Chemical and Process Engineering Research
ISSN 2224-7467 (Paper) ISSN 2225-0913 (Online)
Vol.7, 2013



# HIV-I Protease Based Inhibitor Discovery

Dr. T ANTHONEY SWAMY, M.Sc., M.Phil., Ph.D., I.A.C.T
Department of Chemistry
University of Eastern Africa, Baraton, Post Box 2500-30100, Eldoret, Kenya
Email: drtanthony2011@yahoo.com

Dr. GUMMADI SAMUEL MOSES, M.Sc., Ph. D., PGDEEM., MISTE., FICC., FICS.,
Head, Senior Reader, P.G Dept. of Chemistry
Government College (NAAC 'A' Grade)
Rajahmundry, Andhra Pradesh
Email: drgsmoses@yahoo.com

**ABSTRACT**:

We can devise a drug (inhibitor) to restrain the activity of gene. As a gene engenders protein/enzyme, so to circumvent the development of any disease causing proteins, we have to stop the activity of that gene. With the aid of different bioinformatics tools and software's we can do this. A protease is an enzyme that smites proteins to their constituent peptides. The HIV-I Protease (PR) hydrolyses viral polyproteins into functional protein products that are vital for viral assembly and subsequent activity. HIV-I protease activity is decisive for the terminal maturation of infectious virions. Once HIV enters the cell, viral RNA experiences reverse transcription to generate double-stranded DNA (a step inhibited by nucleoside analogues such as zidovudine, didanosine, zalcitabine, stavudine, and lamivudine).

In the presence of HIV-I protease inhibitors, the virion is incapable to mature and is quickly cleared by inadequately comprehended mechanisms. Figure 1, left, is a photomicrograph of normal budding virions from an infected cell, while Figure 1, right, determines the effect of bathing these cells with the protease inhibitor, saquinavir. The consequent lack of a dense core for these "ghosted" particles is the feature of noninfectious HIV virions. By applying ncbi we can acquire the nucleotide and protein sequence of HIV-I Protease. By tool and softwares like pfam, clustalw, gold, blast, we designed the inhibitor "SKF 108737"for HIV-I protease.

**Keywords**: Inhibitor (Drug), HIV-I protease

## Introduction:

HIV-I protease (HIV PR) is an aspartic protease that is essential for the lifecycle of HIV, the retrovirus that causes AIDS. [1][2] HIV PR cleaves lately synthesized polyproteins at the apposite places to generate the mature protein components of an infectious HIV virion. Without effectual HIV PR, HIV virions remain uninfectious.[3][4] Thus, mutation of HIV PR's active site or inhibition of its activity disturbs HIV's capability to replicate and infect additional cells,[5] making HIV PR inhibition the subject of much pharmaceutical research. [6]

## STRUCTURAL FEATURES:

HIV-I Protease is a homodimer (chain A, chain B). Each monomer consists of 99 amino acids and is alike in conformation. The position of each monomer in the ardent protease builds an axis of symmetry. The secondary structure of each monomer constitutes, one alpha helix and two anti-parallel beta sheets. The two Asp artifacts (one from each chain) function as the catalytic residues. According to the procedure for HIV PR protein cleavage considered by Jaskolski et al., water acts as a nucleophile, which acts in simultaneous conjunction with a well-placed aspartic acid to hydrolyze the scissile peptide bond.[7] beside, HIV PR has the two molecular

“flaps” which move a distance of up to 7 Å when the enzyme becomes associated with a substrate.[8] Aliphatic residues constrain each monomer in a hydrophobic core. in addition, the dimmer is constrained by no covalent interactions, hydrophobic packing of side chains and interactions composing the catalytic residues. Each monomer comprises two cysteine residues, but these do not form disulfide bonds.

The active site forms at the dimmer interface .It is conceived in a cleft between the two domains as part of four stranded beta turn. The substitute view ascertains the position of the active site nestled in approximately in the center of the molecule.

