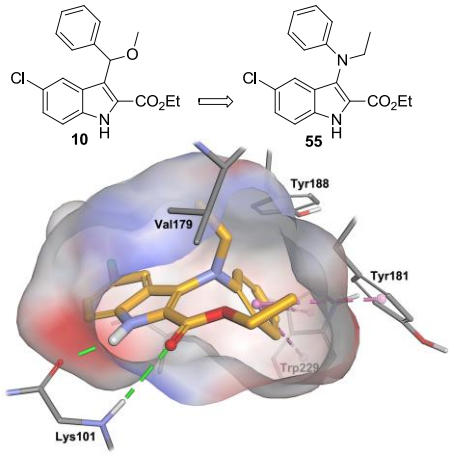


Figure 14: Bioisosteric replacement of the methyl ether for a tertiary amine showed promising results from molecular modelling from Brigg et al, 2016

Having found tertiary amine derivative **55** was not active, we turned to a different strategy. Using a bioisostere which was a worse Lewis base than the methyl ether, we hoped to counteract the problem of stability. We had several alternatives to choose from, either a group which had virtually no base capabilities such as an ethyl group in place of the methyl ether or a group which was a worse Lewis base compared to the methyl ether such as a phosphine or sulfur. Since we

were interested in whether the basicity would influence the stability, we chose to look at replacing the oxygen with a sulfur atom, the least basic option, forming a sulfide (Figure 15). Modelling results were again promising and the planar structure of the overall molecule would not be disrupted even if the sulfur atom was to react with a species in solution. A closer look at the modelling results showed that although the sulfur atom can be accommodated in the Val179 pocket, probably due to the larger size, it does not fit quite as well as the oxygen alternative so we were prepared for the efficacy results to reflect this outcome.

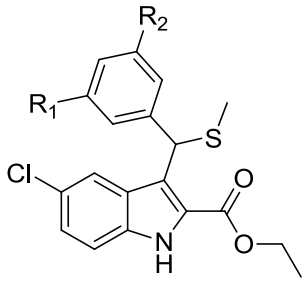
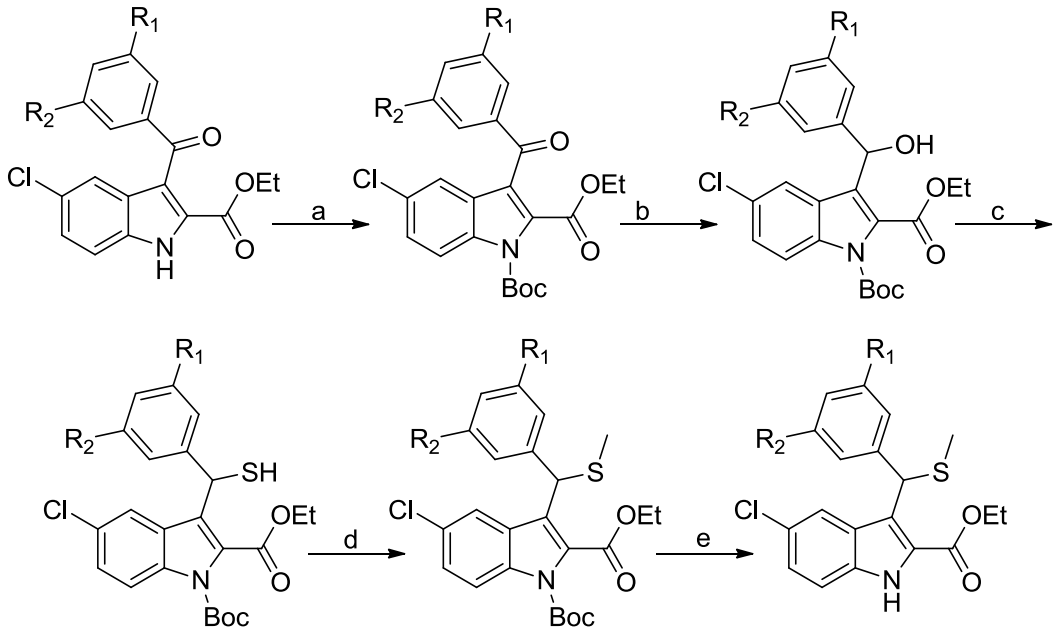


Figure 15

3.11 Introducing the synthetic route for the sulfide derivatives

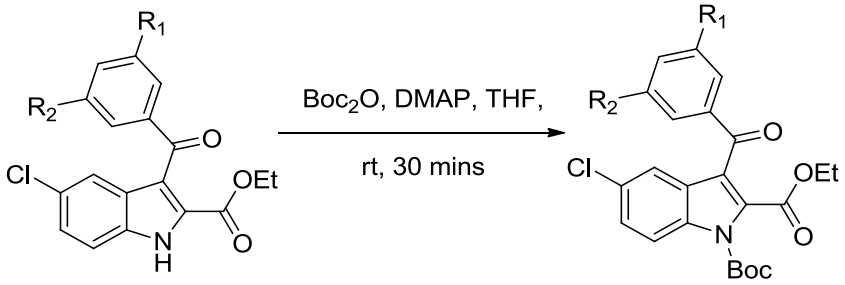
The synthetic route was much the same as for the methyl ether derivatives, only using Boc as a protecting group (Scheme 5). Starting with the acylated indoles **29** - **35**, the nitrogen on the indole was protected with a Boc group followed by reduction of the ketone to the alcohol. Here we branched off from the previous synthesis, substituting the alcohol for a thiol using Lawesson's reagent. The thiol was then methylated before the Boc protecting group was removed to yield the final compounds. With a synthetic strategy at hand we set out to synthesise several derivatives, again introducing the same substituents on the phenyl ring as for the methyl ether derivatives to evaluate for efficacy against mutant strains.



Scheme 19: a) Boc_2O, DMAP, THF, 3 h, rt b) $NaBH_4$, EtOH, THF, 3h, 0°C - rt c) Lawesson's Reagent, toluene, 3 h, reflux d) MeI, Et_3N, DCM, 18h, rt e) K_3PO_4, EtOH, 1 h, 100 °C

3.12 Synthesis of Boc protected indole derivatives

Having decided to use Lawesson's reagent in our synthesis, the Boc protecting group appeared to be a better choice than the tosyl group. Boc protected amines are seen in several syntheses where Lawesson's reagent has been utilised, and the protecting group remains untouched during the reaction. This reaction was carried out as before using Boc_2O and a catalytic amount of DMAP in THF to protect the indole nitrogen (Scheme 20).

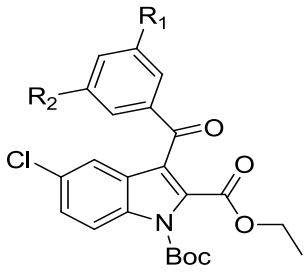


Scheme 20

The Boc protection was achieved in good yield across the board (Table 9), and furthermore the Boc protection was a much easier reaction to carry out compared to the tosyl protection. Not only was the reaction complete in only 30 minutes, there was no work-up involved and no unwanted

side products formed during the reaction. Furthermore, since the Boc group has no aromatic moieties, the ^{1}H and ^{13}C NMR spectra for the products are much easier to interpret in the aromatic region. In the ^{1}H NMR spectrum, the large singlet at 1.62 ppm integrating for 9 protons is characteristic and easy to identify to confirm that the indole nitrogen is protected with a Boc group. All derivatives were fully characterised and all analysis confirmed they were the desired products.

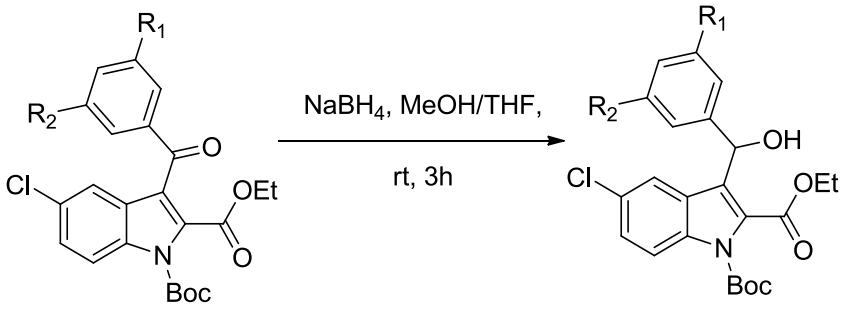
Table 9



Compound	R_1	R_2	%yield
56	H	H	88
57	Me	Me	75
58	Me	H	84
59	Cl	H	71
53	Br	H	94

3.13 Synthesis of the Boc protected alcohol derivatives

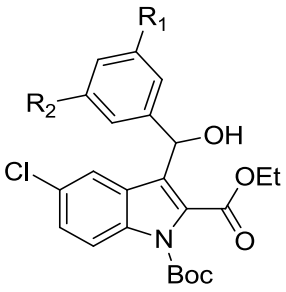
Having successfully protected our derivatives to give compounds **53** and **56-59**, we could now perform the reduction of the ketone to an alcohol (Scheme 21). This was done as before, using $NaBH_4$ as a reducing agent and MeOH as the solvent.



Scheme 21

The first noticeable difference between the tosyl protected and Boc protected derivatives for this reaction was the time it took for all starting material to be consumed. We assume this is due to the less electron withdrawing Boc group pulling less electron density from the indole and adjusting the reactivity at the benzylic position. Yields for the reduction (Table 10) were much improved with the Boc protecting group, but there is still a noticeable difference in yield for the halogen substituted phenyl ring derivatives **54** and **63**. These differences will be discussed further in Section 3.19 . All derivatives were fully characterised. However, we did not collect IR data for compound 63. In the ^{1}H NMR we see benzylic C – H signal at around 6.15 ppm, as well as the O – H signal confirmed the presence of an alcohol in the benzylic position. Once again, we saw doublets for these signals with matching coupling constants, and the O – H signal is very well resolved. This is contrary to the tosyl protected alcohol derivatives which showed only a singlet for the benzylic C – H and a broad signal for the O – H.

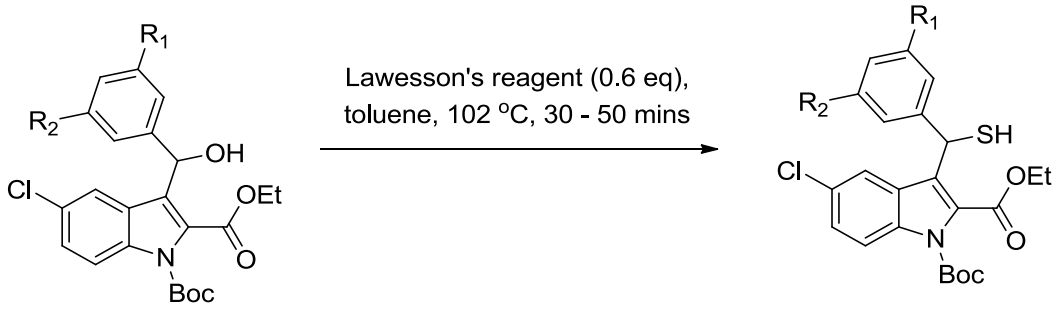
Table 10



Compounds	R_1	R_2	Reduction %yield
60	H	H	94
61	Me	Me	94
62	Me	H	96
63	Cl	H	76
54	Br	H	70

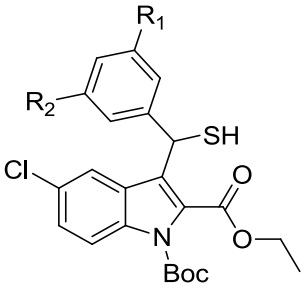
3.14 Synthesis of the thiol derivatives

In 1978, 2,4-*bis*-(4-methoxyphenyl)-1,3,2,4-dithiadiphosphetane 2,4-disulfide started being investigated for thionation properties on numerous carbonyl based functional groups.[48] The group which initiated this research was led by Sven-Olov Lawesson, hence the name Lawesson's reagent. However, it was not until 1993, that Lawesson's reagent was first used by Takehiko Nishio for the direct conversion of an alcohol to a thiol, which is what we wanted to achieve for our derivatives.[49] The one problem we could foresee was the presence of the ester, but since the ester was less reactive towards Lawesson's reagent than the alcohol we hoped that using only 0.6 equivalents of the Lawesson's reagent in the reaction would reduce the probability of the ester being attacked. The reaction was carried in dry toluene, under nitrogen. The alcohol derivatives **54** and **60-63** were added followed by the Lawesson's reagent and the reaction mixture was refluxed at 102 °C for 30 – 50 minutes. At this point the reaction mixture started turning light green which appeared to signal the appearance of side products, as seen by TLC analysis. Leaving for any longer amount of time only resulted in more spots on TLC, in particular a darkening spot on the base line. No work-up was necessary, the reaction mixture was simply concentrated *in vacuo* and purified by column chromatography.



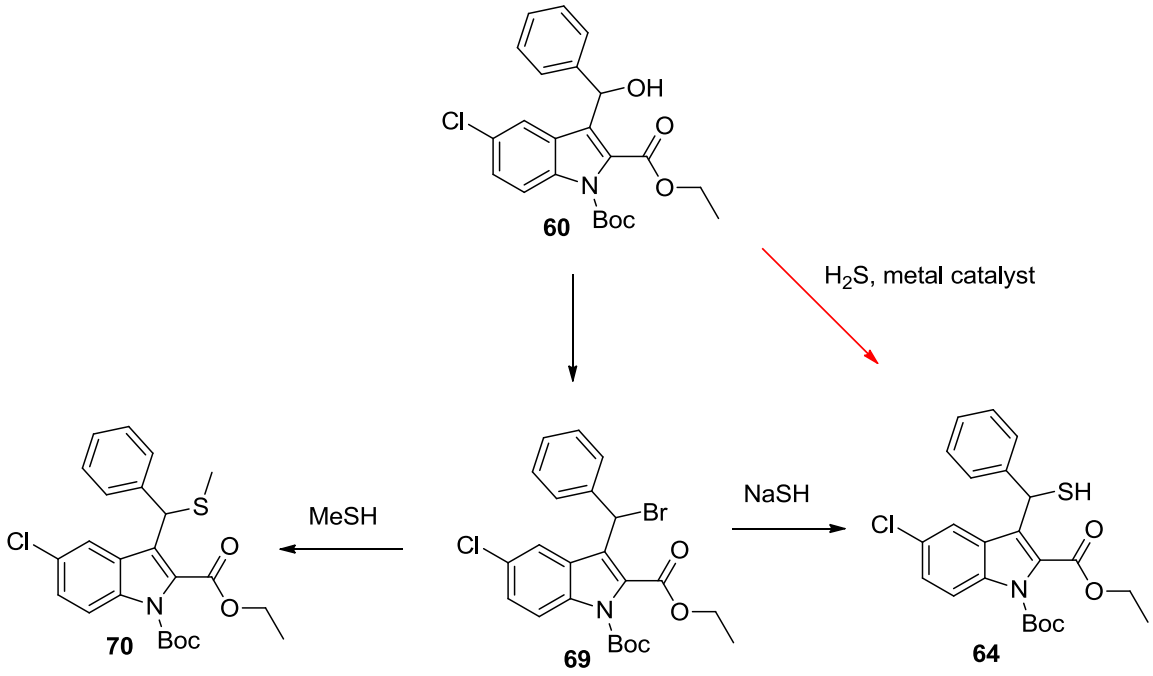
Scheme 22

Table 11



Compound	R_1	R_2	% Yield
64	H	H	49
65	Me	Me	49
66	Me	H	49
67	Cl	H	27
68	Br	H	-

Overall, the reaction did not give good yields (Table 11), nevertheless it was the more attractive route. Other options (Scheme 23) included converting a halogen into the thiol, but this would involve first converting the alcohol, such as for compound **60,** to a bromine **69** first and then using NaSH to produce the thiol **64**, adding an extra step to the synthesis. Using MeSH to give the thioether **70** directly from substitution of the bromine was not considered since MeSH is a gas and is not pleasant to work with, and likewise for the direct conversion of the alcohol **60** to the thiol **64** using H_2S with a metal catalyst.



Scheme 23

Lawesson's reagent was able to convert alcohol derivatives **60-63** to the thiols **64-67** successfully and we could continue with the next reaction. For the compound **54** we were unable to isolate compound **68**, and we discontinued working with this derivative. All other derivatives were characterised by IR, MS, NMR and melting point analysis. The presence of the thiol was confirmed the absence of a broad O – H stretch in the IR spectra and the presence of a band at about 2500 cm^{-1}, which corresponds to the frequency observed for thiols. In the ^{1}H NMR spectra there was a shift of the doublet, which was previously identified as the signal pertaining to the alcohol proton, to a more up field position, an indicator that the O – H had successfully been converted to an S – H. As a final confirmation, the sulfur atom was also detected by MS analysis where all calculated isotopic masses corresponded to the masses detected by MS.

3.15 Synthesis of the Boc protected sulfide derivatives

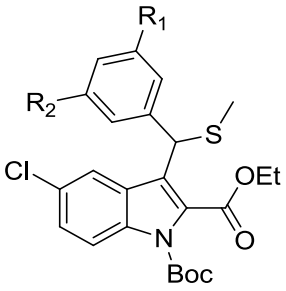
A final alteration to the functional group on the benzylic position had to be made, which was methylation of the sulfur (Scheme 24). This was accomplished by deprotonating the thiols **64-67** with NEt_3 and using MeI as a methylating agent. The reaction was performed at rt and was left overnight not necessarily due to the reaction taking this long, but on TLC the starting material and

product had the same R_f, so to ensure all starting material was consumed the reaction was left for much longer than necessary. The reaction was quenched with saturated ammonium chloride, and stirred for 2h to ensure all the MeI was reacted. The product was then extracted with ethyl acetate and purified by column chromatography.

Scheme 24

Yields were moderate to good, and we saw good consistency across the board for compounds **70**-**73** (Table 12), unlike for many of the other reactions. All derivatives were fully characterised. The signal for the methyl group on the sulfur was seen between 1.6 and 1.7 ppm in the ^{1}H NMR spectra and at between 45 and 50 ppm in the ^{13}C NMR spectra for the different derivatives.

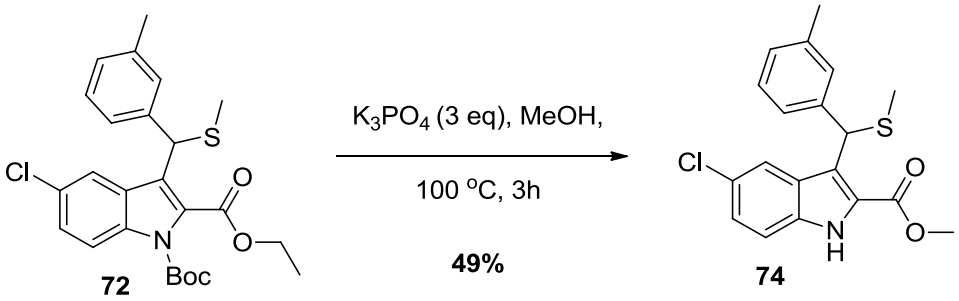
Table 12



Compound	R_1	R_2	% Yield
70	H	H	75
71	Me	Me	64
72	Me	H	83
73	Cl	H	84

3.16 Attempted synthesis of ethyl 5-chloro-3-((methylthio)(*m*-tolyl)methyl)-1H-indole-2-carboxylate **74**

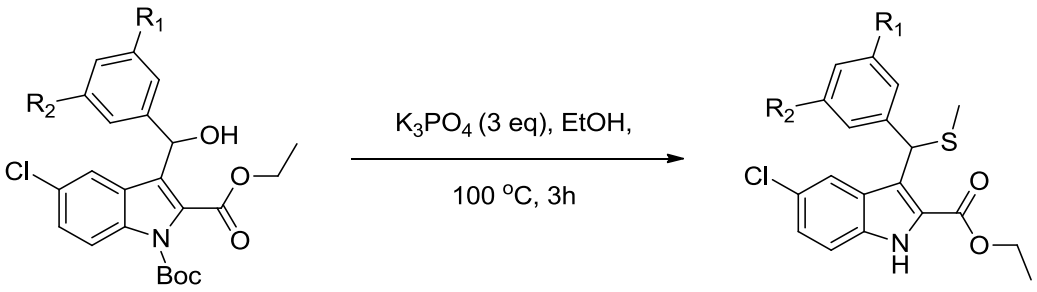
The final step in this synthetic strategy was the removal of the Boc protecting group from the indole nitrogen. We initially tried to use MeOH as a solvent for the deprotection of compound **72** (Scheme 25). This reaction was done by refluxing the indole in dry MeOH and K_3PO_4 at 100 °C. To our dismay, the Boc was removed but there was a second reaction which occurred on the molecule, which was believed to be a transesterification to the methyl ester, giving an overall yield of 49% of compound **75**. This compound was not fully characterised but the ^{1}H NMR spectrum showed a signal at 3.95 ppm integrating for 3 protons and there were no signals corresponding to the ethyl ester. The transesterification reaction, although a slight irritation in this regard, opened a new route for the formation of a variety esters in this position under mild reaction conditions and we would keep this in mind for future reactions.



Scheme 25

3.17 Synthesis of the deprotected sulfide derivatives

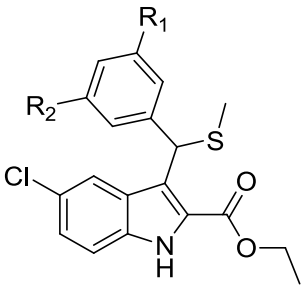
Since it was not the methyl ester which we required, but rather the ethyl ester, to maintain this functionality we had to change the solvent for the Boc deprotection to EtOH. The reaction was again carried out with K_3PO_4, this time with EtOH, and refluxed at 100 °C for 3h (Scheme 26).



Scheme 26

Yields were moderate for this reaction, nevertheless we had our sulfide derivatives **76-78** at hand and they could be sent for biological testing and we could test our hypothesis regarding these derivatives providing improved acid stability. All derivatives were characterised in full by IR, NMR, MS and melting point analysis.

Table 13



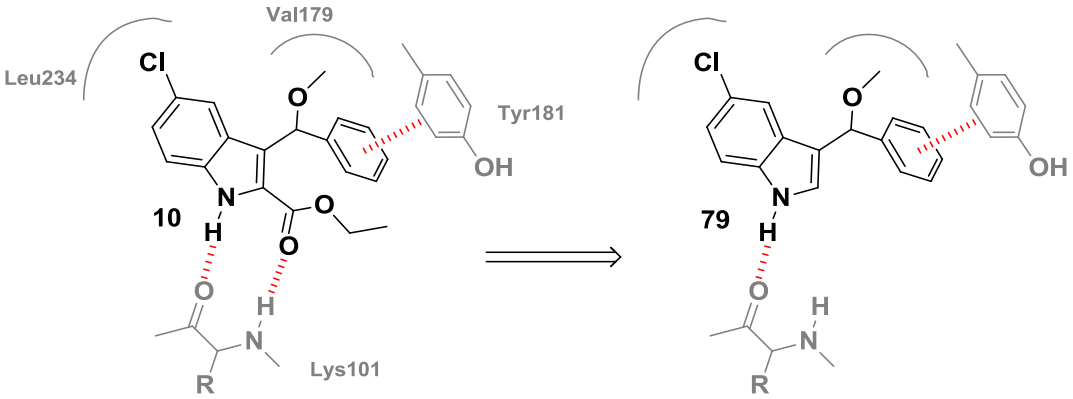
Compound	R_1	R_2	% Yield
76	H	H	59
77	Me	Me	81
74	Me	H	88
78	Cl	H	69

3.18 To what extent does the ester functionality contribute to the activity of the derivatives?

The ester functionality appeared to be essential for binding of the lead compound in the NNIBP. The ester has two interactions which contribute to the binding. First and most importantly, there is a hydrogen bonding interaction with Lys101. This is the first hydrogen bonding interaction with this residue, the indole N – H contributing the second interaction, both of which are strong

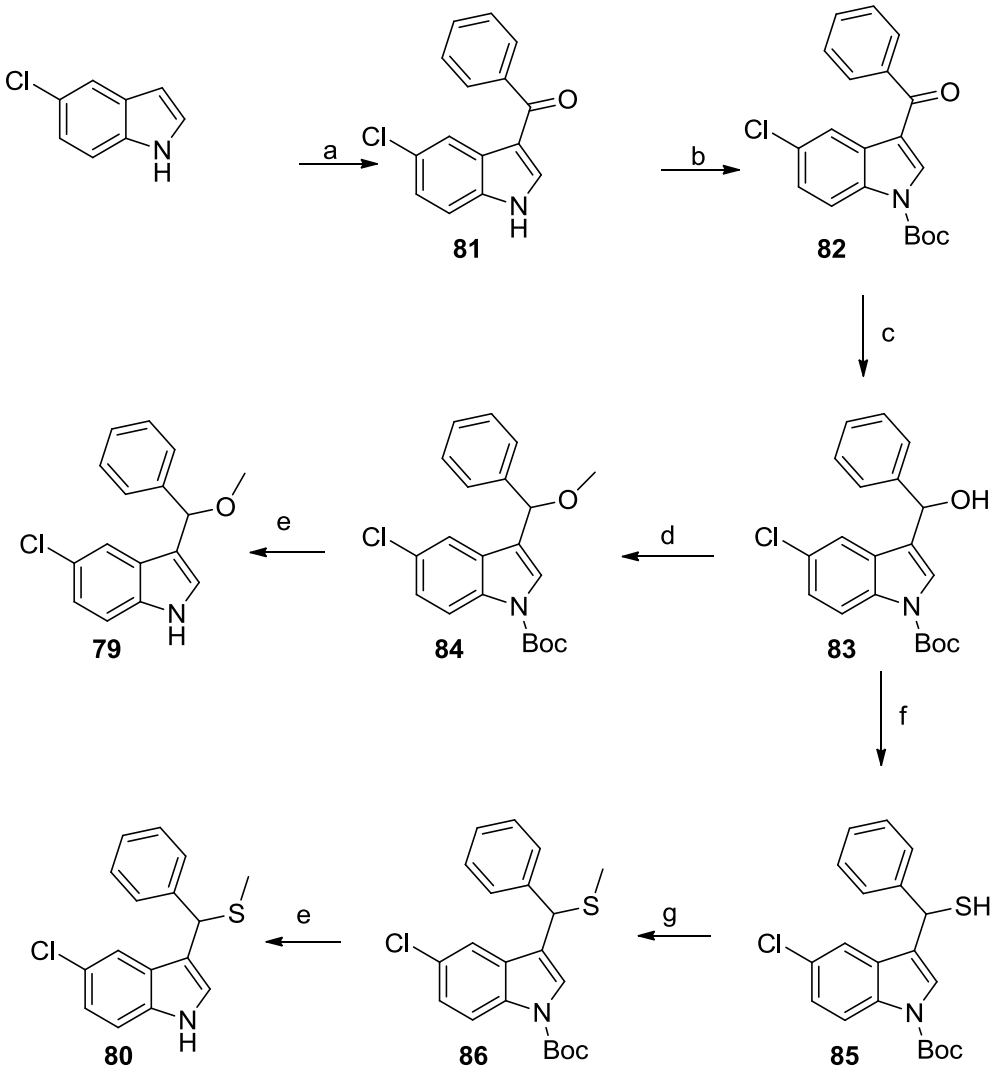
interactions which contribute greatly to the efficacy of these derivatives. Secondly, a more abstract interaction is thought to be the ester side chains interaction with the entrance to the NNIBP. The entrance is lined with water and polar residues and although the hydrophobic ester side chain appears to be a mismatched interaction, it contributes to the compounds ability to cross cell membranes which the carboxylate is unable to achieve.

To establish the extent to which ester contributed to the efficacy of the lead compound, we set out to synthesise the methyl ether derivative **79** and sulfide derivative **80**, lacking the ester moiety (Scheme 27). Without the ester moiety, the compounds would lack the key hydrogen bond interaction with Lys101, and therefore should be significantly less active.



Scheme 27

Due to the success of the synthesis of compound **21** (the bromo derivative), Boc was used as a protecting group for the synthesis of both the methyl ether and sulfide derivatives. This required starting out the synthesis with a Friedel-Crafts acylation to give compound **81**, this would be followed by a Boc protection to yield compound **82** and reduction of the ketone to an alcohol to obtain compound **83**. From here we branched off to focus on the synthesis towards the methyl ether derivative. This would be done by performing a substitution of the alcohol for a methyl ether using MeOH and *p*-TsOH to yield compound **84**. Finally, deprotection of this compound would give the final methyl ether compound **79**. For the sulfide derivative, the alcohol would be converted to a thiol using Lawesson's reagent to give compound **85**, methylation of the thiol would yield compound **86** and finally deprotection of the indole would yield the final sulfide derivative **80**.

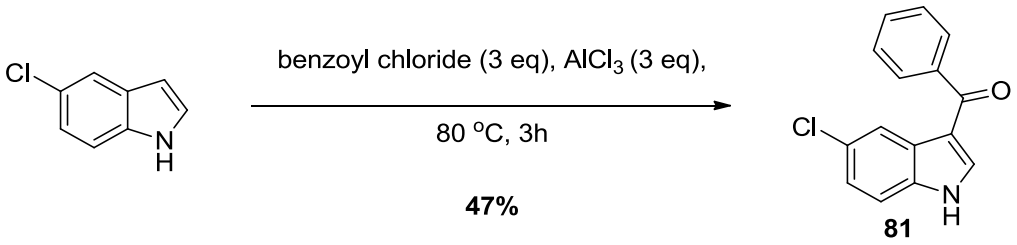


Scheme 28: Reagents and conditions: a) aryl acid chloride, $AlCl_3$, DCE, 2 h, 0°C - 85°C; b) Boc_2O, DMAP, THF, 3 h, r.t. c) $NaBH_4$, EtOH, THF, 3 h, 0°C - r.t. d) Lawesson's Reagent, toluene, 3 h, reflux e) MeI, Et_3N, DCM, 18 h, r.t. f) K_3PO_4, EtOH, 1 h, 70°C

3.18.1 Synthesis of (5-chloro-1H-indol-3-yl)(phenyl)methanone **81**

Friedel-Crafts acylation was once again used to introduce the acyl group to the 3-position of the ester. This was done in the same way as done previously, using benzoyl chloride as the acyl source and $AlCl_3$ as a Lewis acid to catalyse the reaction. The acylation gave compound **81** in 47% yield, which was lower than we expected since we anticipated much improved reactivity on the 3-position of the indole without the ester in the 2-position drawing electron density from the system. Compound **81** was fully characterised. The presence of the ketone was confirmed

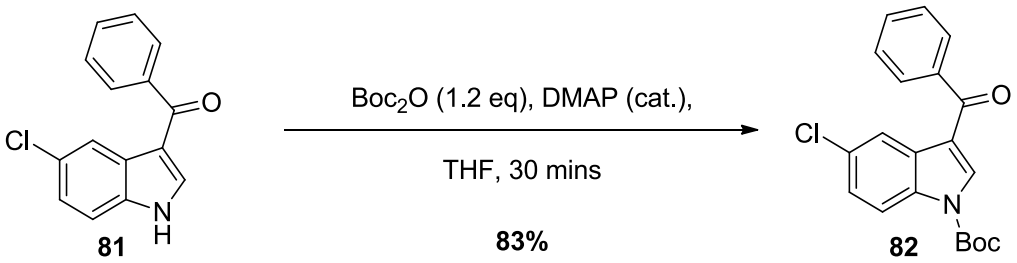
especially by the ^{13}C NMR spectrum which showed a carbonyl peak at 195 ppm. The IR spectrum also showed the characteristic C=O stretch at around 1700 cm^{-1}.



Scheme 29

3.18.2 Synthesis of *tert*-butyl 3-benzoyl-5-chloro-1H-indole-1-carboxylate 82

Boc protection of the compound **81** was achieved without any complications and compound **82** was synthesized in 83% yield (Scheme 30). Full characterization of this compound was performed and the peak at 1.68 ppm integrating for 9 protons, along with the absence of an N – H signal in the ^{1}H NMR spectrum served to confirm that the indole nitrogen had been successfully protected with a Boc group.

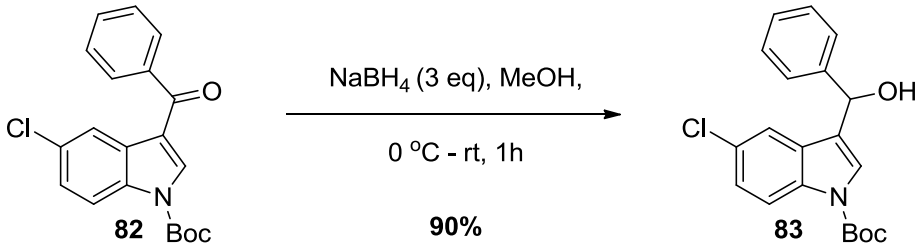


Scheme 30

3.18.3 Synthesis of *tert*-butyl 5-chloro-3-(hydroxy(phenyl)methyl)-1H-indole-1-carboxylate 83

Reduction of the ketone **82** was performed using $NaBH_4$ as a reducing agent. The reaction was completed after only 1h, testifying to the considerable effect the ester has on the reactivity of the indole. Following an acid workup, compound **83** was isolated by column chromatography in 90% yield. The compound was fully characterised and the benzylic C – H and O – H signals in the ^{1}H NMR spectrum confirmed reduction had been successful. Interestingly, we once again see only a

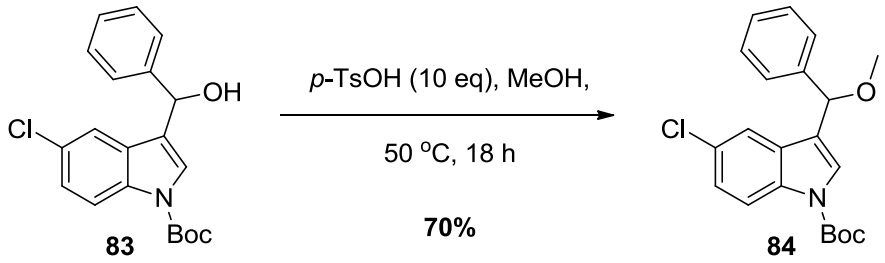
singlet for these signals and no coupling as for some of the other alcohol derivatives. A discussion regarding the differences observed for these signals is carried out in Section 3.19



Scheme 31

3.18.4 Synthesis of *tert*-butyl 5-chloro-3-(methoxy(phenyl)methyl)-1H-indole-1-carboxylate 84

The synthetic scheme now branched towards the synthesis of the methyl ether derivative. To install the methyl ether moiety itself we performed the substitution reaction using MeOH and *p*-TsOH to obtain compound **84**. Without the ester in place we saw a dramatic increase in the yield (70%) for this reaction. This was due to electron density no longer being withdrawn by the ester from the benzylic position, overall increasing the rate of this reaction. Compound **84** was fully characterised, and the presence of a signal in the ^{1}H NMR spectrum at 3.41 ppm integrating for 3 protons confirmed the formation of the methyl ether.

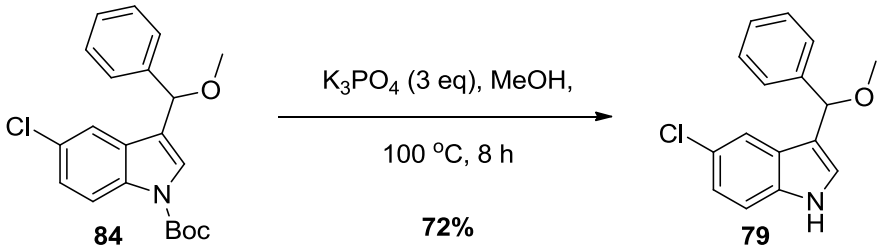


Scheme 32

3.18.5 Synthesis of 5-chloro-3-(methoxy(phenyl)methyl)-1H-indole 79

The final step to afford the methyl ether derivative **79** was a deprotection step. Removal of the Boc group from compound **84** was once again performed in K_3PO_4 and MeOH afforded compound **79** in 72% yield. This compound was fully characterised. All analysis confirmed that we had

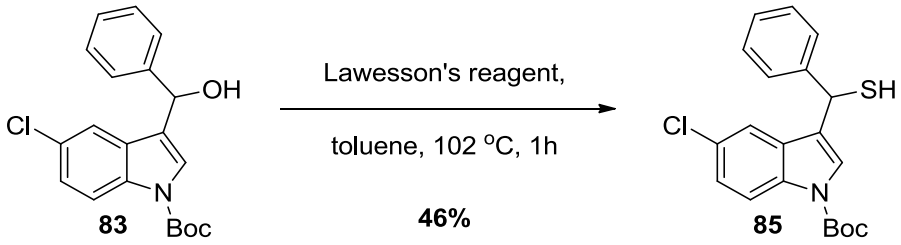
successfully synthesised compound **79** in an overall yield of 17%. This compound was sent to be tested for its efficacy against HIV-1.



Scheme 33

3.18.6 Synthesis of *tert*-butyl 5-chloro-3-(mercapto(phenyl)methyl)-1H-indole-1-carboxylate 85

For the conversion from the alcohol **83** to the thiol to produce compound **85** Lawesson's reagent was again utilized (Scheme 34). This reaction gave a 46% yield of compound **85**, much lower than would be expected, especially without the ester to interfere with the regioselectivity of the reaction. Unlike the ester bearing derivatives, which seemed to mostly reach completion for this reaction, 24% of the starting material was recovered during purification of compound **85**. The compound was fully characterised.

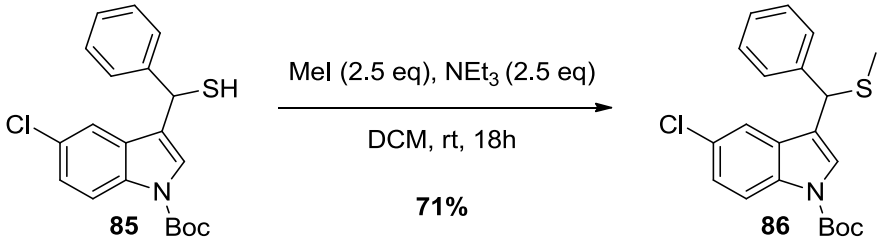


Scheme 34

3.18.7 Synthesis of *tert*-butyl 5-chloro-3-((methylthio)(phenyl)methyl)-1H-indole-1-carboxylate 86

To obtain the sulfide we needed to perform a methylation of the sulphur of compound **85**. This was once again done using MeI as a methylating agent and afforded compound **86** in 71% yield. Due to the starting material and the product having such similar R_f values, we found purification particularly difficult, and the purified product still had starting material, even though a reaction

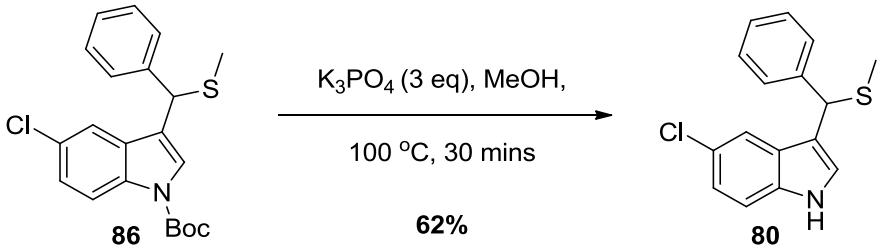
time of 18h was performed, which should have consumed all starting material. Compound **86** was fully characterised and the presence of the methyl peak on the sulfur was confirmed by a signal in the ^{1}H NMR spectrum at 1.65 ppm integrating for 3 protons.



Scheme 35

3.18.8 Synthesis of 5-chloro-3-((methylthio)(phenyl)methyl)-1H-indole 80

Finally, deprotection of the sulfide derivative **86** was carried out using K_3PO_4 in MeOH (Scheme 36). The yield for this reaction was 62%, making the overall yield for this synthesis 7%. Final compound **80** was fully characterised and was sent together with the other derivatives for biological testing.

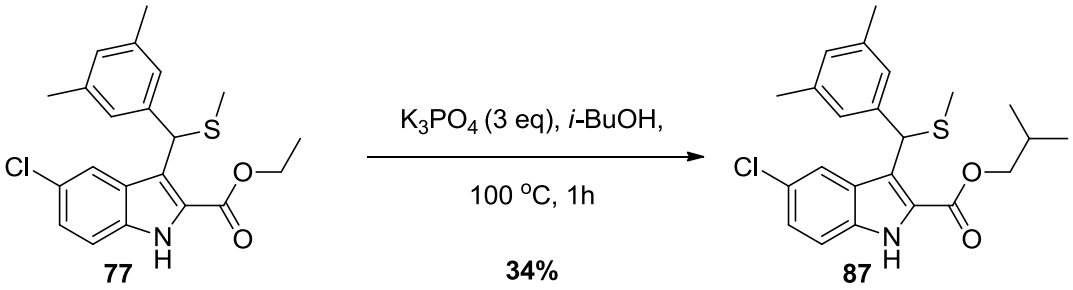


Scheme 36

3.18.9 Synthesis of isobutyl 5-chloro-3-((3,5-dimethylphenyl)(methylthio)methyl)-1H-indole-2-carboxylate 87

The transesterification in the presence of K_3PO_4 along with MeOH to give compound **75**, discussed earlier, presented us with the idea of using this reaction to convert the ethyl ester to a bulkier ester. For this we chose to look at an isobutyl ester to establish a boundary for chain lengthening and then we could synthesise the propyl, isopropyl and butyl esters later if this proved to be too large for the NNIBP. Since we had a fair amount of compound **77** available we used this sample to apply the transesterification from ethyl ester to isobutyl ester (Scheme 37). Compound **77** was

refluxed in isobutanol at 100 °C for 1h with K_3PO_4 to give two products of equal amounts. The desired product **87** was isolated in only 34% yield and was fully characterised to confirm the transesterification had been successful. The identity of the second product was not determined, however with compound **87** at hand it could be sent for biological testing.



Scheme 37

3.19 A more detailed discussion regarding the reactivity of these derivatives

Throughout the syntheses of all the different derivatives, it became apparent that there were some patterns emerging regarding the behaviour of the derivatives for different reactions. For the most part we see variations between the reactivities of the derivatives with different substituents on the phenyl ring. We also saw considerable differences for indoles bearing different protecting groups, in our case Boc and tosyl protecting groups. Finally, the presence of the ester also played a role in the reactivities. Since these differences were observed fairly consistently throughout the different reactions, a more thorough discussion is warranted.

3.19.1 The reactivities based on the substituents on the phenyl ring

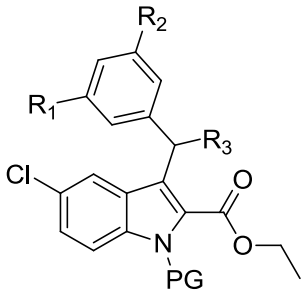


Figure 16

Looking at the effects of different substituents on the phenyl ring (Figure 16), we see that the unsubstituted phenyl derivatives have much the same yields as the 3- and 3,5- methyl substituted

phenyl derivatives for the reduction and substitution reactions. In contrast, we see a significant decrease in yields for most reactions when there is a halogen substituent in the 3-position on the ring.

For the reduction of the ketone to the alcohol, for example for the conversion of compound **38** to compound **12,** the ketone forms part of an extended conjugated system. This means that having an electron withdrawing (EW) substituent on the phenyl ring could quite easily increase the electrophilic nature of the carbon of the ketone by further pulling electron density away from this system making the carbon centre more reactive towards the hydride nucleophile. This would mean the presence of the EW substituent on the phenyl ring could increase the yields for the reduction. However, the opposite is seen which indicates that there are other factors in play. For example, the halogens may also act as electron donating groups using the Pi electrons on these atoms and this effect may create a less electrophilic carbonyl through resonance.

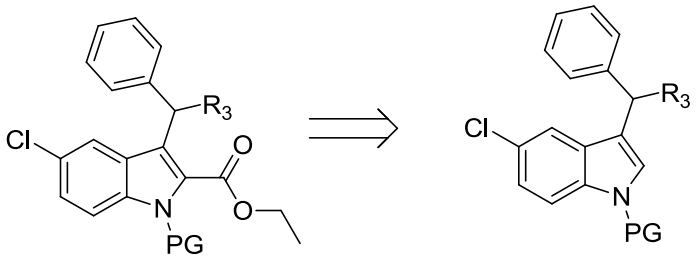
For the substitution reaction, we again see the same pattern of reactivity, with the halogen substituted derivatives giving much lower yields than the other derivatives. We propose that the electron donating (ED) substituents on the phenyl ring result in there being slightly more electron density on the benzylic carbon which makes this carbon more susceptible to the formation of the 'pseudo carbocation', which initiates the S_N1 reaction. The indole nitrogen's participation in this reaction further assists the elimination of the alcohol and since this is facilitated by the lone pair of electrons on the indole nitrogen, any groups having an EW effect on the indole, such as the ester in the 2-position, hinder the reaction. Finally looking at the Lawesson's reaction, we again see the ED substituents on the phenyl ring resulted in higher yields than the EW substituents. Since the mechanism for this reaction also involves breaking the C – O bond we would expect similar reactivity to the substitution, which is exactly what is observed.

3.19.2 Reactivities based on the protecting groups

Comparing the reactivities of the indoles protected with different groups, we notice that for the tosyl protected indoles we often get lower yields for several reactions compared to the Boc protected indole. This is seen for the reduction, which again may be due to resonance effects as discussed above. Unfortunately, we cannot compare any other reactions in this regard since only the bromo derivative underwent the substitution with both a tosyl and a Boc group and this was not a direct comparison either since we had deprotection and substitution occurring in a one pot

reaction for the Boc protected derivative. Nevertheless, if we were to compare these we would most likely find that the substitution from the alcohol to the methyl ether moiety occurs in elevated yields for Boc protected derivatives.

3.18.3 Reactivities based on the presence of an ester group in the 2-position of the indole



Scheme 38

The lack of the electron withdrawing ester group in the 2-position had a considerable effect on the substitution reaction in particular (Scheme 38). It was clear to see that substitution of the alcohol group **83** for a methyl ether **84** was much more efficient without the ester and this was seen by the sizeable decrease in reaction time and an increase in yield for the reaction. This of course also means that having no group in the 2-position we have an even more unstable compound, which would have increased susceptibility to acid degradation.

3.19.3 A closer look at the reasons for inconsistencies in the signals for the O – H and benzylic C – H signals of the alcohol derivatives as observed by ^{1}H NMR

When comparing the ^{1}H NMR spectra for the tosyl protected and Boc protected alcohol derivatives we see a major difference between the two. For the tosyl protected derivatives we observe a broad signal for the alcohol which indicates a fair amount of chemical exchange of the alcohol proton is occurring. We also only see a singlet for the benzylic C – H. For the Boc protected derivatives we see a sharp doublet, which have a coupling constants of around 5.4 Hz, matching a doublet corresponding to the benzylic C – H. An explanation for this observation, could be the electron withdrawing tosyl group or ester resulting in a more acidic O – H proton which can undergo more chemical exchange of this proton with the deuterated solvent.

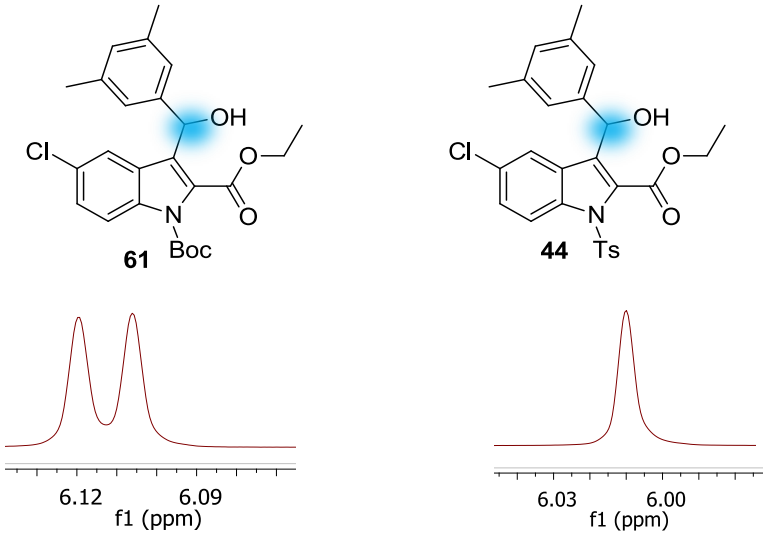
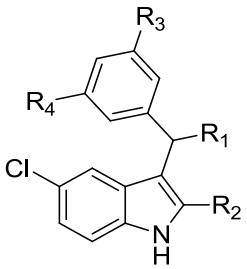


*Figure 17: Signals in the ^{1}H NMR spectra for the benzylic C – H proton (highlight in blue) seen for the tosyl protected derivative **44** and Boc protected derivative **61***

3.20 Biological evaluation of the compound series

Having synthesised a small library of compounds, we could send these for biological testing to establish their efficacy against both wild-type and mutant strains of HIV-1. The biological testing was conducted by collaborators at the National Institute for Communicable diseases (NICD) located in Johannesburg, South Africa. The procedure for the testing involves an *in vitro* single-cycle, non-replicative phenotypic assay and a toxicity assay. A vector system isolated from HIV-1 was utilised to produce virus-like particles (VPLs) which were incubated with the test compounds along with 293T cells, kidney cells manipulated to express T antigens, for 48h.[50] To assess the inhibition of HIV-1, luminescence measurements were conducted.[50] The results are reported Table 14 as IC_{50} and CC_{50} values. IC_{50} is the minimum concentration of the test compound which results in 50% inhibition of a specific response, in this case growth of VPLs. The CC_{50} value is minimum concentration which results in 50% cell viability, in this assay this would mean the viability of the 293T cells, which translates to the toxicity value. The results from the previous docking studies are included in Table 14, to allow for comparison to the assay results.

Table 14: Results for biological assay against wild-type HIV-1



Compound	R_1	R_2	R_3	R_4	CDocker Energy	IC_{50} /μM	CC_{50} /μM
Nevirapine						0.138	>6
10	OMe	CO_2Et	H	H	-49.8	0.016	26.7
13	OMe	CO_2Et	Me	Me	-51.8	0.034	35.9
19	OMe	CO_2Et	Me	H	-53.1	0.020	27.4
20	OMe	CO_2Et	Cl	H	-51.1	0.033	26.8
21	OMe	CO_2Et	Br	H	-50.6	0.020	26.4
79	OMe	H	H	H	-36.6	2.516	48.0
76	SMe	CO_2Et	H	H	-43.7	0.039	26.1
77	SMe	CO_2Et	Me	Me	-41.5	0.060	32.0
74	SMe	CO_2Et	Me	H	-44.5	0.038	29.6
78	SMe	CO_2Et	Cl	H	-41.9	0.042	28.6
80	SMe	H	H	H	-32.5	9.95	52.7
87	SMe	CO_2i-Bu	Me	Me	*i*-Bu too large for pocket, did not dock in correct pose	0.408	49.1

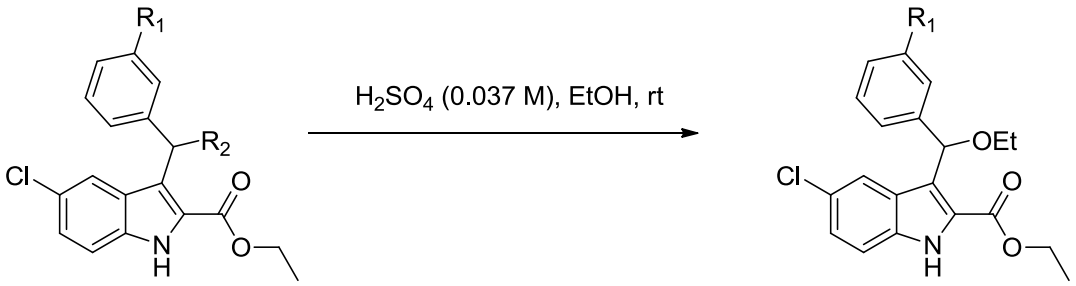
A more negative CDocker energy value indicates a better binding affinity for the pocket, and thus we would expect these compounds to have better efficacy results. As can be seen, the docking results correctly predicted that the sulfide derivatives would have slightly lower efficacy, probably due to the larger sulfur atom not being as well accommodated in the Val179 pocket. Although the sulfide derivatives are still ten times more potent than nevirapine, the loss of potency suggests that introducing a smaller atom in place of the oxygen, such as a carbon, might improve potency, or at least maintain similar potency to the methyl ether derivatives. In fact, introducing an ethyl chain in place of the methyl ether would completely eliminate the stability problem, and this is investigated in Chapter 4. Comparing the efficacies for the derivatives with different substituents on the phenyl ring, we see that the unsubstituted derivatives 10 and **76** still performed best against wild-type HIV-1. Having substituents on the phenyl ring slightly lowered the potency of the derivatives, although mutant studies may prove that they have better resistance profiles and hence are more advantageous overall. The derivatives lacking the ester moiety, compounds **79** and **80**, proved to be one hundred times less potent than the ester bearing derivatives, which confirms the importance of the ester for maintaining potency. The formation of the second hydrogen bond with Lys101 is therefore necessary for potency. The isobutyl ester derivative **87** showed only a ten times loss of potency, which may indicate that although the isobutyl group is too large there may be some room to accommodate a slightly longer ester side chain such as a propyl, isopropyl ester or even a butyl ester since the presence of two bulky methyl groups on the end of the isobutyl could also be responsible for loss in potency. The maintenance of some amount of activity despite the large isobutyl group again emphasises the importance of the ester carbonyl moiety for binding.

All compounds, excluding **79**, **80** and **87**, were sent for mutant studies. Unfortunately, we have yet to receive these results and thus cannot yet discuss how introducing substituents on the phenyl ring has affected the activity again mutant strains of the virus.

Overall, we had managed to develop several potent inhibitors of HIV-1, and these results also confirmed the accuracy of our binding model used for modelling. A reliable binding model can be invaluable in the design of inhibitors and can save a huge amount of time that would have gone in to synthesising inactive compounds.

3.21 Stability Testing

Having established that the sulfide derivatives **74** and **76-78** had good activity against wild-type HIV, we now needed to confirm if they were indeed more stable than their methyl ether analogues **10**, **13** and **19-21**. To this end, we chose to set up an experiment using three derivatives which offered good comparisons, compounds **19**, **21** and **74**. The experiment was done by setting up three identical reactions. 0.103 mmol of each compound was dissolved in exactly 10 mL of EtOH in three separate 20 mL flasks (Scheme 39). Since the reactions had to be set up at the Central Analytical Facility (CAF), part of Stellenbosch University, we did not have access to nitrogen to keep air and moisture out of the reactions; however, this did not appear to hinder the reactions. At t_0 we added 20 µL of concentrated sulfuric acid to each flask, translating to a concentration of 0.037 M H_2SO_4, giving a pH of around 1.4. Although this did not correctly mimic the conditions of the stomach, we were more interested in establishing the relative stabilities of these compounds to confirm that we had increased the acid stability by introducing the sulfide moiety. The presence of ethanol meant that the derivatives, **19**, **21** and **74** were converted to ethyl ether derivatives, **88** and **89**.



19 R_1 = Me, R_2 = OMe
21 R_1 = Br, R_2 = OMe
74 R_1 = Me, R_2 = SMe

88 R_1 = Me
89 R_1 = Br

Scheme 39

Once the reactions were set up, aliquots of 200 µL were taken from each reaction mixture at set time intervals and diluted with 400 µL of MeCN and 400 µL of milli-Q water. The samples were then injected into a Waters 1525 Binary HPLC coupled to a Waters 2487 Dual λ absorbance detector. The run time was 8 minutes using a split volume of 0.65 ACN/0.15 water as the eluent.

The results from the HPLC analysis are shown in Figure 18. The sulfide derivative **74** showed no conversion to the ethyl ether after 4h, whereas the methyl ether counterpart **19** showed

complete conversion after just 2h. This was a very exciting result for us, and showed that using moiety which is a worse Lewis base than the methyl ether moiety was a successful way of improving the acid stability of the compound.

The bromo derivative **21** also showed enhanced stability compared to compound 19, which was a particularly interesting result and again emphasises the large impact remote groups can have on the reactivity of different compounds. Nevertheless, there was some conversion of compound **21** to the ethyl ether so overall the sulfide was more successful in improving the acid lability of the lead compound. This also meant that our hypothesis regarding the replacement of the methyl ether for a worse Lewis base to increase the acid stability was indeed correct.

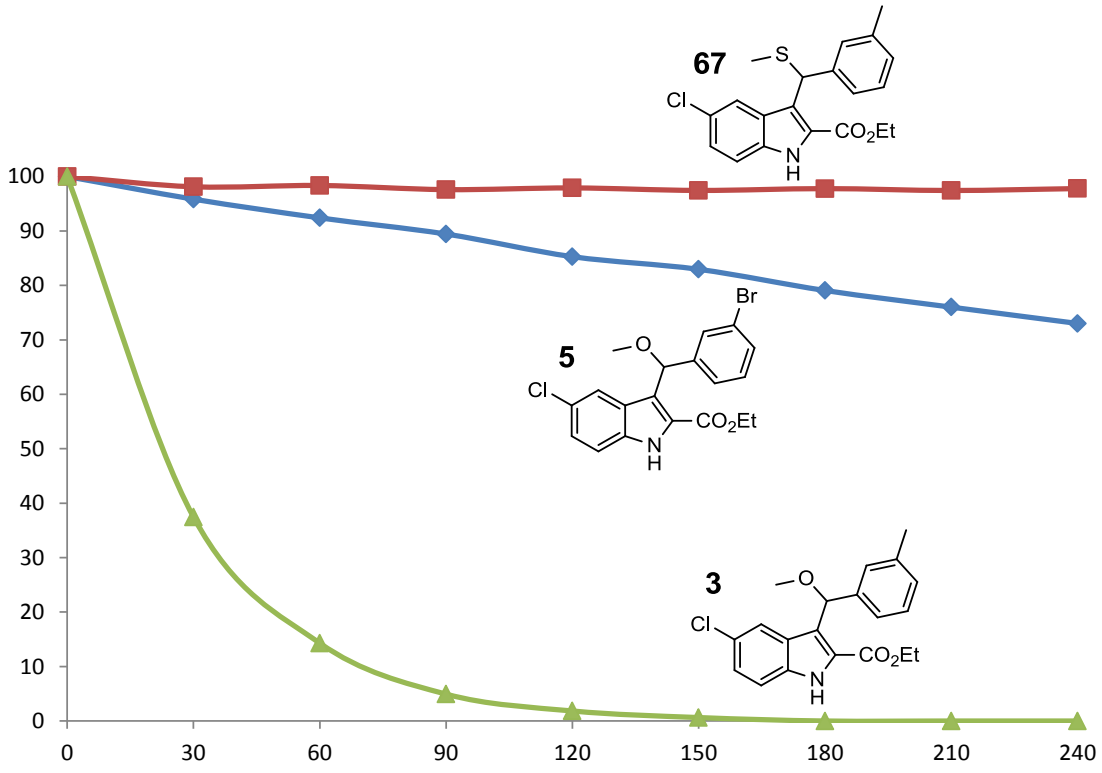


Figure 18: The graph shows the %conversion to the ethyl ether derivative vs. time

3.22 Concluding remark pertaining to Chapter 3

The work done in this chapter was based on two aims of the project; to improve the acid stability and to improve the resistance profiles of these compounds. Introducing substituents on the 3 and/or 5 position of the phenyl ring of the lead compound was, according to literature and previous research in our group, meant to improve the resistance profiles of these compounds. For

this aim we were able to successfully synthesise five methyl ether derivatives and four sulfide derivatives, of which seven were novel, to test for activity against wild-type HIV-1 and several mutant strains. Unfortunately, despite their successful synthesis and good potency against HIV-1 we have not yet received the results of the mutant studies; thus, we have not fulfilled this aim as yet.

The aim to improve on the acid stability of the compounds was very successful and we showed by HPLC analysis that the sulfide derivative **74** could withstand more than 4h in an acidic environment without degrading. Compared to the methyl ether derivative **19** this was a major triumph, since we had seen complete degradation after just 2h for this compound. The increase in stability of the bromo derivative **21** was an interesting result and serves to emphasise what a huge impact remote groups have on the reactivity of a molecule. Although the sulfide derivatives were more potent than nevirapine, we did see a slight loss of potency compared to the methyl ether derivatives; however, the added advantage of stability in acidic environment still gives these derivatives the edge over the original lead compound **10**.

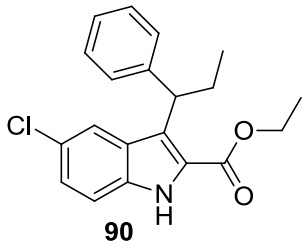
Finally, the investigation on the extent to which the ester moiety contributes to binding was also successful. The derivatives lacking the ester moiety showed a hundred times loss in potency, proving the need for the second hydrogen bond with Lys101 and the ester for the activity of these compounds. The isobutyl ester derivative showed some maintenance of potency; however, the isobutyl group was clearly not well accommodated in the NNIBP.

Overall from a synthetic and biological perspective this work was a major success and offered an improvement on the lead compound.

Chapter 4: Eliminating the problem with acid stability by replacement of the methyl ether moiety for an ethyl group

4.1 The successes and failures of the sulfide derivatives and an introduction to the ethyl derivative

Although the sulfide derivatives proved to have much better stability than the methyl ether derivatives, their anti-HIV activity was slightly reduced due to the larger sulfur atom not being accommodated as comfortably in the Val179 pocket. Furthermore, despite improvement in the stability, the problem was not completely eradicated which would be a more ideal situation. Replacement of the methyl ether for an ethyl group, to give ethyl derivative **90** (Figure 19)**,** could potentially be a much more prudent choice. The carbon atom is smaller than either the oxygen or sulfur, therefore we could be confident that it would be accommodated in the Val179 pocket.

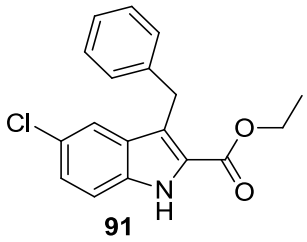


90

Figure 19

Calculated binding energies for the methyl ether derivative **10** (-71.6 kcal/mol) and ethyl derivative **90** (-66.3 kcal/mol) indicated that the ethyl derivative would not bind quite as effectively which may be reflected in the efficacy results.[14] The modelling results suggested that the ethyl group may be too small to maintain as good van der Waals interactions as the methyl ether moiety has within the Val179 pocket. Efficacy testing previously done in our group revealed that a derivative lacking a group in the benzylic position, such as compound **91 (**Figure 20), resulted in up to ten times reduction in potency, emphasising the need for a small group in this position to occupy the Val179 pocket, and we may have found our threshold for the size of this group when we designed the methyl ether derivative **10**.[20] Despite the possibility that ethyl

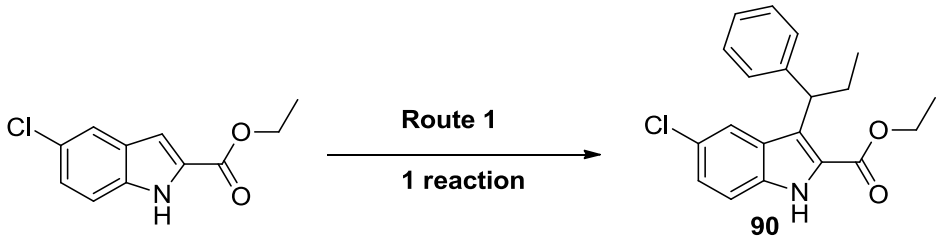
derivative **90** may be less active than the original lead compound **10**, it offers a very important advantage over both the methyl ether and sulfide derivatives – it is not possible for the ethyl group to be eliminated and then substituted under conditions, such as an aqueous environment, that led to the production of an alcohol in this position in other derivatives.



91

Figure 20

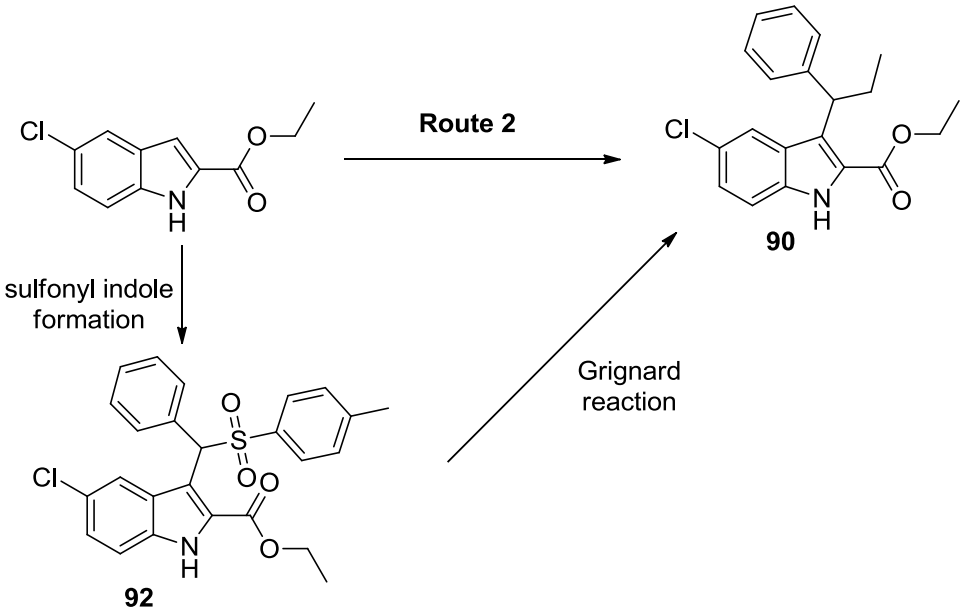
After scouring the literature for possible ways of introducing a propylbenzene group into the 3-position of the indole to synthesise compound **90**, several possible strategies arose. The first envisaged route was introducing a propylbenzene group directly to the ethyl-5-chloro-*1H*-indole-2-carboxylate (Scheme 40). This was a more attractive route since it allowed the group to be introduced into the 3-position of the indole without any further requirement for modification and only involved one step. A review of the literature revealed many different procedures which had been used to install alkylbenzene groups onto the 3-position of an indole directly and we attempted several these procedures on our system as will be discussed below.



90

Scheme 40

Another route, inspired by literature, involved synthesis of 3-sulfonyl indole **92** followed by substitution of the sulfonyl group for an alkyl group using a Grignard reaction to give compound **90**.[51,52] This was also a desirable route since it only involved two steps to reach the desired product and we envisaged that it may work with the ester in place if we kept the equivalents of the Grignard reagent low.

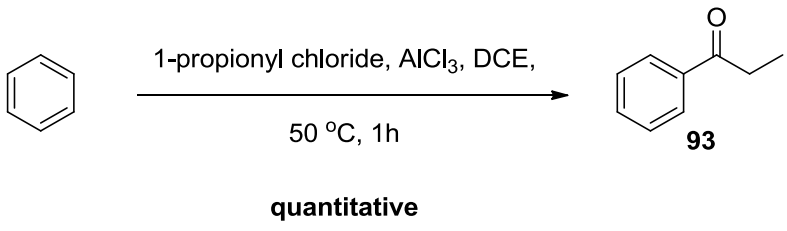


Scheme 41

4.2 Employing route 2 – Direct introduction of the propylbenzene group to the 3-position of the indole

4.2.1 Synthesis of 1-propiophenone **93**

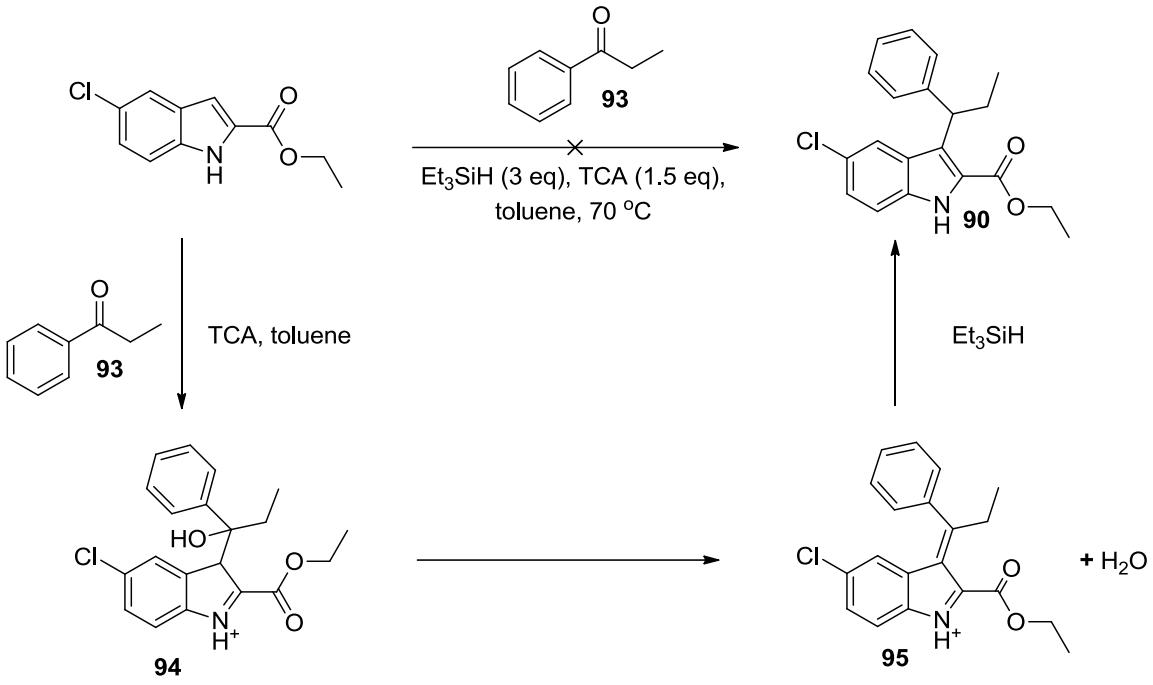
A few papers outlined a procedure using 1-alkylphenone derivatives to introduce the corresponding alkylbenzene group to the 3-position of the indole. To test these different procedures the indole starting material, we first needed to synthesise 1-propiophenone. This was achieved by a Friedel-Crafts acylation using benzene and 1-propionyl chloride. After a basic workup, the product was extracted with EtOAc and concentrated *in vacuo to* afford a yellow oil which was isolated in quantitative yield. Confirmation that we had obtained compound **93** was by means of ^{1}H and ^{13}C NMR spectral analysis and all signals corresponded to the desired compound.



Scheme 42

4.2.2 Attempted synthesis of ethyl 5-chloro-3-(1-phenylpropyl)-1H-indole-2-carboxylate 90 using reductive elimination with 1-propiophenone

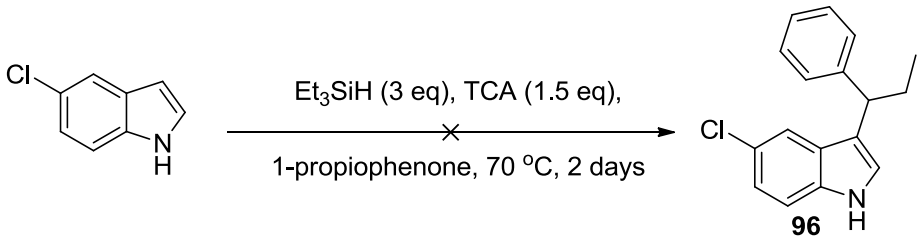
A paper by Rizzo *et al* (2008) described a procedure for introducing alkylbenzene groups onto the 3-position of an indole by reductive elimination with a ketone as an electrophile.[53] The reaction involved stirring the indole with the corresponding ketone, trichloroacetic acid (TCA) and triethylsilane in toluene (Scheme 43). In the paper, they investigated the reaction on unsubstituted and various substituted indoles. The reaction was most effective for a 5-bromo indole using cyclopentanone, which was promising for our indole system with the chloro group in the 5-position. Additionally, the reaction had been successful for introducing an ethylbenzene group onto the 3-position, again similar to the propylbenzene group we wanted to introduce. The mechanism proposed by Rizzo *et al* is shown in Scheme 43. Protonation of the carbonyl oxygen of the ketone of compound **93** activates the carbonyl for nucleophilic attack from the indole to yield intermediate **94**. The alcohol which is formed is then protonated to form a good leaving group, which is eliminated to afford intermediate **95**. Finally, Et_3SiH is used to reduce the indole and aromaticity is regained to form the desired compound **90**. Our concern regarding the success of this procedure was the presence of the ester group in the 2-position of the indole starting material. The electron withdrawing ester makes the 3-position less reactive towards electrophiles. Nevertheless, we first attempted the reaction on ethyl 5-chloro-*1H*-indolecarboxylate. The flask was charged with toluene followed by TCA and triethylsilane. The indole and 1-propiophenone were dissolved in toluene and introduced by dropping funnel at rt. Once everything was added the flask was heated to 70 °C. After two days of heating no new spots were observed on TLC and we reasoned that the ester was preventing the reaction proceeding.



Scheme 43: Reaction procedure and mechanism using reductive elimination with ketones to introduce a group to the 3-position of the indole from Rizzo et al, 2008[53]

4.2.3 Attempted synthesis of 5-chloro-3-(1-phenylpropyl)-1H-indole 96 using reductive elimination with 1-propiophenone

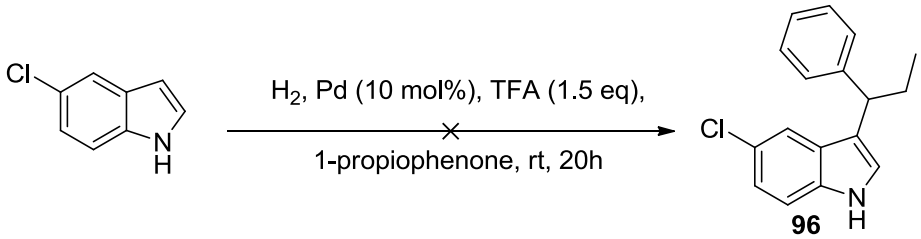
Since no acylation was occurring when there was an ester in the 2-position of the indole we attempted the reaction again under identical conditions but now using 5-chloroindole. Again, after two days of heating we did not observe any new product being formed. We imagined that the more electronegative chloro group in the 5-position of the indole may be hindering the reaction, since the literature procedure had used a bromo group and had been successful using a slightly different ketone. Since there were no other products which seemed to form, as seen by TLC analysis, we assumed that the nucleophilic attack was not occurring at all.



Scheme 44

4.2.4 Attempted synthesis of 5-chloro-3-(1-phenylpropyl)-1H-indole **96** using reductive elimination with 1-propiophenone and H_2 with Pd/C as a reducing agent

A similar procedure, used by Cao *et al* (2011), utilised H_2 and Pd/C (Scheme 45) to perform the final reduction step in the mechanism (Scheme 43) and replaced trichloroacetic acid with the more acidic trifluoroacetic acid.[54] In literature the procedure worked very well for introducing a variety of different benzylic groups into the 3-position. However, the paper had utilised 2-substituted indoles, with either methyl or phenyl groups in the 2-position. Despite this red flag, we thought that we could still attempt the reaction on our indole system. The 5-chloroindole and 1-propiophenone were dissolved in DCM and TFA was added along with Pd/C. The flask was placed under a hydrogen atmosphere and left at rt. Unfortunately, after 20h there was still no product observed. It seemed that without any electron donating substituents on the indole, using 1-propiopehenone as an electrophile was not a viable route for installing the propylbenzene group.

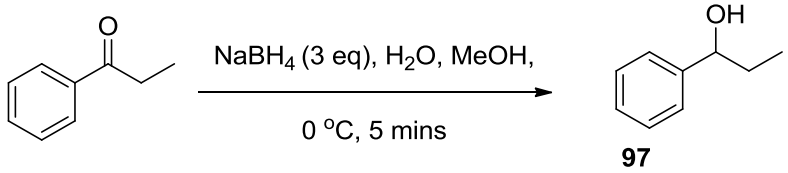


Scheme 45

4.2.3 Synthesis of 1-phenylpropan-1-ol **97**

Since all attempts to make compounds **90** or **96** had been unsuccessful using 1-propiophenone as an electrophile, we had to return to the literature to find a different route. Literature research revealed that an alcohol, along with I_2 to activate this species as an electrophile could be utilised to install alkylbenzene groups onto the 3-position of the indole.[55] To this end, we set out to synthesise 1-phenyl-propan-1-ol, compound **97**. Our first attempt at making this compound involved dissolving 1-propiophenone in MeOH and adding $NaBH_4$ at 0°C, then allowing the mixture to reach rt. The reaction was not complete after 18h and only 270 mg was synthesised from 500 mg of starting material. In an attempt to optimise this reaction we tried it again, but this time we first dissolved the $NaBH_4$ in distilled water before adding it dropwise to the 1-propiophenone dissolved in MeOH at 0 °C. This time the reaction was complete after just 5 mins

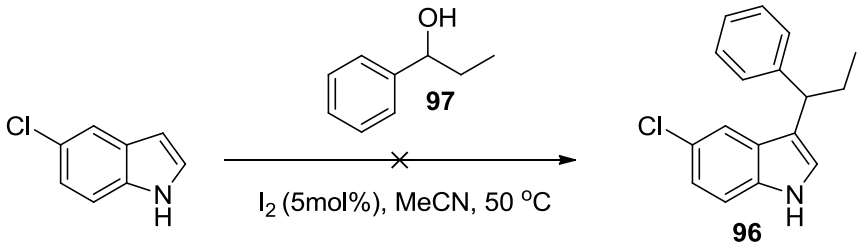
and gave an 80% yield. ^{1}H and ^{13}C NMR spectra with confirmed we had successfully synthesised compound **97**.



Scheme 46

4.2.4 Attempted synthesis of 5-chloro-3-(1-phenylpropyl)-1H-indole 96 by reductive elimination using 1-phenylpropan-1-ol

A reaction procedure outlined by Srihari *et* al (2008) described substitution on the 3-position of the indole using I_2 to activate the alcohol as a nucleophile and as a good leaving group for the elimination step (Scheme 47).[55] The procedure had been successful for introducing alkylbenzene groups to the 3-position of the indole, however this procedure had worked best when there were *para* substituents on the phenyl ring. Despite concerns that our indole system had no substitution on the phenyl ring the reaction was still attempted on the indole starting material (Scheme 47). Unfortunately, once again this reaction would not proceed. After a full 5 days at 50 °C, in the presence of compound **97** and I_2 in MeCN, there was only starting material present.



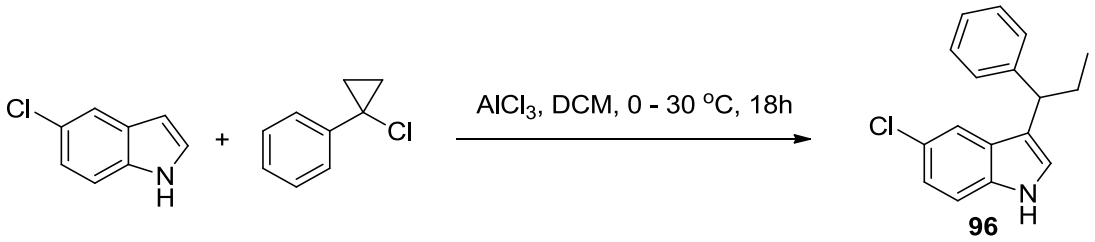
Scheme 47

4.3 Direct introduction of the propylbenzene group onto the 3-position of the indole by Friedel-Crafts alkylation

4.3.1 Inspiration from previous research

Prior to the discovery of the very active lead compound 10, our group was investigating cyclopropyl derivatives and during an attempt to introduce this moiety to the 3-position of the

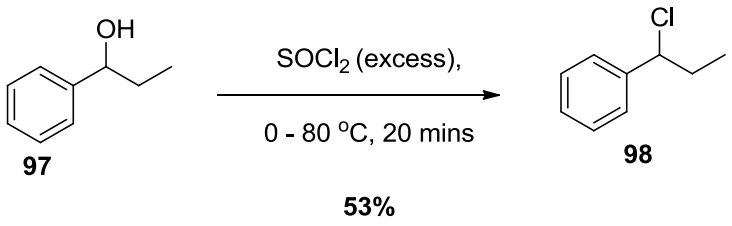
indole by a Friedel-Crafts alkylation, the cyclopropyl group cleaved and gave the undesired, at the time, ethyl derivative **96** (Scheme 48). This gave precedence for this derivative to be obtained by a Friedel-Crafts alkylation, and we hoped that we could optimise this reaction using (1-chloropropyl)benzene, compound **98**, to give a moderate yield.



Scheme 48: Attempted synthesis of the cyclopropyl derivative resulted in the formation of the ethyl derivative by Müller, 2013

4.3.1 Synthesis of (1-chloropropyl)benzene 98

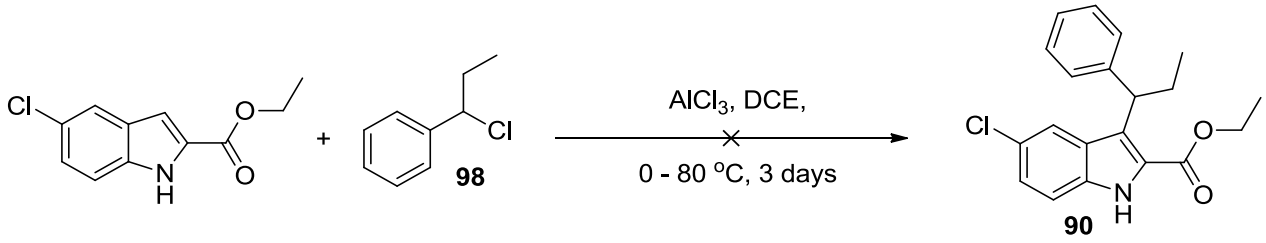
In order to carry out the Friedel-Crafts alkylation it was first necessary to synthesise (1-chloropropyl)benzene **98** (Scheme 49) by using a standard literature procedure to obtain this product.[56] Since we already had 1-phenylpropan-1-ol **97** at hand, we were able to convert this to compound **98** by dropwise addition of thionyl chloride to the alcohol **97** at 0°C, followed by refluxing at 80 °C. Once cooled, the reaction was neutralised with ice and saturated sodium bicarbonate and the product was extracted with ether and purified by column chromatography to give a 53% yield of compound **98**. ^{1}H and ^{13}C NMR spectral analysis showed all peaks corresponded to the desired product.



Scheme 49

4.3.2 Attempted synthesis of ethyl 5-chloro-3-(1-phenylpropyl)-1H-indole-2-carboxylate 90 by Friedel-Crafts alkylation

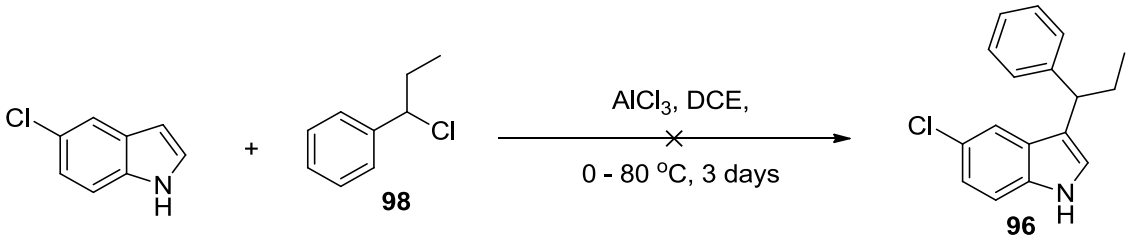
Since we had previously used a Friedel-Crafts acylation to install an acyl group onto the 3-position of ethyl 5-chloro-*1H*-indolecarboxylate, we were fairly confident this could also work for a Friedel-Crafts alkylation. The reaction was performed in DCE, instead of DCM which was used for the procedure by Müller (2013), to allow for a higher refluxing temperature. Compound **98** was added to the flask containing DCE followed by aluminium chloride at 0°C. After stirring for 30 mins at 0 °C, the indole was introduced and the reaction mixture was heated to 80 °C. After refluxing the reaction for 3 days there was still no product seen on TLC and we abandoned the idea of performing this reaction with the ester already in place.



Scheme 50

4.3.3 Attempted synthesis of 5-chloro-3-(1-phenylpropyl)-1H-indole 96 by Friedel-Crafts acylation

Since the Friedel-Crafts alkylation had been unsuccessful for an indole bearing an ester in the 2-position, we then attempted the same reaction using 5-chloroindole as our starting material. Again after 3 days of refluxing, no product was observed.



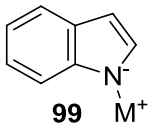
Scheme 51

Our inability to afford any product explained why there was no precedence in literature for this classical Friedel-Crafts procedure. Previous groups had successfully performed modified Friedel-

Crafts alkylations using catalysts other than $AlCl_3$, including copper, zinc, iron, scantium based catalysts and even organocatalysts.[51] Unfortunately, there was still no precedence for these reactions working on an indole with an ester in the 2-position and these catalysts were expensive, thus we would only try these if other routes had been exhausted.

4.4 Attempted synthesis of 5-chloro-3-(1-phenylpropyl)-1H-indole 96 using reductive elimination with (1-chloropropyl)benzene and an activated magnesium indole

Having failed so far to produce any desired product at all, we turned to activating the indole as a better nucleophile. It has been well established that the indolyl anion **99** (Scheme 52), with its metal counter-ion, is a better nucleophile than the neutral indole species.[57]



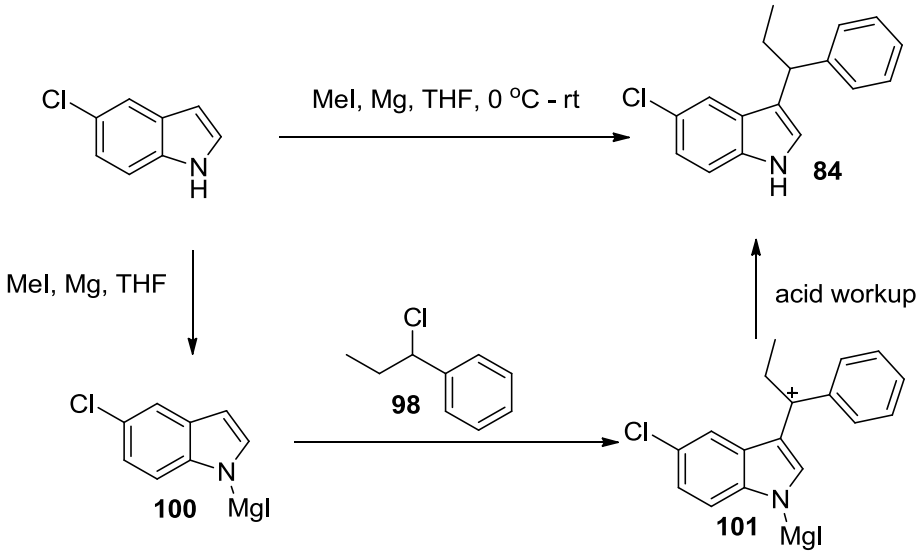
99

Scheme 52

The most reactive position on the indolyl anion, either the 1 or 3 position, is determined by a number of factors. The electrophile is important, with allylic or benzylic halides being the best electrophiles to use for the Michael type addition that occurs at the 3-position, whereas alkyl halides tend to result in alkylation of the nitrogen.[57] The metal counter-ion has also been found to be hugely important. The two most frequently used counter-ions are sodium and magnesium halide. Sodium forms a salt with the nitrogen which dissociates fairly easily, particularly in solvents such as DMF.[57] This means that for the most part the negative charge of the indolyl is free in solution and is able to perform a nucleophilic attack on any electrophile which is available in solution – for example *p*-toluenesulfonyl chloride which is seen in the tosyl protection of an indole which forms a sodium indolyl intermediate.[57] Contrastingly, magnesium halides form a fairly strong ionic bond with the indolyl and have a much lower dissociation constant than for the sodium indolyl salt.[57] This means that for the most part the indolyl is not available to perform a nucleophilic attack and instead the indole will more likely undergo the Michael type addition, which results in alkylation of the 3-position[57]. Using this insight, we set out to apply this theory to

indole system in this study. We decided to first try this on the 5-chloroindole and if it was successful we would move on to attempt it on the indole with an ester already in place.

The formation of the indolyl magnesium iodide species **100** (Scheme 53) was achieved by addition of MeMgI prepared in THF, followed directly by addition of compound **98**. Nucleophilic attack of compound **100** on compound **98** was expected to yield intermediate **101**, which would reveal the desired compound **96** after an acid workup. Once again on TLC no product was observed. With hindsight from a later reaction, diethylether would have been a better solvent for preparing the Grignard reagent. Furthermore, altering the addition of reagents for example adding the Grignard to a solution of indole in Et_2O and allowing the formation of the indolyl magnesium iodide to go to completion before adding the electrophile may have been more successful in producing a reaction. There were several factors we did not consider in this reaction, with more optimisation the reaction may eventually have worked. However, with lack of any positive results so far, we decided to move away from trying to introduce the propylbenzene group directly.

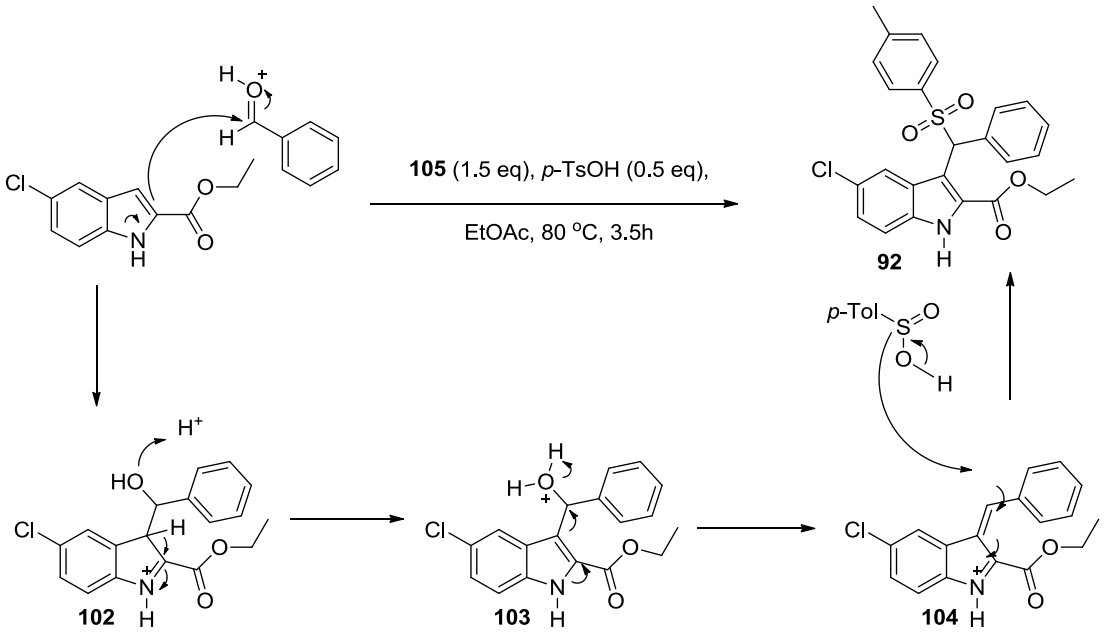


Scheme 53: Proposed mechanism for the magnesium indole concept

4.5 Employing strategy 3 – Utilising sulfonyl indoles to install the propylbenzene group onto the 3-position of the indole

Two separate papers, Ballini *et al*, 2008 and Palmeiri *et al*, 2007, detailed a procedure involving synthesising a 3-sulfonyl indole, such as for compound **92**, and then utilising a Grignard reaction to substitute the sulfonyl group for an alkyl group, which would give compound **90** in our synthesis

(Scheme 41).[51,52] This gave precedence for this procedure to be used to install the propyl benzene group in the 3-position of the indole, but we were unsure how the procedure would fare with an ester present in the 2-position of the indole.[51] We hoped that by using only one equivalent of the Grignard reagent we would avoid attacking the ester during this reaction, since the sulfonyl group is a far more reactive species. The proposed mechanism is shown in Scheme 54. Benzaldehyde is protonated by the *p*-TsOH and the indole performs a nucleophilic attack of this species to form intermediate **102**. The alcohol of this intermediate is protonated and elimination of the proton on the 3-position of the indole allows aromaticity to be regained yielding intermediate **103**. Through an indole mediated S_N1 substitution there is first elimination of water to form intermediate **104**, which is then attacked by the *p*-toluenesulfinic acid, compound **105**, to yield compound **92**.

