KADIR HAS UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING
DEPARTMENT OF BIOINFORMATICS AND GENETICS

IN SILICO SCREENING OF POTENT HIV-1 INTEGRASE INHIBITORS FOR THE TREATMENT OF HUMAN IMMUNODEFICIENCY VIRUS (HIV)

Augustine S. Samorlu

MASTER'S THESIS

ISTANBUL, APRIL, 2018

In silico Screening of Potent HIV-1 Integrase Inhibitors for the Treatment of Human Immunodeficiency Virus (HIV)

Augustine S. Samorlu

MASTER'S THESIS

Submitted to the Graduate School of Science and Engineering of Kadir Has University in partial fulfillment of the requirements for the degree of Master's in the Program of Computational Biology and Bioinformatics

ISTANBUL, APRIL, 2018

DECLARATION OF RESEARCH ETHICS

I, Augustine S. Samorlu, hereby declare that;

- this Master's Thesis is my own original work and that due references have been appropriately provided on all supporting literature and resources;
- this Master's Thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- I have followed "Kadir Has University Academic Ethics Principles" prepared in accordance with the "The Council of Higher Education's Ethical Conduct Principles"

In addition, I understand that any false claim in respect of this work will result in disciplinary action in accordance with University regulations.

Furthermore, both printed and electronic copies of my work will be kept in Kadir Has Information Center under the following condition as indicated below:

❖ The full content of my thesis will be accessible from everywhere by all means.

Augustine S. Samorlu

_____

KADIR HAS UNIVERSITY

GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

ACCEPTANCE AND APPROVAL

This work entitled ***IN SILICO*** **SCREENING OF POTENT HIV-1 INTEGRASE INHIBITORS FOR THE TREATMENT OF HUMAN IMMUNODEFICIENCY VIRUS (HIV)**

Prepared by **AUGUSTINE S. SAMORLU** has been judged to be successful at the defense exam held on **13, April, 2018** and accepted by our jury as **MASTER'S THESIS**.

APPROVED BY:

Prof. Dr. Kemal Yelekçi (Advisor) Kadir Has University

Prof. Dr. Safiye ERDEM Marmara University

Dr. Hatice Bahar Şahin Kadir Has University

I certify that the above signatures belong to the faculty members named above.

Doç. Dr. Ebru D. Akten

Dean of Graduate School of Science and Engineering

DATE OF APPROVAL: 13/04/2018

TABLE OF CONTENTS

IN SILICO SCREENING OF POTENT HIV-1 INTEGRASE INHIBITORS FOR THE TREATMENT OF HUMAN IMMUNODEFICIENCY VIRUS (HIV)

ABSTRACT

The purpose of this research is to obtain potent inhibitors for the treatment of HIV-1, which causes a ceaseless and depressive disease of the human immune system known as AIDS.

HIV-1 integrase inhibitors are very essential in the treatment of HIV-1 infection. Inhibiting the enzyme integrase (IN) results in the termination of the HIV-1 replicative process, thus, putting an end to its life cycle. The *in silico approach* was employed for the purpose of obtaining these inhibitors. Basically, the Otava's chemical library was screened as well as a systematic approach of designing an inhibitor was employed, thus, leading to the yielding of four potent IN's inhibitors. The potency of inhibitors was measured through two docking programs, namely PyRx and AutoDock 4.2. For a chemical structure to be considered as a potent inhibitor for this study, it must generate a calculated free energy of binding = negative 8.00 kcal/mol or less and also interact with any of the three key important amino acids of IN. The Discovery Studio Visualizer was used to draw the structure of inhibitors at the same time generating pictures of enzyme inhibitor complexes, displaying 2D and 3D structures which enable us to determine the type of interaction between the enzyme and the inhibitor.

Of the four potent inhibitors obtained, one was designed and yielded an estimated inhibition constant (Ki) of 652.83 nanomolar and a free energy of binding negative 8.44kcal/mol. The remaining three inhibitors were screened from the OTAVA's chemical library, listed in parenthesis with OTAVA's code, Ki and binding energy respectively; (107320240=131.7nm, -9.39kcal/mol; 109750115=44.19nm, -10.03kcal/mol; 111150115=395.19nm, -8.74kcal/mol).

Keywords: Integrase, HIV/AIDS, Docking, AutoDock, PyRx, Inhibitor, Otava chemical library, Protease, Reverse transcriptase, *In silico* screening.

İNSAN İMMÜNYETMEZLIK VIRÜSÜ (HIV) TEDAVISINE YÖNELIK *IN SILICO* TARAMAYLA POTANSIYEL İNTEGRAZ İNHIBITÖRLERININ BULUNMASI

ÖZET

Bu araştırmanın amacı, AIDS olarak bilinen insan bağışıklık sistemine etki eden, duraksamayan ve depresif bir hastalığa neden olan HIV-1'in tedavisi için potansiyel inhibitörleri elde etmektir.

HIV-1 integraz inhibitörleri, HIV-1 enfeksiyonunun tedavisinde çok önemlidir. İntegraz enziminin (IN) inhibe edilmesi HIV-1 virüsünün çoğalma işleminin sonlandırılmasına neden olur. Böylece yaşam döngüsüne son verir. Bu inhibitörleri elde etmek için bilgisayar destekli *in silico* yaklaşım kullanılmıştır. Temelde, Otava Kimya Kütüphanesi tarandı ve inhibitör tasarımında kullanılan sistematik yaklaşımlar uygulandı, böylece dört güçlü integraz inhibitörü bulundu. İnhibitörlerin enzime bağlanma değerleri PyRx ve AutoDock 4.2 doklama programları kullanılarak gerçekleştirildi. Çalışmada bir kimyasalın güçlü bir inhibitör olabilmesi için hesaplanan serbest bağ enerjisi = -8.00 kcal / mol veya daha az olması ve integrazın aktif bölgesinde bulunan 3 önemli amino asidinden herhangi biri ile de etkileşimde bulunması kriterine uyulmuştur. Discovery Studio Visualizer, inhibitörlerin yapısını çizmekte, inhibitörü komplekslerinin resimlerini üretmekte, enzim ve inhibitör arasındaki etkileşimin türünü belirlememizi sağlayan 2D ve 3D yapıları görüntülemek için kullanıldı.

Elde edilen dört güçlü inhibitörden, kendimizin tasarladığı moleküllerden (Ki= 652.83 nanomolar bir ve bağlanma serbest enerjisi -8.44kcal / mol), kalan üç inhibitörde, Otava Kimya Kütüphanesi'nde tarandı ve Otava koduyla parantez içerisinde listelenmiştir. Bunların inhibisyon sabiti ve bağlanma enerjileri sırasıyla; 107320240, Ki=131.7nm, -9.39kcal/ mol; 109750115, Ki= 44.19nm, -10.03kcal / mol; 111150115 Ki = 395.19nm, -8.74kcal / mol olarak bulunmuştur.

AnahtarSözcükler: Integrase, HIV/AIDS, Docking, AutoDock, PyRx, Inhibitor, Otava chemical library, Protease, Reverse transcriptase, *In silico* screening.

ACKNOWLEDGEMENT

I want to take this opportunity to thank The Scientific and Technological Research Council of Turkey (TÜBİTAK 2235) for funding my studies under the program 'Graduate Scholarship for the Least Developed Countries. Your generosity has led me to acquiring an education that will help me contribute to the educational sector of my country (Liberia).

Secondly, my sincere appreciation goes to my able supervisor (Prof. Dr. Kemal Yelekci) for his guidance during my studies and the entire faculty members of the department of Bioinformatics and Computational Biology of Kadir Has University for affording me the opportunity to achieve my dream. At the same time, I will like to appreciate all of my friends who aided me one way or another through my studies.

To My Parents Mr. Edward Samorlu And Mrs. Sarah Samorlu.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

RT……..…Reverse Transcriptase

IN…….…….Integrase

PR…….Protease

WHO……..……World Health Organization

UNAIDS…….……..United Nation Programme on HIV/AIDS

DNA……..……Deoxyribonucleic acid

RNA…….…….Ribonucleic acid

Mg……Magnesium

Mn……Manganese

ATP……..Adenosine Triphospate

NHEJ……Non-homologous end joining

BER…….…….Base excision repair

ADMET……..……Absorption, Distribution, Metabolism, Excretion and Toxicity

CCD……..…Catalytic core domain

FDA……Food and Drug Administration

D or Asp…….Aspartate

E or Glu……..Glutamate

Cys……..Cysteine

GP120……..Glycoprotein

INI……Integrase Inhibitors

ART……..Antiretroviral therapy

INSTIs……..…… Integrase strand transfer inhibitors

PBD……..Protein Data Bank

HTS……..High-Throughput Screening

CADD……..Computer Aided Drug Design

KI……Inhibition constant

PDBQT……Protein Data Bank Partial charge and Atom Type

DPF……..Docking Parameter File

GPF……..Grid Parameter File

1. OBJECTIVE AND STATEMENT OF THE PROBLEM

1.1 Objective

Human Immunodeficiency Virus (HIV-1) is a pathogen that causes a continuous and depressive disease of the human immune system known as human acquired immunodeficiency syndrome (AIDS) (Vyas, Shah, & Ghate, 2017). One of the major epidemics of the world that is affecting many people especially in Africa, of which social factor is serving as one of its major contribution 'AIDS', was first reported in June 1981 (Pommier, Johnson, & Marchand, 2005).

HIV was discovered in 1983 and has led to the death of approximately 33 million people in the world. As a result of the introduction of the combination antiretroviral therapy (cART) in the 20^{th} century, reduction in the death of people living with HIV has been observed (Bilodeau et al., 2014), (Hosseini, Mollica, & Mirzaie, 2016).

Currently, reverse transcriptase (RT), integrase (IN) and protease (PR) are the three enzymes that play important roles in HIV-1 virus replication cycle. With a complete inhibition of the HIV-1 IN enzyme, the treatment of HIV infection can be successful (Gupta, Garg, & Roy, 2012)(Kim, 2002).

In view of the above, this research was conducted for the purpose of

1. Obtaining potent inhibitor(s) for the HIV-1 integrase enzyme using an *in silico* approach.
2. Screening a library of compounds for the selection of ligands with a free binding energy (ΔG) of -8.0kcal /mol or lower.

3. Identifying ligands that can be easily modified by computational biologists or drugs designers, thus producing a potent inhibitor.
4. This study is also aimed at inserting the missing residues in the target crystal structure of integrase enzyme (code: 1QS4).

1.2 Statement Of The Problem

HIV continues to be a major global health issue, destroying many lives every year. Drugs industries are working very hard in developing potential inhibitors against the three enzymes of HIV. Unfortunately, some of the inhibitors are either having side effects and not tolerant or the enzymes are constantly mutating, thus, causing the drug to be less potent. In this light, there is a need for computational biologists and bioinformaticians to constantly produce potent inhibitors for those enzymes.

As a member of the HIV-1 enzymes, IN plays a major role in the replication of this virus. After the process of reverse transcriptase, IN transfers the double-stranded DNA into the host genome (Kim, 2002). Blocking this enzyme will lead to interruption of further processes, therefore resulting in the reduction of the morbidity and mortality rates. Based on this, the research is being conducted.

2. INTRODUCTION

HIV/AIDS remains the primary killer disease in Sub-Saharan Africa, thus, accounting for majority deaths in adults. It was established by the Joint United Nation Programme on HIV/AIDS (UNAIDS) along with the World Health Organization (WHO) in 2013, that 35 million people were living with HIV in the world, while 2.1 million people were newly infected. Majority of the people living with HIV were reported in sub-Saharan Africa, approximately 24.7 million with 1.1 million deaths taking place in a year (Lutambi, 2015).

There are generally two types of HIV, namely HIV-1 and HIV-2; they can be distinguished genetically and pathophysiologically. Clinically, the progression of AIDS in HIV-1 infection is much faster when compared with HIV-2 infection (Chupradit et al., 2017).

2.1 Integrase

From the earliest stage, the development of HIV inhibitors was focused on basically two enzymes (HIV-1 RT and HIV-1 PR). However, the ability of HIV-1 to constantly develop drug resistance along with toxicity problems obliged researchers to develop a new target for the treatment of HIV-1, known as IN (Yehudagoldgur et al., 1999), (Bilodeau et al., 2014). Currently, all FDA approved IN's drugs accurately block the strand transfer stage of HIV, therefore, inhibitors of this nature are called IN strand-transfer inhibitors (INSTIs) (Varadarajan et al., 2016).

For the human immunodeficiency virus to successfully overcome the human genome, it must transfer its double-stranded DNA obtained from the enzyme 'reverse transcriptase' into the host chromosome. The enzyme responsible for this action is known as HIV-1 integrase (IN) (Gupta et al., 2012) (Greene & Peterlin, 2008). Apart from just transferring the double-stranded DNA into the host genome, IN also carries on HIV-1 particle production, nuclear import of the pre-integration complex as well as cooperation in reverse transcription (Mulder, Chakrabarti, & Muesing, 2002). IN undergoes two stages in order to complete its task. At the beginning, it removes a dinucleotide from each 3'-ends by binding to the viral DNA in the cytoplasm. Finally, the complex is transported into the nucleus where the strand transfer takes place by covalently connecting the 3'-ends of the viral DNA to the 5'-ends of the host cell DNA, Figure 2.1 and Figure 2.3d &e, (Bilodeau et al., 2014; Inhibitor & Bailey, 2015). During the integration reaction, only one of the following divalent cations (Mg^{2+} or Mn^{2+}) is needed for the catalytic activity instead of a source of energy like ATP (Barreca et al., 2006). To complete these intricate processes, IN interacts with different cellular factors and exploits their functions. Some of those factors include non-homologous end joining (NHEJ), the base excision repair (BER) and the enzyme polymerase-1 which work in maintaining the genome integrity. For example, NHEJ and BER work by repairing the retroviral integration gaps while polymerase accelerates the access of the repaired gaps to the integration site (Mulder et al., 2002).

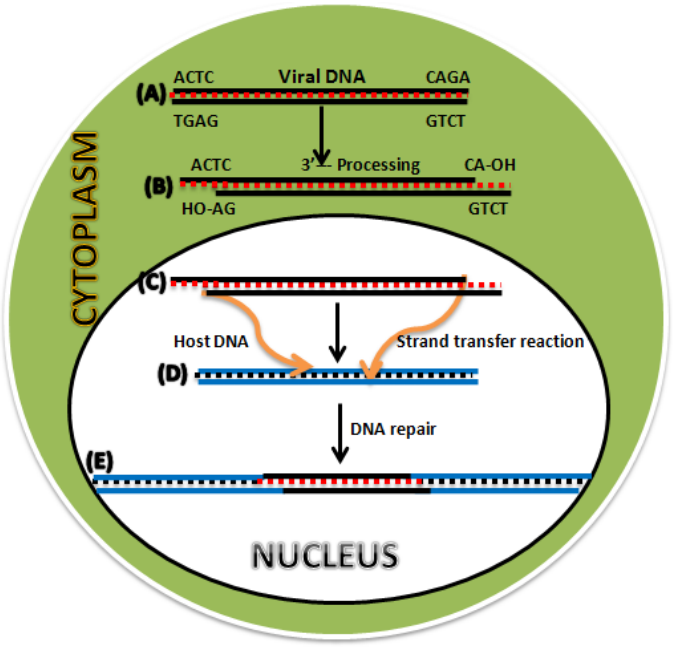


Figure 2.1: Enzymatic illustration of HIV-1 integrase in an infected cell (Inhibitor & Bailey, 2015):

In this **Figure**, (**Fig. 2.1A)** the viral DNA has the nucleotide bases; Thymine, Guanine, Adenine, Guanine (TGAG) left and Cytosine, Adenine, Guanine, Adenine (CAGA) right on its 3' ends, with complementary base pairs on each 5' ends. (**Fig. 2.1B)** based on the successful removal of a dinucleotide from each 3' ends of the viral DNA, (**Fig. 2.1C)** a successful transfer to the nucleus and (**Fig. 2.1D)** attachment to the host DNA was achieved. Finally (**Fig. 2.1E)** a complete DNA repair takes place in which the host DNA becomes fully infected by the viral DNA.

