

Union College

Union | Digital Works

Honors Theses

Student Work

6-2022

Blocking The Entry Of HIV Into Host Cells Through Co-receptor Inhibition

Garyfallia Ralli

Union College - Schenectady, NY

Follow this and additional works at: <https://digitalworks.union.edu/theses>



Part of the [Biochemistry Commons](#)

Recommended Citation

Ralli, Garyfallia, "Blocking The Entry Of HIV Into Host Cells Through Co-receptor Inhibition" (2022). *Honors Theses*. 2643.

<https://digitalworks.union.edu/theses/2643>

This Open Access is brought to you for free and open access by the Student Work at Union | Digital Works. It has been accepted for inclusion in Honors Theses by an authorized administrator of Union | Digital Works. For more information, please contact digitalworks@union.edu.

Blocking the entry of HIV into host cells through
co-receptor inhibition

By

Garyfallia Ralli

* * * * *

Submitted in partial fulfillment of the requirements for
Honors in the Biochemistry Program in the Department of
Chemistry.

UNION COLLEGE
June 2022

ABSTRACT

RALLI, GARYFALLIA. Blocking the entry of HIV by co-receptor inhibition.

ADVISOR: LEE PEDZISA

The goal of our project is to synthesize a dual CXCR4/CCR5 co-receptor inhibitor to block the entry of HIV into host cells. In the early phase of the HIV-1 replication cycle, HIV-1 binds to host cells through the CD4 protein present on the host cell surface. To infect the cell, HIV-1 requires further interactions that promote fusion of the viral and cellular membranes. This can occur through binding to the chemokine co-receptors such as CXCR4 and CCR5. We are choosing dual inhibition since under selective pressure of a CCR5 antagonist, CXCR4-using strains have been shown to predominate. Through dual inhibition, we want to minimize possible resistance development that could occur when inhibiting only one of the two co-receptors.

Using a computational screen, a compound predicted to bind to both CCR5 and CXCR4 was identified (Hit 1). We established a reaction scheme to synthesize an analog of this compound (target compound 4) through solid phase peptide synthesis. We confirmed the successful synthesis of this compound through LCMS and HPLC.

We also made progress towards synthesizing two analogs of target compound (4) by varying the last Fmoc-phenylalanine. The first analog contained Fmoc-4-fluoro-phenylalanine while the second analog contained Fmoc-4-methoxy-phenylalanine.

Lastly, we ran a docking screen to determine FDA approved compounds that could be repurposed as potential CCR5 inhibitors. In this, 40 compounds were identified to bind CCR5 with a higher affinity than FDA-approved CCR5 inhibitor, Maraviroc.

DEDICATION

To my parents, I couldn't have done this without you

ACKNOWLEDGMENTS

I would like to thank my thesis advisor, Lee Pedzisa, for supporting and guiding me not only throughout this project but also in planning my future after college. I would also like to acknowledge our lab team, Camille Parker and Sophie Danziger, for making our working environment very fun.

I would also like to thank all the professors that shaped my education and made me grow intellectually through the years.

TABLE OF CONTENTS

ABSTRACT	II
DEDICATION	III
ACKNOWLEDGEMENTS	IV
CHAPTER 1: INTRODUCTION	
1.1 AIDS AND HIV	1
1.2 HIV REPLICATION CYCLE	1
1.3 HIV INFECTION PROCESS	3
1.4 HIV TREATMENTS	6
1.5 HIV ENTRY INHIBITORS	9
1.6 MARAVIROC AND RESISTANCE MECHANISMS	10
1.7 REPURPOSING SCREENING	11
1.8 PROJECT OBJECTIVES	12
CHAPTER 2: RESULTS AND DISCUSSION	
2.1 SOLID PHASE SYNTHESIS: TARGET COMPOUND (4)	17
2.2 SOLID PHASE SYNTHESIS: FURTHER ANALOGS BY VARYING AMINO ACIDS	24
2.3 AUTO DOCK VINA VIRTUAL DOCKING OF POSITIVE CONTROL LIGAND MARAVIROC	25
2.4 AUTO DOCK VINA VIRTUAL SCREENING OF FDA APPROVED LIGANDS	27
2.5 HIGHEST CCR5 BINDING AFFINITY FDA APPROVED LIGANDS ANALYSIS	29

CHAPTER 3: EXPERIMENTAL

3.1 SOLID PHASE SYNTHESIS: TARGET COMPOUND (4)	38
3.2 SOLID PHASE SYNTHESIS: FURTHER ANALOGS BY VARYING AMINO ACIDS	42
3.3 REPURPOSING SCREENING:	
3.3.1 NECESSARY FILES FOR SCREENING	44
3.3.2 HARDWARE AND SOFTWARE	44
3.3.3 RECEPTOR PROTEIN PREPARATION AND GRID BOX CREATION	45
3.3.4 CREATING THE CONFIGURATION FILE (NOTEPAD OR OTHER TEXT EDITOR) ...	47
3.3.5 PREPARATION OF POSITIVE CONTROL LIGANDS (MARAVIROC)	47
3.3.6 AUTODOCK VINA VIRTUAL SCREENING OF ONE LIGAND	48
3.3.7 PREPARATION OF FDA TEST LIGANDS	48
3.3.8 AUTODOCK VINA VIRTUAL SCREENING OF FDA TEST LIGANDS	49
3.3.9 ANALYSIS OF VIRTUAL SCREENING RESULTS FOR FDA TEST LIGANDS	50
CHAPTER 4: CONCLUSIONS AND FUTURE WORK	53
APPENDIX	54

CHAPTER 1: INTRODUCTION

1.1 AIDS and HIV

Acquired immunodeficiency syndrome (AIDS) caused by the human immunodeficiency virus (HIV), is a disease of the human immune system that poses a great public health threat (Sharp and Hahn, 2011). AIDS, a chronic disease, is triggered by HIV infection and ultimately leads to the collapse of the immune system, signaled by CD4-T-lymphocyte cell depletion (Chatterjee et al., 2012). The course of the infection varies among individuals. Some individuals remain clinically and immunologically healthy for ten or more years after the infection while others progress to AIDS as early as one year after HIV infection (Chatterjee et al., 2012). Since its discovery, HIV has infected approximately 79.3 million people and has claimed around 36.3 million lives (UNAIDS, 2020). In 2020, 37.7 million people were living with HIV and 68 thousand lives were lost to an AIDS-related disease (UNAIDS, 2020). Therefore, a preventive or therapeutic treatment for AIDS is of big importance.

HIV is a retrovirus with a complex genome and a cone-shaped capsid core particle. Its genome is encoded by RNA and it is reversely transcribed to DNA by reverse transcriptase (RT) upon entering a host cell (Turner and Summers, 1999). The virus falls into two categories, HIV-type 1 (HIV-1) and HIV-type 2 (HIV-2), with HIV-1 being the main agent of AIDS. However, the viral structure is similar for both types of HIV (Fanales-Belasio et. al, 2010).

1.2 HIV replication cycle

During the early phase of the HIV replication cycle, HIV-1 binds to host cells through the CD4 protein (receptor protein) present on the host cell surface, which is

otherwise important for the immune system (Figure 1). To infect the cell, HIV-1 requires further interactions that promote fusion of the viral and cellular membranes. This can occur through binding to the chemokine co-receptors such as CXCR4 and CCR5 (co-receptor proteins). Upon entering the host cell, reverse transcription of the viral RNA is catalyzed in the cytosol by reverse transcriptase. The viral DNA is then transported to the nucleus; a complex process aided by various proteins. Finally, the viral DNA is integrated into the host genome by integrase. The late phase of the HIV-1 life cycle starts as the mRNA transcripts are transported out of the nucleus for translation. The structural proteins resulting from this process rearrange to form the virus, a process called maturation. (Turner and Summers, 1999).

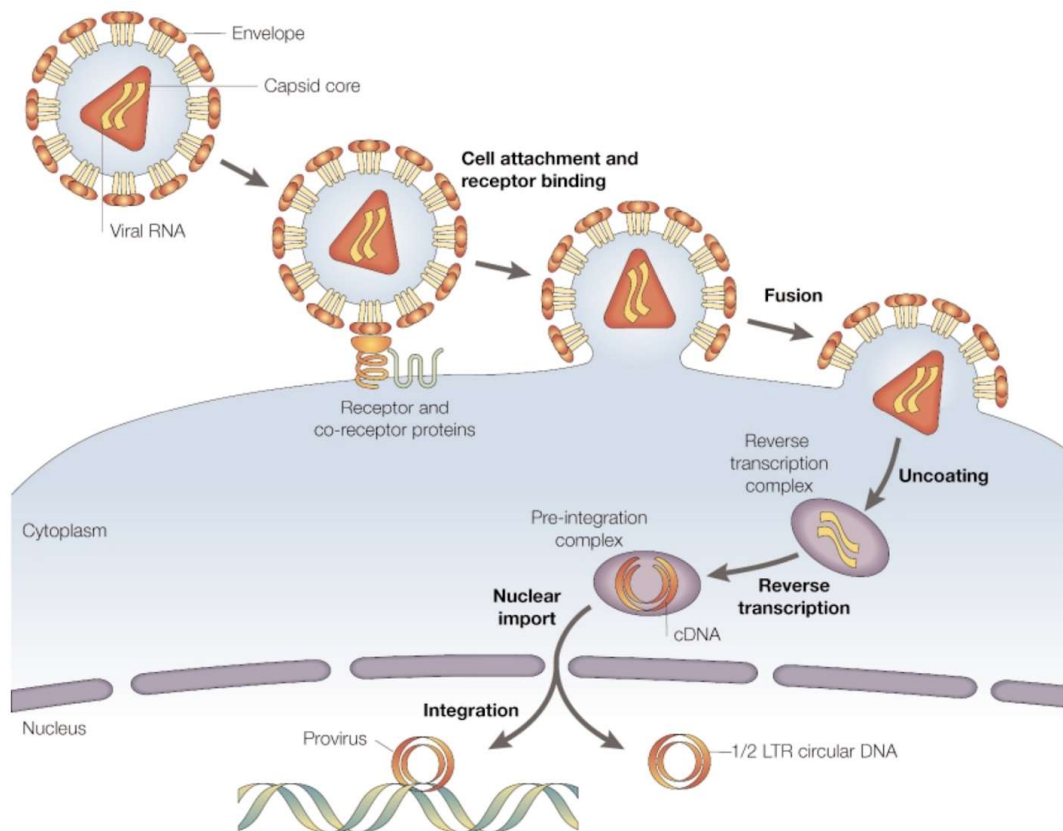


Figure 1. Summary of the early phase of the HIV replication cycle (Moore and Stevenson, 2000).

1.3 HIV infection process

The infection process is initiated when HIV binds to the CD4 protein on the surface of the host cell (Figure 2). The binding occurs through interactions between the virus's surface glycoprotein gp120 and CD4. The envelope glycoprotein is synthesized after the cleaving of gp160, an oligomeric precursor protein. The cleaving results in gp120, which is the surface subunit responsible for binding to receptor CD4 and co-receptors CXCR4/CCR5, and in gp41, responsible for membrane fusion (Feng et al., 1996). Gp120 includes an inner domain interacting with gp41 and an outer domain, which interacts with CD4/ CXCR4 & CCR5 for host cell infection (Kwong et al., 1998). Gp41 has an extracellular, transmembrane, and cytoplasmic domain with the extracellular domain including the fusion peptide (FP) (Garg et al., 2011).

Binding to CD4 causes a conformational change in the viral gp120, which exposes the chemokine receptor binding sites and, thus, promotes further interaction with the chemokine co-receptors (CXCR4 or CCR5). Consequently, the interactions with the chemokine co-receptors trigger a series of conformational changes that encourage the fusion of the viral and cellular membranes (Chen, 2019; Turner and Summers, 1999). In this process, the fusion peptide is necessary for successful infection. Due to its hydrophobic character, it can be embedded into the host cell membrane upon the triggering of fusion. Gp120 binding to CD4 causes the exposure of the co-receptor binding site and the displacement of the fusion peptide towards the membrane of the host cell (Blumenthal et al., 2012; Contarino et al., 2013; Eckert and Kim, 2001). A series of conformational changes in gp41 bring the viral and host cell membranes together (Blumenthal et al., 2012; Contarino et al., 2013; Eckert and Kim, 2001).

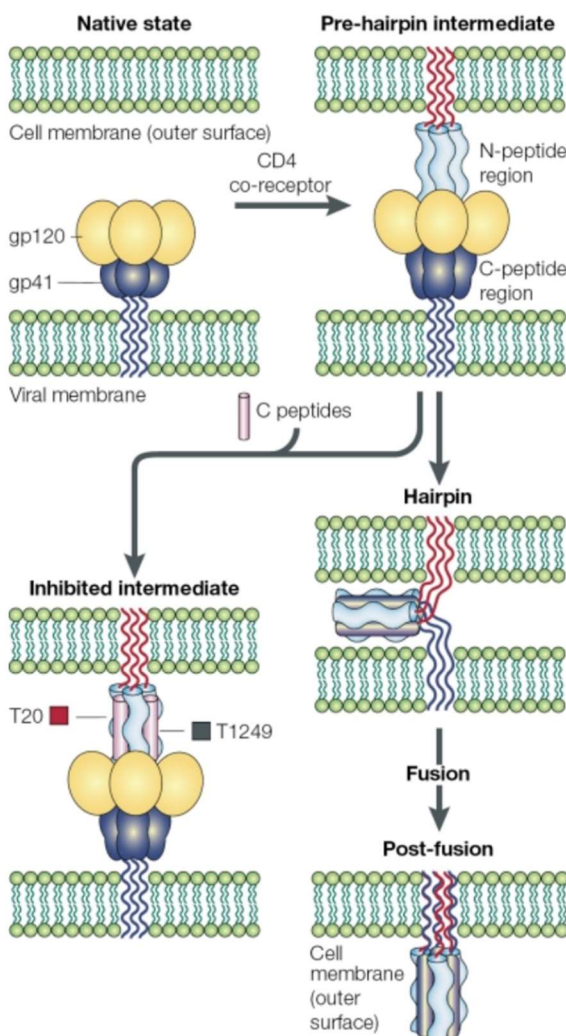


Figure 2. The fusion process for the trimeric HIV-1 envelope glycoprotein complex is depicted in schematic form. The sequential binding of the gp120 moieties (yellow ovals) to CD4 and a co-receptor on the cell membrane (not shown) drives conformational changes in the gp41 moieties (dark blue ovals). These changes cause the N-peptide region of gp41 (light blue cylinders) to translocate, and the fusion peptide (red lines) to insert into the cell membrane, forming the 'pre-hairpin intermediate'. Subsequent conformational changes in gp41 may be necessary to create the 'hairpin form' in which the viral and cellular membranes are brought into close enough proximity for membrane coalescence and fusion to occur (adapted from Moore and Stevenson, 2000).

Interestingly, some HIV-1 strains bind preferentially to the CCR5 receptor (M-tropic or R5-tropic) while others use CXCR4 (T-tropic or X4-tropic) (Figure 3). However, there are also some strains that can bind to both receptors, CCR5 and CXCR4 (dual tropic)

(Fanales-Belasio et. al, 2010). Although years after transmission the CXCR4 co-receptor can be used to facilitate infection too, during early infection R5-tropic viruses dominate (Connell et al., 2020) (Figure 4). R5-tropic HIV-1 viruses have a higher affinity for the CD4 receptor. In early infection CD4⁺ memory T-cells, which have a high expression of CCR5 on their cell surface, are the primary target of HIV-1. As the disease progresses, memory T-cells become depleted which allows targeting of naive CD4⁺ T-cells with predominantly CXCR4 co-receptors on their cell surface. In this way, it becomes clear that X4-tropic viruses are not the dominant viral phenotype. However, they can be present as a minority and rapidly multiply if R5-tropic viruses are inhibited. Maraviroc, a CCR5 antagonist, presents an example of this as X4-tropic viruses rapidly increase when the drug is present and are reduced to a minority again when the drug is discontinued (Connell et al., 2020).

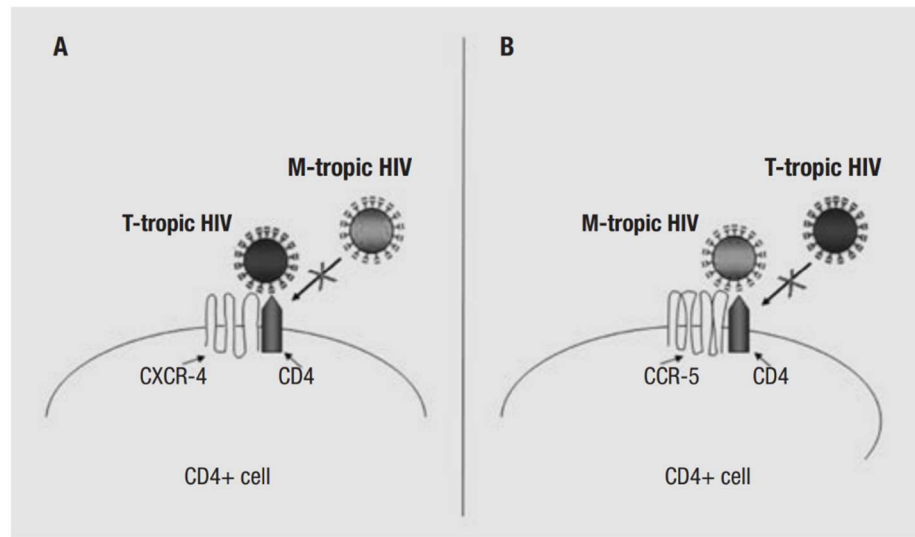


Figure 3. HIV tropism. In panel A a T-tropic (X4 tropic) HIV virus binds to a CXCR-4 coreceptor while a M-tropic (R5 tropic) HIV virus is unable to bind. In panel B a M-tropic HIV virus binds to a CCR-5 coreceptor while a T-tropic HIV virus is unable to bind (Fanales et al., 2010).

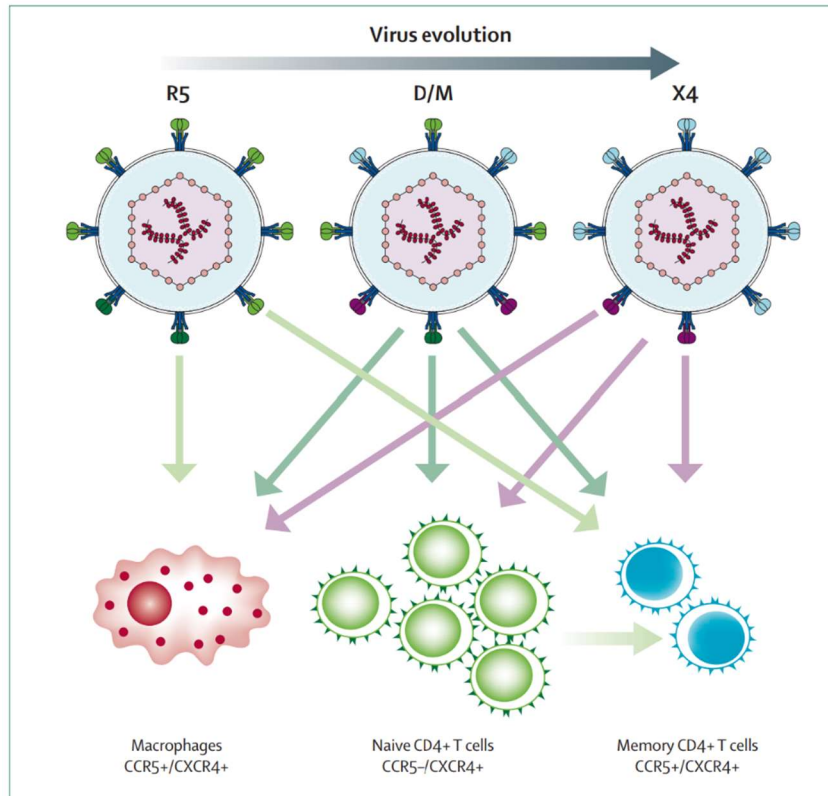


Figure 4. Cellular tropism Naive CD4+ T cells are CXCR4+ and CCR5–, and rarely infected by R5 virus. R5 HIV-1 strains are commonly transmitted and persist during the course of infection but might evolve to dual tropic (D) or mixed (M) virus populations, and to X4 phenotype at late stages of disease. Of treatment-naïve individuals, 12–19% have dual or mixed viral populations, and less than 1% are infected with X4 viruses. Of treatment-experienced individuals, 22–48% have dual or mixed viral populations, and 2–4% are infected with X4 viruses (adapted from Este and Talenti, 2007).

1.4 HIV Treatments

When it comes to the treatment of HIV, its variability allows it to overcome host immunity and drug and vaccine effects. This variability occurs through the error-prone mechanism of reverse transcriptase, the rapid viral replication, and the possibility of recombination of two or more different HIV viruses (Fanales-Belasio et. al, 2010). However, the error-prone mechanism of the enzyme reverse transcriptase has been recognized as the main cause of drug resistance development in HIV (Moyle, 1997).

In general, mutations can significantly reduce or increase viral fitness, but they can also turn out not to have any effects on the virus. Particularly, for drug resistance to occur due to a mutation, the coding region of the therapeutic agent's target must mutate in a way that the region's resulting protein remains active even in the presence of the drug. Mutations can also be transmitted from one person to another through transmission of mutated HIV (Najera et al., 1995). Importantly, mutations that result in resistance to multiple therapeutic agents can pose a significant problem. Specifically, this became apparent in the early days of HIV treatment when monotherapy, treatment with a single agent, was used. Monotherapy is more likely to result in resistant isolates. In this way, the importance of combination therapy for HIV-1 became apparent (Hirsch et al, 1998).

Combining antiretroviral drugs can have several potential benefits and drawbacks. Two drugs in combination may allow for reducing the dosage needed for each drug while still achieving the desired effects. On the other hand, the combination of two drugs can lead to side-effects not associated with any of the two drugs individually. Combination therapy can also influence drug resistance development. More specifically, the two drugs can be effective while using different mechanisms of action. In this way, resistance development may be less likely to occur and in case it occurs, it is likely to have a lesser impact than in the case of monotherapy. Lower dosages may also contribute to resistance development being less likely. However, a mutation that leads to resistance to both drugs can be detrimental to the treatment (Pirrone et al., 2011).

Highly active antiretroviral therapy (HAART) is an HIV treatment approach, which includes the administration of a combination of three or more antiretroviral drugs. In doing so, the main goal is to inhibit HIV replication through multiple different mechanisms so

that resistance to one of the administered agents can be counteracted by the other ones present (Cunningham et al., 1999; Kitahata et al., 1996; Rackal et al., 2011; Shafer et al., 1999).

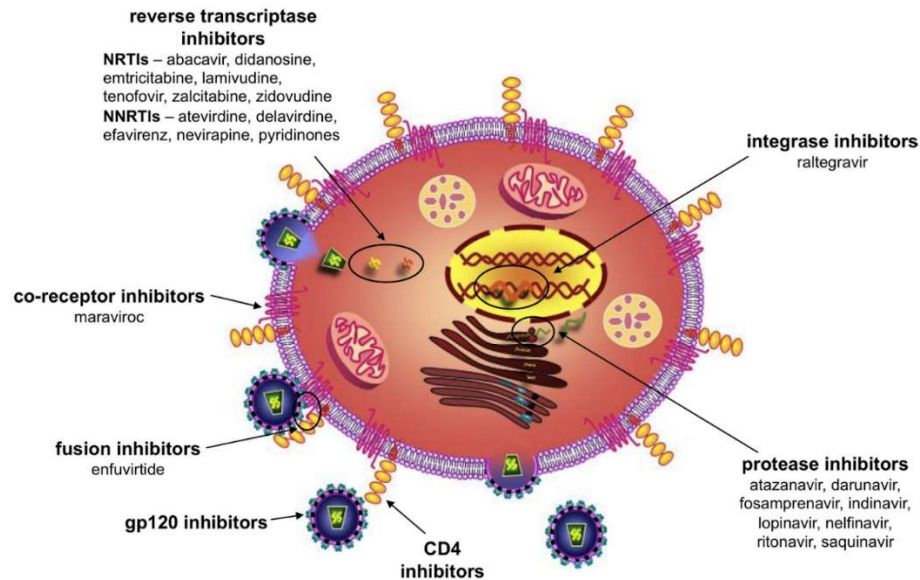


Figure 5. Summary of the HIV antiretroviral drug classes along with the drugs that fall under each class (Pirrone et al., 2011).

There are five classes of combination antiretroviral therapy drugs that impact different phases of the HIV-1 life cycle (Figure 5). The first class contains agents that block viral entry into the host cell (entry inhibitors) by interfering with chemokine receptors or membrane fusion. The second class terminates viral replication using nucleoside reverse transcription inhibitors (NRTIs). The third one, non-nucleoside reverse transcription inhibitors (NNRTIs), also terminates viral replication by binding to a different site than NRTIs. The agents of the fourth class, integrase strand transfer inhibitors (INSTIs) inhibit integration of the viral DNA into the host cell genome and the agents of the last class interfere with protease activity, which is responsible for assembling the new virus (Protease inhibitors, PIs) (Atta et al., 2019). The standard HAART combination for most patients

includes two NRTIs and one NNRTI or INSTI. (Günthard et al., 2016; Ford et al., 2018; Thompson, et al., 2012).

1.5 HIV entry inhibitors

Currently there are four FDA approved entry inhibitors. Enfuvirtide (2003) is a fusion inhibitor, which blocks HIV from entering CD4 cells (NIH, 2021). It does so by mimicking amino acids 127-162 on gp41 and interfering with membrane fusion of the HIV and host-cell membranes (Este and Telenti, 2007). Fostemsavir (2020) is an attachment inhibitor that binds to gp120 and prevents HIV from entering CD4 cells. Ibalizumab-uiyk (2018), a post-attachment inhibitor, binds to CD4 receptors and, thus, does not allow HIV to bind to coreceptors CCR5/CXCR4. Lastly, maraviroc (2007) is a CCR5 coreceptor inhibitor. It acts as an allosteric, non-competitive inhibitor and blocks the entry of HIV into the host cells (Este and Talenti, 2007; Latinovic et. al, 2019). A complete timeline of the medicines approved by the FDA for HIV-treatment can be seen in Figure 6. As becomes clear in this figure, when compared to the variety of protease and reverse transcriptase inhibitors, there are very few entry and integrase inhibitors. As discussed, HAART relies on attacking multiple pathways in the HIV cycle, therefore, there remains a need to diversify HIV drugs.

FDA Approval of HIV Medicines

1981: First AIDS cases are reported in the United States.				
'85-'89	1987 Zidovudine (NRTI)			
'90-'94	1991 Didanosine* (NRTI)	1992 Zalcitabine* (NRTI)	1994 Stavudine* (NRTI)	
'95-'99	1995 Lamivudine (NRTI) Saquinavir (PI)	1996 Indinavir* (PI) Nevirapine (NNRTI) Ritonavir (PI)	1997 Combivir (FDC) Delavirdine* (NNRTI) Nelfinavir* (PI)	1998 Abacavir (NRTI) Efavirenz (NNRTI)
'00-'04	2000 Didanosine EC* (NRTI) Kaletra (FDC) Trizivir (FDC)	2001 Tenofovir DF (NRTI)	2003 Atazanavir (PI) Emtricitabine (NRTI) Enfuvirtide (FI) Fosamprenavir (PI)	2004 Epzicom (FDC) Truvada (FDC)
'05-'09	2005 Tipranavir (PI)	2006 Atripla (FDC) Darunavir (PI)	2007 Maraviroc (CA) Raltegravir (INSTI)	2008 Etravirine (NNRTI)
'10-'14	2011 Complera (FDC) Nevirapine XR (NNRTI) Rilpivirine (NNRTI)	2012 Stribild (FDC)	2013 Dolutegravir (INSTI)	2014 Cobicistat (PE) Elvitegravir* (INSTI) Triumeq (FDC)
'15-'19	2015 Evotaz (FDC) Genvoya (FDC) Prezcobix (FDC)	2016 Descovy (FDC) Odefsey (FDC)	2017 Juluca (FDC)	2018 Biktarvy (FDC) Cimduo (FDC) Delstrigo (FDC) Doravirine (NNRTI) Ibalizumab-uiyk (PAI) Symfi (FDC) Symfi Lo (FDC) Symtuza (FDC) Temixys (FDC)
'20-'24	2020 Fostemsavir (AI)	2021 Cabenuva (FDC) Cabotegravir (INSTI)		

Drug Class Abbreviations:

AI: Attachment Inhibitor; CA: CCR5 Antagonist; FDC: Fixed-Dose Combination; FI: Fusion Inhibitor; INSTI: Integrase Inhibitor; NNRTI: Non-Nucleoside Reverse Transcriptase Inhibitor; NRTI: Nucleoside Reverse Transcriptase Inhibitor; PE: Pharmacokinetic Enhancer; PI: Protease Inhibitor; PAI: Post-Attachment Inhibitor

*Note: Drugs with an asterisk are no longer available and/or are no longer recommended for use in the United States by the HHS HIV/AIDS medical practice guidelines. These drugs may still be used in fixed-dose combination formulations.



Figure 6. A timeline of the medicines approved by the FDA for HIV treatment (NIH, 2021).

1.6 Maraviroc and resistance mechanisms

Maraviroc is currently the only coreceptor inhibitor approved by the FDA. It acts by locking CCR5 into a conformation that HIV cannot bind and, thus, inhibits entry of R5 tropic HIV into host cells. However, maraviroc does not have any major effects in patients

infected with X4 and dual tropic HIV (Este and Talenti, 2007; Hardy et al., 2010; Latinovic et al., 2019; Sierra-Madero et al., 2010). One resistance mechanism to Maraviroc involves selection for a minority of HIV viruses present in the organism that utilize CXCR4 for host cell entry (Westby et al., 2006). As described above, Maraviroc therapy presents an example of how X4-tropic viruses can rapidly expand under selection pressure and are reduced to a minority again when the pressure is discontinued (Connell et al., 2020). Another resistance mechanism selects for mutations that alter coreceptor use from CCR5 to CXCR4 (Nedellec et al., 2011). Therefore, the importance of dual inhibition of both CXCR4 and CCR5 coreceptors has been emphasized in order to improve therapeutic strategies against HIV (Horuk, 2009; Princen et al., 2004).

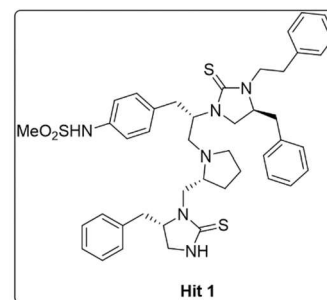
1.7 Repurposing screening

Another way of developing novel therapeutics is drug repurposing. Drug repurposing has the potential to utilize medications with established safety profiles to target new patient populations with conditions different than what the medication was originally meant to address (Cha et al., 2018). Drug repurposing can reduce the high cost and time requirements associated with drug discovery. Additionally, in drug discovery and development, there is approximately 45% chance of failure connected to safety and toxicity issues. All the above can be mitigated through drug repurposing. This process also offers significant advantages to patients as they are able to gain access to novel treatments with well analyzed safety profiles faster (Asburn and Thor, 2004). An example of a repurposed drug is azidothymidine, which was originally designed as a chemotherapy agent but developed to be an HIV medication instead (Volberding et al., 1990).

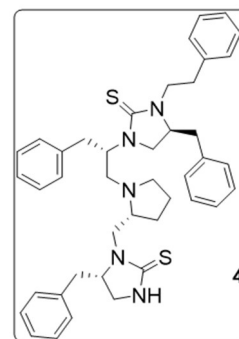
Repurposing approaches can be divided into two categories; experimental screening approaches and in silico approaches. Molecular docking is an important in silico method in predicting how chemical compounds can interact with a target. Therefore, this method has proven to be useful in drug discovery and drug development. More specifically, molecular docking utilizes the 3D structure of a compound of interest and predicts its orientation when interacting with a biological target, such as a protein. Then, the complementarity between the two is scored (Kitchen et al., 2004). In this way, based on the docking score, we can determine which compounds are predicted to have a high binding affinity for the protein of interest. Additionally, we can analyze the favorable ligand-receptor interactions.

1.8 Project Objectives

Using a docking approach, Professor Lee Pedzisa identified a compound that is predicted to bind to both CCR5 and CXCR4 (Hit 1). We plan to synthesize analogs of Hit 1, starting with target compound (4), using solid phase synthesis.



After achieving a satisfying number of compounds, we will evaluate our inhibitors for biological activity if time permits. Specifically, we will evaluate how well the ligands bind to CCR5 and CXCR4, how specific the synthesized target compounds are to these co-receptors, whether the compounds inhibit the co-receptors and ultimately whether the compounds reduce viral entry into cells.



Additionally, we want to run a new docking screen to analyze how FDA approved compounds could potentially be repurposed as CXCR4 and CCR5 inhibitors.

Goals of the project summarized:

1. Synthesize target compound (4) for CCR5 & CXCR4 dual inhibition.
2. Synthesize analogs of the target compound.
3. Test for biological activity:
 - Evaluate ligand binding to CCR5 and CXCR4.
 - Evaluate the specificity of the synthesized target compounds.
4. Screen FDA approved compounds to be repurposed as potential CXCR4 and CCR5 inhibitors.

References

Ashburn T.T and Thor K.B. Drug repositioning: identifying and developing new uses for existing drugs. *Nat Rev Drug Discov*, 2004, 3, 673-683.

Atta Mohamed G.; De Seignux Sophie and Lucas Gregory M. Clinical Pharmacology in HIV Therapy. *Clin J Am Soc Nephrol*. 2019, 14, 435-444.

Blumenthal, Robert et al. HIV Entry and Envelope Glycoprotein-mediated Fusion. *Journal of Biological Chemistry*, 2012, 287, 40841-40849.

Cha Y et al. Drug repurposing from the perspective of pharmaceutical companies. *BJP*, 2018, 175, 168-180.

Chatterjee Animesh et al. Chemokines and Chemokine Receptors in susceptibility to HIV-1 infection and progression to AIDS. *Disease Markers*, 2012, 32, 143-151.

Chen, Brian. Molecular Mechanism of HIV-1 Entry. *Trends Microbiol.*, 2019, 10, 878-891.

Connel, Bridgette J. et al. Immune activation correlates with and predicts CXCR4 co-receptor tropism switch in HIV-1 infection. *Scientific Reports*, 2020, 10.

Cunningham WE et al. The effect of hospital experience on mortality among patients hospitalized with acquired immunodeficiency syndrome in California. *Am J Med*. 1999, 107, 137-43.

Eckert, D.M. and Kim, P.S. Mechanism of viral membrane fusion and its inhibition. *Annu. Rev. Biochem.*, 2001, 70, 777-810.

Este Jose A and Telenti Amalio. HIV entry inhibitors. *Lancet*. 2007, 370, 81-88.

Fanales-Belasio Emanuele; Raimondo Mariangela; Suligoi Barbara and Butto Stefano. HIV virology and pathogenic mechanisms of infection: a brief overview. *Ann. Ist. Super Sanita*. 2010, 46, 5-14.

HIVinfo.NIH.gov. FDA-Approved HIV Medicines. (October 24, 2021)
<https://hivinfo.nih.gov/understanding-hiv/fact-sheets/fda-approved-hiv-medicines>

Feng, Y. et al. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane G protein-coupled receptor. *Science*, 1996, 272, 872-877.

Ford N et al. Benefits and risks of rapid initiation of antiretroviral therapy. *AIDS*. 2018, 32, 17-23.

Garg H. Targeting HIV-1 gp41-induced fusion and pathogenesis for antiviral therapy. *Curr. Top. Med. Chem.*, 2011, 11, 2947-2958

Günthard HF et al. Antiretroviral Drugs for Treatment and Prevention of HIV Infection in Adults: 2016 Recommendations of the International Antiviral Society-USA Panel. *JAMA*, 2016, 316, 191-210.

Hardy WD et al. Two-year safety and virologic efficacy of maraviroc in treatment-experienced patients with CCR5-tropic HIV-1 infection: 96-week combined analysis of MOTIVATE 1 and 2. *J Acquir Immune Defic Syndr* 2010; 55, 558-564.

Hirsch M et al. Antiretroviral drug resistance testing in adults with HIV infection: implications for clinical management. *JAMA*, 1998, 279, 931-939.

Horuk R. Promiscuous drugs as therapeutics for chemokine receptors. *Expert Rev. Mol. Med.*, 2009, 11.

Kitahata MM et al. Physicians' experience with the acquired immunodeficiency syndrome as a factor in patients' survival. *N Engl J Med*. 1996, 334, 701-706.

Kitchen, D.B et al. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov*, 2004, 3, 935–949.

Kwong, P.D. et al. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*, 1998, 393, 648-659.

Latinovic et al. CCR5 Inhibitors and HIV-1 Infection. *J AIDS HIV Treat*. 2019, 1, 1-5.

Montaner, Julio S G et al. Association of highly active antiretroviral therapy coverage, population viral load, and yearly new HIV diagnoses in British Columbia, Canada: a population-based study. *Lancet*, 2010, 376, 532-9.

Moore, J. and Stevenson, M. New targets for inhibitors of HIV-1 replication. *Nat Rev Mol Cell Biol*, 2000, 1, 40-49.

Moyle, G. J. Current knowledge of HIV-1 reverse transcriptase mutations selected during nucleoside analogue therapy: the potential to use resistance data to guide clinical decisions. *J. Antimicrob. Chemother.* 1997, 40, 765-777.

Najera et al. Pol gene quasispecies of human immunodeficiency virus: mutations associated with drug resistance in virus from patients undergoing no drug therapy. *J. Virol*, 1995, 69, 23-31.

Nedellec R et al. Resistance to the CCR5 inhibitor 5P12-RANTES requires a difficult evolution from CCR5 to CXCR4 coreceptor use. *PLoS One*, 2011, 6.

Pirrone Vanessa; Thakkar Nina et al. Combinational Approaches to the Prevention and Treatment of HIV-1 Infection. *Antimicrobial Agents and Chemotherapy*, 2011, 55, 1831-1842.

Princen et al. Inhibition of human immunodeficiency virus replication by a dual CCR5/CXCR4 antagonist. *J. Virol.*, 2004, 78, 12996-13006.

Rackal JM et al. Provider training and experience for people living with HIV/AIDS. *Cochrane Database Syst Rev*. 2011, 6.

Salzwedel et al. A Conserved Tryptophan-Rich Motif in the Membrane-Proximal Region of the Human Immunodeficiency Virus Type 1 gp41 Ectodomain Is Important for Env-Mediated Fusion and Virus Infectivity. *Journal of Virology*, 1999, 73, 2469-2480.

Shafer RW and Vuitton DA. Highly active antiretroviral therapy (HAART) for the treatment of infection with human immunodeficiency virus type 1. *Biomed Pharmacother.* 1999, 53, 73-86.

Sharp, Paul M, and Beatrice H Hahn. Origins of HIV and the AIDS pandemic. *Cold Spring Harbor perspectives in medicine*, 2011, 1.

Sierra-Madero J et al. Efficacy and safety of maraviroc versus efavirenz, both with zidovudine/lamivudine: 96-week results from the MERIT study. *HIV Clin Trials* 2010, 11, 125-132.

Thompson MA et al. Antiretroviral treatment of adult HIV infection: 2012 recommendations of the International Antiviral Society-USA panel. *JAMA*, 2012, 308, 387-402.

Turner Brian G.; Summers Michael F. Structural Biology of HIV. J. Mol. Biol. 1999, 285, 1-32.

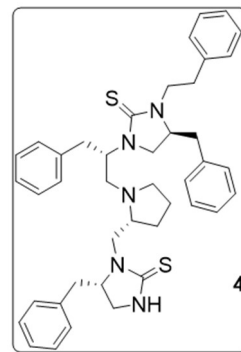
UNAIDS. Global HIV & AIDS statistics – 2020 fact sheet.

Westby M et al. Emergence of CXCR4-using human immunodeficiency virus type 1 (HIV-1) variants in a minority of HIV-1 infected patients following treatment with the CCR5 antagonist maraviroc is from a pretreatment CXCR4-using virus reservoir. J virol, 2006, 5, 4909-20.

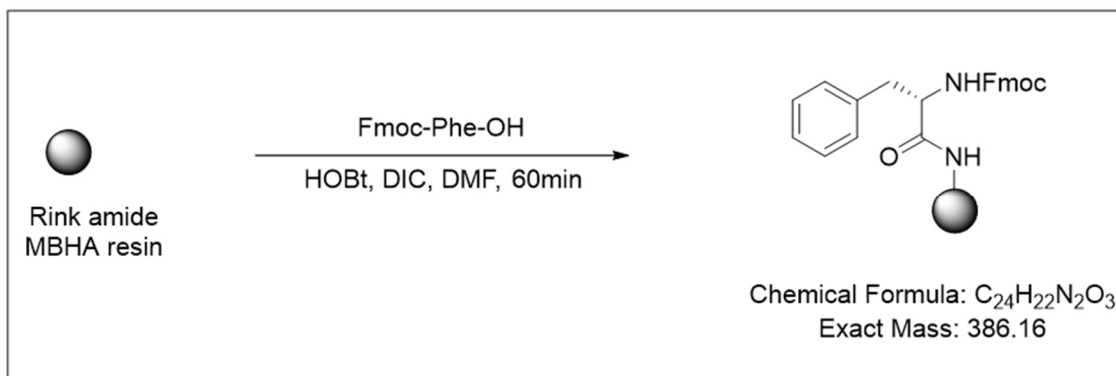
CHAPTER 2: RESULTS AND DISCUSSION

2.1 Solid phase synthesis: target compound (4)

Our first goal was to establish a robust and effective solid phase synthesis method of compound (4), which is an analog of the small polypeptide predicted to bind both CXCR4 and CCR5. To do this, we used rink amide MBHA resin beads as our basic building block where all the amino acids were added one after the other. We based our procedure on Nefzi and Santos (2006).

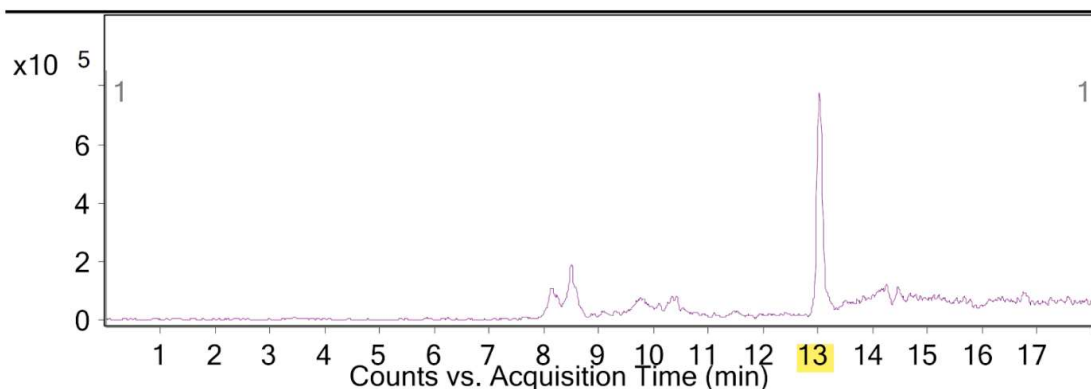


We started by neutralizing our rink amide MBHA resin beads with 10% DIEA/DMF overnight. In the original paper by Nefzi and Santos (2006), neutralization with 5% DIEA/DCM overnight was suggested. However, we determined that 5% DIEA/DMF overnight did not lead to bead neutralization based on a p-chloranil amine test (Vojtkovsky, 1995). Upon neutralization, the first amino acid, Fmoc-Phe-OH, was added as summarized in Reaction Scheme 1.



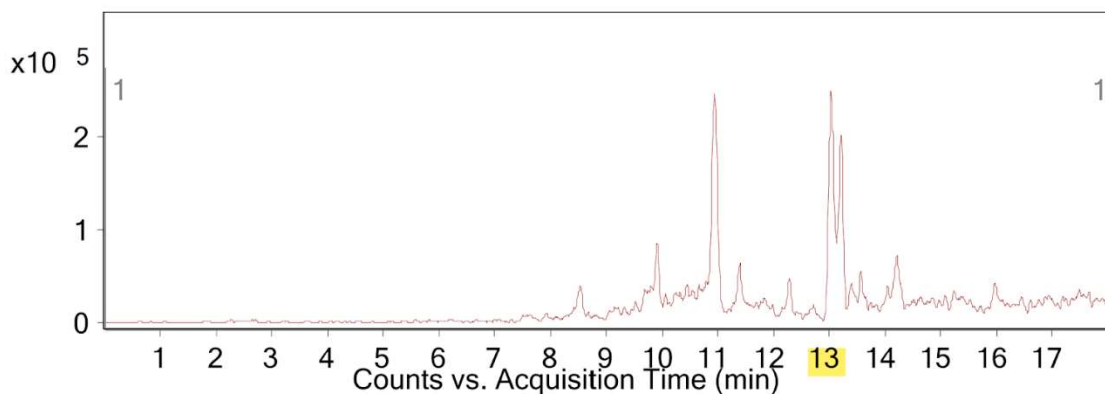
Reaction Scheme 1. Addition of the first amino acid, Fmoc-Phe-OH. The reaction conditions are indicated above and below the arrow. The chemical formula and exact mass are reported.

To confirm that the desired product was obtained from the solid phase synthesis, we collected and cleaved a small sample of our beads to analyze by LCMS and HPLC. Since our product is non-polar, a gradient of 5%-95% ACN with 0.1% formic acid was selected to increase retention time. In spectrum 1 below, a peak can be identified at approximately 13 minutes. The identified peak corresponds to the exact mass of our desired product plus 1 (Reaction Scheme 1). This indicated that the product picked up a hydrogen cation (+1 charge) resulting in a peak corresponding to 387. Thus, successful addition of the first amino acid was confirmed.



Spectrum 1. Spectrum obtained in the LCMS after addition of the first amino acid (Fmoc-Phe-OH, Reaction Scheme 1). The prominent peak at 13 min (highlighted yellow) corresponds to 387.