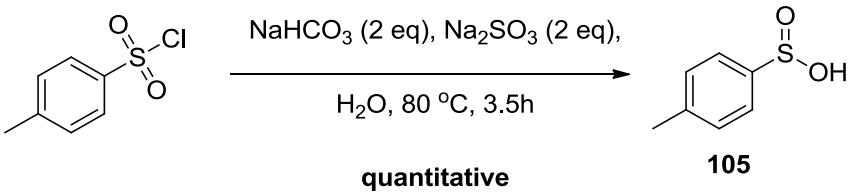


Scheme 54: Proposed mechanism for the formation of the sulfonyl indole from Ballini et al, 2006 and Palmieri and Petrini, 2007[51,52]

4.5.1 Synthesis of *p*-toluenesulfinic acid 105

For the formation of the sulfonyl indole, *p*-toluenesulfinic acid **105** was required and this compound needed to be synthesised. This was done by adding sodium bicarbonate and sodium sulfite to distilled water, heating for 2h at 80 °C and then introducing the *p*-toluenesulfonyl chloride and refluxing for a further 1h. The use of the sodium bicarbonate is to keep the solution basic, which prevents any acid degradation products forming from both the sodium sulfite and *p*-

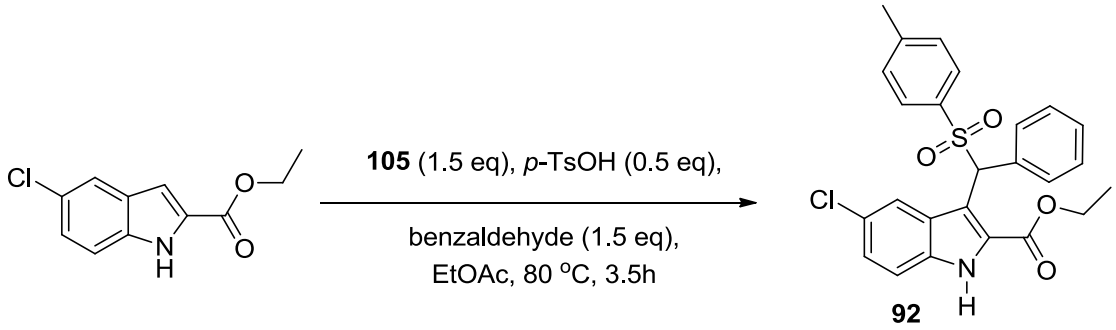
toluenesulfonyl chloride. Once the reaction has cooled, it was acidified with HCl and product **105** was left to crash out of solution overnight. The crystals were filtered off and dried under vacuum, but after just a few minutes an odour was given off from the sample and the bright white crystals started turning brown. After 1 day at rt, the sample had turned dark brown and had a strong odour. The reaction was repeated and the crystals were stored at -20 °C until they were needed for the next reaction, they appeared to store well at this temperature for a day or so but for any longer amount of time degradation started occurring so the sample had to be used almost immediately for the next reaction. Due to the instability of compound **105**, it was not easily characterised. The only form of characterisation that could be done was melting point analysis and the melting point we obtained corresponded within 5 °C of the literature value.



Scheme 55

4.5.2 Attempted synthesis of the ethyl 5-chloro-3-(phenyl(tosyl)methyl)-1H-indole-2-carboxylate 92

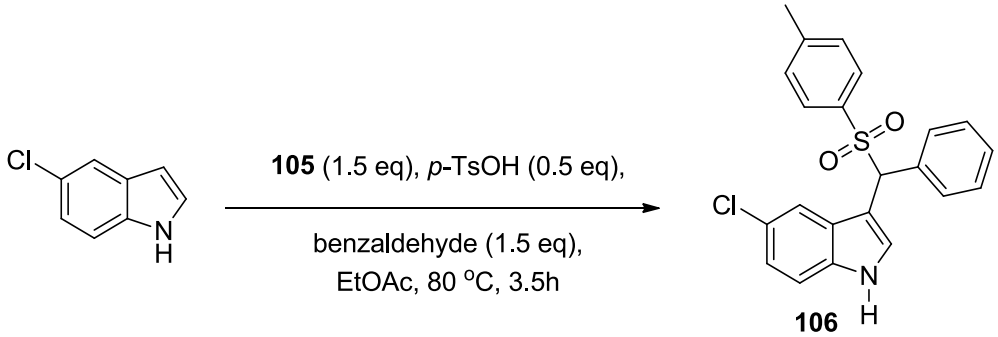
Having successfully made the sulfinic acid **105**, our first attempt to make the 3-sulfonyl indole **92** was using ethyl 5-chloroindolecarboxylate as the starting material (Scheme 56). The reaction was carried out in EtOAc, and all components were added together at rt. Refluxing at 80 °C resulted in the solution turning orange and on TLC a streak was seen close to the baseline. After 3.5h the solution was basified and extracted with EtOAc. Unfortunately, purification was not successful. Two peaks seen under the UV detector on the auto column machine would not separate, collectively giving only 100 mg of material, and we were unable to interpret them by NMR spectral analysis due to the impurities. Furthermore, we were unsure how the ester would fare in the following reaction with the Grignard reagent, thus it seemed more effective to install the sulfonyl group onto an indole lacking the ester in the 2-position.



Scheme 56

4.5.3 Synthesis of 5-chloro-3-(phenyl(tosyl)methyl)-1H-indole 106

With the failure to form the sulfonyl on the indole bearing the ester, we tried the same reaction using 5-chloroindole as our starting material and EtOAc as a solvent. The same equivalents were used and after refluxing at 80 °C for 3h this time we saw the starting material was consumed and a bright red spot was seen close to the baseline. Basifying with conc. sodium bicarbonate, extraction with EtOAc and purification by column chromatography revealed the desired compound as a bright red solid in 62% yield. The product was fully characterised. In the ^{1}H NMR spectrum, the methyl signal of the sulfonyl group was seen at 2.35 ppm, and a singlet for the benzylic C – H was observed at 5.57 ppm. All aromatic signals could be accounted and the indole N – H signal was also observed at 8.72 ppm. Additionally, IR analysis also revealed strong bands at 1360 and 1200 cm^{-1} indicating the presence of S = O bonds.

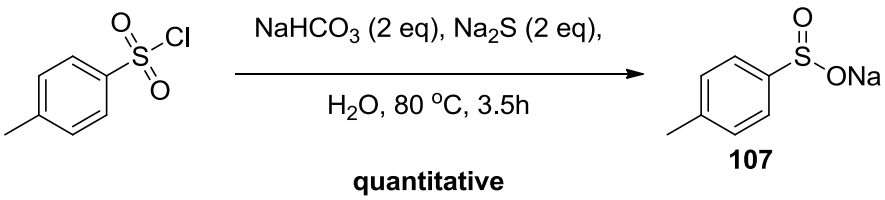


Scheme 57

4.5.4 Synthesis of sodium *p*-toluenesulfinate 107

Having been successful in synthesising sulfonyl indole **106** we set out to attempt a few optimisations for this reaction. We wondered if it was necessary to use the *p*-toluenesulfinic acid

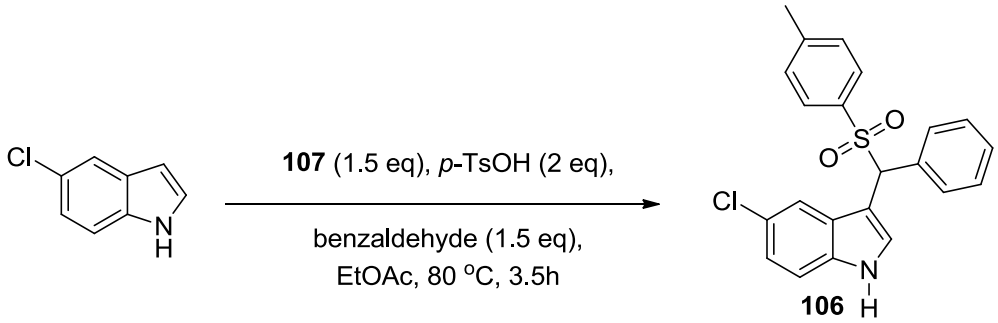
105 or if the sodium salt **107** could be used in its place when forming the 3-sulfonyl indole **106** (Scheme 58). The acid **105** could then be formed *in situ* by adding an extra 1.5 equivalents of *p*-TsOH. The reason the sodium salt would be preferred was due to the benefit of it being stable enough to be made in bulk and stored, whereas the acid degrades in a few days and therefore has to be remade each time the synthesis of the sulfonyl indole is repeated. In terms of making several derivatives with different alkyl groups and substituents on the phenyl ring, having this reagent in bulk simplifies the situation. The sodium salt **107** was synthesised in the same way as the acid **105,** however no acidification was performed once the reaction was complete. The reaction mixture was simply left to cool overnight and the compound **107** crashed out of solution and could be filtered it off. No other form of purification was performed, however melting point analysis corresponded to literature melting point within 10 °C.



Scheme 58

4.5.5 Synthesis of 5-chloro-3-(phenyl(tosyl)methyl)-1H-indole 106

Having formed the sodium *p*-tolenesulfinate **107**, the compound could now be tested to see if it could be used instead of the *p*-toluenesulfinic acid **105** in the reaction to form compound **106**. We repeated the reaction under the same conditions, this time using sodium *p*-tolenesulfinate, compound **107**, and 1.5 equivalents of *p*-TsOH. This resulted in a slight reduction in yield (52%) but avoided the complications which arose from using *p*-toluenesulfinic acid **105**, so we concluded since it was much easier to handle thus would be a better option for future reactions.

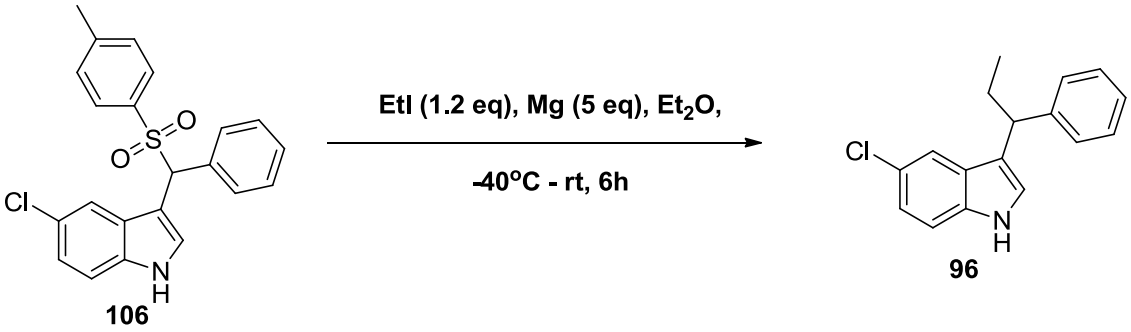


Scheme 59

4.5.6 Synthesis of 5-chloro-3-(1-phenylpropyl)-1H-indole **96**

Now that we were successfully synthesising the sulfonyl indole, the next step was to substitute the sulfonyl group for an ethyl group. A flask was charged with dry Et_2O, magnesium turnings were added and along with EtI to form EtMgI. Once most of the magnesium had been used up and the cloudy solution had cooled, approximately after 1h, this Grignard reagent was transferred drop wise by syringe into a second flask containing the sulfonyl indole **106** in Et_2O at -35 °C. As the Grignard was being added the solution went from clear red to murky orange/brown and a precipitate was visible. This precipitate was most likely the indolyl magnesium iodide salt being formed, which is insoluble in the organic solvent. The reaction mixture was allowed to warm up to rt overnight and TLC showed the formation of a new spot, although there was still starting material visible. Nevertheless, the reaction was quenched with distilled water, acidified with 4M HCl and extracted with EtOAc. Purification by column chromatography revealed compound **96** in 66% yield and a small amount of the unreacted starting material, compound **106** could be recovered. In subsequent optimisation attempts, we found that changing the equivalents of Grignard reagent did not allow for all starting material to be consumed; this appeared to be a problem caused by the insoluble precipitate formed when adding to Grignard reagent to the indole solution. We tried changing the solvent to THF but this did not prevent the formation of the clumpy precipitate and additionally caused a large reduction of the yield. Nevertheless, we had compound **96** at hand and this was characterised. ^{1}H NMR spectroscopy revealed the characteristic splitting pattern which would be expected for the ethyl group in the benzylic position, including a triplet integrating for one proton, a doublet of quartets integrating for two protons and a triplet integrating for three protons (Figure 21). Unfortunately, the mass spectrum

showed that degradation of some kind had occurred and we were unable to find a molecular ion to confirm this compound had been synthesised.



Scheme 60

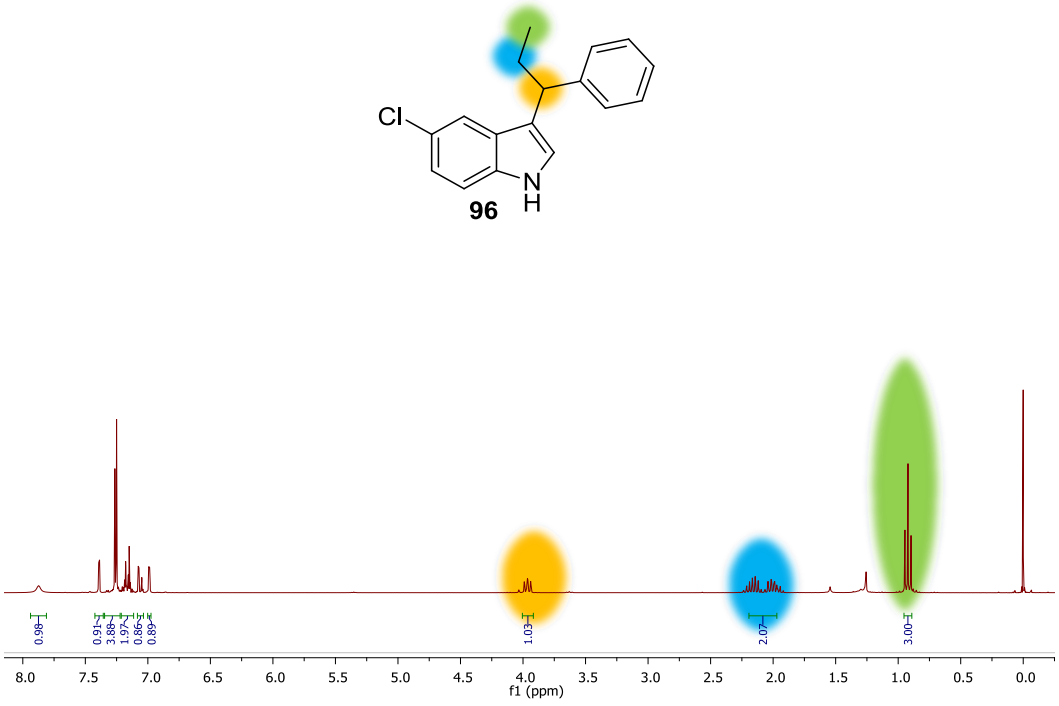


Figure 21

4.5.7 Synthesis of *tert*-butyl 5-chloro-3-(phenyl(tosyl)methyl)-1H-indole-1-carboxylate 108

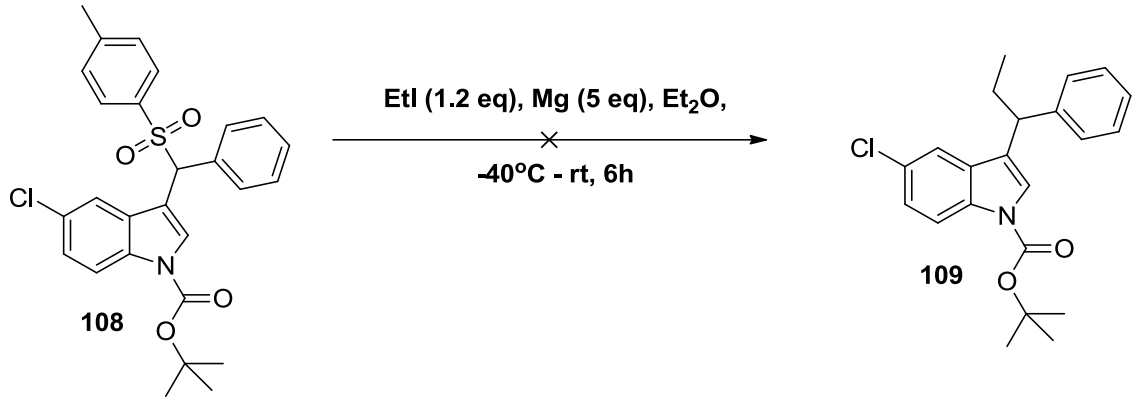
Although we had successfully obtained indole **96**, the reaction was still hindered by the formation of insoluble aggregates while adding the indole. We suspected that the intermediate responsible for this was an indolyl magnesium iodide species, which was insoluble in the Et_2O. Since changing the solvent was not an option, we decided to protect the indole nitrogen with a Boc group to see if the precipitate would no longer form and if this would increase the yield. To this end compound

106 first was protected with a Boc group using Boc_2O and catalytic DMAP in THF to afford compound **108** in 80% yield (Scheme 61). The product was fully characterised and the characteristic signal at 1.71 ppm integrating for 9 protons along with the absence of an N – H signal in the ^1H NMR spectrum indicated we had successfully Boc protected the sulfonyl indole **108**.

Scheme 61

4.5.8 Attempted synthesis of *tert*-butyl 5-chloro-3-(1-phenylpropyl)-1H-indole-1-carboxylate 109

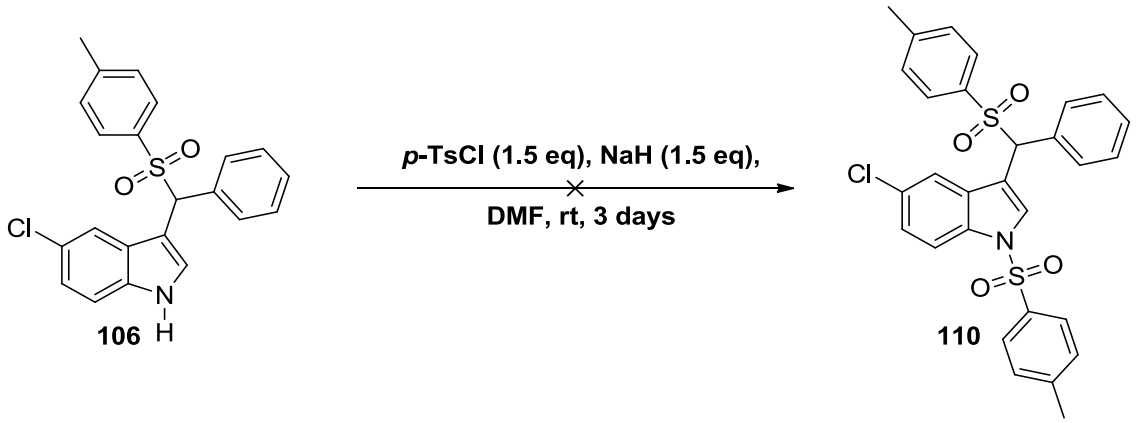
Having protected the indole nitrogen of compound **106** with a Boc group to yield compound **108**, we attempted to substitute the sulfonyl group for an ethyl group to afford compound **109**. The reaction was repeated as before and at first, we were excited to see that no precipitate formed when adding the Grignard reagent, which was the main motivation for protecting the indole in the first place. Unfortunately, on TLC there were multiple spots observed and this translated to a messy purification process in which we were not able to isolate the desired product. Further research into literature revealed that Boc groups are not tolerant of Grignard reagents and this explains the large number of products formed during the reaction.



Scheme 62

4.5.8 Attempted synthesis of 5-chloro-3-(phenyl(tosyl)methyl)-1-tosyl-1H-indole 110

Having been unsuccessful in introducing the ethyl group on the Boc protected indole **108**, we decided to attempt to use a tosyl protecting group and synthesise compound **110** (Scheme 63). We hoped that this group would be more tolerant to the Grignard reagent. With this in mind, we attempted to protect indole **106**, using NaH and *p*-TsCl in DMF. After 3 days, not all starting material had been consumed and multiple spots could be observed on TLC. The reaction was worked-up, however purification was not successful and only starting material could be successfully isolated by column chromatography.



Scheme 63

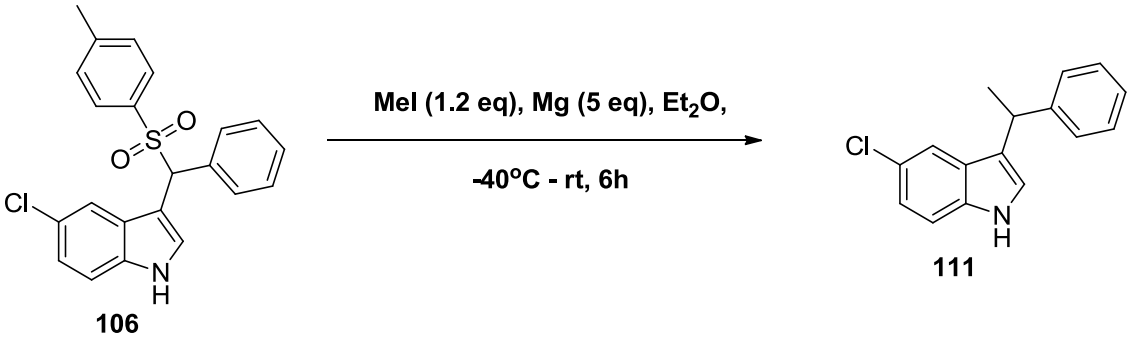
4.6 Investigating the robustness of this reaction

Despite being unable to optimise the installation of the ethyl group using a Grignard reaction any further, the yield obtained for compound **96** was still good. We wanted to investigate if this same

procedure would succeed using different alkyl groups, so we could probe the boundaries within the Val179 pocket. If this was successful, it could open the possibility of developing another small library of alkyl derivatives.

4.6.1 Synthesis of 5-chloro-3-(1-phenylethyl)-1H-indole 111

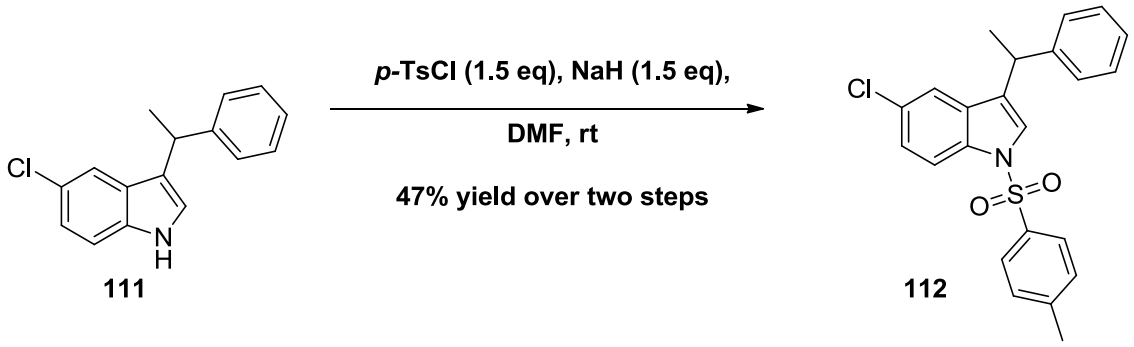
The next alkyl derivative we attempted to synthesise was the methyl derivative **111.** The reaction was performed in Et_2O, using MeI and Mg to form the Grignard reagent and adding this to sulfonyl indole **106** (Scheme 64). Following purification, which was not entirely successful due to coelution, the product was analysed by NMR. From ^{1}H NMR analysis it appeared we had made the product, owing to a quartet at 4.34 ppm integrating for one proton and a doublet at 1.72 ppm integrating for three protons, indicating a methyl group in the benzylic position. Unfortunately, the sample we had obtained was not pure so we did not perform further analysis on this compound.



Scheme 64

4.6.2 Synthesis of 5-chloro-3-(1-phenylethyl)-1-tosyl-1H-indole 112

Since we were unable successfully purify the sample of **111** which was obtained, we envisaged that protecting the indole nitrogen may allow for easier separation of the coeluting products so that we could establish that we had the desired product at hand. Tosyl protection was performed on our sample and purification afforded compound **112** in 47% yield over two steps (Scheme 65). The compound was fully characterised. Once again, the ^{1}H NMR spectrum showed a doublet for the CH_3 protons and a quartet for the benzylic CH proton, confirming that the methyl group was present in the benzylic position and that the previous reaction had been successful (Figure 22). The yield over two steps was also good enough for this route to be employed to make further alkyl derivatives.



Scheme 65

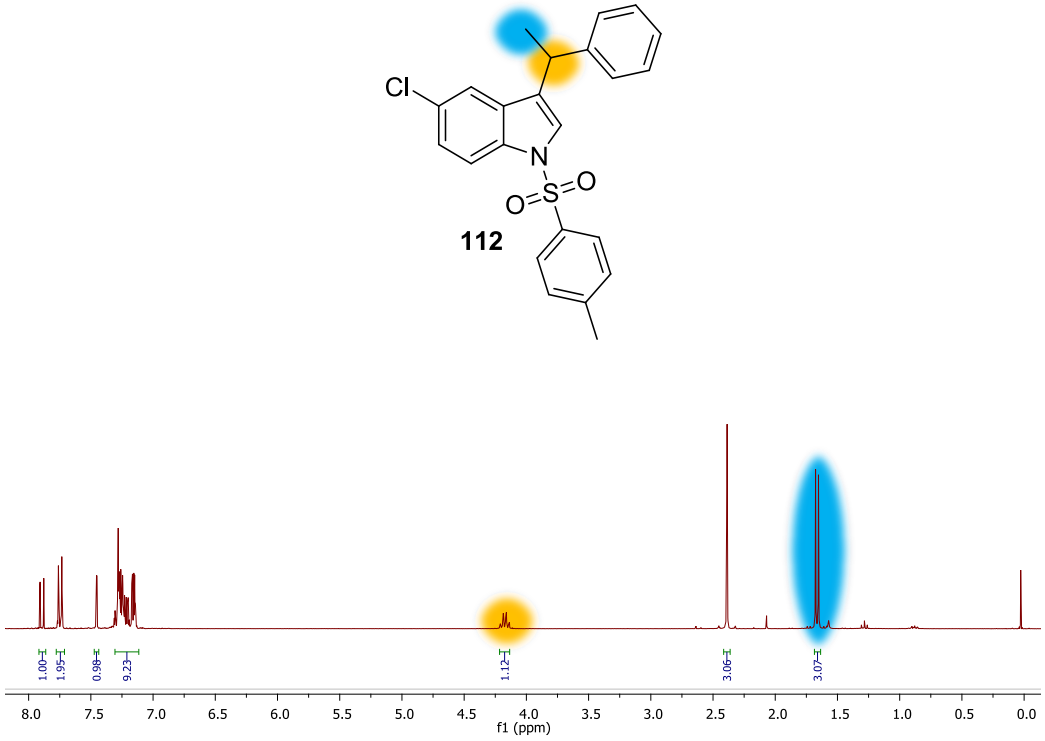
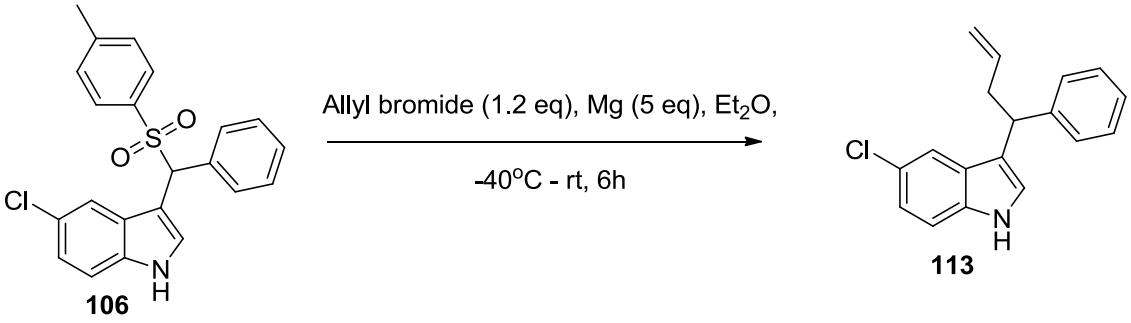


Figure 22

4.6.3 Synthesis of 5-chloro-3-(1-phenylbut-3-en-1-yl)-1H-indole **113**

To extend on possible alkyl derivatives, an allyl group was also installed onto the benzylic position (Scheme 66). The Grignard reagent was prepared with allyl bromide and Mg, and was added to the sulfonyl indole **106**. Unfortunately the yield was very low, affording compound **113** in only 18% yield. The inferior reactivity of an alkyl bromide, in comparison to an alkyl iodide, in the formation of the Grignard and its reactions was most likely the main reason for this reaction not giving as high yields as the ethyl derivative. The compound was fully characterised. In the aliphatic

region of the ^{1}H NMR spectrum, the signals for all protons corresponding to the allyl group in the benzylic position are accounted for, along with their characteristic splitting patterns (Figure 23).



Scheme 66

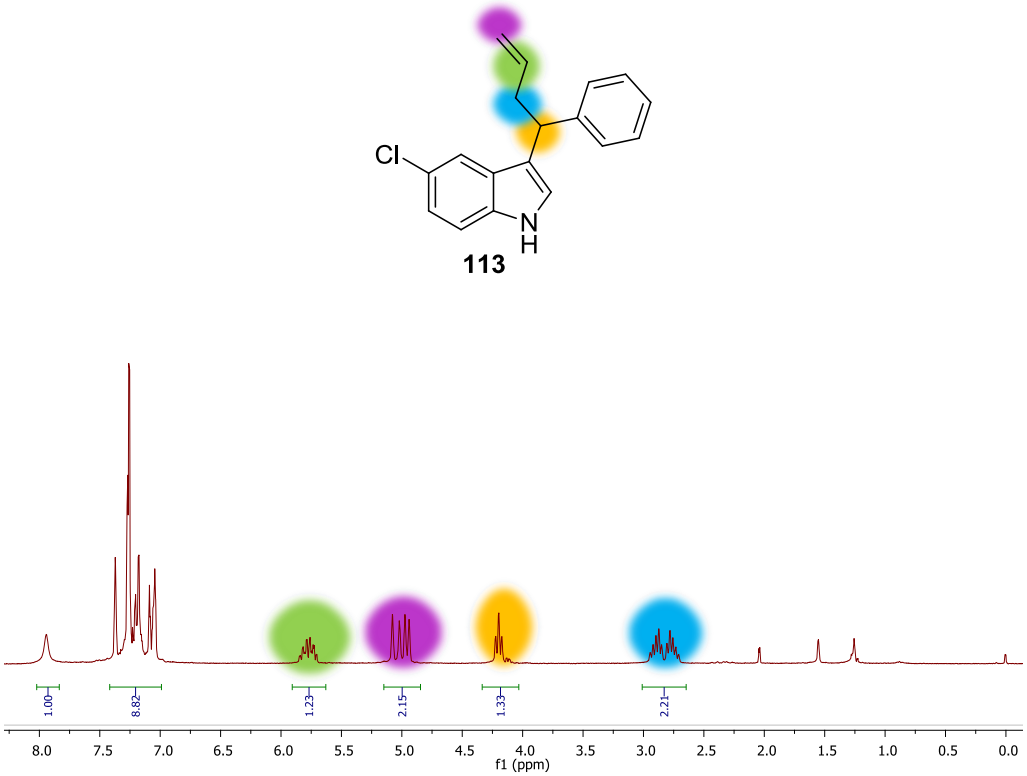


Figure 23

4.6.4 Comments on the robustness of this procedure

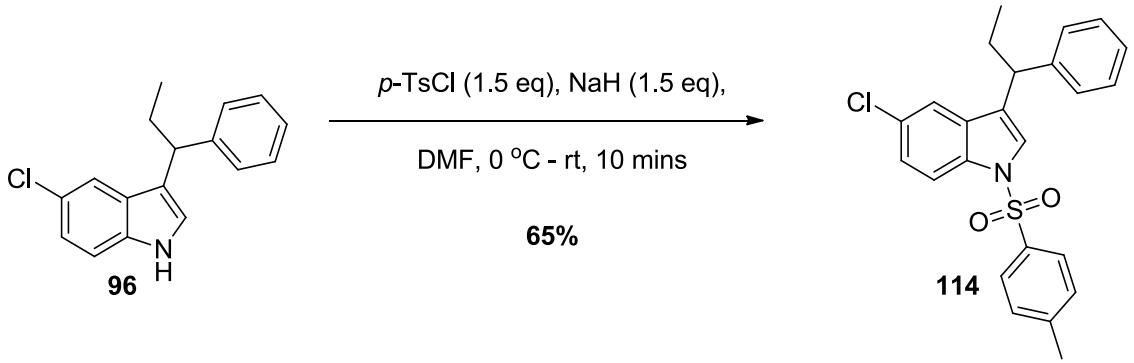
It appears that this procedure may be successful for introducing a variety of different alkylbenzene groups onto the 3-position of the indole. Iodoalkanes should be used to produce the Grignard

reagent and more optimisation may be required for this reaction in the form of changing the temperature and equivalents used. Overall, this procedure opens the possibility of synthesising a new library of different alkyl derivatives, however first there should be more attempts to introduce the alkylbenzene group to the indole with an ester already in the 2-position.

4.7 Installing an ester and an amide onto the 2-position of the indole

4.7.1 Synthesis of 5-chloro-3-(1-phenylpropyl)-1-tosyl-1H-indole **114**

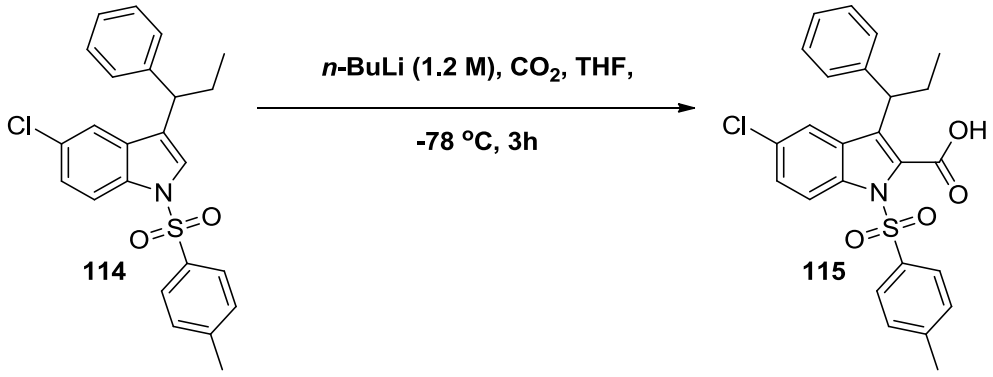
Having installed our desired group into the 3-position to give **96**, the only modification which was left to do was to install the ester into the 2-position. Since we had previously tested the effect of having an amide in this position and it had shown to have equal activity to the ester, we envisaged installing a carboxylic acid into the 2-position and then modifying this to both the ethyl ester and the amide. The most common way of introducing a carboxylic acid into the 2-position of an indole is using an alkyl lithium reagent. Several literature procedures outlined elaborate procedures of using alkyl lithium reagents to install the carboxylic acid, using a number of chelating agents and bulky modifiers such as TMEDA to aid the reaction. Although some literature procedures for installing a carboxylic acid into the 2-position of the indole were performed on unprotected indoles, more commonly a protecting group on the indole nitrogen was utilised. A tosyl protecting group was used by a few literature procedures and was not cleaved during the reaction. To this end, we set out to protect the indole nitrogen of compound **96** with a tosyl group. The reaction was accomplished in a moderate yield of 65% to afford compound **114**. What was interesting to note was the speed at which the starting material was consumed. In previous tosyl protection reactions we had protected an indole with an ester installed in the 2-position and in this case, we had seen reaction times between 18h and a few days. This highlights the considerable effect the ester has on the reactivity of the system. The compound was fully characterised and the presence of the tosyl group was confirmed by a CH_3 signal at 2.39 ppm, signals in the aromatic region spectrum and absence of an N – H signal in the ^{1}H NMR spectrum.



Scheme 67

4.7.2 Synthesis of 5-chloro-3-(1-phenylpropyl)-1-tosyl-1H-indole-2-carboxylic acid 115

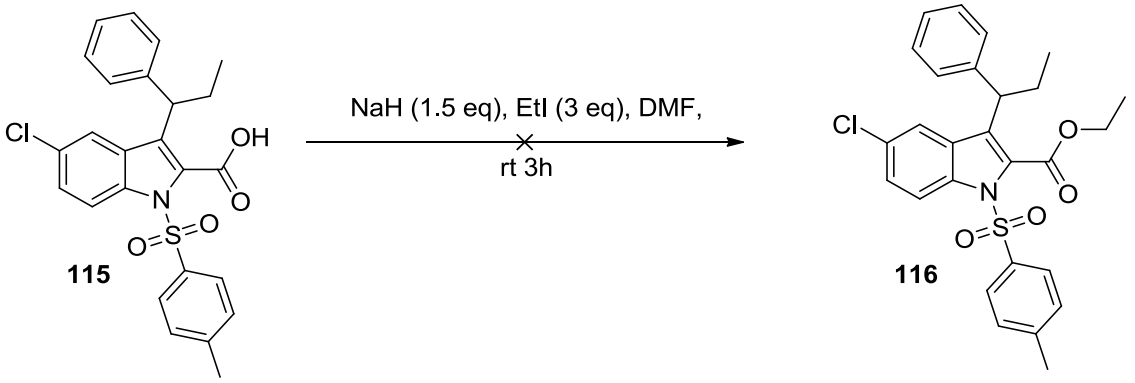
Before attempting to install the carboxylic acid using all the elaborate techniques described in literature, we first wanted to try a more simplified procedure using only n-BuLi with CO_2. The indole **114** was dissolved in dry THF and n-BuLi was added at -78 °C. This was stirred for 1h before dry ice was added directly to the reaction mixture and it was allowed to warm up. After 2h all starting material had been consumed and the reaction was quenched with water, acidified and extracted with EtOAc. No further purification was performed, and the crude sample of compound **115** was used directly in the next reaction. Although, we did not fully characterise this compound, TLC indicated we most likely had formed the carboxylic acid, since there was a new spot formed which was too polar to move off the baseline. Additionally, a peak at 167 ppm in the ^{13}C NMR spectrum indicated the presence of a carboxylic acid derivative.



Scheme 68

4.7.3 Attempted synthesis of ethyl 5-chloro-3-(1-phenylpropyl)-1-tosyl-1H-indole-2-carboxylate 116

Now that we had our carboxylic acid **115** at hand we wanted to modify this functionality to the ester derivative **116**. We first attempted to form this using a simple ethylation with EtI in DMF using NaH to deprotonate the carboxylic acid (Scheme 69). Unfortunately, NaH turned out to be strong enough to remove the tosyl group at rt – exposing the indole nitrogen for attack and yielding a number of unwanted side products. We concluded that a less harsh base such as potassium carbonate could be used instead of NaH to prevent the tosyl group being removed during the reaction.

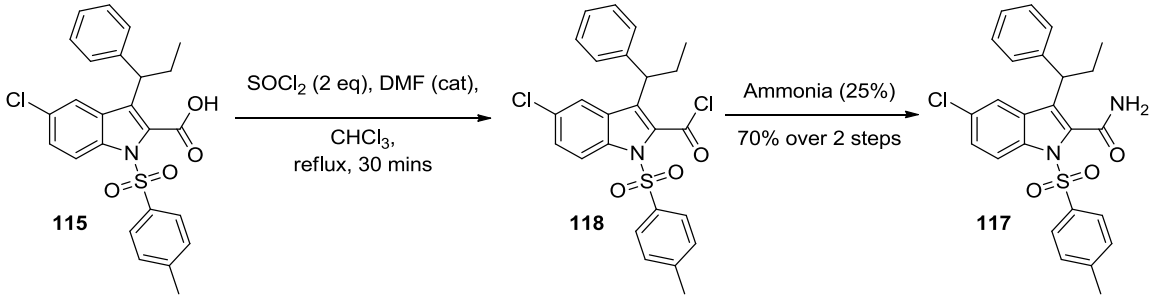


Scheme 69

4.7.4 Synthesis of 5-chloro-3-(1-phenylpropyl)-1-tosyl-1H-indole-2-carboxamide 117

With no luck so far with forming the ester, we temporarily turned our attention to forming the amide derivative, compound **117**. A procedure used by Csomós *et al* (2007) afforded an amide from a carboxylic acid by refluxing the indole with thionyl chloride in chloroform and quenching the acid chloride **118** formed with aqueous ammonia to yield the amide (Scheme 70).[58] The reaction was set up in a flask which we charged with chloroform and compound **115** followed by thionyl chloride at 0 °C. The reaction was refluxed, but after 1 h no change was observed by TLC. Since acid chlorides are often synthesised using a catalytic amount of DMF, we added a drop of DMF to our reaction mixture. After 30 mins, all the starting material had been consumed. The reaction was cooled and then poured into a 25% aqueous ammonia solution which had been cooled on ice. A murky solution formed from which we extracted our product using EtOAc. Purification by column chromatography afforded compound **117** in 70% yield based on the crude

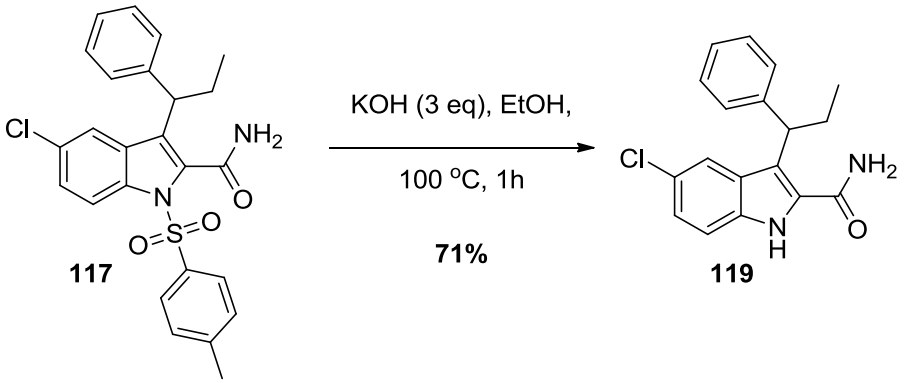
starting material. The compound was characterised, however the mass spectrum did not show the molecular ion, most likely due to degradation of the compound prior to performing the analysis – although the degradation product is hard to speculate. A signal in the ^{1}H NMR spectrum integrating for 2 protons at 5.87 ppm corresponding to the NH_2 was used to confirm the formation of the amide, along with a signal in the ^{13}C NMR spectrum at 164 ppm corresponding to the carbonyl of the amide.



Scheme 70

4.7.5 Synthesis of 5-chloro-3-(1-phenylpropyl)-1H-indole-2-carboxamide 119

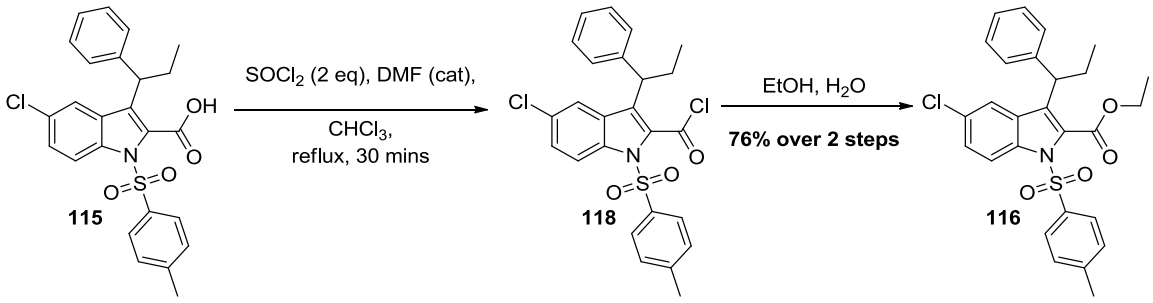
The success of the amide formation left only the tosyl deprotection to be carried out in order for our final amide derivative **119** to be revealed. We were concerned that the KOH in EtOH would hydrolyse the amide, however after refluxing the reaction for just 1h, all starting material had been consumed and only one new spot was seen on TLC. An acid work-up followed by purification by column chromatography revealed compound **119** in 71% yield. The compound was fully characterised and then sent for biological testing to assess this derivatives activity against HIV-1.



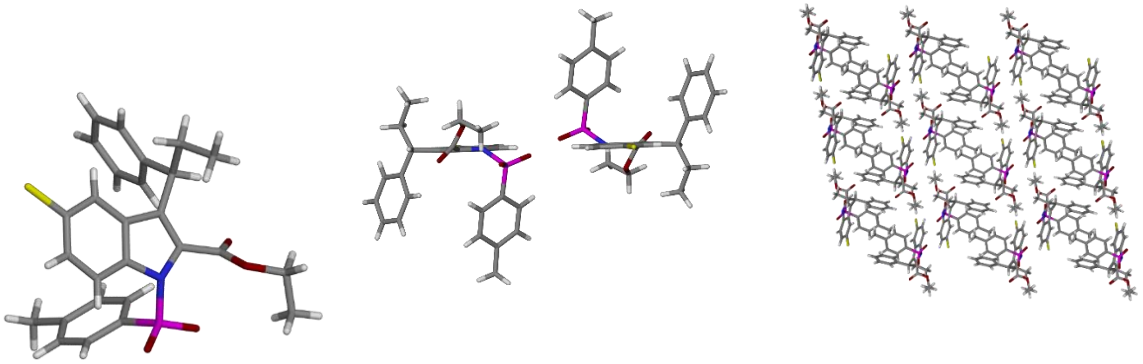
Scheme 71

4.7.6 Synthesis of ethyl 5-chloro-3-(1-phenylpropyl)-1-tosyl-1H-indole-2-carboxylate 116

With the success of forming the amide from the acid chloride **118**, we believed that this strategy may work using ethanol to quench the reaction (Scheme 72). The acid chloride **118** was again formed by refluxing the carboxylic acid **115** in excess thionyl chloride for 1h. Thereafter the reaction mixture was cooled on ice before quenching with ethanol, and after it appeared that nothing was happening as determined by TLC, water was added in an effort to recover the carboxylic acid **115**. The solution was left for 5 days, and crystals formed, which were filtered off. To our surprise, the crystals which had formed were not in fact the carboxylic acid **115**, but we had obtained the desired compound **116** in 76% yield. This was confirmed by single crystal X-ray crystallography which was performed to confirm that the groups had been installed in the desired 2- and 3- positions of the indole (Figure 24). The crystal structure data revealed that compound **116** packs in the triclinic space group P1. Both enantiomers are present in the crystal structure in a 1:1 ratio, producing a racemic mixture. However, only one of the molecules (with S-configuration) is present in the asymmetric unit and the second molecule (with R-configuration) is generated through the center of inversion. This pattern is then repeated throughout the crystal structure. The crystal structure confirmed that we had successfully installed the propylbenzene group onto the 3-position of the indole and the ester onto the 2-position of the indole. No further characterisation was done, since we only had 24 mg with which to work with to obtain our final compound **90**.



Scheme 72

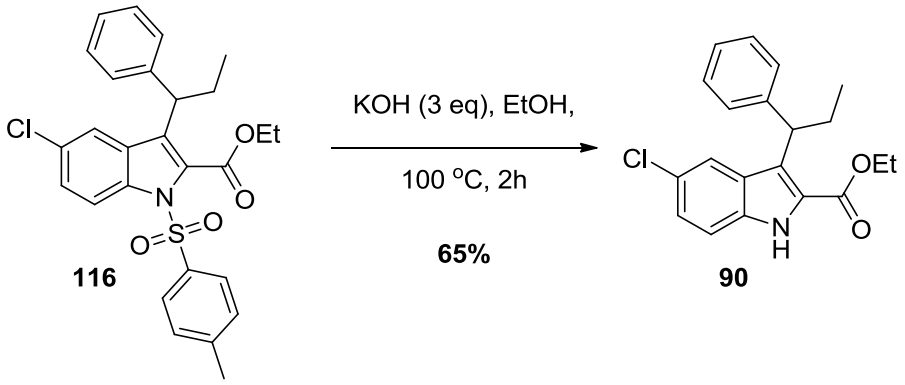


a) b) c)

*Figure 24: (a) Asymmetric unit of compound **116** (b) both R and S molecules; (c) packing of compound **116** viewed down the a axis.*

4.7.7 Synthesis of ethyl 5-chloro-3-(1-phenylpropyl)-1H-indole-2-carboxylate 90

The final step towards compound **90** was removal of the tosyl group (Scheme 73). This was performed by refluxing compound **116** in KOH and EtOH for 2h followed by an acid work-up. Purification by column chromatography afforded our final compound in 65% yield. As seen previously, this reaction does not hydrolyse the ester. The stability of this ester on the 2-position of the indole may be due to a number of reasons, however the most obvious explanation is the resonance which would be possible due to the conjugated system which is present. The compound was fully characterised and the presence of the ethyl ester could now be confirmed by a quartet at 4.51 ppm along with a 0.84 ppm. The absence of a singlet at 2.34 ppm along with the presence of an N – H signal at 8.9 ppm confirmed that the tosyl group was removed and we had our desired compound **90** at hand and this, along with the amide derivative **119**, could be sent for biological testing to assess their activity against HIV-1.



Scheme 73

4.6 Pending biological efficacy results for derivatives 90 and 119

Unfortunately, due to our collaborator Dr Adri Basson moving to a new facility, new assays needed to be set up at his new laboratories. We were unable to acquire the efficacy results in time to include them in this thesis. The assays are expected to be up and running again by the end of March 2017, after which the compounds can be tested for their efficacy against wild-type HIV-1. Based on the modelling results we would expect the ethyl derivatives **90** and **119** to have less activity against HIV-1 than the original lead compound 10 but this remains to be seen.

4.7 Concluding remarks pertaining to Chapter 4

In our endeavour to introduce the propylbenzene moiety to the 2-position of the indole, the effect of remote groups on the compound which represent a change from the literature example has been emphasised repeatedly. Having the 5-chloro substitution, and in some instances also the ester in the 2-position, appeared to greatly reduce the reactivity of the indole system. Reactions which were successful in literature on unsubstituted indoles had no reaction on our 5-chloroindole. After attempting several procedures including reductive eliminations, Friedel-Crafts alkylations and more abstract reactions to introduce the propylbenzene group directly to the 3-position of the indole, we eventually turned our attention to a procedure which made use of a 3-sulfonyl indole as an intermediate. Thankfully, we could introduce the propylbenzene group to afford compound **96** using this two-step procedure, albeit not for the derivative with the ester group already in place. Introduction of the ester and amide to the 2-position of the indole was achieved by forming the acid chloride **118** from the carboxylic acid and quenching this with either

ethanol or ammonia respectively. Confirmation that we had introduced these substituents onto the desired positions performed on compound **116** by single crystal X-ray crystallography.

In conclusion, we successfully synthesised compounds **90** and **119** and these compounds are ready for biological testing to assess their activity against HIV-1. Unfortunately, these results are not available in time for the submission of this thesis however we expect to have obtained them by end of March 2017. Additionally, this procedure has been shown to be applicable for introducing other alkylbenzene groups to the 3-position of the indole and therefore could potentially be used to make another library of derivatives with varying alkyl groups on the benzylic position.

Chapter 5: Enantiomer resolution using classical crystallisation techniques and HPLC

5.1 Introduction to chiral drugs and methods for producing pure enantiomers

It is widely acknowledged the importance of chirality in the drug development industry and that generally enantiomers' biological activity differs in a number of ways.[59] In some cases, the one enantiomer may be completely inactive and in some rare cases the one enantiomer may have several undesirable side effects.[60] Asymmetric drug receptors, and enantioselective biological processes contribute to the differences in behavior *in vitro* and *in vivo* between the enantiomers. Adsorption, distribution, metabolism and elimination (ADME) profiles may be hugely different and this often plays an equally important role in the efficacy of a drug as the ability to bind to the drug receptor. The propensity for the body to discriminate between the two enantiomers emphasizes the importance of stereo-pharmacokinetic and pharmacodynamic studies and single enantiomer drug assays – all of which require the purification of enantiomers. In some countries, pharmaceutical regulatory bodies are now insisting that chiral pharmaceuticals are administered only in their optically pure form.[61] Above 50% of active pharmaceuticals are chiral, with more than 70% of current drug candidates exhibiting some form of chirality thus making the field of developing methods for purification of enantiomers indispensable. There are two main strategies for obtaining pure enantiomers; firstly, by asymmetric synthesis and secondly by separation of enantiomers by physical means.

There has been remarkable progress in developing enantioselective procedures for asymmetric synthesis in the last fifty years. In 2001, the Nobel Prize for Chemistry was awarded to Knowles, Noyori and Sharpless for their contribution towards asymmetric catalysis. Regrettably, a major criticism of this strategy for its use in industry is the cost of reagents used in the procedures. This is especially problematic when metal containing catalysts are utilized which often contain expensive metals such as titanium, palladium and rhodium. Enzymes have also been utilized in this endeavor and are also very costly, especially on a large scale since they are not as hardy as metal catalysts or tolerant to higher temperatures thus need to be replaced after every few

batches. Asymmetric synthesis procedures are not always robust, often requiring the development of a customized catalyst which only works for a specific substrate. It may also not be highly reproducible on a large scale. Furthermore, the enantiomeric excess obtained using these techniques do not always meet the requirements of regulatory bodies. Nevertheless, there are several pharmaceutical compounds which are synthesized as pure enantiomers using this technique including naproxen and L-Dopa.

In 1853, Pasteur discovered the method of classical resolution by diastereomeric salt crystallization. This method involves an acid-base reaction between a racemate (the drug) with one enantiomer of a chiral resolving agent yielding a 1:1 mixture of diastereomeric salts which can then be resolved using crystallization or filtration if the one enantiomer is not soluble in the solvent.[62] Once separated the desired diastereomeric salt may then be treated with an acid or base to obtain the pure enantiomer. The classical method is used in the purification of enantiomers of several synthetic amino acids and pharmaceutical drugs and works brilliantly on a large scale. Building on this method, in place of ionically bonded chiral resolving agents, a covalently bonded chiral auxiliary may be utilized. The downfall of this procedure is the need to perform a reaction to attach the chiral auxiliary and, once the diastereomers are separated, an additional reaction is needed to remove it. The attachment and removal of the chiral auxiliary add extra steps to the synthetic pathway and contribute to a lower overall yield of the final drug compound. Nevertheless, this procedure may allow for preferential crystallization of a single diastereomer giving a highly pure product using a method is easy to upscale. These crystallization techniques are by far the most widely used form of enantiomer purification on an industrial scale. Drugs which are purified in this manner include Citalopram (antidepressant), Zoloft (antidepressant), Sulphostin (diapeptidyl peptidase IV inhibitor) and Plavix (for treatment of cardiovascular disease) to name a few. Although crystallization methods are often cheap, the downfall is that there is no ‘one size fits all’ strategy and to achieve separation by crystallization a huge amount of trial and error is involved. This may take years to be successful and for many researchers this is not always an option.

As for the use of HPLC for separation of enantiomers, there are two ways in which this is carried out. The first method, known as a direct HPLC method involves the use of chiral columns. Chiral columns along with running costs of the machine can be immensely expensive and the columns need to be replaced on a regular basis when using this method for mass production. Also, there

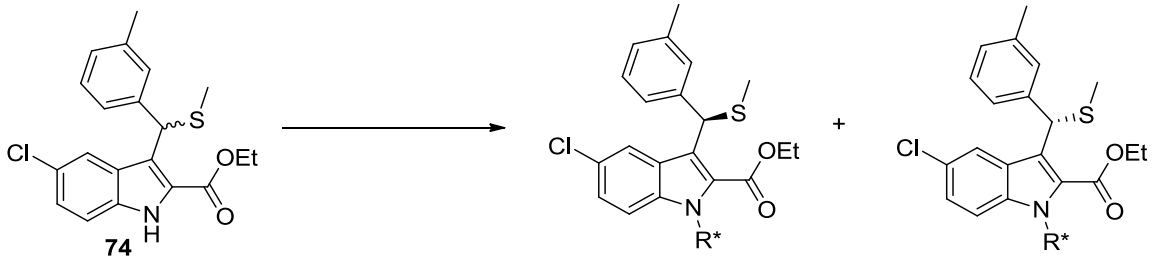
is currently no universal chiral column, so the technique often requires a trial and error approach to get started, and this can often be just as tedious and a whole lot more expensive than classical crystallization techniques. The second method, known as indirect HPLC involves derivatization of the racemate to form covalent or ionic diastereomers. This method relies on the two diastereomeric salts or compounds which are formed to have chemical and physical properties which are different enough to allow for separation. It is very rarely used on a large scale since it is not very effective in large batches, but due to its sensitivity and reliability it is a common technique for small scale separation of enantiomers.

5.2 A closer look at the chirality of compounds developed in this project and strategies for enantiomer purification

With a chiral center present in the benzylic position of all the compounds which have been synthesized in this project, it is inevitable that this needed to be addressed at some stage. To date, we have only tested racemic mixtures of our compounds, and although the activity of the compounds tested so far have been very good, we are still unsure if both enantiomers are equally active and there is a chance that one is completely inactive, thereby rendering our IC_{50} values half as good as they could be. Previous docking studies of the methyl ether **10** and sulfide derivative **76** in the NNIBP show that the *(S)*-enantiomer binds preferentially to the NNIBP than the *(R)*-enantiomer, based on binding scores. Furthermore, the *(S)*-enantiomer of the ethyl derivative **90** binds in the desired position, whilst the *(R)*-enantiomer docks in an inverted position in the NNIBP and lacks several important interactions. The opposite is seen for compound **13** where the *(R)*-enantiomer docks in the desired position but the *(S)*-enantiomer occupies an entirely different space. Without efficacy results available for the ethyl derivative, the sulfide derivative was at this time the most optimized form of the lead compound – having improved acid stability properties compared to the methyl ether derivatives whilst remaining active. For this reason, we decided to perform the enantiomer resolution experiments on a sulfide derivative, and compound **74** was chosen as our test compound for the different resolution strategies since this had the highest overall yield. Due to limited funding and the desire to develop a robust method which could be up scaled in future, our main focus was to develop a crystallization method for purifying the enantiomers. In the case that crystallization failed, the compounds were also sent for HPLC analysis.

5.3 Strategy 2A: Forming a covalent diastereomer with a chiral auxiliary on the nitrogen of the indole

One of the easiest functional groups to derivatize on compound **74** is the indole nitrogen. A pure enantiomer of a chiral group may be installed onto the indole nitrogen to create two diastereomers (Scheme 74). There are several requirements for what makes a good candidate to use as a chiral auxiliary but one of the most important is the ease of installing the group as well as removing it once the diastereomers have been separated. For this reason, it is often a good idea to choose chiral auxiliaries which have similar functionalities as common protecting groups such as Boc or tosyl protecting groups. In this way, the chiral auxiliary may be removed in a similar manner to which the corresponding protecting group is removed.

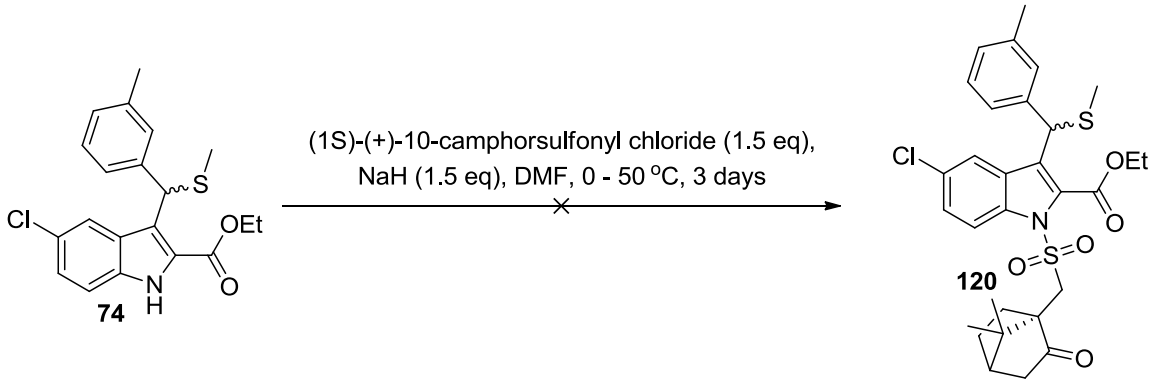


Scheme 74

5.3.1 Attempted synthesis of ethyl 5-chloro-1-((((*1S*)-7,7-dimethyl-2-oxobicyclo[2.2.1]heptan-1-yl)methyl)sulfonyl)-3-((3-methylphenyl)(methylthio)methyl)-1H-indole-2-carboxylate 120

Due to the natural abundance of (*1R*)-(+)-camphor, it has been used in a number of applications in organic chemistry throughout the years.[63] The ease of derivatization makes it extremely versatile in the syntheses of several enantiomerically pure compounds. In particular, the sulfonation of (*1R*)-(+)-camphor to yield (*1S*)-(+)-camphorsylfonic acid, which was first carried out in 1882, allows it to be used to make diastereomeric salts.[63] Unfortunately, the indole N-H is not basic enough to be used in this regard, however (*1S*)-(+)-camphorsulfonyl chloride can be used to form a covalent bond to the nitrogen in the same way that we use a tosyl protecting group. By protecting the indole nitrogen with this camphorsulfonyl group we create two diastereomers which in theory we could separate using crystallisation or HPLC.[64]

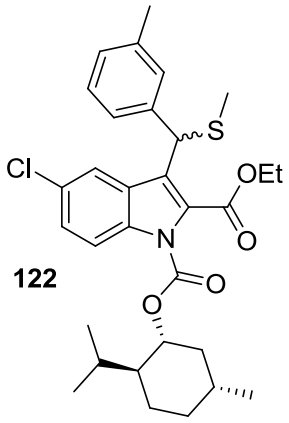
To this end, we set out to synthesise compound **120** (Scheme 75). The reaction was done in the same way as a tosyl protection, using DMF as a solvent and NaH was used to deprotonate the indole nitrogen at 0 °C. (*1S*)-(+)-camphorsulfonyl chloride was added to the deprotonated compound **74** and the reaction was then heated to 50 °C and after 3 days all starting material had been consumed. The reaction was worked-up and purified by column chromatography, however, characterisation revealed we had not formed the desired product **120**. The ^{1}H NMR spectrum indicated the nitrogen of the indole had not been protected, however there appeared to be signals which corresponded to the presence of the camphor sulfonyl group such as several complex multiplets in the aliphatic region. MS analysis showed m/z of 587.2427 which did not match any of the products that could be predicted.



Scheme 75

5.3.2 Synthesis of (*1S,2R,5S*)-(-)-menthyl-chloroformate 121

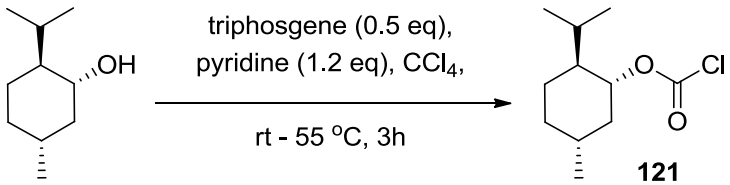
With the failure of introducing the camphor sulfonyl group to the nitrogen of the indole, we turned to another chiral auxiliary which has also been commonly used in enantiomer resolution. Menthol derivatives have been commonly used as chiral auxiliaries on the indole nitrogen to promote resolution of enantiomers. The most common way of installing the chiral menthyl functionality on the nitrogen is by utilising the chloroformate derivative, thereby forming a carbamate such as in compound **122** (Figure 25).



122

Figure 25

This was an attractive strategy for us, since we had experience dealing with carbamates, in the form of a Boc protecting group, and had a well-established strategy for removing this functionality from our system, and we envisaged that this would still be feasible for the menthyl carbamate group. To this end we set out to synthesise (*1R*,2S,5R)-(-)-menthyl-chloroformate, compound **121**, using a procedure outlined in Hajra *et al* (2007) (Scheme 76).[65] A dropping funnel was charged with carbon tetrachloride followed by triphosgene. Pyridine was carefully added to the dropping funnel which resulted in a milky gas forming and the solution turning milky yellow in colour. This solution was added all at once at rt to a solution of (*1R*,2S,5R)-(-)-menthol dissolved in carbon tetrachloride and this was heated to 55 °C for 3h. The reaction was quenched with ice water and extracted with DCM. The DCM layer was concentrated *in vacuo* to reveal a clear oil. The oil was characterised by ^{1}H and ^{13}C NMR spectroscopy, whereby all signals corresponded to literature. In addition, to establish there was no racemisation which occurred during the reaction, the optical activity of the compound was also measured. The experimental α_D value at 30 °C was found to be - 81°, which corresponded moderately well with the literature value of -74.4° at 25 °C. Since the measurement was not taken at exactly 25 °C, the slight difference in optical activity to that of literature is probably due to this factor, along with slight impurities within the sample.



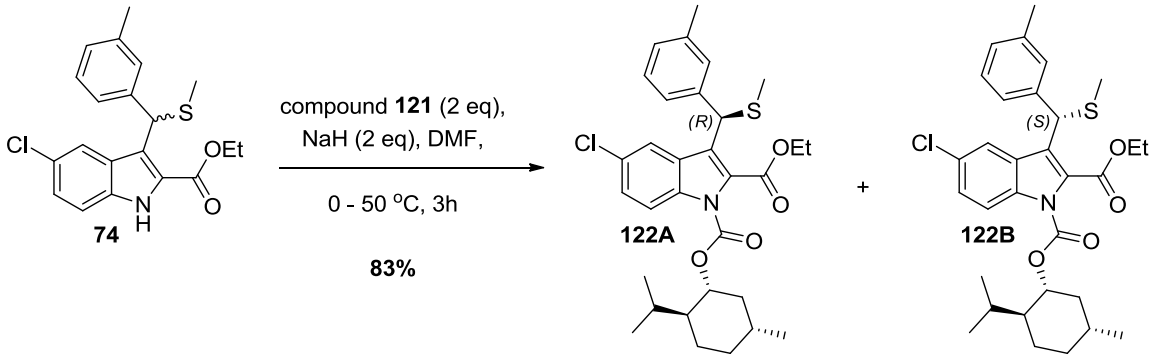
121

89%

Scheme 76

5.3.3 Synthesis of 2-ethyl 1-((1R,2S,5R)-2-isopropyl-5-methylcyclohexyl) 5-chloro-3-(*(R)*-(methylthio)(m-tolyl)methyl)-1H-indole-1,2-dicarboxylate **122A** and 2-ethyl 1-((*1S,2R,5S*)-2-isopropyl-5-methylcyclohexyl) 5-chloro-3-(*(S)*-(methylthio)(*m*-tolyl)methyl)-1H-indole-1,2-dicarboxylate **122B**

Having successfully synthesised (*1S*)-(-)menthyl-chloroformate, compound **121**, we were now able to install this group onto the indole nitrogen. This was achieved by deprotonating the indole nitrogen of compound **74** using NaH in DMF, and then introducing compound **121**. The reaction gave compound **122A** and **122B** in 83% yield.



Scheme 77

Compounds **122A** and **122B** were fully characterised by MS, ^{1}H and ^{13}C NMR spectral analysis. The mass spectrum did not contain the molecular ion, since the positive ionisation results in elimination of the thioether group. The ^{1}H NMR spectrum was very complex, and the splitting patterns often show broad signals which are difficult to identify. The aliphatic region is shown in the ^{1}H NMR spectrum below (Figure 26). The methyl group of the thioether, adjacent to the chiral centre, shows a doublet in the ^{1}H NMR spectrum at 2.31 ppm, which is commonly seen for racemic mixtures of diastereomers, and is an indication that we do have a racemic mixture which was the outcome we were looking for in this reaction. Although the signals for the menthyl moiety are not very well resolved, we do see a perfect doublet of triplets for a CH at 1.80 ppm which is further downfield and we have predicted this to be the proton labelled in the spectrum below. The rest of the signals can only be assigned based on integration since the splitting patterns are not discernible.

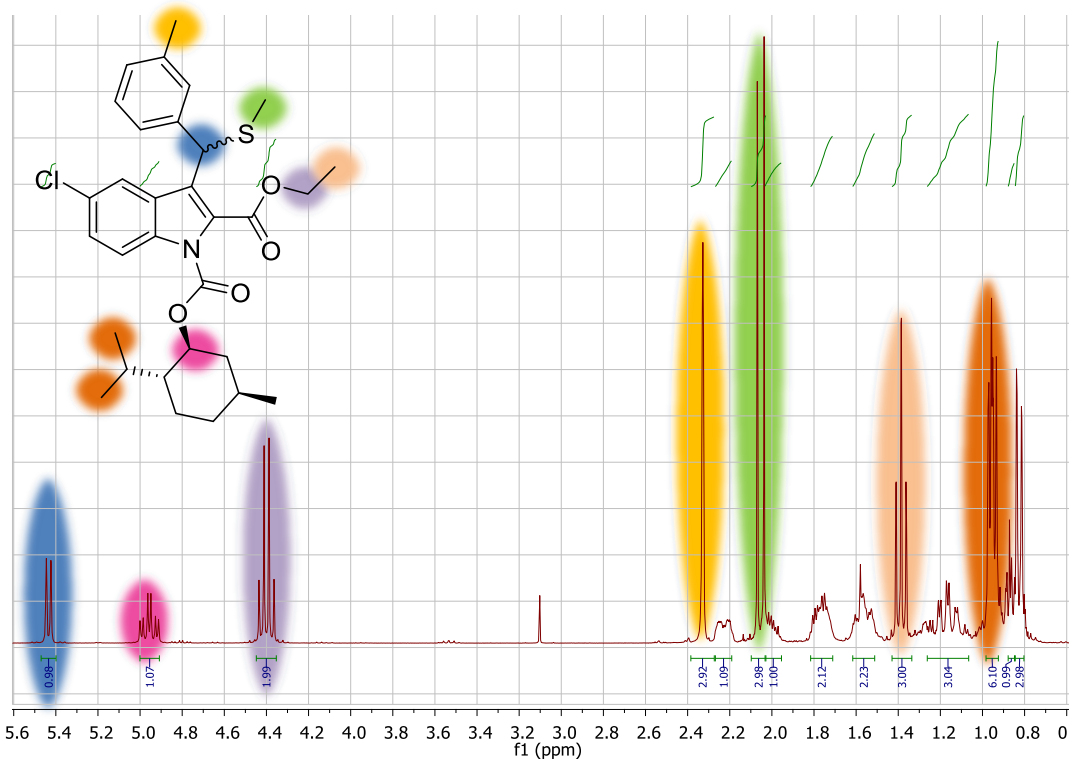


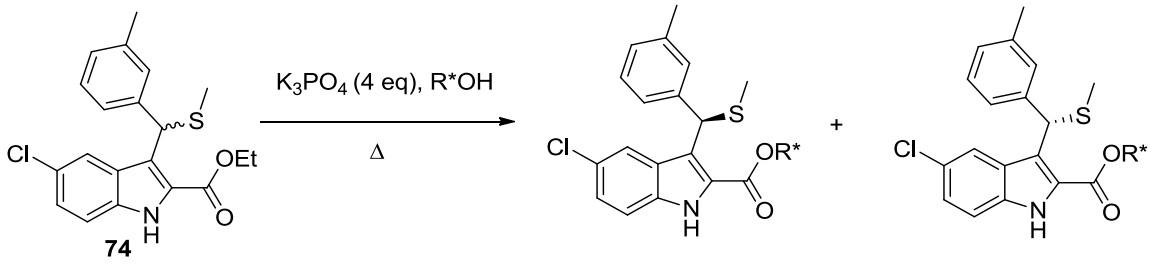
Figure 26

Having characterized compound **122** and confidently shown that we had successfully synthesized it, we now attempted to crystallize this sample to try to resolve the two diastereomers which were produced. We were disappointed that the compound appeared to be an oil however it has been shown that with the correct solvent ratios, oils can be made to crystallize. We even considered that this might be in our favor, since there was more of a chance of one diastereomer remaining soluble in the solvent system whilst the other one crashed out. Unfortunately, after extensive attempts using multiple solvent systems we were unable to form crystals. Unable to achieve crystallization, we turned to LC-MS analysis. LC-MS analysis only showed one major peak on the UV trace, with no evidence of diastereomers. A chiral column may be used in future to attempt the separation by LC-MS again.

5.4 Strategy 2B: Forming a covalent diastereomer with a chiral auxiliary on the ester in the 2-position of the indole

Another functional group on compound **74** which is easy to derivatize is the ester. Since we had managed to transesterify this ester using K_3PO_4 in the corresponding alcohol, we thought we could use this procedure to introduce a chiral ester side chain to the ester functionality (Scheme 78).

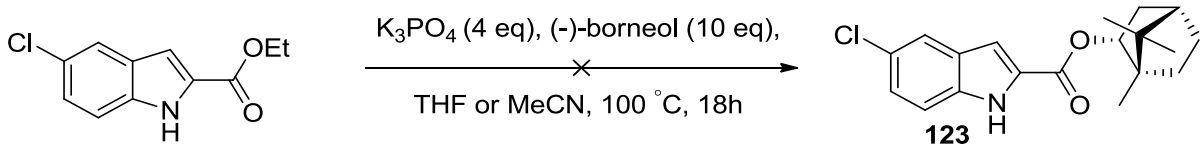
We had several different chiral alcohols available in our stores which could be used to test this strategy. (*1R,2S,5R*)-(-)-menthol and (*1S,3R,4S*)-(-)-borneol were in abundant supply, both of which contain secondary alcohols. There were no primary alcohols available, however, we did have L-proline at hand which could be modified to obtain a primary alcohol functionality. We suspected that the secondary alcohols may be less reactive in this procedure, thus it would be preferential to obtain a primary alcohol to test this strategy.



Scheme 78

5.4.1 Attempted synthesis of (*1S,2R,4S*)-1,7,7-trimethylbicyclo[2.2.1]heptan-2-yl 5-chloro-1H-indole-2-carboxylate **123**

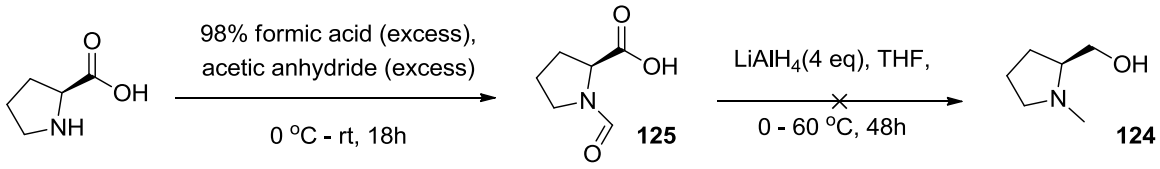
The first attempt at installing a chiral ester side chain was performed using (*1S,3R,4S*)-(-)-borneol. To establish if this procedure would work for compound **74**, we first attempted the reaction on the starting material indole ethyl 5-chloroindole-2-carboxylate which would yield compound **123** (Scheme 79). The indole was dissolved in THF and (*1S,3R,4S*)-(-)-borneol was added along with K_3PO_4. The reaction was refluxed at 100 °C. After 18h there was no sign of any product being formed. The reaction was worked-up and starting material was recovered by column chromatography. There were several reasons why this reaction may not have been successful. Firstly, this procedure had previously been used with the alcohol as both the reagent and solvent. Lack of success in this adapted procedure may be attributed to the use of THF as a solvent with 10 equivalents of solid borneol. Secondly, the less reactive secondary alcohol may have prevented the transesterification.



Scheme 79

5.4.2 Attempted synthesis of *(S)*-(1-methylpyrrolidin-2-yl)methanol **124**

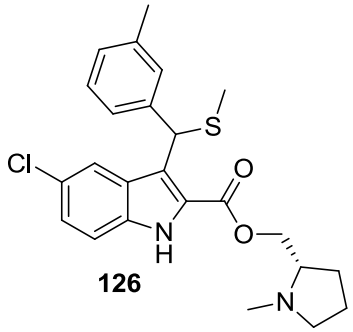
Since we had previously been successful in transesterifiying the ester in the two position of our indole using primary alcohols such as methanol and isobutanol, we believed that using a chiral primary alcohol may be more successful than the chiral secondary alcohol. Unfortunately, chiral alcohols are generally very expensive. The cheapest chiral material available is often in the form of amino acids, since they are naturally produced as only the L-enantiomer in nature. To this end we decided to derivatize L-proline to afford a primary alcohol **124** using a procedure outlined in a paper by Leete *et al*, 1991.[66] Formic acid and acetic anhydride were added together in a flask and stirred for 2h. The reaction mixture was cooled to 0 °C and L-proline was added. This was allowed to warm to rt and stirred overnight. The solution was quenched with water and concentrated *in vacuo* to reveal a yellow oil. This intermediate, assumed to be **125**, was then dissolved in THF and cooled to 0 °C. $LiAlH_4$ was added slowly. Once bubbling had stopped and all the $LiAlH_4$ had been added the mixture was heated to 60 °C for two days. The reaction mixture was cooled and quenched with ice water and dilute aqueous NaOH. A white solid precipitated out of solution but literature shows that the desired product compound **124** is a liquid. After filtering off this solid which was formed and performing EtOAc extraction on the water layer, we obtained 200 mg of an oil. ^{1}H NMR spectralanalysis did not correspond to literature values for compound **124** so we also analyzed the solid which was collected but this also did not correspond to the desired product either.



Scheme 80

In future, if this reaction is repeated no extraction should be performed on the quenched reaction mixture. This should simply be concentrated *in vacuo* before dissolving in DCM and drying over $MgSO_4$. Once concentrated again under vacuum, the oil mixture which is obtained should be distilled to obtain the product. Alternatively, column chromatography could be performed using a stain to observe the eluents on TLC since the product is not UV active and the product is much less polar than the starting materials. Unfortunately, we did not have time to repeat this reaction, but transesterification of the ethyl ester of compound **74** to the chiral proline ester derivative **126**

(Figure 27), remains a good strategy for attempting to resolve the enantiomers and should not be disregarded in future attempts.



126

Figure 27

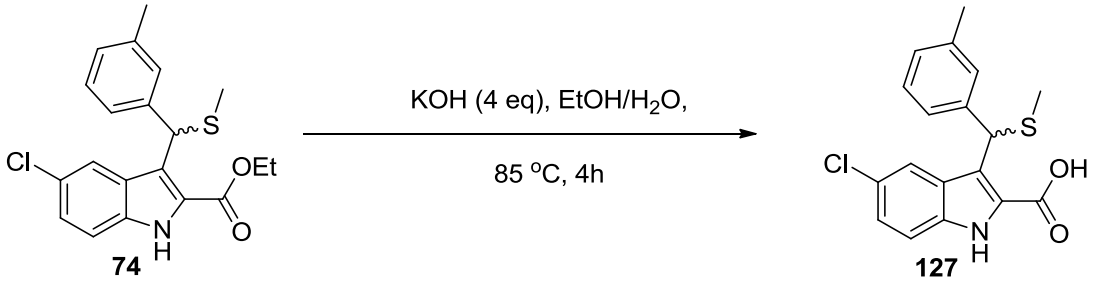
5.5 Strategy 1: Diastereomeric salt resolution

5.5.1 Synthesis of 5-chloro-3-((methylthio)(*m*-tolyl)methyl)-1H-indole-2-carboxylic acid 127

In conjunction with the covalent diastereomer resolution, we worked on forming the diastereomeric salt. This strategy involves forming the carboxylic acid derivative, compound **127** (Scheme 81), by hydrolysis of the ethyl ester of compound **74**. The carboxylic acid can then be heated in a solvent such as MeOH together with a chiral amine, to allow an acid-base reaction to occur which results in the formation of a diastereomeric salt. There are several chiral amines that are used for this purpose in synthesis. The most common are either enantiomers of ephedrine, (-)-brucine, (+)-quinine among a number of others.[60] For this technique to be successful, the salt must crystallize fairly easily in common solvents and the diastereomers produced must have a large enough difference in solubility that one crystallizes preferentially to the other in certain solvent mixtures.[60] This is very much a trial and error technique in spite of many physical chemistry papers describing ways of predicting the outcomes of these procedures. It is very rare for the first solvent system and resolving agent to give you the result which you desire.

In order to undertake the acid-base reaction to form the diastereomeric salt, we first needed to have the carboxylic acid **127** at hand, thus it was necessary to hydrolyze the ester of compound **74**. This was performed by dissolving compound **74** in EtOH and adding KOH and distilled water. The reaction mixture was heated to 85 °C and after 4h all starting material had been consumed. The reaction mixture was cooled and quenched with ice before acidifying with HCl. Although a

precipitate formed, the reaction was too small scale (265 mg) for filtration to be a viable means of purification, as it would result in a reduced yield, so the product was extracted with EtOAc and dried over $MgSO_4$. No other form of purification was performed. We assumed that since on TLC the product remained on the baseline, we had successfully made compound **127**. Furthermore, a peak at 11.80 ppm in the ^1H NMR spectrum served to confirm the presence of the carboxylic acid of compound **127**. Unfortunately, the sample was not very pure and no further analysis was performed, however we were fairly confident that the reaction had been successful. In this case the presence of water allowed for the hydrolysis to occur in these conditions, unlike what was seen earlier in pure EtOH.



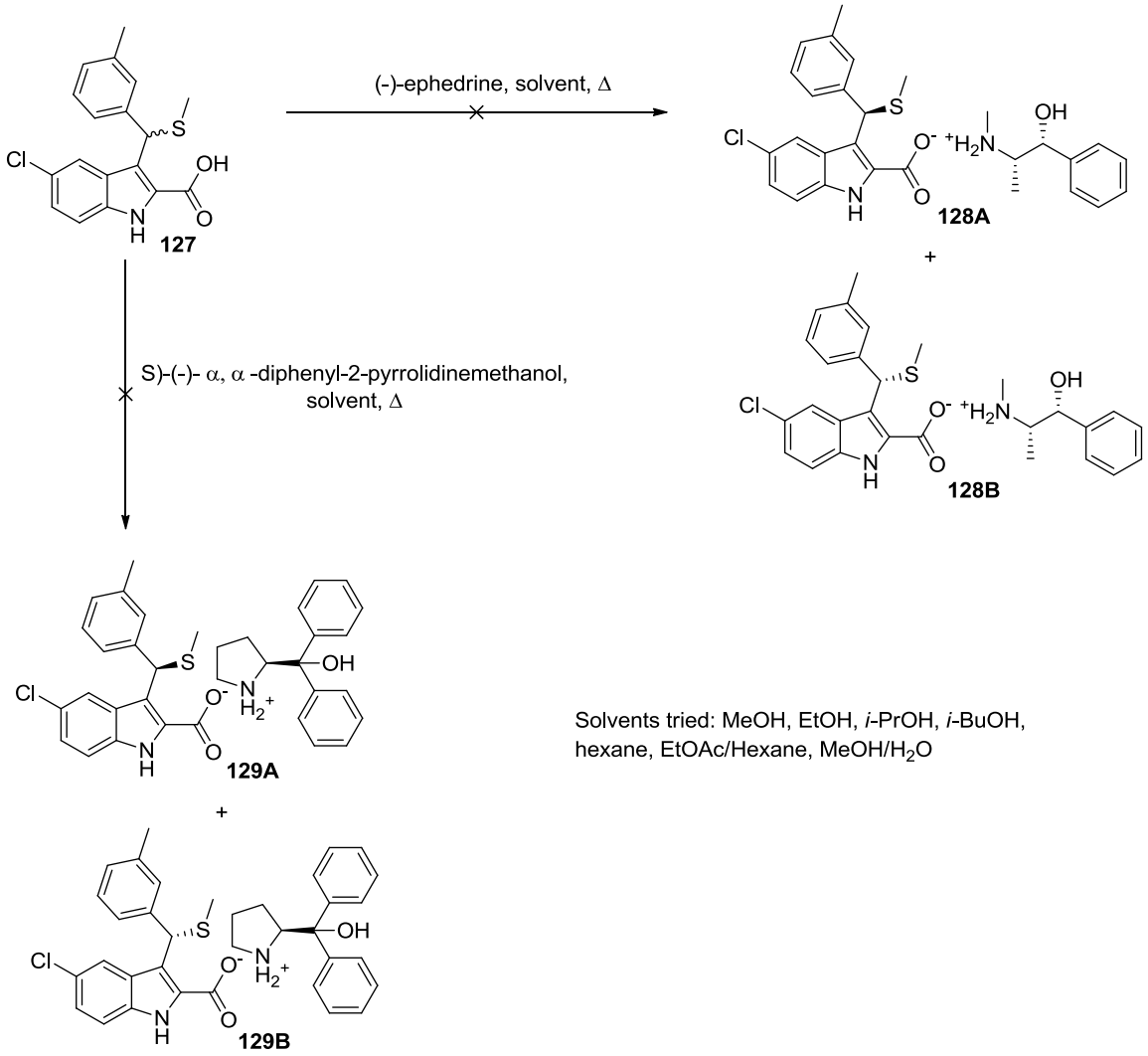
Scheme 81

5.5.2 Attempted synthesis of diastereomeric salts

With compound **127** at hand, we now wanted to see if we could prepare diastereomeric salts using different chiral resolving agents. We chose to use two chiral resolving agents which we had readily available in bulk, (-)-ephedrine and *(S)*-(-)- α,α-diphenyl-2-pyrrolidinemethanol, which would form the salts **128** and **129** respectively. One half of an equivalent of the chiral resolving agent was added to a vile along with minimum amount of solvent needed to dissolve the sample while heating. The sample was then allowed to cool to rt and if no crystals were seen after a few hours this was moved to the fridge.

Unfortunately, this strategy was not successful for us for any solvent system for either of the bases used. The reagents were either not soluble at all, the product/reagents were too soluble to form crystals, or precipitation occurred too rapidly. We could not be sure if the salt was even forming at all for these samples without being able to purify them since all reagents and products showed up as streaks on the baseline of TLC plates regardless of how polar we made the mobile phase. Furthermore, ^1H NMR spectral analysis was not useful since the samples had a mixture of different

compounds and this showed up as a large range of peaks the spectrum. We propose that if this strategy is tried again in future, a larger batch of compound **127** needs to be synthesised and purified properly before doing the trial and error run. MeOH/H_2O showed the most promise and it may be that using a much smaller ratio of H_2O to MeOH may result in crystals forming slowly as opposed to almost immediate precipitation. Also, more resolving agents may be trialed for this strategy.



Scheme 82

5.6 Concluding remarks pertaining to Chapter 5

Unfortunately, we were not successful in achieving this this aim in the project – to form a strategy for resolving the enantiomers for our lead compounds. This was an ambitious aim to begin with

and had we been successful in the short amount of time that this became a focus of the project we would have indeed been very lucky.

In future, there are a number of ways forward:

1. Continue with using a chiral auxiliary to form either covalent or ionic diastereomers that have preferential crystallisation.
2. Use an inclusion compound such as β-cyclodextrin to selectively include and crystallise one diastereomer. This technique has been shown to work with oils.
3. Using more specialized HPLC analysis with chiral columns, which may even be successful in separating the enantiomers themselves, and translating this to a preparative HPLC method.
4. Asymmetric synthesis to form one enantiomer with a high enantiomeric excess.

Although the chiral auxiliary route would be most useful if this synthesis was ever up scaled for industrial use, at this stage we are more concerned about getting efficacy data back for one or both enantiomers of any of our lead compounds. For this reason, the chiral HPLC method would most likely be our next plan of action.

Chapter 6: Conclusion

During this project, we could address the aims presented in Chapter 2 and we had considerate amount of success in working towards completing these aims.

The first aim of this project was to synthesize derivatives which would allow us to investigate the effect of changing substituents in the meta positions of the phenyl group, which interacts with Tyr181 of the NNIBP, on the ability to inhibit common resistant strains. We successfully synthesized six derivatives **10**, **13**, **19-21**, **79** which can be used towards this investigation. Unfortunately, we did not receive the results of the mutant studies before the completion of the project.

The second aim was to improve on the acid stability of the group occupying this Val179 pocket by using suitable bioisosteres in place of the acid labile methyl ether moiety. We successfully synthesized sulfide derivatives **74**, **76-78** and **80**, which were shown to have only a slight reduction in biological activity against HIV-1 compared to our original methyl ether derivative. Pleasingly, the sulfide derivatives also proved to be considerably more stable in acidic media. This indicated that the hypothesis proposed regarding use of a worse Lewis acid as a bioisostere to improve stability was correct. Interestingly, we were also able to show that the substitution on the meta position of the phenyl ring had a sizeable effect on the acid stability, in particular using a halogen in place of a methyl group. Furthermore, we synthesized the ethyl derivatives **90** and **119** which offered the advantage of eliminating the problem with acid stability entirely, thereby taking over as the lead compound if they perform well in the efficacy tests which still need to be performed.

Our third aim was to design a strategy for separating these enantiomers so that their efficacies could be attained separately. Although we tried a variety of strategies, we were not successful in resolving the enantiomers and there is still a lot more trial and error work necessary for this to be a success.

As a side investigation, we also looked into modifying the 2-position of the indole. In the original lead compound **10**, there was an ester in this position. We synthesized derivatives lacking this moiety and as expected this resulted in a significant loss of potency. Furthermore, we modified the ethyl ester to an isobutyl ester on the sulfide derivative 77 to yield compound **87**. Here we saw a reduction in potency, due to the larger isobutyl group not being as well accommodated in the NNIBP, however there was some maintenance of potency due to the presence of the ester.

Finally, we introduced an amide to this position on the ethyl derivative, however we are still awaiting the results from biological testing.

Overall, although we have not received all our biological testing results, the project was successful in synthesizing a large variety of derivatives which all contribute to our SAR data and can be used in future to further optimize the lead compound.

Chapter 7: Future Work

7.1 Review of biological activity for compounds not tested to date

Due to time constraints, a few of the compounds synthesised in the project have not yet been tested for their efficacy against HIV-1 and mutant strains of the virus. Synthesis of compound methyl ether **10**, **13**, **19-21**, **79** and sulfide derivatives **74**, **76-78**, **80** and **87** with varying meta substitution on the phenyl ring was successful and they were found to be potent against wild-type HIV-1. However, we are still awaiting mutant studies for these compounds to determine whether certain substituents on the phenyl ring influence the resistance profiles of these compounds. Furthermore, compounds **90** and **119**, the ethyl derivatives were also synthesised successfully, however we are still awaiting the assay results for these compounds to determine if they are active against wild-type HIV-1.

Once these results are in, we can use them along with our other SAR data to further optimise the lead compound and from here we may develop a suitable candidate to enter the drug development pipeline.

7.2 Extending the investigation into the ester side chain of the 2-position of the indole

In this project, we have so far determined the activity of the sulfide derivatives with an ethyl ester **74**, **76-78**, **80** and with an isobutyl ester **87**. We have established that the isobutyl ester is not well accommodated in the NNIBP, however there is still room for investigating the efficacy for derivatives containing several alternative alkyl side chains on the ester including methyl, propyl, isopropyl, butyl, *tert*-butyl and *sec*-butyl may be investigated (Figure 28). Since we have a simple method for transesterifiying our ester, refluxing with K_3PO_4 in the corresponding alcohol, this may quickly generate a small library for further developing our SAR data. The group in the 3-position of the indole would be decided based on the pending biological assay results.

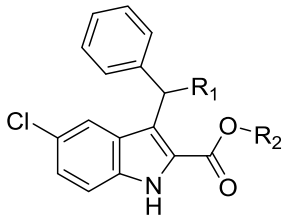


Figure 28

7.3 Developing derivatives with varying benzylic alkyl groups on the 2-position of the indole

If the ethyl derivative is found to be potent against HIV-1, varying the alkyl group in this position to probe the space available within the Val179 pocket could be achieved using the synthetic scheme outlined in chapter 4. Although this synthesis is not particularly high yielding, it does allow for fairly easy derivatisation of the alkyl group. We have already shown there is precedence for allyl and methyl derivatives to be synthesised, however, the lowest yielding reaction in the synthesis, the Grignard reaction, may need to be optimised before setting out to synthesise a full library. Methyl, propyl, isopropyl and allyl groups are all possible alkyl groups which could be installed in this position to probe the limits of this hydrophobic Val179 pocket.

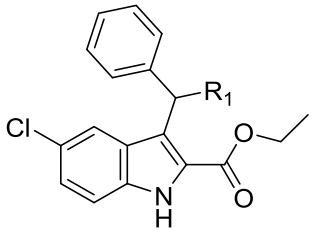


Figure 29

7.4 Resolving the enantiomers

7.4.1 Searching for a more immediate solution

As we are still currently in the development stage of this project, instead of developing a strategy for separating the enantiomers of a compound which may later be discarded as a lead compound, it may be more effective to use preparative HPLC with a chiral column. Although finding a chiral column which is suitable for separating the specific enantiomer is once again a trial and error strategy, it is far simpler and less variable than trying to selectively crystallise one diastereomer

as we tried when employing the chiral auxiliary method. To undertake this strategy, we would need to form a collaboration with one of the analytical laboratories in South Africa since the current HPLC facility that we have been utilising does not specialise in chiral chromatography, thus does not have many chiral columns available.

7.4.2 Continuing the chiral auxiliary route

Although this strategy has so far not been successful for resolving the enantiomer, once a candidate for the drug development pipeline has been established, this is a worthwhile endeavour due to its ability to be used on an industrial scale. This is an aim which should be approached as a main focus for a project and not as it was approached here, as a side focus. More success may be gained by using inclusion compounds to enforce crystallisation of oils. Using either α- or β-cyclodextrin, for example, has been successful in selectively crystallising a single enantiomer. This technique was used by Smith *et al*, 2010 to selectively crystallise four phenylurea herbicides.[67] Selective crystallisation of a single diastereomer may be easier to achieve using this technique, especially since we have so far only encountered compounds with a tendency to form oils.

Chapter 8: Experimental

8.1. General Procedures

8.1.1 Reagents and solvents

Chemicals were purchased from Merck or Sigma Aldrich. Solvents used for chromatography or work-ups were distilled by regular fractional distillation methods. Solvents used in reactions were distilled under nitrogen using appropriate drying agents. Et_2O (≥98% purity) was dried over activated molecular sieves and we did not perform distillation.

8.1.2 Chromatography

Thin layer chromatography was performed on Merck silica gel 60 F254 coated on aluminium sheets and compounds were visualised under UV light or *p*-anisaldehyde, $KMNO_4$ or ninhydrin stains.

Merck silica gel (particle size 0.063-0.200 mm, 60 Å) was used for column chromatography.

Waters Synapt G2 on a Waters BEH C18, 2.1 x 100 mm column was utilised for any HPLC analysis used in this project.

8.1.3 Spectroscopic and physical data

^{1}H and ^{13}C NMR spectra were recorded at 25 °C on one of three machines: 300 MHz Varian VNMRS (75 MHz for ^{13}C), 400 MHz Varian Unity Inova (101 MHz for ^{13}C), or 600 MHz Varian Unity Inova (150 MHz for ^{13}C). Chemical shifts were referenced using either an external reference or residual solvent peaks. The data was processed using MestReNova LITE.

A Waters SYNAPT G2 was used to perform mass spectral analysis. Thermo Nicolet Nexus 470 in Attenuated Total Reflectance (ART) mode was used to record IR data. Gallenkamp Melting Point Apparatus was used to obtain melting points which are uncorrected.

8.1.4 Other general procedures

Reaction were performed under a positive pressure of nitrogen gas, unless the solvent was water or this was otherwise stipulated. Solvents were degassed and purged with nitrogen before

reagents were introduced. All glass wear was pre-dried in a 90 °C oven for at least 18h before reactions were performed. Under vacuum refers to solvents being removed firstly by using a standard rotary evaporator followed by use of a high vacuum pump set as approximately 0.08 mm Hg.

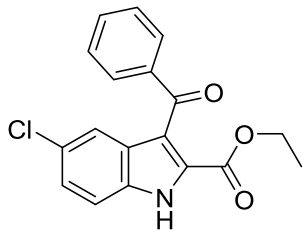
8.2. Experimental Procedures Pertaining to Chapter 3

8.2.1 3,5-dimethyl benzoyl chloride 30

3,5-dimethylbenzoic acid (1.00 g, 6.65 mmol) was added to a 50 mL RBF containing 10 mL of $SOCl_2$. The solution was refluxed at 80 °C for 24 h then concentrated under vacuum. The product was used crude in the next reaction.

8.2.2 Ethyl 3-benzoyl-5-chloro-1*H*-indole-2-carboxylate 29

Benzoyl chloride was added to a 100 mL RBF charged with 60 mL of DCE (3 eq., 1.60 mL, 13.5 mmol). $AlCl_3$ (3 eq., 1.80 g, 13.5 mmol) was added at 0 °C and this solution was stirred at 0 °C for 30 minutes. Following the addition of ethyl-5-chloro-1*H*-indole-2-carboxylate (1.00 g, 4.50 mmol), the reaction mixture was refluxed at 90 °C for 3 hours. Once cooled to rt, the reaction mixture was quenched with sat. $NaHCO_3$. The resulting emulsion was filtered through celite, whereupon the filtrate was extracted with EtOAc (3 × 40 mL). The organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude product was purified by column chromatography (5-30% EtOAc/Hex). The reaction yielded compound **29** as a yellow solid, 898 mg, 2.74 mmol, 61% (R_f = 0.37, 20% EtOAc/Hex).

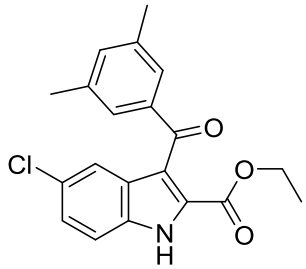


^{1}H NMR (600 MHz, $CDCl_3$) δ 9.56 (s, 1H, NH), 7.86 (dd, J = 7.1, 1.1 Hz, 2H, ArH), 7.72 (d, J = 2.0 Hz, 1H, ArH), 7.59 – 7.55 (m, 1H, ArH), 7.46 – 7.40 (m, 3H, ArH), 7.33 (dd, J = 8.8, 2.0 Hz, 1H, ArH), 4.05 (q, J = 7.1 Hz, 2H, CH_2), 0.87 (t, J = 7.1 Hz, 3H, CH_3). **^{13}C NMR (151 MHz, $CDCl_3$) δ** 192.53 (CO), 161.19 (COO), 139.34, 133.95, 133.18, 129.58, 128.55, 128.35, 128.29, 127.71, 126.98, 121.49, 119.40, 113.26, 61.94 (CH_2), 13.49 (CH_3).

8.2.3 Ethyl 5-chloro-3-(3,5-dimethylbenzoyl)-1*H*-indole-2-carboxylate 32

The same procedure was used as per the synthesis of **29**. The following equivalents were used: Compound **30** (2.20 eq., 1.10 g, 6.60 mmol), ethyl-5-chloro-1*H*-indole-2-carboxylate (700 mg, 3.10

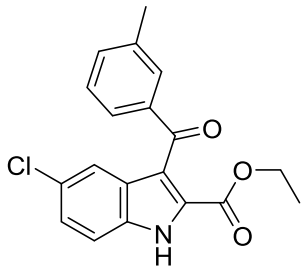
mmol), $AlCl_3$ (3 eq., 887 mg, 9.50 mmol). The reaction yielded compound **32** as an off-white powder, 624 mg, 1.75 mmol, 56% (R_f = 0.35, 20% EtOAc/Hex).



^{1}H NMR (300 MHz, $CDCl_3$) δ 9.80 (s, 1H, NH), 7.69 – 7.59 (m, 1H, ArH), 7.41 – 7.30 (m, 3H, ArH), 7.27 – 7.16 (m, 1H, ArH), 7.13 (s, 1H, ArH), 4.07 – 3.93 (m, 2H, CH_3), 2.25 (s, 6H, 2 × CH_3), 0.88 – 0.79 (m, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 192.93 (CO), 161.33 (COO), 139.26, 138.17, 138.07, 135.38, 134.77, 133.92, 128.20, 127.98, 127.90, 128.47, 127.25, 126.70, 121.27, 119.74, 113.24, 61.81 (CH_2), 21.15 (CH_3), 13.37 (CH_3).

8.2.4 Ethyl 5-chloro-3-(3-methylbenzoyl)-1*H*-indole-2-carboxylate **33**

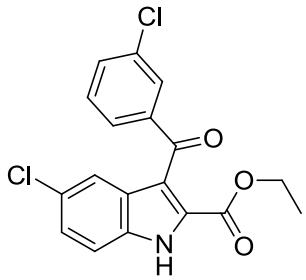
The same procedure was used as per the synthesis of **29**. The following equivalents were used: *m*-toluoyl chloride (3 eq., 1.80 mL, 13.5 mmol), ethyl-5-chloro-1*H*-indole-2-carboxylate (1.00 g, 4.50 mmol), $AlCl_3$ (3 eq., 1.70 g, 13.5 mmol). The reaction yielded compound **33** as a yellow solid, 991 mg, 2.90 mmol, 65% (R_f = 0.35, 30% EtOAc/Hex).



Mp 156-162 °C **IR (ATR, cm^{-1}):** 3293 (N-H str), 1693 (C=O str), 1647 (C = O str), 1598 (N – H bend), 1256, 1246, 1178, 801 (C – Cl str) **^{1}H NMR (400 MHz, $CDCl_3$) δ** 9.40 (s, 1H, NH), 7.72 – 7.68 (s, 2H, ArH), 7.61 (d, *J* = 7.6 Hz, 1H, ArH), 7.44 – 7.36 (m, 2H, ArH), 7.35 – 7.29 (m, 2H, ArH), 4.07 (q, *J* = 7.2 Hz, 2H, CH_2), 2.39 (s, 3H, CH_3), 0.90 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 192.70 (CO), 161.15 (COO), 139.32, 138.38, 133.99, 133.87, 129.86, 128.42, 128.24, 127.65, 127.04, 126.95, 121.51, 119.66, 113.21, 61.90 (CH_2), 21.43 (CH_3), 13.52 (CH_3). **HRMS:** Calculated for $C_{19}H_{16}ClNO_3$ $[M + H]^+$, 342.0897, found 342.0891.

8.2.5 Ethyl 5-chloro-3-(3-chlorobenzoyl)-1*H*-indole-2-carboxylate **34**

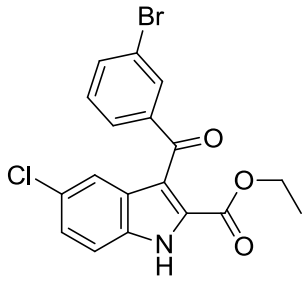
The same procedure was used as per the synthesis of **29**. The following equivalents were used: 3-chlorobenzoyl chloride (3 eq., 2.57 mL, 20.1 mmol), ethyl-5-chloro-1*H*-indole-2-carboxylate (1.50 g, 6.70 mmol), $AlCl_3$ (3 eq., 2.70 g, 20.1 mmol). The reaction yielded **34** as a yellow solid, 1.52 g, 4.20 mmol, 62%, (R_f = 0.22, 20% EtOAc/Hex).



Mp 180-182 °C **IR (ATR, cm^{-1}):** 3288 (N-H str), 1693 (C=O str), 1633 (C = O str), 1597 (N – H bend), 1259, 762 (C-Cl str). **^{1}H NMR (400 MHz, $CDCl_3$) δ** 9.41 (s, 1H, NH), 7.86 – 7.84 (m, 1H, ArH), 7.75 – 7.68 (m, 2H, ArH), 7.57 – 7.53 (m, 1H, ArH), 7.45 – 7.33 (m, 3H, ArH), 4.09 (q, *J* = 7.1 Hz, 2H, CH_2), 0.94 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 191.17 (CO), 160.87 (COO), 141.13, 134.96, 133.94, 133.05, 129.93, 129.47, 128.68, 128.41, 127.96, 127.72, 127.28, 121.49, 118.74, 113.38, 62.14 (CH_2), 13.67. **HRMS:** Calcd for $C_{18}H_{13}Cl_2NO_3$ $[M + H]^+$, 362.0351, found 362.0344.

8.2.6 Ethyl 3-(3-bromobenzoyl)-5-chloro-1*H*-indole-2-carboxylate **35**

The same procedure was used as per the synthesis of **29**. The following equivalents were used: 3-bromobenzoyl chloride (3.0 eq., 2.70 mL, 20.1 mmol), ethyl-5-chloro-1*H*-indole-2-carboxylate (1.50 g, 6.70 mmol), $AlCl_3$ (3 eq., 2.70 g, 20.1 mmol). The reaction yielded compound **35** as a yellow solid, 1.20 g, 2.95 mmol, 44 % (R_f = 0.21, 20% EtOAc/Hex).



Mp 172-174 °C **IR (ATR, cm^{-1}:** 3309-3290 (N-H str), 1692 (C=O str), 1632 (C=O str), 1562 (N – H str), 1258, 760 (C – Cl str). **^{1}H NMR (400 MHz, $CDCl_3$) δ** 9.63 (s, 1H, NH), 8.02 – 7.98 (m, 1H, ArH), 7.78 – 7.67 (m, 3H, ArH), 7.44 (d, *J* = 8.8 Hz, 1H, ArH), 7.37 – 7.30 (m, 2H, ArH), 4.09 (q, *J* = 7.2 Hz, 2H, CH_2), 0.94 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 191.02 (CO), 160.89 (COO), 141.27, 135.89, 133.91, 132.33, 130.12, 128.62, 128.33, 128.09, 127.91, 127.20, 122.84, 121.41, 118.59, 113.34, 62.11 (CH_2), 13.60 (CH_3). **HRMS:** Calcd for $C_{18}H_{13}BrClNO_3$ $[M + H]^+$, 405.9846, found 405.9857.

8.2.7 Attempted synthesis of ethyl 5-chloro-3-((3,5-dinitrophenyl)(methoxy)methyl)-1H-indole-2-carboxylate **36**

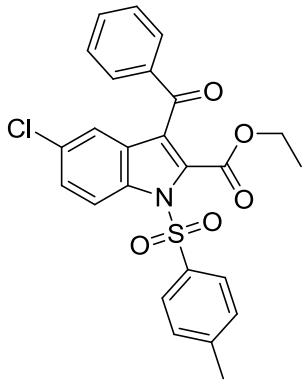
The same procedure was used as per the synthesis of **29**. The following equivalents were used: 3,5-dinitrobenzoyl chloride (3 eq., 3.11 g, 13.5 mmol), ethyl-5-chloro-1*H*-indole-2-carboxylate (1.00 g, 4.50 mmol), $AlCl_3$ (3 eq., 1.70 g, 13.5 mmol). Following purification, the reaction yielded 100 mg of a yellow solid. The characterization did not correspond to the desired product **36**.

8.2.8 Attempted synthesis of ethyl 3-((3,5-*bis*-(trifluoromethyl)phenyl)(methoxy)methyl)-5-chloro-1H-indole-2-carboxylate **37**

The same procedure was used as per the synthesis of **29**. The following equivalents were used: 3,5-*bis*-(trifluoromethyl)benzoyl chloride (3 eq., 2.50 mL, 13.5 mmol), ethyl-5-chloro-1*H*-indole-2-carboxylate (1.00 g, 4.50 mmol), $AlCl_3$ (3 eq., 1.70 g, 13.5 mmol). The reaction yielded 800 mg of a cream coloured solid. The characterization did not correspond to the desired product **37**.

8.2.9 Ethyl 3-benzoyl-5-chloro-1-tosyl-1H-indole-2-carboxylate 38

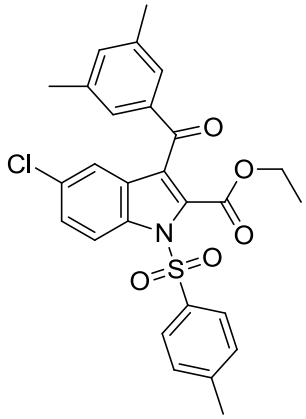
Compound **29** (1.00 g, 3.10 mmol) was added to a 100 mL RBF charged with 40 mL of DMF. NaH (1.50 eq., 60%, 155 mg, 4.60 mmol) was added to the flask at 0 °C where upon a color change from yellow to bright orange was observed. The reaction mixture was stirred for 15 minutes at 0 °C before *p*-TsCl (1.50 eq., 880 mg, 4.60 mmol) was added and a color change back to yellow was observed. The reaction mixture was heated to 40 °C. TLC analysis showed all starting material had been consumed after 18 hours. The reaction mixture was then quenched with a sat. NH_4Cl (50 mL) and extracted with EtOAc (3 × 30 mL). The crude extract was purified by column chromatography (10-30% EtOAc/Hex) to afford compound **38** as a viscous yellow oil, 1.34 g, 2.78 mmol, 90 % (R_f = 0.46, 30% EtOAc/Hex).



^{1}H NMR (400 MHz, $CDCl_3$) δ 8.04 – 7.93 (m, 3H, ArH), 7.79 – 7.72 (m, 2H, ArH), 7.65 – 7.58 (m, 1H, ArH), 7.56 – 7.53 (m, 1H, ArH), 7.50 – 7.44 (m, 2H, ArH), 7.39 (dd, J = 9.0, 2.1 Hz, 1H, ArH), 7.34 – 7.30 (m, 2H, ArH), 4.08 (q, J = 7.2 Hz, 2H, CH_2), 2.41 (s, 3H, CH_3), 1.18 (t, J = 7.2 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 190.56 (CO), 161.01 (COO), 146.30, 138.13, 133.53, 130.87, 130.21, 129.83, 129.34, 129.23, 129.10, 128.82, 128.69, 127.93, 127.31, 121.93, 115.84, 115.13, 63.06 (CH_2), 21.87 (CH_3), 13.63 (CH_3).

8.2.10 Ethyl 5-chloro-3-(3,5-dimethylbenzoyl)-1-tosyl-1*H*-indole-2-carboxylate 39

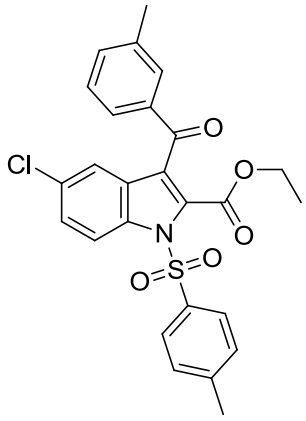
The same procedure was used as per the synthesis of **38**. The following equivalents were used: Compound **32** (270 mg, 0.760 mmol), NaH (2 eq., 40.0 mg, 1.20 mmol), *p*-TsCl (2 eq., 217 mg, 1.14 mmol). The reaction yielded compound **39** as a cream-colored oil, 342 mg, 0.67 mmol, 88 % (R_f = 0.50, 30% EtOAc/Hex).



^{1}H NMR (400 MHz, $CDCl_3$) δ 8.02 – 7.93 (m, 3H, ArH), 7.55 (d, *J* = 1.9 Hz, 1H, ArH), 7.41 – 7.34 (m, 3H, ArH), 7.31 (d, *J* = 8.2 Hz, 2H, ArH), 7.23 (s, 1H, ArH), 4.09 (q, *J* =7.2 Hz, 2H, CH_2), 2.39 (s, 3H, CH_3), 2.34 (s, 6H, 2 × CH_3), 1.19 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 190.87 (CO), 161.05 (COO), 146.23, 138.40, 138.10, 135.18, 134.84, 134.38, 133.84, 130.83, 130.16, 128.72, 127.86, 127.23, 127.13, 122.28, 121.95, 115.84, 63.00 (CH_2), 21.82 (CH_3), 21.24 (CH_3), 13.60 (CH_3).

8.2.11 Ethyl 5-chloro-3-(3-methylbenzoyl)-1-tosyl-1*H*-indole-2-carboxylate 40

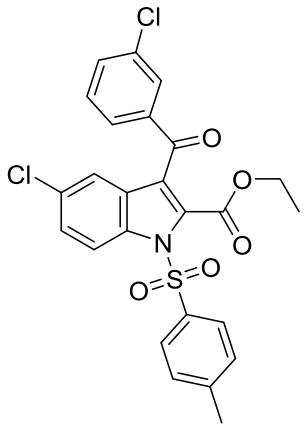
The same procedure was used as per the synthesis of **38**. The following equivalents were used: Compound **33** (630 mg, 1.80 mmol), NaH (2 eq., 127 mg, 3.60 mmol), *p*-TsCl (1.5 eq., 686 mg, 2.70 mmol). The reaction yielded compound **40** as a yellow oil, 780 mg, 1.57 mmol, 87% (R_f = 0.34, 20% EtOAc/Hex).



IR (ATR, cm^{-1}): 3309-2924 (aromatic str), 1733 (C=O str), 1648 (C = O str), 1372 (S = O str), 1174 (S=O str), 764 (C – Cl str). **^{1}H NMR (400 MHz, $CDCl_3$)** 8.02 – 7.94 (m, 3H, ArH), 7.63 – 7.52 (m, 3H, ArH), 7.44 – 7.29 (m, 5H, ArH), 4.10 (q, *J* = 7.2 Hz, 2H, CH_2), 2.44 – 2.40 (m, 6H, 2 × CH_3) 1.18 (t, *J* = 7.2 Hz, 2H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 190.73 (CO), 161.05 (COO), 146.27, 138.63, 138.12, 134.92, 134.48, 134.32, 133.86, 130.87, 130.20, 129.78, 128.66, 128.55, 127.92, 127.28, 126.68, 122.06, 121.97, 115.85, 63.04 (CH_2), 21.87 (CH_3), 21.40 (CH_3), 13.63 (CH_3). **HRMS:** Calcd for $C_{26}H_{22}ClNO_5S$ $[M + H]^+$, 496.0907, found 496.0985.

8.2.12 Ethyl 5-chloro-3-(3-chlorobenzoyl)-1-tosyl-1*H*-indole-2-carboxylate 41

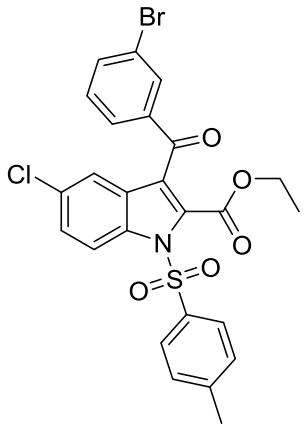
The same procedure was used as per the synthesis of **38**. The following equivalents were used: Compound **34** (1.00 g, 2.76 mmol), NaH (2 eq., 146 mg, 6.08 mmol), *p-TsCl* (1.5 eq., 756 mg, 3.97 mmol). The reaction yielded compound **41** as a yellow oil, 1.27 g, 2.46 mmol, 89% (R_f = 0.35, 20% EtOAc/Hex).



IR (ATR, cm^{-1}: 2980 (aromatic str), 1738 (C=O str), 1646 (C=O str), 1374 (S = O str), 1203 (S = O str), 803 (C – Cl str). **^{1}H NMR (400 MHz, $CDCl_3$) δ** 7.98 (dd, *J* = 19.9, 8.7 Hz, 3H, ArH), 7.77 – 7.72 (m, 1H, ArH), 7.64 – 7.56 (m, 3H, ArH), 7.45 – 7.37 (m, 2H, ArH), 7.32 (d, *J* = 8.2 Hz, 2H, ArH), 4.12 (q, *J* = 7.2 Hz, 2H, CH_2), 2.41 (s, 3H, CH_3), 1.23 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 189.25 (CO), 160.96 (COO), 146.46, 139.74, 135.39, 135.02, 134.30, 133.75, 133.35, 131.13, 130.27, 130.03, 129.14, 128.36, 127.92, 127.47, 127.47, 121.85, 120.97, 115.89, 63.29 (CH_2), 21.89 (CH_3), 13.66 (CH_3). **HRMS:** Calcd for $C_{25}H_{19}Cl_2NO_5S$ [M + H]$^+$, 516.0439, found 516.0432.

8.2.13 Ethyl 3-((3-bromophenyl)(methoxy)methyl)-5-chloro-1-tosyl-1H-indole-2-carboxylate **42**

The same procedure was used as per the synthesis of **38**. The following equivalents were used: Compound **35** (1.18 g, 2.90 mmol), NaH (1.5 eq., 60%, 146 mg, 4.35 mmol), *p*-TsCl (1.5 eq., 830 mg, 4.35 mmol). The reaction yielded compound **42** as a yellow oil, 487 mg, 0.868 mmol, 30% (R_f = 0.40, 20% EtOAc/Hex).



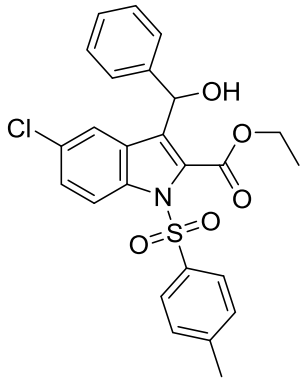
IR (ATR, cm^{-1}): 3064-2924 (aromatic str), 1736 (C=O str), 1652 (C=O str), 1372 (S = O str), 1177 (S = O str), 795 (C – Cl str), 666 (C- Br str). **^{1}H NMR (400 MHz, $CDCl_3$) δ** 8.00 (d, J = 8.7 Hz, 1H, ArH), 7.95 (d, J = 8.4 Hz, 2H, ArH), 7.90 (dd, J = 3.4, 1.7 Hz, 1H, ArH), 7.76 – 7.71 (m, 1H, ArH), 7.67 – 7.63 (m, 1H, ArH), 7.58 (d, J = 2.1 Hz, 1H, ArH), 7.40 (dd, J = 9.0, 2.1 Hz, 1H), 7.37 – 7.29 (m, 3H), 4.12 (q, J = 7.2 Hz, 2H, H_9), 2.41 (s, 3H, H_{22}), 1.24 (t, J = 7.2 Hz, 3H, H_{10}). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 189.13 (CO), 160.95 (COO), 146.45, 139.93, 136.26, 135.42, 134.29, 133.73, 132.04, 131.13, 130.27, 130.25, 128.35, 127.92, 127.90, 127.47, 122.94, 121.85, 120.89, 115.88, 63.32 (CH_2), 21.89 (CH_3), 13.68 (CH_3). **HRMS:** Calculated for $C_{25}H_{20}BrClNO_5S$[M+H]$^+$, 559.9929, found 559.9934.

8.2.14 Attempted synthesis of Ethyl 5-chloro-3-(3,5-dinitrobenzoyl)-1-tosyl-1H-indole-2-carboxylate 43

The same procedure was used as per the synthesis of **38**. The following equivalents were used: Compound **17** (1.00 g, 2.39 mmol), NaH (1.5 eq., 60%, 120 mg, 3.60 mmol), *p*-TsCl (1.5 eq., 683 mg, 3.60 mmol). The reaction yielded a bright yellow solid. The characterization did not correspond to the desired product, compound **43**.

8.2.15 Ethyl 5-chloro-3-(hydroxy(phenyl)methyl)-1-tosyl-1*H*-indole-2-carboxylate 12

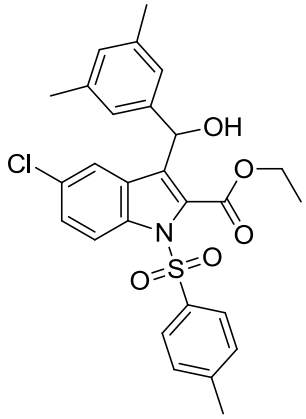
Compound **38** (1.25 g, 2.6 mmol) was added to a 100 mL RBF charged with 5 mL of THF and MeOH (30 mL) was added to this solution. $NaBH_4$ (3 eq., 295 mg, 7.80 mmol) was added at 0 °C and the reaction mixture was allowed to reach r.t. TLC analysis confirmed all starting material had been consumed after 18 h. The reaction mixture was quenched with a sat. NH_4Cl (50 mL) and extracted with EtOAc (3 × 40 mL). The organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude mixture was purified by column chromatography (5 – 20% EtOAc/Hex). The reaction yielded compound **12** as a yellow solid, 971 mg, 2.00 mmol, 77% (R_f = 0.22, 20% EtOAc/Hex).



^{1}H NMR (300 MHz, $CDCl_3$) δ 7.84 (d, *J* = 8.9 Hz, 1H, ArH), 7.69 – 7.62 (m, 2H, ArH), 7.36 (d, *J* = 1.9 Hz, 1H, ArH), 7.33 – 7.28 (m, 2H, ArH), 7.25 – 7.15 (m, 4H, ArH), 7.12 – 7.07 (m, 2H, ArH), 6.02 (d, *J* = 4.8 Hz, 1H, CH), 4.36 (q, *J* = 7.1 Hz, 2H, CH_2), 2.74 (d, *J* = 4.8 Hz, 1H, OH), 2.25 (s, 3H, CH_3), 1.30 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 162.91 (COO), 145.59, 141.26, 135.18, 134.07, 130.17, 129.89, 129.84, 129.56, 129.19, 128.67, 128.63, 128.53, 127.99, 127.35, 127.10, 126.23, 121.80, 116.51, 69.02, 63.03 (CH_2), 21.77 (CH_3), 14.08 (CH_3).

8.2.16 Ethyl 5-chloro-3-((3,5-dimethylphenyl)(hydroxy)methyl)-1-tosyl-1*H*-indole-2-carboxylate **44**

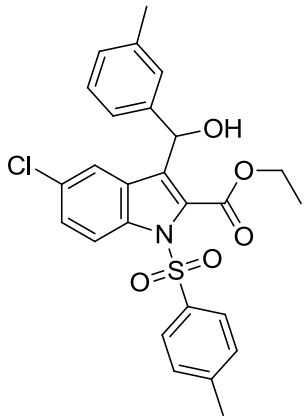
The same procedure was used as per the synthesis of **12.** The following equivalents were used: Compound **39** (500 mg, 0.980 mmol), $NaBH_4$ (3 eq., 111 mg, 2.94 mmol). The reaction yielded compound **44** as a light yellow solid, 386 mg, 0.754 mmol, 77% (R_f = 0.25, 20% EtOAc/Hex).



Mp 150—156 °C **^{1}H NMR (400 MHz, $CDCl_3$) δ** 7.93 (d, *J* = 8.9 Hz, 1H, ArH), 7.79 – 7.74 (m, 2H, ArH), 7.43 (d, *J* = 1.9 Hz, 1H, ArH), 7.28 (dd, *J* = 8.73, 2.15, 1H, ArH), 7.19 (d, *J* = 8.1 Hz, 2H, ArH), 7.01 (s, 2H, ArH), 6.89 (s, 1H, ArH), 6.01 (s, 1H, CH), 4.52 – 4.37 (m, 2H, CH_2), 2.46 – 2.36 (m, 1H, OH), 2.33 (s, 3H, CH_3), 2.27 (s, 6H, 2 × CH_3), 1.40 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 163.08 (COO), 145.55, 141.17, 138.24, 135.01, 134.14, 130.10, 129.84, 129.75, 129.25, 128.41, 127.37, 126.99, 124.07, 121.74, 116.38, 77.48, 77.16, 76.84, 69.24 (CH), 62.97 (CH_2), 21.75 (CH_3), 21.51 (CH_3), 14.07 (CH_3), 9.39 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{27}H_{25}ClNO_4S$ $[M - OH]^+$, 494.1193, found 494.1100.

8.2.17 Ethyl 5-chloro-3-(hydroxy(*m*-tolyl)methyl)-1-tosyl-1*H*-indole-2-carboxylate **45**

The same procedure was used as per the synthesis of **12.** The following equivalents were used: Compound **40** (139 mg, 0.271 mmol), $NaBH_4$ (3 eq., 46.0 mg, 1.20 mmol). The reaction yielded compound **45** as a cream colored solid, 113 mg, 0.226 mmol, 84% (R_f = 0.32, 20% EtOAc/Hex).

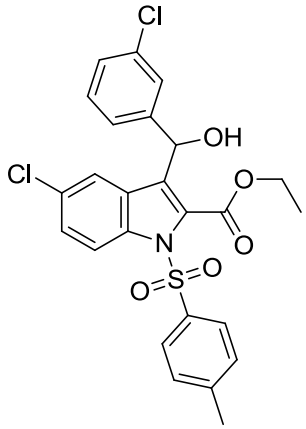


Mp 80-82 °C **IR (ATR, cm-1):** 3457 (O – H str), 3100 – 2920 (C – H str), 1741 (C = O str), 1441 (C – H bend), 1366, (C – H bend), 1229 (S = O str), 1206 (S = O str), **^{1}H NMR (600 MHz, $CDCl_3$) δ** 7.93 (d, *J* = 8.9 Hz, 1H, ArH), 7.75 (d, *J* = 8.3 Hz, 2H, ArH), 7.43 (d, *J* = 1.7 Hz, 1H, ArH), 7.31 – 7.15 (m, 6H, ArH), 7.09 – 7.04 (m, 1H, ArH), 6.06 (s, 1H, CH), 4.45 (q, *J* = 7.1 Hz, 2H, CH_2), 2.67 (s, 1H, OH), 2.34 (s, 3H, CH_3), 2.31 (s, 3H, CH_3) 1.40 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (151 MHz, $CDCl_3$) δ** 162.98 (COO), 145.57, 141.21, 138.35, 135.10, 134.15, 130.15, 129.85, 129.23, 128.82, 128.51, 127.38, 127.05, 126.94, 123.33, 121.78, 116.45, 69.15, 63.00 (CH), 31.06 (CH_2), 21.82 (CH_3), 21.65 (CH_3), 16.81 (CH_3), 14.08 (CH_3). **HRMS:** Note: No parent molecular ion was

present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{26}H_{23}ClNO_4S$ $[M - OH]^+$, 480.1036, found 480.1035.

8.2.18 Ethyl 5-chloro-3-((3-chlorophenyl)(hydroxy)methyl)-1-tosyl-1*H*-indole-2-carboxylate **46**

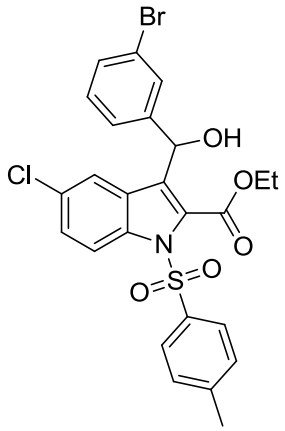
The same procedure was used as per the synthesis of **12.** The following equivalents were used: Compound **41** (1.50 g, 2.42 mmol), $NaBH_4$ (3 eq., 275 mg, 7.26 mmol). The reaction yielded compound **46** as a yellow solid, 796 mg, 1.54 mmol, 64% (R_f = 0.32, 20% EtOAc/Hex).



Mp 90-95 °C **IR (ATR, cm-1):** 3480 – 3444 (O – H str), 1716 (C = O str), 1371 (S = O str), 1181 (S = O str), 667 (C- Cl str) **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.85 (d, *J* = 8.9 Hz, 1H, ArH), 7.65 – 7.57 (m, 2H, ArH), 7.38 - 7.32 (m, 2H, ArH), 7.25 – 7.09 (m, 6H, ArH), 5.99 (d, *J* = 4.8 Hz, 1H, CH), 4.38 (q, *J* = 7.1 Hz, 2H, CH_2), 2.83 (d, *J* = 4.8 Hz, 1H, OH), 2.26 (s, 3H, CH_3), 1.33 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 162.78 (COO), 145.75, 143.30, 135.27, 134.61, 133.88, 130.36, 130.15, 129.92, 129.89, 128.92, 128.27, 128.11, 127.34, 127.29, 126.39, 124.24, 121.67, 116.67, 68.17 (CH), 63.19 (CH_2), 60.55, 21.79, 14.35, 14.09. **HRMS:** Calcd for $C_{25}H_{21}Cl_2NO_5S$ $[M + Na]^+$, 540.0415, found 540.0401.

8.2.19 Ethyl 3-((3-bromophenyl)(hydroxy)methyl)-5-chloro-1-tosyl-1H-indole-2-carboxylate **47**

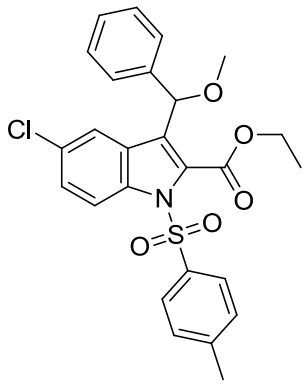
The same procedure was used as per the synthesis of **12.** The following equivalents were used: Compound **42** (450 mg, 0.80 mmol), sodium borohydride (3 eq., 92.0 mg, 2.40 mmol). The reaction yielded compound **47** as an off white solid, 196 mg, 0.35 mmol, 44 %)(R_f = 0.24, 20% EtOAc/Hex).



Mp 82-85 °C. **IR (ATR, cm^{-1}):** 3454 (O – H str), 1714 (C = O str), 1370 (S = O str), 1180 (S = O str), 668 (C – Br str) **^{1}H NMR (300 MHz, CDCl$_3$) δ** 7.93 (d, *J* = 8.9 Hz, 1H), 7.70 (d, *J* = 8.3 Hz, 2H), 7.64 – 7.09 (m, 9H), 6.07 (s, 1H), 4.48 (q, *J* = 7.3 Hz, 2H), 2.41(s, 1H), 2.35 (s, 3H), 1.27 (t, *J* = 7.3 Hz, 3H). **^{13}C NMR (75 MHz, CDCl$_3$) δ** 162.79 (COO), 145.72, 143.58, 135.18, 133.81, 130.96, 130.71, 130.29, 130.21, 130.13, 129.97, 129.26, 128.86, 128.21, 127.28, 125.42, 124.68, 122.73, 121.69, 116.56, 68.00, 63.19 (CH), 60.57 (CH_2), 21.77 (CH_3), 14.06 (CH_3). Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_2H_{20}ClNO_5S$ [M - OH]$^+$, 543.9985, found 543.9984.

8.2.20 Synthesis of ethyl 5-chloro-3-(methoxy(phenyl)methyl)-1-tosyl-1H-indole-2-carboxylate **48**

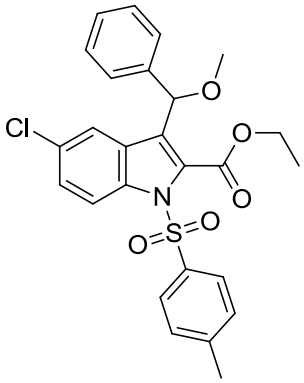
Compound **12** (100 mg, 0.21 mmol) was dissolved in DMF (10 mL) in a 100 mL RBF. NaH (2 eq., 8.5 mg, 0.25 mmol) was added at 0 °C and the reaction mixture was stirred on ice for 15 mins. MeI (10 eq., 0.13 mL, 20.1 mmol) was added and the solution was allowed to reach rt. After 20h there were multiple spots on TLC, so the reaction was quenched with sat. NH_4Cl and stirred for 1h before extracting with EtOAc (3 x 20 mL). The crude product was purified by column chromatography (10-30% EtOAc/Hex) to afford compound **48** as an off white solid, 4.0 mg, 0.008 mmol, 4% (R_f = 0.30, 20% EtOAc/Hex).



^{1}H NMR (600 MHz, CDCl$_3$) δ 7.84 (d, *J* = 8.9 Hz, 1H, ArH), 7.66 (d, *J* = 8.3 Hz, 2H, ArH), 7.44 (d, *J* = 1.8 Hz, 1H, ArH), 7.31 (d, *J* = 7.6 Hz, 2H, ArH), 7.23 – 7.13 (m, 4H, ArH), 7.09 (d, *J* = 8.1 Hz, 2H, ArH), 5.51 (s, 1H, CH), 4.44 (q, *J* = 7.2 Hz, 2H, CH_2), 3.24 (s, 3H, CH_3), 2.26 (s, 3H, CH_3), 1.37 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (151 MHz, CDCl$_3$) δ** 162.39 (COO), 145.52, 139.83, 135.35, 133.96, 131.02, 130.18, 129.78, 129.05, 128.48, 127.80, 127.37, 127.03, 126.56, 126.47, 122.34, 116.47, 77.84, 62.87 (CH), 57.27 (CH_2), 21.76 (CH_3), 14.22 (CH_3).

8.2.21 Synthesis of ethyl 5-chloro-3-(methoxy(phenyl)methyl)-1-tosyl-1H-indole-2-carboxylate **48**

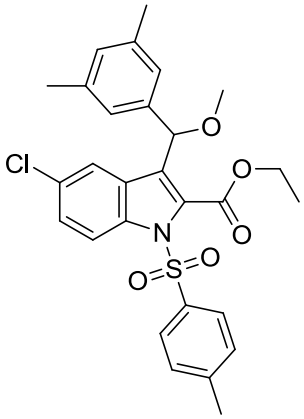
Compound **12** (670 mg, 1.39 mmol) was dissolved in THF (5 mL) in a 100 mL three necked RBF. MeOH (30 mL) was added followed by *p*-TsOH (3 eq., 793 mg, 4.20 mmol) and the reaction was refluxed at 50 °C for 18 hours. The reaction mixture was quenched with sat. $NaHCO_3$ and extracted with EtOAc (3 × 20 mL). The crude product was purified by column chromatography (5-30% EtOAc/Hex) to afford compound **48** as an off white solid, 430 mg, 0.860 mmol, 62% (R_f = 0.30, 20% EtOAc/Hex).



^{1}H NMR (600 MHz, $CDCl_3$) δ 7.84 (d, *J* = 8.9 Hz, 1H, ArH), 7.66 (d, *J* = 8.3 Hz, 2H, ArH), 7.44 (d, *J* = 1.8 Hz, 1H, ArH), 7.31 (d, *J* = 7.6 Hz, 2H, ArH), 7.23 – 7.13 (m, 4H, ArH), 7.09 (d, *J* = 8.1 Hz, 2H, ArH), 5.51 (s, 1H, CH), 4.44 (q, *J* = 7.2 Hz, 2H, CH_2), 3.24 (s, 3H, CH_3), 2.26 (s, 3H, CH_3), 1.37 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (151 MHz, $CDCl_3$) δ** 162.39 (COO), 145.52, 139.83, 135.35, 133.96, 131.02, 130.18, 129.78, 129.05, 128.48, 127.80, 127.37, 127.03, 126.56, 126.47, 122.34, 116.47, 77.84, 62.87 (CH), 57.27 (CH_2), 21.76 (CH_3), 14.22 (CH_3).

8.2.22 Ethyl 5-chloro-3-((3,5-dimethylphenyl)(methoxy)methyl)-1-tosyl-1*H*-indole-2-carboxylate **50**

The same procedure was used as per the synthesis of **48**. The following equivalents were used: Compound **44** (340 mg, 0.660 mmol), *p*-TsOH (10 eq., 1.01 g, 5.31 mmol). The reaction yielded compound **50** as a viscous milky oil, 156 mg, 0.297 mmol, 45 % (R_f = 0.21, 20% EtOAc/Hex).



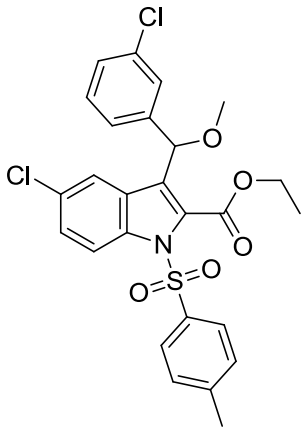
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.94 (d, *J* = 8.9 Hz, 1H, ArH), 7.79 (d, *J* = 8.3 Hz, 2H, ArH), 7.56 – 7.51 (m, 1H, ArH), 7.31 – 7.24 (m, 1H, ArH), 7.18 (d, *J* = 8.4 Hz, 2H, ArH), 7.03 (s, 2H, ArH), 6.87 (s, 1H, ArH), 5.53 (s, 1H, CH), 4.53 (q, *J* = 7.17, 2H, CH_2), 3.31 (s, 3H, CH_3), 2.33 (s, 3H, CH_3), 2.27 (s, 6H, 2 × CH_3), 1.46 (t, *J* = 7.15, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 162.42 (COO), 145.45, 139.61, 137.91, 135.12, 133.96, 130.84, 130.05, 129.72, 129.46, 129.04, 127.31, 126.88, 126.26, 124.15, 122.24, 116.28, 77.84, 62.75 (CH), 57.17 (CH_2), 21.64 (2 x CH_3), 21.46 (CH_3), 14.16 (CH_3).

8.2.23 Ethyl 5-chloro-3-(methoxy(*m*-tolyl)methyl)-1-tosyl-1*H*-indole-2-carboxylate **49**

The same procedure was used as per the synthesis of **48**. The following equivalents were used: Compound **45** (340 mg, 0.660 mmol), *p*-TsOH (10 eq., 1.01 g, 5.31 mmol). Compound **49** was used crude in the next reaction.

8.2.24 Ethyl 5-chloro-3-((3-chlorophenyl)(methoxy)methyl)-1-tosyl-1*H*-indole-2-carboxylate **51**

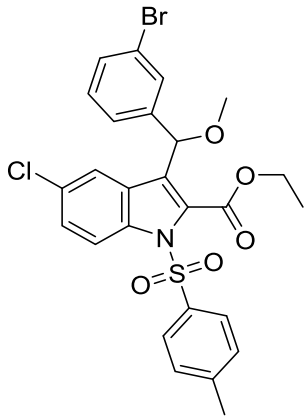
The same procedure was used as per the synthesis of **48**. The following equivalents were used: Compound 46 (790 mg, 1.52 mmol), *p*-TsOH (3 eq., 870 mg, 4.50 mmol). The reaction yielded compound **51** as a white solid, 90 mg, 0.17 mmol, 11 % (R_f = 0.35, 20% EtOAc/Hex).



Mp 120-122 °C. **^{1}H NMR (600 MHz, d_6-DMSO) δ** 8.00 (d, *J* = 8.9 Hz, 1H, ArH), 7.79 (d, *J* = 8.3 Hz, 2H, ArH), 7.52 (d, *J* = 2.0 Hz, 1H, ArH), 7.49 – 7.41 (m, 2H, ArH), 7.40 – 7.36 (m, 4H, ArH), 7.35 – 7.31 (m, 1H, ArH), 5.66 (s, 1H, CH), 4.46 (q, 7.2 Hz, 2H, CH_2) 3.24 (s, 3H, CH_3), 2.30 (s, 3H, CH_3), 1.37 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (151 MHz, d_6-DMSO) δ** 161.37 (COO), 146.18, 142.02, 133.99, 133.12, 132.63, 130.37, 130.12, 130.07, 129.15, 128.25, 127.84, 127.03, 126.82, 126.15, 125.01, 124.48, 120.83, 116.63, 76.05, 62.63 (CH), 56.81 (CH_2), 21.03 (CH_3), 13.74 (CH_3). **HRMS:** Calcd for $C_{26}H_{23}Cl_2NO_5S$ $[M - H]^-$, 530.0596, found 530.0587.

8.2.25 Ethyl 3-((3-bromophenyl)(methoxy)methyl)-5-chloro-1-tosyl-1H-indole-2-carboxylate **52**

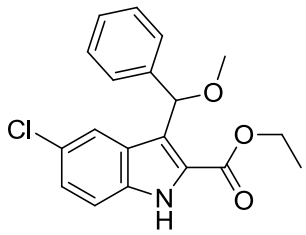
The same procedure was used as per the synthesis of **48**. The following equivalents were used: Compound **47** (195 mg, 0.35 mmol), *p*-TsOH (8 eq., 530 mg, 2.80 mmol). The reaction yielded compound **52** as a cream-colored oil, 16.8 mg, 0.03 mmol, 9% (R_f = 0.42, 20% EtOAc/Hex).



IR (ATR, cm-1): 2918 (C – H str), 1715 (C = O str), 1442 (C – H bend) 1373 (S = O str), 1307 (C – O str) 1184 (S = O str), 1094 (C-Br), 812 (Ar – H oop bend), 666 (C- Cl str) **^{1}H NMR (400 MHz, $CDCl_3$) δ** 7.86 (d, *J* = 8.8 Hz, 1H, ArH), 7.65 (d, *J* = 8.4 Hz, 2H, ArH), 7.51 (s, 1H, ArH), 7.40 (d, *J* = 1.8 Hz, 1H, ArH), 7.31 – 7.18 (m, 3H, ArH), 7.15 – 7.04 (m, 3H, ArH), 5.45 (s, 1H, CH), 4.45 (q, *J* = 7.1 Hz, 2H, CH_2), 3.23 (s, 3H, CH_3), 2.26 (s, 3H, CH_3), 1.38 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 162.31 (COO), 145.68, 142.24, 135.36, 133.76, 131.37, 130.92, 130.34, 130.05, 129.88, 129.53, 128.73, 127.29, 127.24, 125.89, 124.88, 122.70, 122.09, 116.58, 76.96, 63.04 (CH), 57.33 (CH_2), 21.79 (CH_3), 14.23 (CH_3), 9.16 (CH_3). Calculated for $[M+H]^+$, 576.0247, found 576.0054.

8.2.26 Synthesis of Ethyl 5-chloro-3-(methoxy(phenyl)methyl)-1*H*-indole-2-carboxylate 10

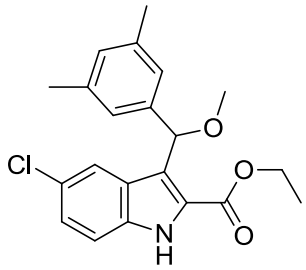
Compound **48** (85.0 mg, 0.806 mmol) was dissolved in THF (2 mL) and EtOH (5 mL) in a 100 mL RBF. KOH (8 eq., 76.0 mg, 3.20 mmol) was added and the reaction was stirred at rt for 3.5 hours. TLC showed all starting material was consumed. The reaction mixture was quenched with a sat. NH_4Cl and extracted with EtOAc (3 × 15 mL). The organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude product was purified by column chromatography (5-30%) to afford compound 10 as a white powder, 56.3 mg, 0.164 mmol, 96% (R_f = 0.30, 20% EtOAc/Hex).



^{1}H NMR (300 MHz, $CDCl_3$) δ 8.97 (s, 1H, NH), 8.07 – 8.03 (m, 1H, ArH), 7.58 – 7.51 (m, 2H, ArH), 7.37 – 7.20 (m, 5H, ArH), 6.40 (s, 1H, CH), 4.48 (q, *J* = 7.1 Hz, 2H, CH_2), 3.46 (s, 3H, CH_3), 1.46 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 161.76 (COO), 141.90, 134.48, 128.41, 127.45, 126.89, 126.78, 126.44, 126.39, 124.93, 123.61, 123.07, 112.90, 78.74 (CH), 61.50 (CH_2), 57.32 (CH_3), 14.56 (CH_3).

8.2.27 Ethyl 5-chloro-3-((3,5-dimethylphenyl)(methoxy)methyl)-1*H*-indole-2-carboxylate 13

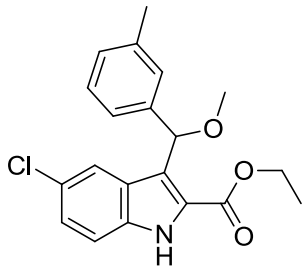
The same procedure was used as per the synthesis of 10. The following equivalents were used: Compound **50** (138 mg, 0.260 mmol), KOH (4 eq., 59.0 mg, 1.05 mmol). The reaction yielded compound **13** as a white solid, 69.5 mg, 0.187 mmol, 72 % (R_f = 0.58, 50% EtOAc/Hex).



^{1}H NMR (300 MHz, $CDCl_3$) δ 9.08 (s, 1H, NH), 8.07 (s, 1H, ArH), 7.26 (d, *J* = 1.7 Hz, 2H, ArH), 7.15 (s, 2H, ArH), 6.89 (s, 1H, ArH), 6.33 (s, 1H, CH), 4.53 – 4.45 (m, 2H, CH_2), 3.45 (s, 3H, CH_3), 2.30 (s, 6H, 2× CH_3), 1.47 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 161.90 (COO), 141.73, 137.84, 134.54, 129.16, 126.95, 126.28, 124.93, 124.53, 123.59, 123.14, 112.88, 78.84 (CH), 61.47 (CH_2), 57.29 (CH_3), 21.53 (2 x CH_3), 14.55 (CH_3).

8.2.28 Ethyl 5-chloro-3-(methoxy(*m*-tolyl)methyl)-1*H*-indole-2-carboxylate 19

The same procedure was used as per the synthesis of 10. The following equivalents were used: Compound **49** (50.0 mg, 0.10 mmol), KOH (4 eq., 22.0 mg, 0.400 mmol). The reaction yielded compound **19** as an off white solid, 36 mg, 0.10 mmol (R_f = 0.60, 50% EtOAc/Hex).

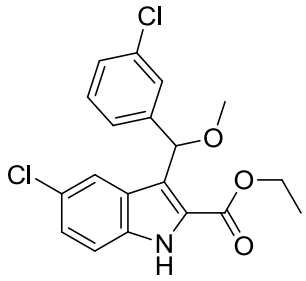


Mp 122 – 126°C **IR (ATR, cm-1):** 3289 (N – H str), 2999 – 2787 (C – H str), 1697 (C = O str), 1289 (C – O str), 667 (C- Cl str) **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.95 (s, 1H, NH), 8.08 – 8.01 (m, 1H, ArH), 7.40 – 7.14 (m, 5H, ArH), 7.10 – 6.99 (m, 1H, ArH), 6.35 (s, 1H, CH), 4.47 (q, *J* =7.1 Hz, 2H, CH_2), 3.44 (s, 3H, CH_3), 2.33 (s, 3H, CH_3), 1.46 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 161.79 (COO), 141.80, 137.97, 134.48, 128.31, 128.23, 127.48, 126.93, 126.41, 126.36, 124.94, 123.80, 123.63, 123.13, 112.87, 78.77 (CH), 61.48 (CH_2), 57.32 (CH_3), 21.68 (CH_3), 14.57 (CH_3). **HRMS:** Calcd for $C_{20}H_{20}ClNO_3$ $[M + H]^+$, 358.1132, found 358.1035.

8.2.29 Ethyl 5-chloro-3-((3-chlorophenyl)(methoxy)methyl)-1*H*-indole-2-carboxylate 20

The same procedure was used as per the synthesis of 10. The following equivalents were used: Compound **51** (75.0 mg, 0.140 mmol), KOH (4 eq., 31.0 mg, 0.560 mmol). The reaction yielded

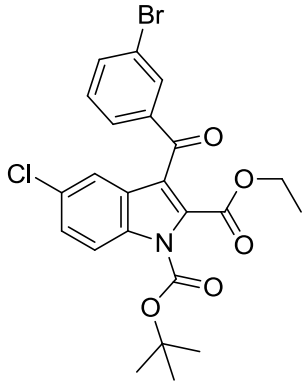
compound **20** as a cream colored solid, 53 mg, 0.14 mmol, 15% over two steps (R_f = 0.41, 50% EtOAc/Hex).



Mp 124-126 °C**. IR (ATR, cm-1):** 3288 (N – H str), 1693 (C = O str), 1633 (N – H bend), 1259 (C – O str), 762 (C- Cl str) **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.97 (s, 1H, NH), 7.89 (d, *J* = 1.3 Hz, 1H, ArH), 7.47 – 7.44 (m, 1H, ArH), 7.37 – 7.32 (m, 1H, ArH), 7.24-7.20 (m, 2H, ArH), 7.19-7.13 (m, 2H, ArH), 6.29 (s, 1H, CH), 4.45 (q, *J* = 7.2 Hz, 2H, CH_2), 3.37 (s, 3H, CH_3), 1.40 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 161.67 (COO), 143.99, 134.48, 134.22, 129.70, 127.54, 126.92, 126.69, 126.62, 126.55, 125.18, 124.72, 122.75, 122.70, 113.02, 77.88 (CH), 61.64 (CH_2), 57.34 (CH_3), 14.55 (CH_3). **HRMS:** Calcd for $C_{19}H_{19}Cl_2NO_3$ $[M - H]^-$, 376.0507, found 376.0503.

8.2.30 1-*Tert*-butyl 2-ethyl 3-(3-bromobenzoyl)-5-chloro-1*H*-indole-1,2-dicarboxylate 53

Compound **35** (1.00 g, 2.46 mmol) was dissolved in THF in a 100 mL RBF. Boc_2O (1.2 eq., 0.610 mL, 2.95 mmol) and DMAP (cat.) were added at rt. The reaction was complete after 0.5h after which it was concentrated *in vacuo* and purified by column chromatography to yield compound **53** as a viscous yellow oil, 1.18 g, 2.32 mmol, 94% (R_f = 0.43, 20% EtOAc/Hex).

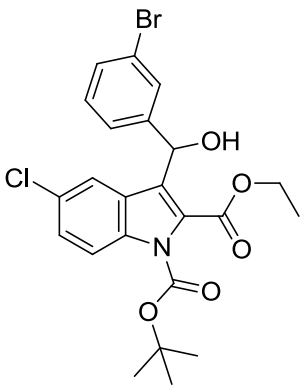


Mp 53 – 57 °C **IR (ATR, cm-1):** 2980 (C – H str), 1740 (C=O str), 1569 (C=C str), 1474 (C – H str), 1203 (C – O str), 844 (Ar C – H bend), 760 (C – Cl str) **^{1}H NMR (400 MHz, $CDCl_3$) δ** 8.09 (d, *J* = 9.0 Hz, 1H, ArH), 7.95 – 7.92 (m, 1H, ArH), 7.74 – 7.64 (m, 3H, ArH), 7.42 – 7.31 (m, 2H, ArH), 3.96 (q, *J* = 7.2 Hz, 2H, CH_2), 1.63 (s, 9H, 3 × CH_3), 1.13 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 189.60 (COO), 161.29, 148.36, 140.44, 136.15, 134.18, 132.09, 130.51, 130.33, 127.93, 127.88, 127.51, 122.99, 121.51, 120.31, 116.58, 86.96, 62.67 (CH_2), 27.99 (3 x CH_3), 13.72 (CH_3). **HRMS:** Calcd for $C_{23}H_{21}BrClNO_5$ $[M + H]^+$, 506.0370, found 506.0365.

8.2.31 1-*Tert*-butyl 2-ethyl 3-((3-bromophenyl)(hydroxy)methyl)-5-chloro-1*H*-indole-1,2-dicarboxylate 54

The same procedure was used as per the synthesis of Error! Reference source not found.. The following equivalents were used: Compound **53** (983 mg, 1.97 mmol), sodium borohydride (3 eq.,

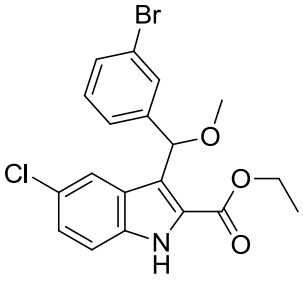
224 mg, 5.91 mmol). The reaction yielded compound **54** as an off white solid, 703 mg, 1.38 mmol, 70 % (R_f = 0.30, 20% EtOAc/Hex).



Mp: 50 – 54°C. **IR (ATR, cm^{-1}):** 3457 (O – H str), 2935 (C – H str.), 1736 (C=O str) 1593 (C=C str), 1474 (C – H str), 1447 (C – H bend), 1369 (C – H bend), 1317 (C – N str), 1259 (C – O str), 1113 (C – Br str), 842 (Ar C – H oop bend), 807 (Ar C-H bend), 765 (C-Cl str) **^{1}H NMR (600 MHz, $CDCl_3$) δ** 7.96 (d, *J* = 9.0 Hz, 1H, ArH), 7.65 – 7.62 (s, 1H, ArH), 7.56 (d, *J* = 2.0 Hz, 1H, ArH), 7.41 – 7.35 (m, 2H, ArH), 7.31 (dd, *J* = 8.9, 2.1 Hz, 1H, ArH), 7.21 – 7.14 (m, 1H, ArH), 6.14 (d, *J* = 6.10 Hz, 1H, CH), 4.37 – 4.30 (m, 2H, CH_2), 3.24 (d, *J* = 6.10 Hz, 1H, OH) 1.61 (s, 9H), 1.29 (t, *J* = 7.2 Hz, 3H, H_{10}). **^{13}C NMR (151 MHz, $CDCl_3$) δ** 165.43 (COO), 151.37, 146.71, 137.23, 133.25, 132.57, 131.69, 131.62, 131.16, 130.17, 129.66, 128.60, 127.23, 125.21, 123.54, 118.90, 88.19, 70.28 (CH), 64.92 (CH_2), 30.56 (3 x CH_3), 16.60 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{23}H_{22}BrClNO_4$ $[M - OH]^+$, 490.0421, found 490.0413.

8.2.32 Ethyl 3-((3-bromophenyl)(methoxy)methyl)-5-chloro-1*H*-indole-2-carboxylate **21**

The same procedure was used as per the synthesis of **48**. The following equivalents were used: Compound **54** (357 mg, 0.700 mmol), *p*-TsOH (5 eq., 666 mg, 3.51 mmol). The reaction yielded compound **21** as a cream solid, 188 mg, 0.440 mmol, 63% (R_f = 0.30, 20% EtOAc/Hex).



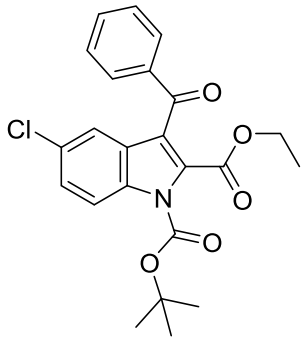
Mp: 125 – 128°C. **IR (ATR, cm^{-1}):** 3311 (N – H str), 1697 (C=O str) 1449, 1380 (C – H bend), 1330 (C – N str), 1250 (C – O str), 1091 (C – Br str), 801 (C – Br str), 708 (Ar C – H bend), 768 (C – Cl str) **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.88 (s, 1H, NH), 7.95 – 7.90 (m, 1H, ArH), 7.70-7.66 (m, 1H, ArH), 7.47 – 7.25 (m, 4H, ArH), 7.22-7.15 (m, 1H, ArH), 6.33 (s, 1H, CH), 4.47 (qd, *J* = 7.1, 1.1 Hz, 2H, CH_2), 3.41 (s, 3H, CH_3), 1.44 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 161.72 (COO), 144.23, 134.51, 130.44, 129.99, 129.81, 126.64, 126.58, 126.50, 125.18, 125.16, 122.70, 122.57, 122.46, 113.04, 77.83 (CH), 61.65 (CH_2), 57.32 (CH_3), 14.53 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd $C_{18}H_{14}BrClNO_2$ $[M - OCH_3]^+$, 389.9896, found 398.9895.

8.2.33 Attempted synthesis of ethyl 5-chloro-3-((3-cyanophenyl)(methoxy)methyl)-1H-indole-2-carboxylate 22

Compound 21 (50.0 mg, 0.120 mmol) was added to a 2-necked RBF fitted with a condenser along with 3 mL of NMP. The compound did not dissolve in this solvent, however CuCN (2 eq., 21.5 mg, 0.240 mmol) was added and the reaction was heated to 163 °C. After 3h no reaction had occurred and the starting material was still not dissolved. A further 1 mL of NMP was added along with 2 mL of DMF. This helped solubilize the starting material and we saw this clear solution undergo several color changes from clear to red to yellow and finally orange. After 3 days, there was still starting material present but a new spot was visible on TLC, thus we continued to work-up the reaction. The reaction mixture was cooled, poured over ice water (30 mL) and aqueous NH_3 (10 mL) was added to this. The product was extracted with DCM (3 x 15 mL). The organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude material was purified by column chromatography (10 – 60% EtOAc/Hex) to obtain 29 mg of a brown solid. ^{1}H and ^{13}C analysis indicated this was not the desired product and was most likely a decomposition product from the prolonged refluxing.

8.2.34 1-*tert*-butyl 2-ethyl 3-benzoyl-5-chloro-1H-indole-1,2-dicarboxylate 56

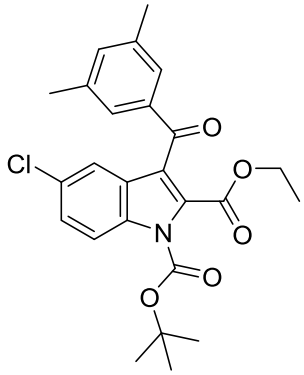
The same procedure was used as per the synthesis of **53**. The following equivalents were used: Compound **29** (648 mg, 1.98 mmol), Boc_2O (1.2 eq., 0.540 ml, 2.40 mmol) and DMAP (3-4 crystals). The reaction yielded compound 56 as a yellow oil, 730 mg, 1.70 mmol, 88% (R_f = 0.68, 30% EtOAc/Hex).



^{1}H NMR (300 MHz, $CDCl_3$) δ 8.11 (dd, *J* = 9.0, 0.5 Hz, 1H, ArH), 7.82-7.78 (m, 2H, ArH), 7.64 – 7.56 (m, 2H, ArH), 7.51-7.44 (m, 2H, ArH), 7.39 (dd, *J* = 9.0, 2.1 Hz, 1H, ArH), 3.93 (q, *J* = 7.2 Hz, 2H, CH_2), 1.63 (s, 9H, 3 × CH_3), 1.08 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 191.01 (CO), 161.38 (COO), 148.41, 138.53, 134.24, 133.40, 130.22, 129.32, 128.69, 127.99, 127.31, 121.52, 121.18, 116.44, 86.64, 62.38 (CH_2), 27.91 (3 x CH_3), 13.61 (CH_3).

8.2.35 1-*tert*-butyl 2-ethyl 5-chloro-3-(3,5-dimethylbenzoyl)-1H-indole-1,2-dicarboxylate 57

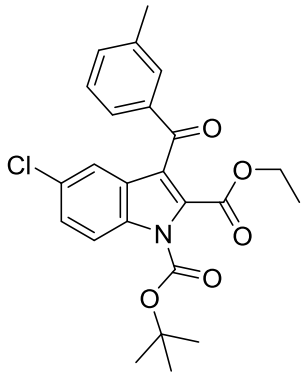
The same procedure was used as per the synthesis of **53**. The following equivalents were used: Compound **32** (800 mg, 2.25 mmol), Boc_2O (1.2 eq., 0.620 ml, 2.70 mmol), DMAP (cat.). The reaction yielded compound 57 as a yellow solid, 756 mg, 1.66 mmol, 75% (R_f = 0.49, 20% EtOAc/Hex).



^{1}H NMR (400 MHz, $CDCl_3$) δ 8.10 (d, *J* = 9.0 Hz, 1H, ArH), 7.66 (d, *J* = 2.0 Hz, 1H, ArH), 7.42 – 7.36 (m, 3H, ArH), 7.22 (s, 1H, ArH), 3.90 (q, *J* = 7.2 Hz, 2H, CH_2), 2.34 (s, 6H, 2 × CH_3), 1.62 (s, 9H, 3 × CH_3), 1.06 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 191.33 (CO), 161.43 (COO), 148.46, 138.57, 138.41, 134.96, 134.25, 130.16, 128.12, 127.23, 127.08, 121.58, 121.49, 116.38, 86.52, 62.32 (CH_2), 27.90 (3 x CH_3), 21.25 (2 x CH_3), 13.54 (CH_3).

8.2.36 1-*tert*-butyl 2-ethyl 5-chloro-3-(3-methylbenzoyl)-1H-indole-1,2-dicarboxylate 58

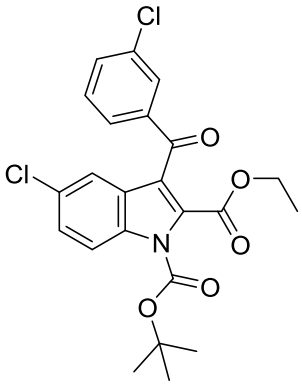
The same procedure was used as per the synthesis of **53**. The following equivalents were used: Compound **33** (913 mg, 2.67 mmol), boc_2O (1.2 eq., 0.74 ml, 3.21 mmol), DMAP (3-4 crystals). The reaction yielded compound 58 as a yellow solid, 987 mg, 2.23 mmol, 84% (R_f = 0.39, 20% EtOAc/Hex).



Mp 98 – 100°C. **IR (ATR, cm^{-1}):** 2982 (C – H str), 1652 (C=O str), 1585 (C=C str), 1441 (C – H str), 1305 (C - N str), 1216 (C – O str), 808 (Ar C – H bend), 762 (C – Cl str) **^{1}H NMR (400 MHz, $CDCl_3$) δ** 8.11 (d, *J* = 9.0 Hz, 1H, ArH), 7.66 – 7.30 (m, 6H, ArH), 3.92 (q, *J* = 7.2 Hz, 2H, CH_2), 2.40 (s, 3H, CH_3), 1.63 (s, 9H, 3 × CH_3), 1.07 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 190.86 (CO), 161.09 (COO), 148.12, 138.26, 138.22, 133.93, 133.81, 133.43, 129.87, 129.45, 128.25, 127.74, 126.95, 126.27, 121.24, 116.09, 86.26, 62.03 (CH_2), 27.59 (3 x CH_3), 21.06 (CH_3), 13.26 (CH_3). **HRMS:** Calcd for $C_{24}H_{25}ClNO_5$ $[M + H]^+$, 442.1421, found 442.1430.

8.2.37 1-*tert*-butyl 2-ethyl 5-chloro-3-(3-chlorobenzoyl)-1H-indole-1,2-dicarboxylate 59

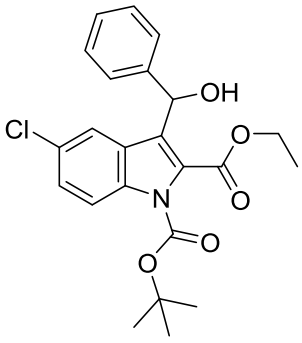
The same procedure was used as per the synthesis of **53**. The following equivalents were used: compound **34** (750 mg, 2.10 mmol), Boc_2O (1.2 eq., 0.570 ml, 2.50 mmol), DMAP (cat.). The reaction yielded compound 59 as a yellow oil, 691 mg, 1.49 mmol, 71% (R_f = 0.75, 30% EtOAc/Hex).



IR (ATR, cm^{-1}): 2970 (C – H str), 1734 (C=O str), 1317 (C – N str), 1259 (C=C str), 829 (Ar C – H bend), 707 (C – Cl str), 647 (C – Cl str) **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.10 (dd, *J* = 9.0, 0.5 Hz, 1H, ArH), 7.81 – 7.77 (m, 1H, ArH), 7.68 – 7.63 (m, 2H, ArH), 7.57 (ddd, *J* = 8.0, 2.1, 1.1 Hz, 1H, ArH), 7.45 – 7.36 (m, 2H, ArH), 3.97 (q, *J* = 7.2 Hz, 2H, CH_2), 1.63 (s, 9H, 3 × CH_3), 1.13 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 189.64 (CO), 161.22 COO), 148.29, 140.15, 134.99, 134.24, 134.11, 133.17, 130.43, 130.01, 129.12, 127.79, 127.44, 127.42, 121.43, 120.32, 116.50, 86.87, 62.56 (CH_2), 27.91 (3 x CH_3), 13.62 (CH_3). **HRMS:** Calcd for $C_{23}H_{21}Cl_2NO_5$ $[M + H]^+$, 462.0875, found 462.0859.

8.2.38 1-*tert*-butyl 2-ethyl 5-chloro-3-(hydroxy(*m*-tolyl)methyl)-1H-indole-1,2-dicarboxylate 60

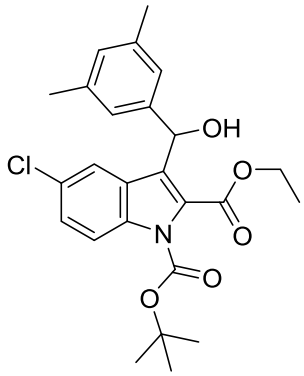
The same procedure was used as per the synthesis of **12.** The following equivalents were used: compound 56 (710 mg, 1.65 mmol), $NaBH_4$ (3 eq., 187 mg, 4.95 mmol). The reaction yielded compound 60 as a yellow oil, 664 mg, 1.54 mmol, 94% (R_f = 0.42, 20% EtOAc/Hex).



^{1}H NMR (400 MHz, $CDCl_3$) δ 7.98 (d, *J* = 8.6 Hz, 1H, ArH), 7.57 (d, *J* = 1.9 Hz, 1H, ArH) 7.48 (dd, *J* = 4.8, 4.3 Hz, 2H, ArH), 7.36 – 7.22 (m, 4H, ArH), 6.20 (d, *J* = 4.5 Hz, 1H, CH), 4.40 – 4.25 (m, 2H, CH_2), 1.63 (s, 9H, 3 × CH_3), 1.31 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 163.03 (COO), 149.00, 141.89, 134.81, 129.02, 128.59, 128.56, 127.99, 127.77, 127.02, 126.78, 126.17, 121.28, 116.36, 85.53, 68.72 (CH), 62.27 (CH_2), 28.09 (3 x CH_3), 14.14 (CH_3).

8.2.39 1-*tert*-butyl 2-ethyl 5-chloro-3-((3,5-dimethylphenyl)(hydroxy)methyl)-1H-indole-1,2-dicarboxylate 61

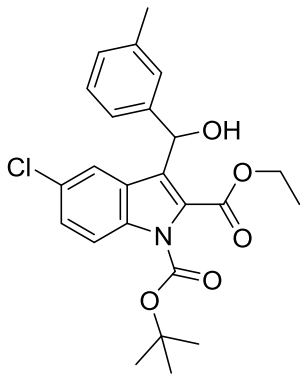
The same procedure was used as per the synthesis of **12.** The following equivalents were used: Compound 57 (716 mg, 1.59 mmol), $NaBH_4$ (3 eq., 181 mg, 4.78 mmol). The reaction yielded compound 61 as a yellow oil, 673 mg, 1.47 mmol, 94% (R_f = 0.32, 20% EtOAc/Hex).



^{1}H NMR (400 MHz, $CDCl_3$) δ 7.99 (dd, *J* = 5.8, 3.1 Hz, 1H, ArH), 7.58 (d, *J* = 1.7 Hz, 1H, ArH), 7.31 (dd, *J* = 8.9, 2.1 Hz, 1H, ArH), 7.07 (s, 2H, ArH), 6.89 (s, 1H, ArH), 6.12 (d, *J* = 5.4 Hz, 1H, CH), 4.37 – 4.30 (m, 2H, CH_2), 2.92 (d, *J* = 5.4 Hz, 1H, OH), 2.28 (s, 6H, 2 × CH_3), 1.64 (s, 9H, 3 × CH_3), 1.34 – 1.30 (m, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 163.11 (COO), 149.03, 141.79, 138.13, 134.74, 129.47, 128.94, 128.47, 126.90, 126.70, 123.96, 121.33, 116.32, 85.46, 68.86 (CH), 62.20 (CH_2), 28.09 (3 x CH_3), 21.52 (2 x CH_3), 14.11 (CH_3).

8.2.40 1-*tert*-butyl 2-ethyl 5-chloro-3-(hydroxy(*m*-tolyl)methyl)-1H-indole-1,2-dicarboxylate 62

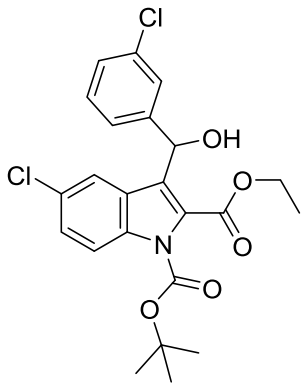
The same procedure was used as per the synthesis of **12.** The following equivalents were used: Compound 58 (875 mg, 1.98 mmol), $NaBH_4$ (3 eq., 228 mg, 6.05 mmol). The reaction yielded compound 62 as a yellow oil, 843 mg, 1.90 mmol, 96% (R_f = 0.23, 5% EtOAc/Hex).



IR (ATR cm^{-1}): 3452 (O – H str), 2980 – 2935 (C – H str), 1736 (C=O str) 1606 (C=C str), 1447 (C – H bend), 1350 (C – N str), 1318 (C – O str), 809 (Ar C – H bend), 765 (C-Cl str) **^{1}H NMR (400 MHz, $CDCl_3$) δ** 7.98 (d, *J* = 8.9 Hz, 1H, ArH), 7.57 (d, *J* = 2.0 Hz, 1H, ArH), 7.33 – 7.04 (m, 4H, ArH), 6.15 (d, *J* = 5.1 Hz, 1H, CH), 4.33 (q, *J* = 7.1 Hz, 2H, CH_2), 2.97 (d, *J* = 5.3 Hz, 1H, OH), 2.32 (s, 3H, CH_3), 1.63 (s, 9H, 3 × CH_3), 1.31 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 163.07 (COO), 149.00, 141.82, 138.23, 134.76, 128.96, 128.54, 128.48, 128.03, 126.94, 126.85, 126.72, 123.23, 121.31, 116.33, 85.49, 68.77 (CH), 62.23 (CH_2), 28.08 (3 x CH_3), 21.65 (CH_3), 14.12 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{24}H_{26}ClNO_5$ $[M - OH]^+$, 426.1472, found 426.1471.

8.2.41 1-*tert*-butyl 2-ethyl 5-chloro-3-((3-chlorophenyl)(hydroxy)methyl)-1H-indole-1,2-dicarboxylate 63

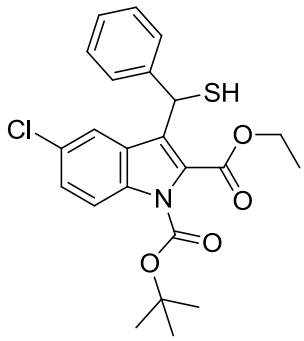
The same procedure was used as per the synthesis of **12.** The following equivalents were used: compound 59 (691 mg, 1.49 mmol), sodium borohydride (3 eq., 173 mg, 4.56 mmol). The reaction yielded compound 63 as a yellow oil, 529 mg, 1.14 mmol, 76% (R_f = 0.26, 20%).



^{1}H NMR (300 MHz, DMSO) δ 7.98 (d, *J* = 8.8 Hz, 1H, ArH), 7.72 (d, *J* = 2.9 Hz, 1H, ArH), 7.56-7.53 (m, 1H, ArH), 7.49 – 7.21 (m, 4H, ArH), 6.43 (d, *J* = 2.8 Hz, 1H, CH), 6.07 (s, 1H, OH), 4.44 – 4.30 (m, 2H, CH_2), 1.58 (s, 9H, 3 × CH_3), 1.32 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, DMSO) δ** 161.58 (COO), 148.27, 145.74, 134.10, 132.97, 130.24, 127.47, 127.27, 127.06, 126.58, 126.09, 125.62, 124.44, 121.42, 116.29, 85.67, 66.30 (CH), 61.89 (CH_2), 27.35 (3 x CH_3), 13.86 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{23}H_{22}Cl_2NO_4$ $[M - OH]^+$, 446.0926, found 446.0923.

8.2.42 1-*Tert*-butyl 2-ethyl 5-chloro-3-(mercapto(phenyl)methyl)-1*H*-indole-1,2-dicarboxylate 64

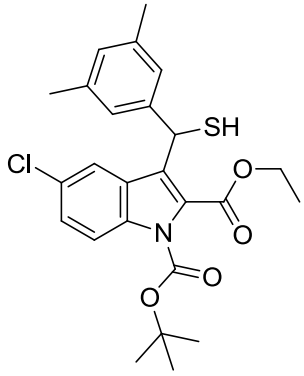
Compound 60 (624 mg, 1.48 mmol) was dissolved in dry toluene (20 mL) in a three necked RBF fitted with a condenser. Lawesson's reagent (0.6 eq., 360 mg, 0.890 mmol) was added and the reaction mixture was heated to 100 °C for 30 mins. The reaction mixture was cooled and concentrated *in vacuo* before the crude product was purified by column chromatography (2 – 20% EtOAc/Hex). The reaction yielded compound **64** as a yellow oil (130 mg, 0.280 mmol, 49% (R_f = 0.41, 5% EtOAc/Hex).



IR (ATR cm^{-1}): 2981 (C – H str), 1727 (C=O str), 1351 (C – N str), 1152 (C – O str), 810 (ArH oop bend), 696 (C – Cl str) **^{1}H NMR (400 MHz, $CDCl_3$) δ** 8.02 (d, *J* = 9.0 Hz, 1H, ArH), 7.54 (d, *J* = 2.0 Hz, 1H, ArH), 7.52 (d, *J* = 7.4 Hz, 2H, ArH), 7.36 – 7.29 (m, 3H, ArH), 7.28 – 7.22 (m, 1H, ArH), 5.89 (d, *J* = 5.8 Hz, 1H, CH), 4.38 (q, *J* = 7.2 Hz, 2H, CH_2), 2.41 (d, *J* = 5.8 Hz, 1H, SH), 1.63 (s, 9H, 3 × CH_3), 1.40 – 1.34 (m, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 162.21 (COO), 148.78, 139.91, 134.90, 128.57, 128.50, 127.91, 127.55, 127.52, 127.37, 126.76, 125.58, 121.60, 116.40, 85.41, 62.03 (CH_2), 37.78 (CH), 27.94 (3 x CH_3), 14.09 (CH_3). **HRMS:** Calcd for $C_{23}H_{25}ClNO_4S[M + H]^+$, 446.1193, found 446.1179.

8.2.43 1-*Tert*-butyl 2-ethyl 5-chloro-3-((3,5-dimethylphenyl)(mercapto)methyl)-1*H*-indole-1,2-dicarboxylate 65

The same procedure was used as per the synthesis of **64**. The following equivalents were used: Compound 61 (650 mg, 1.44 mmol), Lawesson's reagent (0.6 eq., 350 mg, 0.86 mmol). The reaction yielded compound **65** as a yellow oil, 335 mg, 0.71 mmol, 49% (R_f = 0.48, 20% EtOAc/Hex).

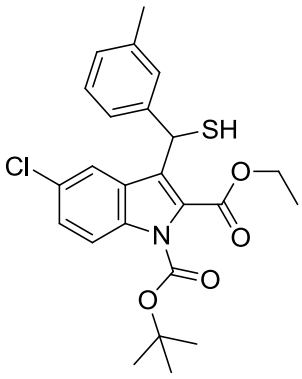


IR (ATR cm^{-1}):^{1}H NMR (400 MHz, $CDCl_3$) δ 8.02 (d, *J* = 8.9, 1H, ArH), 7.59 (d, *J* = 1.9 Hz, 1H, ArH), 7.31 (dd, *J* = 8.9, 2.10 Hz, 1H, ArH), 7.12 (s, 2H, ArH), 6.88 (s, 1H, ArH), 5.82 (d, *J* = 5.8 Hz, 1H, CH), 4.39 (q, *J* = 7.3 Hz, 2H, CH_2), 2.38 (d, *J* = 5.8 Hz, 1H, SH), 2.29 (s, 6H, 2 × CH_3), 1.64 (s, 9H, 3 × CH_3), 1.38 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 162.38 (COO), 148.95, 139.82, 138.16, 135.00, 129.18, 128.63, 127.92, 127.78, 126.81, 125.86, 125.39, 121.82, 116.48, 85.47, 62.11 (CH_2), 37.83 (CH), 28.08 (3 x CH_3), 21.50 (2 x CH_3), 14.18 (CH_3). **HRMS:** Calcd for $C_{25}H_{28}ClNO_4S$ $[M + H]^+$, 474.1506, found 474.1522.

8.2.44 1-*Tert*-butyl 2-ethyl 5-chloro-3-(mercapto(*m*-tolyl)methyl)-1*H*-indole-1,2-dicarboxylate 66

The same procedure was used as per the synthesis of **64**. The following equivalents were used: Compound 62 (510 mg, 1.17 mmol), Lawesson's reagent (0.5 eq. 237 mg, 0.590 mmol). The

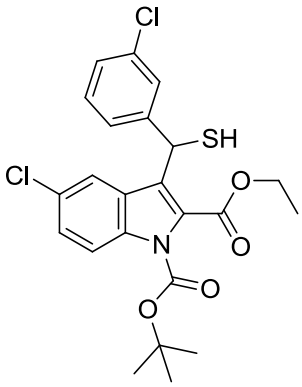
reaction yielded compound **66** as a yellow oil, 129.7 mg, 0.28 mmol, 49% (R_f = 0.41, 5% EtOAc/Hex).



^{1}H NMR (400 MHz, $CDCl_3$) δ 8.05 (d, *J* = 9.0 Hz, 1H, ArH), 7.61 (d, *J* = 2.0 Hz, 1H, ArH), 7.35 – 7.30 (m, 3H, ArH), 7.23 – 7.19 (m, 1H, ArH), 7.06 (d, *J* = 8.9 Hz, 1H, ArH), 5.89 (d, *J* = 5.7 Hz, 1H, CH), 4.41 (q, *J* = 7.2 Hz, 2H, CH_2), 2.43 (d, *J* = 5.7 Hz, 1H, SH), 2.36 (s, 3H, CH_3), 1.67 (s, 9H, 3 × CH_3), 1.41 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, CDCl3) δ** 162.31 (COO), 148.88, 139.88, 138.25, 134.97, 128.62, 128.46, 128.30, 128.24, 127.94, 127.69, 126.80, 125.75, 124.61, 121.73, 116.47, 85.46, 62.09 (CH_2), 37.82 (CH), 28.02 (3 x CH_3), 21.58 (CH_3), 14.18 (CH_3). **HRMS:** Calcd for $C_{24}H_{26}ClNO_4S$ $[M + H]^+$, 460.1349, found 460.1349.

8.2.45 1-*Tert*-butyl 2-ethyl 5-chloro-3-((3-chlorophenyl)(mercapto)methyl)-1*H*-indole-1,2-dicarboxylate 67

The same procedure was used as per the synthesis of **64**. The following equivalents were used: Compound 63 (480 mg, 1.03 mmol), Lawesson's reagent (0.6 eq., 255 mg, 0.630 mmol). The reaction yielded compound **67** as a yellow oil, 134 mg, 0.279 mmol, 27% (R_f = 0.45, 20% EtOAc/Hex).



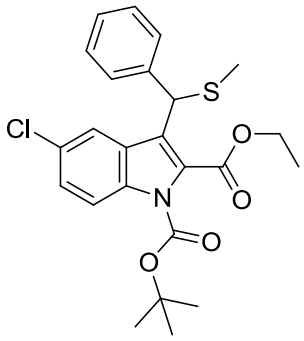
^{1}H NMR (300 MHz, $CDCl_3$) δ 8.04 (d, *J* = 9.0 Hz, 1H, ArH), 7.58 – 7.51 (m, 2H, ArH), 7.45 – 7.39 (m, 1H, ArH), 7.33 (dd, *J* = 9.0, 2.1 Hz, 1H, ArH), 7.28 – 7.19 (m, 2H, ArH), 5.84 (d, *J* = 6.1 Hz, 1H, CH), 4.39 (q, *J* = 7.2 Hz, 2H, CH_2), 2.48 (d, *J* = 6.1 Hz, 1H, SH), 1.65 (s, 9H, 3 × CH_3), 1.38 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 162.12 (COO), 148.73, 142.21, 134.87, 134.49, 129.78, 128.77, 128.18, 127.88, 127.66, 127.35, 126.96, 125.81, 124.81, 121.25, 116.57, 85.61, 62.18 (CH_2), 37.35 (CH), 27.97 (3 x CH_3), 14.05 (CH_3). **HRMS:** Calcd for $C_{23}H_{22}Cl_2NO_4$ $[M - SH]^+$, 446.0926, found 446.0929.

8.2.46 Attempted synthesis of 1-(*tert*-butyl) 2-ethyl 3-((3-bromophenyl)(mercapto)methyl)-5-chloro-1H-indole-1,2-dicarboxylate 68

The same procedure was used as per the synthesis of **64**. The following equivalents were used: Compound **54** (200 mg, 0.38 mmol), Lawesson's reagent (0.6 eq., 93 mg, 0.23 mmol). Multiple spots were observed on TLC and attempts to isolate compound **68** were not successful.

8.2.47 1-*Tert*-butyl 2-ethyl 5-chloro-3-((methylthio)(phenyl)methyl)-1*H*-indole-1,2-dicarboxylate 70

Compound **64** (300 mg, 0.690 mmol) was dissolved in DCM (20 mL) in a 100 mL RBF. NEt_3 (2.5 eq., 0.240 mL, 1.70 mmol) was added and the reaction mixture was stirred for 15 mins. MeI (2.5 eq., 0.107 mL, 1.72 mmol) was added and the reaction mixture was heated to 30 °C for 18 hours (since the R_f of the starting material and product are identical the long reaction time ensure complete conversion to the product). The reaction mixture was quenched with aqueous sat. NH_4Cl (50 mL) and extracted with EtOAc (3 × 30 mL). The organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude product was purified by column chromatography (10% EtOAc/Hex). The reaction yielded compound **70** as a clear oil, 232 mg, 0.504 mmol, 745% (R_f = 0.47, 20% EtOAc/Hex).

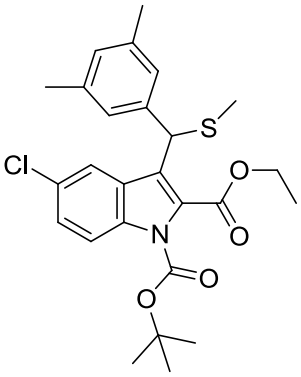


^{1}H NMR (300 MHz, d_6-DMSO) δ 8.02 (d, *J* = 9.0 Hz, 1H, ArH), 7.78 – 7.45 (m, 5H, ArH), 7.39-7.23 (m, 3H, ArH), 5.52 (s, 1H, CH), 4.36 (q, *J* = 7.1 Hz, 2H, CH_2), 1.99 (s, 3H, CH_3), 1.55 (s, 9H, 3 × CH_3), 1.28 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 162.38 (COO), 148.78, 138.84, 134.86, 129.54, 128.54, 128.48, 127.91, 127.35, 126.70, 122.87, 122.07, 116.29, 85.36, 61.98 (CH_2), 46.26 (CH), 27.96 (3 x CH_3), 15.84 (CH_3), 14.16 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{23}H_{23}ClNO_4$ $[M - SCH_3]^+$, 412.1316, found 412.1321.

8.2.48 1-*Tert*-butyl 2-ethyl 5-chloro-3-((3,5-dimethylphenyl)(methylthio)methyl)-1*H*-indole-1,2-dicarboxylate 71

The same procedure was used as per the synthesis of **70**. The following equivalents were used: Compound **65** (283 mg, 0.610 mmol), NEt_3 (2.5 eq., 0.100 mL, 1.50 mmol), MeI (2.5 eq., 0.210 mL,

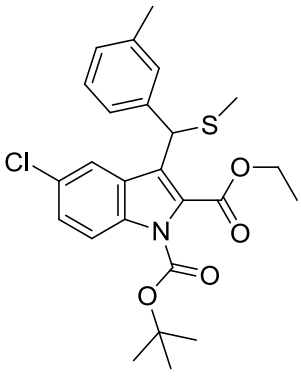
1.50 mmol). The reaction yielded compound **71** as a clear oil, 189 mg, 0.387 mmol, 64% (R_f = 0.48, 20% EtOAc/Hex).



^{1}H NMR (300 MHz, $CDCl_3$) δ 8.05 (d, *J* = 9.2, 1H, ArH), 7.85 (d, *J* = 2.1 Hz, 1H, ArH), 7.33 (dd, *J* = 9.0, 2.1 Hz, 1H, ArH), 7.16 (d, *J* = 0.6 Hz, 2H, ArH), 6.88 (s, 1H, ArH), 5.48 (s, 1H, CH), 4.44 (q, *J* = 7.1 Hz, 2H, CH_2), 2.30 (s, 6H, 2 × CH_2), 2.06 (s, 3H, CH_3), 1.67 (s, 9H, 3 × CH_3), 1.41 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 162.44 (COO), 148.83, 138.71, 138.01, 134.88, 129.48, 129.10, 128.45, 128.09, 126.66, 125.70, 123.06, 122.17, 116.30, 85.31, 61.95 (CH_2), 46.22 (CH), 27.98 (3 x CH_3), 21.41 (2 x CH_3), 15.86 (CH_3), 14.18 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{25}H_{27}ClNO_4$ [M - SCH_3]$^+$, 440.1629, found 440.1624.

8.2.49 1-*Tert*-butyl 2-ethyl 5-chloro-3-((methylthio)(*m*-tolyl)methyl)-1*H*-indole-1,2-dicarboxylate **72**

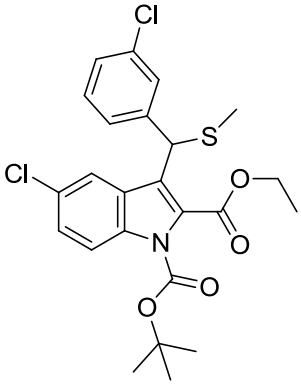
The same procedure was used as per the synthesis of **70**. The following equivalents were used: Compound **66** (520 mg, 1.15 mmol), NEt_3 (2.5 eq., 0.180 mL, 2.90 mmol), MeI (2.5 eq., 0.400 mL, 2.90 mmol). The reaction yielded compound **72** as a clear oil, 445 mg, 0.939 mmol, 83% (R_f = 0.41, 5% EtOAc/Hex).



^{1}H NMR (600 MHz, $CDCl_3$) δ 8.05 (d, *J* = 9.0 Hz, 1H, ArH), 7.83 (d, *J* = 2.1 Hz, 1H, ArH), 7.37 – 7.31 (m, 3H, ArH), 7.27 – 7.20 (m, 1H, ArH), 7.06 (d, *J* = 7.5 Hz, 1H, ArH), 5.51 (s, 1H, CH), 4.44 (q, *J* = 7.2 Hz, 2H, CH_2), 2.34 (s, 3H, CH_3), 2.06 (s, 3H, CH_3), 1.66 (s, 9H, 3 × CH_3), 1.40 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (151 MHz, $CDCl_3$) δ** 165.02 (COO), 151.42, 141.39, 140.77, 137.50, 132.13, 131.28, 131.08, 131.03, 130.78, 130.65, 129.29, 127.53, 125.60, 124.75, 118.92, 87.95, 64.58 (CH_2), 48.86 (CH), 30.60 (3 x CH_3), 24.14 (CH_3), 18.46 (CH_3), 16.79 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{24}H_{25}ClNO_4$ [M - SCH_3]$^+$, 426.1472, found 426.1478.

8.2.50 1-*Tert*-butyl 2-ethyl 5-chloro-3-((3-chlorophenyl)(methylthio)methyl)-1*H*-indole-1,2-dicarboxylate 73

The same procedure was used as per the synthesis of **70**. The following equivalents were used: Compound **67** (125 mg, 0.260 mmol), MeI (2.5 eq., 0.040 mL, 0.660 mmol), NEt_3 (2.5 eq., 0.0920 mL, 0.660 mmol). The reaction yielded compound **73** as a yellow oil, 108 mg, 0.218 mmol, 84% (R_f = 0.45, 20% EtOAc/Hex).



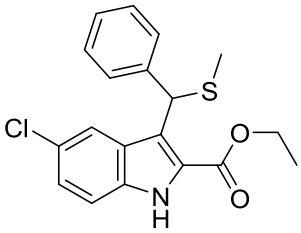
^{1}H NMR (300 MHz, $CDCl_3$) δ 8.06 – 8.01 (m, 1H, ArH), 7.76 – 7.73 (m, 1H, ArH), 7.55 – 7.50 (m, 1H, ArH), 7.45 – 7.38 (m, 1H, ArH), 7.33 (dd, *J* = 9.0, 2.1 Hz, 1H, ArH), 7.29 – 7.19 (m, 2H, ArH), 5.47 (s, 1H, CH), 4.42 (q, *J* = 7.2 Hz, 2H, CH_2), 2.05 (s, 3H, CH_3), 1.65 (s, 9H, 3 × CH_3), 1.39 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$)** δ 162.34 (COO), 148.77, 141.07, 134.88, 134.51, 129.86, 129.81, 128.70, 128.28, 127.79, 129.70, 126.92, 126.14, 122.12, 121.82, 116.49, 85.60, 62.16 (CH_2), 45.84 (CH), 28.04 (3 x CH_3), 15.94 (CH_3), 14.21 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{23}H_{22}Cl_2NO_4$ $[M - SCH_3]^+$, 446.0926, found 446.0917.

8.2.51 Attempted synthesis of ethyl 5-chloro-3-((methylthio)(*m*-tolyl)methyl)-1*H*-indole-2-carboxylate 74

Compound **72** (357 mg, 0.940 mmol) was dissolved in dry MeOH (10 mL) in a 100 mL two-necked round bottom flask. K_3PO_4 (3 eq., 598 mg, 2.82 mmol) and the reaction was heated to 100 °C for 2 hours. Once all starting material was consumed, the reaction mixture was cooled and neutralized with an aqueous sat. NH_4Cl (30 mL). The product was extracted with EtOAc (3 × 30 mL) and the organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude product was purified by column chromatography (5-20% EtOAc/Hex). A 65.0 mg of a white solid was isolated and ^{1}H NMR spectral analysis of this compound revealed that we had not obtained our desired product. The product obtained was thought to be compound **75**, the methyl ester equivalent of the desired compound.

8.2.52 Ethyl 5-chloro-3-((methylthio)(phenyl)methyl)-1*H*-indole-2-carboxylate 76

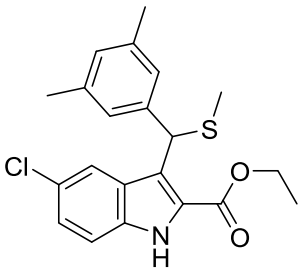
The same procedure was used as per the attempted synthesis of **74**, with EtOH used in place of MeOH as a solvent. The following equivalents were used: Compound **70** (212 mg, 0.46 mmol), K_3PO_4 (3 eq., 293 mg, 1.38 mmol). The reaction yielded compound 76 as a white solid, 97.2 mg, 0.270 mmol, 59% (R_f = 0.50, 30% EtOAc/Hex).



Mp 129-132°C **^{1}H NMR (400 MHz, $CDCl_3$) δ** 9.07 (s, 1H, NH), 8.16 – 8.09 (m, 1H, ArH), 7.58-7.52 (m, 2H, ArH), 7.34 – 7.15 (m, 5H, ArH), 6.37 (s, 1H, CH), 4.52 – 4.39 (m, 2H, CH_2), 2.03 (s, 3H, CH_3), 1.44 – 1.40 (m, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 161.90 (COO), 140.66, 134.56, 128.54, 128.14, 127.36, 127.12, 126.33, 125.93, 124.85, 122.92, 122.61, 113.14, 61.49 (CH_2), 46.46 (CH), 16.11 (CH_3), 14.49 (CH_3). **HRMS:** calcd for $C_{19}H_{18}ClNO_2S$ $[M - H]^-$, 358.0669, found 358.0685.

8.2.53 Ethyl 5-chloro-3-((3,5-dimethylphenyl)(methylthio)methyl)-1*H*-indole-2-carboxylate 77

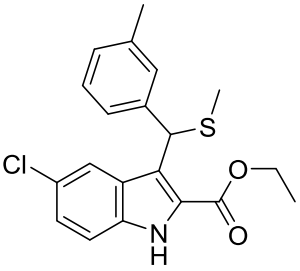
The same procedure was used as per the attempted synthesis of **74**, with EtOH used in place of MeOH as a solvent. The following equivalents were used: Compound **71** (168 mg, 0.340 mmol), K_3PO_4 (3 eq., 223 mg, 1.05 mmol). The reaction yielded compound 77 as a white powder, 107 mg, 0.276 mmol, 81% (R_f = 0.26, 20% EtOAc/Hex).



Mp 122-126 °C **^{1}H NMR (400 MHz, $CDCl_3$) δ** 8.82 (s, 1H, NH), 8.15 – 8.12 (m, 1H, ArH), 7.30 – 7.24 (m, 2H, ArH), 7.16-7.12 (m, 2H, ArH), 6.85 (s, 1H, ArH), 6.28 (s, 1H, CH), 4.44 (q, *J* = 7.1 Hz, 2H, CH_2), 2.26 (s, 6H, 2 × CH_3), 2.02 (s, 3H, CH_3), 1.43 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 161.85 (COO), 140.48, 138.06, 134.54, 128.94, 127.54, 126.42, 125.94, 124.87, 123.17, 122.97, 113.03, 61.46, 46.41, 21.56 (2 x CH_3), 16.19, 14.58, 9.41. **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{20}H_{19}ClNO_2$ $[M - SCH_3]^+$, 340.1104, found 340.1102.

8.2.54 Ethyl 5-chloro-3-((methylthio)(*m*-tolyl)methyl)-1*H*-indole-2-carboxylate **74**

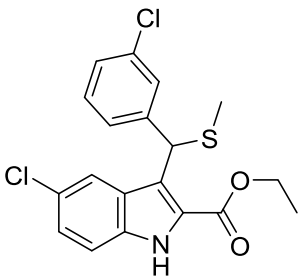
The same procedure was used as per the attempted synthesis of **74**, with EtOH used in place of MeOH as a solvent. The following equivalents were used: Compound **72** (130 mg, 0.284 mmol), K_3PO_4 (3 eq., 181 mg, 0.853 mmol). The reaction yielded compound **74** as a white solid, 89.0 mg, 0.249 mmol, 88% (Rf = 0.51, 20% EtOAc/Hex).



Mp: 108 – 112 °C. **^{1}H NMR (600 MHz, $CDCl_3$) δ** 9.07 (s, 1H, NH), 8.13 – 8.08 (m, 1H, ArH), 7.37 – 7.29 (m, 2H, ArH), 7.25 – 7.14 (m, 3H, ArH), 7.01 (d, *J* = 7.6 Hz, 1H, ArH), 6.36 (s, 1H, CH), 4.47 (q, *J* = 6.5 Hz, 2H, CH_2), 2.34 (s, 3H, CH_3), 2.06 (s, 3H, CH_3), 1.46 (t, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (151 MHz, $CDCl_3$) δ** 164.47 (COO), 143.10, 140.68, 137.12, 131.46, 130.99, 130.49, 129.97, 128.88, 128.47, 127.67, 127.39, 125.56, 125.29, 115.65, 64.03 (CH_2), 48.99 (CH), 24.17 (CH_3), 18.68 (CH_3), 17.06 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{19}H_{17}ClNO_2$ $[M - SCH_3]^+$, 326.0948, found 326.0944.

8.2.55 Ethyl 5-chloro-3-((3-chlorophenyl)(methylthio)methyl)-1*H*-indole-2-carboxylate **78**

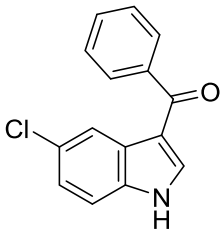
The same procedure was used as per the attempted synthesis of **74**, with EtOH used in place of MeOH as a solvent. The following equivalents were used: Compound **73** (94.0 mg, 0.190 mmol), K_3PO_4 (3 eq., 123 mg, 0.580 mmol). The reaction yielded compound **78** as an opaque white oil, 49.9 mg, 0.131 mmol, 69% (R_f = 0.24, 20% EtOAc/Hex).



^{1}H NMR (400 MHz, $CDCl_3$) δ 8.86 (s, 1H, NH), 8.06 – 8.02 (s, 1H, ArH), 7.56 – 7.12 (m, 6H, ArH), 6.32 (s, 1H, CH), 4.44 (q, *J* = 7.1 Hz, 2H, CH_2), 2.04 (s, 3H, CH_3), 1.44 (t, *J* = 7.1, 3H, CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 161.69 (COO), 142.85, 134.56, 129.92, 128.54, 127.47, 127.32, 126.70, 126.33, 125.05, 124.80, 122.81, 122.07, 113.28, 113.06, 61.67 (CH_2), 46.07 (CH), 16.27 (CH_3), 14.64 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{18}H_{14}Cl_2NO_2$ $[M - SCH_3]^+$, 346.0402, found 346.0394.

8.2.56 (5-chloro-1H-indol-3-yl)(phenyl)methanone 81

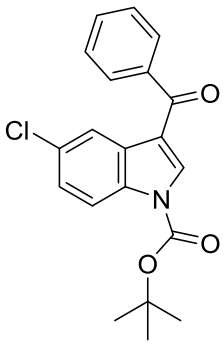
The same procedure was used as per the synthesis of **29**. The following equivalents were used: 5-chloro-1*H*-indole (1.00 g, 6.60 mmol), benzoyl chloride (3 eq., 2.30 mL, 20.0 mmol), $AlCl_3$ (3 eq., 2.63 g, 20.0 mmol). The acylation product was precipitated from EtOAc/hexane to reveal compound **81** as a cream coloured solid, 799 mg, 3.12 mmol, 47% (R_f = 0.07, 20% EtOAc/Hex).



Mp 140 – 146 °C **^{1}H NMR (400 MHz, DMSO) δ** 8.25 (d, *J* = 2.0 Hz, 1H, NH), 8.04 (d, *J* = 3.1 Hz, 1H, ArH), 7.82 – 7.77 (m, 2H, ArH), 7.65 – 7.50 (m, 5H, ArH), 7.29 (m, 1H, ArH). **^{13}C NMR (101 MHz, DMSO) δ** 189.81 (CO), 140.06, 137.03, 135.22, 131.29, 128.48, 128.40 127.45, 126.65, 123.19, 120.55, 114.53, 113.93. **HRMS:** Calcd for $C_{15}H_{10}ClNO$ $[M + H]^+$, 256.0529, found 256.0529.

8.2.57 *Tert*-butyl 3-benzoyl-5-chloro-1H-indole-1-carboxylate 82

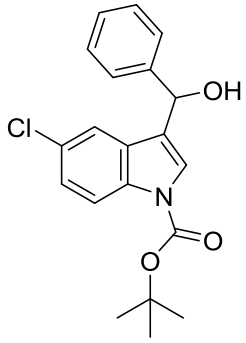
The same procedure was used as per the synthesis of **53**. The following equivalents were used: Compound **81** (500 mg, 1.96mmol), Boc_2O (1.2 eq., 0.55 mL, 2.4 mmol), DMAP (cat.). The reaction yielded compound **53** as a brown solid, 987 mg, 2.23 mmol, 84% (R_f = 0.39, 20% EtOAc/Hex).



Mp 118-136 °C **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.37 (dd, *J* = 2.2, 0.5 Hz, 1H, ArH), 8.10 – 8.03 (m, 2H, ArH), 7.88 – 7.84 (m, 2H, ArH), 7.64 – 7.58 (m, 1H, ArH), 7.52 (m, 2H, ArH), 7.40 – 7.34 (m, 1H, ArH), 1.69 (s, 9H, 3 × CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 190.99 (CO), 148.99, 139.41, 134.75, 134.04, 132.34, 130.47, 129.64, 129.01, 128.71, 126.11, 122.36, 118.86, 116.18, 86.06, 28.21 (3 x CH_3). **HRMS:** Calcd for $C_{20}H_{18}NO_3Cl$ $[M + H]^+$, 356.1053, found 356.1045.

8.2.58 *Tert*-butyl 5-chloro-3-(hydroxy(phenyl)methyl)-1*H*-indole-1-carboxylate 83

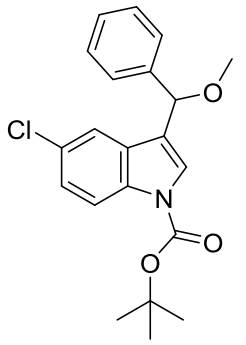
The same procedure was used as per the synthesis of **12.** The following equivalents were used: Compound **82** (500 mg, 1.3 mmol), $NaBH_4$ (3 eq., 163 mg, 4.30 mmol). The reaction yielded compound **83** as a white solid, 451 mg, 1.26 mmol, 90% (R_f = 0.27, 20% EtOAc/Hex).



Mp 48-50 °C **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.04 (d, *J* = 8.8 Hz, 1H, ArH), 7.52 – 7.42 (m, 4H, ArH), 7.41 – 7.29 (m, 3H, ArH), 7.26 – 7.22 (m, 1H, ArH), 6.01 (s, 1H, CH), 2.34 (s, 1H, OH), 1.65 (s, 9H, 3 × CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 148.67, 142.77, 133.66, 129.71, 128.55, 127.43, 127.22, 126.99, 125.29, 124.54, 123.36, 119.55, 116.39, 84.49, 37.76 (CH), 27.62 (3 x CH_3). **HRMS:** Calcd for $C_{20}H_{20}NO_3Cl$ $[M - H]^-$, 356.1053, found 356.1045.

8.2.59 *Tert*-butyl 5-chloro-3-(methoxy(phenyl)methyl)-1H-indole-1-carboxylate **84**

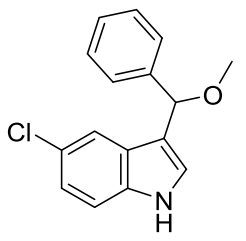
The same procedure was used as per the synthesis of **12.**The following equivalents were used: Compound **83** (200 mg, 0.57 mmol), *p*-TSOH (8 eq., 870 mg, 4.6 mmol). The reaction yielded compound **84** as a cream coloured oil, 144 mg, 0.39 mmol, 70% (R_f = 0.58, 20% EtOAc/Hex).



^{1}H NMR (400 MHz, $CDCl_3$) δ 8.03 (d, *J* = 8.5 Hz, 1H, ArH), 7.49 (d, *J* = 2.0 Hz, 1H, ArH), 7.46 – 7.41 (m, 3H, ArH), 7.40 – 7.34 (m, 2H, ArH), 7.34 – 7.28 (m, 1H, ArH), 7.24 (m, 1H, ArH), 5.43 (s, 1H, CH), 3.42 (s, 3H,), 1.65 (s, 9H). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 140.24, 128.62, 128.39, 128.08, 127.18, 125.21, 124.76, 121.50, 120.00, 116.35, 84.29 79.25, 77.43, 77.11, 76.80, 57.02 (CH_3), 28.25 (3 x CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{20}H_{19}NO_2Cl$ $[M – OCH_3]^-$, 340.1104, found 340.1094.

8.2.60 5-chloro-3-(methoxy(phenyl)methyl)-1H-indole **79**

The same procedure was used as per the attempted synthesis of **74**. The following equivalents were used: Compound **84** (120 mg, 0.310 mmol), K_3PO_4 (3eq., 199 mg, 0.940 mmol). The reaction yielded compound **79** as a cream coloured solid, 60.9 mg, 0.22 mmol, 72% (R_f = 0.23, 20% EtOAc/Hex).

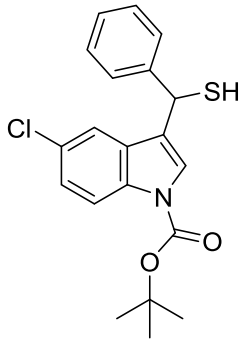


Mp 102-108 °C **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.05 (s, 1H, NH), 7.52 (m, 1H), 7.39 – 7.32 (m, 2H), 7.32 – 7.17 (m, 3H), 7.09 – 6.97 (m, 2H), 6.65 (m, 1H), 5.43 (s, 1H), 3.33 (s, 3H). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 141.19, 135.01, 128.52, 127.77, 127.38, 127.20, 125.58, 124.57, 122.70, 119.37, 117.73, 112.29, 79.52, 77.58, 77.16, 76.74, 56.90. **HRMS**: Note: No parent molecular ion

was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{15}H_{11}ClN$ $[M - OCH_3]^-$, 240.0580, found 240.0573.

8.2.61 *Tert*-butyl 5-chloro-3-(mercapto(phenyl)methyl)-1*H*-indole-1-carboxylate **85**

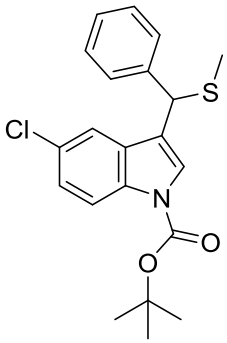
The same procedure was used as per the synthesis of 64. The following equivalents were used: Compound **83** (620 mg, 1.77 mmol), Lawessons' reagent (0.6 eq., 430 mg, 1.06 mmol). The reaction yielded compound **85** as a yellow oil, 240 mg, 0.642 mmol, 36% (R_f = 0.54, 20% EtOAc/Hex).



^{1}H NMR (300 MHz, d_6-DMSO) δ 8.01 (d, *J* = 8.9 Hz, 1H, ArH), 7.66 (d, *J* = 1.0 Hz, 1H, ArH), 7.52 – 7.45 (m, 3H, ArH), 7.38 – 7.29 (m, 3H, ArH), 7.29 – 7.21 (m, 1H, ArH), 5.68 (d, *J* = 6.7 Hz, 1H, CH), 3.72 (d, *J* = 6.7 Hz, 1H, SH), 1.62 (s, 9H, 3 × CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 148.67, 142.77, 133.66, 129.71, 128.55, 127.43, 127.22, 126.99, 125.29, 124.54, 123.36, 119.55, 116.39, 84.49, 37.76, 27.62 (3 x CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{20}H_{19}ClNO_2$ $[M - SH]^+$, 340.1104, found 340.1096.

8.2.62 *Tert*-butyl 5-chloro-3-((methylthio)(phenyl)methyl)-1*H*-indole-1-carboxylate **86**

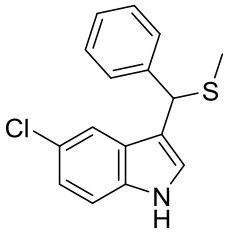
The same procedure was used as per the synthesis of **70**. The following equivalents were used: Compound **85** (200 mg, 0.55 mmol), NEt_3 (2.5 eq., 0.190 mL, 1.40 mmol). MeI (2.5 eq., 0.0850 mL, 1.40 mmol) The reaction yielded compound **86** as a clear oil, 151 mg, 0.389 mmol, 71% (R_f = 0.54, 20% EtOAc/Hex).



^{1}H NMR (300 MHz, $CDCl_3$) δ 8.05 (d, *J* = 8.3 Hz, 1H, ArH), 7.66 – 7.17 (m, 8H, ArH), 5.16 (s, 1H, CH), 2.04 (s, 3H, CH_3), 1.67 (s, 9H, 3 × CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 149.30, 142.20, 139.99, 128.81, 128.67, 128.25, 127.65, 127.56, 127.48, 124.75, 119.67, 116.37, 84.30, 47.35, 28.15 (3 x CH_3), 15.69 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{20}H_{19}ClNO_2$ $[M - SCH_3]^+$, 340.1104, found 340.1094.

8.2.63 5-chloro-3-((methylthio)(phenyl)methyl)-1*H*-indole **80**

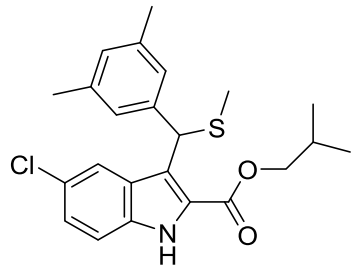
The same procedure was used as per the attempted synthesis of **74**. The following equivalents were used: Compound **86** (107 mg, 0.28 mmol), K_3PO_4 (3 eq., 179 mg, 0.84 mmol). The reaction yielded compound **80** as a white solid, 47.8 mg, 0.166 mmol, 62% (R_f = 0.28, 20% EtOAc/Hex).



Mp 94-96 °C **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.98 (s, 1H, NH), 7.65 – 7.60 (m, 1H, ArH), 7.50 – 7.42 (m, 2H, ArH), 7.38 – 6.99 (m, 6H, ArH), 5.23 (s, 1H, CH), 2.00 (s, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 141.24, 134.96, 128.63, 128.36, 127.53, 127.35, 125.38, 124.62, 122.74, 119.32, 116.54, 112.34, 47.96, 15.80 (CH_3). **HRMS:** Calcd for $C_{16}H_{14}ClNS$ $[M - H]^-$, 286.0457, found 286.0469.

8.2.64 Synthesis of isobutyl 5-chloro-3-((3,5-dimethylphenyl)(methylthio)methyl)-1H-indole-2-carboxylate **87**

The same procedure was used as per the attempted synthesis of **74**, using *i*-BuOH in place of MeOH as a solvent. The following equivalents were used: Compound 77 (65.8 mg, 0.180 mmol), K_3PO_4 (3 eq., 112 mg, 0.530 mmol). The reaction yielded compound **87** as a yellow oil, 25.6 mg, 0.064 mmol, 34% (R_f = 0.31, 20% EtOAc/Hex).



^{1}H NMR (400 MHz, $CDCl_3$) δ 8.83 (s, 1H, NH), 8.12 (s, 1H, ArH), 7.34 – 7.20 (m, 2H, ArH), 7.14 (s, 2H), 6.85 (s, 1H, ArH), 6.30 (s, 1H, CH), 4.18 (d, *J* = 6.6 Hz, 2H), 2.26 (s, 6H, 2 x CH_3), 2.12 (m, 1H, CH), 2.02 (s, 3H, CH_3), 1.26 (s, 1H), 1.04 (m, 6H, 2 x CH_3). **^{13}C NMR (101 MHz, $CDCl_3$) δ** 162.11, 140.33, 138.06, 134.59, 128.92, 127.56, 126.43, 125.97, 125.93, 124.92, 123.19, 122.79, 113.04, 71.63, 46.36, 29.85, 28.02, 21.55, 19.44, 19.41, 16.16. **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{22}H_{23}ClNO_2$ $[M - SCH_3]^+$, 368.1417, found 368.1418.

8.2.65 Stability testing

Compound **19**, **21** and **74** were made up into three separate 10.3 µM solutions in 10 mL of EtOH in three 20 mL RBFs. 20 µL of H_2SO_4 was added to each RBF at t_o and thereafter every 30 minutes a 200 µL sample was taken from each reaction and diluted with 400 µL of ACN and 400 µL of H_2O. The diluted samples were then injected into a Waters 1525 Binary HPLC coupled to a Waters 2487

Dual λ absorbance detector. The run time was 8 minutes using a split volume of 0.65 ACN/0.15 water as the eluent. The peak areas were normalised in order to allow for simplified comparison, the results are shown in

Table *15*.

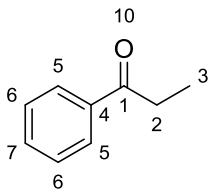
Table 15

	Peak area starting material normalised (%)		
time (min.)	**Compound** 19	**Compound 74**	**Compound** 21
0	100.00	100.00	100.00
30	37.47	97.60	95.85
60	14.30	97.83	92.40
90	4.94	97.08	89.44
120	1.84	97.39	85.28
150	0.63	96.93	82.96
180	0.00	97.26	79.07
210	0.00	96.92	76.01
240	0.00	97.27	73.02
1260	0.00	97.83	15.92

8.3 Experimental Procedures Pertaining to Chapter 4

8.3.1 1-propiophenone **93**

1-propionyl chloride (3 eq., 6.70 mL, 76.8 mmol) was added to a two-necked RBF fitted with a condenser charged with 30 mL of DCE. The flask was placed on ice and $AlCl_3$ (3 eq., 6.83 g, 76.8 mmol). After stirring for 15 mins, benzene (2.28 mL, 25.6 mmol) was added and the reaction mixture was heated to 50 °C for 3h. There was a colour change observed from colourless to dark red. After cooling, the reaction mixture was poured over a mixture of ice and sat. $NaHCO_3$ and the product was extracted with EtOAc (3 x 50 mL). The organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude product was purified by column chromatography (5-30% EtOAc/Hex). The reaction yielded compound **93** as a yellow oil, 2.65 g, 19.7 mmol, 77% (R_f = 0.44, 20% EtOAc/Hex).



^{1}H NMR (300 MHz, $CDCl_3$) δ 7.95 (m, 2H, H_5), 7.56 – 7.49 (m, 1H, H_7), 7.48 – 7.38 (m, 2H, H_6), 2.98 (q, *J* = 7.2 Hz, 2H, H_2), 1.21 (t, *J* = 8.3 Hz, 3H, H_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 200.83 (C_1), 136.98 (C_4), 132.92 (C_7), 128.60 (C_5), 128.01 (C_6), 31.82 (C_2), 8.29 (C_3).

8.3.2 Attempted synthesis of ethyl 5-chloro-3-(1-phenylpropyl)-1H-indole-2-carboxylate **90** by reductive elimination with Et_3SiH

Et_3SiH (3eq., 0.43 mL, 2.70 mmol) and TCA (1.5 eq., 219 mg, 1.34 mmol) were added to a three-necked RBF fitted with a condenser charged with toluene. In a dropping funnel, ethyl 5-chloro-*1H*-indole carboxylate (200 mg, 0.89 mmol) was dissolved in toluene and compound **93** (1.1 eq., 0.13 mL, 0.98 mmol) was added. The contents of the dropping funnel was added dropwise to the RBF at rt and once everything was added the reaction mixture was heated to 70 °C. After two days only starting material was present, and this was recovered to use in a different strategy.

8.3.3 Attempted synthesis of 5-chloro-3-(1-phenylpropyl)-1H-indole **96** by reductive elimination with Et_3SiH

The same procedure as per 8.3.2 was utilised. The following equivalents were used: 5-chloroindole (100 mg, 0.66 mmol), Et_3SiH (3 eq., 0.32 mL, 1.96 mmol), TCA (1.5 eq., 162 mg, 0.99

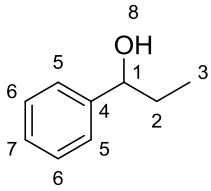
mmol), compound **93** (1.1 eq., 0.1 mL, 0.73 mmol). Again only starting material was present after two days of refluxing at 70 °C and this was recovered for use in a different strategy.

8.3.4 Attempted synthesis of 5-chloro-3-(1-phenylpropyl)-1H-indole 96 by reductive elimination with H_2 and Pd/C

Pd/C (10 mol%, 15 mg, 0.11 mmol) and TFA (1.5 eq., 0.08 mL, 0.99 mmol) were added to a three-necked RBF charged with DCM (10 mL). 5-chloroindole (100 mg, 0.66 mmol) and compound **93** was added to the RBF before the flask was placed under a H_2 atmosphere. After 24 h only starting material was present and this was recovered for use in a different strategy.

8.3.5 1-phenylpropan-1-ol 97

Compound **93** (4.10 g, 30.3 mmol) was dissolved in 40 mL of EtOH in a 150 mL RBF. $NaBH_4$ (3 eq., 3.50 g, 91.2 mmol) was dissolved in a 50 mL beaker in 10 mL of distilled water and swirled before adding this solution dropwise to the RBF at 0 °C. The reaction was complete in complete in 5 mins after which it was slowly transferred into a mixture of sat. NH_4Cl and ice. After the fizzing had stopped, the product was extracted with EtOAc (3 x 40 mL) and the organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude product was purified by column chromatography (5-20% EtOAc/Hex). The reaction yielded compound **97** as a yellow oil, 4.10 g, 30.1 mmol, 99% (R_f = 0.39, 20% EtOAc/Hex).



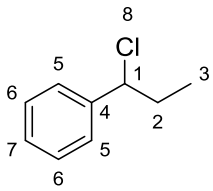
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.37 – 7.23 (m, 4H, H_{5-7}), 4.59 (m, 1H, H_1), 1.90 – 1.86 (m, 1H, H_8), 1.85 – 1.67 (m, 2H, H_2), 1.01 – 0.79 (m, 3H, H_3).

8.3.6 Attempted synthesis of 5-chloro-3-(1-phenylpropyl)-1H-indole 96 by reductive elimination with I_2

Compound **97** (1 eq., 230 mg, 1.70 mmol) and 5-chloroindole (267 mg, 1.70 mmol) were added to a two-necked RBF fitted with a condenser charged with 10 mL of ACN. The flask was lowered into an ice bath and I_2 (5 mol%, 18.6 mg, 0.07 mmol) was added. The reaction was heated to 50 °C and stirred for 5 days, after which only starting material was present and this was recovered for use in a different strategy.

8.3.7 (1-chloropropyl)benzene 98

Compound **97** (1.33 g, 9.76 mmol) was added to a two-necked RBF fitted with a condenser. $SOCl_2$ (1.5 eq, 1.16 g, 14.6 mmol) was added dropwise at 0 °C. The reaction mixture was then refluxed at 80 °C for 5 mins before leaving to cool and this was quenched with ice first and then sat. $NaHCO_3$ was added. The product was extracted with Et_2O (3 x 20 mL) and the organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude product was purified by column chromatography (2-10% EtOAc/Hex). The reaction yielded compound **98** as a yellow oil, 805 mg, 5.20 mmol, 53% (R_f = 0.75, 10% EtOAc/Hex).



^{1}H NMR (300 MHz, $CDCl_3$) δ 7.41 – 7.26 (m, 4H, $H_{5\text{-}7}$), 4.78 (m, 1H, H_1), 2.11 (m, 2H, H_2), 1.00 (t, *J* = 9.3, 3H, H_3)**. ^{13}C NMR (75 MHz, $CDCl_3$) δ** 128.71, 128.32, 127.12, 77.58, 77.16, 76.74, 65.63 (C_1), 33.36 (C_2), 11.87 (C_3).

8.3.8 Attempted synthesis of ethyl 5-chloro-3-(1-phenylpropyl)-1H-indole-2-carboxylate 90 by Friedel-Crafts alkylation

Compound **98** (1 eq., 0.070 mL, 0.45 mmol) was dissolved in 30 mL of DCE in a 100 mL RBF. $AlCl_3$ (1 eq., 60 mg, 0.45 mmol) was added at 0 °C The reaction mixture was stirred for 15 minutes before ethyl 5-chloro-*1H*-indole carboxylate (100 mg, 0.45 mmol) was added and the reaction was then heated to 90 °C for 24 h after which the reaction mixture was dark red in colour. A number of spots were observed on TLC, but there was still a prominent starting material spot. The reaction was cooled and quenched with aqueous sat. $NaHCO_3$, extracted with EtOAc (3 x 20 mL) and the organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude material was purified by column chromatography, however the only considerable amount of compound which was isolated was starting material.

8.3.9 Attempted synthesis of 5-chloro-3-(1-phenylpropyl)-1H-indole 96 by Friedel-Crafts alkylation

The same procedure was followed as per 8.3.8. The following equivalents were used: 5-chloroindole (100 mg, 0.66 mmol), $AlCl_3$ (1 eq., 88 mg, 0.66 mmol), compound **98** (1 eq., 102 mg, 0.66 mmol). Once again only starting material was recovered from this reaction after refluxing for 24h.

8.3.10 Attempted synthesis of 5-chloro-3-(1-phenylpropyl)-1H-indole **96** using Mg as an activator

MeMgI was prepared by adding Mg (1.5 eq, 24 mg, 0.99 mmol) and MeI (1.5 eq, 0.062 mL, 0.99 mmol) to a 20 mL RBF charged with 5 mL of THF. This was stirred for an hour before 5-chloroindole (100 mg, 0.66 mmol) was added followed directly by compound **98** (1.5 eq, 0.13 mL, 0.99 mmol) at 0 °C. After 18h only starting material was present and this was recovered for later use.

8.3.11 *p*-toluenesulfinic acid 105

$NaHCO_3$ (2 eq., 1.44 g, 17.2 mmol) and Na_2SO_3 (2 eq., 2.17 g, 17.2 mmol) were dissolved in distilled water in a 100 mL RBF fitted with a condenser and this was heated to 80 °C for 2h. *p*-TsCl (3.3 g, 18 mmol) was added and the heating was continued for another 1h. The reaction mixture was left to cool before acidifying with 4M HCl. This mixture was left overnight and the crystals that were formed were filtered off and dried to be used without further purification in the next reaction.

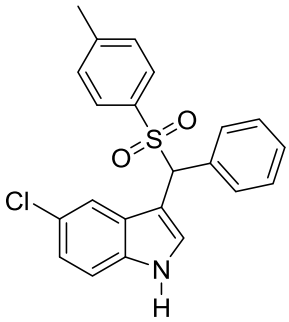
8.3.12 Attempted synthesis of the ethyl 5-chloro-3-(phenyl(tosyl)methyl)-1H-indole-2-carboxylate **92**

Ethyl 5-chloro-*1H*-indole carboxylate (500 mg, 2.20 mmol), compound **105** (1.2 eq., 465 mg, 2.6 mmol), *p*-TsOH (0.5 eq, 210 mg, 1.10 mmol) and benzaldehyde (1.2 eq., 0.3 mL, 2.6 mmol) were all added to a two-necked RBF fitted with a condenser charged with EtOAc (30 mL). The reaction mixture was refluxed at 80 °C for 3.5h and a colour change was observed from clear to bright red. The reaction mixture was quenched with sat. $NaHCO_3$, extracted with EtOAc (3 x 20 mL) and the organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude material was purified by column chromatography (2-10% EtOAc/Hex). Although there were several products isolated none of them were the desired compound **92** and 60% of the starting material was recovered.

8.3.13 5-chloro-3-(phenyl(tosyl)methyl)-1H-indole **106** using *p*-toluenesulfinic acid

The same procedure was followed as per 8.3.12. The following equivalents were used: 5-chloroindole (182 mg, 1.2 mmol), compound **105** (1.1 eq., 200 mg, 1.3 mmol), *p*-TsOH (0.5 eq.,

107 mg, 0.6 mmol), benzaldehyde (1.1 eq., 0.13 mL, 1.3 mmol). The reaction yielded compound **106** as a red solid, 292 mg, 0.73 mmol, 62% (R_f = 0.07, 20% EtOAc/Hex).



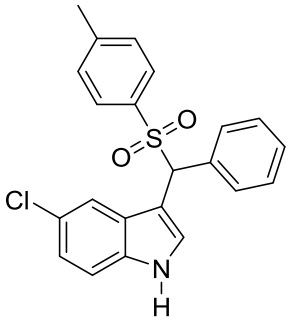
Mp 70 – 75 °C **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.64 (d, *J* = 2.7 Hz, 1H, ArH), 7.53 – 7.41 (m, 4H, ArH), 7.31 – 7.23 (m, 4H, ArH), 7.17 – 7.08 (m, 3H, ArH), 7.03 (m, 1H, ArH), 5.57 (s, 1H, CH), 2.35 (s, 3H, CH_3). **^{13}C NMR (75 MHz, $CdCl_3$)** δ 144.79, 135.13, 133.98, 133.37, 130.12, 129.48, 129.13, 128.77, 128.70, 128.24, 126.50, 125.92, 122.82, 117.78, 112.69, 107.01, 69.10 (CH), 21.72 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{15}H_{11}ClN$ [M - sulfonyl]$^+$, 240.0580, found 240.0576.

8.3.14 sodium *p*-toluenesulfinate 107

$NaHCO_3$ (2 eq., 1.44 g, 17.2 mmol) and Na_2SO_3 (2 eq., 2.17 g, 17.2 mmol) were dissolved in distilled water in a 100 mL RBF fitted with a condenser and this was heated to 80 °C for 2h. *p*-TsCl (3.3 g, 18 mmol) was added and the heating was continued for another 1h. The reaction mixture was left left overnight and the crystals that were formed were filtered off and dried to be used without further purification in the next reaction.

8.3.15 5-chloro-3-(phenyl(tosyl)methyl)-1H-indole 106 using sodium *p*-toluenesulfinate

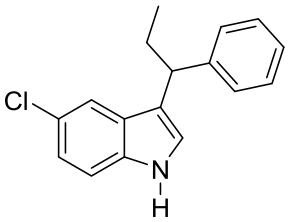
The same procedure was followed as per 8.3.12. The following equivalents were used: 5-chloroindole (1 g, 6.6 mmol), compound **107** (1.2 eq., 1.4 g, 7.9 mmol), *p*-TsOH (1.5 eq., 1.9 g, 9.9 mmol), benzaldehyde (1.2 eq., 0.81 mL, 7.9 mmol). The reaction yielded compound **106** as a red solid, 1.35 g, 3.4 mmol, 52% (R_f = 0.07, 20% EtOAc/Hex).



Mp 70 – 75 °C **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.64 (d, *J* = 2.7 Hz, 1H, ArH), 7.53 – 7.41 (m, 4H, ArH), 7.31 – 7.23 (m, 4H, ArH), 7.17 – 7.08 (m, 3H, ArH), 7.03 (m, 1H, ArH), 5.57 (s, 1H, CH), 2.35 (s, 3H, CH_3). **^{13}C NMR (75 MHz, $CdCl_3$)** δ 144.79, 135.13, 133.98, 133.37, 130.12, 129.48, 129.13, 128.77, 128.70, 128.24, 126.50, 125.92, 122.82, 117.78, 112.69, 107.01, 69.10 (CH), 21.72 (CH_3). **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{15}H_{11}ClN$ [M - sulfonyl]$^+$, 240.0580, found 240.0576.

8.3.16 5-chloro-3-(1-phenylpropyl)-1H-indole 96

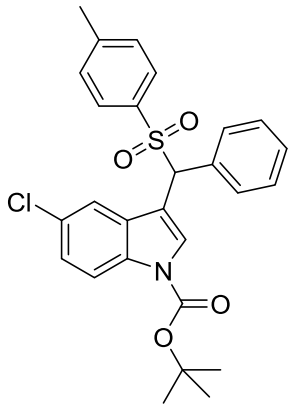
EtMgI was prepared by adding Mg (3.5 eq, 20.1 mg, 0.825 mmol) and EtI (2.2 eq, 0.04 mL, 0.55 mmol) to a 20 mL RBF charged with 5 mL of Et_2O. This was stirred for an hour before it was transferred to a 100 mL RBF containing compound **106** dissolved in 25 mL of Et_2O at -35 °C. The reaction was left to heat up to rt for 18h before it was quenched with sat. NH_4Cl and extracted with Et_2O (3 x 15 mL) and the organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude material was purified by column chromatography (10% EtOAc/Hex). The reaction yielded compound **96** as a cream coloured solid, 44.5 mg, 0.16 mmol, 66% (R_f = 0.36, 20% EtOAc/Hex).



Mp 111 – 114 °C **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.87 (s, 1H, NH), 7.41 – 7.36 (m, 1H, ArH), 7.29 – 7.22 (m, 4H, ArH), 7.20 – 7.13 (m, 2H, ArH), 7.09 – 7.02 (m, 1H, ArH), 7.00 – 6.95 (m, 1H, ArH), 3.96 (t, *J* = 7.78, 1H, CH), 2.27 – 1.88 (m, 2H, CH_2), 0.92 (t, *J* = 7.3 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 144.97, 134.91, 128.46, 128.33, 128.01, 126.21, 125.00, 122.46, 122.33, 120.39, 119.06, 112.12, 44.75 (CH), 29.16 (CH_2), 12.92 (CH_3). **HRMS:** Sample had degraded before this analysis could be performed.

8.3.17 *Tert*-butyl 5-chloro-3-(phenyl(tosyl)methyl)-1H-indole-1-carboxylate 108

The same procedure was used as per the synthesis of **56**. The following equivalents were used: Compound **106** (1.0 g, 2.53 mmol), Boc_2O (1.2 eq., 0.70 mL, 3.0 mmol), DMAP (cat.). The reaction yielded compound **108** as a cream coloured solid, 1.0 g, 2.02 mmol, 80% (R_f = 0.36, 20% EtOAc/Hex).



Mp 129 – 136 °C **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.17 – 8.11 (m, 1H, ArH), 8.08 – 8.01 (m, 1H, ArH), 7.55 – 7.47 (m, 2H, ArH), 7.45 – 7.37 (m, 2H, ArH), 7.33 – 7.11 (m, 7H, ArH), 5.47 (s, 1H, CH), 2.37 (s, 3H, CH_3), 1.68 (s, 9H, CH_3). **^{13}C NMR (151 MHz, $CDCl_3$) δ** 149.17, 144.96, 134.98, 133.41, 132.45, 130.93, 130.11, 129.52, 129.22, 129.05, 128.80, 128.67, 127.23, 125.09, 118.38, 116.47, 112.00, 84.87, 68.32 (CH), 28.26 (3 x CH_3), 21.72 (CH_3). **HRMS:** Calcd for $C_{27}H_{26}ClNO_4Na$ $[M + Na]^+$, 518.1169, found 518.1174.

8.3.18 Attempted synthesis of *tert*-butyl 5-chloro-3-(1-phenylpropyl)-1H-indole-1-carboxylate 109

The same procedure was followed as per 8.3.12. The following equivalents were used: Compound **108** (250 mg, 0.500 mmol), Mg (5 eq, 60.8 mg, 2.50 mmol), EtI (1.2 eq, 0.500 mL, 0.6 mmol). Column chromatography was not successful in isolating the desired compound **109**.

8.3.19 Attempted synthesis of 5-chloro-3-(phenyl(tosyl)methyl)-1-tosyl-1H-indole 110

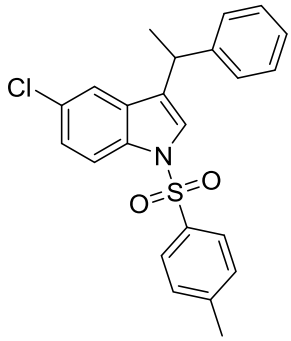
The same procedure was used as per the synthesis of **38**. The following equivalents were used: Compound **106** (330 mg, 0.83 mmol), NaH (1.5 eq., 60%, 45 mg, 1.25 mmol), *p*-TsCl (1.5 eq., 238 mg, 1.25 mmol). Column chromatography was not successful in isolating the desired compound **110**.

8.3.20 5-chloro-3-(1-phenylethyl)-1H-indole 111

The same procedure was used as per the synthesis of **96**. The following equivalents were used: Compound **106** (300 mg, 0.78 mmol), Mg (4 eq., 73.7 mg, 3.03 mmol), MeI (2.5 eq, 0.12 mL, 1.95 mmol). The reaction yielded 112 mg of a brown solid, however problems with purification led us to use this sample of compound **111** directly in the next reaction without performing characterization.

8.3.21 5-chloro-3-(1-phenylethyl)-1-tosyl-1H-indole 112

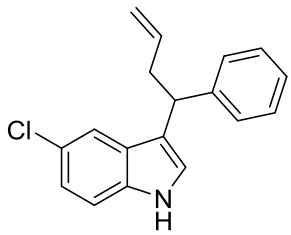
The same procedure was used as per the synthesis of **38**. The following equivalents were used: Compound **111** (110 mg, 0.43 mmol), NaH (1.5 eq., 60% in mineral oil, 24 mg, 0.65 mmol), *p*-TsCl (1.5 eq., 123 mg, 0.65 mmol). The reaction yielded compound **112** as a brown oil, 152 mg, 0.37 mmol, 47% over two steps (R_f = 0.83, 50% EtOAc/Hex).



^{1}H NMR (300 MHz, $CDCl_3$) δ 7.89 – 7.84 (m, 1H, ArH), 7.75 – 7.68 (m, 2H, ArH), 7.44 – 7.41 (m, 1H, ArH), 7.27 – 7.10 (m, 10H), 4.15 (q, *J* = 7.1 Hz, 1H, CH), 2.36 (s, 3H, CH_3), 1.64 (d, *J* = 7.1 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 145.22, 144.51, 135.10, 134.18, 131.85, 130.04, 129.06, 128.78, 127.37, 127.30, 126.86, 126.75, 125.02, 124.36, 120.12, 114.95, 36.86 (CH), 22.13 (CH_3), 21.74 (CH_3). **HRMS:** Calcd for $C_{23}H_{20}ClNO_2SNa$ $[M + Na]^+$, 432.0803, found 432.0804.

8.3.22 5-chloro-3-(1-phenylbut-3-en-1-yl)-1H-indole **113**

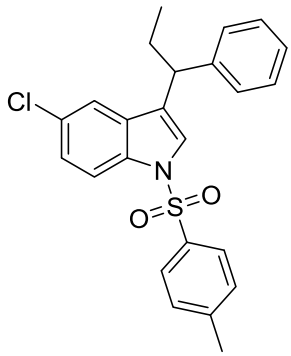
The same procedure was used as per the synthesis of **96**. The following equivalents were used: Compound **106** (300 mg, 0.78 mmol), Mg (4 eq., 73.7 mg, 3.03 mmol), allyl bromide (2.5 eq, 0.17 mL, 1.95 mmol). The reaction yielded compound **113** as a brown solid, 39.2 mg, 0.14 mmol, 18% (R_f = 0.55, 20% EtOAc/Hex).



Mp 143 – 148 °C **^{1}H NMR (300 MHz, cdcl$_3$) δ** 7.94 (s, 1H, NH), 7.47 – 6.93 (m, 9H, ArH), 5.91 – 5.59 (m, 1H, CH), 5.00 (m, CH_2), 4.18 (m, 1H, CH), 3.02 – 2.58 (m, 2H, CH_2). **^{13}C NMR (75 MHz, CDCl$_3$) δ** 144.41, 137.11, 134.89, 128.51, 128.22, 127.98, 126.38, 125.11, 122.83, 122.46, 119.61, 119.05, 116.33, 112.15, 77.59, 77.16, 76.74, 42.91, 40.59. **HRMS:** Calcd for $C_{18}H_{17}ClN$ $[M + H]^+$, 282.1050, found 282.1030.

8.3.23 5-chloro-3-(1-phenylpropyl)-1-tosyl-1H-indole **114**

The same procedure was used as per the synthesis of **38**. The following equivalents were used: Compound **96** (150 mg, 0.55 mmol), NaH (1.5 eq., 60%, 30 mg, 0.83 mmol), *p*-TsCl (1.5 eq., 160 mg, 0.83 mmol). The reaction yielded compound **114** as a clear oil, 152 mg, 0.36 mmol, 65% (R_f = 0.44, 20% EtOAc/Hex).



^{1}H NMR (300 MHz, CDCl$_3$) δ 7.88 – 7.82 (m, 1H, ArH), 7.70 (d, *J* = 8.3 Hz, 2H, ArH), 7.44 (d, *J* = 11.0 Hz, 1H, ArH), 7.30 – 7.11 (m, 9H, ArH), 3.87 – 3.79 (m, 1H, CH), 2.35 (s, 3H, CH_3), 2.16 – 1.93 (m, 2H, CH_2), 0.91 (t, *J* = 7.3 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, CDCl$_3$) δ** 145.19, 142.96, 135.06, 134.12, 132.15, 130.02, 129.10, 128.67, 127.91, 126.83, 126.75, 126.49, 125.01, 124.21, 119.97, 114.98, 44.49 (CH), 28.70 (CH_2), 21.73 (CH_3), 12.65 (CH_3). **HRMS:** Calcd for $C_{24}H_{22}ClNO_2Na$ $[M + Na]^+$, 446.0957, found 446.0948.

8.3.24 5-chloro-3-(1-phenylpropyl)-1-tosyl-1H-indole-2-carboxylic acid **115**

Compound **114** was dissolved in degassed THF (10 mL) in a 100 mL RBF. The flask was placed in a -78 °C bath and *n*-BuLi (1.5 eq., 0.35 mL, 0.42 mmol) was added causing the solution to change from clear to yellow in colour. The reaction mixture was removed from the -78 °C bath and stirred for 1h. The flask was again placed in the -78 °C bath before 60 mL of dry ice was added directly to

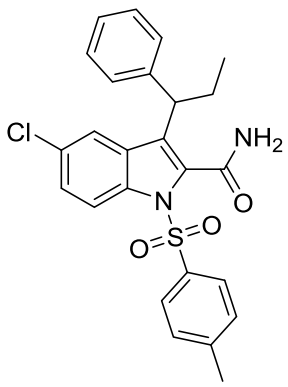
the reaction mixture. The reaction was left to reach rt overnight before it was quenched with sat. NH_4Cl and extracted with Et_2O (3 x 20 mL). The organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The product was used directly in the next reaction without further purification.

8.3.25 Attempted synthesis of ethyl 5-chloro-3-(1-phenylpropyl)-1-tosyl-1H-indole-2-carboxylate 116

The crude sample of compound **115** (80 mg, 0.17 mmol) was dissolved in DMF (5 mL) in a 20 mL RBF. NaH (1.5 eq, 10 mg, 0.25 mmol) was added at 0 °C resulting in a colour change from clear to bright yellow. The reaction was stirred for 15 minutes at 0 °C before EtI (1.5 eq, 0.02 mL, 0.25 mmol) was introduced and the reaction was left to reach rt and stirred for 3h. The reaction was then quenched with sat. NH_4Cl and extracted with EtOAc (3 x 20 mL). The organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude material was purified by column chromatography (10% EtOAc/Hex). ^1H NMR analysis revealed that we had not isolated the desired product compound **116**.

8.3.26 5-chloro-3-(1-phenylpropyl)-1-tosyl-1H-indole-2-carboxamide 117

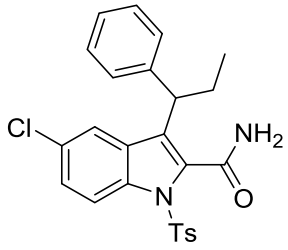
The crude sample of compound **115** (80 mg, 0.17 mmol) was added to a two-necked RBF fitted with a condenser charged with $CHCl_3$ (7 mL). $SOCl_2$ (2 eq. 0.025 mL, 0.33 mmol) was added and the reaction was refluxed for 1h. TLC analysis showed that no reaction was occurring thus DMF (cat.) was added. After a further 3h of refluxing, all starting material had been consumed. The reaction was left to cool and then poured over a mixture of 25% NH_3 and ice. This solution was extracted with EtOAc (3 x 20 mL) and the organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude material was purified by column chromatography (80% EtOAc/Hex). The reaction yielded compound **117** as a white solid, 56 mg, 0.13 mmol, 70% (R_f = 0.33, 100% EtOAc).



Mp 124 – 127 °C **^{1}H NMR (300 MHz, $CDCl_3$) δ** 7.82 (d, *J* = 8.9 Hz, 1H, ArH), 7.57 (d, *J* = 8.2 Hz, 1H, ArH), 7.40 – 7.10 (m, 10H, ArH), 6.75 (s, 2H, NH_2), 4.23 (m, 1H, CH), 2.33 (s, 3H, CH_3), 2.31 – 2.10 (m, 2H, CH_2), 0.90 – 0.85 (m, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 169.85, 164.68, 145.65, 142.31, 136.22, 135.07, 133.18, 132.39, 130.61, 129.69, 129.62, 128.64, 128.57, 127.78, 126.65, 126.24, 125.54, 121.59, 116.07, 43.22 (CH), 25.66 (CH_2), 21.55 (CH_3), 12.90 (CH_3). **HRMS:** Sample had degraded before this analysis could be performed.

8.3.27 5-chloro-3-(1-phenylpropyl)-1H-indole-2-carboxamide 119

The same procedure was used as per the synthesis of 10. The following equivalents were used: Compound **117** (30.6 mg, 0.07 mmol), KOH (3 eq., 11.6 mg, 0.21 mmol). The reaction yielded compound **119** as a white solid, 15.5 mg, 0.049 mmol, 71 % (R_f = 0.28, 50% EtOAc/Hex).



Mp 118 – 122 °C **^{1}H NMR (300 MHz, $CDCl_3$) δ** 9.61 (s, 1H, NH), 7.74 (d, *J* = 1.9 Hz, 1H, ArH), 7.41 – 7.15 (m, 7H, ArH), 5.86 (s, 2H, NH_2), 4.47 (m, 1H, CH), 2.47 – 2.18 (m, 2H, CH_2), 0.91 (t, *J* = 7.6 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 164.24 (CO), 143.92, 134.17, 129.19, 129.14, 128.57, 127.69, 127.00, 126.01, 125.36, 121.24, 120.05, 113.40, 44.19 (CH), 28.16 (CH_2), 13.30 (CH_3). **HRMS:** Calcd for $C_{18}H_{18}ClN_2O$ $[M + H]^+$, 313.1108, found 313.1106.

8.3.29 Ethyl 5-chloro-3-(1-phenylpropyl)-1-tosyl-1H-indole-2-carboxylate 116

The crude sample of compound **115** (30 mg, 0.07 mmol) was added to a 100 mL RBF. $SOCl_2$ (2 eq, 0.12 mL) was added dropwise and the reaction was refluxed for 3h. TLC anlaysis showed all starting material had been consumed. The reaction was cooled and quenched with EtOH (20 mL). After 3h, no change was seen on TLC so 5 mL of distilled water was added in an attempt to reform compound **115**. After 2 days crystals had formed and these were sent for analysis by X-ray crystallography (Table 16). The analysis showed that the reaction had yielded compound **116** as a cream coloured crystalline solid, 24.0 mg, 0.048 mmol, 76 % (R_f = 0.45, 10% EtOAc/Hex).

The crystal structure was obtained by our collaborator at the Crystallography Department, University of Stellenbosch. Single crystals of diffraction quality were obtained, mounted in oil and data were collected using a Bruker SMART Apex III X-ray diffractometer equipped with a Mo fine-focus sealed tube (λ = 0.71073 ♦). Data collections were performed at 298 K using an Oxford Cryostream cryostat (700 series Cryostream Plus) attached to the diffractometer. Reduction of data, adsorption corrections as well as unit cell determination was carried out using Bruker diffraction software APEXIII.[1, 2] All structures were solved using SHELXS-97 and refined using SHELXL-97[3] within the X-Seed[4, 5] graphical user interface. Non-hydrogen atoms were refined anisotropically and hydrogen atoms were placed on calculated positions, except those on oxygen and nitrogen atoms, which were located using electron density maps. Figures were generated using POVRay[6] within the X-Seed graphical user interface.

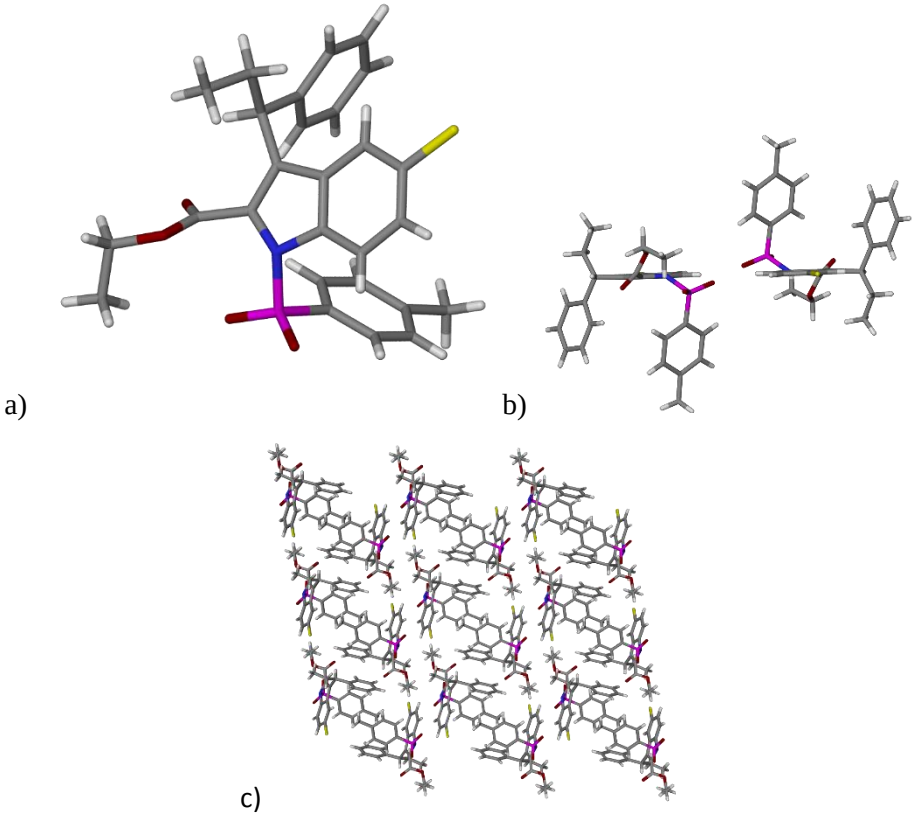


*Figure 30: Figure 1: (a) Asymmetric unit of compound **109**; (b) both R and S molecules; (c) packing of CM_SB_1Ee viewed down the a axis.*

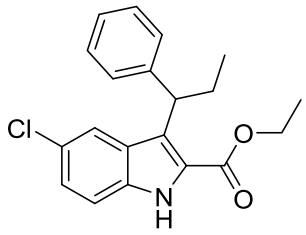
Table 16

	Compound 116

Molecular formula	$0.5(C_{27}H_{26}NO_4SCl)$
***Mr*/g mol^{-1}**	248.01
Temperature/K	298(2)
Crystal system	Triclinic P
Space group	*P 1*
***a*/ ♦**	9.723(3)
***b*/ ♦**	10.466(4)
***c*/ ♦**	12.988(5)
Σ/∀	99.65
T/∀	102.91
Υ/∀	93.05
***V* / ♦^{3}**	1255.8(8)
Z	4
}/mm^{-1}	0.27
R_1 [I > 2′(I)]	0.0525
wR_2 (F^2)	0.1004

8.3.30 Ethyl 5-chloro-3-(1-phenylpropyl)-1H-indole-2-carboxylate 90

The same procedure was used as per the synthesis of 10. The following equivalents were used: Compound **116** (22 mg, 0.04 mmol), KOH (4 eq., 11 mg, 0.19 mmol). The reaction yielded compound **90** as a white solid, 13.8 mg, 0.049mmol, 100% (R_f = 0.41, 20% EtOAc/Hex).



Mp 109 – 114 °C **^{1}H NMR (300 MHz, $CDCl_3$) δ** 8.80 (s, 1H, NH), 7.58 (d, *J* = 6.8 Hz, 1H, ArH), 7.43 – 7.34 (m, 2H, ArH), 7.31 – 7.09 (m, 5H, ArH), 5.10 (q, *J* = 9.5, Hz, 1H, CH), 4.49 – 4.38 (m, 2H, CH_2), 2.46 – 2.15 (m, 2H, CH_2), 1.42 (t, *J* = 7.2 Hz, 3H, CH_3), 0.96 – 0.85 (t, *J* = 7.2 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 162.25, 144.47, 134.60, 128.40, 127.92, 127.71, 126.33, 126.10, 125.91, 125.71, 125.05, 122.17, 113.09, 61.22, 42.92 (CH_2), 29.85 (CH), 27.08 (CH_2), 14.57 (CH_3), 13.03 (CH_3). **HRMS**: Calcd for $C_{20}H_{21}ClNO_2$ $[M + H]^+$, 342.1261, found 342.1263.

8.4 Experimental procedures pertaining to Chapter 5

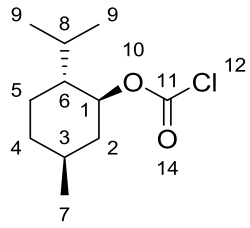
8.4.1 Attempted synthesis of ethyl 5-chloro-1-((((1S)-7,7-dimethyl-2-oxobicyclo[2.2.1]heptan-1-yl)methyl)sulfonyl)-3-((3-methylphenyl)(methylthio)methyl)-1H-indole-2-carboxylate 120

Compound **74** (54 mg, 0.14 mmol) was dissolved in 10 mL of DMF in a 100 mL RBF. NaH (1.5 eq., 60%, 7.1 mg, 0.21 mmol) was added at 0 °C and stirred for 15 minutes before adding (*1S*)-(+)-10-camphorsulfonyl chloride. The reaction mixture was heated to 50 °C for 3 days after which all starting material had been consumed. After cooling, he reaction was quenched with sat. NH_4Cl and the product was extracted with EtOAc (3 x 15 mL). The organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude material was purified by column chromatography (5 - 20% EtOAc/Hex). The reaction yielded 41 mg of a yellow solid which we were unable to characterize.

8.4.2 (*1S*,*2R*,*5S*)-(-)-menthyl-chloroformate 121

In a three-necked RBF, (*1R*,2S,5R)-(-)-menthol (1.00g, 6.41 mmol) dissolved in CCl_4 (30 mL). A dropping funnel was charged with CCl_4 (10 mL) followed by triphosgene (0.5 eq., 780 mg, 2.63 mmol). Pyridine (1.2 eq., 0.60 mL, 7.45 mmol) was added to the dropping funnel which resulted

in a milky gas forming and the solution in the dropping funnel turning milky yellow in colour. This solution was added all at once at rt to the RBF and the reaction was heated to 55 °C for 3h. Once cooled, the reaction was quenched with ice water and extracted with DCM (3 x 30 mL). The organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuumto reveal compound **121** as a clear oil, 1.25 g, 5.76 mmol, 89% (R_f = 0.77, 20% EtOAc/Hex).



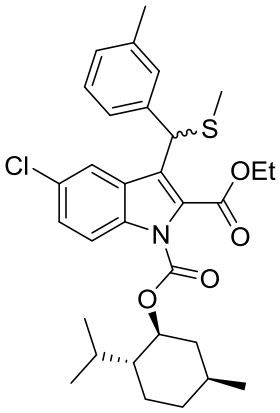
^{1}H NMR (300 MHz, $CDCl_3$) δ 4.73 (td, *J* = 11.0, 4.5 Hz, 1H, H_1), 2.18 – 1.86 (m, 2H, H_{3+7}), 1.77 – 1.37 (m, 4H), 1.21 – 0.97 (m, 3H), 0.96 – 0.88 (dd, *J* = 6.6, 3.8 Hz, 6H, H_9), 0.83 – 0.76 (d, *J* = 7.0 Hz, 3H, H_7). **^{13}C NMR (75 MHz, $CDCl_3$) δ** 150.09 (C_{11}), 84.12 (C_1), 46.98, 40.28, 33.96, 31.63, 26.45, 23.57, 21.97, 20.66, 16.42 (C_7).

8.4.3 2-ethyl 1-((*1S,2R,5S*)-2-isopropyl-5-methylcyclohexyl) 5-chloro-3-(*(R)*-(methylthio)(*m*-tolyl)methyl)-1H-indole-1,2-dicarboxylate **122A** and 2-ethyl 1-((*1S,2R,5S*)-2-isopropyl-5-methylcyclohexyl) 5-chloro-3-(*(S)*-(methylthio)(*m*-tolyl)methyl)-1H-indole-1,2-dicarboxylate **122B**

Compound **74** (50 mg, 0.13 mmol) was dissolved in 5 mL of DMF in a 100 mL RBF. NaH (2 eq., 7 mg, 0.27 mmol) was added at 0 °C, inducing a colour change from clear to bright yellow. This was stirred for 15 minutes before compound **121** was added and the reaction mixture was heated to 50 °C for 3h after which all starting material had been consumed. Once cooled, the reaction was quenched with sat. NH_4Cl (20 mL) and extracted with EtOAc (3 x 15 mL). The organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum. The crude material was purified by column chromatography (2-30% EtOAc/Hex) to reveal a racemic mixture of compounds **122A** and **122B** as a yellow oil, 59.8 mg, 0.108 mmol, 83% (R_f = 0.53, 20% EtOAc/Hex). We then attempted to crystallise the compound in a number of different solvent systems by adding the entire sample to a 10 mL RBF and adding 2 mL of the solvent/s and then heating this solution gently to evaporate off some of the solvent before allowing it to cool to rt. The results from these successive attempts are shown in Table 17.

Table 17

Solvent	Result
MeOH	Soluble at all temperatures
EtOH	Soluble at all temperatures
***i*-PrOH**	Soluble at all temperatures once dissolved
***i*-BuOH**	Oiled out
Hexane	Oiled out
EtOAc/Hex 3:1	Oiled out
MeOH/H_2O 19:1	Solution turned cloudy



^{1}H NMR (300 MHz, cdcl$_3$) δ 8.01 (d, *J* = 9.0 Hz, 1H, ArH), 7.80 (d, *J* = 2.1 Hz, 1H, ArH), 7.38 – 7.25 (m, 3H, ArH), 7.20 (td, *J* = 7.6, 3.4 Hz, 1H, ArH), 7.05 (d, *J* = 7.5 Hz, 1H, ArH), 5.43 (d, *J* = 6.5 Hz, 1H, CH), 4.96 (td, *J* = 11.0, 4.4 Hz, 1H, CH), 4.40 (q, *J* = 7.2 Hz, 2H, CH_2), 2.31 (s, 3H, CH_3), 2.28 – 2.16 (m, 1H, CH), 2.05 (d, *J* = 9.7 Hz, 3H, CH_3), 2.04 -1.96 (m, 1H, CH), 1.89 – 1.67 (m, 2H, CH_2), 1.65 – 1.48 (m, 2H, CH_2), 1.44 – 1.33 (m, 3H, CH_3), 1.27 – 1.08 (m, 2H, CH_2), 1.00 – 0.89 (m, 6H, 2 x CH_3),0.88 – 0.86 (m, 1H, CH), 0.90 – 0.77 (m, 3H, CH_3) . **^{13}C NMR (75 MHz, $CDCl_3$) δ** Note: Splitting of signals are observed in the ^{13}C NMR spectrum due to the presence of diastereomers, therefore there are more signals then carbon atoms. 162.59, 162.56, 150.19, 150.17, 138.79, 138.70, 138.32, 138.30, 134.72, 134.70, 129.73, 128.81, 128.76, 128.55, 128.54, 128.42, 128.37, 128.31, 126.81, 125.04, 125.01, 123.13, 122.98, 122.23, 116.64, 79.33, 62.26, 47.42, 47.40, 46.45, 46.37, 41.00, 34.18, 31.65, 26.29, 26.23, 23.44, 23.40, 22.09, 21.64, 20.98, 20.96, 16.35, 16.31, 16.01, 14.23. **HRMS:** Note: No parent molecular ion was present in the mass spectrum for this compound however the daughter ion was found. Calcd for $C_{30}H_{35}ClNO_4$ $[M - SCH_3 - H]^+$, 508.2255, found 508.2245.

8.4.4 Attempted synthesis of (*1S,2R,4S*)-1,7,7-trimethylbicyclo[2.2.1]heptan-2-yl 5-chloro-1H-indole-2-carboxylate 123

(*1S*,3R,4S)-(-)-borneol ethyl 5-chloroindole-2-carboxylate (40 mg, 0.18 mmol) was dissolved in THF and (*1S*,3R,4S)-(-)-borneol (10 eq., 246 mg, 1.79 mmol) was added along with K_3PO_4 (3 eq., 115 mg, 0.54 mmol). The reaction was refluxed at 100 °C for 18h, however TLC analysis showed that no product had formed. Starting material was recovered for future reactions.

8.4.5 Attempted synthesis of *(S)*-(1-methylpyrrolidin-2-yl)methanol 124

Formic acid (98%, 18.4 mL) was added to a 100 mL RBF followed by acetic anhydride (50 eq., 6.0 mL, 0.50 mmol). This was stirred for 2h at rt before L-proline (1 g, 0.01 mmol) was added at 0 °C. The reaction was stirred at rt overnight before quenching with distilled water (20 mL) and this was concentrated under vacuum to reveal a yellow oil which was used directly in the next reaction. The oil was dissolved in THF (30 mL) in a two-necked RBF fitted with a condenser and the flask was placed into an ice bath. $LiAlH_4$ (4 eq., 1.64 g, 0.04 mmol) was added portion wise and once fizzing had stopped, the reaction was heated to 60 °C for 2 days. Once cooled, the reaction was quenched with ice water (10 mL) and basified with 1M NaOH. A white precipitate was formed which was filtered off and the filtrate was extracted with EtOAc (3 x 30 mL) and concentrated under vacuum to reveal 200 mg of a clear oil. ^{1}H and ^{13}C NMR analysis did not correspond to literature.

8.4.6 Synthesis of 5-chloro-3-((methylthio)(*m*-tolyl)methyl)-1H-indole-2-carboxylic acid 127

Compound **74** (265 mg, 0.71 mmol) was dissolved in 20 mL of EtOH in a two-necked RBF fitted with a condenser. 8 mL of distilled water was added followed by KOH (4 eq., 157 mg, 2.8 mmol), which resulted in the reaction mixture changing from clear to bright yellowin colour, and the reaction was refluxed at 85 °C for 4h. TLC analysis showed all starting material had been consumed. The reaction mixture was quenched with distilled water before acidifying with 4M HCl and extracted with EtOAc (3 x 20 mL). The organic layer was washed with brine, dried over $MgSO_4$ and filtered then concentrated under vacuum to reveal 248 mg of a yellow oil. No other purification was performed. The product was used crude in the next set of reactions.

2 x 100 mg of crude sample (0.29 mmol) was placed in two 20 mL RBFs. (-)-Ephedrine (0.5 eq., 83 mg, 0.15 mmol) and *(S)*-(-)- α,α-diphenyl-2-pyrrolidinemethanol (0.5 eq., 127 mg, 0.15 mmol) were each added to one of the RBFs. Apprximately 5 mL of the solvent system was added to the RBFs and the sample was heated for several minutes to attempt to dissolve the solids and then left to cool to rt. If the sample had dissolved and no crystals were seen after a few hours this was moved to the fridge. If the solvent system was unsuccessful, the solvent was removed under vacuum before trying the next solvent system. The results from these attempted crystallisations are shown below (Table 18).

Table 18

	Base used to form the diastereomeric salt	
Solvent	(-)-ephedrine	*(S)*-(-)- α,α-diphenyl-2-pyrrolidinemethanol
MeOH	Soluble at all temperatures	Soluble at all temperatures
EtOH	Soluble at all temperatures	Soluble at all temperatures
***i*-PrOH**	Soluble at all temperatures once dissolved	Insoluble even when heated
***i*-BuOH**	Soluble at all temperatures once dissolved	Insoluble even when heated
Hexane	Insoluble even when heated	Insoluble even when heated
EtOAc/Hex 3:1	Oiled out	Oiled out
MeOH/H_2O 19:1	Precipitate formed too rapidly	Precipitate formed too rapidly

References

(1) Hu, W.-S.; Hughes, S. H. HIV-1 Reverse Transcription. *Cold Spring Harb. Perspect. Med.* **2012**, *2* (10), a006882–a006882.

(2) Gallo, R. C.; Montagnier, L. The Discovery of HIV as the Cause of AIDS. *N. Engl. J. Med.* **2003**, *349* (24), 2283–2285.

(3) Sharp, P. M.; Hahn, B. H. The Evolution of HIV-1 and the Origin of AIDS. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2010**, *365* (1552), 2487–2494.

(4) De Clercq, E. Anti-HIV Drugs: 25 Compounds Approved within 25 Years after the Discovery of HIV. *Int. J. Antimicrob. Agents* **2009**, *33* (4), 307–320.

(5) Botes, L.; van den Heever, W. M. J.; Pretorius, G. H. J. The Structural Biology of HIV. *Med. Technol. SA* **2007**, *21* (1), 13–18.

(6) Gao, F.; Bailes, E.; Robertson, D. L.; Chen, Y.; Rodenburg, C. M.; Michael, S. F.; Cummins, L. B.; Arthur, L. O.; Peeters, M.; Shaw, G. M.; Sharp, P. M.; Hahn, B. H. Origin of HIV-1 in the Chimpanzee Pan Troglodytes Troglodytes. *Nature* **1999**, *397* (6718), 436–441.

(7) WHO. Participant Manual. Module 1; Overview of HIV Infection. *Www.Who.Int/Diagnostics.../Pm_Module1.Pdf, Accessed on 02/06/2015* **2005**, 1–12.

(8) Reynolds, C.; de Koning, C. B.; Pelly, S. C.; van Otterlo, W. a. L.; Bode, M. L. In Search of a Treatment for HIV – Current Therapies and the Role of Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs). *Chem. Soc. Rev.* **2012**, *41* (13), 4657.

(9) Unaids. Fact Sheet 2015 | UNAIDS. *Unaids* **2015**, 1–8.

(10) Pribut, N. The Design and Synthesis of Novel HIV-1 Non-Nucleoside Reverse Transcriptase Inhibitors by. **2015**, No. March.

(11) Engelman, A.; Cherepanov, P. The Structural Biology of HIV-1: Mechanistic and Therapeutic Insights. *Nat. Rev. Microbiol.* **2012**, *10* (4), 279–290.

(12) Zhu, P.; Liu, J.; Bess, J.; Chertova, E.; Lifson, J. D.; Grisé, H.; Ofek, G. a; Taylor, K. a; Roux, K. H. Distribution and Three-Dimensional Structure of AIDS Virus Envelope Spikes. *Nature* **2006**, *441* (7095), 847–852.

(13) Craig B. Wilen , John C. Tilton, and R. W. D. Molecular Mechanisms of HIV Entry. In *Viral Molecular Machines*; 2012; pp 223–242.

(14) Muller, R. The Rational Design and Synthesis of Novel HIV Non-Nucleoside Reverse Transcriptase Inhibitors. **2013**, No. December.

(15) Martins, S.; Ramos, M. J.; Fernandes, P. A. The Current Status of the NNRTI Family of Antiretrovirals Used in the HAART Regime against HIV Infection. *Curr. Med. Chem.* **2008**, *15* (11), 1083–1095.

(16) Pawlica, P.; Berthoux, L. Cytoplasmic Dynein Promotes HIV-1 Uncoating. *Viruses* **2014**, *6* (11), 4195–4211.

(17) Jonckheere, H.; Anné, J.; Clercq, E. De. (RT) Process as Target for RT Inhibitors. 129–154.

(18) Mehellou, Y.; Clercq, E. De. Twenty-Six Years of Anti-HIV Drug Discovery : Where Do We Stand and Where Do We Go ? **2010**, 521–538.

(19) Perry, C. M.; Simpson, D. Tenofovir Disoproxil Fumarate In Chronic Hepatitis B. **2009**, *69* (16), 2245–2256.

(20) Müller, R.; Mulani, I.; Basson, A. E.; Pribut, N.; Hassam, M.; Morris, L.; Van Otterlo, W. A. L.; Pelly, S. C. Novel Indole Based NNRTIs with Improved Potency against Wild Type and Resistant HIV. *Bioorganic Med. Chem. Lett.* **2014**, *24* (18), 4376–4380.

(21) Hassam, M.; Basson, A. E.; Liotta, D. C.; Morris, L.; Van Otterlo, W. A. L.; Pelly, S. C. Novel Cyclopropyl-Indole Derivatives as HIV Non-Nucleoside Reverse Transcriptase Inhibitors. *ACS Med. Chem. Lett.* **2012**, *3* (6), 470–475.

(22) Lesbats, P.; Engelman, A. N.; Cherepanov, P. Retroviral DNA Integration. *Chem. Rev.* **2016**, *63* (4), 836–843.

(23) Craigie, R.; Bushman, F. D. HIV DNA Integration. *Cold Spring Harb. Perspect. Med.* **2012**, *2* (7), 1–18.

(24) Wong, E.; Trustman, N.; Yalong, A. HIV Pharmacotherapy. *J. Am. Acad. Physician Assist.* **2016**, *29* (2), 36–40.

(25) Karn, J.; Stoltzfus, C. M. Transcriptional and Posttranscriptional Regulation of HIV-1 Gene Expression. *Cold Spring Harb. Perspect. Med.* **2012**, *2* (2), 1–17.

(26) Freed, E. O. HIV-1 Assembly, Release and Maturation. *Nat Rev Microbiol* **2015**, *13* (8), 484–496.

(27) Sundquist, W. I.; Kräusslich, H. G. HIV-1 Assembly, Budding, and Maturation. *Cold Spring Harb. Perspect. Med.* **2012**, *2* (7), 1–24.

(28) Neil, S. J. D.; Zang, T.; Bieniasz, P. D. Tetherin Inhibits Retrovirus Release and Is Antagonized by HIV-1 Vpu. *Nature* **2008**, *451* (7177), 425–430.

(29) Arien, K. K.; Abraha, A.; Quinones-Mateu, M. E.; Kestens, L.; Vanham, G.; Arts, E. J.; Ariën, K. K.; Abraha, A.; Quiñones-Mateu, M. E.; Kestens, L.; Vanham, G.; Arts, E. J. The Replicative Fitness of Primary Human Immunodeficiency Virus Type 1 (HIV-1) Group M, HIV-1 Group O, and HIV-2 Isolates. *Journal of Virology*. 2005, pp 8979–8990.

(30) Treatmentcare, H. I. V. Hiv Treatment and Care. **2015**, No. November 2015.

(31) Li, D.; Zhan, P.; De Clercq, E.; Liu, X. Strategies for the Design of HIV-1 Non-Nucleoside Reverse Transcriptase Inhibitors: Lessons from the Development of Seven Representative Paradigms. *J. Med. Chem.* **2012**, *55* (8), 3595–3613.

(32) Hopkins, A. L.; Ren, J.; Milton, J.; Hazen, R. J.; Chan, J. H.; Stuart, D. I.; Stammers, D. K. Design of Non-Nucleoside Inhibitors of HIV-1 Reverse Transcriptase with Improved Drug Resistance Properties. 1. *J. Med. Chem.* **2004**, *47* (24), 5912–5922.

(33) Ren, J.; Esnouf, R.; Garman, E.; Somers, D.; Ross, C.; Kirby, I.; Keeling, J.; Darby, G.; Jones, Y.; Stuart, D. High Resolution Structures of HIV-1 RT from Four RT-Inhibitor Complexes. *Nat. Struct. Biol.* **1995**, *2* (4), 293–302.

(34) Zhan, P.; Chen, X.; Li, D.; Fang, Z.; De Clercq, E.; Liu, X. HIV-1 NNRTIs: Structural Diversity, Pharmacophore Similarity, and Impliations for Drug Design. *Med. Res. Rev.* **2013**, *33* (SUPPL. 1).

(35) Lansdon, E. B.; Brendza, K. M.; Hung, M.; Wang, R.; Mukund, S.; Jin, D.; Birkus, G.; Kutty, N.; Liu, X. Crystal Structures of HIV-1 Reverse Transcriptase with Etravirine (TMC125) and Rilpivirine (TMC278): Implications for Drug Design. *J. Med. Chem.* **2010**, *53* (10), 4295–4299.

(36) Rhee, S.-Y.; Gonzales, M. J.; Kantor, R.; Betts, B. J.; Ravela, J.; Shafer, R. W. Human Immunodeficiency Virus Reverse Transcriptase and Protease Sequence Database. *Nucleic*

Acids Res. **2003**, *31* (1), 298–303.

(37) Ren, J.; Milton, J.; Weaver, K. L.; Short, S. A.; Stuart, D. I.; Stammers, D. K. Structural Basis for the Resilience of Efavirenz (DMP-266) to Drug Resistance Mutations in HIV-1 Reverse Transcriptase. *Structure* **2000**, *8* (10), 1089–1094.

(38) Lipton, S. A. Neuronal Injury Associated with HIV-1: Approaches to Treatment. *Annu. Rev. Pharmacol. Toxicol.* **1998**, *38*, 159–177.

(39) Williams, T. M.; Ciccarone, T. M.; MacTough, S. C.; Rooney, C. S.; Balani, S. K.; Condra, J. H.; Emini, E. A.; Goldman, M. E.; Greenlee, W. J.; Kauffman, L. R.; et al. 5-Chloro-3-(Phenylsulfonyl)indole-2-Carboxamide: A Novel, Non-Nucleoside Inhibitor of HIV-1 Reverse Transcriptase. *J. Med. Chem.* **1993**, *36* (9), 1291–1294.

(40) Silvestri, R.; Artico, M. Indolyl Aryl Sulfones (IASs): Development of Highly Potent NNRTIs Active against Wt-HIV-1 and Clinically Relevant Drug Resistant Mutants. *Curr. Pharm. Des.* **2005**, *11* (29), 3779–3806.

(41) Zhao, Z.; Wolkenberg, S. E.; Lu, M.; Munshi, V.; Moyer, G.; Feng, M.; Carella, A. V.; Ecto, L. T.; Gabryelski, L. J.; Lai, M. T.; Prasad, S. G.; Yan, Y.; McGaughey, G. B.; Miller, M. D.; Lindsley, C. W.; Hartman, G. D.; Vacca, J. P.; Williams, T. M. Novel Indole-3-Sulfonamides as Potent HIV Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs). *Bioorganic Med. Chem. Lett.* **2008**, *18* (2), 554–559.

(42) Castellino, S.; Groseclose, M. R.; Sigafoos, J.; Wagner, D.; de Serres, M.; Polli, J. W.; Romach, E.; Myer, J.; Hamilton, B. Central Nervous System Disposition and Metabolism of Fosdevirine (GSK2248761), a Non-Nucleoside Reverse Transcriptase Inhibitor: An LC-MS and Matrix-Assisted Laser Desorption/Ionization Imaging MS Investigation into Central Nervous System Toxicity. *Chem. Res. Toxicol.* **2013**, *26* (2), 241–251.

(43) Brigg, S.; Pribut, N.; Basson, A. E.; Avgenikos, M.; Venter, R.; Blackie, M. A.; van Otterlo, W. A. L.; Pelly, S. C. Novel Indole Sulfides as Potent HIV-1 NNRTIs. *Bioorg. Med. Chem. Lett.* **2016**, *4* (6), 1–5.

(44) Famiglini, V.; La Regina, G.; Coluccia, A.; Pelliccia, S.; Brancale, A.; Maga, G.; Crespan, E.; Badia, R.; Clotet, B.; Esté, J. A.; Cirilli, R.; Novellino, E.; Silvestri, R. New Indolylarylsulfones as Highly Potent and Broad Spectrum HIV-1 Non-Nucleoside Reverse Transcriptase

Inhibitors. *Eur. J. Med. Chem.* **2014**, *80*, 101–111.

(45) Romines, K. R.; Freeman, G. a; Schaller, L. T.; Cowan, J. R.; Gonzales, S. S.; Jeffrey, H.; Iii, C. W. A.; Stammers, D. K.; Hazen, R. J.; Ferris, R. G.; Steven, a; Chan, J. H.; Boone, L. R. Structure Activity Relationship Studies of Novel Benzophenones Leading to the Discovery of a Potent, Next Generation HIV Non-Nucleoside Reverse Transcriptase Inhibitor. *J. Med. Chem.* **2006**, 727–739.

(46) Strauss, M. J. Technical Review. **1979**, *18* (3).

(47) Rueping, M.; Nachtsheim, B. J. A Review of New Developments in the Friedel-Crafts Alkylation - From Green Chemistry to Asymmetric Catalysis. *Beilstein J. Org. Chem.* **2010**, *6*, 1–24.

(48) Ozturk, T.; Ertas, E.; Mert, O. Use of Lawesson ' S Reagent in Organic Syntheses. **2007**, 5210–5278.

(49) Nishio, T. Direct Conversion of Alcohols into Thiols. **1993**, 2–6.

(50) Mosmann, T. Rapid Colorimetric Assay for Cellular Growth and Survival: Application to Proliferation and Cytotoxicity Assays. *J. Immunol. Methods* **1983**, *65* (1–2), 55–63.

(51) Ballini, R.; Palmieri, A.; Petrini, M.; Torregiani, E.; Chimiche, S.; Agostino, V. S. Solventless Clay-Promoted Friedel − Crafts Reaction of Indoles with R-Amido Sulfones : Unexpected Synthesis of 3- (1-Arylsulfonylalkyl) Indoles. *Tetrahedron* **2006**, No. 4, 4567–4570.

(52) Palmieri, A.; Petrini, M. Simplified Synthesis of 3-(1-Arylsulfonylalkyl) Indoles and Their Reaction with Reformatsky Reagents. *J. Org. Chem.* **2007**, *72* (5), 1863–1866.

(53) Rizzo, J. R.; Alt, C. A.; Zhang, T. Y. An Expedient Synthesis of 3-Substituted Indoles via Reductive Alkylation with Ketones. *Tetrahedron Lett.* **2008**, *49* (48), 6747–6749.

(54) Cao, L.; Wang, D.; Jiang, G.; Zhou, Y. An Efficient Route to 2 , 3-Disubstituted Indoles via Reductive Alkylation Using H 2 as Reductant. *Tetrahedron Lett.* **2011**, *52* (22), 2837–2839.

(55) Srihari, P.; Bhunia, D. C.; Sreedhar, P.; Yadav, J. S. Iodine-Catalyzed Nucleophilic Substitution Reactions of Benzylic Alcohols. *Synlett* **2008**, No. 7, 1045–1049.

(56) Hoster, P. Separation of an Enantiomorph and Its Racemate by Sublimation. **1967**, *227* (June).

(57) The Chemistry of Indoles - Richard Sundberg - Google Books.

(58) Fodor, L. An Efficient Route for the Synthesis of 2-Arylthiazino [5 , 6- B] Indole Derivatives. **2007**, *63*, 4983–4989.

(59) Wang, Y.; Chen, A. M. Re V Iews Enantioenrichment by Crystallization Abstract : **2008**, *12* (2), 282–290.

(60) International, Y. Diastereomeric Crystallisation - the " Classical " Chiral Technology.

(61) Lorenz, H.; Perlberg, A.; Sapoundjiev, D.; Elsner, M. P.; Seidel-morgenstern, A. Crystallization of Enantiomers. **2006**, *45*, 863–873.

(62) Collet, Â. Resolution of Racemates : Did You Say [a]Classical[o] ? **1998**, 3239–3241.

(63) Oppolzer, W. Camphor Derivatives as Chiral Auxiliaries in Asymmetric Synthesis. *Tetrahedron* **1987**, *43* (9), 1969–2004.

(64) Isherwood, M. L.; Guzzo, P. R.; Henderson, A. J.; Hsia, M. M.; Kaur, J.; Nacro, K.; Narreddula, V. R.; Panduga, S.; Pathak, R.; Shimpukade, B.; Tan, V.; Xiang, K.; Qiang, Z.; Ghosh, A. An Efficient Synthesis of (7S,10R)-2-Bromo-5,6,7,8,9,10-Hexahydro-7,10-Epiminocyclohepta[b]indole: Application in the Preparation and Structural Confirmation of a Potent 5-HT6 Antagonist. *Tetrahedron: Asymmetry* **2012**, *23* (22–23), 1522–1527.

(65) Hajra, S.; Bhowmick, M.; Maji, B.; Sinha, D. Design and Synthesis of Chiral N-Chloroimidodicarbonates : Application to Asymmetric Chlorination of Silyl Enol Ethers. **2007**, No. 2, 4872–4876.

(66) Leete, E.; Bjorklund, J. A.; Couladis, M. M.; Kim, S. H. Late Intermediates in the Biosynthesis of Cocaine - 4-(1-Methyl-2-Pyrrolidinyl)-3-Oxobutanoate and Methyl Ecgonine. *J. Am. Chem. Soc.* **1991**, *113* (24), 9286–9292.

(67) Smith, V. J.; Bogdan, D.; Caira, M. R.; Bogdan, M.; Bourne, S. A.; Fărcaş, S. I. Cyclodextrin Inclusion of Four Phenylurea Herbicides: Determination of Complex Stoichiometries and Stability Constants Using Solution 1H NMR Spectroscopy. *Supramol. Chem.* **2010**, *22* (3), 172–177.