In recent studies, *in silico* screening is serving as a great means for the identification and optimization of very potent HIV-1 IN inhibitors with a good ADMET property. Although many HIV-1 IN inhibitors have been developed over the years, only two of them have been successfully tested for HIV-1 IN. They include Raltegravir and Elvitegravir which are mainly used in combination with other antiretroviral drugs (Gupta, Garg, & Roy, 2013).

Integrase is a very attractive target for the development of new HIV-1 inhibitor because it lacks a cellular counterpart and plays a major role in the replication cycle (Srivastav

& Tiwari, 2013). Another privilege IN has over reverse transcriptase and protease is the ability to use a single active site to house DNA substrates in two different configurations. As a result of this, HIV finds it very difficult to develop resistance against IN's inhibitors (Salam Pradeep Singh, 2012). However, obtaining IN inhibitors through structure-based drug design is very difficult because of its shallow substrate binding site, the unavailability of a crystal structure for the protein's full structure (Rihn, Hughes, Wilson, & Bieniasz, 2015) and the absence of a lead compound (Zeinalipour-Loizidou, Nicolaou, Nicolaides, & Kostrikis, 2007).

The ability for HIV-1 IN to insert a double-stranded DNA into the host genome makes it a very important enzyme in the life cycle of HIV-1. Its 288 amino acids, measuring 32kDa can be grouped into three structural and functional domains (Han et al., 2013); namely, catalytic core domain (CCD, ranging from 50-212 residues), N-terminal domain (NTD, ranging from 1-50 residues) and the C-terminal domain (CTD, ranging from 213-288 residues) as well as Mg^{2+} or Mn^{2+} (Righi et al., 2017), (Vyas et al., 2017), (Ercan, 2017). Both 3' processing and strand transfer reactions are done by CCD which also contributes to the host target DNA binding (Rihn et al., 2015).

IN contains three important acidic amino acids located in the catalytic core domain (residues 50 to 212) (Diamond & Bushman, 2006). If any of these three amino acids (D64, D116 or E152) including Mg^{2+} interacts accurately with a ligand, the enzymatic activities of IN are terminated as well as the ability for the virus to replicate (Righi et al., 2017) (**Figure 2.2a**). Another important amino acid located in the N-terminal domain of IN that is involved in zinc coordination is Cysteine (Cys). The act of mutating the N-terminal Cys residues of IN results in the disruption of its 3'-end processing. Of the six Cys residues of IN, three are located in the core domain, one in the C-terminal domain and the remaining two are located in the N-terminal domain (Zhu, Dobard, & Chow, 2004).

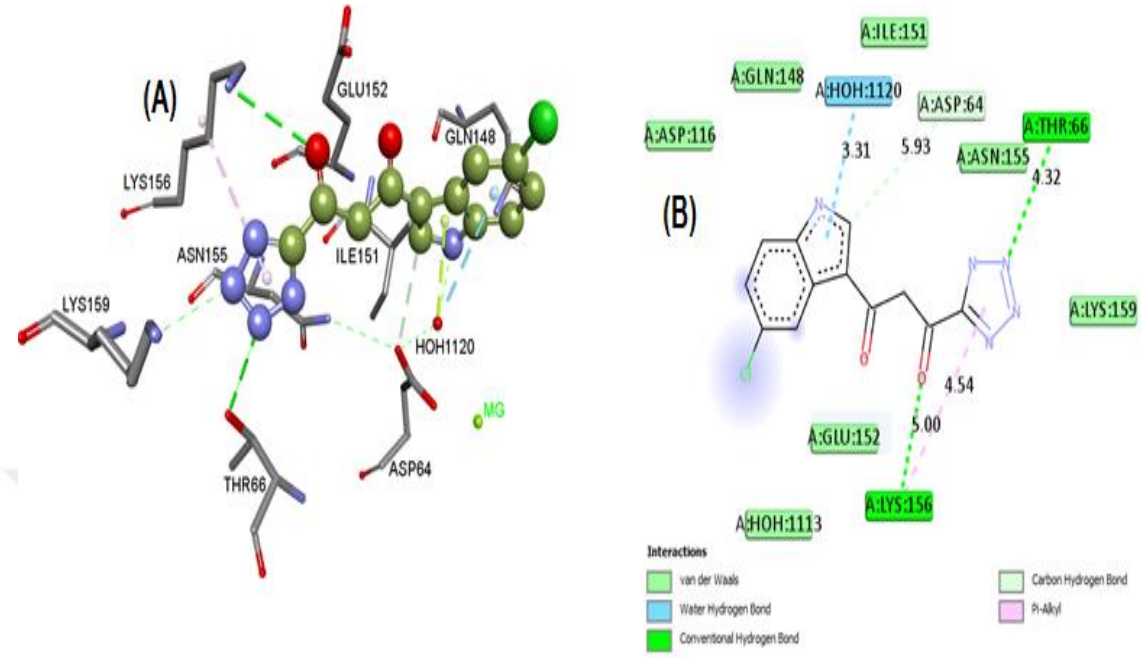


Figure 2.2 A & B: Crystal structure of IN in complex with its native ligand (3D and 2D structures respectively):

(Figure 2.2a) the ligand is displayed as a scaled ball and stick and **(Figure 2.2b)** specific interactions are shown along with their distances; carbon-hydrogen bond interaction is displayed by ASP 64 while THR66 and GLU152 are having a conventional hydrogen bond and van der Waals interactions respectively. The water molecule (HOH 1120) is forming a water hydrogen bond interaction, GLN 148, ILE 151, ASN 155, LYS 159, HOH 1113 and ASP 116 are having van der Waals interactions with the ligand while LYS 156 is forming a conventional Hydrogen Bond and Pi-Akyl interactions.

2.2 Reverse Transcriptase (RT)

Before the virus can be integrated into the host genome by IN, it must be converted from a single-stranded viral RNA into a double-stranded pro-viral DNA, which is the primary function of HIV-1 reverse transcriptase (HIV-1 RT) (Kim, 2002). HIV-1 RT replicates the single-stranded RNA genome into a double-stranded DNA, thus making it

one of the main targets of antiviral therapy. It focuses on the use of chain-terminating nucleoside analogs and non-nucleoside inhibitors (Jaeger, Restle, & Steitz, 1998).

RT is a versatile enzyme, possessing DNA-directed DNA polymerase, RNA-directed DNA polymerase, and ribonuclease activities. Its molecule (a heterodimer) contains two chains, a 560 residue chain known as p66 and a second chain consisting of the initiatory 440 residues of p66 called p51. The discovered inhibitors of HIV-1 RT can be placed into two categories, namely; nucleoside analogs which serve as DNA chain terminator and the non-nucleoside analog inhibitors (Ren et al., 1995). The structure of RT has been crystallized in many different ways with different resolutions. Some of those crystal structures are complexed with Nevirapine, non-nucleoside inhibitors, duplex DNA or without an inhibitor (Jaeger et al., 1998).

2.3 Protease (PR)

HIV-1 Protease plays an important role in the life cycle of the causative agent of AIDS, (HIV). It slices the newly created polyproteins in the right places for the purpose of developing a mature protein. HIV-1 PR is a homodimer belonging to the aspartyl-protease family with about 99 amino acids in each chain (Cherqaoui et al., 2017). It is another essential target for anti-HIV remedy and has approximately nine inhibitors approved by FDA. Those inhibitors include saquinavir, amprenavir, lopinavir, ritonavir, indinavir, nelfinavir, atazanavir, tipranavir and darunavir (da Cunha et al., 2009), (Vyas et al., 2017). The latest and most potent of these drugs is Darunavir (DRV). It occupies four of the eight subsites the HIV-1 protease active site can be characterized into, thus maki8ng it very potent (Paulsen, Leidner, Ragland, Kurt Yilmaz, & Schiffer, 2017).

2.4 Overview of HIV-1

The acquired immunodeficiency syndrome (AIDS) is being introduced as a result of a life threating disease that targets the human immune system known as the type-1 Human Immunodeficiency Virus (HIV-1). HIV measures 0.1 µm in diameter and has a spherical shape, with an outer coat (envelope) composing of two layers of fatty molecules called lipids. It has the ability to enter and exit the host cell through the Lipid

Raft (a special area of the cell membrane) (S., S., A., & N.S.H.N., 2007). The virus begins its life cycle by attaching its gp120 envelope protein to the CD4 molecule of the host cell with a very high affinity. The three specific enzymes (RT, IN, and PR) of the virus work very closely in order to achieve a complete life cycle, thus, they are known to be the key focus for the development of potent inhibitors, **Figure 2.3** (Cherqaoui et al., 2017)(Pommier et al., 2005). The viral capsid of the HIV virus particle which contains the primary capsid proteins p24, the nucleo-capsid protein p7/p9, the enzymes reverse transcriptase, integrase, and protease, as well as the single-stranded RNA genome, is encircled by the matrix and the virion envelope, (Zeinalipour-Loizidou et al., 2007).

The primary target of HIV-1 is the memory $CD4^+$ T cells because of their ability to express C-C chemokine receptor type 5 (CCR5) – the second target for HIV-1 treatment. In the absence of the CCR5 gene, HIV-1 does exhibit a high affinity with the target cells, thus, individuals living with a mutation due to deletion in their CCR5 gene, cannot be affected by HIV-1 (Hosseini et al., 2016).

HIV transmission can be cataloged into three stages in the life of an infected individual. The stage that results immediately after the initial infection is termed as the primary stage. The second and third stages are the asymptomatic and symptomatic or AIDS stages, respectively. Without treatment in the third stage, an infected individual can live for at most one or two years (Lutambi, 2015).

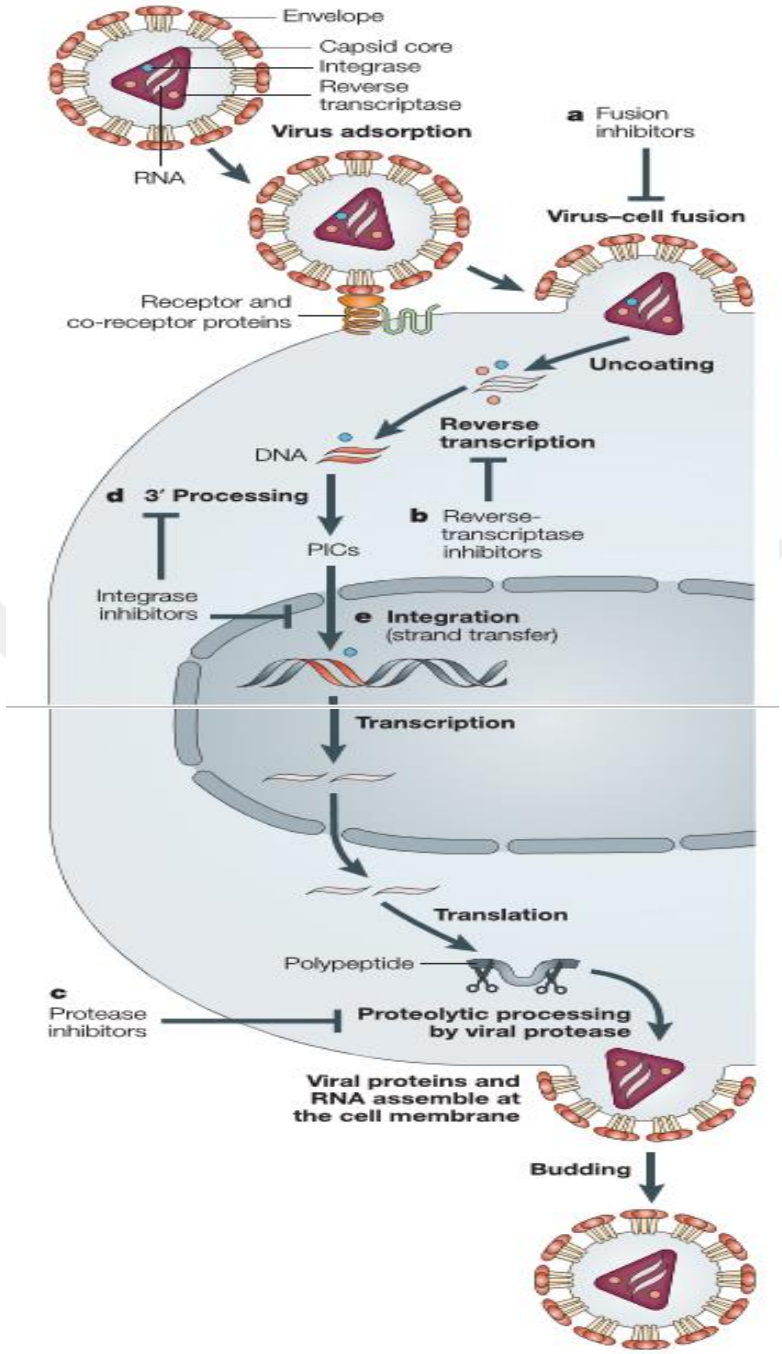


Figure 2.3: HIV-1 life cycle and drug targets, extract from a 2005 article (Pommier et al., 2005):

Figure (**2.3a**) portrays the fusion of the virus with the host cell in the absence of the fusion inhibitor, (**2.3b**) upon entry, the virus makes a copy of itself in the absence of reverse transcriptase inhibitor. (**2.3d**) It subsequently enters the cell genome if not

blocked by integrase inhibitor and (2.3**c**) finally completes its life cycle by leaving the host affected cell in the absence of protease inhibitor.

2.5 Treatment

Currently, four chemotherapeutic agents are mainly responsible for the prevention of the replication of HIV-1. They consist of reverse transcriptase inhibitors (RTI), protease inhibitors (PI), integrase inhibitors (INI) and fusion inhibitors (Maddali, Kumar, Marchand, Pommier, & Malhotra, 2011).

Early antiretroviral therapy (ART) has a very high beneficial effect for people living with HIV/AIDS on a very large scale. With the WHO recommended early ART and the joint United Nations Programme on HIV/AIDS 90-90-90 (status of 90% of people living with HIV (PLHIV) to be known, 90% PLHIV to be on ART and 90% of PLHIV to be suppressed virologically) targets to be achieved in 2020, another great reduction in the percentage of PLHIV is expected (Dorward et al., 2017)(Oliveira, 2017). Inconceivable progress has been achieved since the launch of the 90-90-90 targets to present. Based on the foundation laid by high-income countries across the world as well as large-scale projects in rural and urban settings in eastern and southern Africa, this target can be achieved globally by 2020 (Joint United Programme on HIV/AIDS (UNAIDS), 2017).

Many HIV-1 drugs have been developed over the year through computer-aided drug design, making it a key contributor to modern drug development. Currently, more than thirty HIV-1 drugs have been approved by FDA targeting different stages of the HIV life cycle. Some basic targets and inhibitors include a viral entry (fusion inhibitors), reverse transcription (RT inhibitors), integration (IN inhibitor) and viral maturation (PR inhibitors). Combinations of these inhibitors play a very important role in controlling HIV infection. On the other hand, most of these drugs result in failure due to cross-resistant strands of the virus and toxicity of the drug (Zhan, Pannecouque, De Clercq, & Liu, 2016). At present, approximately 30 percent of people living with HIV are still unaware of their status despite the presence of the universal access to knowledge of HIV

status by the World Health Organization and the Joint United Nations program on HIV/AIDS (Who & Unaids, n.d.).

The shame that comes as a result of being aware may be responsible for the huge number of unaware individuals living with HIV/AIDS. Shame serves as a barrier to effective treatment, good care, and even public health policy. The fear of being stigmatized after a positive HIV test result, on the other hand, may also serve as a contributing factor to unawareness. The implementation of policy to eliminate shame and stigmatization will aid in the reduction of the huge number of people living with HIV/AIDS that are unaware of their status (Hutchinson & Dhairyawan, 2017).

The opportunistic disease 'Tuberculosis (TB)' is known to be the most common reason for which people living with HIV worldwide are dying today. TB normally develops as a result of a weak immune system mainly caused by HIV infection (Setiawaty, 2015).

Interestingly today, HIV-negative individuals can follow the two developed platforms of antiviral administration meant for the reduction of HIV-1 infection risk. They include, pre-exposure prophylaxis (PrEP), which is for seronegative individuals who are at risk of contracting the virus (male or female workers) and post-exposure prophylaxis (PEP), which can be taken immediately after having unsafe sex with a partner, or after coming in contact with an HIV positive sample (Chupradit et al., 2017).

2.6 Approved HIV-1 drug

Raltegravir (RAL), elvitegravir (EVG) and dolutegravir (DTG) are current inhibitors that block specific proteins responsible to transfer the virus DNA into a healthy cell's DNA, which therefore prevent HIV from replicating. These inhibitors are known as integrase strand transfer inhibitors (INSTIs) (Oliveira, 2017). RAL and EVG were the first IN inhibitors to be approved but over the years have shown significant cross-resistance. However, RAL is normally taken two times a day while EVG is taken along with a pharmacokinetic booster like cobicistat or ritonavir. On the other hand, DTG being the latest inhibitor of IN is still being observed. With its long plasma half-life, it is normally taken without a booster, according to pharmacokinetic studies. It inhibits the

process by which the infected $CD4^+$ lymphocyte is being inserted into the host chromosomes. In combination with other antiretroviral drugs, the ability for the virus to create resistance becomes almost impossible (Jiang et al., 2016).

There are many others reverse transcriptase, protease as well as fusion drugs approved by FDA over the years, **Table 2.1** (Pommier et al., 2005) (S. et al., 2007).

Table 2.1 Available HIV-1 Inhibitors Approved By FDA:

FDA approval	Brand name	Generic name	Manufacturer
Fusion inhibitors			
2003	Fuzeon	Enfuvirtide(T-20)	Roche Pharmaceuticals & Trimeris
Protease Inhibitors			
1995	Invirase	Saquinavir	Roche Pharmaceuticals
1996	Norvir	Ritonavir	Abbott Laboratories
1996	Crixivan	Indinavir (IDV)	Merck
1997	Viracept	Nelfinavir	Pfizer
1997	Fortovase	Saquinavir Mesylate	Roche Pharmaceuticals
1999	Agenerase	Amprenavir	GlaxoSmithKline
2000	Kaletra	Lopinavir +Ritonavir	Abbott Laboratories
2003	Reyataz	Atazanavir	Bristol-Myers Squibb
2003	Lexiva	Fosamprenavir	GlaxoSmithKline
Nucleoside reverse transcriptase inhibitors			
1987	Retrovir	Zidovudine (AZT)	GlaxoSmithKline
1991	Videx	Didanosine	Bristol-Myers Squibb
1992	Hivid	ZaLCITABINE	Roche Pharmaceuticals
1994	Zerit	Stavudine	Bristol-Myers Squibb
1995	Epivir	Lamivudine	GlaxoSmithKline
1997	Combivir	Lamivudine+Zidovudine	GlaxoSmithKline
1998	Ziagen	Abacavir	GlaxoSmithKline
2000	Trizivir	Abacarvir+Iamivudine+Zidovudine	GlaxoSmithKline
2000	Videx EC	Didanosine	Bristol-Myers Squibb
2001	Viread	Tenofovir disoproxil	Gilead Sciences
2003	Emtriva	Emtricitabine	Gilead Sciences
2004	Epzicom	Abacavir+Lamivudine	GlaxoSmithKline
2004	Truvada	Emtricitabine+Tenofovir	Gilead Sciences
Non-nucleoside reverse transcriptase inhibitors			
1996	Viramune	Nevirapine	Boehringer Ingelheim

1997	Rescriptor	Delavirdine	Pfizer
1998	Sustiva	Efavirenz	Bristol-Myers Squibb

Table 2.1 shows the year of approval, brand name, generic name, and manufacturer extracted from a 2005 article (Pommier et al., 2005). It categorized the inhibitors according to the points of inhibitions as well as the year of approval by FDA, brand name, generic name and manufacturer.

2.7 Computer-Aided Drug Design

The introduction of computer-aid in the development of drugs is currently serving as a boost to pharmaceutical industries and researchers. In recent studies, the computer-based method is aiding researchers by eliminating the screening of irrelevant compounds, thus, saving time and money. Today, ligands are obtained from large library of compounds such as the Zinc, Otava Chemical Libraries among others or by using the computational approach in joining atoms (Shore, 2012).

2.8 Docking Studies

One of the purposes of using the computational approach of designing a drug is to predict the binding affinity of a complex through docking. There are many scoring functions that are aimed at predicting the interaction between a ligand and its protein via docking. Some of which include, CHEM Score, AUTODOCK4, FRESNO, GOLD SCORE and many more. Computationally, docking requires finding the orientation and conformation of a molecule that applies to the global minimum free binding energy. Docking may take place in the following forms, flexible protein-ligand docking, flexible protein-protein flexible ligand-rigid protein or hydrophobic docking (Shore, 2012).

2.9 Toxicity and ADME Properties

Toxicity and ADME analysis play a major role in drug discovery today. With accurate predictions, the chance for drug candidates to pass clinical trials is very high (Gupta et al., 2013).

Accepting compounds that obey the toxicity and ADME properties can be shown to be promising drug candidates. Filtering molecules to determine their drugs like status can help in early preclinical development which aid in avoiding cost at the late-stage preclinical and clinical failures (Lipinski, 2004).

3. MATERIALS AND METHOD

The objectives of this research are to be achieved through a computer simulation (*In Silico*) approach. The RSCB protein data bank (PDB) was used to obtain the crystallized structure of our targeted HIV-1 integrase enzyme. Using the docking protocol in Accelrys Discovery Studio, the enzyme was prepared for docking.

During this study, two approaches were applied in order to obtain promising inhibitors for HIV-1 IN. At the beginning, ligands were designed using discovery studio visualizer. Secondly, the OTAVA's Chemical Library was used to obtain ligands that could serve as promising inhibitors of IN.

Thousands of ligands were obtained from the OTAVA's Chemical Library and converted to PDB files using discovery studio for the purpose of docking via PyRx and Autodocktools.

PyRx was first used to dock the converted ligands; those with a binding affinity of -8.00 Kcal/mol or lower were selected for further docking using autodocktools. Ligands not remaining in the confine of the threshold (-8.00 Kcal/mol) were not further processed. Inhibitors with a low inhibition constant (Ki) were taken to discovery studio for the generation of 2D and 3D structures.

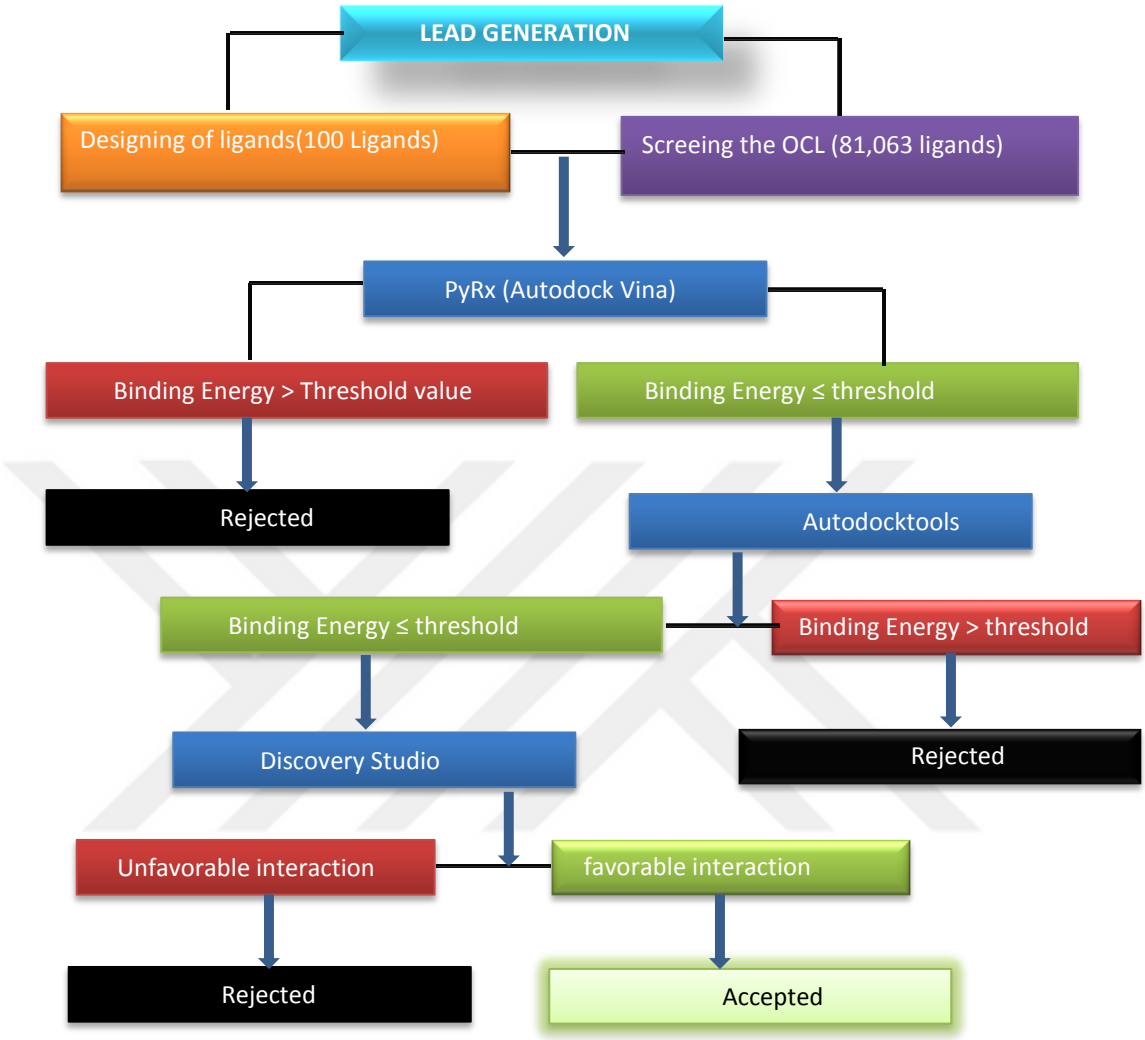


Figure 3.1: Procedure for the selection of potent inhibitor:

For a ligand to be considered as a potent inhibitor for IN, it must meet all requirements listed in **Figure 3.1**

3.1 Method 1: Designing of Ligands

All ligands were sketched and saved with a PDB extension using the ligand [1-(5-chloroindol-3-yl)-3-hydroxy-3-(2H-tetrazol-5-yl)-propenone] (**Figure 3.2**) as a lead compound obtained from our targeted enzyme. Systemically, elements and fragments were added as well as removed from this lead compound for the purpose of interacting

with basically important residues of IN and also with the aim of yielding lowest possible binding energy. Newly designed ligands were optimized using the ligand preparatory protocol in Discovery Studio Visualizer after the addition of hydrogen. Each prepared ligand was tested using Lipinski Rule of Five, a drug-likeness tool that takes into account, number of hydrogen bond donors and acceptors, molecular weight, molar refractivity, and lipophilicity (http://www.scfbio-iitd.res.in/software/drugdesign/lipinski.jsp).

A total of 100 ligands were designed and tested for their drug-likeness property and finally docked.

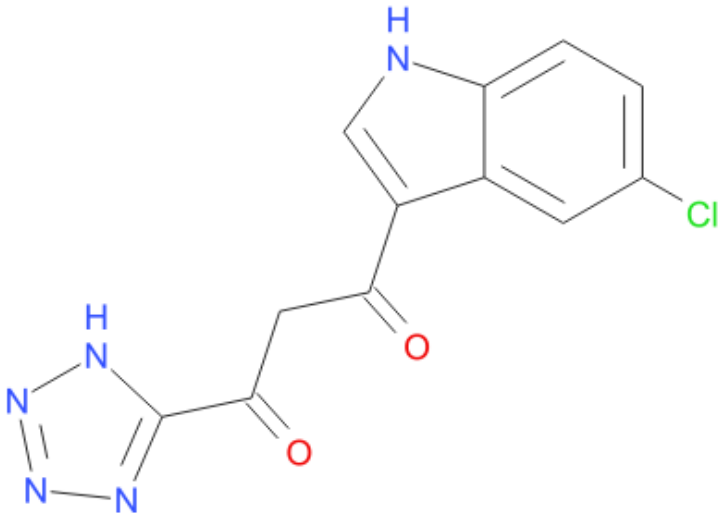


Figure 3.2: Native ligand extracted from PDB code 1qs4 to be used as a lead compound for designing of inhibitors.

3.2 Method 2: Screening the OTAVA's Chemical Library

About 81,063 lead-like compounds were taken from the OTAVA's Chemical Library (a chemical library that was established in 1997 and offers more than 270,000 compounds for high-throughput screening (HTS) including libraries of fragments, lead-like, drug-like and other compounds) for docking. The lead-like library was selected based on its low molecular weight, lipophilicity and fewer hydrogen bond acceptor, thus enabling the compounds to possess drug-like properties at the same time the ability to undergoing further lead optimization (http://www.otavachemicals.com/products/compound-libraries-for-hts/lead-like-library). All compounds were converted to PDB extension,

while the clean geometry option in discovery studio was used to optimize its geometry with a fast, dreiding-like forcefield.

3.3 Enzyme Preparation

The enzyme used in this study (1QS4.PDB; resolution: 2.1 ^{o}A (Yehudagoldgur et al., 1999) was obtained from the Protein Data Bank (https://www.rcsb.org/pdb/explore/explore.do?structureId=1qs4), an online database that contains protein and nucleic acids. This data bank was developed by the RCSB PDB, a global resource that is aimed at improving research and education in biology and medicine (Berman, 2000). The enzyme contains three chains (A, B and C) with one ligand in chain A and three Mg^{2+} in each of the chains. Additionally, it has a total of 154 amino acid residues ranging from 56 to 209, forming three domains with the following missing residues, -indicated by residue name, chain and sequence number respectively; ILE A 141, PRO A 142, TYR A 143, ASN A144, **Figure 3.3** ILE B 141, PRO B 142, ILE C 141, PRO C 142, TYR C 143.

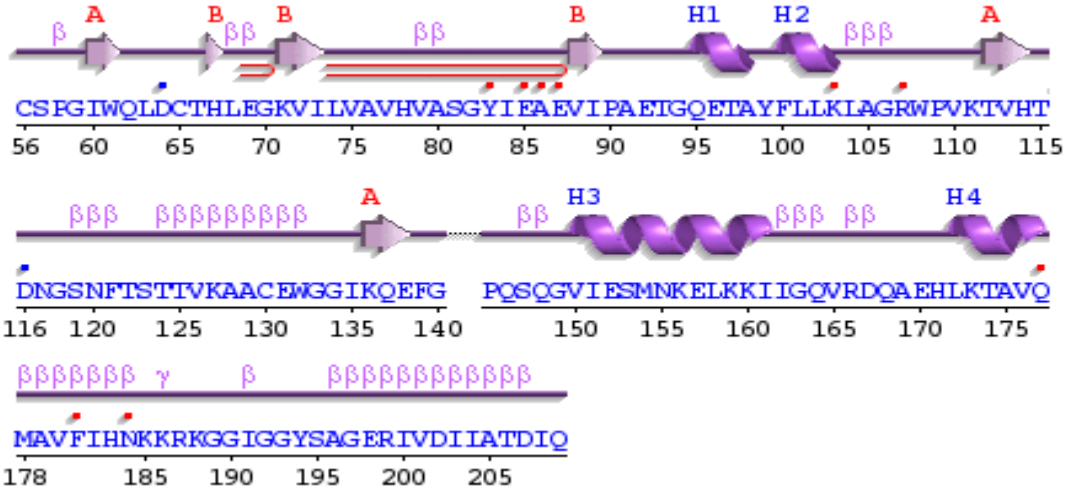


Figure 3.3: Secondary Structure Of Chain-A With Missing Residues.

This secondary structure of chain-A shows all missing residues ranging from 141 to 144; it was generated using PatchDock (a geometry-based molecular docking algorithm that is aimed at obtaining docking transformations thus, producing good molecular shape complementarity (Schneidman-Duhovny, Inbar, Nussinov, & Wolfson, 2005)) and PDBsum Generate (https://www.ebi.ac.uk/thornton-

srv/databases/pdbsum/Generate.html) respectively. The figure contains helices (H1,H2), and strands by their sheets A; Motifs: beta turn, gamma turn and beta hairpin; Residue contacts: red dots (to ligand) and blue dots (to metal).

To maintain the conformation of the protein, all three chains along with the Mg ions were prepared.

During the minimization process, water molecules were removed along with the native ligand from the original structure, while hydrogen atoms were added. The protein was therefore cleaned, the missing residuals were inserted, loops were minimized, and finally, the protein was protonated using discovery studio (**Figure 3.4**).

The insertion of the missing residues (especially residues 141-144) is important for this study because these residues are believed to be near the substrate during the integration process (Schames et al., 2004).

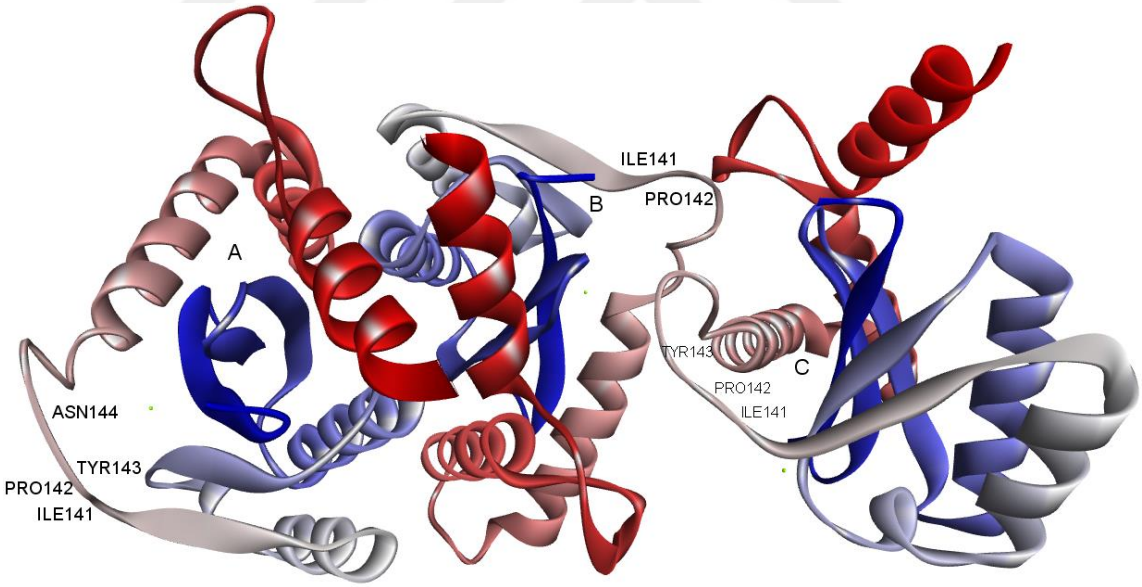


Figure 3.4: Minimized IN Enzyme With Inserted Missing Residues:

It was btained from DSV diagrammatized the full minimized structure of IN with inserted missing residues in chains A,B and C.

3.4 PyRx

In order to obtain the best protein-ligand binding affinity, AutoDock Vina of PyRx (https://pyrx.sourceforge.io)-a free academic virtual screening software was used to dock the screened ligands to the IN enzyme. The easily used docking wizard of PyRx and its ability to predict the binding affinity of a complex makes it an important tool in Computer Aided Drug Design (CADD) (Paper, 2017).

A set of fifty (50) ligands were uploaded to PyRx along with the enzyme for docking. The grid box size (XYZ) was increased to 60 Å X 60 Å X 60 Å dimensions and moved to the site defined for binding of the ligand in the crystal structure (1QS4). Of all the ligands (sketched and those obtained for the OTAVA's Library) a total of 1500 passed the first phase and were taken to autodock for further evaluation.

3.5 Autodocktools

In an effort to obtain inhibitors that meet the requirement of this study, AutoDockTools was used to generate the necessary files meant to calculate the energy scores and Ki values. The tool is designed to predict the binding energy of small molecules like substrates or drug to a enzyme or a known 3D structures (Prakash, 2017). AutoDock is one of the molecular modeling simulation software that is effective for protein-ligand docking and uses AutoDockTools for setting up and running. The newest version of AutoDock (AutoDock 4.2) was used during this study because its semi-empirical force field contains an improved unbound state model to estimate the free energy of binding as well as its updated desolvation term (Kobeissy & Nemer, 2017). Docked files (gpf and dpf) along with the map files were produced using the default sets in AutoDockTools while autogrid4 and autodock4 were used to calculate the free energy of binding and the inhibition constant (Ki) of each compound using the command line interface terminal. The XYZ coordinates of the grid box center and grid box size were given as; -17, 29.51, 66.53 and 70 Å X 70 Å X 70 Å dimensions respectively. The XYZ coordinates of the grid box center were generated using an online docking server (https://www.dockingserver.com/web/docking/). The default settings for the others were considered.

AutoDockTools however, does not consider magnesium to have 2^+ charge, instead zero charge. Because of the importance of magnesium ion in IN, the plus two charge was added manually to Mg in the PDBQT file before generating the dpf file. We used the Lamarckian genetic algorithm (LGA) to find the optimal ligand binding conformation which is known to be one of the standards and efficient methods implemented in AutoDock 4.2(Kobeissy & Nemer, 2017).

Upon completion of the 1500 compounds obtained from the previous docking in PyRx, 25 compounds from the Otava's Chemical Library along with 1 designed compound were selected based on their energy values in addition to their mode of interactions with the enzyme.

With further evaluation in Discovery Studio visualizer, three compounds from Otava chemical library and one designed compound finally met the study requirement.

4. RESULTS AND DISCUSSION

A complex with docked result having low estimated free energy of binding (-8.00Kcal/mol or below) is considered a good inhibitor for the enzyme 1QS4. However, the ligand [1-(5-chloroindol-3-yl)-3-hydroxy-3-(2H-tetrazol-5-yl)-propenone] **Figure 3.2** which was extracted from the crystallized structure of our targeted enzyme yielded a binding energy greater than -8.00Kcal/mol (the threshold) after being prepared **Figure 4.1** and docked to its native enzyme. According to a research published in the journal 'Natural Product Research' in July 2017 (Righi et al., 2017), the estimated free energy of binding of the compound with the best value was -6.55 kcal/mol, corresponding to an inhibition constant of 15.72 μM. This implies that our selected threshold of compounds will make a better contribution to inhibiting the HIV IN enzyme because of low free energy of binding and k_i.

4.1 Result

In the first approach of developing potent inhibitors for our enzyme, fragments were added as well as removed from the native ligand in order achieve interactions with some specific amino acids. After multiple of modifications to the native ligand, **Figure 4.2** was obtained, docked and tested for its drugs-like properties.

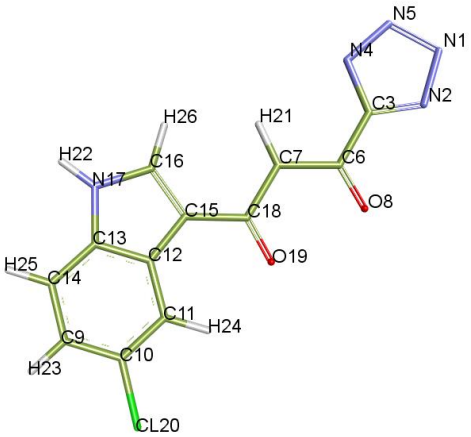


Figure 4.1: Prepared native ligand for docking and modification

A very careful and sensitive modification was done on **Figure 4.1** which led to the generation of a potent HIV-1 IN inhibitor. Specific modifications done can be seen in **Table 4.1.**

Table 4.1 Modification Done On Figure 4.1

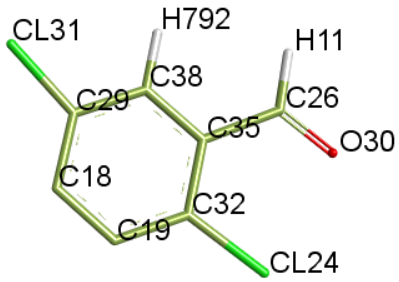
Element removed	Element removed and replaced		Fragment added to C9 and C14
C18 and O19	Removed	Replaced	
	C7	**Oxygen**	
	Cl 20	**Aldehyde** 	

Table 4.1 shows the summary of the approach used in designing a potent ligand (method 1) where specific structures were added removed or replaced. The result obtained after the removal of C18 and O19, the replacement of C7 with oxygen and Cl

20 with an aldehyde, as well as the addition of a ring containing chlorine atoms along with an aldehyde show a significant improvement in the free energy of binding.

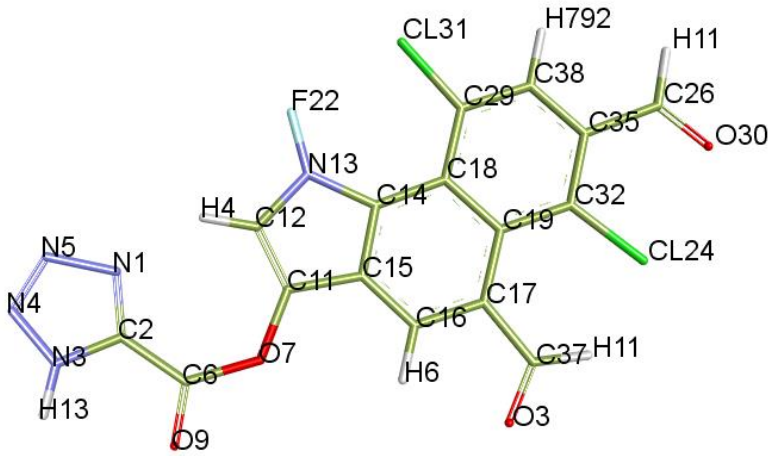


Figure 4.2: Designed IN Inhibitor:

This figure shows a modified structure of **figure 4.1** meant for the inhibition of HIV-1 IN. It is named (**6,9-dichloro-1-fluoro-5,7-diformyl-1H-benzo[g]indol-3-yl 1H-tetrazole-5-carboxylate).** This name was generated using the free academic version of ChemSketch (http://www.acdlabs.com/resources/freeware/chemsketch/).

When docked into the cavity of the minimized structure of 1QS4, using autogrid4 and autodock4 respectively, **Figure 4.2** obtained a calculated free energy of binding value of -8.44kcal/mol at temperature 298.15 K with an estimated inhibition constant, Ki= 652.83 nM (nanomolar). A complete 2D and 3D structures showing important interactions with major amino acids of IN can be seen in **Figures 4.3** and **11a&b** respectively.

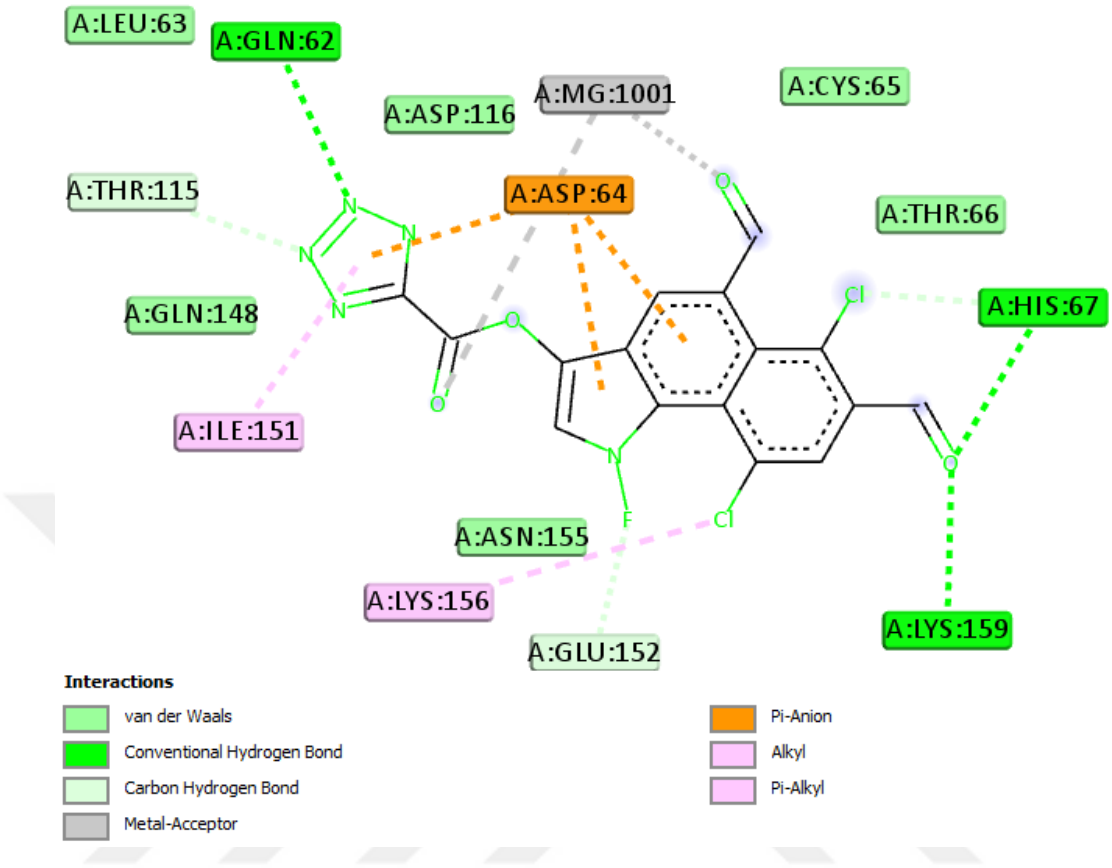


Figure 4.3: 2D structure of our designed inhibitor in complex with the targeted enzyme:

It depicts the 2D structure of a complex with multiple interactions.

It can be seen that the most important amino acids of IN (GLU 152, ASP 64 and ASP 116) are having carbon hydrogen bond, Pi-Anion and Van der waals interactions respectively. At the same time, the co-factor (MG 1001) is observed to have a metal-Acceptor interaction. On the other hand, ILE 151 and LYS 156 are showing Pi-Alkyl and Alkyl interactions respectively. LYS 159, HIS 67, GLN 62 are indicating conventional Hydrogen Bond interactions; THR 115 has Carbon Hydrogen Bond interaction while THR 66, CYS 65, LEU 63' GLN 148 and ASN 155 are forming Van der waals interactions. The ligand is displayed as a line with carbon colored black and all other elements colored in green.

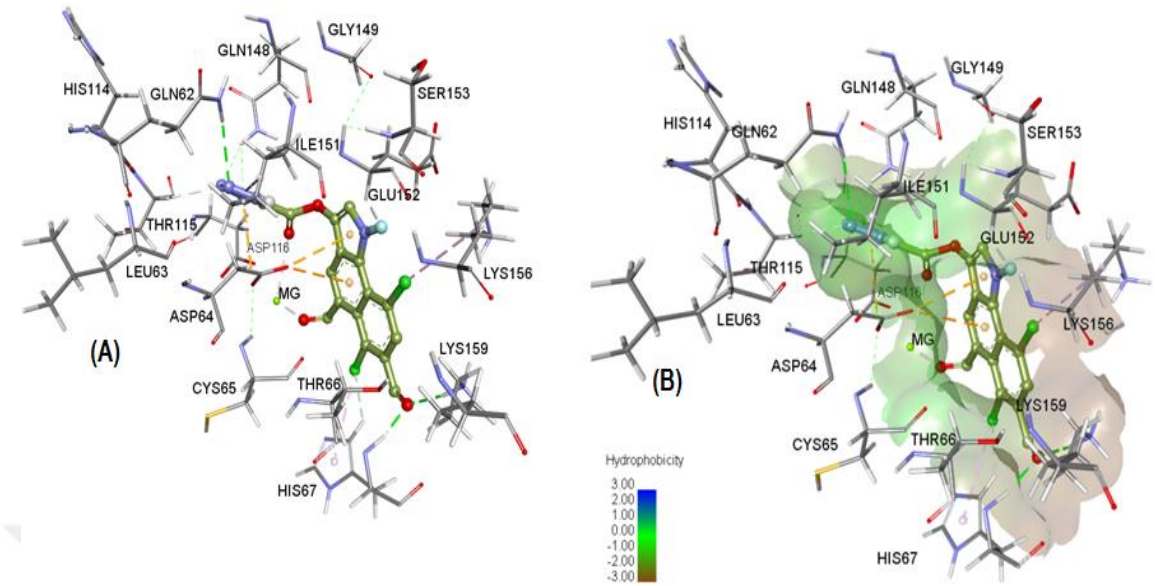
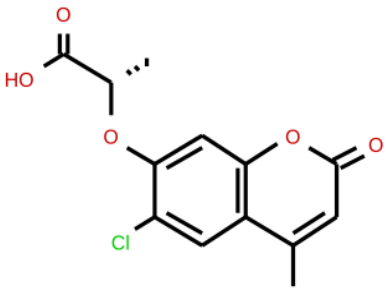
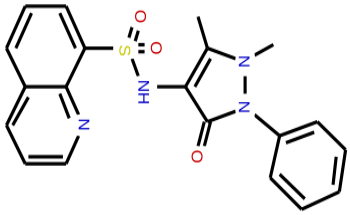
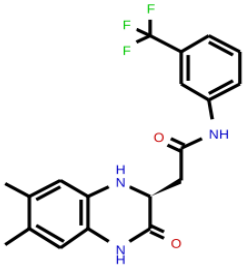
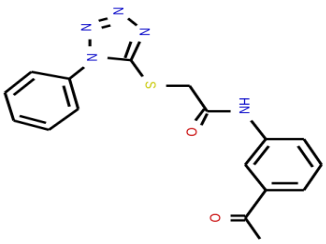
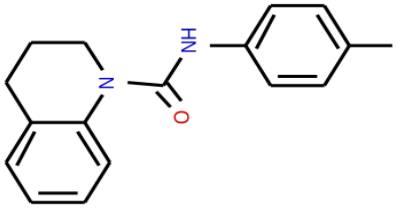


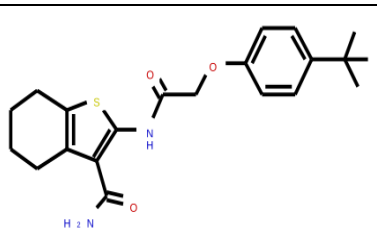
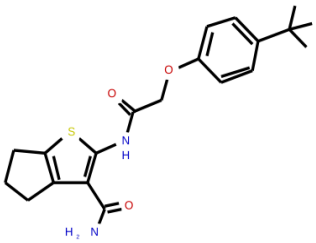
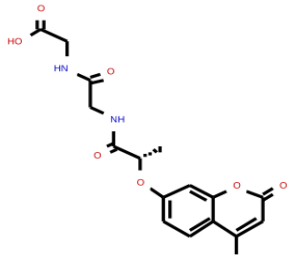
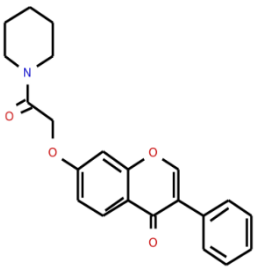
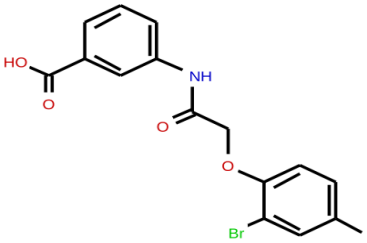
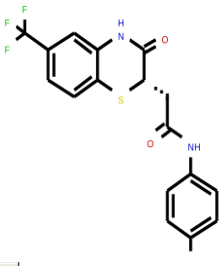
Figure 4.4 A& B: 3D Structures of Figure 4.3:

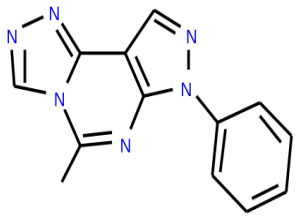
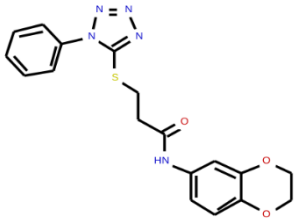
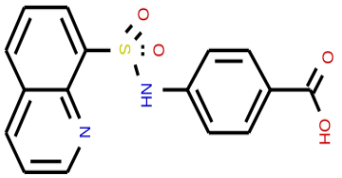
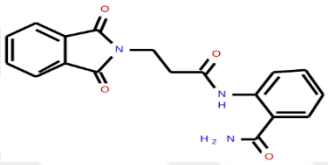
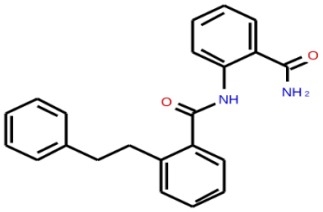
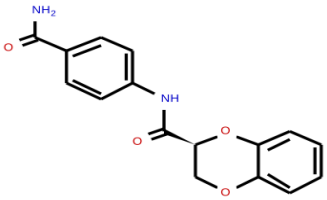
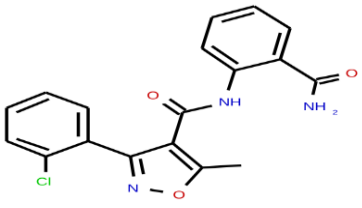
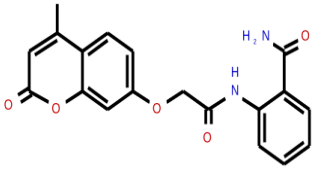
The inhibitor is displayed as a ball and stick while all interacting amino acids are displayed as stick and labeled in black. **Figure 4.4 B** shows the hydrophobicity of the ligand; the scale shows the ligand to be between 0.00 and 1.00, indicating that the inhibitor is hydrophobic (an acceptable result).

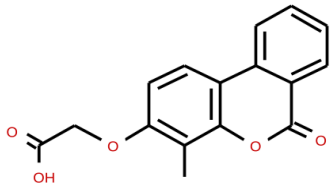
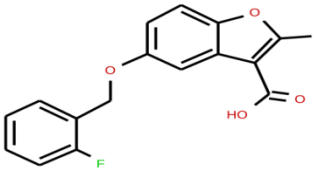
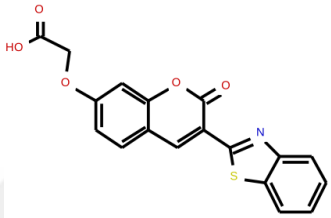
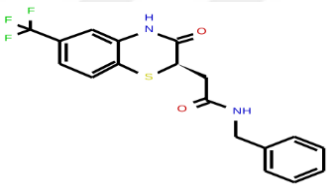
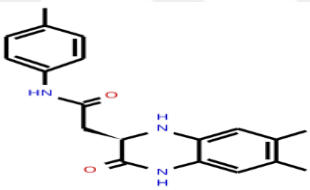
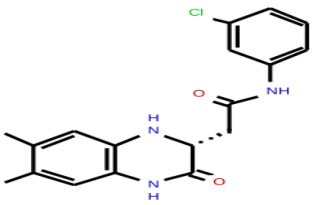
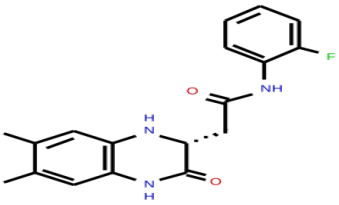
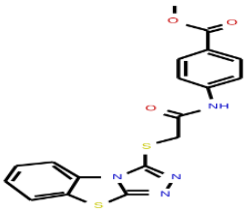
Using the Otava's Chemical Library for further generation of inhibitors, the best 87 ligands of the 1,500 lead-like compounds obtained from PyRx (Autodock Vina) are displayed in **Table 4.2** along with their energy values and Otava's code.

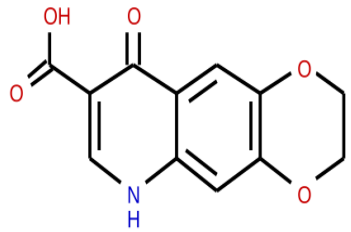
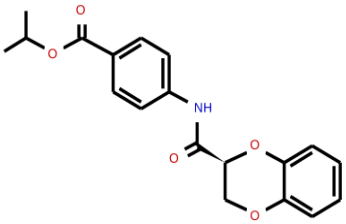
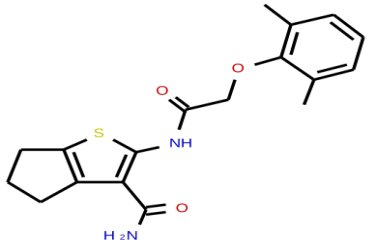
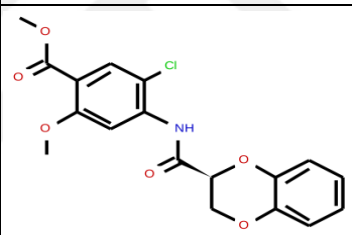
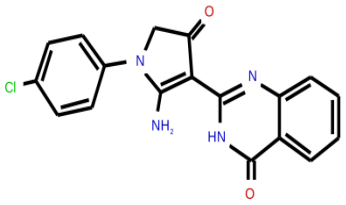
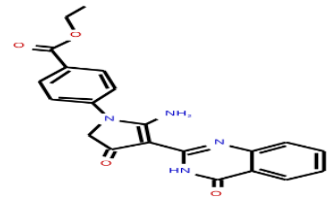
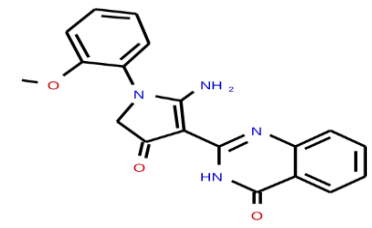
Table 4.2 Best 87 Of 1500 Selected Compounds

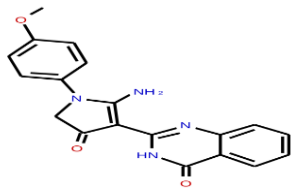
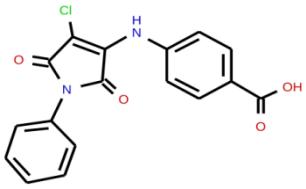
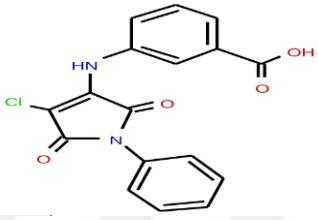
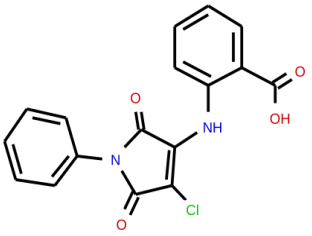
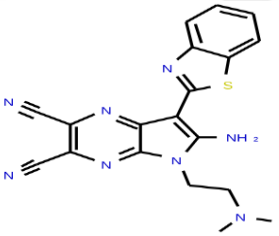
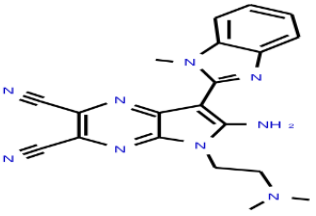
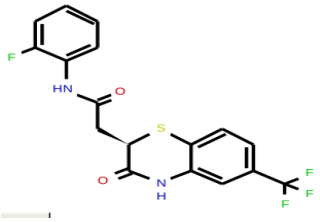
No	Otava's Code	Binding Energy -in Kcal/mol	2D Structure
1	105980018	-8.6	
2	107050035	-8.8	
3	107050060	-8.9	
4	107070112	-8.9	
5	107210181	-8.6	

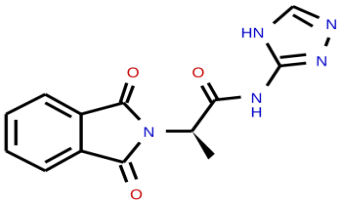
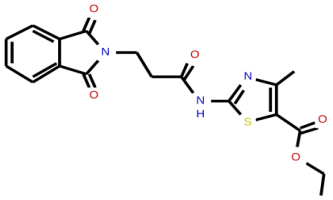
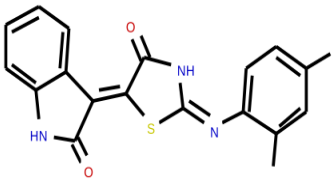
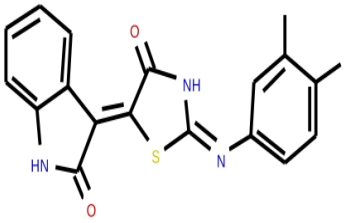
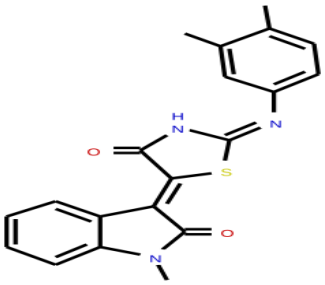
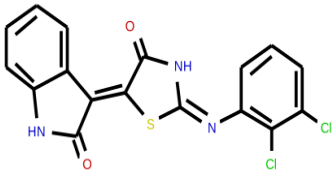
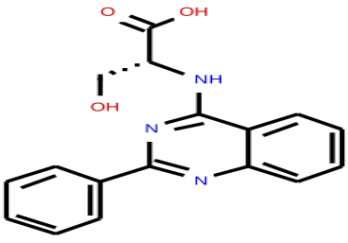
6	107210806	-9.0	
7	107210807	-9.0	
8	107240106	-8.4	
9	107300032	-9.2	
10	107320240	-8.5	
11	107490019	-9.1	

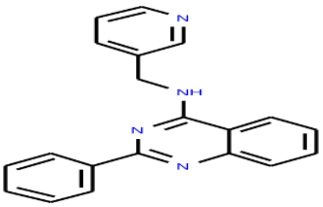
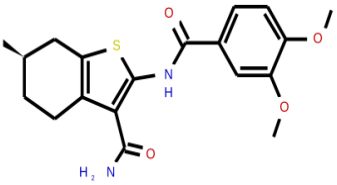
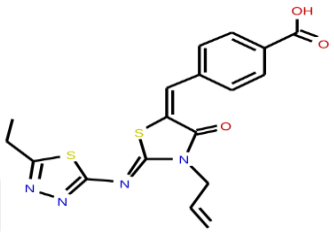
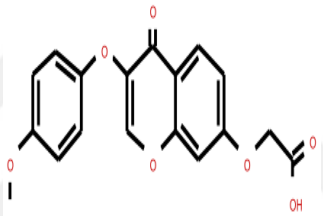
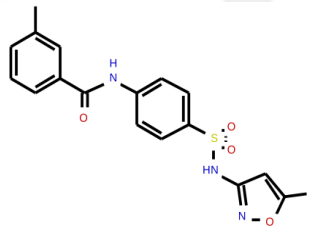
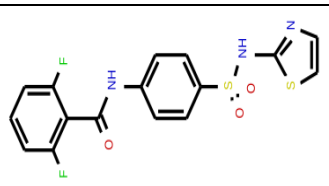
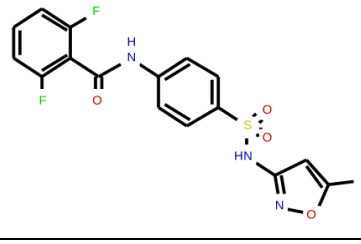
12	107510012	-8.5	
13	107510017	-8.7	
14	107600046	-8.7	
15	107620099	-8.6	
16	107620110	-8.7	
17	107620482	-8.8	
18	107620499	-8.9	
19	107620500	-10.2	

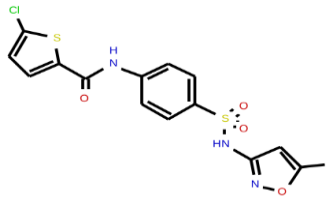
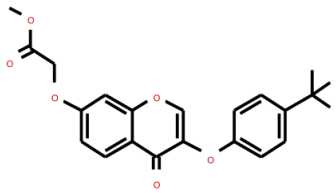
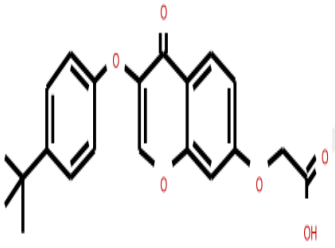
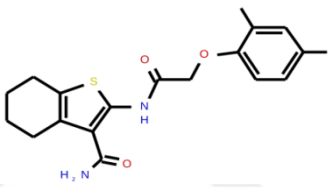
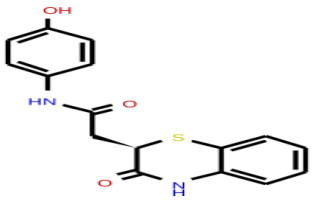
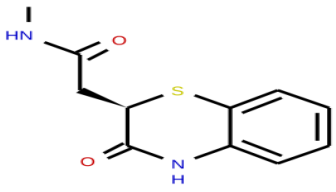
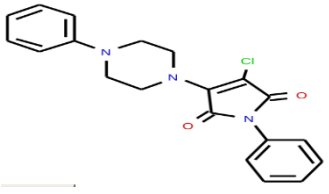
20	107640012	-8.1	
21	107710049	-8.0	
22	107710124	-9.0	
23	107780007	-8.6	
24	107780027	-8.6	
25	107780028	-8.8	
26	1077800029	-8.8	
27	107790026	-8.6	

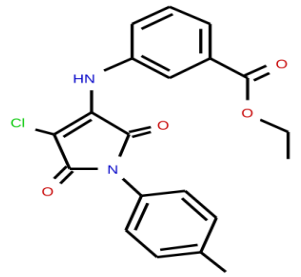
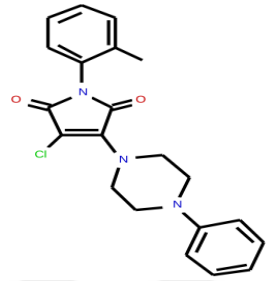
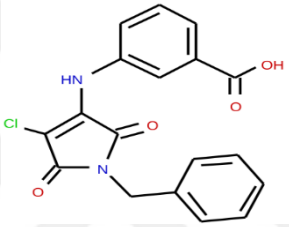
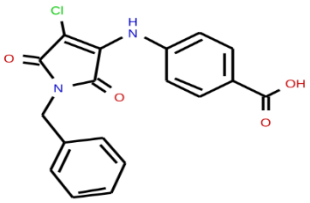
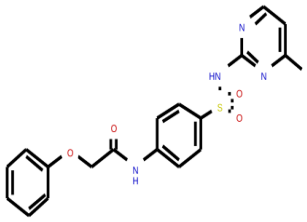
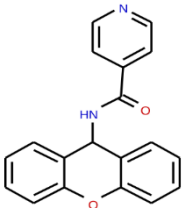
28	107830117	-8.0	
29	107930183	-8.9	
30	107930667	-8.9	
31	107930994	-8.7	
32	108470140	-8.9	
33	108470145	-9.4	
34	108470147	-9.5	

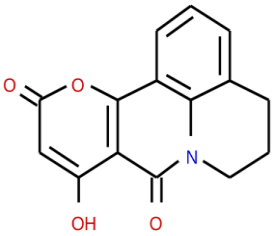
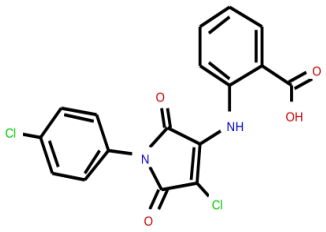
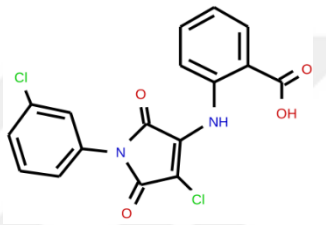
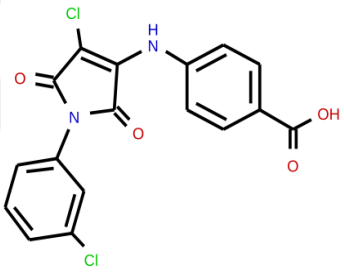
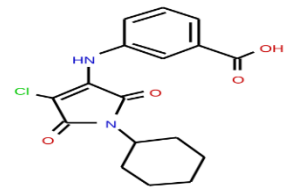
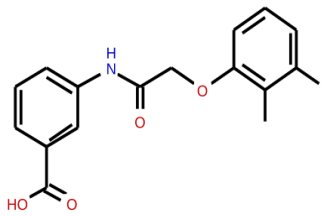
35	108470148	-8.9	
36	108470168	-8.3	
37	108470169	-8.5	
38	108470170	-8.2	
39	108470244	-8.9	
40	108470246	-9.0	
41	108480037	-9.1	

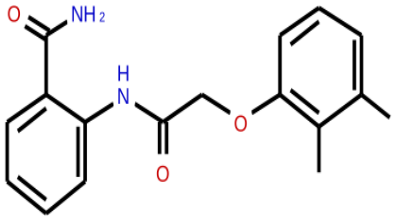
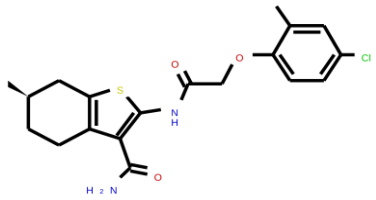
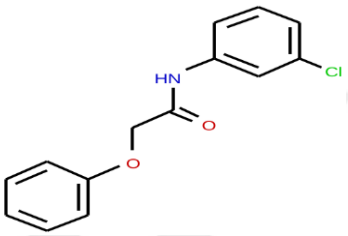
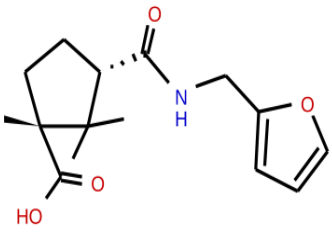
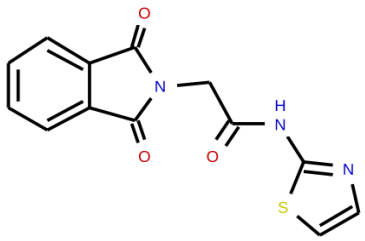
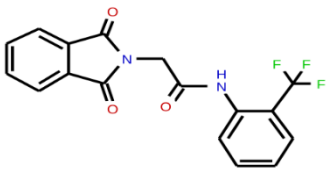
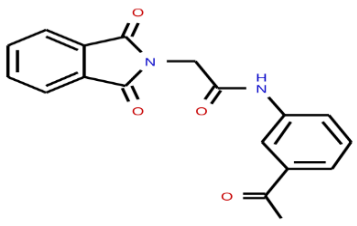
42	108560159	-8.6	
43	108560562	-8.7	
44	108580147	-9.4	
45	108580161	-9.4	
46	108580162	-9.5	
47	108580176	-9.2	
48	108610023	-8.0	

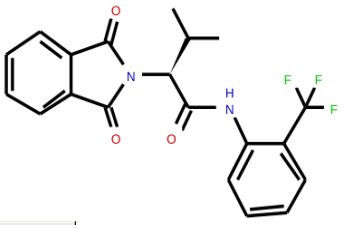
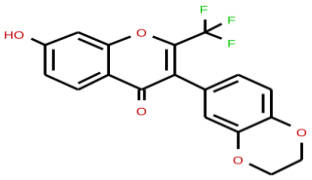
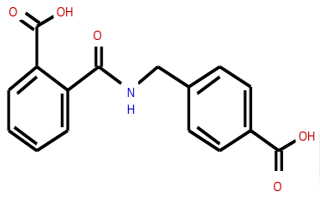
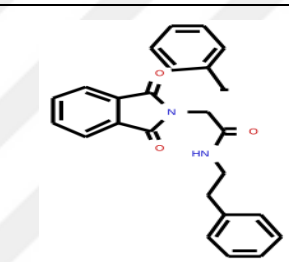
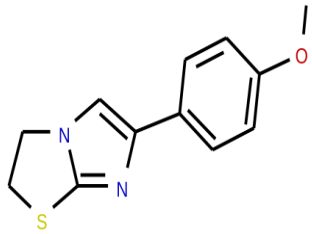
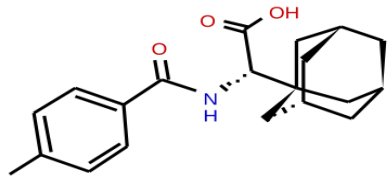
49	109080052	-8.9	
50	109160389	-8.8	
51	109240128	-8.4	
52	109250044	-8.2	
53	109280008	-9.1	
54	109280050	-8.6	
55	109280054	-9.0	

56	109280104	-8.6	
57	109350043	-8.8	
58	109350045	-9.0	
59	109370805	-8.6	
60	109390004	-8.8	
61	109390008	-9.1	
62	109420046	-8.6	

63	109420132	-8.7	
64	109420282	-8.9	
65	109420342	-8.4	
66	109420344	-8.3	
67	109450020	-8.6	
68	109740015	-8.6	

69	109740023	-8.9	
70	109750155	-8.3	
71	109750264	-8.6	
72	109750269	-8.5	
73	109750322	-8.4	
74	109770138	-8.2	

75	109770150	-8.2	
76	109770706	-8.9	
77	110420022	-8.6	
78	110420076	-8.8	
79	110420131	-8.6	
80	110490073	-8.6	
81	110490075	-9.3	

82	110490097	-8.6	
83	110510042	-8.8	
84	110520176	-8.1	
85	110830248	-9.4	
86	111050005	-8.0	
87	111150115	-8.3	

All compounds listed in **Table 4.2** successfully passed through the first stage of this study and can therefore be further modified to produce a better inhibition constant. Interestingly, twenty five (25) of these compounds were capable of passing through the

second stage of the study and can be seen in **Table 4.3** along with their inhibition constant (Ki) which were not calculated by PyRx.

Table 4.3 Best 25 Ligands Exhibiting Promising Characteristics Of Being Potent Inhibitors

No	Otava's code	Lowest Binding Energy (Kcal/mol)	Estimated Inhibition constant Ki, in (nM)
1	105980018	-9.45	117.50
2	107240106	-10.61	16.74
3	107320240 *	-9.39	131.70
4	107600046	-10.12	38.18
5	107640012	-9.71	76.21
6	107710049	-9.43	122.16
7	107710124	-9.28	158.12
8	107830117	-8.74	389.26
9	108470168	-10.03	44.71
10	108470169	-10.74	13.45
11	108470170	-9.56	98.94
12	108610023	-10.92	9.90
13	109240128	-8.73	400.3
14	109250044	-8.54	553.86
15	109350045	-9.74	72.71
16	109420342	-9.74	72.31
17	109420344	-8.99	257.24
18	109750155 *	-10.03	44.19
19	109750264	-9.98	48.30
20	109750269	-10.00	46.77
21	109750322	-10.37	24.92
22	109770138	-8.59	503.83
23	110520176	-9.94	51.54
24	111050005	-9.41	125.73
25	111150115 *	-8.74	395.19

Among the best 87 ligands obtained from PyRx, twenty five (25) were selected (**Table 4.3)** based on their consistency in yielding minimum binding energy as well as their

ability to interact with specific amino acids or the co-factor of IN. (* = inhibitor with best interactions).

Of the 25 inhibitors, the best three displaying interactions with at least two of three most important amino acids are shown in **Figures 12a&b, 13a&b and 14a&b.**

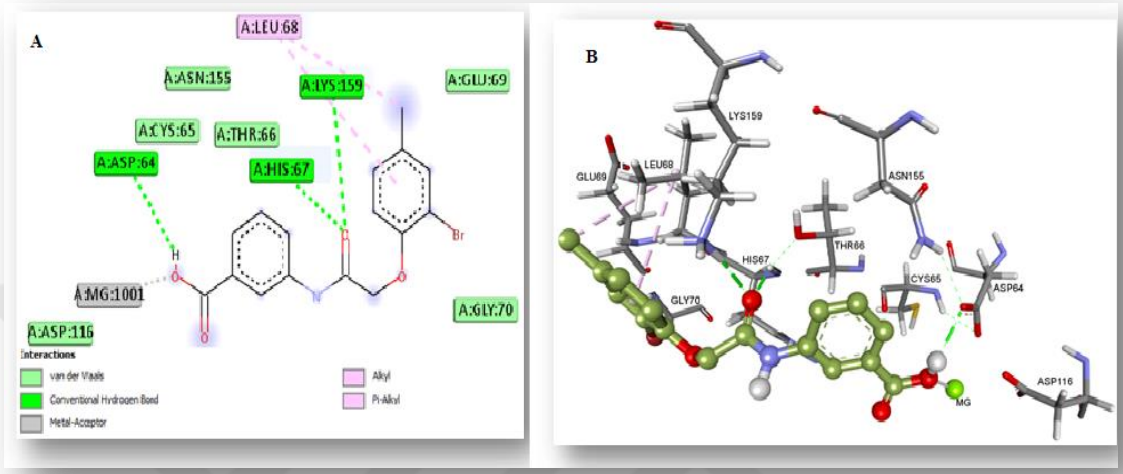


Figure 4.5 A and B: 2D and 3D structures of otava's lead-like library code: 107320240 respectively

Figure 4.5A: This structure was selected as a result of its direct interaction with ASP 64 and Mg 1001. ASP 116, another important amino acid is also observed to have a Van Der Waals interaction. ASP and Mg 1001 are observed to have conventional hydrogen bond interaction and mental-acceptor interaction respectively. The remaining amino acids which play significant role in obtaining a better inhibition constant satisfactory for the study are known to have alkyl and pi-alkyl interactions (LEU 68), conventional hydrogen bond interactions (HIS 67 & LYS 159), and van der Waals interactions (CYS 65, THR 66, GLU 69, GLY 70 & ASN 155).

Figure 4.5b: 3D structure of 107320240**:** Shows the inhibitor (ball and stick) in the active site of IN with important interactions labeled. The thick bonds indicate ligand-amino acid interactions while the thin binds indicate interactions between amino acids. The inhibitor is shown as a ball and stick, with each element given a specific color while the amino acids are displayed as sticks and labeled.

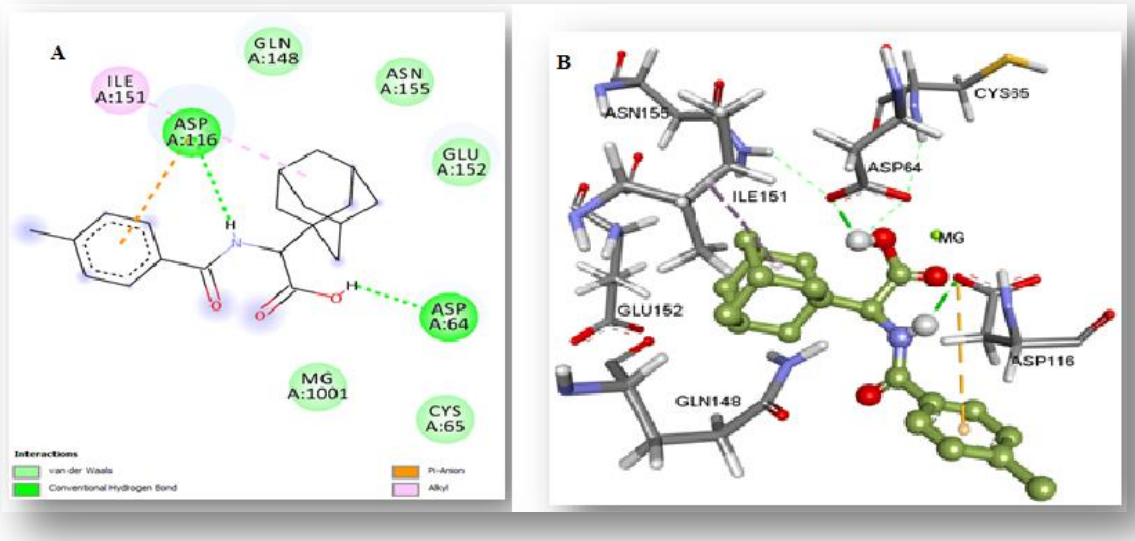


Figure 4.6 A and B: 2D and 3D structures of otava's lead-like library code: 111150115 respectively

Figure 4.6a: 2D structure of 111150115 in complex with the enzyme**:** Illustrates the direct interactions of two wanted amino acids (ASP 116 and ASP 64). Both are shown to have pi-anion and conventional hydrogen bond interactions (ASP 116) and a conventional hydrogen bond interaction (ASP 64). On the other hand, GLU 152 along with GLU 148, ASN 155, Mg 1001, and CYS 65 are displaying a Van Der Waals interaction while ILE 151 is showing alkyl interaction.

Figure 4.6b | 3D structure of 111150115: In this diagram, the inhibitor is displayed as (ball and stick) in the active site of the protein while the important amino acids are labeled in black in stick representations.

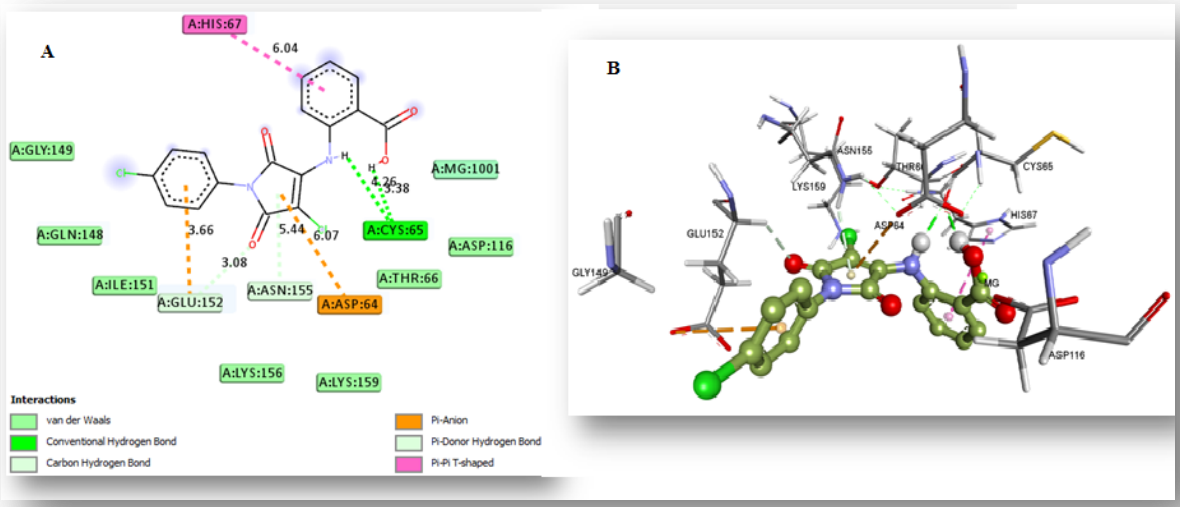


Figure 4.7 A and B: 2D and 3D structures of otava's lead-like library code: 109750155 respectively

Figure 4.7a: 2D structure of 109750155 in complex with the enzyme**:** Illustrates the direct interactions of two desired acid amino acids (ASP 64 and GLU 152). ASP 64 is displaying pi-anion interaction while GLU 152 is showing a Carbon Hydrogen Bond and Pi-Anion interactions. HIS 67, ASN 155 and CYS 65 are displaying Pi-Pi T-shaped, carbon Hydrogen Bond and conventional Hydrogen Bond interactions respectively. The remaining amino acids (GLY 149, GLN 148, ILE 151, LYS 156, LYS 159, ASP 116) and Mg 1001 are showing van der Waals interactions.

Figure 4.7b: 3D structure of 109750155**:** The inhibitor is displayed as a ball and stick with atoms represented by specific colors. All amino acids are displayed as sticks and labeled.

4.1.1 Lipinski rule of five

In addition to providing the potency of our selected inhibitors, the Lipinski rule of five was used to determine whether those ligands have drug-like properties or not. For a molecule to be classified as drug-like, it must possess the following; molecular mass less than 500 Dalton, high lipophilicity (expressed as LogP less than 5), less than 5 hydrogen bond donors, less than 10 hydrogen bond acceptors, and molar refractivity

should be between 40-130 (Lipinski, 2004). The filter was done for the best four inhibitors at pH value 7 and can be seen in **Table 4.4.**

Table 4.4 Lipinski Rule Of Five Drugs-Like Predictions

No	Inhibitor	mass	Hydrogen bone donor	Hydrogen bond acceptor	LogP	Molar refractivity
1	Designed	381.0	1	7	2.5	95.34
2	107320240	364.0	1	5	-0.240	75.77
3	111150115	326.0	1	4	2.06	88.05
4	109750155	335.0	1	6	1.62	90.77

Table 4.4 further proved that the selected molecules possess drugs-like properties as a result of complying with all the rules of Lipinski. An online server was used to obtain these results (http://www.scfbio-iitd.res.in/software/drugdesign/lipinski.jsp#anchortag).

4.2 Discussion

Comparatively, all selected inhibitors of this study seem to be better than the native inhibitor of our targeted enzyme. Specifically, many amino acids are observed to have interacted with the four selected inhibitors and have provided docked calculated estimated energy of binding greater than that of the native inhibitor (both the experimental value from the literature and that obtained from our docking). On the other hand, inhibitions of some specific amino acids and co-factor that can lead to the termination of the HIV-1 life cycle were observed to have a perfect inhibition with our selected ligands. For example, the native inhibitor of IN interacted with just one of the key amino acids (ASP:64) while those of ours; the designed inhibitor interacted with at least two of the three key amino acids (specifically, ASP:64 and GLU:152) along with the co-factor (Mg:1001). The remaining three selected inhibitors obtained from the Otava's chemical library are observed to have interacted with the enzyme as shown in **Table 4.5:**

Table 4.5: Important Observed Interactions

No	Inhibitor's code	Significant amino acid(s) or co-factor
1	107320240	ASP:64 and Mg:1001
2	111150115	ASP:64 and ASP:116
3	109750155	ASP:64 and GLU:152
4	Native ligand	ASP:64

According to **Table 4.5,** all selected inhibitors interacted with ASP: 64 along with one other important amino acid or the co-factor. This result indicates that the selected inhibitors are more potent and could be good drugs candidates of integrase.

5. CONCLUSION

Conclusively, HIV-1 integrase is shown to be one of the best targets of the four points of inhibition for HIV-1. Obtaining a potent integrase's inhibitor seems to be very difficult and time-consuming; however, we are certain that the four selected compounds will serve as a possible treatment for HIV-1 if taken for further processing.

Many compounds were designed but could not meet the requirement of the study with the exception of one, (**Figure 4.2**). This compound is shown to be a promising inhibitor for IN based on those great interactions which were not displayed by the native ligand of our selected enzyme.

On the other hand, generating many promising inhibitors was of major concern, which led to the screening a huge library of compounds. Of the many screened compounds, from PyRx to AutoDock4.2, three compounds were capable of meeting the study's objective (**Figure 4.5a, Figure 4.6a and Figure 4.7a**).

Although the remaining 22 compounds of the best 25 compounds obtained from autodock4.2 were not finally selected because of limited interaction or inability to interact with the key amino acids of focus, they are still considered promising IN's inhibitors. Other researchers can easily modify or test these compounds by using different protocol and yield a very good result.

The best three inhibitors obtained during this study went through many stages of screening yet proved to be promising and potent inhibitors. The Lipinski rule of five results, **Table 4.4** proved that these compounds can easily pass through early preclinical development of drug. As a result of this, these inhibitors can be taken for advance studies for the development of HIV-1 drug.

REFERENCES

Barreca, M. L., Luca, L. De, Ferro, S., Rao, A., Monforte, A., & Chimirri, A. (2006). Computational and synthetic approaches for the discovery of HIV-1 integrase inhibitors. *Arkivoc*, *2*(v ii), 224–244.

Berman, H. M. (2000). The Protein Data Bank. *Nucleic Acids Research*, *28*(1), 235–242. https://doi.org/10.1093/nar/28.1.235

Bilodeau, F., Bousquet, Y., Pesant, M., Brochu, C., Chabot, C., Halmos, T., … Yoakim, C. (2014). Discovery of BI 224436, a noncatalytic site integrase inhibitor (NCINI) of HIV-1, (August 2016). https://doi.org/10.1021/ml500002n

Cherqaoui, D., A. F. Martis, E., Joseph, B., C. Coutinho, E., Hdoufane, I., & Bjij, I. (2017). *Flexibility of Important HIV-1 Targets and in silico Design of anti-HIV Drugs*. *Current Chemical Biology* (Vol. 11). https://doi.org/10.2174/2212796811666170807154035

Chupradit, K., Moonmuang, S., Nangola, S., Kitidee, K., Yasamut, U., Mougel, M., & Tayapiwatana, C. (2017). Current peptide and protein candidates challenging HIV therapy beyond the vaccine Era. *Viruses*, *9*(10). https://doi.org/10.3390/v9100281

da Cunha, E. F. F., Sippl, W., de Castro Ramalho, T., Ceva Antunes, O. A., de Alencastro, R. B., & Albuquerque, M. G. (2009). 3D-QSAR CoMFA/CoMSIA models based on theoretical active conformers of HOE/BAY-793 analogs derived from HIV-1 protease inhibitor complexes. *European Journal of Medicinal Chemistry*, *44*(11), 4344–4352. https://doi.org/10.1016/j.ejmech.2009.05.016

Diamond, T. L., & Bushman, F. D. (2006). Role of metal ions in catalysis by HIV integrase analyzed using a quantitative PCR disintegration assay, *34*(21), 6116–6125. https://doi.org/10.1093/nar/gkl862

Dorward, J., Garrett, N., Quame-Amaglo, J., Samsunder, N., Ngobese, H., Ngomane, N., … Nigel Garrett1, 2, Justice Quame-Amaglo3, Natasha Samsunder1, Hope Ngobese4, Noluthando Ngomane4, Pravikrishnen Moodley5, Koleka Mlisana 6, 7, Torin Schaafsma3, Deborah Donnell3, Ruanne Barnabas3, 8, 9, Kogieleum Naidoo1, 10, Salim Abdool Karim1, 10, 11, 9. (2017). Protocol for a randomized controlled implementation trial of point-of-care viral load testing and task shifting: the Simplifying HIV TREAtment and Monitoring (STREAM) study. *BMJ Open*, (September). https://doi.org/http://dx.doi.org/10.1136/bmjopen-2017-017507

Ercan, S. (2017). Docking and Molecular Dynamics Calculations of Some Previously Studied and Newly Designed Ligands to Catalytic Core Domain of HIV-1 Integrase and an Investigation to Effects of Conformational Changes of Protein on Docking Results, *4*(1), 243–270.

Greene, W. C., & Peterlin, B. M. (2008). Molecular Biology of HIV: Implications for New Therapies. *Global HIV/AIDS Medicine*, (July 2014), 23–38. https://doi.org/10.1016/B978-1-4160-2882-6.50007-1

Gupta, P., Garg, P., & Roy, N. (2012). Identification of Novel HIV-1 Integrase Inhibitors Using Shape-Based Screening, QSAR, and Docking Approach. *Chemical Biology and Drug Design*, *79*(5), 835–849. https://doi.org/10.1111/j.1747-0285.2012.01326.x

Gupta, P., Garg, P., & Roy, N. (2013). In silico screening for identification of novel HIV-1 integrase inhibitors using QSAR and docking methodologies. *Medicinal Chemistry Research*, *22*(10), 5014–5028. https://doi.org/10.1007/s00044-013-0490-y

Han, Y.-S., Xiao, W.-L., Quashie, P. K., Mesplède, T., Xu, H., Deprez, E., … Wainberg, M. A. (2013). Development of a fluorescence-based HIV-1 integrase DNA binding assay for identification of novel HIV-1 integrase inhibitors. *Antiviral Research*, *98*(3), 441–448. https://doi.org/10.1016/j.antiviral.2013.04.001

Hosseini, Y., Mollica, A., & Mirzaie, S. (2016). Structure-based virtual screening efforts against HIV-1 reverse transcriptase to introduce the new potent non-nucleoside reverse transcriptase inhibitor. *Journal of Molecular Structure*, *1125*(January 2017), 592–600. https://doi.org/10.1016/j.molstruc.2016.07.040

Hutchinson, P., & Dhairyawan, R. (2017). Shame and HIV: Strategies for addressing the negative impact shame has on public health and diagnosis and treatment of HIV. *Bioethics*, (August). https://doi.org/10.1111/bioe.12378

Inhibitor, N. I., & Bailey, M. D. (2015). Preclinical Profile of BI 224436 , a Novel HIV-1, (March 2014). https://doi.org/10.1128/AAC.02719-13

Jaeger, J., Restle, T., & Steitz, T. A. (1998). The structure of HIV-1 reverse transcriptase complexed with an RNA pseudoknot inhibitor. *EMBO Journal*, *17*(15), 4535–4542. https://doi.org/10.1093/emboj/17.15.4535

Jiang, J., Xu, X., Guo, W., Su, J., Huang, J., Liang, B., … Liang, H. (2016). Dolutegravir(DTG, S/GSK1349572) combined with other ARTs is superior to RAL- or EFV-based regimens for treatment of HIV-1 infection: a meta-analysis of randomized controlled trials. *AIDS Research and Therapy*, *13*(1), 30. https://doi.org/10.1186/s12981-016-0115-x

Joint United Programme on HIV/AIDS (UNAIDS). (2017). Ending AIDS: Progress towards the 90-90-90 targets, 134. Retrieved from http://www.unaids.org/sites/default/files/media_asset/Global_AIDS_update_2017_en.pdf

Kim, H. J. (2002). Inhibition of HIV-1 reverse transcriptase and HIV-1 integrase and antiviral activity of Korean seaweed extracts, (July 2014). https://doi.org/10.1023/A

Kobeissy, F. H., & Nemer, G. (2017). Neuroproteomics, *1598*(May). https://doi.org/10.1007/978-1-4939-6952-4

Lipinski, C. A. (2004). Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discovery Today: Technologies*, *1*(4), 337–341. https://doi.org/10.1016/j.ddtec.2004.11.007

Lutambi, A. M. (2015). Modelling the Impact of Stages of HIV Progression on Estimates, (September), 101–113.

Maddali, K., Kumar, V., Marchand, C., Pommier, Y., & Malhotra, S. V. (2011). Biological evaluation of imidazolium- and ammonium-based salts as HIV-1 integrase inhibitors. *MedChemComm*, *2*(2), 143. https://doi.org/10.1039/c0md00201a

Mulder, L. C. F., Chakrabarti, L. A., & Muesing, M. A. (2002). Interaction of HIV-1 integrase with DNA repair protein hRad18. *Journal of Biological Chemistry*, *277*(30), 27489–27493. https://doi.org/10.1074/jbc.M203061200

Oliveira, M. (2017). HIV-1 strains belonging to large phylogenetic clusters show accelerated escape from integrase inhibitors in cell culture compared ..., (May). https://doi.org/10.1093/jac/dkx118

Paper, C. (2017). Cross-docking study of flavonoids against tyrosinase enzymes using PyRx 0 . 8 virtual screening tool, (March), 0–4.

Paulsen, J. L., Leidner, F., Ragland, D. A., Kurt Yilmaz, N., & Schiffer, C. A. (2017). Interdependence of Inhibitor Recognition in HIV-1 Protease. *Journal of Chemical Theory and Computation*, *13*(5), 2300–2309. https://doi.org/10.1021/acs.jctc.6b01262

Pommier, Y., Johnson, A. A., & Marchand, C. (2005). Integrase inhibitors to treat HIV/Aids. *Nature Reviews Drug Discovery*, *4*(3), 236–248. https://doi.org/10.1038/nrd1660

Prakash, J. S. (2017). IN SILICO MOLECULAR DOCKING OF KAEMPFEROL

WITH CANCER TARGET , HUMAN ESTROGEN RECEPTOR (2IOK) International Conference on Translational Medicine and Imaging (ICTMI-2017), (August).

Ren, J., Esnouf, R., Hopkins, A., Ross, C., Jones, Y., Stammers, D., & Stuart, D. (1995). The structure of HIV-1 reverse transcriptase complexed with 9-chloro-TIBO: lessons for inhibitor design. *Structure*, *3*(9), 915–926. https://doi.org/10.1016/S0969-2126(01)00226-X

Righi, G., Pelagalli, R., Isoni, V., Tirotta, I., Marini, M., Palagri, M., … Bovicelli, P. (2017). Synthesis of potential HIV integrase inhibitors inspired by natural polyphenol structures View supplementary material ... Synthesis of potential HIV integrase inhibitors inspired by natural polyphenol structures. *Natural Product Research*, (July), 1–9. https://doi.org/10.1080/14786419.2017.1354191

Rihn, S. J., Hughes, J., Wilson, S. J., & Bieniasz, P. D. (2015). Uneven Genetic Robustness of HIV-1 Integrase. *Journal of Virology*, *89*(1), 552–567. https://doi.org/10.1128/JVI.02451-14

S., G., S., R., A., G., & N.S.H.N., M. (2007). AIDS: A review of targets and approaches for treatment. *Indian Journal of Pharmaceutical Sciences*, *69*(2), 173–179. https://doi.org/10.4103/0250-474X.33139

Salam Pradeep Singh. (2012). In silico structure assessment analysis of core domain of six protein data bank entries of HIV-1 Integrase. *Journal of Computational Biology and Bioinformatics Research*, *4*(1), 1–7. https://doi.org/10.5897/JCBBR11.017

Schames, J. R., Henchman, R. H., Siegel, J. S., Sotriffer, C. A., Ni, H., & McCammon, J. A. (2004). Discovery of a Novel Binding Trench in HIV Integrase. *Journal of Medicinal Chemistry*. https://doi.org/10.1021/jm0341913

Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., & Wolfson, H. J. (2005). PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids*

Research, *33*(SUPPL. 2). https://doi.org/10.1093/nar/gki481

Setiawaty, V. (2015). The stages of HIV infection and the risk of opportunistic Tuberculosis infection, *6*(2), 121–125. https://doi.org/10.22435/hsji.v6i2.4771.121-125

Shore, H. C. (2012). Computer Aided Drug Design & QSAR, *12072*, 2012.

Srivastav, V. K., & Tiwari, M. (2013). QSAR and docking studies of coumarin derivatives as potent HIV-1 integrase inhibitors QSAR and docking studies of coumarin derivatives as potent HIV-1 integrase inhibitors. *ARABIAN JOURNAL OF CHEMISTRY*, (August 2014). https://doi.org/10.1016/j.arabjc.2013.01.015

Varadarajan, J., Mcwilliams, M. J., Mott, B. T., Thomas, C. J., Smith, S. J., & Hughes, S. H. (2016). Drug resistant integrase mutants cause aberrant HIV integrations. *Retrovirology*, 1–10. https://doi.org/10.1186/s12977-016-0305-6

Vyas, V. K., Shah, S., & Ghate, M. (2017). CHEMISTRY Generation of new leads as HIV-1 integrase inhibitors : 3D QSAR , docking and molecular dynamics simulation. *Medicinal Chemistry Research*, 532–550. https://doi.org/10.1007/s00044-016-1772-y

Wang, W., Wang, W., & Zaki, M. (2005). 5th International Workshop on Data Mining in Bioinformatics Workshop Chairs :, (September).

Who, & Unaids. (n.d.). WHO, UNAIDS statement on HIV testing services: new opportunities and ongoing challenges. Retrieved from http://aidsdatahub.org/sites/default/files/publication/UNAIDS_WHO-UNAIDS_statement_HIV-testing-services_2017.pdf

Yehudagoldgur, R., Cohen, G., Tamiofujiwara, Tomokazuyoshinaga, Toshiofujishita, Hirohikosugimoto, T., … Davies. (1999). StructureoftheHIV-1integrasecatalyticdomain complexedwithaninhibitor:Aplatformfor

antiviraldrugdesign. https://doi.org/10.1073/pnas.96.23.13040

Zeinalipour-Loizidou, E., Nicolaou, C., Nicolaides, A., & Kostrikis, L. G. (2007). HIV-1 integrase: from biology to chemotherapeutics. *Current HIV Research*, *5*(4), 365–88. https://doi.org/10.2174/157016207781023965

Zhan, P., Pannecouque, C., De Clercq, E., & Liu, X. (2016). Anti-HIV Drug Discovery and Development: Current Innovations and Future Trends. *Journal of Medicinal Chemistry*, *59*(7), 2849–2878. https://doi.org/10.1021/acs.jmedchem.5b00497

Zhu, K., Dobard, C., & Chow, S. A. (2004). Requirement for integrase during reverse transcription of human immunodeficiency virus type 1 and the effect of cysteine mutations of integrase on its interactions with reverse transcriptase. *Journal of Virology*, *78*(10), 5045–55. https://doi.org/10.1128/JVI.78.10.5045

CURRICULUM VITAE

Augustine S. Samorlu

Graduate Student
Kadir Has University
Istanbul, Turkey
+905396940324
as.samorlu@gmail.com

Personal Information

Gender: **Male**
Nationality: **Liberian**
Marital Status: **Single**
Language: **English**

Education

2016 – Present: **MSc:** Kadir Has University, Istanbul Turkey
Department of Bioinformatics and Genetics
Major: Bioinformatics and Computational Biology

Some Courses: Essential Cell Biology, Introduction to Bioinformatics Molecular Modeling and Graphics, Introduction to Programming and Algorithm, Computational Structural Biology, System Biology Computer Aided Drug Design, Organic Chemistry of Drug Design

Supervisor: Prof. Dr. Kemal Yelekçi

Thesis: *In silico* Screening of Potent HIV-1 Integrase Inhibitors for the *Treatment of Human Immunodeficiency Virus (HIV)*

Expected date of Graduation: June 30, 2018

2015 – 2016: **PGD:** International Academic and Management Association Center of Excellence (IAMA-COE), Haryana (India)
Major: Epidemiology and Biostatistics

Some Courses: Introduction to Public Health, Introduction to Epidemiology Introduction to Statistical Methods, Advanced Epidemiology Probability, Introduction to Excel etc

Supervisor: Dr. Mogana Flomo Jr.

2010 – 2014: **BSc:** Cuttington University, Bong County, Liberia
College of Natural Sciences
Major: Biology

Some Courses: Biology 101-421, Chemistry 101-371, English 101-210
Math 101-211, History, Science Research Method, Senior Thesis

Certificate

October 3 -- November 20, 2017: Certificate of Achievement, *AuthorAid INASP*

Course: Research Communication
Grade: Merit

http://moodle.inasp.info/badges/badge.php?hash=bc3717cfe94cd9c047465b4f4c17e3ee7dde7b48

April 21--22, 2015: Certificate of Completion, *International Medical Crops*

Course: Mental Health In High Schools

Honors and Awards

October 9, 2015: 2235-Graduate Scholarship for the Least Developed Countries
Scientific and Technological Research Council of Turkey (TÜBITAK)

2010 September: The President's Prestigious Scholarship, Cuttington University

2009/2010: Valedictorian…..William V. S. Tubman Gray High School

Working Experience

2014—2016: Instructor (Biology and Chemistry)
Cuttington University Campus School
Grade: 10—12

2012—2016: Instructor (Biology)
William V.S. Tubman Gray High School
Grade: 10—12

Others

National Services

December 20, 2014: Presiding Officer
National Election Commission, Liberia

November 7, 2010: Voter Identification Officer
National Election Commission, Leleh Town, Bong County, Liberia

2009/2010: Student Council President
William V.S. Tubman Gray High School

Skills

- Computer software, especially programming languages
- Relate easily to people

References

1. Prof. Dr. Kemal Yelekçi
   Department Head and Thesis Advisor
   Department of Bioinformatics and Genetics, Kadir Has University
   Contact: yelekci@khas.edu.tr +902125336532/1332
2. Asst. Prof. Dr. Hatice Bahar Şahin
   Assistant Professor, Kadir Has University
   Contact: hbahar.sahin@khas.edu.tr 02125336532
   Web: http://sites.khas.edu.tr/bioinformatics/research/h-bahar-sahin/
3. Dr. Mogana Flomo Jr.
   Executive Director
   International Academic and Management Association Center of Excellence (IAMA-COE)
   Contact: moganajr@gmail.com
4. Mr. Christain Bob Nicol
   Dean/ College of Natural Sciences, Cuttington University
   Contact: cbnicol2009@yahoo.com