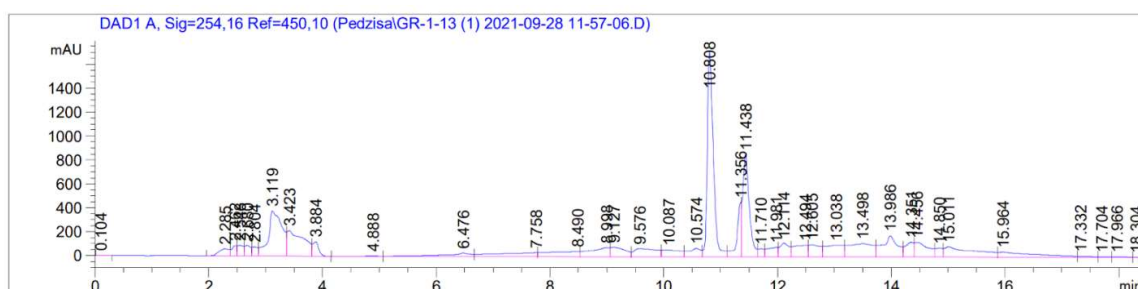
We also checked for a charge of +23, which can result from picking up a sodium cation. Based on spectrum 1, we would expect a peak at approximately the same retention time (13 minutes) to confidently say that some of our product also picked up sodium ions. As shown in spectrum 2, a peak at approximately 13 minutes can be identified in this case as well supporting, thus, our expectations.



Spectrum 2. Spectrum obtained in the LCMS after addition of the first amino acid (Fmoc-Phe-OH, Reaction Scheme 1). The peak at 13 min (highlighted yellow) corresponds to 409.

In both cases, the peak of interest in our spectra is not the only peak present. This could point to multiple ionizable molecules present in our solution.

The purity of our product can be evaluated upon obtaining HPLC spectra. As seen in Spectrum 3, there are two dominant peaks. Both of them are close to the retention time of 13min. Therefore, either one of them could represent the resulting product after addition of the first Fmoc-PheOH (Reaction Scheme 1).

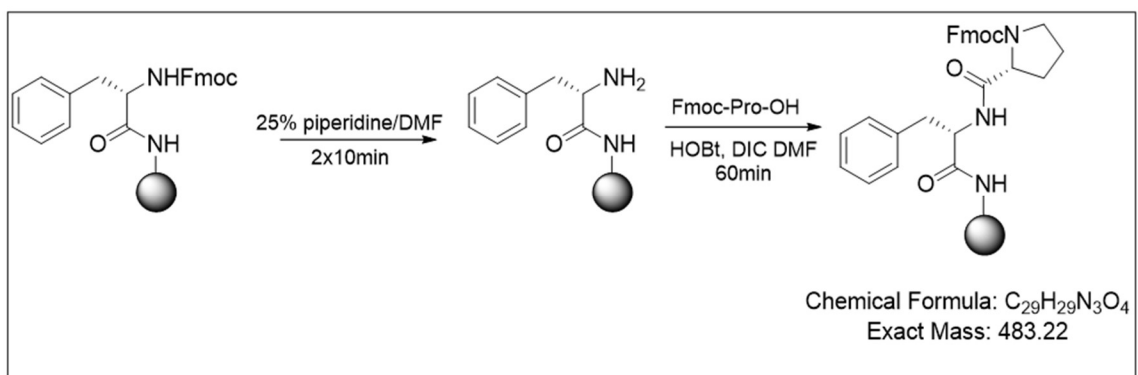


Spectrum 3. Spectrum obtained in the HPLC after addition of the first amino acid (Fmoc-Phe-OH, Reaction Scheme 1).

Additionally, the presence of two peaks points to possible impurities associated with our product synthesis. An explanation could be found in the way our sample for the

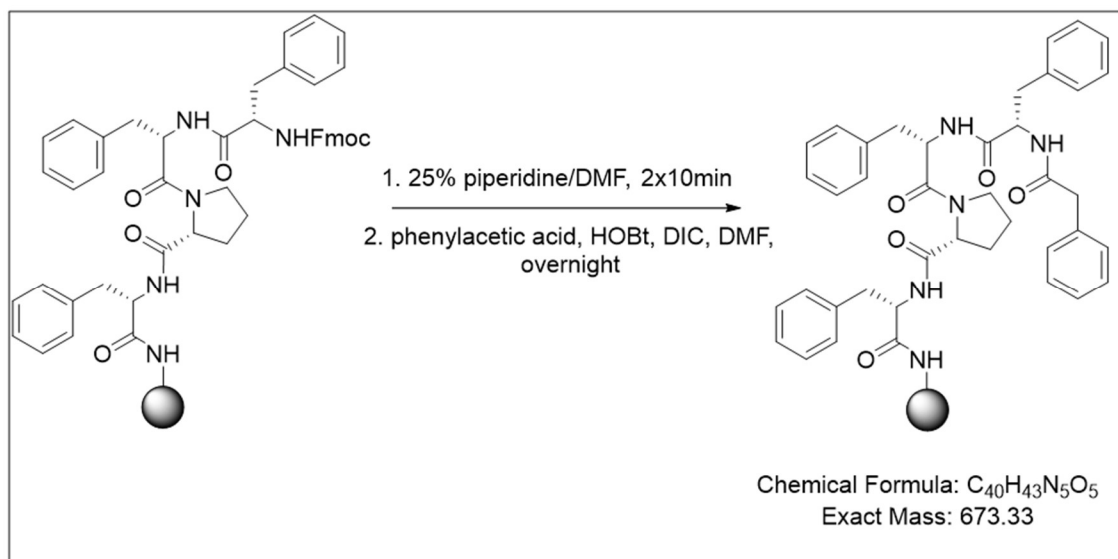
LCMS/HPLC was prepared. We cleaved our beads with TFA, separated the solution from the beads and left the solution open to air to evaporate and avoid corrosion of any instrument parts that could be caused by the TFA. After, we dissolved the resulting precipitate using DMF. However, the DMF used was not LCMS/HPLC grade, leading us to the conclusion that it could contain impurities. To test this possible explanation in the future, LCMS/HPLC grade DMF can be used. Additionally, we could determine the mass of the purity and predict what this impurity may be.

Upon completing the analysis of our first step in synthesis, the Fmoc group was removed using 25% piperidine/DMF before adding the next amino acid. Following the procedure described above and in the Experimental section (Chapter 3), we added three more amino acids (Fmoc-proline, Fmoc-PheOH and Fmoc-PheOH). The steps for adding these amino acids are summarized in Reaction Scheme 2 using the example of Fmoc-proline. For the addition of Fmoc-proline, the LCMS spectrum can be found in the Appendix. Specifically, a peak at approximately 13min corresponding to 484 (exact mass +1) can be identified confirming successful synthesis. Additionally, both the LCMS and HPLC spectrum for the addition of the second Fmoc-Phe-OH can be found in the Appendix. In the LCMS spectrum, a peak for both 631 (exact mass +1) and 653 (exact mass +23) can be found at approximately 22min. Similarly, in the HPLC, a prominent peak was identified at 21.5 min.



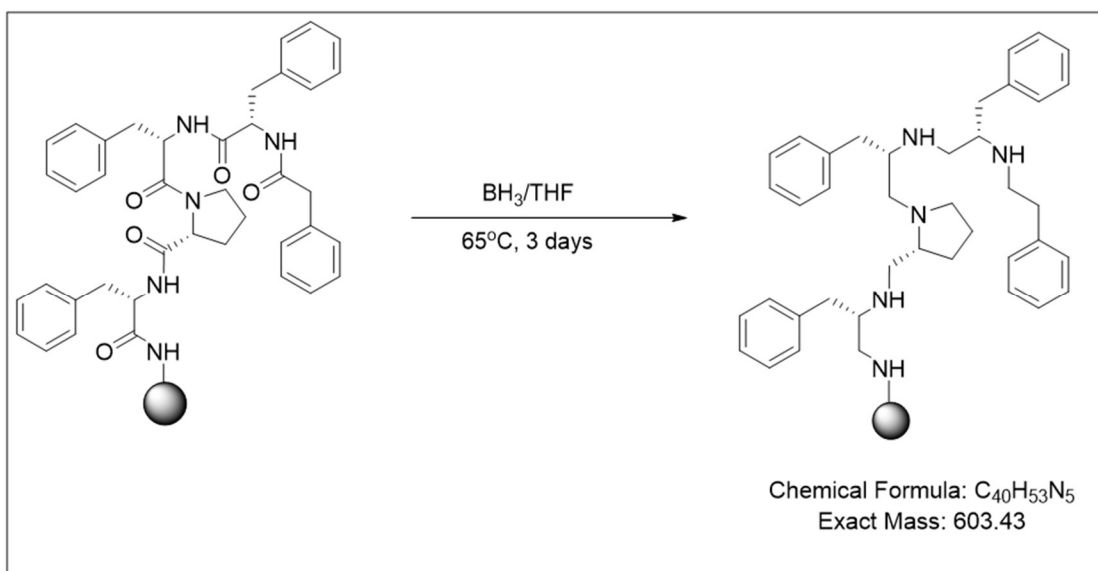
Reaction Scheme 2. Deprotection and addition of Fmoc-proline. The reaction conditions are indicated above and below the arrow. The chemical formula and exact mass are reported.

Upon addition of the last amino acid, our product was *N*-acetylated with phenylacetic acid as presented in Reaction Scheme 5. The addition of phenylacetic acid had a longer duration as it was run overnight.



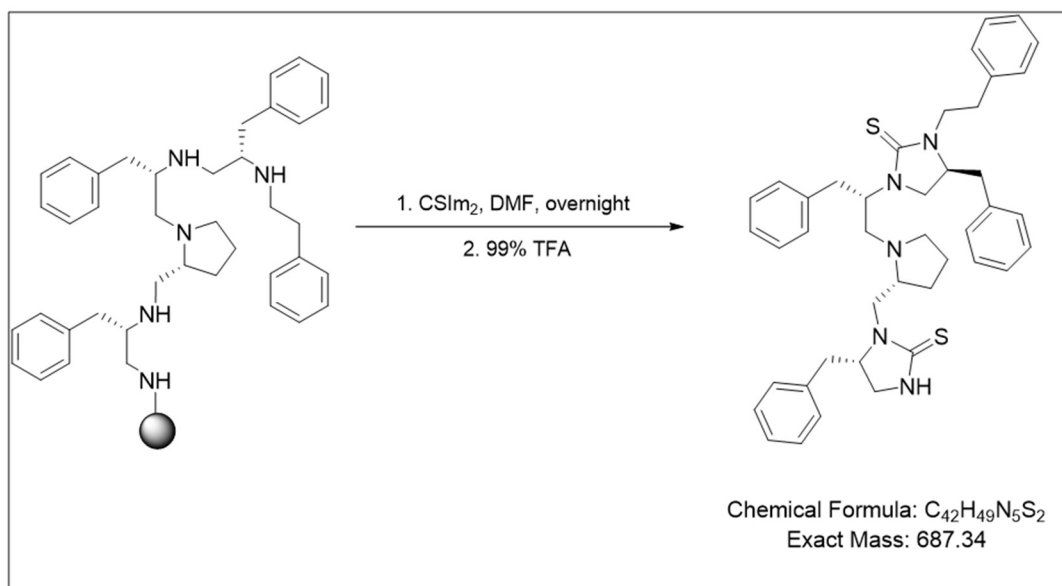
Reaction Scheme 3. *N*-acetylation with phenylacetic acid. The reaction conditions are indicated above and below the arrow. The chemical formula and exact mass are reported.

Then, the amide bonds were reduced using BH_3/THF (Reaction Scheme 6). This reaction proved to be challenging. Reaction time was very important. Specifically, when the beads were left in BH_3/THF for longer than 72 hours, we observed decomposition. This becomes apparent due to the presence of a white cloud in the otherwise yellow solution. Additionally, the reduction reaction is extremely sensitive to oxygen and moisture since BH_3 is unstable and very reactive and THF is anhydrous. For this reason, we left the glassware used for the reaction to dry overnight in the oven to reduce moisture. Moreover, during the preparation for the reaction and the reaction run, we had to make sure that BH_3 had minimal contact with oxygen by constantly purging the reaction with nitrogen and sealing the glassware fast. Another challenge was transferring the beads from the syringe used for solid phase synthesis to the round bottom flask for the reduction reaction. To do this, we used THF suspend the beads, transfer most of them and then wash the syringe. After completion of the reduction reaction, leftover BH_3 needed to be quenched with methanol and disproportioned with piperidine, as described in detail in the Experimental section.



Reaction Scheme 4. Amide reduction with BH₃/THF. The reaction conditions are indicated above and below the arrow. The chemical formula and exact mass are reported.

Then, the resin-bound polyamines were treated with thiocarbonyldiimidazole for cyclization (Reaction Scheme 7, step 1). Finally, they were cleaved from the resin using 99% TFA (Reaction Scheme 7, step 2). The resulting solution containing the polypeptide was diluted with acetonitrile and water and lyophilized to obtain a powder-like, white substance. The successful synthesis of the desired polypeptide was confirmed through LCMS (Appendix). A peak corresponding to 710 (exact mass +23) can be seen at approximately 19min and confirms the synthesis of target compound 4. Unfortunately, we were not able to obtain LCMS and HPLC spectra for all our intermediate steps. In the future, this is something that we would like to address in order to further support the reliability of our synthesis procedure.

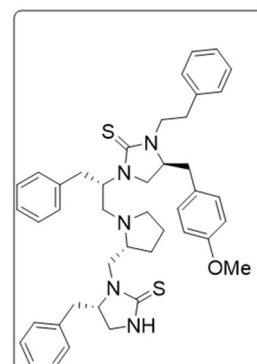
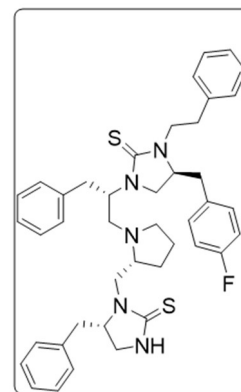


Reaction Scheme 5. Polyamines treated with thiocarbonyldiimidazole. The reaction conditions are indicated above and below the arrow. The chemical formula and exact mass are reported.

2.2 Solid phases synthesis: further analogs by varying amino acids

Upon successful synthesis of the aforementioned analog, we proceeded by planning the synthesis of two more analogs through varying the last Fmoc-phenylalanine. The goal of synthesizing analogs is to optimize CXCR4/CCR5 dual inhibition.

For the first analog, we decided to substitute the last Fmoc-Phe-OH with a 4-fluoro-Fmoc-phenylalanine. For our second analog, we chose to substitute with 4-methoxy-Fmoc-phenylalanine. Both the fluoro- and the methoxy-group could enhance chemokine receptor binding due to the polypeptide occupying more space and due to additional hydrogen bonding.



The procedure used to synthesize these analogs was the same as the one used to synthesize target compound 4, outlined above and in the Experimental section. We faced some challenges while synthesizing the two analogs. Both analogs decomposed during the reduction reaction with BH_3/THF . A possible explanation for this can be inferred from the observation that solvent evaporated during the 72h of the reaction. Specifically, after 72h, one of the round bottom flasks used had no solvent left in it while the other had very little solvent left. This led us to conclude that the septum was not tight enough and, thus, the solvent evaporated and escaped. During the 72h, the constantly decreasing solvent volume due to evaporation led to increasing concentration of BH_3 and this could have decomposed the beads. This incident further underlines the challenging nature of the reduction reaction for this synthesis.