An additional turn, a beta hairpin loop, of a beta sheet encases the active site. This flap persists to be adaptable and apportions for hinge like mobility. It enables substrate access to the active site by opening and folding the tips into hydrophobic pockets thus exercising a central role in protease activity.

The overall shape of protease is oblong and relatively flat. This surface contour evinces where potential binding or protein interaction might occur: several binding pockets exist inside the hollow cleft. The structure at left is a ligand.

**MECHANISM OF ACTION:**

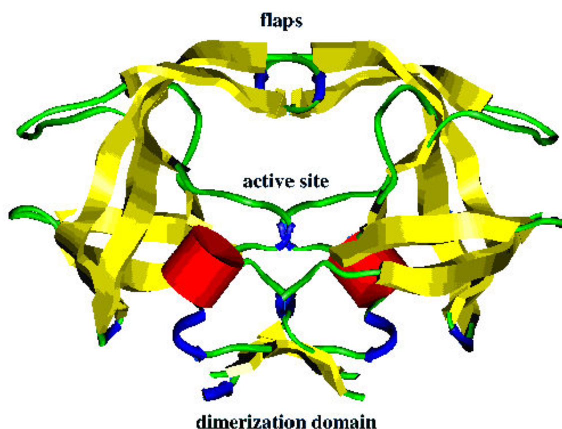


**Figure 1: HIV-I Protease**

HIV-I protease activity is captious for the terminal maturation of infectious virions. Once HIV enters the cell, viral RNA endures reverse transcription to generate double-stranded DNA (a step inhibited by nucleoside analogues such as zidovudine, didanosine, zalcitabine, stavudine, and lamivudine). This viral DNA is incorporated into the host genome and, circumstantially, transcribed and translated by cellular enzymes to create large, nonfunctional polypeptide chains, insinuated to as polyproteins. After these poly-proteins are comprised and packaged at the cell surface, immature virions are generated and released into the plasma. At this point, HIV-I protease, functioning as a "molecular scissors," cleaves the poly-proteins into smaller, functional proteins, thereby apportioning the virion to mature. In the existence of HIV-I protease inhibitors, the virion is unable to mature and is quickly cleared by insufficiently assented mechanisms.

Figure 1, left, is a photomicrograph of normal budding virions from an infected cell, while Figure 1, right, displays the effect of bathing these cells with the protease inhibitor, saquinavir. The consequent lack of a cohesive core for these "ghosted" particles is characteristic of noninfectious HIV virions.

## HIV-I PROTEASE ACTIVE BINDING:

HIV-I protease consists of two protein chains. The chains are indistinguishable to one another, and each contains amino acids. When the two chains amass, a long tunnel is created (as seen from the side view on above left). Protein "flaps" envelop the tunnel and open up to allow the enzyme to connect to a protein chain. After connection, the flaps then close around the protein chain, thus holding it in the tunnel and allowing the chain to be degraded.

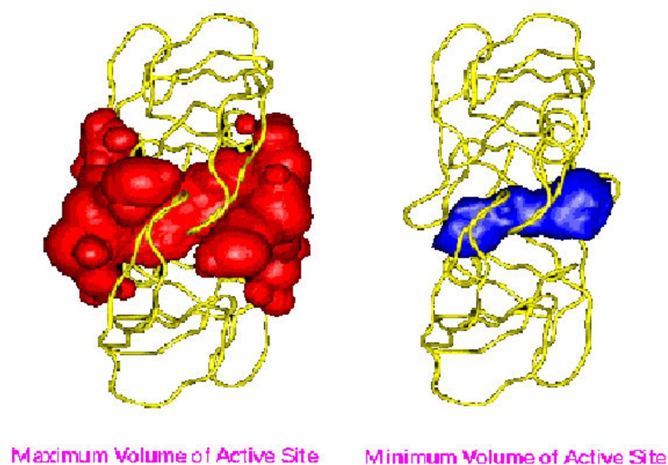


**Figure 2:HIV-I protease**

The illustration in the center demonstrates the flaps in an open conformation and an inhibitor bound to the active site. This is analogous to how a protein chain would be bound during protease activity. When the inhibitor is removed as in the illustration on the right, two aspirate residues can be seen (denoted with asterisks), which attack the protein chain and do all the work.

## CHEMINFORMATICS:

The term Chemoinformatics was defined by F.K. Brown [9, 10] in 1998: Chemoinformatics is the mixing of those information resources to transform data into information and information into knowledge for the considered assay of deriving beneficial decisions faster in the area of drug lead detection and optimization.

Since then, both spellings have been used, and some have evolved to be ascertained as heminformatics, while European Academia settled in 2006 for Chemoinformatics. The current establishment of the Journal of Cheminformatics is an energetic push towards the shorter variant.

Cheminformatics blends the scientific working fields of chemistry and computer science for example in the area of chemical graph theory and mining the chemical space.[11, 12] Cheminformatics can also be employed to data analysis for distinct industries like paper and pulp,dyes and such allied industries.

## COMPUTER AIDED DRUG DISCOVERY:

Artificial Intelligence based drug design supporting systems are employed in this study to identify and in analyzing the structure activity data. The understanding of the specificity of the biological function is based on the principles of molecular recognition. The binding and action of a drug are controlled by the patterns of molecular fields found in the vicinity of the contact surface of the receptor.

Computational methods employed in this study are on the following patterns.

1. Structure Based Drug Designing.

2. Rational (Analogue) Based Drug Designing.

## STRUCTURE BASED DRUG DESIGNING:

A technique in which 3 dimensional structure of the disease causing molecule is used to design drugs that specifically inactivates its function. Structure Based Drug Design is based on a firm understanding of molecular recognition between active site groups and interacting molecules and is a strategy that has become an integral part of modern drug discovery.

Latest advances in Structure Based Drug Design methodologies including flexible, faster docking techniques, virtual screening and library design. In Structure Based Drug Design, the 3D structure of a drug target interacting with small molecules is used to guide drug discovery. “Structure Based Drug Design represents the idea that how molecule interacts with its target protein.” Structure Based Drug Design can help lead to better compounds more quickly.

Structure Based Drug Design is a process whereby the 3D structure of the active site of a biomolecule associated with a disease, as well as the relationship between small molecular structures and their bioactivities are studied. To design a drug against a specific molecule, the exact structure of that molecule must first be found.

Using computers we can scan through the PDB and find compounds, which can dock with the active sites and thus inhibit the activities of these molecules. Certain computer programs can even design molecules, which fit directly into the active sites of the molecules such as enzymes or receptors, thus making very effective inhibitors.

## Materials and Research Methodology:

## PROTEOMICS

- HIV-I Protease Protein Sequence

PQITLWQRPLVTIKIGGQLKEALLDTG

ADDTVLEENSLPGRWKPKMIGGIGGFI

KVRQYDQILIEICGHKAIGTVLVGPTP

VNIIGRNLLTQIGCTLNFAAFXFVVX

ClustalW: Multiple Sequence alignment of HIV-I Protease

Pfam: Pfam is a large collection of multiple alignments takes HIV-I Protease protein sequences from Swissprot database, by using SRS retrieval tool and searches for multiple sequence alignment, against the pfam databases.

Similarity structure or sub structure searches:- Searches the similarity structure or sub structure of HIV-I Protease inhibitor (SKF 108738) by using NCI-3D search database.

Draw the HIV-I Protease Inhibitor:

We have constructed the HIV-I Protease inhibitor 3d structure by using Hyperchem model builder.

Computational methods for HIV-I Protease Inhibitors (SKF 108738):-

We are exercising the hyperchem molecular modeling package for computational methods of SKF 108737

Total energy of molecule:

Submit the 3D structure of SKF 108737 to Hyperchem tool, and then set up the molecular mechanics force field on Hyperchem set up menu. Then click the single point calculations on compute menu of Hyperchem.

Energy minimizations or Geometry optimizations:

We have curtailed the energy of SKF 108737 through by altering the geometry and set up the RMS Gradient at 0.01 in Hyper chem Package. Then click the Geometry Optimization in compute menu of hyper chem.

Molecular Dynamics Simulations:

The solvated molecules of SKF 108737 by using periodic boundary conditions in hyper chem. Then exert the molecular dynamics in compute menu of hyper chem.

- Set up the system
- Starting temperature 100k
- Resulting temperature 300k
- Time step 30k

Monte Carlo Simulations:

- Applied to the Monte Carlo Simulations.
- Starting temperature 100k
- Resulting temperature 300k
- Time step 30k

Conformational Analysis:

We have comported the different conformationals structures of SKF 108738 by using Hyper chem. Molecular modeling package.

Activation energy or Transition state structure:

We have assessed the transition state of SKF 108738 by using Hyper chem. Molecular modeling package.

QSAR Studies:

We have calculated all the QSAR properties of SKF 108738 by using Hyper chem. Molecular modeling packages.

## STRUCTURE BASED DRUG DESIGN:

Structure based drug design is tentative on the existence of a model of the receptor. Many drug receptor interactions are administered by a few key receptors groups. The object pf SBDD is to strive and unite with these groups collectively to engender conformationally rational, pretentiously target molecules. SBDD experimental part is carted out by Ligand Docking (Gold). Taken some chemical compounds (sk-108738, Paraaminobenzoate, Iodopyrazole, Guanine) from GenomeNet Database Service (www.genoe.jp/ligand) by using citation from journal of medicinal chemistry and protein ligand database. Any peculiar bonds (disulfide bridges etc.) should have CONECT records.

If a metal ion is existent, affirm that all bonds between the ion and co-coordinating protein or water molecules are deleated (GOLD will find them automatically). Metals should be within bonding distance of at least two proteins and /or water molecules in the active site so that GOLD can anticipate presumable coordination geometries.

### GOLD: For Docking Set up the GOLD Parameters:-

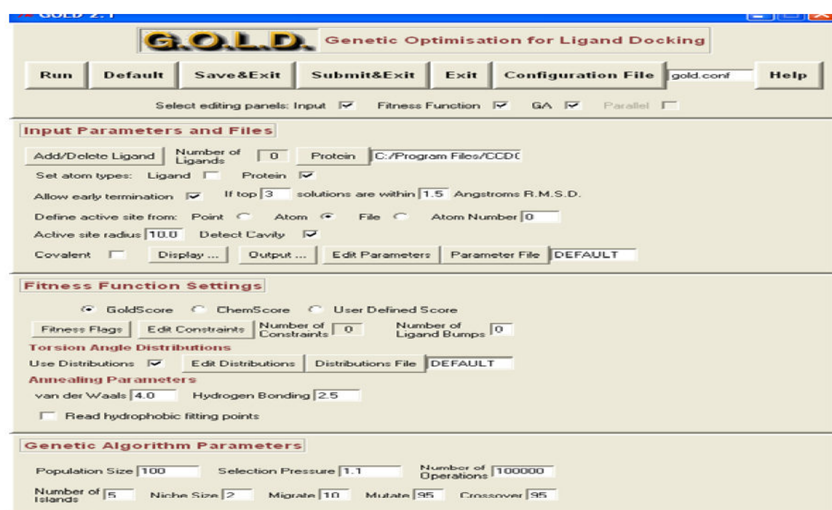


**Figure 4: G.O.L.D. home screen**

General settings

- External energy weight : 1.37
- Flood fill probe radius : 1
- Flood fill point spacing : 0.5
- Length of H bond : 2.7
- Maximum Number of operations :100000

- Doing simplex minimization after every GA run
- Angles coded in 8 bits
- Using ionization dispersion time externally
- Flattening input amide and trigonal nitrogen's
- Solvated point distance : -0.200
- Number of islands : 6
- Population size :100
- Selection pressure :1.2
- Flood fill radius :10
- Flood fill origin
- Cross weight (one string) :90
- Mutate wt (one string): 90
- Migration weight : 10
- Niche size : 2

**Result: MULTIPLE SEQUENCE**

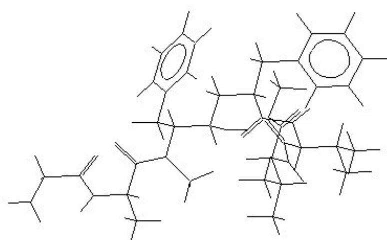

**Figure 4: 3D STRUCTURE OF SKF 108737**

TOTAL ENERGY OF SKF 108738 BY MOLECULAR MECHANICS:

Total Energy = 543887.312500 Gradient = 758719.187600

Molecule Properties:-

RMS Gradient = 7.577e + 005 k/cal/Amol

(Gradient x = 536721.4000 k/cal/Amol, Gradient y = 6134162.5000k/cal/Amol,

Gradient z = 1013716.000k/cal/Amol)

Dipole Moment

Total dipole = 88.85417 Dipole x =87.13437, Dipole y = -11.45754, Dipole z = -13.09165

ENERGY MINIMIZATION OF SKF 108737:-

Energy = 83.005364

Gradient = 0.095556, Converged = yes, Cycles = 962, Points = 2035

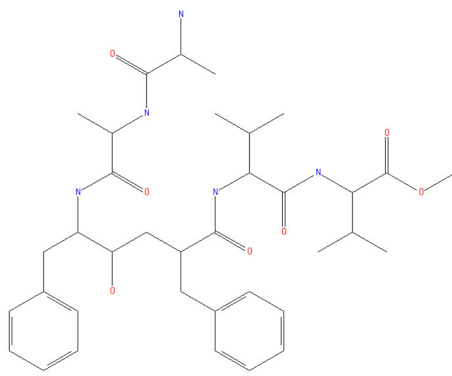

**Figure 5: MINIMISED ENERGY MODEL**

MOLECULAR DYNAMICS SIMULATIONS:

Time = 1.6ps

Total Energy = 88.38042 Kcal/mol, T = 283.467K

At t1=0

Ekin = 29.2076Kcal/mol, Epot = -66.43Kcal/mol

EToT = -52.5204 Kcal/mol, T = 99.98592K

At t2 = 1.6

Ekin = 13.65412Kcal/mol, Epot = -83.00537Kcal/mol

ETOT = -53.79776Kcal/mol, T = 46.74193K

 

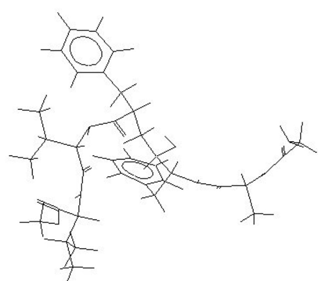

Figure 6: DYNAMICS MOLECULAR MODEL

QSAR STUDIES:-

Surface Area (Appr) = 928.77A

Surface Area (Grid) = 1000.19 A

Volume of the Molecule = 1874.72 A

Log P of the molecule = 2.35

Hydration Energy = 13.52 Kcal/mol

Refractivity of Molecule = 171.18A

Polarisability of Molecule = 70.76 A

Mass of the Molecule = 653.82 amu

Active site Analysis SKF 108738:-

LIGAND CHEMISTRY ANALYSIS

(SKF 108738):-

Donor atoms: 1 6 11 21 33 40

No donor atoms 6

Acceptor atoms: 4 9 21 25 36 43

DOCKING HIV-1 PROTEASE WITH

SKF 108738

Doing GA no population(s) 3 size 100

selection pressure 1.100000

Operation Fitness S (hb_ext) S

 

(vdw_ext) S (hb_int) S (vdw_int)

0 43.91 15.04 53.95

0.00 -226.51

497 45.27 12.35 50.35 0.00 -

177.71

532 46.09 4.95 49.72 0.00 -

133.25

Conclusion:

BEST DOCKING COMPOUND “SKF 108737”

Reference:

1. Davies DR (1990). "The structure and function of the aspartic proteinases". Annu Rev Biophys Biophys Chem 19: 189–215. doi:10.1146/annurev.bb.19.060190.001201. PMID 2194475.

2. Brik A, Wong CH (January 2003). "HIV-I protease: mechanism and drug discovery". Org. Biomol. Chem. 1 (1): 5–14. doi:10.1039/b208248a. PMID 12929379.

3. Kräusslich HG, Ingraham RH, Skoog MT, Wimmer E, Pallai PV, Carter CA (February 1989). "Activity of purified biosynthetic proteinase of human immunodeficiency virus on natural substrates and synthetic peptides". Proc. Natl. Acad. Sci. U.S.A. 86 (3): 807–11. doi:10.1073/pnas.86.3.807. PMID 2644644.

4. Kohl NE, Emini EA, Schleif WA, et al. (July 1988). "Active human immunodeficiency virus protease is required for viral infectivity"

5. Seelmeier S, Schmidt H, Turk V, von der Helm K (September 1988). "Human immunodeficiency virus has an aspartic-type protease that can be inhibited by pepstatin A". Proc. Natl. Acad. Sci. U.S.A. 85 (18): 6612–6. doi:10.1073/pnas.85.18.6612. PMID 3045820.

6. McPhee F, Good AC, Kuntz ID, Craik CS (October 1996). "Engineering human immunodeficiency virus 1 protease heterodimers as macromolecular inhibitors of viral maturation". Proc. Natl. Acad. Sci. U.S.A.

93 (21): 11477–81.

7. Jaskólski M, Tomasselli AG, Sawyer TK, Staples DG, Heinrikson RL, Schneider J, Kent SB, Wlodawer A (February 1991). "Structure at 2.5-A resolution of chemically synthesized human immunodeficiency virus type 1 protease complexed with a hydroxyethylene-based inhibitor". Biochemistry 30 (6): 1600–9. doi:10.1021/bi00220a023. PMID 1993177.

8. Miller M, Schneider J, Sathyanarayana BK, Toth MV, Marshall GR, Clawson L, Selk L, Kent SB,  lodawer A (December 1989). "Structure of complex of synthetic HIV-I protease with a substrate-based inhibitor at 2.3 A resolution". Science (journal) 246 (4934): 1149–52. PMID 2686029. http://www.sciencemag.org/cgi/pmidlookup?view=long&pmid=2686029.

9. F.K. Brown (1998). "Chapter 35. Chemoinformatics: What is it and How does it Impact Drug Discovery". Annual Reports in Med. Chem. 33: 375. doi:10.1016/S0065-7743(08)61100-8.

10. Brown, Frank (2005). "Editorial Opinion: Chemoinformatics – a ten year update". Current Opinion in Drug Discovery & Development 8 (3): 296–302.

11. Gasteiger J.(Editor), Engel T.(Editor): Chemoinformatics : A Textbook. John Wiley & Sons, 2004, ISBN 3-527-30681-1

12. A.R. Leach, V.J. Gillet: An Introduction to Chemoinformatics. Springer, 2003, ISBN 1-4020-1347-7

This academic article was published by The International Institute for Science, Technology and Education (IISTE). The IISTE is a pioneer in the Open Access Publishing service based in the U.S. and Europe. The aim of the institute is Accelerating Global Knowledge Sharing.

More information about the publisher can be found in the IISTE's homepage: http://www.iiste.org

## CALL FOR PAPERS

The IISTE is currently hosting more than 30 peer-reviewed academic journals and collaborating with academic institutions around the world. There's no deadline for submission. **Prospective authors of IISTE journals can find the submission instruction on the following page:** http://www.iiste.org/Journals/

The IISTE editorial team promises to the review and publish all the qualified submissions in a **fast** manner. All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Printed version of the journals is also available upon request of readers and authors.

**IISTE Knowledge Sharing Partners**

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar