HOMOLOGY MODELING OF HUMAN CCR5, 2D QSAR, AND DOCKING STUDIES OF HIV-1 INHIBITORS TARGETING CCR5 PROTEIN

M.RAVI RAJA TEJASVI **(08 – 124 -401)**

Department of Chemistry
NIZAM COLLEGE(autonomous)
Bhasheer bagh , Hyderabad.

NIZAM COLLEGE

(Bhasheer bagh , Hyderabad)

*This is to certify that the dissertation work entitled "**HOMOLOGY MODELLING OF HUMAN CCR5, 2D QSAR ,AND DOCKING STUDIES OF HIV-1 INHIBITORS TARGETING CCR5 PROTEIN**" , is a bonafide project work done by **M.RAVI RAJA TEJASVI** H.T NO: 08-124-401 from **BACHELOR OF SCIENCES in MICROBIOLOGY ,GENETICS AND CHEMISTRY, NIZAM COLLEGE**.*

The work embodied in this dissertation work has not been submitted to any other university or institution for the award of any degree.

PROJECT GUIDE
(Dr. M. Vijjulatha)

HEAD OF THE DEPARTMENT
(Chemistry)

Principal
(Nizam College)

ACKNOWLEDGEMENT

It is indeed a pleasant task to thank the people who have contributed towards the successful completion of this project.

Firstly I take this opportunity to thank my guide ***Dr. M.Vijjulatha ,*** *for her dynamic invaluable technical guidance and constant encouragement without which the project should not have been possible.*

I profoundly thank ***Mr. S. Sreekanth****, for his support, cooperation and his valuable suggestions for the successful completion of the project.*

And I would like to thank Department of Chemistry and Cheminformatics for Providing the Lab and Software facility and UGC for granting the Financial support.

Last but not least I would like to thank ***my parents*** *for their co-operation and support.*

M.RAVI RAJA TEJASVI

DECLARATION

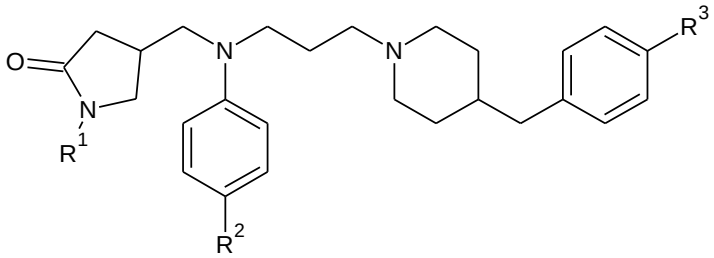
I, ***M.RAVI RAJA TEJASVI*** hereby declare that the project entitled ***"HOMOLOGY MODELLING OF HUMAN CCR5, 2D QSAR ,AND DOCKING STUDIES OF HIV-1 INHIBITORS TARGETING CCR5 PROTEIN "*** is carried out under the guidance of **Dr. M. VIJJULATHA** . This project work is submitted as UGC project work of **"BACHELOR OF SCIENCES".** This report has not been submitted for any other degree of any university or institution.

M.RAVI RAJA TEJASVI

Index

Abstract:

The derivatives of novel lead compound, N-{3-[4-(4-fluorobenzoyl)piperidin-1-yl]propyl}-1-methyl-5-oxo-N-phenylpyrrolidine-3-carboxamide , was identified as a CCR5 antagonists with an Systematic modification of the lead , focused on a series of 5-oxopyrrolidine-3- carboxamides, resulted in the Identification of compounds with improved binding affinity .The most potent CCR5 antagonists reported previously inhibited HIV-1 envelope-mediated membrane fusion with IC50 values of 0.44, 0.19, and 0.49µM, respectively. This gave us an interest in developing the new compounds by systemic modification on the skeleton of these compounds. For this, we have done 2-dimensional QSAR, and new molecules were constructed based on the structural properties of the old molecules .The activity of the newly designed molecules were predicted based on the structural–activity relation of the existing molecules. The existing molecules (38) were divided into training set and test set. The developed model based on training set containing 29 molecules gave R^2_{PRED} value of 0.836 and cross validation Q^2 value of 0.54 and standard error of estimate 0.333 The cross validation on test set gave R^2 value of 0.837. Here, we report comparative docking studies of newly constructed molecules with that of the previously existing CCR5 antagonists using homology based model of CCR5 based on the crystal structures of bovine Rhodopsin (1F88.pdb).The homology modeling was done using MODELLER 8v2 and the Docking studies were done using GLIDE 4.0. Thus, from the entire 52 newly constructed molecules which were docked, we got best 4 of them with optimal GLIDE Score. (Compound 37: -5.30, Compound 39: -5.24, Compound 44: -4.84 and Compound 47: -4.67). Thus from the Complex scoring and binding ability its deciphered that these newly constructed oxo pyrolidine derivatives could be promising inhibitors for HIV-1 using CCR5 as Drug target .Yet pharmacological studies have to confirm it.



Core Structure of 5- oxo pyrollidine derivatives.

1. INTRODUCTION :

1.1 What is AIDS ?

In 1981 an increased occurrence of unusual cases of *Pneumocystis carinii* pneumonia and Kaposi´s cancer together with other opportunistic infections, was observed among previously healthy homosexual men and intravenous drug abusers in the USA. An underlying immunosuppression was found to be promoter of these rare diseases. This syndrome became known as Acquired Immune Deficiency Syndrome (AIDS). In 1983 the causative agent of AIDS was identified as a human retrovirus, first isolated in France from a patient with multiple lymphadenopathies, a condition linked to AIDS, and subsequently in 1984, from AIDS patients. Initially, three different names were given to the virus isolated from AIDS patients; human T lymphotropic virus III (HTLVIII), lymphadenopathy-associated virus (LAV),7 and AIDS-associated retrovirus (ARV). Eventually the AIDS-causing virus was in 1986 given an alternative name, human immunodeficiency virus (HIV).A few years later a second similar virus, HIV-2, was isolated from patients in West Africa. Both HIV subtypes can lead to AIDS, although the pathogenic course with HIV-2 might be longer. The genome homology of HIV-1 and HIV-2 are approximately 40%.

Retrospective studies indicate that the first documented case of AIDS occurred in Central Africa in 1959 and the source of the virus is proposed to come from the same geographic area. The origin of the two viruses has now been shown to be derived from two African monkeys, the chimpanzee (*Pan troglodytes troglodytes*) for HIV-1 and the sooty mangabay (*Cercocebus atys*) for HIV-2.

1.2 Introduction to HIV types, groups and subtypes :

HIV is a highly variable virus which mutates very readily. This means there are many different strains of HIV, even within the body of a single infected person. Based on genetic similarities, the numerous virus strains may be classified into types, groups and subtypes.

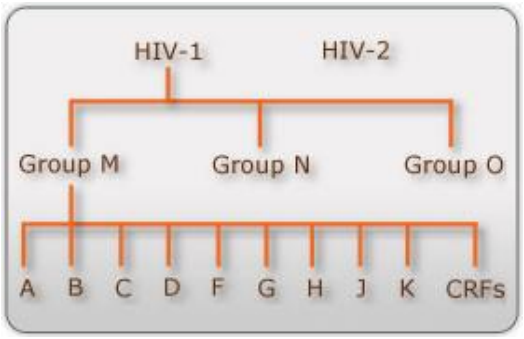
There are two **types** of HIV:

i. HIV-1 (discovered in 1983)

ii. HIV-2(discovered in 1986)

Both types are transmitted by sexual contact, through blood, and from mother to child,and they appear to cause clinically indistinguishable AIDS. However, it seems that HIV-2 is less easily transmitted, and the period between initial infection and illness is longer in the case of HIV-2. Worldwide, the predominant virus is HIV-1, and generally when people refer to HIV without specifying the type of virus they will be referring to HIV-1. The relatively uncommon HIV-2 type is concentrated in West Africa and is rarely found elsewhere.

1.2.1 How many subtypes of HIV-1 are there?



This diagram illustrates the different levels of HIV classification. Each type is divided into groups, and each group is divided into sub types and CRFs.

The strains of HIV-1 can be classified into three groups: the "major" group M, the "Outlier" group O and the "new" group N. These three groups may represent three separate introductions of simian immunodeficiency virus into humans. Group O appears to be restricted to west-central Africa and group N - discovered in 1998 in Cameroon - is extremely rare. More than 90% of HIV-1 infections belong to HIV-1 group M. Within group M there are known to be at least nine genetically distinct **subtypes** (or clades) of HIV-1. These are subtypes A, B, C, D, F, G, H, J and K. Occasionally, two viruses of different subtypes can meet in the cell of an infected person and mix together their genetic material to create a new hybrid virus (a process similar to sexual

reproduction, and sometimes called "viral sex"). Many of these new strains do not survive for long, but those that infect more than one person are known as "circulating recombinant forms" or **CRFs**. For example, the CRF A/B is a mixture of subtypes A and B. The classification of HIV strains into subtypes and CRFs is a complex issue and the definitions are subject to change as new discoveries are made. Some scientists talk about subtypes A1, A2, A3, F1 and F2 instead of A and F, though others regard the former as sub subtypes.

1.2.2 What about subtypes E and I.

One of the CRFs is called A/E because it is thought to have resulted from hybridization between subtype A and some other "parent" subtype E. However, no one has ever found a pure form of subtype E. Confusingly, many people still refer to the CRF A/E as "subtype E" (in fact it is most correctly called CRF01_AE). A virus isolated in Cyprus was originally placed in a new subtype I, before being reclassified as a recombinant form A/G/I. It is now thought that this virus represents an even more complex CRF comprised of subtypes A, G, H, K and unclassified regions. The designation "I" is no longer used.

1.2.3 Where are the different subtypes and CRFs found?

The HIV-1 subtypes and CRFs are very unevenly distributed throughout the world, with the most widespread being subtypes A and C. Subtype A and CRF A/G predominate in West and Central Africa, with subtype A possibly also causing much of the Russian epidemic. Historically, subtype B has been the most common subtype/CRF in Europe, the Americas, Japan and Australia. Although this remains the case, other subtypes are becoming more frequent and now account for at least 25% of new infections in Europe. Subtype C is predominant in Southern and East Africa, India and Nepal. It has caused the world's worst HIV epidemics and is responsible for around half of all infections. Subtype D is generally limited to East and Central Africa. CRF A/E is prevalent in South-

East Asia, but originated in Central Africa. Subtype F has been found in Central Africa, South America and Eastern Europe. Subtype G and CRF A/G have been observed in West and East Africa and Central Europe. Subtype H has only been found in Central Africa; J only in Central America; and K only in the Democratic Republic of Congo and Cameroon.

1.3 What's the Difference between HIV-1 and HIV-2

According to the Center for Disease Control...

- HIV-2 is transmitted in the same ways as HIV-1: Through exposure to bodily fluids such as blood, semen, tears and vaginal fluids.
- Both HIV-1 and HIV-2 can develop into AIDS or what the Center for Disease Control calls "opportunistic infections" (otherwise milder infections that can become serious or deadly against a weakened immune system).
- Immunodeficiency (or a weakened immune system) develops more slowly and is milder in persons with HIV-2.
- People infected with HIV-2 are less infectious in the early stages of the virus than those with HIV-1.
- The infectiousness of HIV-2 increases as the virus progresses. Most reported cases of HIV-2 are found in West Africa.
- There are few reported cases of HIV-2 in the United States.

1.4 Symptoms of AIDS and HIV Infection

Immediately following infection with HIV, most individuals develop a brief, nonspecific “viral illness” consisting of low grade fever, rash, muscle aches, headache and/or fatigue. Like any other viral illness, these symptoms resolve over a period of five to 10 days. Then for a period of several years (sometimes as long as several decades), people infected with HIV are asymptomatic (no

symptoms). However, their immune system is gradually being destroyed by the virus. It can take as short as a year to as long as 10 to 15 years to go from being infected with HIV to "full-blown" AIDS.

According to the Center for Disease Control and Prevention, a person is considered to have AIDS when they have a T cell count (also called CD4 cell count) of 200 or less (healthy T cell levels range from 500 to 1500) or they have an AIDS-defining condition. The AIDS-defining conditions are:

- Candidiasis
- Cervical cancer (invasive)
- Coccidioidomycosis , Cryptococcosis, Cryptosporidiosis
- Cytomegalovirus disease
- Encephalopathy (HIV-related)
- Herpes simplex (severe infection)
- Histoplasmosis
- Isosporiasis
- Kaposi's sarcoma
- Lymphoma (certain types)
- Mycobacterium avium complex
- Pneumocystis carinii pneumonia
- Pneumonia (recurrent)
- Progressive multifocal leukoencephalopathy
- Salmonella septicemia (recurrent)
- Toxoplasmosis of the brain
- Tuberculosis
- Wasting syndrome

People who are not infected with HIV may also develop these diseases; the presence of any one of these conditions does not mean the person has AIDS. To be diagnosed with AIDS, a person must be infected with HIV. Some people infected with HIV may develop a disease that is less serious than AIDS, referred to as AIDS Related Complex (ARC). ARC is a condition caused by the AIDS virus in which the patient tests positive for AIDS infection and has a specific set of clinical symptoms. However, ARC patients' symptoms are often less severe than those with classic AIDS because the degree of destruction of the immune system has not progressed as far as it has in patients with classic AIDS. Symptoms of ARC may include loss of appetite, weight loss, fever, night sweats, skin rashes, diarrhea, and tiredness, lack of resistance to infection or swollen lymph nodes.

***Note**:* Not everyone who has been infected with HIV develops AIDS. Very rarely, some individuals can be infected with HIV yet maintain normal immune function and general good health even after 20 years of infection.

1.5 Transmission Factors

Exposure Route	**Estimated infections per 10,000 exposures**
Blood Transfusion	9,000
Childbirth	2,500
Needle-sharing injection drug use	67
Percutaneous needle stick	30
Receptive anal intercourse*	50
Insertive anal intercourse*	6.5
Receptive penile-vaginal intercourse*	10
Insertive penile-vaginal intercourse*	5
Receptive oral intercourse*§	1
Insertive oral intercourse*§	0.5
* assuming no condom use § source refers to oral intercourse performed on a man	

Three main transmission routes for HIV have been identified. HIV-2 is transmitted much less frequently by the mother-to-child and sexual route than HIV-1.

1.5.1 Other routes

HIV has been found at low concentrations in the saliva, tears and urine of infected *individuals*, but there are no *recorded* cases of infection by these secretions and the potential risk of transmission is negligible.

1.5.2 Multiple infections

Unlike some other viruses, infection with HIV does not provide immunity against additional infections, particularly in the case of more genetically distant viruses. Both inter- and intra-clade multiple infections have been reported, and even associated with more rapid disease progression. Multiple infections are divided into two categories depending on the timing of the acquisition of the second strain. Co infection refers to two strains that appear to have been acquired at the same time (or too close to distinguish). Reinjection (or super infection) is infection with a second strain at a measurable time after the first. Both forms of dual infection have been reported for HIV in both acute and chronic infection around the world.

1.6 Structure of HIV

HIV-1 and HIV-2 are RNA viruses and belong to the family of retroviruses, *Retroviridae* (*retro*, backwards). The genome of retroviruses consists of duplicate copies of positive single-stranded RNA. Once a cell has become infected with a retrovirus the viral genetic information will be transformed from RNA to DNA, catalyzed by viral enzyme reverse transcriptase. The name retrovirus is derived from this unique event, which is completely opposite to the normal process where RNA is transcribed from DNA. Retroviruses are divided into seven genera, where the genus *Lentivirus* (*lenti*, slow), is characterized by the slow development of disease after infection. HIV is a typical lentivirus, since it usually has a disease latency of several years. A schematic drawing of the mature HIV virion is shown in Figure 1.6.1. The virion is almost spherical and is about one ten-thousandth of a millimeter across (ca. 100 nm). The virus is enveloped by a lipid bilayer that is derived from the infected host cell. The outer surface is studded with surface glycoproteins (gp120) that are anchored

to the virus via interactions with the transmembrane protein (gp41). These surface proteins play a crucial role when HIV binds to and enters the host cells. A shell of the matrix protein (p17) lines the inner surface of the viral membrane, and a conical capsid core particle constructed out of the capsid protein (p24) is located in the center of the virus. The capsid particle encapsulates two copies of the viral genome, stabilized by the nucleocapsid protein (p7), and also contains three essential virally encoded enzymes: protease (PR), reverse transcriptase (RT), and integrase (IN).

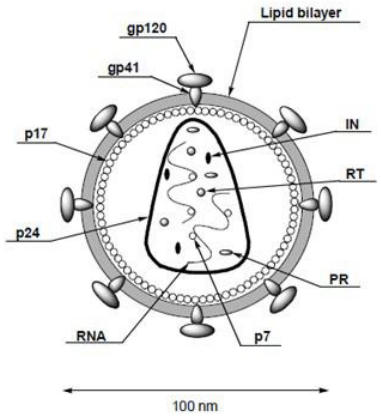


Figure 1.6.1

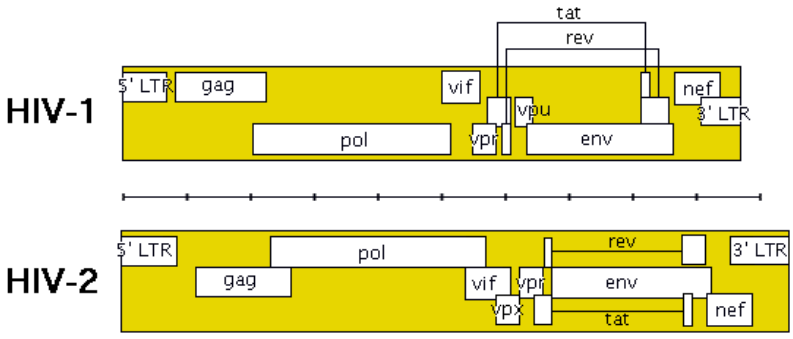
1.7 HIV genome

The integrated form of HIV-1, also known as the provirus, is approximately 9.8 kilo bases in length. Both ends of the provirus are flanked by a repeated sequence known as the long terminal repeats (LTRs). The genes of HIV are located in the central region of the proviral DNA and encode at least nine proteins. These proteins are divided into three classes:

1. The major structural proteins, Gag, Pol, and Env
2. The regulatory proteins , Tat and Rev
3. The accessory proteins, Vpu, Vpr, Vif, and Nef

The full HIV genome is encoded on one long strand of **RNA**. (In a free virus particle, there are actually two separate strands of **RNA**, but they're exactly the same!) This is the form it has when it is a free virus particle. When the virus is integrated into the host's **DNA** genome (as a **provirus**) then

its information too is encoded in **DNA**. The following image shows roughly how the genes are laid out in HIV (remember that **HIV-1 and HIV-2** are quite different).



Picture 1.7.1

The genes in HIV's genome are as follows:

- **gag** (coding for the viral capsid proteins)
- **pol** (notably, coding for **reverse transcriptase**)(NB. **gag** and **pol** together can be
- expressed in one long strand called "**gag-pol**")
- **env** (coding for HIV's envelope-associated proteins) And the regulatory genes.
- **tat** (regulatory gene which accelerates the production of more HIV virus)
- **rev** (It stimulates the **production of HIV proteins**, but suppresses the expression of HIV's regulatory genes)
- **nef** (encodes a protein which hangs around in the cytoplasm of the cell, and retards HIV replication)
- **vif** (codes for "virion infectivity factor", a protein that increases the infectivity of the HIV particle.)
- **vpr** ("Viral protein R" accelerates the production of HIV proteins)
- **vpu** (helps with the assembly of new virus particles, and helps them to bud from the host cell, not PRESENT IN hiv-1)
- **vpx** (Its role in the life of HIV is not entirely clear and not present in HIV-1)

1.8 Life Cycle of HIV

The attachment of the viral surface protein (gp 120) to the CD4-receptor, located on various cells within the immune system, initiates the lifecycle of HIV (Figure 1.8.1). Attached virions utilize several additional cell-surface proteins to promote the fusion of the viral and host cellular membranes. Membrane fusion is followed by a poorly understood uncoating event of the capsid that allows the release of the viral content into the host-cell cytosol .The single-stranded viral RNA complexes with reverse transcriptase, which catalyses the reverse transcription to yield a double stranded DNA molecule. The double-stranded viral DNA is then transported into the cell nucleus and is permanently integrated into the host genome by the catalytic activity of the viral integrase. The integrated viral DNA is designated *provirus*.

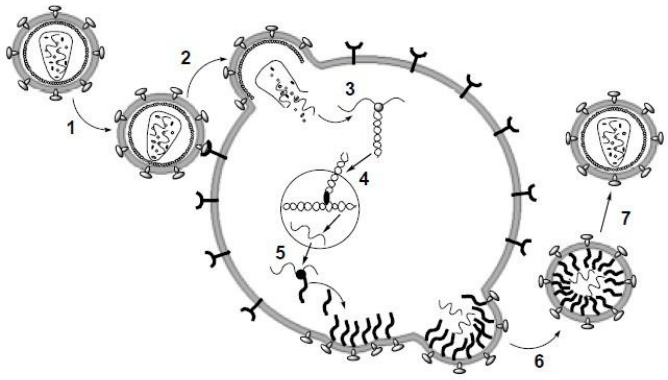


Figure1.8.1 Schematic drawing of the replicative cycle: 1. Attachment to the host cell CD4-receptor, 2.Viral fusion and uncoating, 3. Reverse transcription, 4. Integration of viral DNA to the host genome, 5.Translation, 6. Viral budding, 7. Maturation via protease activities.

By an unknown activation process the cell initiates the transcription of the proviral DNA by the host cellular RNA polymerase II. Initially, short spliced RNA species that encode the regulatory proteins Tat, Rev, and Nef are synthesized. Tat acts as a stimulator of the transcription of the proviral DNA to enhance the production of viral RNA. Full-length and singly spliced RNA is needed in the cytoplasm for the synthesis of Gag and Gag-Pol polyproteins, and for packing into new virions. Rev binds to the full length and singly spliced RNA in the nucleus and protects it from further splicing and actively transports it to the cytosol. In this manner, Rev acts as a switch between the

early synthesis of highly spliced RNAs and the later synthesis of unspliced and singly spliced RNAs. Nef acts as a down-regulator of the number of CD4 receptors on the surface of the infected cell. Translations of the unspliced RNA by the ribosomes produce the polyproteins Gag and Gag-Pol. These polyproteins are transported to the plasma membrane with two molecules of viral RNA. They assemble together with the envelope protein to form an immature virus particle that is released from the cell by budding from the cell surface. To become infectious, the virion has to pass through a maturation process where the enzyme HIV protease cleaves the polyproteins into functional enzymes and structural proteins. The mature HIV virion is now ready to infect a new cell and start a new life cycle.

2. DIAGNOSIS AND TREATMENT

2.1 Diagnosis

Screening for HIV infection is most commonly done by testing blood for HIV antibodies. A newer test, the Orasure test, involves collecting secretions between the cheek and gum and evaluating them for HIV antibodies. Orasure is essentially as accurate as a blood test because it doesn't involve a needle stick, it is favored by many individuals. Orasure is available through physicians' offices and many public health clinics .Finally, a new urine test available for screening, although if the test is positive, blood tests need to be performed for confirmation of the presence of HIV. In 1996, a home HIV blood test (called Home Access) became available to the public. These home kits are available in pharmacies and by mail. The kit contains a few sharp tools called lancets, a piece of blotting paper marked with a unique identification number and a prepaid return envelope with a protective pouch. After pricking the finger with the lancet, a few drops of blood are blotted onto the paper, sealed into the envelope and sent to the address on the envelope. In about a week, the person calls a toll-free number to get the results of the test.

2.2 Treatment of AIDS and HIV Infection

Anti-HIV (also called antiretroviral) medications are used to control the reproduction of the virus and to slow or halt the progression of HIV-related disease. When used in combinations, these medications are termed Highly Active Antiretroviral Therapy (HAART). HAART combines three or more anti-HIV medications in a daily regimen, sometimes referred to as a "cocktail". Anti-HIV medications do not cure HIV infection and individuals taking these medications can still transmit HIV to others. Anti- HIV medications approved by the U.S. Food and Drug Administration (FDA) fall into four classes:

1. Nonnucleoside Reverse Transcriptase Inhibitors (NNRTIs), such as nevirappine (Viramune) and efavirenz (Sustiva), bind to and block the action of reverse transcriptase, a protein that HIV needs to reproduce.

2. Nucleoside Reverse Transcriptase Inhibitors (NRTIs), such as zidovudine (Retrovir), tenofovir DF (Viread), and stavudine (Zerit), are faulty versions of building blocks that HIV needs to make more copies of itself. When HIV uses an NRTI instead of a
normal building block, reproduction of the virus is stalled.

3. Protease Inhibitors (PIs), such as lopinavir/ritonavir (Kaletra), disable protease, a protein that HIV needs to reproduce itself.

4. Fusion Inhibitors, such as enfuvirtide (Fuzeon), are newer treatments that work by blocking HIV entry into cells.

How many pills you will need to take and how often you will take them depends on what medications you and your doctor choose. The treatment of HIV infection and AIDS is in a highly dynamic state. Individuals with this condition are advised to seek out experts in their local community who are current with the latest modes of therapy and ongoing clinical trials for evaluating newer therapies.

2.3 HIV Therapy

There is no cure for AIDS. Although antiretroviral treatment can suppress HIV – the virus that causes AIDS – and can delay illness for many years, it cannot clear the virus completely. There is no confirmed case of a person getting rid of HIV infection. Sadly, this doesn't stop countless quacks and con artists touting unproven, often dangerous "AIDS cures" to desperate people. It is easy to see why an HIV positive person might want to believe in an AIDS cure. Access to antiretroviral treatment is scarce in much of the world. When someone has a life-threatening illness they may clutch at anything to stay alive. And even when antiretroviral treatment is available, it is far from an easy solution. Drugs must be taken every day for the rest of a person's life, often causing unpleasant side effects. A one-off cure to eradicate the virus once and for all is much more appealing. Distrust of Western medicine is not uncommon, especially in
developing countries. The Internet abounds with rumors of the pharmaceutical industry or the U.S. government suppressing AIDS cures to protect the market for patented drugs. Many people would prefer a remedy that is "natural" or "traditional".

2.3.1 Why is it so difficult to cure AIDS?

Curing AIDS is generally taken to mean clearing the body of HIV, the virus that causes AIDS. The virus replicates (makes new copies of itself) by inserting its genetic code into human cells, particularly a type known as CD4 cells. Usually the infected cells produce numerous HIV particles and die soon afterwards. Antiretroviral drugs interfere with this replication process, which is why the drugs are so effective at reducing the amount of HIV in a person's body to extremely low levels. During treatment, the concentration of HIV in the blood often falls so low that it cannot be detected by the standard test, known as a viral load test. Unfortunately, not all infected cells behave the
same way. Probably the most important problem is posed by "resting" CD4 cells. Once infected with HIV, these cells, instead of producing new copies of the virus, lie dormant for many years or even decades. Current therapies cannot remove HIV's genetic material from these cells. Even if someone

takes antiretroviral drugs for many years they will still have some HIV hiding in various parts of their body. Studies have found that if treatment is removed then HIV can re-establish itself by leaking out of these “viral reservoirs”. A cure for AIDS must somehow remove every single one of the infected cells.

2.3.2 Reputable research on curing AIDS

2.3.2.a Activating resting immune cells

Many researchers believe the best hope for eradicating HIV infection lies in combining antiretroviral treatment with drugs that flush HIV from its hiding places. The idea is to force resting CD4 cells to become active, whereupon they will start producing new HIV particles. The activated cells should soon die or be destroyed by the immune system, and the antiretroviral medication should mop up the released HIV. Early attempts to employ this technique used interleukin-2 (also known as IL-2 or by the brand name Proleukin). This chemical messenger tells the body to create more CD4 cells

and to activate resting cells. Researchers who gave interleukin-2 together with antiretroviral treatment discovered they could no longer find any infected resting CD4 cells. But interleukin-2 failed to clear all of the HIV; as soon as the patients stopped taking antiretroviral drugs the virus came back again. There is a problem with creating a massive number of active CD4 cells: despite the antiretroviral drugs, HIV may manage to infect a few of these cells and replicate, thus keeping the infection alive. Scientists are now investigating chemicals that don’t activate all resting CD4 cells, but only the tiny minority that are infected with HIV. One such chemical is valproic acid, a drug already used to treat epilepsy and other conditions. In 2005 a group of researchers led by David

Margolis caused a sensation when they reported that valproic acid, combined with antiretroviral treatment, had greatly reduced the number of HIV-infected resting CD4 cells in three of four

patients. They concluded that: “This finding, though not definitive, suggests that new approaches will allow the cure of HIV in the future.”

Sadly, it seems such optimism was premature; more recent studies suggest that valproic acid has no long term benefits. In fact it’s quite possible that all related approaches are flawed because the virus has other hiding places besides resting CD4cells. There is a lot about HIV that remains unknown.

2.3.2.b Bone marrow transplants and gene therapy

In November 2008, pair German doctors made headlines by announcing they had cureda man of HIV infection by giving him a bone marrow transplant. The transplant - given as a treatment for leukemia - used cells from a donor with a rare genetic mutation known as Delta 32 that confers resistance to HIV infection. Twenty months after the procedure researchers reported they could find no trace of HIV in the recipient's bone marrow, blood and other organ tissues. Other experts have said that more tests are required to verify this cure claim, though it is not the first of its kind. Of more than thirty HIV positive patients given bone marrow transplantation prior to 1996, two appeared to have been cured of their infection based on molecular testing and post-mortem biopsy samples. Even assuming it can be effective, bone marrow transplantation is too dangerous and costly for widespread use as a cure. Many patients die as a result of chemotherapy or reactions to the transplant, which is usually a last resort in treating life-threatening diseases. As Anthony Fauci, director of the National Institute of Allergy and Infectious Diseases put it: “It’s very nice, and it’s not even surprising. But it’s just off the table of practicality” Nevertheless the German transplant does raise hope for related approaches. If scientists can find another way - such as gene therapy - to confer the same sort of protection against HIV as Delta 32 provides, then they may be able to stop the virus replicating. Research in this area is in its very early stages; it may be many years before a useful treatment is found, if at all.

3. WHY CCR5 PROTEIN AS TARGET?

CCR5, short for chemokine (C-C motif) receptor 5 is a protein which in humans is encoded by the CCR5 gene which is located on chromosome 3 on the short (p) arm at position 21. CCR5 has also recently been designated CD195 (cluster of differentiation 195).The CCR5 protein functions as a chemokine receptor in the CC chemokine group. The natural chemokine ligands that bind to this receptor are RANTES, MIP-1α and MIP- 1β. CCR5 is predominantly expressed on T cells, macrophages, dendritic cells and microglia. It is likely that CCR5 plays a role in inflammatory responses to infection, though its exact role in normal immune function is unclear.

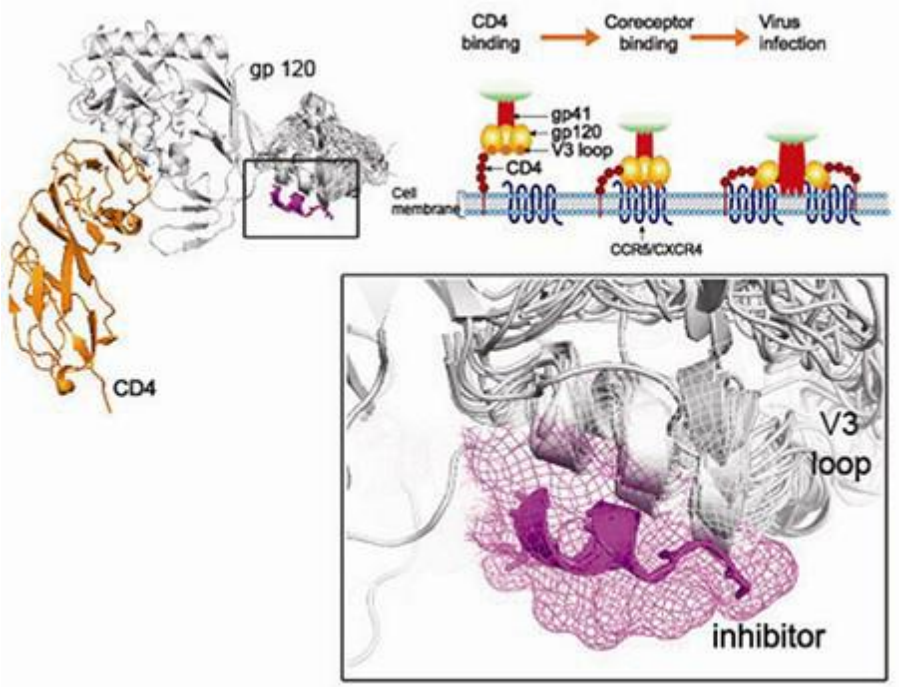


Figure 3. Mode of attachment of HIV virus to Human T cells by CCR5 co receptor.

METHODOLOGY

4. QUANTITATIVE STRUCTURAL-ACTIVITY RELATIONSHIP OF 5-OXOPYROLIDINE-3-CARBOXAMIDE DERIVATIVES

Quantitative structure-activity relationship (QSAR) (sometimes **QSPR**: quantitative structure-property relationship) is the process by which chemical structure is quantitatively correlated with a well defined process, such as biological activity or chemical reactivity. For example, biological activity can be expressed quantitatively as in the concentration of a substance required to give a certain biological response. Additionally, when physicochemical properties or structures are expressed by numbers, one can form a mathematical relationship, or quantitative structure-activity relationship, between the two. The mathematical expression can then be used to predict the biological response of other chemical structures. QSAR's most general mathematical form is:

Activity = f(physiochemical properties and/or structural properties)

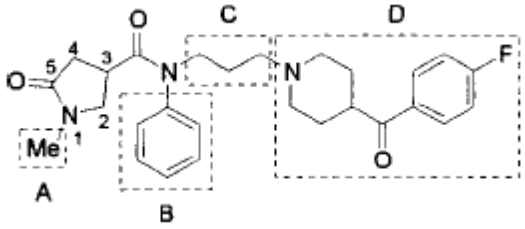


Figure 4.1 LEAD COMPOUND

QSAR analyses of CCR5 targeting HIV-1 inhibitors to explore the correlation between physicochemical and biological activity of 5-oxopyrrolidine-3- carboxamide derivatives using modeling software MOLECULAR MODELLIND PRO PLUS DEMO Version 6.2.5 (molecular modeling software, Published by ChemSw.Inc.) and statistical software SPSS win version 10.0 In the present work we have taken 49 oxopyrrolidine-3- arbaxamide compounds and their HIV-1 inhibitory activity from the reported work . Many of these compounds inhibited wild type HIV-1 with IC_{50} values between 0.033 µM and 5.2 µM. There is high structural diversity and a sufficient range of the

biological activity in the selected series of oxopyrrolidine-3-carbaxamide derivatives. It insists us to select these series of compounds for our QSAR studies. All the HIV-1 inhibitory activities used in the present study were expressed as $pIC_{50} = -\log_{10} IC_{50}$. Where IC_{50} is the micro molar concentration of the compounds producing 50% reduction in the HIV-1 activity is stated as the means of at least two experiments. The compounds which did not show confirmed HIV-1 inbitory activity in the above cited literature have not been taken for our study. We carried out QSAR analysis and established a QSAR model to guide further structural optimization and predict the potency and physiochemical properties of clinical drug candidates.

4.1 Structures of derivatives of 5-oxopyrolidine derivatives and their antagonistic activity

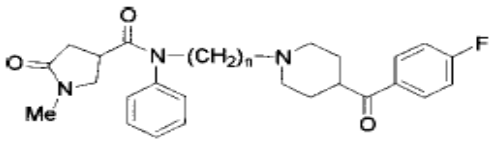


Table 4.1:

Compound	*N*	CCR5[a] IC50 (μM)
1	3	1.9
5a	4	3.2
5b	5	5.0

a)Inhibition of [^{125}I]RANTES binding to CCR5-expressing CH) cells. b) Prevent inhibition at 10 μM

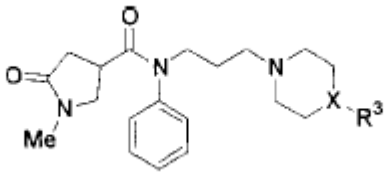
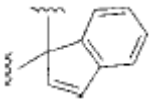
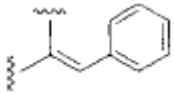


TABLE 4.2:

Compound	X	R^3	CCR5 IC_{50} (μM)
1	CH	CO(4-F-Ph)	1.9
5e	CH	H	3.6
5f	COH	4-Cl-Ph	5.2
5g			2.1
5h	CH	CH_2Ph	0.48
5i	N	CH_2Ph	1.9
7d	CH	OPh	1.6
7a*			1.2
7c*	CH	$(CH_2)_2Ph$	1.1
7e*	CH	OCH_2Ph	1.5
11a*	CH	CH_2(4-F-Ph)	0.31

a)Inhibition of [125I]RANTES binding to CCR5-expressing CH) cells. b) Prevent inhibition at 10 μM. * indicates test set.

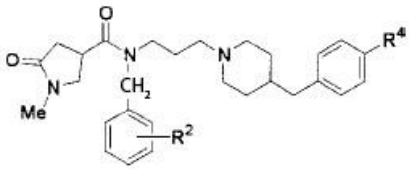
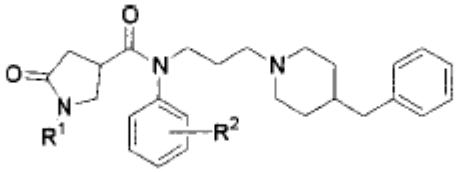


TABLE 4.3:s

Compound	R^2	R^4	CCR5 IC_{50} (μM)
10b	3-Me	H	0.16
10d	4-t-Bu	H	0.16
10f	4-MeO	H	0.12
10h	4-Cl	H	1.2
11b	3,4-diCl	F	0.050
10g*	3-Cl	H	0.12
10i*	3,4-diCl	H	0.057

)Inhibition of [125I]RANTES binding to CCR5-expressing CH) cells. b) Prevent inhibition at 10 μM. * indicates test set.

TABLE 4.4:

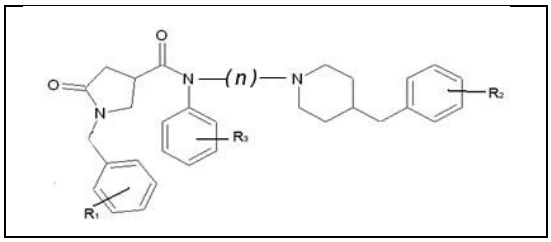


Compound	R^1	R^2	CCR5 IC_{50} (μM)
5h	Me	H	0.48
12a	n-Bu	H	0.13
12b	c-Hex	H	0.11
12c	c-HexCH_2	H	0.23
12d	Ph	H	0.23
12e	PhCH_2	H	0.038
12f	2-Cl-PhCH_2	H	0.033
12g	3-Cl-PhCH_2	H	0.0866
12h	4-Cl-PhCH_2	H	0.22
12i	4-Me-PhCH_2	H	0.33
13	H	H	0.57
14a	CF_3CH_2	H	0.075
14b	2-Me-PhCH_2	H	0.034
15a	PhCH_2	3-Cl	0.044
15b	PhCH_2	3,4-diCl	0.043
12l*	(pyridine-4-yl)CH_2	H	0.24

a)Inhibition of [125I]RANTES binding to CCR5-expressing CH) cells. b) Prevent inhibition at 10 μM. * indicates test set.

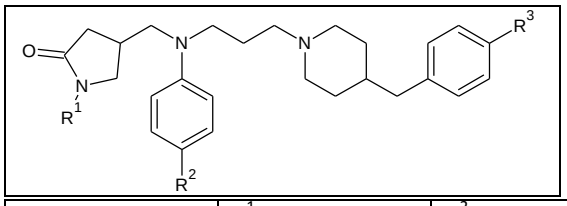
Constructed MOLECULES

TABLE 4.5

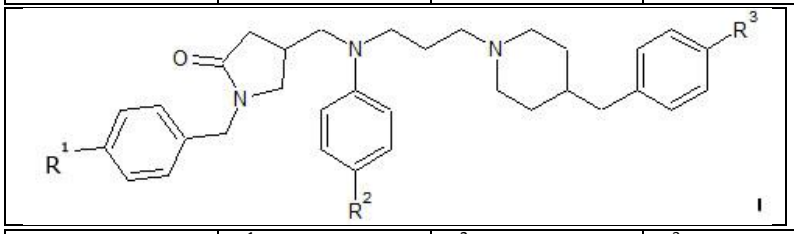


R_1 R_2 R_3- substitutions on different Benzene rings
n –variable chain lengths

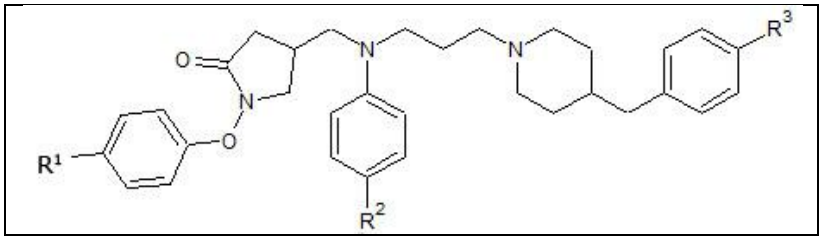
Comp	R^1	R^2	R^3	*n*
1	2-F,6-Cl	-	-	3
2	2,6-di F	-	-	3
3	2,6-di Cl,4-NH_2	-	-	3
4	6-I	-	-	3
5	-	3-Cl	-	3
6	-	3-F	-	3
7	-	2,6-diCl	-	3
8	-	2,4,6-tri Cl	-	3
9	2,4,6- tri Cl	2,6-di Cl	-	3
10	2-OH,5-Cl	-	-	3
11	2,6 di I	4-Cl	-	3
12	2,4,6-tri I	4-I	-	3
13	2,4,6-tri I	4-OCH_3	-	3
14	2,4,6-tri I	4-OCl	-	3
15	2,4,6-tri I	4-OI	-	3
16	2,4,6-tri I	4-I	4-I	3
17	2,4,6-tri I	4-I	4-Cl	3
18	2,4,6-tri I	4-I	4-NH_2	3
19	2,4,6-tri I	4-I	4-CN	3
20	2,4,6-tri I	4-CN	4-CN	3
21	2,4,6-tri I	2,4,6-tri I	4-I	3
22	2,3,4,5,6 -pent I	4-I	4-CF_3	3
23	2,4,6-tri I	2,4,6-tri CF_3	4-I	3
24	2,4,6-tri OMe	2,4,6-tri CF_3	4-CF_3	3
25	Ph	2,4,6-tri CF_3	4-CF_3	3
26	2,4,6-tri CCl_3	2,4,6-tri CCl_3	4- CCl_3	3
27	2,4,6-tri CF_3	2,4,6-tri CF_3	2,4,6-tri CF_3	3
28	2-OH	-	-	3
29	4-OH	-	-	3
30	2,4,6-tri CH_3	2,4,6-tri CH_3	4- CH_3	1
31	4-(2-ClPh) 2,6-di CF_3	2,4,6-tri CF_3	4-CF_3	1
32	4-CH_2Ph	2,4,6-tri CF_3	4-CF_3	1
33	4-sulphonyl	-	-	3
34	4-sulphonyl	4-sulphonyl	4-sulphonyl	3
35	2,3,4,5,6-Cl	-	-	3



Compound	R^1	R^2	R^3
36	CF_3	$(CH_2)_2CH_3$	NH_2
37	OH	COOH	OH
38	CN	$(CH_2)_2COOH$	OH
39	OH	$(CH_2)_2COOH$	COOH
40	CH_2OH	ONH_2	COOH
41	CH_3	OCH_3	OCH_3



Compound	R^1	R^2	R^3
42	F	COOH	OH
43	COOH	NH_2	OH
44	COOH	OH	$COONH_2$
45	OCH_3	NH_2	COOH
46	NO_2	NO_2	CH_2COOH
47	OH	OH	$(CH_2)_2COOH$



Compound	R^1	R^2	R^3
48	COOH	CH_2OH	ONH_2
49	SO_2NH_2	CH_2OH	$(CH_2COOH$
50	$(CH_2)_2COOH$	CH_2NO_2	$(CH_2)_2COOH$
51	OCH_3	$OC(CF_3)_3$	$(CH_2)_2COOH$
52	SO_2NH_2	CH_2NH_2	$(CH_2)_2COOH$

4.2 Modeling of Lignads :

All 37 oxo pyrrolidine derivatives structures were built on workspace of molecular modeling plus pro demo version 6.25 and energy minimization of the molecules was performed until the root mean square gradient value becomes smaller than 0.001 kcal/mol A. Low energy conformation for each compound was generated and used for calculating various physico-chemical descriptors like

density, Hansen polarity, percent hydrophilic surface, mr, LogP-Crippen, Moriguchi Log P of descriptors. All the calculated descriptors were considered as independent variable and biological activity as dependent variable. SPSS 10.0 software was used to generate QSAR models by applying stepwise multiple linear regression analysis. Statistical measures used were n- Number of compounds in regression, r-correlation coefficient, r^2-squared correlation coefficient, F-test (Fischer's value) for statistical significance, SEE-standard error of estimation, q^2-cross validated Correlation coefficient and correlation matrix to show correlation among the parameters.

5. HOMOLOGY MODELING OF CCR5

The amino acid sequence (FASTA sequence) Of the Human CCR5 protein was taken from the Swiss-Prot database, accession no. P51681. Sequence homology searches were carried out using BLAST (Basic Local Alignment Search Tool) algorithm against protein data bank(pdb). From the given sequence homology models 1f88 was selected as the template which has the sequence identity of 23% . The sequence alignment was done using Clustal X 2.0.11. And further modeling was carried out using MODELLER 8v2 taking Bovine Rhodopsin as target template.

Modeller is a computer program that models the 3-D structures of proteins and their assemblies by satisfying of special restrains. "Modeller " is most frequently used for homology or comparative protein structure modelling. The alignment of the sequence obtained from the clustalx2.0.11 was taken as the input for the "Modeller " for modelling the structures with Bovine Rhodopsin(1F88.pdb). The "Modeller" automatically calculates and binds the model with all non-hydrogen atoms. More generally , the inputs to the program were restrained on the spatial structure of amino acid sequence(s) to be modelled. The output of "Modeller" was the 3-D structure of the protein that satisfies these restraints in the best possible manner. "Modeller" automatically derives the restraints from the known related structures and their alignment with the target sequence .A 3-D model was obtained by optimization of a molecular probability density function

(pdf). The molecular pdf for comparative modelling was optimized with the variable target function procedure in Cartesian space that employs methods to conjugate gradients and molecular dynamics with simulated annealing . The models generated by MODELLER were evaluated with PROCHECK. The active sites residues of Human CCR5 reported by Yan Zhang *et al* were used for docking analysis.

6. DOCKING STUDIES ON CCR5.

Molecular docking can fit molecules together with favourable configuration to form a complex system.The 3D structures of the new Human CCR5 structures were modeled with MODELLER modeling for tracking the interactive mode of Human CCR5 and molecules, the advanced Docking program(GLIDE 4.0) was used .

This Docking program leads to the designing of new molecules using Ligand-based drug design. Ligand-based drug design (or indirect drug design) relies on knowledge of other molecules that bind to the biological target of interest. These molecules which bind to the target may be used to derive a pharmacophore which defines the minimum necessary structural characteristics a molecule should possess. In other words, the modelled biological target may be used based on the knowledge of the molecules which binds to it and this in turn may be used to design new molecular entities that interact with the target effectively.

6.1 Ligand preparation:

The ligands were built using maestro build panel and prepared by Ligprep2.0 application. The structures of the ligands are given in table- 4.1 ,4.2 ,4.3 ,4.4, 4.5 . The main objective of Lig prep is to take 2D or 3D structures to use in further computational studies. The Low energy conformers of the ligands were obtained using the OPLS_2005 force field. The low energy conformers of the molecules with drug like property were used for the docking studies on Human CCR5 protein.

6.2 Protein Preparation:

The Homology Model of Human CCR5 with Swiss-Prot ID: P51681 was generated using MODELLER with ClustalX alignment of Bovine Rhodopsin sequence [accession number P02699] and its crystal structure [PDB: 1F88] as Template chosen from NCBI BLAST hit. The obtained Model was validated using PROCHECK and final Energy minimized using GROMOS96 to obtain stable structure for further studies.

6.3 Grid Generation:

The active site of the protein was defined and a grid was prepared for the ligands to be dock in, this was done by the receptor grid generation facility of the glide. The receptors Vander waals scaling of 0.9(which makes the protein site "roomier "by moving back the surface of Non-polar regions of the protein and Ligand. This kind of adjustments emulate to some extent the effect of breathing motions to the protein site it is a kind of giving breathing to receptors). The receptors active site is defined either by selecting the co-crystallized Ligand, or by mentioning the active side residues, Asp2,pro8, Asn13, Tyr14, Tyr15, Thr16, Gln22 and Lys 26.

6.4 Ligand docking:

The low energy conformations of the ligands were selected and docked into the prepared grid. Standard precision mode was used as the docking mode. The results of the docking are given in Table- 7.3.1.

7. RESULTS AND DISCUSSION

7.1 Homology modeling:

7.1.1 Sequence alignment

Amino acid sequence of CCR5 was obtained from Swiss-Prot database , accession no.P51681. Sequence homology searches were carried out using BLAST(Basic Local Alignment Search Tool) algorithm against Protein Data Bank. Target –template alignment was created using Clustal X 2.0.11 win version. Bovine Rhodopsin (pdb code:1F88) amino acid sequence was used for the sequence alignment. It gave the alignment as given below.

```
CLUSTAL 2.0.11 multiple sequence alignment


sp|P51681|CCR5_HUMAN      MDYQVSSPIYDINYYTS-------EPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVIL
1F88_A                    MNGTEGPNFYVPFSNKTGVVRSPFEAPQYYLAEPWQFSMLAAYMFLLIMLGFPINFLTLY
                          *:   .. :*     .:       *. *   .:     :*..   *::::**  *:*.:

sp|P51681|CCR5_HUMAN      ILINCKRLKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQ--WDFGNTMCQLLTGLYFIG
1F88_A                    VTVQHKKLRTPLNYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGFFATLG
                          : :: *:*::  :  *****::***:::     : *:: :  : ** * *:*   :  :*

sp|P51681|CCR5_HUMAN      FFSGIFFIILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGII-FTRSQ
1F88_A                    GEIALWSLVVLAIERYVVVCKPMSNFRFG-ENHAIMGVAFTWVMALACAAPPLVGWSRYI
                             .:: :::*:*:**:.* :.:  ::    ...::  .:***:*: .: * :: ::*

sp|P51681|CCR5_HUMAN      KEGLHYTCSSHFPYSQYQFWKNFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRN----
1F88_A                    PEGMQCSCGIDYYTPHEETNNESFVIYMFVVHFIIPLIVIFFCYGQLVFTVKEAAASATT
                           **:: :*. .:  .: :  ::  .: :.:: :::**:*:.:**. :: *: ..

sp|P51681|CCR5_HUMAN      EKKRHRAVRLIFTIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLG
1F88_A                    QKAEKEVTRMVIIMVIAFLICWLPYAGVAFY--------IFTHQGSDFGPIFMTIPAFFA
                          :* .:...*::: ::*.::: * **  * :         : . ..*:     * :.  :.

sp|P51681|CCR5_HUMAN      MTHCCINPIIYAFVGEKFRNYLLVFFQKHIAKRFCKCCSIFQQEAPERASSVYTRSTGEQ
1F88_A                    KTSAVYNPVIYIMMNKQFRN-----------------CMVTTLCCGKNPSTTVSKTETSQ
                           * .  **:** ::.::***                 * :    . :..*:. :::  .*

sp|P51681|CCR5_HUMAN      EISVGL
1F88_A                    VAPA--
                            ..
```

Figure 7.1.1 Sequence alignment of Human CCR5 and Bovine Rhodopsin (1F88.pdb) done using Clustal X 2.0.11

7.1.2 Evaluation of the models by PROCHECK

The quality of the models obtained from MODELLER was evaluated by PROCHECK.PROCHECK indicates the percentage of residues located in the favored regions of the Ramachandran plot. This plot gives the main chain conformations as pairs of ψ and ϕ dihedral angles for each residue in the protein. The stereo-chemical parameters showed that more than 84.4% of the residues of the model has the (ψ, ϕ) dihedral angles in the most favored and the allowed regions of the Ramachandran plot as expected for a good model.

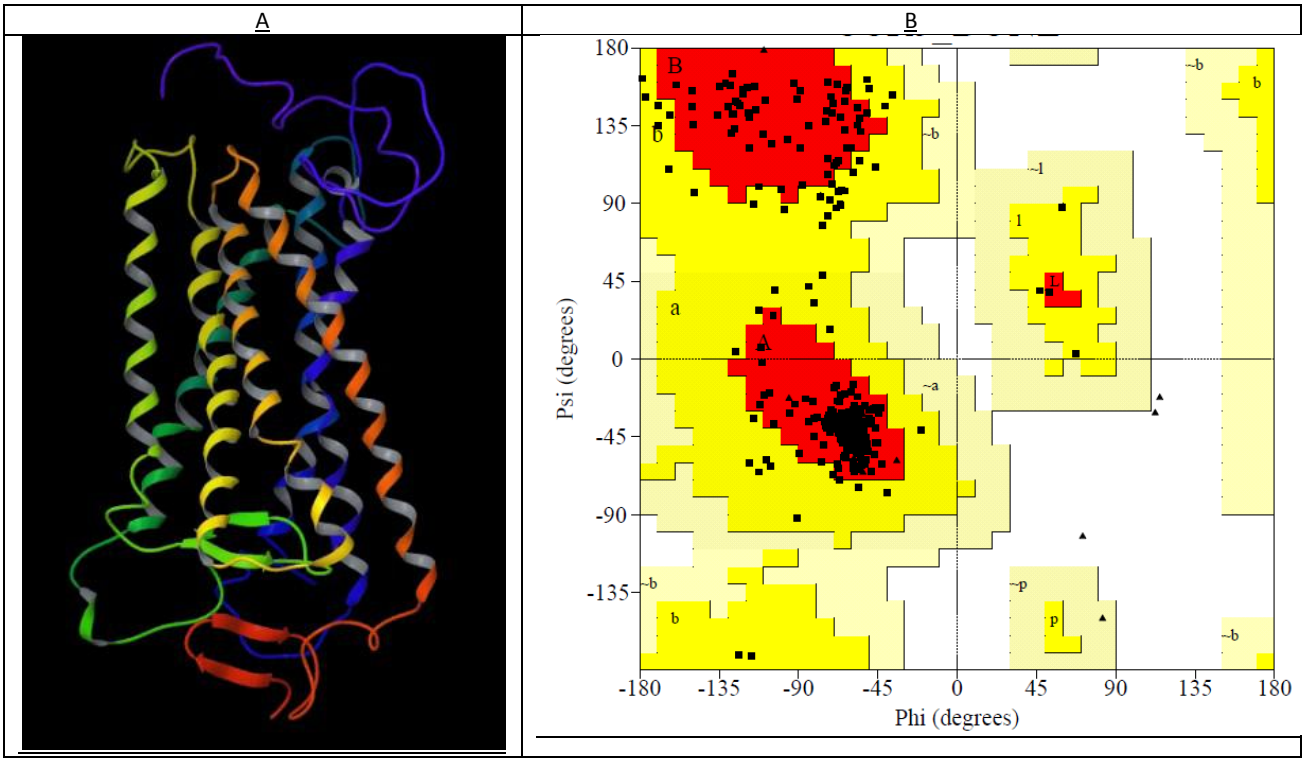


Figure 7.1.2: Picture "A" is the homology modeled protein based on the Bovine rhodopsin and "B" is the Ramachandran plot of the modeled protein.

7.2 2D-QSAR:

A data set of 37 compounds antagonistic to Human CCR5 as anti HIV agents's inhibitory activity was used for the present QSAR study. The QSAR studies of the oxo pyrollidine series resulted in several QSAR equations.The selected descriptors are density, Hansen polarity,percent hydrophilic surface, mr, LogP-Crippen, Moriguchi Log P.

The equation when we considered only one parameter is Eq.1.

Log (1/IC50) = (2.594±2.211)+ DENSITY(3.213±1.831)

n =29, r=0.325, r^2 =0.106, r^2 adj =0.071, SEE=0.704, F=3.078

Eq.2 is considered with two parameters.

Log (1/IC50) = (5.869±1.430)+DENSITY(5.344±1.158)- HANSEN POLARITY

(1.260±0.187)

n =29, r=0.826 , r^2=0.682 , r^2 adj j=0.657 ,SEE=0.4287 ,F=26.807

Eq.3 is considered with three parameters.

Log (1/IC50) = (5.854±1.462)+DENSITY(5.365±1.189)- HANSEN POLARITY(1.283±0.240) + PERCENT HYDROPHILLIC SURFACE(0.003622±0.024)

n =29, r=0.826, r^2=0.682, r^2 adj =0.643, SEE=0.437, F=17.181

Eq.4 is considered with four parameters.

Log (1/IC50) = (3,872±2.689)+ DENSITY(4.778±1.368) –HANSEN POLARITY(1.017±0.386) +PERCENT HYDROPHILLIC SURFACE(0.001544±0.024) +MR(0.0105±0.012)

n =29, r=0.832, r^2=0.693, r^2 adj =0.639, SEE=0.43942, F= 12.958

Eq.5 is considered with five parameters.

Log (1/IC50) = (6.367±2.121)+ DENSITY(2.409±1.177)- HANSEN POLARITY(1.135±0.294)+ PERCENT HYDROPHILLIC SURFACE(0.05913±0.013)- MR(0.0273±0.013) + MORIGUCHI LOG P(1.010±0.240)

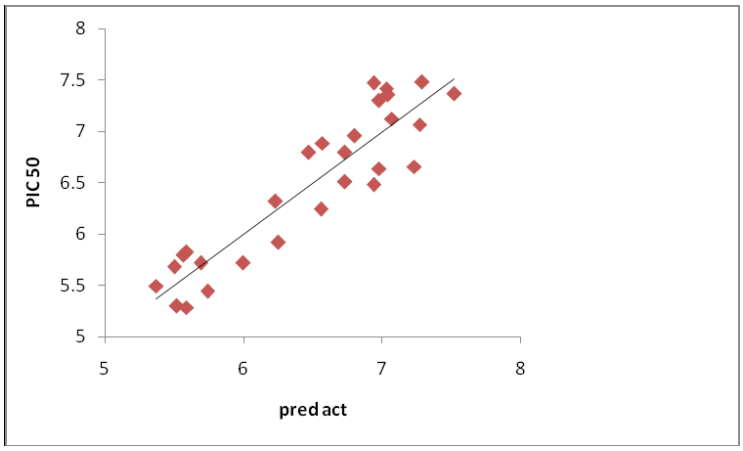
n =29, r=0.912, r^2 =0.831, r^2 adj =0.6793, SEE=0.333, F= 21.651

Eq.6 is considered with six parameters.

Log(1/IC50)=(6.881±2.217) +DENSITY(2.634±1.213) – HANSEN POLARITY(0.750±0.538) + PERCENT HYDROPHILLIC SURFACE(0.0465±0.027) – MR(0.0431±0.022)+ MORIGUCHI LOG P(0.358±0.851)+ LOG P CRIPPEN(0.598±0.698).

n =29, r=0.915, r^2 =0.837, r^2 adj=0.790, SEE=0.335, F= 19.946

The above five equations are not statistically significant one because r^2 value is low and SEE value is high. So we have considered the best equation containing six parameters i.e. Eq.6. Out of above six models, model 6 was selected as the best model on the basis of high r^2 value. Eq. 6 could explain 83.6% and predict 83.6% of the variance of the HIV-1 RT inhibitory activity data. The calculated CCR5 antagonistic HIV- 1 inhibitory activity values by Eq.6 are given in Tables 4.1, 4.2, 4.3, 4,4. This model showed good correlation coefficient (r) between descriptors [Density, Hansen Polarity, Percent hydrophilic surface , MR, Log P- Crippen ,Moriguchi Log P] and HIV-1 inhibitory activity.



Correlation between observed and Predicted P IC_{50}-.

Where Cal .Act - calculated activity, Pred. Act – predicted activity by leave one out method, Pred. Act - predicted activity of test set compounds by Eq.6. F value F=19.946. The predictive ability of the selected model was also confirmed by external r^2 cvext method. According to that, the proposed QSAR model is predictive as it satisfies with values q^2=0.54828, r^2_{pred} =0.836936.

SELECTED PHYSICOCHEMICAL PARAMETERS of OXO PYROLLIDINE DERIVATIVES TABLE 7.2.1

TRAINING SET

s.no	Name	Predact	Density	Hansen polarity	% hydrophilic Surface	MR	LogP- Crippen	Moriguch Log P
1	1	6	1.142644	4.836729	24.27593	130.5919	4.453201	4.377431
2	5a	5.37	1.208781	5.284368	28.46852	136.804	3.952802	4.122196
3	5b	5.52	1.192957	5.062997	26.67854	141.422	4.3491	4.314973
4	5e	5.74	1.124088	5.064099	25.79485	125.9739	4.056901	4.176203
5	5h	6.23	1.232806	4.836729	24.19967	130.5919	4.453202	4.377431
6	5i	5.69	1.198651	5.43031	44.84125	129.914	3.365202	3.404246
7	5f	5.59	1.223559	5.293845	28.87193	132.484	3.918402	4.141474

8	5g	5.5	1.119895	4.981136	24.9106	132.543	4.193101	4.038699
9	7d	5.57	1.260775	5.136687	29.45391	127.6169	3.400402	3.672246
10	10b	6.47	1.2278	4.651731	23.3352	135.21	4.920402	4.575699
11	10d	6.73	1.093409	4.120549	22.02595	149.0641	6.043503	5.154153
12	10f	5.59	1.158003	4.722558	24.30857	136.853	4.200502	4.071742
13	10h	6.25	1.19515	4.99797	23.67762	135.459	4.971201	4.843699
14	11b	6.98	1.379101	5.303861	22.71942	140.176	5.628701	5.767931
15	12a	6.57	1.086143	4.262579	23.38437	144.446	5.660603	4.963931
16	12b	6.8	1.206893	4.120742	21.07969	151.4821	6.038301	4.949936
17	12c	6.98	1.20559	3.968922	21.23779	156.1001	6.352003	5.135386
18	12d	6.73	1.106693	3.968922	21.50341	156.1001	6.352003	5.135386
19	12e	7.03	1.238124	4.211433	24.97491	154.6991	6.229803	5.092387
20	12f	7.29	1.293201	4.359344	24.36882	159.5661	6.747802	5.543602
21	12g	7.27	1.286411	4.359344	24.23878	159.5661	6.747802	5.543602
22	12h	7.23	1.275139	4.359344	23.99299	159.5661	6.747802	5.543602
23	12i	6.94	1.123137	4.070524	24.18621	159.3171	6.697003	5.275603
24	13	6.56	1.256461	4.601098	28.17552	125.9739	4.091801	4.176203
25	14a	7.07	1.267206	4.908071	27.18196	135.21	5.428201	5.154153
26	14b	6.94	1.122562	4.070524	24.20061	159.3171	6.697002	5.275603
27	15a	7.04	1.189928	4.359345	24.63785	159.5661	6.747802	5.543602
28	15b	7.52	1.331521	4.495332	23.5276	164.4332	7.265802	5.992675

TABLE 7.2.2 **TEST SET**

s.no	name	predact	density	Hansen polarity	%_hydrophilic surface	MR	LogP-Crippen	Moriguchi Log_P
1	7c	6.4	1.216051	4.628899	23.15271	135.21	4.849502	4.575699
2	7a	5.6	1.116082	4.936515	25.17915	130.125	4.015901	4.232431
3	7e	5.57	1.252415	4.908436	31.18979	132.235	3.396102	3.605474
4	10g	6.52	1.300444	4.99797	23.58885	135.459	4.971201	4.843699
5	10 i	6.78	1.353253	5.145194	22.81583	140.326	5.489201	5.307153
6	10j	6.7	1.327257	5.164576	23.10216	135.309	5.110701	5.307153
7	10o	5.52	1.258357	5.623432	31.67598	137.426	4.489101	3.658742
8	11a	6.4	1.261035	5.012321	23.66639	130.442	4.592702	4.843699
9	12l	6.01	1.166571	4.687358	32.77093	154.0211	5.598901	3.74298

7.3 Docking Studies :

7.3.1 DOCKING SCORES OF THE NEW MOLECULES

----------Designed Molecules---------- ----------Existing Molecules----------

TABLE 7.3.1

Comp	G score	No of H Bonds
1	-2.89	1
2	-2.73	1
3	-3.08	2
4	-2.94	1
5	-2.77	2
6	-2.79	2
7	-2.44	1
8	-2.58	1
9	-2.78	0
10	-3.28	1
11	-2.68	2
12	-2.93	1
13	-3.27	1
14	-2.95	1
15	-2.7	0
16	-2.7	1
17	-3.14	1
18	-2.96	1
19	-2.89	2
20	-3.77	2
21	-2.76	2
22	-2.69	2
23	-3.3	2
24	-4.16	3
25	-2.78	1
26	-3.36	1

Comp	G score	No of H Bonds
27	-2.7	1
28	-3.55	3
29	-3.08	2
30	-2.95	2
31	-2.91	1
32	-2.66	3
33	-4.31	2
34	-3.6	2
35	-4.64	4
36	-2.51	1
37	-5.3	3
38	-3.89	5
39	-5.24	4
40	-4.27	5
41	-2.8	1
42	-3.61	2
43	-4.46	1
44	-4.84	4
45	-2.7	2
46	-3.52	3
47	-4.67	5
48	-4.58	3
49	-4.32	2
50	-4.42	2
51	-4.17	2
52	-4.45	3

Comp	G score	No of H Bonds
1	-2.3	0
5a	-3.1	2
5b	-2.58	1
5e	-3.28	2
5h	-3.39	0
5i	-2.96	1
5f	-3.27	3
5g	-2.47	2
7d	-2.93	2
10b	-2.31	0
10d	-2.2	1
10f	-3.25	3
10h	-2.98	1
11b	-3.17	2
12a	-2.12	1
12b	-2.06	0
12c	-2.49	1
12d	-2.55	2
12e	-2.69	2
12f	-2.79	1
12g	-2.8	1
12h	-2.31	0
12i	-2.53	1
13	-2.41	2
14a	-3.6	2
14b	-3.11	0

Comp	G score	No of H Bonds
15a	-3.93	1
15b	-3.75	1
7c	-3.01	2
7a	-3.23	1
7e	-1.91	0
10g	-2.94	2
10 i	-2.9	2
10j	-2.85	3
10o	-2.49	3
11a	-2.13	1
12l	-3.35	2

7.3.2 DOCKING POSES OF DESIGNED MOLECULES

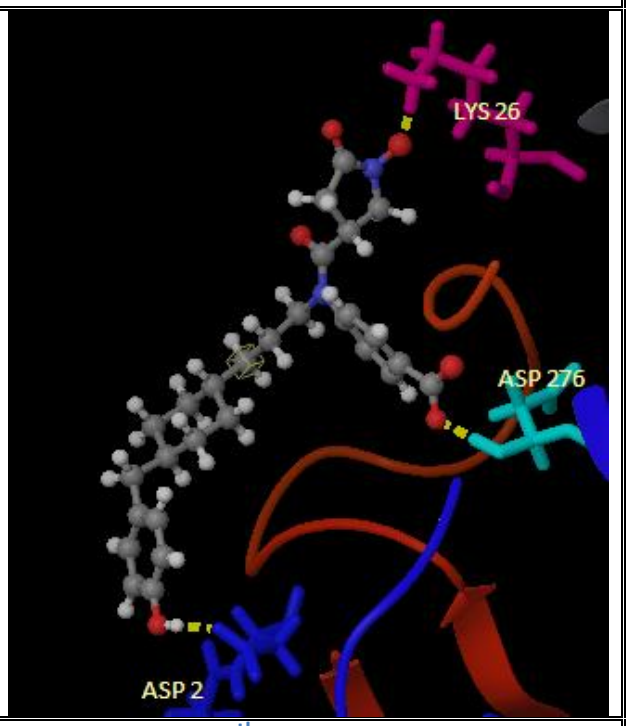
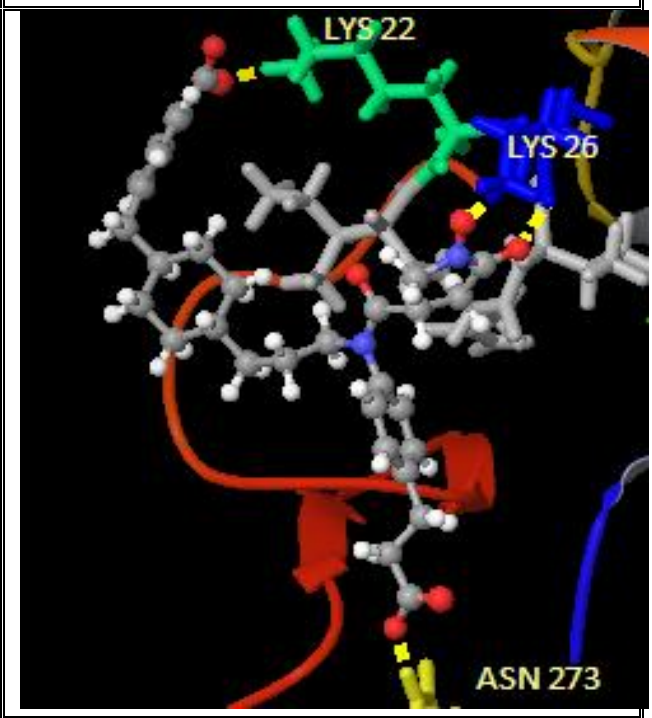
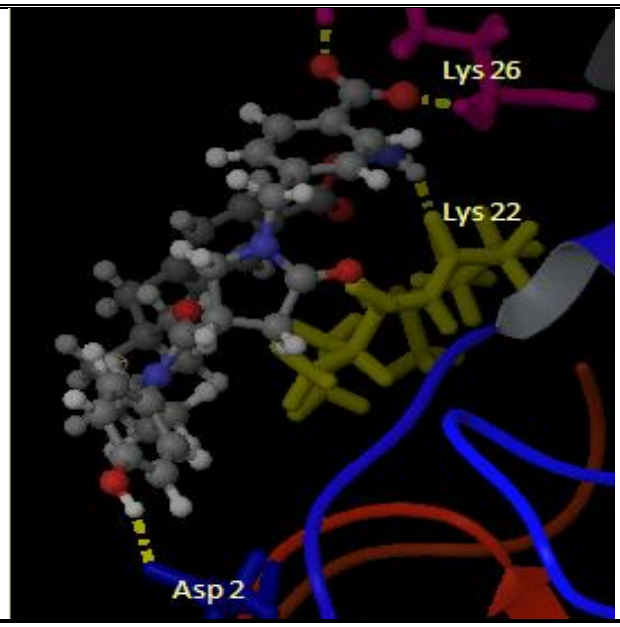
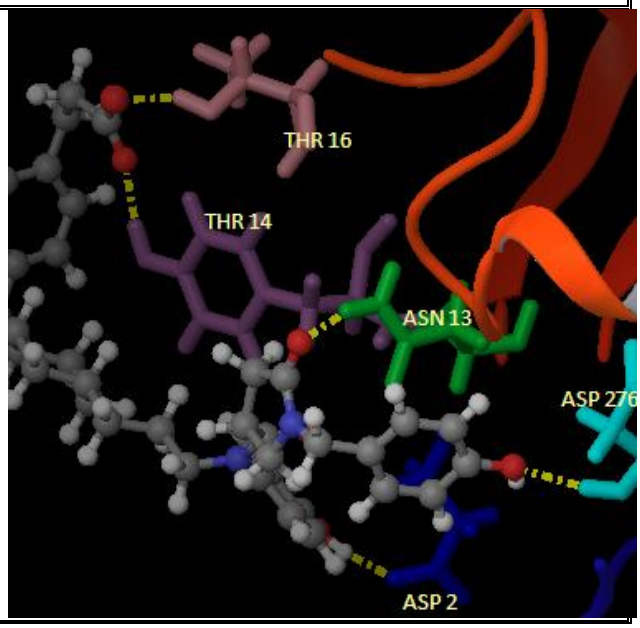
37th Molecule	39th Molecule
	
44th Molecule	**47th Molecule**
	

Figure 7.3.2: Protein Ligand interactions of the 4 best docked molecules with their dock poses, Ash colored (ball and stick)molecule represents the Ligand that was designed and the yellow colored dotted lines represents the hydrogen bonds formed between the Ligand and the active site amino acid residues(colored tubes).

To understand the interaction between CCR5 and the new designed molecules, the molecules from the 2D-QSAR were taken and docked at the active sites amino acid residues- Asp2, pro8,

Asn13, Tyr14, Tyr15, Thr16, Gln22 and Lys 26. The docked poses and the Hydrogen bonding interaction of the ligands with the receptor were given in the Figure 7.3.2. It is evident from the analysis of the Docked complex that these inhibitors are located in the center of the Active site, and is stabilized by hydrogen bonding interactions arresting the co-receptor activity of CCR5 for the entry of M tropic HIV-1. Through the interaction analysis, we concluded that Asp2, Lys 22 and Lys 26 are important anchoring residues for CCR5 and are the main contributors towards compound interaction. Though the interaction energy does not include the contribution from the water or the extended compound structure, this preliminary data along with the list of hydrogen bond interactions between the compounds and the Active site residues clearly supports that Asp2, Lys 22 and Lys 26 are more preferred residues in binding.

8. CONCLUSION:

The 2D QSAR studies performed on the existing CCR5 antagonistic HIV-1 inhibitors gave good statistical value. Docking studies were performed on the modeled structure of CCR5. Modelled structure was validated by Procheck and Verify 3D. Based on 2D QSAR and protein ligand interactions, enabled us to design new molecules which showed better activity than the existing molecules. These can be used for further studies.

9. REFERENCE:

1] Schrödinger, LLC, New York, NY, Glide, Version 4.0, 2005.

2] Richard A Friesner, Jay L Banks, Robert Bm Murphy, Thomas A Halgren, Jasna J Klicic, Daniel T Mainz, Matthew P Repasky, Eric H Knoll, Mee Shelley, Jason K. Perry, David E Shaw, Perry Francis, and Peter S Shenkin *J. Med. Chem***,** 2004, **47***,* 1739-174.

3] S.F.Altschul, T.L.Madden , A.A.Schaffer, J.Zhang , W.Miller, D.J.Lipman, Gapped Blast and PSI-BLAST: a new generation of protein data base search programs,nucleic acids Res.25(1997)3389-3402.

4] A.Sali ,T.L.Blundel, Comparative Protein modelling by satisfaction of spacial restraints, J.Mol.Biol.234(1993)779-815.

5] M.A.Marti-Renon, A.C.Stuart, A.Fiser , R.Sanchez, F.Melo , A.Sali, Comparative protein structure Modelling of genes and genomes, Annu.Rev.Biophys.Biomol.Struct.29(2000)291-325.

6] J.D.Thompson,D.G.Higgins, T.J.Gibbson, CLUSTAL X: improving the senstivity of progressive multiple sequence alignement through sequence weighting, position-specific gap penalities and weight matrix choice, Nucleic acids Res.22(1994)4673-4680.

7] R.A.Laskowski, M.W.Mac Arthur,D.S.Moss , J.M.Thornton, PROCHECK :a program to check the stereochemical quality of protein structures, J.Appl.Cryst.26(1993)283-291.

8] Guo Li, Kendra M. Haney, Glen E. Kellogg, and Yan Zhang*, Comparative Docking Study of Anibamine as the First Natural Product CCR5 Antagonist in CCR5 Homology Models *J. Chem. Inf. Model.* 2009, *49,* 120–132.

9] Shinichi IMAMURA,*, Yuji ISHIHARA, Taeko HATTORI, Osamu KURASAWA, Yoshihiro MATSUSHITA, Yoshihiro SUGIHARA, Naoyuki KANZAKI, Yuji IIZAWA, Masanori BABA, and Shohei HASHIGUCHI , CCR5 Antagonists as Anti-HIV-1 Agents. 1. Synthesis and Biological Evaluation of 5-Oxopyrrolidine-3-carboxamide Derivatives*Chem. Pharm. Bull.* **52**(1) 63—73 (2004) .