2.3 AutoDock Vina Virtual Screening of positive control ligand Maraviroc

In addition to synthesizing a target from a previous computational screen, we wanted to screen FDA approved compounds against chemokine co-receptors CXCR4 and CCR5. The advantage of FDA approved compounds is that they have passed toxicity and stability tests and are ready for administration. In this way, we hoped to identify FDA approved compounds that could be used as potential inhibitors of these two co-receptor proteins.

We started by downloading the CCR5 co-receptor protein with bound Maraviroc from the Protein Data Bank (4MBS). Then, we prepared our protein using Discovery Studio and Autodock as described in Experimental 3.3.3 (Chapter 3). The resulting protein is presented in Figure 7 with its binding site indicated by the yellow sphere.

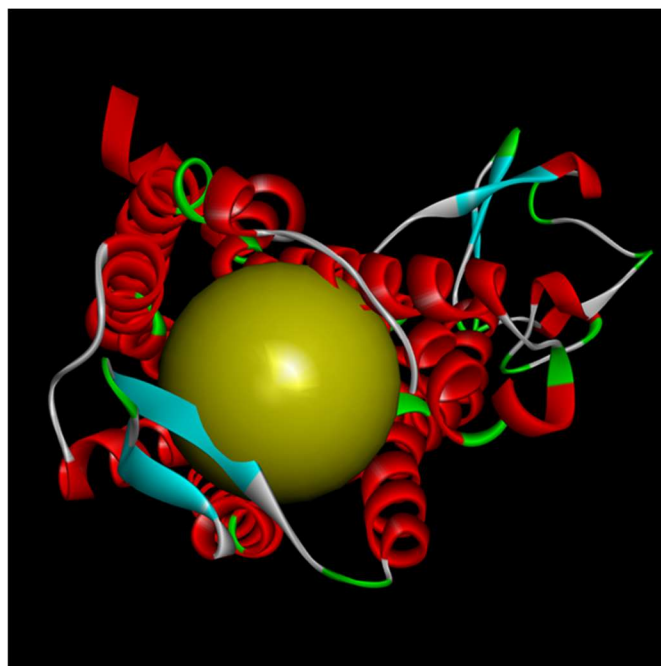


Figure 7. Prepared CCR5 co-receptor protein;
the binding site is indicated by the yellow sphere.

To establish a baseline for the binding affinities in our experiment, we docked a known CCR5 inhibitor, Maraviroc, against our prepared CCR5 co-receptor protein. The ligand was prepared according to Experimental 3.3.5. This yielded a binding affinity of -9.2 kcal/mol. To evaluate the effectiveness of our protein preparation method and to further analyze ligand binding, we compared the orientation of bound Maraviroc from our 4MBS file to the orientation of bound Maraviroc resulting from our docking experiment. In Figure 8, the Maraviroc configuration resulting from docking can be seen in yellow whereas the native Maraviroc configuration (4MBS) can be seen in gray. The upper domain of Maraviroc resulting from screening (yellow), the area indicated by the white line, shows a relatively good overlap with the native Maraviroc (gray). More differences in orientation can be observed in the lower parts of the molecules (indicated by the yellow arrow).

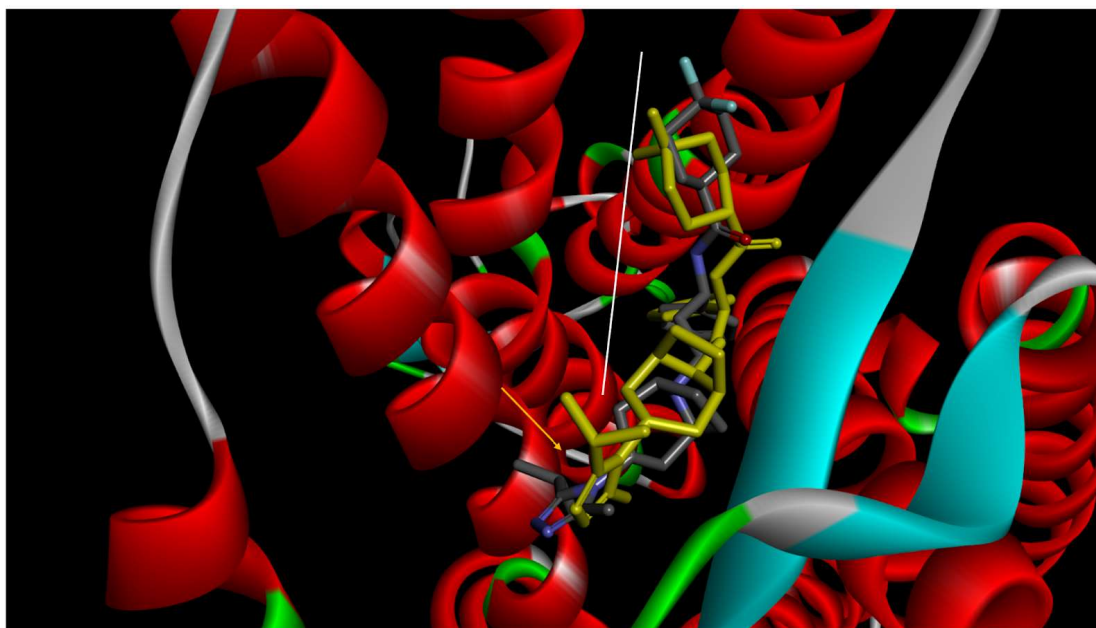


Figure 8. Comparison of 4MBS bound Maraviroc (native conformation - gray) with Maraviroc resulting from screening (yellow).

Although native and screen Maraviroc binding were not identical, we decided to move forward by using the protein preparation outlined in Experimental 3.3.3. Additionally, we decided to include additional positive controls in our FDA ligands screen. We included 16 CCR5 inhibitors with known IC_{50} s such as Vicriviroc. In addition to having known IC_{50} s, some of these compounds have undergone clinical trial indicating that they have efficacious binding affinities. In this way, we hoped to have more points of reference to compare the binding of our FDA ligands to.

2.4 AutoDock Vina Virtual Screening of FDA approved ligands

We proceeded by screening 1614 FDA approved compounds obtained from the Zinc database against our prepared CCR5 co-receptor protein in the Maraviroc binding site. The FDA ligands were prepared according to Experimental 3.3.7. The screen took

approximately a week to complete (Experimental 3.3.8), and the results were analyzed according to Experimental 3.3.9. Our screen yielded 40 compounds that were predicted to bind CCR5 with a higher or equal affinity to Maraviroc (Table 1).

Rank	Binding energy (kcal/mol)	Compound #
1	-11	406
2	-10.9	1276 - Adapalene
3	-10.8	416 - Naldemedine
4	-10	910 - Dihydroergotamine
5	-9.9	271 - Olaparib
6	-9.9	650 - Naftifine
7	-9.8	1264
8	-9.7	1277
9	-9.7	786
10	-9.6	1381
11	-9.6	1468
12	-9.6	839
13	-9.5	1535
14	-9.5	1575
15	-9.5	168
16	-9.5	404
17	-9.5	42
18	-9.5	467
19	-9.5	636
20	-9.5	695
21	-9.4	1017
22	-9.4	1498
23	-9.4	444
24	-9.3	1385
25	-9.3	1591
26	-9.3	165
27	-9.3	273
28	-9.3	307
29	-9.3	382
30	-9.3	539
31	-9.3	579
32	-9.3	689
33	-9.2	1359
34	-9.2	1564
35	-9.2	364
36	-9.2	593
37	-9.1	1058
38	-9.1	107
39	-9.1	131
40	-9.1	144
41	-9.1	1626 - Maraviroc

Table 1. FDA approved compounds identified by virtual screen to bind CCR5 with a higher/equal affinity than/as FDA-approved CCR5 inhibitor, Maraviroc.

The compounds that ranked from the 2nd to the 6th place based on binding affinity were identified as Adapalene (-10.9 kcal/mol), Naldemidine (-10.8 kcal/mol), Dihydroergotamine (-10 kcal/mol), Olaparib (-9.9 kcal/mol) and Naftifine (-9.0 kcal/mol) accordingly. A challenge we faced while trying to identify the top ranked compounds was bonds being altered due to possible file corruption either during the protein preparation process or the run. Due to this, we were unable to identify compound 406 (first place ranking).

2.5 Highest CCR5 binding affinity FDA approved ligands analysis

Rank 2: Adapalene

The first FDA approved compound we identified was adapalene. The structure of adapalene can be found in Figure 8.

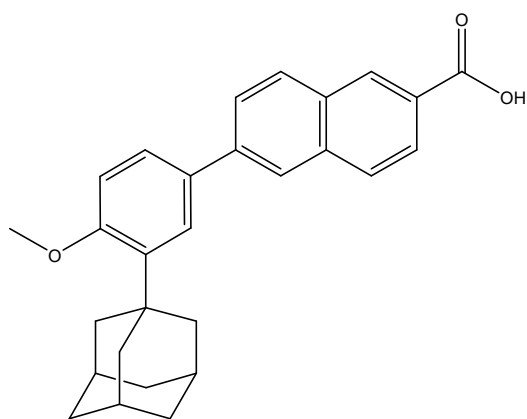


Figure 8. Adapalene structure.

Adapalene is a synthetic retinoid and a very effective acne treatment option (Cunliffe et al., 1997). Adapalene is better tolerated than other synthetic retinoids used to treat acne (Schaefer and Reichert, 1990). Specifically, it acts as an acne treatment by having an effect on cell proliferation and differentiation as well anti-inflammatory properties. It does so by

interacting with retinoic acid receptors (RAR α , RAR β , RAR γ) and with 9cis-retinoic acid receptors (RXRA α , RXRA β , RXRA γ) (Michel et al., 1998). A common brand currently on the market containing adapalene is the Differin Gel. Adapalene has not been tested in the context of HIV entry inhibition.

Our repurposing screening identified adapalene as one of the compounds with the highest CCR5 binding affinity (Table 1). This makes adapalene an interesting compound to further investigate in the context of chemokine co-receptor inhibition. In Figure 9A, adapalene is shown bound to co-receptor CCR5 in the same binding site used by Maraviroc. Panel B shows the interactions between adapalene and the amino acids present in the CCR5 binding site. All interactions between adapalene and CCR5 are hydrophobic.

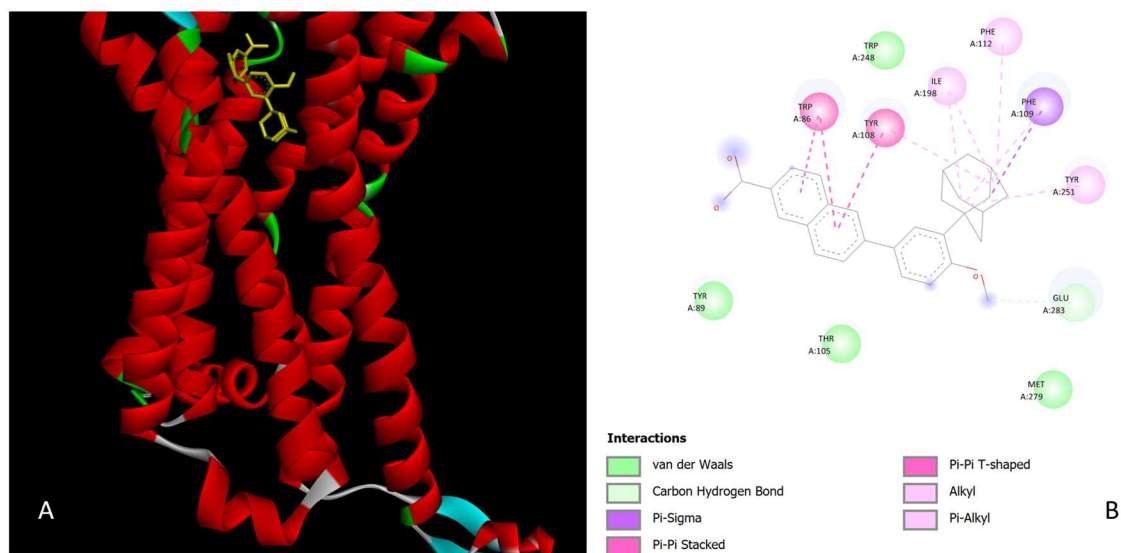


Figure 9. Adapalene bound to chemokine co-receptor CCR5. Panel A shows the orientation of adapalene in the CCR5 binding site. Panel B shows the interactions between adapalene and the amino acids present in the CCR5 binding site.

Rank 3: Naldemedine

Naldemedine was the FDA approved compound with the third highest affinity binding ranking (Table 1, Figure 10).

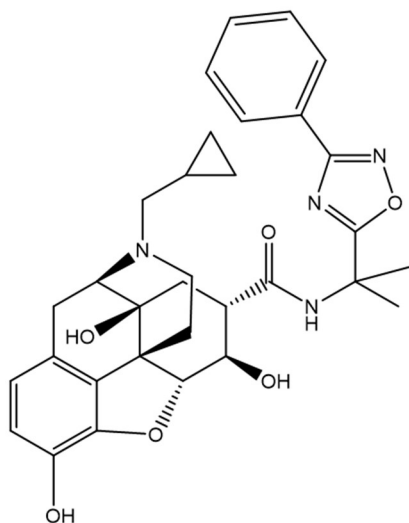


Figure 10. Structure of naldemedine.

Naldemedine (Symproic in Japan/US, Rizmoic in EU) is an μ -opioid receptor antagonist. It is used to increase bowel movement in opioid induced constipation, which is caused by opioid treatment for cancer pain or for chronic non-cancer pain (Blair, 2019). Naldemedine acts on opioid receptors in the gastrointestinal tract and it is generally well-tolerated. Its possible side effects are related to gastrointestinal function (Markham, 2017). Naldemedine has not been investigated in relation to HIV.

Our virtual screen determined naldemedine as a possible CCR5 inhibitor. In Figure 11, panel A shows naldemedine bound to CCR5. Naldemedine interacts with CCR5 through hydrophobic interactions and van der Waals interactions (fade green) (Figure 11B).

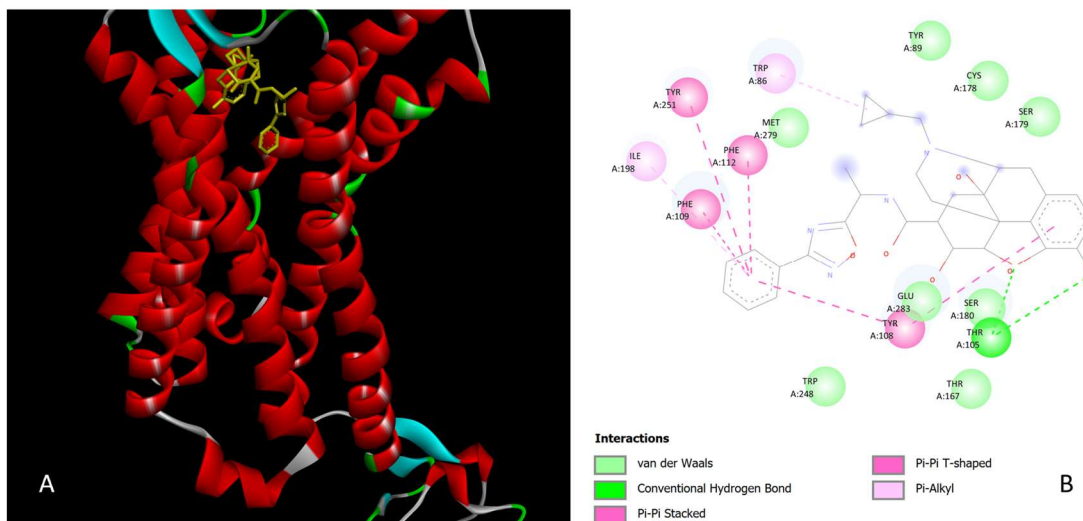


Figure 11. Naldemedine bound to chemokine co-receptor CCR5. Panel A shows the orientation of naldemedine in the CCR5 binding site. Panel B shows the interactions between naldemedine and the amino acids present in the CCR5 binding site.

Rank 4: Dihydroergotamine

Dihydroergotamine was the next FDA approved compound we identified (Table 1, Figure 12).

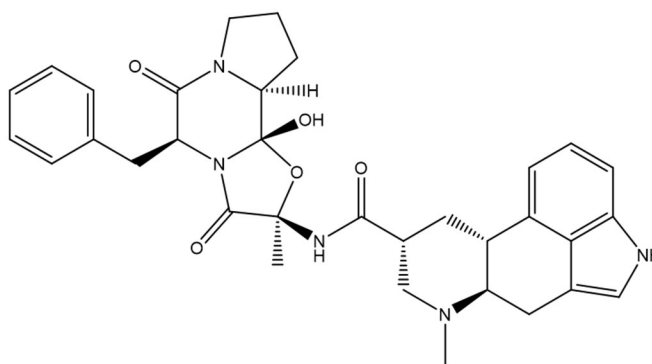


Figure 12. Structure of dihydroergotamine.

Dihydroergotamine belongs to the family of ergot alkaloids, a group with many pharmacological effects (Silberstein, 1997). Dihydroergotamine can be synthesized from ergotamine by reducing an unsaturated bond. Both ergotamine and dihydroergotamine are

used to treat migraines. They do so by acting as antagonists to receptors 5-HT1B, 5-HT1D and 5-HT1F (Silberstein and McCrory, 2002). However, both compounds also have adverse effects due to their interactions with other monoamine receptors. This can cause nausea, vomiting and other reactions (Silberstein and McCrory, 2002). In the context of HIV, it is advised to not co-administer dihydroergotamine with ritonavir, an HIV protease inhibitor, as this can lead to acute ergot toxicity (Mohamedi et al., 2021).

The results showed dihydroergotamine to be the compound with the fourth highest CCR5 binding affinity (Table 1). Figure 13A shows dihydroergotamine bound to CCR5 while Panel B presents the interactions between dihydroergotamine and CCR5 binding site amino acids. Dihydroergotamine interacts hydrophobically with the amino acids shown in pink. It also shows a pi-anion interaction with Glu282 (orange).

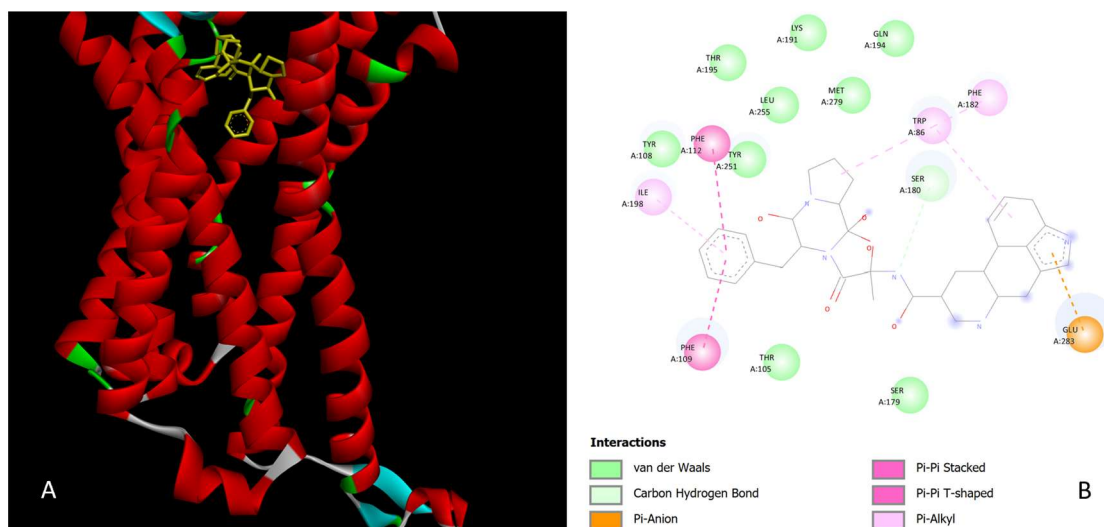


Figure 13. Dihydroergotamine bound to chemokine co-receptor CCR5. Panel A shows the orientation of dihydroergotamine in the CCR5 binding site. Panel B shows the interactions between dihydroergotamine and the amino acids present in the CCR5 binding site.

Rank 5: Olaparib

Olaparib is used as a treatment option in ovarian and other types of cancer. It is a poly-ADP-ribose polymerase inhibitor (Figure 14). Specifically, it acts at the catalytic site of enzymes PARP1 and PARP2 as an NAD⁺ competitive inhibitor (Robson et al., 2017). These enzymes are important in DNA repair. By inhibiting them, olaparib inhibits DNA repair and induces cell death (Bochum et al., 2018). Olaparib treatment had not been reported in the context of HIV.

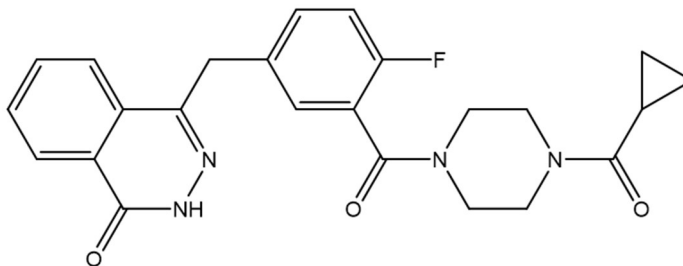


Figure 14. Structure of olaparib.

Our screen ranked olaparib in the fifth place for CCR5 binding affinity (Table 1). Olaparib interacts with CCR5 through hydrophobic interactions. Additionally, it shows pi-anion interactions with Glu283 (orange) (Figure 13). This interaction was previously observed in dihydroergotamine (Figure 15).

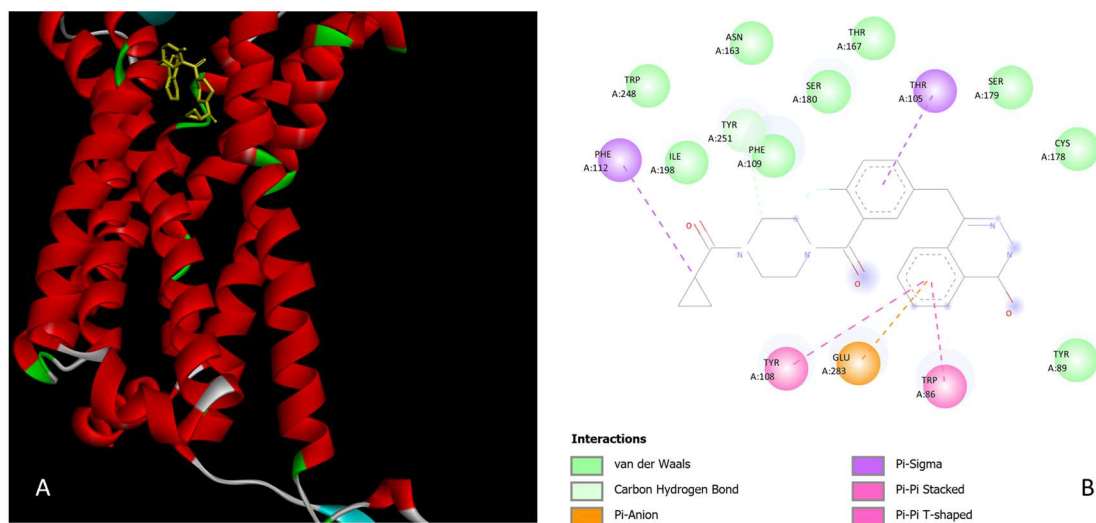


Figure 15. Olaparib bound to chemokine co-receptor CCR5. Panel A shows the orientation of olaparib in the CCR5 binding site. Panel B shows the interactions between olaparib and the amino acids present in the CCR5 binding site.

Rank 6: Naftifine

Naftifine is the last FDA approved ligand we looked further into (Table 1, Figure 16). Naftifine is an allylamine derivative that is used as an antifungal. It can be administered as a topical gel, which is well tolerated with topical irritation occurring in some cases (Monk and Brodgen, 2012; Mühlbacher, 1991). Naftifine works by inhibiting squalene epoxidase, which leads to inhibition of ergosterol synthesis (Monk and Brodgen, 2012). Naftifine has not been studied for HIV inhibition.

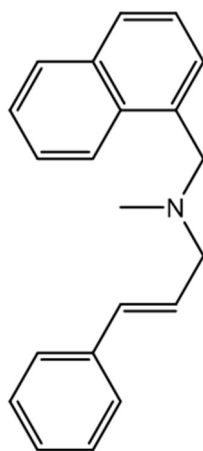


Figure 16. Structure of naftifine.

In Figure 17A, naftifine is shown bound to the CCR5 Maraviroc binding site. As shown in Panel B, naftifine interacts hydrophobically with the amino acids present in the binding site (pink).

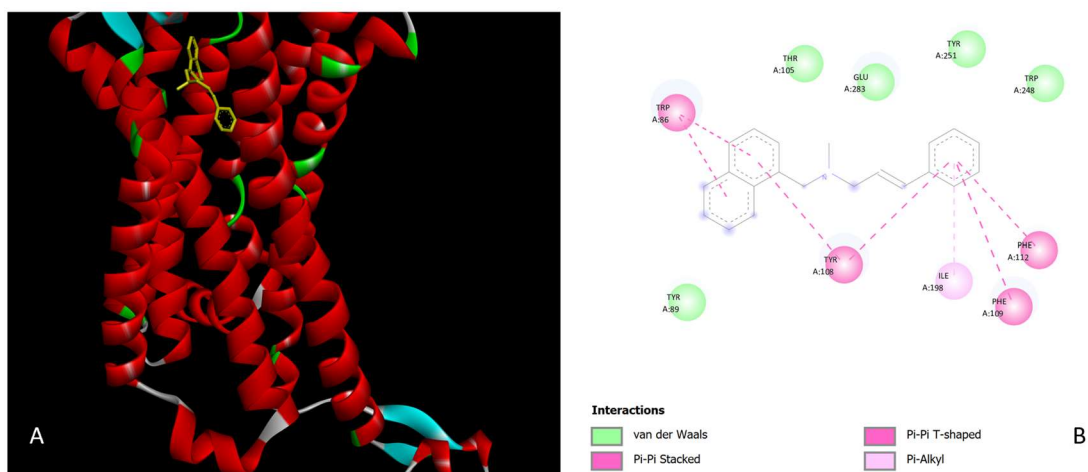


Figure 17. Naftifine bound to chemokine co-receptor CCR5. Panel A shows the orientation of naftifine in the CCR5 binding site. Panel B shows the interactions between naftifine and the amino acids present in the CCR5 binding site.

References

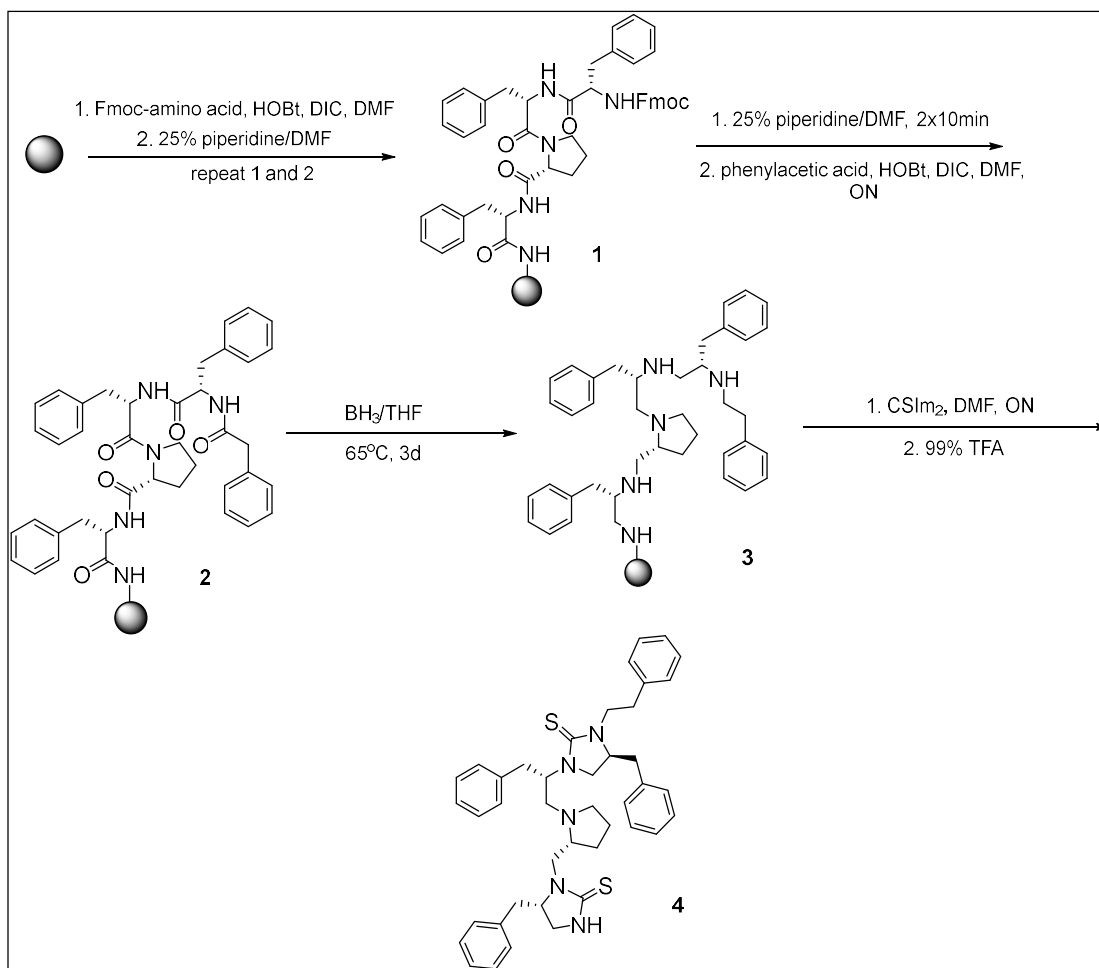
- Blair H. Naldemedine: A Review in Opioid-Induced Constipation. *Drugs*, 2019, 79, 1241–1247.
- Bochum S. et al. Olaparib. *Small molecules in oncology*, 2008, 21, 217-33.
- Cunliffe WJ et al. Efficacy and safety comparison of adapalene (CD271) gel and tretinoin gel in the topical treatment of acne vulgaris. A European multicentre trial. *J Dermatol Treatment*, 1997, 8, 173-8.
- Markham A. Naldemedine: First Global Approval. *Drugs* 2017, 77, 923–927.
- Michel et al. Pharmacology of adapalene. *British Journal of Dermatology*, 1998, 139, 3-7.
- Mohamedi N et al. Ergotism with acute limb ischemia, provoked by HIV protease inhibitors interaction with ergotamine, rescued by multisite transluminal balloon angioplasty. *JMV*, 2021, 46, 13-21.
- Monk J P and Brogden R N. Naftifine. *Drugs*, 2012, 42, 659–672.
- Mühlbacher J M. Naftifine: A topical allylamine antifungal agent. *Clinics in Dermatology*, 1991, 9, 479-485.
- Nefzi, A. and Santos R.T. A versatile access to new macrocyclic oligoheterocycles (MOH). *Bioorganic & Medicinal Chemistry Letters*, 2006, 16, 3358–61.
- Robson M. et al. Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *N Engl J Med* 2017, 377, 523-533.
- Schaefer H, Reichert U. Retinoids and their perspectives in dermatology. *Nouv Dermatol*, 1990, 9, 3-6.
- Silberstein S D. The pharmacology of ergotamine and dihydroergotamine. *Headache*, 1997, 15-25.
- Silberstein S D and Douglas D C. Ergotamine and Dihydroergotamine:History, Pharmacology, and Efficacy. *Headache*, 2003, 43, 144-166.
- Vojkovsky T. Detection of Secondary Amines on Solid-Phase. *Peptide Research*. 1995, 71, 236–237.

CHAPTER 3: EXPERIMENTAL

3.1 Solid phase synthesis: target compound (4)

General information:

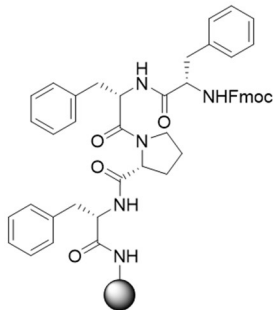
All amino acids, carboxylic acids, reagents and solvents were obtained commercially and used without further purification. Solid phase synthesis was carried out in a polypropylene syringe with a frit.



Synthesis overview for target compound (4)

We based our synthesis procedure on Nefzi and Santos (2006).

Synthesis of (1)



Step 1: 50mg of MBHA (methylbenzylhydramine) resin was contained within a polypropylene syringe with a frit. The amino-resin was neutralized with 10% DIEA (diisopropylethylamine) in DMF (dimethylformamide) overnight (approx. 24h). The resin was washed with DMF (2x10min). Fmoc-Phe-OH (6 equiv, 0.1M) was coupled in the presence of HOBt (hydroxybenzotriazole) (6 equiv., 0.1M) and DIC (diisopropylcarbodiimide) (6 equiv., 0.1M) in anhydrous DMF (2mL) for 60 minutes. A sample was collected to run in the LCMS and HPLC. To do this, the beads were cleaved using 99% TFA (trifluoroacetic acid). Upon filtering, the solution was left open to air for the TFA to evaporate overnight. The resulting precipitate was dissolved in DMF. For the LCMS and HPLC an Agilent C18 column (Synergi Hydro-RP 80A LC Column) was used. A continuous gradient of 5%-95% ACN (acetonitrile) with 0.1% formic acid and H₂O with 0.1% formic acid was set to run for 20min with the percentage of ACN increasing over time. The same procedure was followed for all LCMS and HPLC (Agilent Technologies 1100 and 1200 Series) runs.

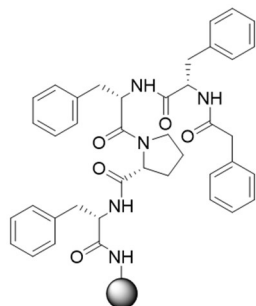
Steps 2-3: The Fmoc group was then removed with 25% piperidine in DMF (2x10min) and the resin was washed with DMF (8x). The free amine was coupled to Fmoc-proline (6equiv, 0.1M) in the presence of HOBt (6 equiv., 0.1M) and DIC (6 equiv., 0.1M) in

anhydrous DMF (2mL) for 60 minutes. A sample was collected to run in the LCMS and HPLC.

Steps 4-5: The Fmoc group was then removed with 25% piperidine in DMF (2x10min) and the resin was washed with DMF (8x). The free amine was coupled to Fmoc-Phe-OH (6equiv, 0.1M) in the presence of HOBt (6 equiv., 0.1M) and DIC (6 equiv., 0.1M) in anhydrous DMF (2mL) for 60 minutes. A sample was collected to run in the LCMS and HPLC.

Steps 6-7: The Fmoc group was then removed with 25% piperidine in DMF (2x10min) and the resin was washed with DMF (8x). The free amine was coupled to Fmoc-Phe-OH (6equiv, 0.1M) in the presence of HOBt (6 equiv., 0.1M) and DIC (6 equiv., 0.1M) in anhydrous DMF (2mL) for 60 minutes. A sample was collected to run in the LCMS and HPLC.

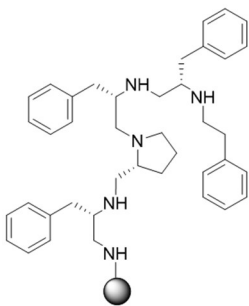
Synthesis of (2)



Steps 8-9: The Fmoc group was then removed with 25% piperidine in DMF (2x10min) and the resin was washed with DMF (8x). The free amine was N-acylated with phenylacetic acid (10 equiv.) in the presence of HOBt (10 equiv.) and DIC (10 equiv.) in anhydrous

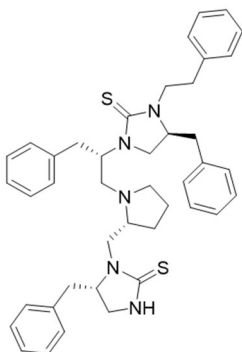
DMF (2mL) overnight (approx. 24h). A sample was collected to run in the LCMS and HPLC.

Synthesis of (3)



Steps 10-11: The amide reduction was performed in a 25mL pear flask under nitrogen. The resin was treated with 1M BH₃-THF (borane–tetrahydrofuran) (40-fold excess over each amide bond). The tube was heated at 65 °C for 72 h, decanted, washed with THF, and any remaining borane quenched with MeOH (methanol). The borane was disproportionated by treatment with piperidine at 65 °C overnight (approx. 24h). The resin was then washed with MeOH (2x), DMF (6x) and left open to air to dry overnight. A sample was collected to run in the LCMS and HPLC.

Synthesis of (4)

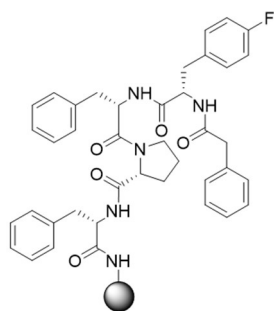


Steps 12-13: The resin-bound polyamines were treated overnight with a 6-fold excess of TCDI (thiocarbonyldiimidazole, 0.05 M) in anhydrous DMF. Following cleavage from the

resin with 99% TFA, the desired product was diluted with acetonitrile/water (50:50) and lyophilized. A sample was collected upon cleaving to run in the LCMS and HPLC. Another sample was collected upon lyophilization (white-yellow powder, 0.0002g) of the desired product to run in the LCMS and HPLC.

3.2 Solid phase synthesis: further analogs by varying amino acids

Synthesis of 4-fluoro-Fmoc-Phe-OH analog

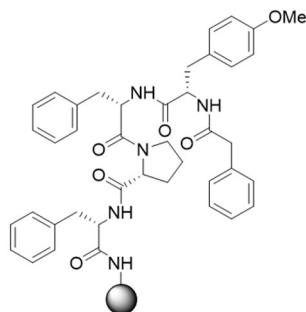


For the synthesis of the 4-fluoro-Fmoc-Phe-OH analog the same procedure as outlined above was used with the following modifications:

In the addition of the last amino acid (steps 6-7), Fmoc-Phe-OH was substituted with 4-fluoro-Fmoc-Phe-OH (6 equiv.).

Unfortunately, our compound decomposed during the reduction reaction. The intermediate we successfully obtained before the reduction reaction is shown above.

Synthesis of 4-methoxy-Fmoc-Phe-OH analog



For the synthesis of the 4-methoxy-Fmoc-Phe-OH analog the same procedure as outlined above was used with the following modifications:

In the addition of the last amino acid (steps 6-7), Fmoc-Phe-OH was substituted with 4-methoxy-Fmoc-Phe-OH (6 equiv.).

Unfortunately, our compound decomposed during the reduction reaction. The intermediate we successfully obtained before the reduction reaction is shown above.

We based our repurposing screening procedure on Danziger (2022).

3.3 Repurposing screening

Overview:

- 3.3.1 Necessary files for screening
- 3.3.2 Hardware and Software
- 3.3.3 Receptor protein preparation and Grid Box creation
- 3.3.4 Creating the configuration file (Notepad or other text editor)
- 3.3.5 Preparation of positive control ligands (Maraviroc)
- 3.3.6 AutoDock Vina Virtual Docking of one ligand
- 3.3.7 Preparation of FDA test ligands
- 3.3.8 AutoDock Vina Virtual Screening of FDA test ligands
- 3.3.9 Analysis of Virtual Screening Results for FDA test ligands

3.3.1 Necessary files for screening

- Prepared receptor protein in PDBQT format
- Prepared positive control ligands in PDBQT format
- Folder with prepared FDA approved compounds (test ligands)
- FDA ligands text file
- Configuration file with the receptor protein (and ligand if docking only one ligand)
- Vina_windows.pl (to dock multiple ligands in the FDA text file)

(create a text file named Vina_windows.pl and copy the script below)

Script for Vina_windows is as follows:

```
#!/usr/bin/perl
print"Ligand_file:\t";
$ligfile=<STDIN>;
chomp $ligfile;
open (FH,$ligfile)||die "Cannot open file\n";
@arr_file=<FH>;

for($i=0;$i<@arr_file;$i++)
{
    print"@arr_file[$i]\n";
    @name=split(/\./,@arr_file[$i]);
}
for($i=0;$i<@arr_file;$i++)
{
    chomp @arr_file[$i];
    print"@arr_file[$i]\n";
    system("vina.exe --config configFDA.txt --ligand @arr_file[$i] --log
    @arr_file[$i]_log.log");
}
```

3.3.2 Hardware and Software

- Computer/Laptop: Linux, Macintosh, or Windows PC
- Notepad or other text editor
- AutoDock4 (for protein preparation and visualization)
- AutoDock Vina

- AutoDockTools (Part of MGLTools)
- AutoLigand (Part of AutoDockTools)
- BIOVIA Discovery Studio Visualizer (for protein/binding site preparation and visualization)
- Open Babel (for file conversions)
- Ubuntu (for separating multiple ligands in one file)
- Perl (for docking multiple ligands at once)

3.3.3 Receptor protein preparation and Grid Box creation

- Download the receptor protein from the Protein Data Bank (<https://www.rcsb.org/>)
- Type 4MBS (for the CCR5 chemokine receptor) in the search bar and download protein as a PDB file

Discovery Studio (for protein editing and defining the binding site):

- Open PDB receptor protein (4MBS) in Discovery Studio
- *View > Hierarchy* to see protein chains, ligands, water molecules, etc.
- Delete water molecules, cofactors, ions, etc. that should not be included in the protein receptor for the run
- Select ligand in the active site of interest:

Receptor-Ligand Interactions > Define and Edit Binding Site > From Current Selection.

A sphere defining the binding site will appear based on where your selected ligand is located.

- *SBD Site Sphere > (right click) Attributes of SBD_Site_Sphere*
 1. Expand to include residues of interest by editing the radius.

2. Get x, y, z coordinates for the binding site.

For 4MBS: X,Y,Z → 149.854622, 108.414865, 22.290459

Radius = 10

- Copy and paste grid dimensions (xyz and radius) into a new notepad file to be used as the configuration file for the run (see 3.3.4 for further instructions on configuration file preparation)
- Close Attributes of SBD_Site_Sphere window by saving current grid dimensions
- Delete any ligand groups and protein chains that were used for defining the SBD site sphere but are not needed for the run
- *Chemistry > Hydrogens > Add Polar* to add polar hydrogens
- Save as PDB

Autodock (for further protein editing and conversion to PDBQT):

- Open Autodock
- *Grid > Macromolecule > Open* and use the “Files of type” menu to choose “all files” and open the receptor protein file previously edited in Discovery Studio
- Select your file > *Open*.
- To further edit your receptor protein:
 1. Edit > Hydrogens > Edit Histidine Hydrogens > HE2 (for all hydrogens) > Apply > Dismiss
 2. Edit > Charges > Add Kollman Charges
 3. Edit > Misc > Check for Missing Atoms
 4. Edit Charges > Check Totals on Residues > Spread Deficit Over All Atoms in Residues

Protein preparation is now done

- Grid > Macromolecule > Choose > Save as PDBQT and place in docking folder with ligands

3.3.4 Creating the configuration file (Notepad or other text editor)

- Create a Notepad/text editor file (see 3.3.3)

Organize your notepad/text editor configuration file as shown below:

```
receptor = 4MBS.pdbqt
```

```
center_x = -12.661086
```

```
center_y = 31.554174
```

```
center_z = 45.079252
```

```
size_x = 10
```

```
size_y = 10
```

```
size_z = 10
```

```
num_modes = 10 (This specifies the maximum number of binding modes to generate)
```

```
energy_range = 4 (This specifies the energy difference between the best and the worst binding mode)
```

```
exhaustiveness = 24 (This specifies the number of runs)
```

- If docking only one ligand, include the ligand below the receptor:

```
receptor = 4MBS.pdbqt
```

```
ligand = Maraviroc.pdbqt
```

Save (as 4MBS_config.txt) and place in your working folder

3.3.5 Preparation of positive control ligands (here: Maraviroc)

Autodock:

- Open Autodock
- Go to: All Molecules > (right click) Read Molecule > select your ligand file (Maraviroc) > Open

- To prepare your ligand:
 1. Edit > Delete water
 2. Edit > Charges > Add Kollman Charges
 3. Edit > Charges > Compute Gasteiger
 4. Ligand > Input > Choose > click on your ligand (Maraviroc) > Select Molecule for AutoDock4
 5. Ligand > Torsion Tree > Detect Root
- To save the prepared ligand as PDBQT:

Ligand > Output > Save as PDBQT and place in your working folder

3.3.6 AutoDock Vina Virtual Docking for one ligand

- Open the command terminal on your device
- Change the directory to your working folder by typing:
cd C:\Users\garyf\Desktop\Docking\4MBS
- Run vina using the prepared configuration file and save the output as output.txt by typing:
vina.exe --config 4MBS_config.txt --log output.txt

3.3.7 Preparation of FDA test ligands

The compounds were downloaded and minimized by Danziger (2022).

- Create a separate working folder for all the pdbqt ligand files (FDAligands)

Ensure that the following files are also pasted in your working folder:

- Autodock4.exe
- Autogrid4.exe
- 4MBS_config.txt (configuration text file for receptor protein)
- Vina.exe

- Vina_license.rtf
- Vina_split.exe
- Vina_windows.pl
- 4MBS.pdbqt (receptor protein)

3.3.8 AutoDock Vina Virtual Screening for FDA test ligands

FDA test ligands text file:

- Create a Notepad/text editor file (FDAligands) in your working folder
- Open the command terminal on your device
- Change the directory to your working folder by typing:
cd C:\Users\garyf\Desktop\Docking\CCR5 antagonists pdbqt
- Copy the names of everything in your working folder into your FDAligands.txt file
by typing:

dir /B > FDAligands.txt

- Remove the file names you do not need from your text file (only include PDBQT ligands)

To run:

1. Type in command terminal (you have already directed it to your working folder):

perl Vina_windows.pl (this commands will ask you for the ligand file)

FDAligands.txt (this points to the ligand txt file you created above)

The run took approximately a week to complete

A results text file was generated for each FDA ligand

3.3.9 Analysis of Virtual Screening Results for FDA test ligands

Compile all the resulting text files into one new text file:

Please note:

1. The tail command requires a Mac
 2. The tail command reads the last (-n) lines [where n is the number of lines] of all .log files in the folder and writes then onto (>) output file (output.txt)
- Open the command terminal on your computer
 - Navigate to the working folder containing the FDA approved compounds log files as previously shown
 - Compile all the resulting text files into one new text file by typing:
*tail -12 *.log > ccr5_output.txt*

To rank the compounds based on their binding affinities use the following Python code (this creates an Excel document with ranking of the compound, binding affinity and compound number):

```
#reads a text file and extracts binding energies

from openpyxl import Workbook

#Open the file to read
file = "/Users/leepedzisa/Desktop/Research/ccr5_output.txt"

# read by line numbers
def get_lines(fp, line_numbers):
    return (x for i, x in enumerate(fp) if i in line_numbers)

def number_lines(file):
    with open(file, 'r') as fp:
        #count the number of lines
        x = len(fp.readlines())
    return x
    #print(x)

def get_compounds(file):
    compounds = []
```

```

with open(file, 'r') as fp:
    x = number_lines(file)
    #read lines 2, 15, ..., < x
    cmpd_lines = get_lines(fp, [i for i in range(0, x)])
    for line in cmpd_lines:
        cmpd_line = line.split()
        if cmpd_line and cmpd_line[0][0] == "=":
            cmpd_name = cmpd_line[1]
            name_split = cmpd_name.split(".")
            cmpd = name_split[0]
            #cmpd = line.split()[1].split(".")[0]
            compounds.append(cmpd)
    return compounds

def get_energies(file):
    energies = []
    with open(file, 'r') as fp:
        #read lines 2, 15, ..., < x
        lines = get_lines(fp, [i for i in range(1, x)])
        j = 0
        for line in lines:
            cmpd_energy = []
            temp = line.split()
            if temp and temp[0] == "1":
                cmpd_energy.append(int(compounds[j]))
                cmpd_energy.append(float(temp[1]))
                energies.append(cmpd_energy)
                j += 1
    return energies

def print_energies(number):
    if number <= len(energies):
        for i in range(number):
            print("Compound: ", sorted_energy[i][0], "\t\tRank: ",
i+1, "\tBinding energy: ", sorted_energy[i][1])
    else:
        for i in range(len(energies)):
            print("Compound: ", sorted_energy[i][0], "\t\tRank: ",
i+1, "\tBinding energy: ", sorted_energy[i][1])

def write_xl(lst, number):
    #method requires "from openpyxl import Workbook"
    book = Workbook()
    sheet = book.active
    sheet["A1"] = "Rank"
    sheet["B1"] = "Energy"
    sheet["C1"] = "Compound"

    for i in range(number):
        sheet.cell(row = i+2, column = 1).value=i+1
        sheet.cell(row = i+2, column = 2).value=lst[i][1]
        sheet.cell(row = i+2, column = 3).value=lst[i][0]

```

```

book.save("Test_energies.xlsx")

#Run functions
x = number_lines(file)
compounds = get_compounds(file)
energies = get_energies(file)

sorted_energy = sorted(energies, key=lambda x:x[1])

write_xl(sorted_energy, 100)

#print_energies(20)

```

References

Nefzi, A. and Santos R.T. A versatile access to new macrocyclic oligoheterocycles (MOH). *Bioorganic & Medicinal Chemistry Letters*, 2006, 16, 3358–3361.

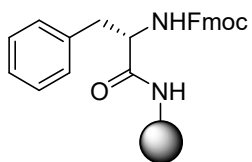
Danziger S. *Discovering a Fatty Acid Synthase Inhibitor for Cancer Treatment*. Union College, 2022.

CHAPTER 4: CONCLUSIONS AND FUTURE WORK

In our project we successfully synthesized target compound (4) and following it to obtain target (4). We also confirmed the synthesis of (4) using LCMS and HPLC. Additionally, we made progress in the synthesis of two analogs through varying amino acids (4-fluoro-Fmoc-Phe-OH analog and 4-methoxy-Fmoc-Phe-OH analog). However, these analogs decomposed during the reduction reaction. Thus, we would like to reattempt analog synthesis and obtain the aforementioned compounds. Due to time constraints, we were unable to test (4) for biological activity by evaluating its binding to CCR4 and CCR5 and its specificity. Therefore, this is one of the steps we would like to investigate in the future. We have already identified some resources that could aid us in this. As the docking screen took longer than expected, we were also unable to screen the FDA approved compounds against CXCR4. In the future, we would like to move forward with docking the FDA ligands to CXCR4 as well. Lastly, it would be interesting to test the FDA ligands that ranked high in CCR5 binding affinity in the lab. We could do this by obtaining these compounds commercially and running a CCR5 binding assay.

APPENDIX

Addition of first Fmoc-Phe-OH



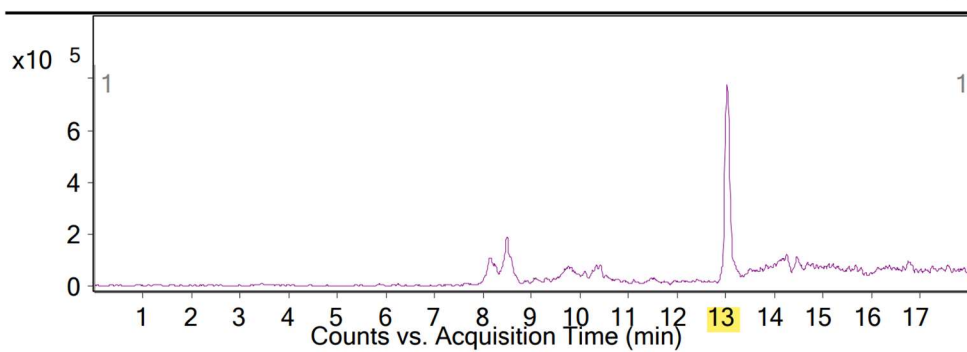
Chemical Formula: $C_{24}H_{22}N_2O_3$

Exact Mass: 386.16

LCMS spectra:

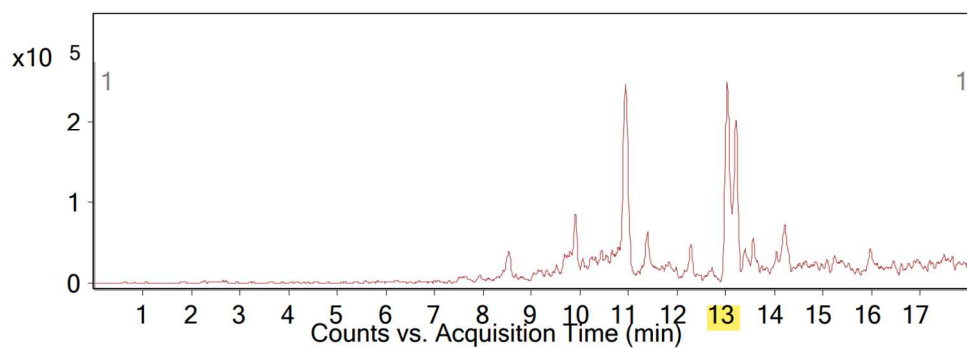
Fmoc-Phe-OH #1

(+1) 387



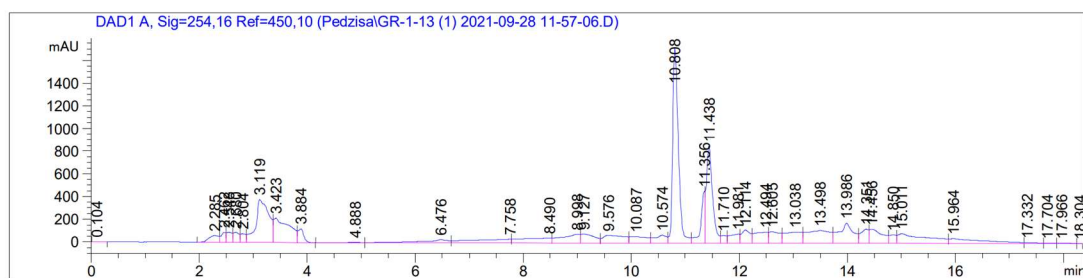
Fragmentor Voltage 135 Collision Energy 0 Ionization Mode ESI

(+23) 409

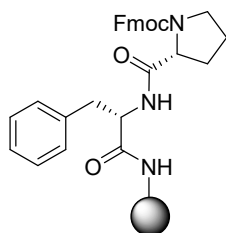


Fragmentor Voltage 135 Collision Energy 0 Ionization Mode ESI

HPLC spectra:



Addition of proline

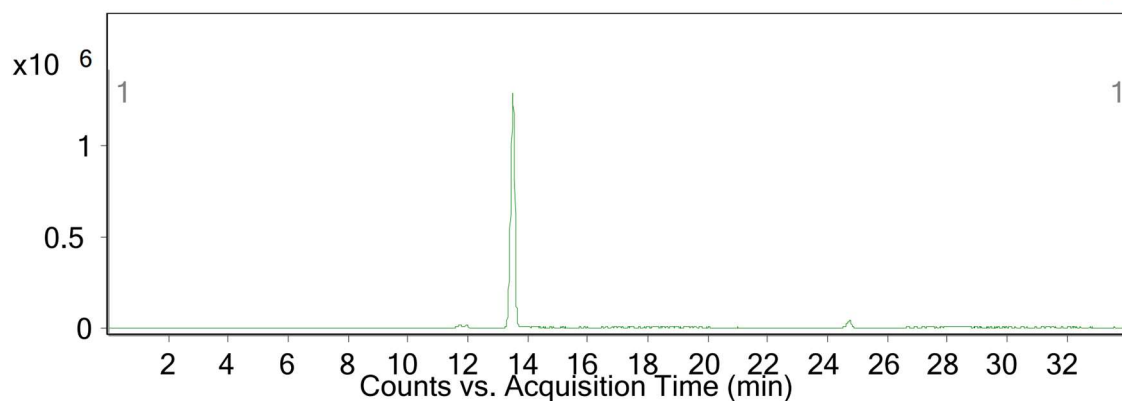


Chemical Formula: $C_{29}H_{29}N_3O_4$
Exact Mass: 483.22

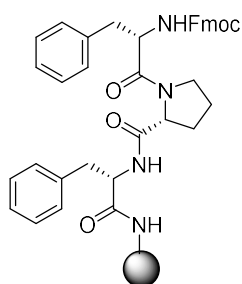
LCMS spectra:

Fragmentor Voltage 135 Collision Energy 0 Ionization Mode ESI

(+1) 484

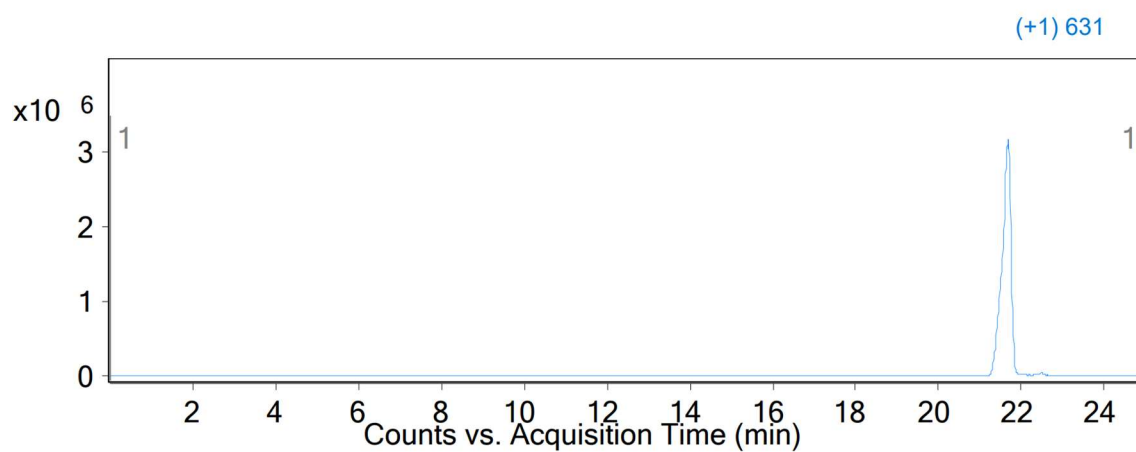


Addition of second Fmoc-Phe-OH

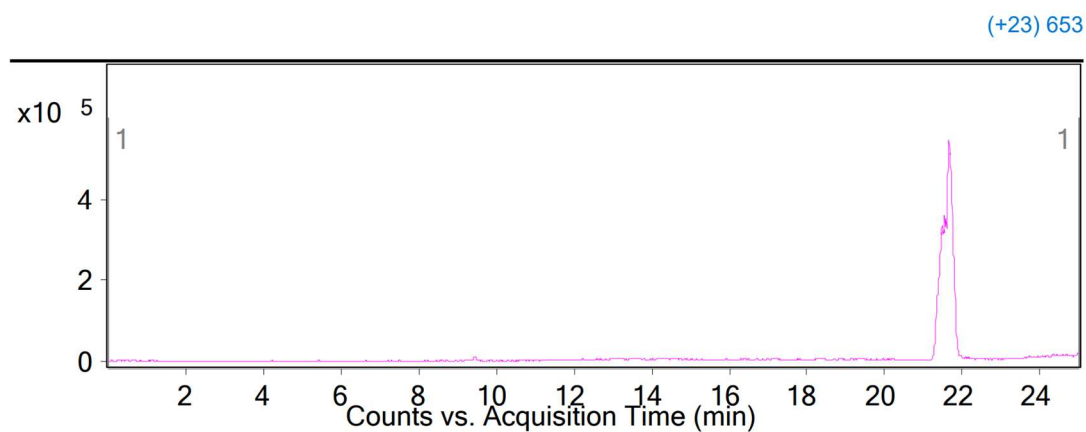


Chemical Formula: $C_{39}H_{40}N_4O_5$
Exact Mass: 644.30

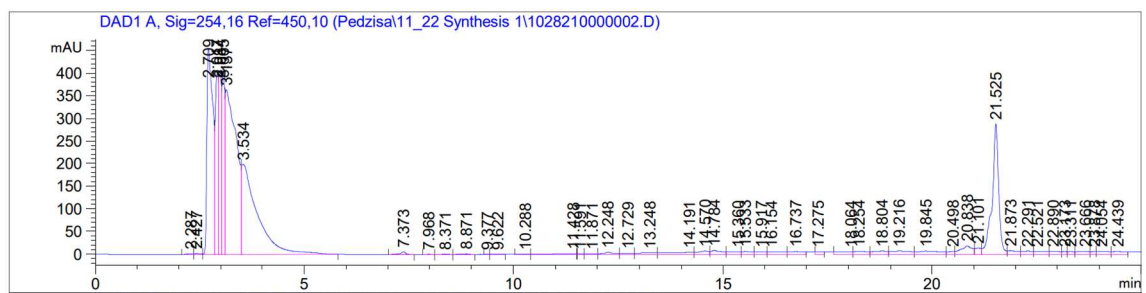
LCMS spectra:



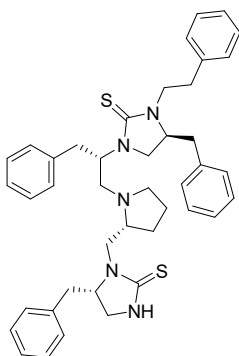
Fragmentor Voltage 135 Collision Energy 0 Ionization Mode ESI



HPLC spectra



Cyclization with CSIm₂ (final step)



Chemical Formula: C₄₂H₄₉N₅S₂
Exact Mass: 687.34

LCMS spectra:

