International Journal of Innovative Research and Advanced Studies (IJIRAS)
Volume 4 Issue 1, January 2017

ISSN: 2394-4404

Molecular Docking Studies; Dietary Compounds As HIV 1 Protease Inhibitors

Jisha Undarathi

Research Scholar, Department of Chemistry,
Kannur University

Abstract: ***The study indicates the importance of diet in the prevention and treatment of diseases. Five molecules that "Docked" well into the active site of the target, HIV 1 Protease are selected. And they are considered to be as the "Hit Molecules". Glibenclamide, Lectin, Resveratrol, Oleuropein, Silymarin are considered as the hit molecules. As they are docked into the active site of the target with a minimum energy (found to be stable). So these molecules can be suggested to be the interesting candidates for further testing in the laboratory. Finally, this study strongly underscores the importance of computational approaches in drug discovery, supplementing classical methods, thus saving enormous amount of time and money.***

Index Terms: ***Acquired Immunodeficiency Syndrome, ArgusLab, Bioinformatics, Computer-Aided Drug Design, Hit Molecules, HIV 1 Protease, Inhibitor, Molecular docking, Pharmacophores, QSAR, Virtual Screening.***

I. INTRODUCTION

A. BIOINFORMATICS

The term bioinformatics was coined by Paulien Hogeweg and Ben Hesper in 1978 for the study of informatic processes in biotic systems. Bioinformatics is a rapidly developing branch of biology that derives knowledge from computer analysis of biological information. Bioinformatics is the science of developing computer databases and algorithms for the purpose of speeding up and enhancing biological research. It is a scientific discipline that comprises all aspects of the gathering, storing, handling, analyzing, interpreting, and spreading of biological information.

It involve the use of techniques from applied mathematics, information, statistics and computer science to solve biological problems. It has many practical applications in different areas of biology and medicine [1][2].

B. BRANCHES OF BIOINFORMATICS

Bioinformatics has mainly three branches: 1) Molecular bioinformatics, 2) Cellular and sub cellular bioinformatics and 3) Organic and community bioinformatics.

Subsidiary branches of these are:

- ✓ Molecular bioinformatics:
  - Genomics
  - Proteomics
  - Drug design
- ✓ Cellular and sub-cellular bioinformatics:
  - Metabolic pathway
  - Epigenetic
  - Neuro bioinformatics
- ✓ Organic and community bioinformatics:
  - Bioinformatics of species diversity
  - Behaviour, evolution and effects of
  - Pollutanats

C. AIMS AND TASKS OF BIOINFORMATICS

Bioinformatics has three components: 1) the creation of databases allowing the storage and management of large biological data sets, 2) the development of algorithms and statistics to determine relationships among members of large datasets and 3) the use of these tools for the analysis and interpretation of various types of biological data [3][4].

AIMS

- ✓ To organize data in a way that follows researches to access existing information and to submit new entries as they are produced.
- ✓ To develop tools and resources that aid in the analysis and management of data.
- ✓ To use these tools to analyse the data and interpret the results in a biologically meaningful manner.

TASKS

The tasks in bioinformatics involve the analysis of sequence information. This involves the following:

- ✓ Identifying the genes in the DNA sequence from various organism.
- ✓ Developing methods to study the structure and function of newly identified sequences and corresponding structural RNA sequences.
- ✓ Identifying families of related sequences and the development of models.
- ✓ Aligning similar sequences and generating phylogenetic trees to examine evolutionary relationships.

Besides these, one of the important dimensions of bioinformatics is identifying drug targets and pointing out lead compounds.

D. APPLICATIONS OF BIOINFORMATICS

Biocomputing has found its application in many areas, These include the following areas:

- ✓ Genome and sequence analysis
- ✓ Protein modeling
- ✓ Molecular medicine
- ✓ Functional genomics
- ✓ Proteomics
- ✓ Personalized medicine and Gene therapy
- ✓ Antibiotic resistance
- ✓ Crop improvement
- ✓ Bioweapons
- ✓ Drug designing [5][6]

E. DRUG DESIGNING

Drug design is the approach of finding drugs, based on their biological targets. Drug discovery can be arrived by two methods. The empirical and rational. The empirical method is a bind and hit or lose method. Thousands of chemical compounds are tested on the disease without even knowing the target on which the drug acts and the mechanism of action. Advent of bioinformatics has changed this paradigm to rational approach. Rational approach starts from the clear knowledge of the target as well as the mechanism by which it is to be attacked [7].

F. BIOINFORMATICS AND DRUG DESIGNING

Drug designing is one of the important area in bioinformatics. The processes of designing a new drug using bioinformatics tools have open a new area of research. Computers are used to gather, store, analyse and integrate biological and genetic information which can then be applied to gene-based drug discovery and development. Bioinformatics helps to identify and validate drug targets, to find and design molecules that uniquely interact with a specific protein target and to optimize molecule [8][9].

G. COMPUTER-AIDED DRUG DESIGN

Computer-aided drug design is a theoretical tool that can be used to identify novel potential drugs. The aim of using the computer for drug design is to analyze the interactions between the drug and its receptor site and to "design" molecules that can give an optimal fit. The principle is that a good fit results from structural and chemical complementary to the target receptor. The techniques provided by computational methods include computer graphics for visualization and the methadology of theoretical chemistry. By means of quantum mechanics the structure of small molecules can be predicted to experimental accuaracy. Statistical mechanics permits molecular motion and solvent effects to be incorporated.

CADD, Computer Aided Drug Design is a discipline that uses computational methods to stimulate drug-receptor interactions. CADD methods are heavily depend on bioinformatics tools, applications and databases. The CADD helps to reduce the time and expense of the drug discovery than the conventional drug discovery methods which used the trail and error method for the discovery of new drugs [10].

BENEFITS OF CADD

CADD methods and bioinformatics tools offer significant benefits for drug discovery programs.

COST SAVINGS: It suggests that the cost of drug discovery and development has reached 800 million for each drug successfully brought to market. Many biopharmaceutical companies now use computational methods [11] and bioinformatics tools to reduce this cost burden. Virtual screening, lead optimization and predictions of bioavailability and bioactivity can help guide experimental research. Only the most promising experimental lines of inquiry can be followed and experimental dead-ends can be avoided early based on the results of CADD simulations [12].

TIME-TO-MARKET: The predictive power of CADD can help drug research program to choose only the most promising drug candidates. By focusing drug research on specific lead candidates and avoiding potential "dead-end" compounds, biopharmaceutical companies can get drugs to market more quickly.

ISSN: 2394-4404

INSIGHT: one of the non-quantifiable benefits of CADD and the use of bioinformatics tools is the deep insight that researches acquire about drug-receptor interactions. Molecular models of drug compounds can reveal intricate, atomic scale binding properties that are difficult to envision in any other way. This is an intangible benefit that can help design research programs. Therefore CADD and bioinformatics together is a powerful combination in drug research and development.

H. CONCEPT OF DRUG AND TARGETS

TARGETS

The first step and a challenging task in drug discovery is to identify the target. Drug targets are the macromolecular structures like proteins and nucleic acids, which play crucial roles in the physiology and metabolism of an organism. Targets are molecules that are strictly involved in affecting process that is crucial for the existance of an organism. It is critical to a disease that may be targeted with potential therapeutic agent. Target classification can be done by using several bioinformatics approaches [13]. It may be single molecular entities where inputs and outputs of the system may be visible, but the actual processes taking place in the system may be not be discernible. Membrane protein is an important class of drug targets that can be analyzed and validated by proteomic technologies.

LIGAND

Ligand is any chemical molecule that acts on a biological molecule to bring some effect, which could be positive or negative [14]. It is a molecule that binds with another through non covalent forces that is usually does not involve chemical bond formation. It should have the ability to inhibit the activity of the specific biological molecule. Ligands can be downloaded from various small molecule databases.

ACTIVE SITE

The active site of an enzyme is the binding site where catalysis occurs. The active site of an enzyme is the region that binds the substrate and contributes the amino acid residues that directly participate in the making and breaking of chemical bonds. The structure and chemical properties of the active site allow the recognition and binding of the substrate. The active site is usually a small pocket at the surface of the enzyme that contains residues responsible for the substrate specificity (charge, hydrophobicity, steric hindrance) and catalytic residues which often act as proton donors or acceptors or are responsible for binding a cofactor. The active site is often the site of inhibition of enzyme.

These are several models of how enzyme work: The lock-and-key model [15] and the induced fit model. Substrates bind to the active site of the enzyme or a specificity pocket by hydrogen bonds, hydrophobic interactions, temporary covalent bond or a combination. Residues of the active site will act as donors or acceptors protons or other groups on the substrate to facilitate the reaction.

An enzyme inhibitor is a molecule that binds to an enzyme and decreases its rate of reaction. Many drugs are enzyme inhibitors.

I. STRUCTURE BASED DRUG DESIGN

The majority of drugs on the market today to treat disorders in humans, animals and plants were discovered either by chance observation or by systematic screening of large series of synthetic and natural substances. The traditional method of drug discovery is now supplemented by methods exploiting the increasing knowledge of the molecular targets assumed to participate in some disorder, computer technology and the physical principle underlying drug-target interactions. Rational drug design- traditional method were or are not irrational-or-better "structure based ligand" continue to increase in importance in the endeavor of promoting a biologically active ligand towards the status of a drug useful in human and veterinary medicine and the phytopharmaceutical world.

Structure based drug design is a very powerful approach in drug design and is most effective, when 3D structure of an existing inhibitor complex with its target is known. This technique has played a major role in designing a number of drug candidates that have progressed to clinical trials. A requirement for this approach is an understanding of the principles of molecular recognition in the protein ligand complexes. It involves proposing or evaluating novel ligand to biologal receptors prior to synthesis, based on structure information and provides a rapid and controlled exploration of the geometric intricacies of the target sites.

"Structure-based drug design represents an idea that we can see exactly how our molecules interact with its target protein". The structural information can be obtained with X-ray crystallography (NMR). Ideally, these two techniques complement one another.

Originally, structure-based drug design was equated with De novo design or building a molecule from ground up [16]. The active site of the protein was a space to be filled with a molecule that complemented in terms of shape, charge and other binding components. One of the driving forces behind structure based drug design is lead optimization. Structure is really good way of quickly getting a handle on how the lead compound bind to the target and what one able to do with chemistry to modify the molecule to get desired properties (Milburn). A major development that structure based method of a place of prominence in drug discovery has been increased speed.

J. LIGAND BASED DRUG DESIGN

Ligand based (analog based) drug design is a computer aided drug designing. It is based on a set of known ligands and there will not be any structural information about the receptor. This mainly uses pharmacophore maps and quantitative structure-activity relationship (QSAR) to identify or modify a lead in the absence of a known 3D structure of a receptor or target.

ISSN: 2394-4404

K. PHARMACOPHORES

It is a hypothesis of the critical features of a ligand. The features include Hydrogen-bond donors and acceptors, charged groups and hydrophobic patterns. These can be used to screen databases for compounds and to refine existing leads.

L. QSAR

The goal of this is to predict the activity of new compounds based on their chemical structure. Information about the strength of interactions of each compound can be calculated by steric, electronic, and hydrophobic descriptors. It is a mathematical relationship used to determine how the structural features of a molecule are related to biological activity [17].

M. VIRTUAL SCREENING

In rational drug design, protein structural data is used to predict the type of ligands that will interact with a given target. The identification of lead compounds is based on the high throughput screening. Virtual screening is a drug discovery method for the identification of new lead compounds against a drug target. Virtual screening is to computationally screen large libraries of chemical compounds that are complement to large structure and experimentally test those that are bind well with target. It involves protein structure based compound screening or docking and chemical similarity search based on small molecules. Virtual screening technologies are used in high throughput docking, homology searching and pharmacophore searches of 3D databases [18].

N. MOLECULAR DOCKING

In the field of molecular modeling, docking is a method which predicts the preferred orientation of one molecule to a second, when bound to each other to form a stable complex [19]. Knowledge of the preferred orientation in turn may be used to predict the strength of association or binding affinity between two molecules.

The association between biologically relevant molecules such as proteins, nucleic acids, carbohydrates, and lipids play a central role in signal transduction. Furthermore, the relative orientation of the two interacting partners may affect the type of signal produced (e.g.,agonism vs antagonism). Therefore docking is useful for predicting both the strength and type of signal produced.

Docking is frequently used to predict the binding orientation of small molecule drug candidates to their protein targets in order to predict the affinity and activity of the small molecule. Hence docking plays an important role in the rational design of drugs [20]. Giving the biological and pharmaceutical significance of molecular docking, considerable efforts have been directed towards improving the methods used to predict docking.

DOCKING GLOSSARY

RECEPTOR OR HOST OR LOCK – The "receiving" molecule, most commonly a protein or other biopolymer.

LIGAND OR GUEST OR KEY – The complementary partner molecule which binds to the receptor. Ligands are most often small molecules but could also be another biopolymer.

DOCKING – Computational simulation of a candidate ligand binding to a receptor.

BINDING MODE – The orientation of the ligand relative to the receptor as well as the conformation of the ligand and receptor when bound to each other.

POSE – A candidate binding mode.

SCORING – The process of evaluating a particular pose by counting the number of favorable intermolecular interactions such as hydrogen bonds and hydrophobic contacts.

RANKING – The process of classifying which ligands are most likely to interact favorably to a particular receptor based on the predicted free energy of binding.

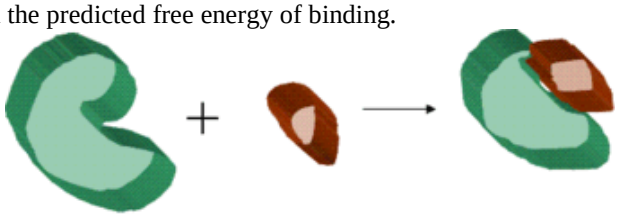

Figure 1: Schematic diagram illustrating the docking of a small molecule ligand (brown) to a protein receptor (green) to produce a complex

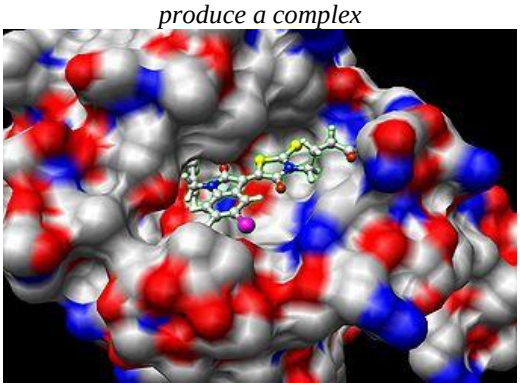

Figure 2: Small molecule docked to a protein

Molecular docking is a study of how two or more molecular structures, for example drug and enzyme or receptor of protein, fit together. In other words, the problem is like solving a 3D puzzle. For example, the action of a harmful protein in human body may be prohibited by finding an inhibitor, which binds to that particular protein. Molecular docking softwares are mainly used in drug research industry. The most important application of docking software is virtual screening. In virtual screening the most interesting and promising molecules are selected from an existing database for further research. This places demands on the used computational method; it must be fast and reliable [21].

Molecular docking is used to predict the structure of the intermolecular complex formed between two or more molecules. The most interesting case is the protein ligand interaction, because of its applications in medicine. Generic protein-protein interactions differ from protein-ligand

interactions due to the small size of ligand [22]. Because of their large size, proteins are usually treated as rigid bodies. However, conformational changes in the protein and the ligand are often necessary for a successful docking process.

THE MECHANICS OF DOCKING

To perform a docking screen, the first requirement is structure of protein. Usually the structure has been determined in the lab using a biophysical technique such as x-ray crystallography, or less often, NMR. This protein structure and databases of potential ligands serve as inputs to a docking program. The success of a docking program depends on two components: the search algorithm and scoring function.

THE SEARCH ALGORITHM

The search space consists of all possible orientations and conformations of the protein paired with the ligand. With present compounding resources, it is impossible to exhaustively explore the search space-this would involve enumerating all possible distortions of each molecule and all possible rotational and translational orientations of the ligand relative to the protein at a given level of granularity. Most docking programs in use account for a flexible ligand, and several are attempting to model a flexible protein receptor. Each "snapshot" [23] of the pair is referred to a pose, there are many strategies for sampling the search space. Here are some examples:

- ✓ Use a coarse-grained molecular dynamics simulation to propose energetically reasonable poses.
- ✓ Use a "linear combination" of multiple structures determined for the same protein to emulate receptor flexibility.
- ✓ Use a genetic algorithm to "evolve" new poses that are successively more and more likely to represent favourable binding interactions.

THE SCORING FUNCTION

The scoring function takes a pose as input and returns a number indicating the likelihood that the pose represent a favourable binding interaction. Most scoring functions are physics-based molecular mechanics force field that estimate the energy of the pose; a low energy indicates a stable system and thus a likely binding interaction. An alternative approach is to derive a statistical potential for interactions from a large database of protein-ligand complex's, such as Protein Data Bank, and evaluate the fit of the pose according to this inferred potential.

All scoring function used in docking will yield a large number of false positive hits, i.e., ligands predicted to bind to the protein that actually donot when placed together in a test tube. One way to reduce the number of false positives is to regulate the energy of the top-hit poses using a higher resolution (and therefore slow) technique like Generalized Born or Poisson-Boltzmann methods [24]. However, typically the researcher will screen a database of ten to hundreds of thousands of compounds and test the top 60 or so in vitro, and to identify any true binders is still considered a success.

O. APPLICATIONS

A binding interaction between a small molecule ligand and an enzyme protein may result in activation or inhibition of the enzyme. If the protein is a receptor, ligand binding may result in agonism or antagonism. Docking is most commonly used in the field of drug design - most drugs are small organic molecules, and docking may be applied to:

- ✓ *HIT IDENTIFICATION* – Docking combined with a scoring function can be used to quickly screen large databases of potential drugs in silico to identify molecules that are likely to bind to protein target of interest [25].
- ✓ *LEAD OPTIMIZATION* – Docking can be used to predict in where and in which relative orientation a ligand binds to a protein. This information may in turn be used to design more potent and selective analogs.
- ✓ *BIOREMEDIATION* – Protein ligand docking can also be used to predict pollutants that can be degraded by enzymes.

P. DOCKING SOFTWARES

A number of softwares are for the process of docking. Some of them are:

ArgusLab: Arguslab is a freely available tool for various biological simultations. It was developed by Dr. Mark T Hompson and Planaria Software LLC. It is designed for windows platform. As the price of commercial software is rising Arguslab is boon for beginners to start up with. IT provides almost all facilities same as that of the commercial softwares but the accuracy level is low. It promotes the idea of open source programming. The entire code of the program can be downloaded and if required changes can be made and can be send back to the company. The users are allowed to fix the bugs. It has got the minimal system requirements. It can be accessed from http//www>arguslab>com. In arguslab before docking we should remove water contents and add hydrogen atom and then minimized the energy of protein. After protein preparation, open the ligand and we can do docking. It can perform calculations like single-point energies, geometry optimization, dipolemoments, electronic absorption spectra etc.

AUTO DOCK: Auto Dock is a suite of automated docking tools. Designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D structure. The program AutoDock was originally written in1990 by DAVID S. Goodsell.

DOCK: Dock is one of the oldest and best known ligand-protein docking program. It is useful for fast docking, but it is not the most accurate software available.

FLEXX: It is extremely fast, robust computer program. It predicts the geometry of the protein-ligand complex. Complex prediction and virtual screening are the two main applications of FlexX.

GOLD: GOLD (Genetic Optimization for Ligand Docking) is a genetic algorithm for docking flexible ligands into protein binding sites. It is a program for calculating the docking modes of small molecules into protein binding sites.

FRED: It is a highly accurate, multiconformer docking program. It examines all possible poses within a protein active site.

II. AIM AND OBJECTIVES

- ✓ To apply computational techniques in drug design against Human Immunodeficiency Virus.
- ✓ To validate the anti-viral potential of dietary compounds.
- ✓ To identify dietary compounds that can structurally bind to the HIV 1 Protease.

III. SCOPE

The primary goal of bioinformatics is to increase the understanding of biological processes. It focus on developing and applying computationally intensive techniques (e.g., pattern recognition, data mining, machine learning algorithms, and visualization) to achieve this goal. Major research efforts in the field include sequence alignment, gene finding, genome assembly, drug design, drug discovery, protein structure alignment, protein structure prediction, prediction of gene expression and protein–protein interactions, genome-wide association studies and the modeling of evolution.

A. ACQUIRED IMMUNODEFICIENCY SYNDROME (AIDS)

In June 5,1981, scientists in the United States and France first recognized the Acquired Immuno Deficiency Syndrome (AIDS)**.** The emergence and pandemic spread of AIDS have posed the greatest challenge to public health in modern times. HIV; the Human Immunodeficiency Virus is the etiological agent of AIDS. AIDS is a medical condition where the immune system cannot function properly and protect the body from disease. It is an infectious disease caused by human immunodeficiency virus that causes immune system failure and debilitation. The virus can lie hidden in the body for up to 10 years without producing any obvious symptoms or before developing into the AIDS disease, and in the meantime the person can unknowingly infect others. Currently, an estimated 40 million people worldwide are HIV carriers, and three million a year are dying of AIDS.

HIV is spread through direct contact with the bodily fluids of an infected person. It is a fatal. AIDS is a measure of how much damage HIV has done to person's immune system. It is not a disease. AIDS can develop after someone gets HIV. After HIV has been inside someone's body for a long time it can weaken or destroy their immune system. The immune system can't fight germs any more. They get different disease or illnesses, also called opportunistic infections [26].

B. STRUCTURE OF HIV

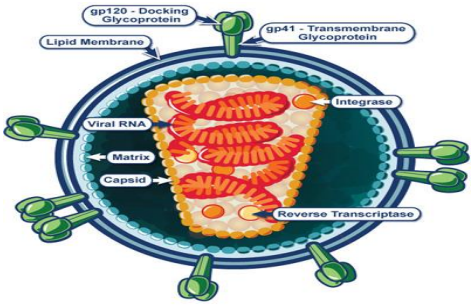

Figure 3

HIV is a spherical enveloped virus, about 90-120 nm in size. The nucleocapside has an outer icosahedral shell and an inner cone shaped core, enclosing the RNA. The genome is diploid, composed of two identical single stranded, positive sense RNA copies. In association with viral RNA, reverse transcriptase enzyme is present in HIV. HIV is a highly mutable and thermolabile. It exhibit frequent antigenic variations [27].

By 1985 HIV was a recognised retrovirus, but then it was observed that the few AIDS cases in West Africa were associated with a different virus from what had become known as HIV. The new virus was called HIV-2.

Both HIV-1 and HIV-2 have similar effects and almost exactly the same set of genes. But HIV-2 is strictly more like certain types of Simian Immunodeficiency Virus (SIV) than it is like HIV-1. Both HIV-2 and SIV carry an extra gene not found in HIV-1, which encodes viral protein X. And comparing the viral proteins of HIV-2 and SIV, only subtle differences in molecular sizes can differentiate them. Once HIV 1 is integrated into cell, faster it produces new virus particles. So HIV-1 is more pathogenic than HIV-2.

C. GENES PRESENT IN HIV

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 how the genes are laid out in HIV [28].

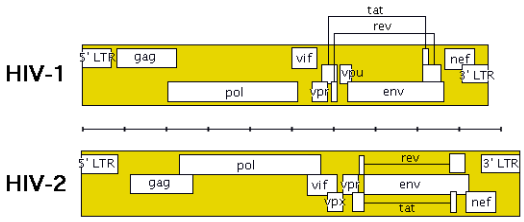

Figure 4

The genome of HIV contains three stuctural genes (gag, pol and env) and nonstructural and regulatory genes (tat, rev, nef, vif, vpr, LTR)

GAG: The group antigen gene is found in all retroviruses. It makes various proteins necessary to protect the virus. In HIV, it has three parts: MA (matrix), CA (capsid), and NC (nucleocapsid).

ISSN: 2394-4404

POL: The polymerase gene is also found in all retroviruses. It makes enzymes necessary for virus replication. In HIV, it also has three parts: PR (protease), IN (endonuclease), and RT (reverse transcriptase).

ENV: The envelope gene is also found in all retroviruses. It makes proteins for the envelope to the virus. In HIV, it has two parts. SU (surface envelope, gp120) and TM (transmembrane envelop, gp41).

TAT: The transactivator gene influences the function of genes some distance away. It controls transactivation of all HIV proteins.

REV: The differential regulator of expression of virus protein genes.

VIF: The virus infectivity factor gene is required for infectivity as cell-free virus.

NEF: The negative regulator factor retards HIV replication.

VPR: The virus protein R gene has an undetermined function.

VPU: The virus protein U gene is required for efficient viral replication and release. It is found only in HIV-1.

VPX: The virus protein X gene has an undetermined function. It is found only in HIV-2 and SIV.

The HIV genome also has a "Long Terminal Repeat" (LTR) at each end of its genome-not quite a gene, but a sequence of RNA/DNA which is the same at either end and which serves some structural and regulatory purposes.

D. HIV 1 PROTEASE AS DRUG TARGET

HIV-1 protease (HIV PR) is a retroviral aspartyl protease that is essential for the lifecycle of HIV, the retrovirus that causes AIDS. HIV PR cleaves newly synthesized polyproteins at appropriate places to create the mature protein components of an infectious HIV virion. Without effective HIV PR, HIV virions remain uninfectious. Thus, mutation of HIV PR's active site or inhibition of its activity disrupts HIV's ability to replicate and infect additional cells, making HIV PR inhibition the subject of much pharmaceutical research.

Due to its time-sensitivity and crucial role, HIV protease is an effective target for drug therapy [29][30].

STRUCTURAL FEATURES

HIV-1 Protease is a homodimer. Each monomer contains 99 amino acids and is identical in conformation. The position of each monomer in the active protease forms an axis of symmetry. The secondary structure of each monomer includes, one alpha helix and two anti parallel beta sheets.

The two Asp25 residues (one from each chain) act as the catalytic residues. According to the mechanism for HIV PR protein cleavage proposed 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. Additionally, HIV PR has two molecular "flaps" which move a distance of up to 7 Å when the enzyme becomes associated with a substrate [31][32].

Aliphatic residues stabilize each monomer in a hydrophobic core. Additionally, the dimmer is stabilized by covalent interactions, hydrophobic packing of side chains and interactions involving the catalytic residues. Each monomer contains two cysteine residues, but these do not form disulfide bonds.

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

E. MECHANISM OF ACTION

STRUCTURE

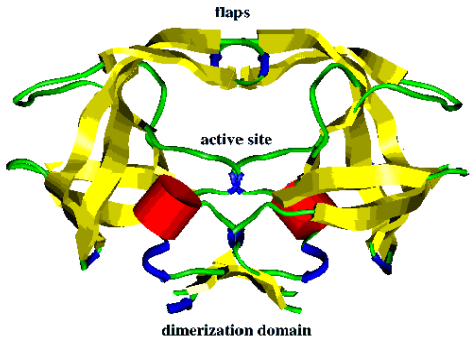

Figure 5

HIV-1 protease activity is critical for the terminal maturation of infectious virions. Once HIV enters the cell, viral RNA undergoes reverse transcription to produce double-stranded DNA. This viral DNA is integrated into the host genome and eventually, transcribed and translated by cellular enzymes to produce large, nonfunctional polypeptide chains, referred to as polyproteins. After these poly-proteins are assembled and packaged at the cell surface, immature virions are produced and released into the plasma. At this point, HIV-1 protease, acting as a "molecular scissors," cleaves the poly-proteins into smaller, functional proteins, thereby allowing the virion to mature**.** In the presence of HIV-1 protease inhibitors, the virion is unable to mature.

F. HIV-1 PROTEASE ACTIVE SITE BINDING

HIV-1 protease consists of two protein chains. When the two chains assemble, a long tunnel is formed protein "flaps" cover the tunnel and open up to allow the enzyme to attach to a protein chain. After attachment, the flaps then close around the protein chain, thereby holding it in the tunnel and allowing the chain to be degraded.

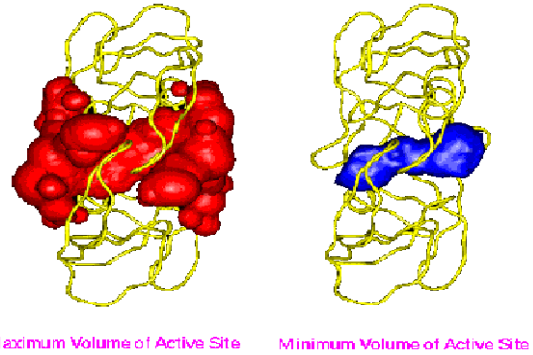

Figure 6

The illustration on the left shows the flaps in an open conformation and an inhibitor bound to the active site. This is similar 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, which attack the protein chain and do all the work.

IV. REVIEW OF LITERATURE

Human Immunodeficiency Virus (HIV) is a lentivirus (a member of the retrovirus family) that causes Acquired Immuno Deficiency Syndrome (AIDS), a condition in humans in which the immune system begins to fail, leading to life-threatening opportunistic infections. To have AIDS, one must have tested positive on an HIV test and have another disease that is known as an "AIDS defining disease". These diseases include: yeast infections, cervical cancer, kaposis sarcoma, tuberculosis, cytomegalovirus, and pneumonia [33]. While some of the "AIDS defining" illnesses and conditions can be fatal, many of them are not. Some of these conditions can actually cause a person to receive a "positive test result". AIDS is now a pandemic. Today, 33.3 million people are living with HIV and AIDS worldwide.

MODE OF TRANSMISSION OF HIV

- ✓ Sexual route: The majority of HIV infections are acquired through unprotected sexual relations. Sexual transmission occurs when there is contact between sexual secretions of one partner with the rectal, genital or oral mucous membrane of another.
- ✓ Mother-to-child transmission: The transmission of the virus from the mother to the child can occur in utero during the last weeks of pregnancy and at childbirth. Breast feeding also presents a risk of infection for the body.
- ✓ Blood or blood product route: This transmission route is particularly important for intravenous drug users, hemophiliacs and recipients of blood transfusions and blood products.

HIV has been found at low concentrations in the saliva, tears and urine of infected.

PATHOGENESIS

HIV contains spikes composed of glycoprotein 120 on its envelop. It enters the cell by specific binding of this glycoprotein to the cell receptor. The receptor for the virus is the CD4 antigen on the surface. The fusion of virus is brought about by the transmembrane gp41. After fusion, the HIV genome uncoated and internalized in to the cell. Viral reverse transcriptase mediates transcription of its RNA to double stranded DNA, which is integrated in to genome of the infected cell through the action of integrase, causing a latent infection. HIV first destroys the CD4 cells. Through the damage cellular immunity, humoral mechanism is also affected. AIDS patients are unable to respond to new antigens [34][35].

HIV SYMPTOMS

Many people have no symptoms when they first get HIV-some have no symptoms for years. It varies from person to person. But some people get a flu-like illness within a month or two after getting HIV. The flu-like-symptoms-fever, headache, fatigue, swollen lymph nodes often go away within a week. Even if there are no symptoms, HIV can still be passed to another person. HIV is never diagnosed by the symptoms. As the HIV infection spreads through the body, it starts to feel sick. For many people, the first symptom they notice is large lymph nodes that may be enlarged for more than 3 months. Other symptoms that follow may include:

- ✓ Being very tired
- ✓ Quick weight loss
- ✓ Fever
- ✓ Night sweats
- ✓ Headache
- ✓ Nausea
- ✓ Vomiting
- ✓ Diarrhea

There are also other symptoms that are more common, serious and harder to treat in women with HIV:

- ✓ Vaginal yeast infections
- ✓ Other vaginal infections such as bacterial vaginosis and STDs like gonorrhea, chlamydia and trichomoniasis
- ✓ Pelvic Inflammatory Disease (PID) or infection of a women's reproductive organs
- ✓ Menstrual cycle changes, like not having periods or having heavy and constant bleeding
- ✓ Human papilloma Virus (HPV) infections that cause genital warts and can lead to cervical cancer

As the immune system continues to weaken, other disease and infections can develop that affect your eyes, digestive system, kidneys, lungs, skin and brain. A healthy body is equipped with CD4 helper lymphocyte cells (CD4 cells). These cells help the immune system function normally and fight off certain kinds of infections. They do this by acting as messengers to other types of immune system cells, telling them to become active and fight against an invading germ [36]. HIV attaches to these CD4 cells, infects them. In doing so, the virus destroys the ability of the infected cells to do their job in the immune system. The body then loses the ability to fight many infections.

TREATMENTS THAT SUPPRESS HIV

When AIDS first surfaced in the United States, no drugs were available to combat the underlying immune deficiency, and few treatments existed for the opportunistic infections [37] that resulted. Over the past 10 years, however, therapies have been developed to fight both HIV infection and its associated infections and cancers.

- ✓ Some medications target HIV itself, to reduce the virus's assault on the immune system, or to even prevent the virus from entering human immune cells.
- ✓ Other treatments are used to treat or prevent specific opportunistic infections that threaten the health of people with HIV-damaged immune systems.

HIV-1 protease (HIV PR) is a homodimeric enzyme, that comprise the human immunodeficiency virus. Competitive inhibitors are used to bind to the protease and block its function, thereby suppressing the virus, which cannot transform to its mature, infectious form. HIV 1 Protease inhibitors are one class of drugs that target the viral enzyme HIV-1 protease. They are drugs used to treat or prevent infection by HIV. It prevent protease from splitting proteins into peptides. That is, they prevent the cleavage of HIV precursor proteins into active proteins. Drugs that interfere with the activity of retrovirus HIV are generally known as antiretrovirals [38]. There is no cure for HIV/AIDS, but a variety of drugs can be used in combination to control the virus. Each of the classes of anti-HIV drugs blocks the virus in different ways. It's best to combine at least three drugs from two different classes to avoid creating strains of HIV that are immune to single drugs. The classes of anti-HIV drugs include [39]:

- ✓ Nucleoside analog reverse transcriptase inhibitors (NRTIs): NRTIs interrupt an early stage of HIV replication. AZT (zidovudine), the first drug approved for treating HIV infection, is an NRTI.
- ✓ Non-nucleoside reverse transcriptase inhibitors (NNRTIs): This class of drugs includes delavirdine, nevirapine, and efavirenz.
- ✓ Protease inhibitors: Protease inhibitors interrupt a later stage of viral replication. This class of drugs includes saquinavir, indinavir, ritonavir, nelfinavir, and amprenavir.
- ✓ Fusion inhibitors: Fusion inhibitors prevent HIV from entering human immune cells. The only fusion inhibitor approved to date is enfuvirtide.

Some, Dietary compounds that are used as drugs; like silymarin, lectin, glibenclamide, resveratrol, oleuropein etc; are used as HIV 1 protease inhibitors. They greatly inhibit the function of the enzyme. Also they have no adverse effect as carcinogenicity and mutagenicity. Various bioactivities of dietary compounds are responsible for their chemopreventive properties and also contribute to their inducing apoptosis by arresting cell cycle, regulating carcinogen metabolism and ontogenesis expression, inhibiting DNA binding and cell adhesion, migration, proliferation or differentiation, and blocking signaling pathways.

A. INHIBITORS OF HIV 1 PROTEASE

HIV protease inhibitors was first invented between 1989 and 1994 by researchers working for the pharmaceutical companies of Hoffmann- La Roche Inc. (of Nutley, New Jersey), Abbott Laboratories and Merck & Co., Inc. HIV protease inhibitors are used in the treatment of patients with AIDS and were considered the first breakthrough in over a decade of AIDS research. HIV protease inhibitors can lower the viral load carried by AIDS patients.

We can design a drug (inhibitor) to inhibit the activity of gene. As a gene produces protein/enzyme, so to avoid the formation of any disease causing proteins, we have to stop the activity of that gene. With the help of different bioinformatics tools we can do this.

HIV 1 Protease inhibitors are one class of drugs that target the viral enzyme HIV-1 protease. They are drugs used to treat or prevent infection by HIV. It prevent protease from splitting proteins into peptides. That is, they prevent the cleavage of HIV precursor proteins into active proteins, a process that normally occurs when HIV replicates.

B. THE DIETARY COMPOUNDS USED AS HIV 1 PROTEASE INHIBITORS

It is well known that combination of foods and dietary constituents plays an important role in prevention and treatment of many diseases. Several types of natural products such as carotenoids, phenolics, dietary fibers and fish oils are successful drugs. Epidemiologic evidence suggests that regular consumption of fruits and vegetables may reduce susceptibility to many diseases.

Various bioactivities of dietary compounds are responsible for their chemopreventive properties (e.g., antioxidant, anticarcinogenic, or antimutagenic and anti-inflammatory effects).

BETA CAROTENE

β-Carotene (from Carrot) is a strongly-coloured red-orange pigment abundant in plants and fruits. It is an organic compound and chemically is classified as a hydrocarbon and specifically as a terpenoid (isoprenoid), reflecting its derivation from isoprene units. β-Carotene is biosynthesized from geranylpyrophosphate. It is a member of the carotenes, which are tetraterpenes, synthesized biochemically from eight isoprene units and thus having 40 carbons. Among this general class of carotenes, β-Carotene is distinguished by having beta-rings at both ends of the molecule.

CAFFEINE

Caffeine (from Coffee) is a bitter, white crystalline xanthine alkaloid and psychoactive stimulant. Caffeine was first isolated from coffee in 1820 by the German chemist Friedlieb Ferdinand Runge and again in 1821 by French chemists Robiquet, Pelletier, and Caventou. Pelletier first coined the word "cafeine", which became the English word "caffeine".

Caffeine is found in varying quantities in the seeds, leaves, and fruit of some plants, where it acts as a natural pesticide that paralyzes and kills certain insects feeding on the plants. In humans, caffeine acts as a central nervous system stimulant, temporarily warding off drowsiness and restoring alertness. Caffeine is the world's most widely consumed psychoactive substance, but, unlike many other psychoactive substances, is legal and unregulated in nearly all jurisdictions.

D-LIMONENE

D-limonene (from Lemon) is a colourless liquid hydrocarbon classified as a cyclic terpene. The more common D isomer posesses a strong smell of oranges. It is used in chemical synthesis as a precursor to carvone and as a renewably-based solvent in cleaning products. Limonene is a

chiral molecule, and biological sources produce one enantiomer: the principal industrial source, citrus fruit, contains D-limonene ((+)-limonene), which is the (R)-enantiomer. Racemic limonene is known as dipentene. D-Limonene is obtained commercially from citrus fruits through two primary methods: centrifugal separation or steam distillation.

ALLIXIN

Allixin (from Garlic), first isolated and characterized in 1989. When garlic is stored for long periods of time, it can form visible accumulations of crystalline allixin on its surface, particularly in areas where tissue has become necrotic. After 2 years of storage, the amount of allixin accumulated can approach 1% of the dry weight of the cloves.

ELLAGIC ACID

Ellagic acid (from Black berries) is a natural phenol antioxidant. The antiproliferative and antioxidant properties of ellagic acid have spurred preliminary research into the potential health benefits of ellagic acid consumption.

EMODIN

Emodin (from Rhubarb) is a purgative resin. It shows promise as an agent that could reduce the impact of type 2 diabetes. It effectively limits the effect of the glucocorticoids, and ameliorates diabetes and insulin resistance. Pharmacological studies have demonstrated that emodin from rhubarb exhibits anti-cancer effects on several human cancers, including human pancreatic cancer.

PHENETHYL ISOTHIOCYANATE

Phenethyl isothiocyanate, PEITC (from Cabbage) inhibit carcinogenesis and tumorigenesis. It is also used for amino acid sequencing in the Edman degradation.

OLEUROPEIN

Oleuropein is a chemical compound found in olive leaf from the olive tree. Oleuropein has powerful antioxidant activity both in vivo and in vitro and give extra-virgin olive oil its bitter, pungent taste. Oleuropein preparations have been claimed for several pharmacological effects among them strengthening of the immune system.

ANTHOCYANIN

Anthocyanins (from Egg plant) are water soluble vacuolar pigments that may appear red, purple, or blue according to pH. They belong to a parent class of molecules called flavonoids synthesized via the phenylpropanoid pathway; they are odourless and nearly flavourless, contributing to taste as a moderately astringent sensation. Anthocyanins are derivatives of anthocyanidins which include pendant sugars.

QUERCETIN

Quercetin (from Onion) is a flavonoid widely distributed in nature. The name has been used since 1857, and is derived from quercetum, It is a naturally occurring polar auxin transport inhibitor.

LECTIN

Lectins (from Banana) are sugar binding proteins that are highly specific for their sugar moieties. They play a role in biological recognition phenomena involving cells and proteins. For example, some viruses use lectins to attach themselves to the cells of the host organism during infection. Lectins may be disabled by specific mono- and oligosaccharides, which bind to them and prevent their attachment to cell membranes.

APIGENIN

Apigenin (from Apple) is a flavone that is the aglycone of several glycosides. It is a yellow crystalline solid that has been used to dye wool. Apigenin is commonly recognised as to mediate at least part of the chemopreventive action of vegetables and fruits in the cancerous process. Recently it was shown that Apigenin induces a process called autophagia (a kind of cellular dormancy) which may well explain it chemopreventive properties. Apigenin acts as a monoamine transporter activator, and is one of the few chemicals demonstrated to possess this property.

COUMARIN

Coumarin (from Tonka bean) is a pleasantly fragrant chemical compound. The name comes from a French word, coumarou. It has a sweet odour, readily recognised as the scent of newly-mown hay, and has been used in perfumes since 1882. It is used in the pharmaceutical industry as a precursor molecule in the synthesis of a number of synthetic anticoagulant pharmaceuticals. Coumarin has clinical medical value by itself, as an edema modifier. It is also used as a gain medium in some dye lasers, and as a sensitizer in older photovoltaic technologies.

SILYMARIN

It is a flavanoid extracted from milk thistle. Both in vitro and animal research suggest that silymarin has hepatoprotective properties that protect liver cells against toxins. Chemically modified form is used in treatment of severe intoxications with hepatotoxic substances, such as death cap.

INDOMETHACIN

Indomethacin (from Pumpkin) uncouples oxidative phosphorylation in cartilaginous (and hepatic) mitochondria, like salicylates.

GLIBENCLAMIDE

Glibenclamide (from Bitter melon) also known as glyburide, is drug in a class of medications known as sulfonylureas, closely related to sulfa drugs.

FISETIN

Fisetin (from Strawberry) is a flavonol, a structurally distinct chemical substance that belongs to the flavonoid group of polyphenols. Its chemical formula was first described by Austrian chemist Josef Herzig in 1891. Fisetin is a potent sirtuin activating compound, an agent that modulates sirtuins. It is therefore a caloric restriction mimetic candidate, a drug that has been shown to be able to alleviate ageing effects in certain model organisms such as the yeast s cerevisiae. It is a potent antioxidant. Its antioxidative activity may be due to its structural properties as well as to its ability to modulate certain cellular signaling pathways.

RESVERATROL

Resveratrol (from Grapes) is a stilbenoid, a type of natural phenol, and a phytoalexin produced naturally by several plants when under attack by pathogens such as bacteria or fungi. It is also produced by chemical synthesis http://en.wikipedia.org/wiki/Resveratrol - cite_note-pmid16401555-6 or by biotechnological synthesis and is sold as a nutritional supplement.

INULIN

Inulins (from Asparagus) are a group of naturally occurring polysaccharides. Inulin is increasingly used in processed foods because it has unusually adaptable characteristics. Its flavour ranges from bland to subtly sweet. It can be used to replace sugar, fat, and flour. This is particularly advantageous because inulin contains a quarter to a third of the food energy of sugar or other carbohydrates and a ninth to a sixth of the food energy of fat. While inulin is a versatile ingredient, it also has health benefits. Inulin increases calcium absorption and possibly magnesium absorption, while promoting the growth of intestinal bacteria. In terms of nutrition, it is considered a form of soluble fiber and is sometimes categorized as a prebiotic. Due to the body's limited ability to process polysaccharides, inulin has minimal increasing impact on blood sugar, and unlike fructose is not insulemic and does not raise triglycerides, making it considered suitable for diabetics and potentially helpful in managing blood sugar related illnesses.

CAPSAICIN

It is the active component of chili peppers, which are plants belonging to the genus Capsicum. It is an irritant for mammals, including humans, and produces a sensation of burning in any tissue with which it comes into contact. Capsaicin and several related compounds are called capsaicinoids and are produced as a secondary metabolite by chili peppers, probably as deterrents against certain herbivores and fungi. Pure capsaicin is a hydrophobic, colourless, odourless, crystalline to waxy compound.

V. METHODOLOGY

Using ArgusLab software we can prepare and run molecular docking calculations. Before docking a molecule, define the atoms that make up the Ligand (drug, inhibitor, etc.) and the Binding Site of the protein where the drug binds. The 3D structure of HIV 1 Protease (PDB ID - 3OXC) can be downloaded from the Brookhaven Protein Databank (PDB).

Steps taken for molecular docking using Arguslab is shown below.

- ✓ Open the PDB file, contains the active site residues attached to the inhibitor
- ✓ Open a ligand from the molecule database
- ✓ Select the active site residues and right click the mouse
- ✓ Click “Make a group from the residues”. Select ‘Binding sit’ and click OK on the dialogue box appeared
- ✓ Now select the original ligand and right click the mouse
- ✓ Click “make a ligand group from the residue”
- ✓ Select the ligand from the small molecule database
- ✓ Right click on it and click “Make a ligand group from the residue”
- ✓ Select calculation->Dock a ligand
- ✓ Select the ligand and active site in the dialogue box appeared. Click the button Calculate size and then click OK. Then docking will starts and number of poses with different energies will be displayed.

Explanation of steps-

DOWNLOAD AND OPEN THE PDB STRUCTURE OF 3OXC

Open the web browser and type in the Bookhaven Protein Databank (www.rcsb.org) web address and in the search field type 3OXC. Download the displayed structure as PDB file preferably to Desktop. Open the 3OXC.pdb located in the folder designated during download in ArgusLab program (Arguslab icon is located on the desktop). This is a PDB file so make sure to select the correct file type in the file open dialog (pdb type). Expand the Molecule Tree View tool of 3OXC (located on the left side on the screen).

We can see something like this:

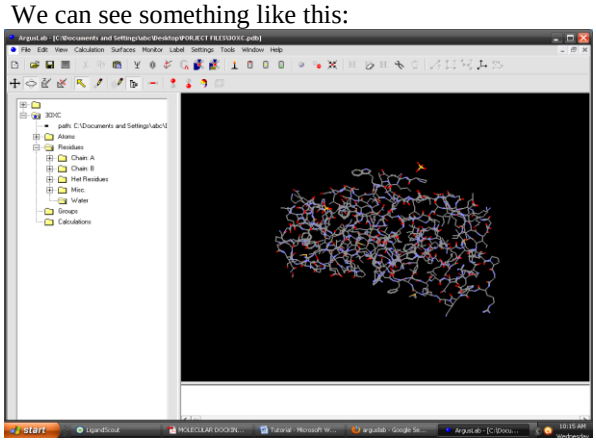

Figure 7

ISSN: 2394-4404

CREATE "LIGAND AND BINDING SITE GROUPS"

Left-click on 30XC in the Tree View. It contains ROC, which is already defined as a ligand and the binding site has been made.

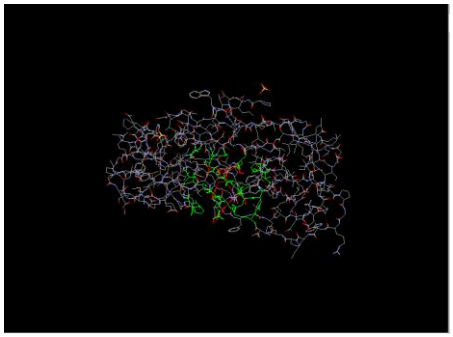

Figure 8: Binding site and ligand shown in green and red respectively

Select the Edit/Hide Unselected menu option to hide all atoms that are not selected.

We can see something like this in the screen:

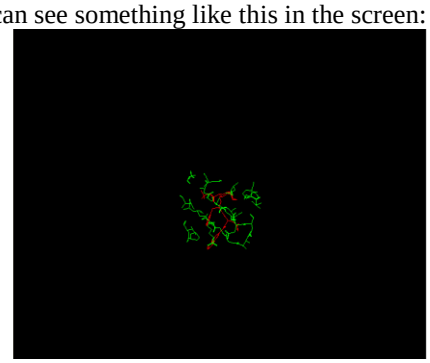

Figure 9

DOCK THE LIGAND INTO DEFINED BINDING SITE

Dietary compounds used as drugs are selected as HIV 1 protease inhibitor. The 3D structure of this compounds are downloaded from PubChem in SDF (Structure Data File) format, which is then converted to pdb using Vega ZZ software (www.vegazz.net). And, finally they are docked into the binding site of 3OXC. For this, bring up the dock settings dialog box by selecting the Calculation/Dock a Ligand menu option or clicking on the  button on the toolbar.

The Dock Settings dialog box looks like:

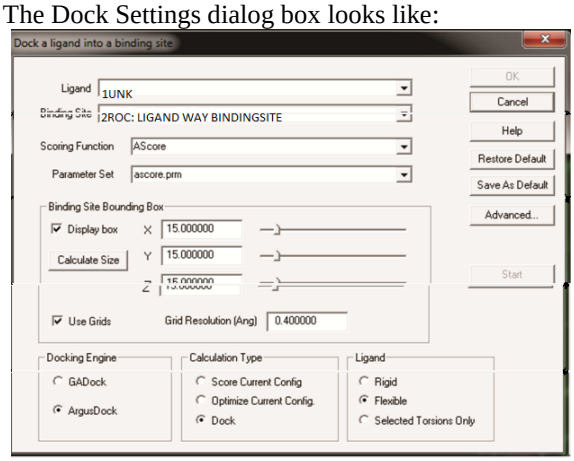

Figure 10

Select the ligand to dock in the "ligand" drop-box. Make sure to select the group named "ligand" group and not the "ligand-xray" group. Click on the "Calculate Size" button and a docking box tailored to the binding site will be made and shown on the screen. Make sure "ArgusDock" is the selected docking engine, the Calculation type = Dock, and the Ligand is Flexible. Click the "Start" button and the docking calculation will begin. Notice the messages that appear in the status line at the bottom of the molecule window. ArgusLab first generate the scoring grids used during the docking, then the various search phases will occur, and finally the candidate poses should be processed until the calculation is done.

When 3OXC is docked with silymarin, we can see something like this:

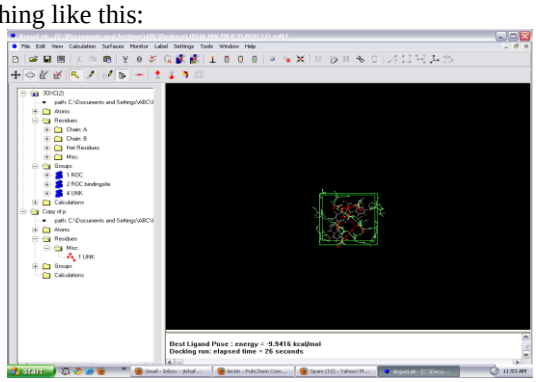

Figure 11: Silymarin shown in violet colour

DOCK ANOTHER COPY OF ROC INTO THE BINDING SITE

Open the file 3OXC.pdb. A ligand group, ROC is already created for this structure and verify this by looking in the groups folder under the 3OXC molecule in the Tree View. Make a copy of this ligand, and is docked into the binding site of 3OXC. This is made as the REFERENCE..

ANALYZE AND SAVE THE RESULTS

There are two sources of information about the performed docking calculation: The separate log file and information located in the Tree View tool. Expand the ArgusDock calculation in the Calculations folder in the Tree View. There will be listing of docking settings and several "poses" starting with the most stable as Pose 1 according to the calculated energy in kcal/mol. (A term "pose" is usually used to designate the specific set of coordinates of a docked structure). Right click on pose 2 and select the "Display" option. The coordinates of ligand will change to this pose and its score will be displayed on the screen.

VI. RESULT AND DISCUSSION

Human Immunodeficiency Virus (HIV) is the retrovirus that causes AIDS. The HIV virus undergoes continuous mutation. As a result new genotypes are evolved. Hence it is relevant that identifying the genotype is important for drug development and disease management. The better way of developing a drug against HIV is to inhibit the activity of HIV 1 Protease, which cleaves newly synthesized polyproteins at

appropriate places to create the mature protein components of an infectious HIV virion. Inhibiting the activity of this protein will essentially supress the virus, which cannot transform to its mature infectious form. So there is a growing interest in the development of HIV 1 Protease inhibitors.

For carrying out molecular docking studies, the 3D structure of the protein is inevitable. The active site of the target molecule, HIV 1 protease (PDB ID - 3OXC) is found out, which is complexed with a known inhibitor ROC. Approximately, twenty dietary compounds that are used as drugs are selected. Various chemical libraries containing the 3D structure of this twenty molecules are screened against it employing molecular docking.

The target ligand docking energy is noted, which is compared with that of the reference one (A copy of ROC is made and docked into the acitve of 3OXC; is considered as reference).

The target ligand docking energy of the reference was found to be -10.194 Kcalmol-1; Running time 372 sec; 10 Number of poses.

The twenty molecules tested are furnished in table A. The docked poses along with their corresponding score are also shown.

NO.	MOLECULES	TARGET LIGAND ENERGY (KCalmol-1)	RUNNING TIME (sec)	POSES
1	Allixin	-7.988	87	44
2	Anthocyanin	-9.578	67	93
3	Apigenin	-9.444	33	10
4	Beta carotene	-8.989	40	28
5	Caffeine	-4.849	26	56
6	Capsaicin	-9.667	135	80
7	Coumarin	-8.105	83	73
8	D- limonene	-8.519	200	92
9	Ellagic acid	-8.346	101	6
10	Emodin	-8.966	24	10
11	Fisetin	-9.383	115	4
12	Glibenclamide	-11.155	80	63
13	Indomethacin	-9.487	19	37
14	Inulin	-9.187	27	1
15	Lectin	10.527	9	16
16	Oleuropein	-9.996	1348	2
17	Phenethyl isothiocyanate	-8.209	283	97
18	Quercetin	-9.388	397	4
19	Resveratrol	-10.082	200	10
20	Silymarin	-9.941	26	1

Table A

The computational experiment undertaken has resulted in the identification of ten molecules, which shows good binding potential. They are listed in table B.

RANK	MOLECULES	TARGET LIGAND ENERGY (KCalmol-1)	RUNNING TIME (sec)	POSES
1	Glibenclamide	-11.155	80	63
2	Lectin	-10.527	9	16
3	Resveratrol	-10.082	200	10
4	Oleuropein	-9.996	1348	2
5	Silymarin	-9.941	26	1
6	Capsaicin	-9.667	135	80
7	Anthocyanin	-9.578	67	93
8	Indomethacin	-9.487	19	37
9	Apigenin	-9.444	33	10
10	Quercetin	-9.388	397	4

Table B

From 10 molecules, first five ranking molecules are selected as "hit molecules". They are presented in table C.

RANK	MOLECULES	TARGET LIGAND ENERGY (KCalmol-1)	RUNNING TIME (sec)	POSES
1	Glibenclamide	-11.155	80	63
2	Lectin	-10.527	9	16
3	Resveratrol	-10.082	200	10
4	Oleuropein	-9.996	1348	2
5	Silymarin	-9.941	26	1

Table C

(List of top 5 ranked molecules)

Table C shows that, each molecule is docked well to the active site of HIV 1 protease with a minimum energy (found to be stable). The molecule with high score is ranked as 1, and so on. This top ranked molecules are selected as the inhibitors of the HIV 1 protease. Also they have no adverse effect as carcinogenicity and mutagenicity [40]. So these effective lead compounds can be suggested to be the interesting candidates for further testing in the laboratory.

FIGURES

3OXC IN COMPLEX WITH A KNOWN LIGAND ROC IS SHOWN BELOW

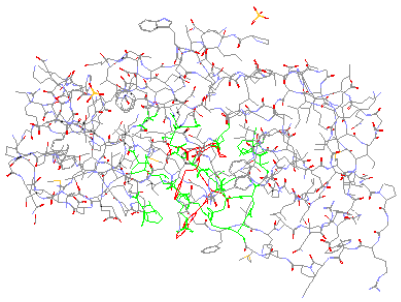

Figure 12: Binding site and ligand shown in green and red respectively

A 2D STRUCTURE OF ROC

[(2S)-N-[(2S,3R)-4-[(2S,3S,4aS,8aS)-3-(tert-butylcarbamoyl)-3,4,4a,5,6,7,8,8a-octahydro- 1H-isoquinolin-2-yl]-3-hydroxy-1-phenyl-butan-2-yl]-2-(quinolin-2-ylcarbonylamino)butanediamide]

Molecular Formula: $C_{38}H_{50}N_6O_5$

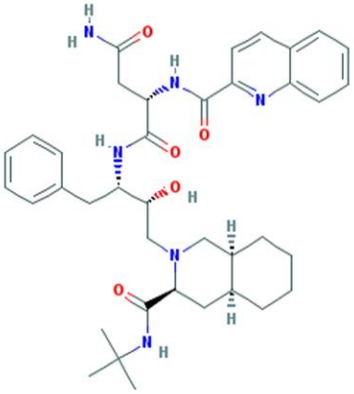

Figure 13

ISSN: 2394-4404

REFERENCE

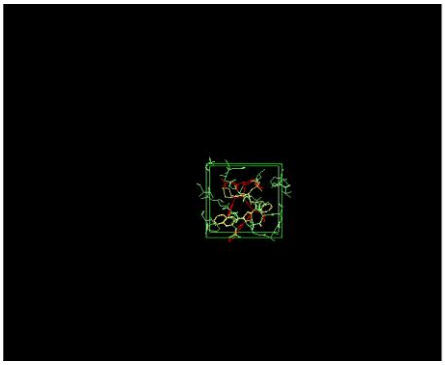

Figure 14: Copy of known ligand, ROC shown in yellow

STRUCTURE OF TOP 5 RANKED MOLECULES

✓ *Glibenclamide*

[5-chloro-N-[2-[4-(cyclohexylcarbamoylsulfamoyl)phenyl]ethyl]-2-methoxybenzamide]

Molecular formula: $C_{23}H_{28}ClN_3O_5S$
Molecular weight: 494.003 g/mol
Compound ID: 3488
H bond donor: 3
H bond acceptor: 5

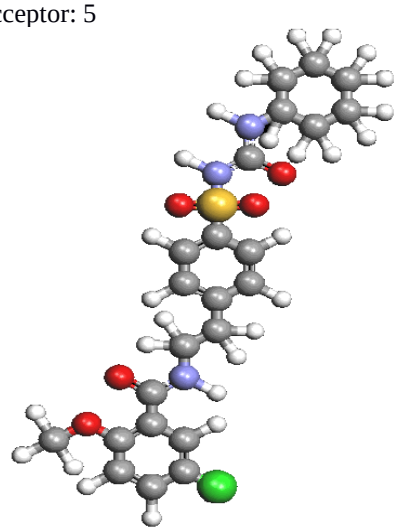

Figure 15

✓ *Lectin*

[9-benzyl-3-methylidene-1,5-bis-(4-methylphenyl)sulfonyl-1,5,9-triazacyclododecane]

Molecular formula: $C_{31}H_{39}N_3O_4S_2$
Molecular weight: 581.789 g/mol
Compound ID: 466371
H bond donor: 0
H bond acceptor: 7

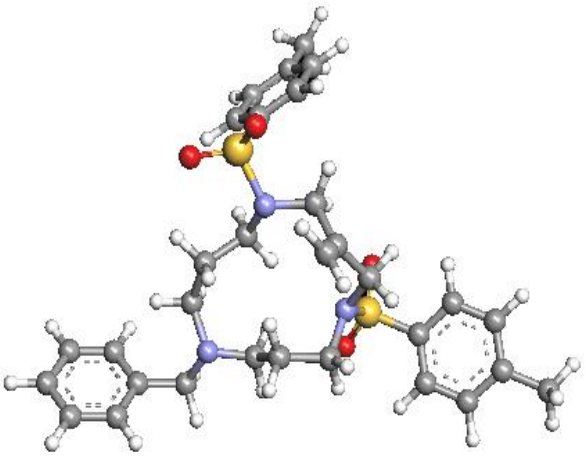

Figure 16

✓ *Resveratrol*

[5-[(E)-2-(4-hydroxyphenyl)ethenyl]benzene-1,3-diol]
Molecular formula: $C_{14}H_{12}O_3$
Molecular weight: 228.243 g/mol
Compound ID: 445154
H bond donor: 3
H bond acceptor: 3

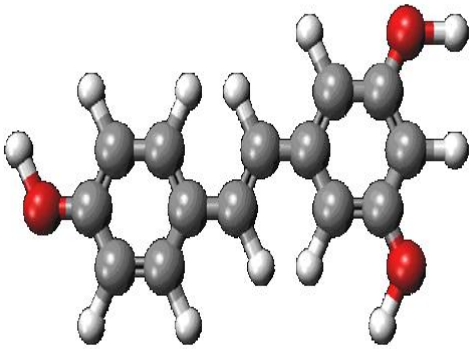

Figure 17

✓ *Oleuropein*

[methyl(4S,5E,6S)-4-[2-[2-(3,4-dihydroxyphenyl)ethoxy]-2-oxoethyl]-5-ethylidene-6-[(2S,3R,4S,5S,6R)-3,4,5- trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy-4H-pyran-3-carboxylate]

Molecular formula: $C_{25}H_{32}O_{13}$
Molecular weight: 540.513 g/mol
Compound ID: 5281544
H bond donor: 6
H bond acceptor: 13

ISSN: 2394-4404

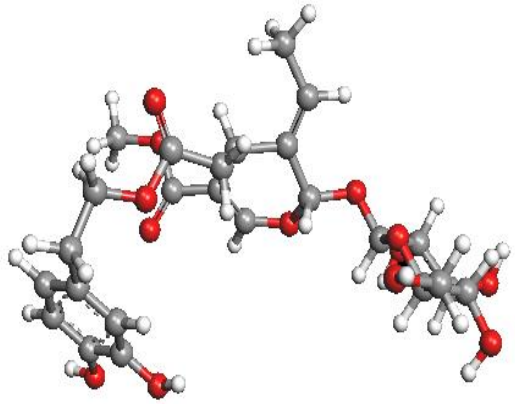

Figure 18

✓ *Silymarin*

[(2R,3R)-3,5,7-trihydroxy-2-[(2R,3R)-3-(4-hydroxy-3-methoxyphenyl)-2-(hydroxymethyl)-2,3-dihydro-1,4-benzodioxin- 6-yl]-2,3-dihydrochromen-4-one]

Molecular formula: $C_{25}H_{22}O_{10}$
Molecular weight: 482.436 g/mol
Compound ID: 31553
H bond donor: 5
H bond acceptor: 10

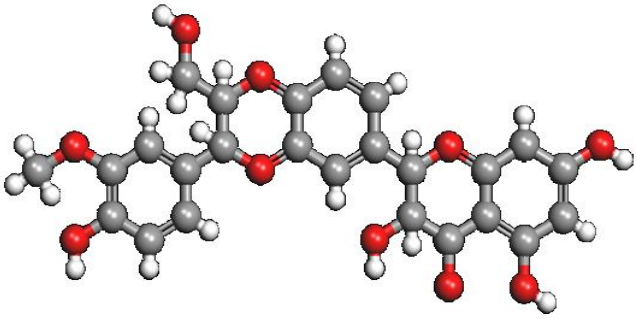

Figure 19

DOCKING OF TOP 5 RANKED MOLECULES INTO THE ACTIVE SITE OF 3OXC

✓ *Docking of Glibenclamide*

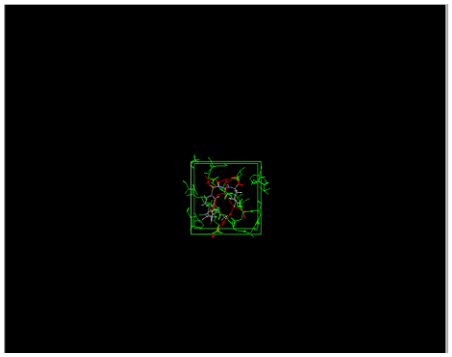

Figure 20: Glibenclamide shown in violet colour

✓ *Docking of Lectin*

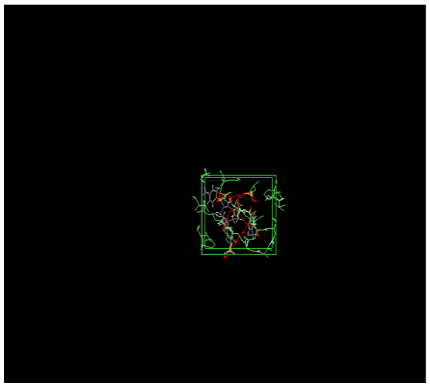

Figure 21: Lectin shown in violet colour

✓ *Docking of Resveratrol*

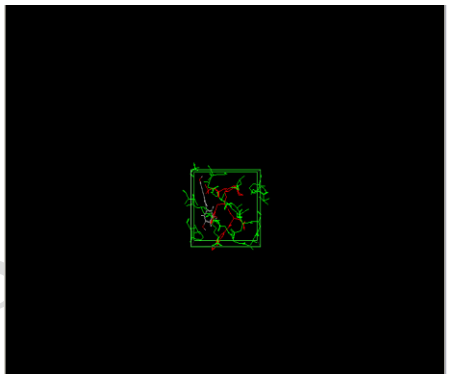

Figure 22: Resveratrol shown in violet colour

✓ *Docking of Oleuropein*

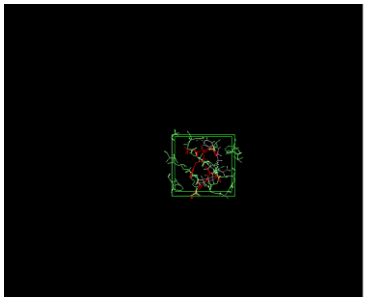

Figure 23: Oleuropein shown in violet colour

✓ *Docking of Silymarin*

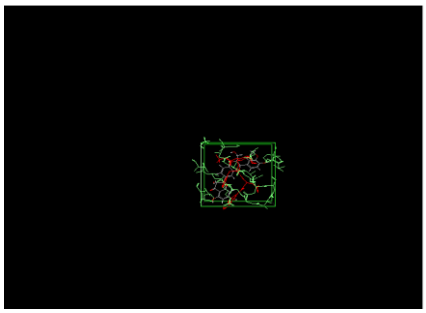

Figure 24: Silymarin shown in violet colour

VII. CONCLUSION

The study indicates the importance of diet in the prevention and treatment of diseases. Five molecules that "Docked" well into the active site of the target, HIV 1 Protease are selected. And they are considered to be as the "Hit Molecules". Glibenclamide, Lectin, Resveratrol, Oleuropein, Silymarin are considered as the hit molecules. As they are docked into the active site of the target with a minimum energy (found to be stable). So these molecules can be suggested to be the interesting candidates for further testing in the laboratory. Finally, this study strongly underscores the importance of computational approaches in drug discovery, supplementing classical methods, thus saving enormous amount of time and money.

REFERENCES

[1] Rajan Sundara S, Balagi K. "Introduction to bioinformatics" (Mumbai: Himalaya publishing house, 2005), 308-317.
[2] Krane E Dan, Raymer L Michael. "Fundamental concepts of Bioinformatics" (Singapore: Person education Pvt.Ltd, 2004),187-197.
[3] V.Kothekar. "Introduction to Bioinformatics" (Delhi: Dhruv publications, 2004),256-263.
[4] Ignacimuthu S. "Basic bioinformatics" (NewDelhi: Narosa publishing house, 2005),1-12,142-147.
[5] http://www.geocities.com/bioinformaticsweb/applications .html
[6] Rastogy S.C, Mendiratta, Rastogy P. "Bioinformatics: Methods and Applications" (Delhi: Prentice-Hall of India Pvt.Ltd,2004),1-25,334-364.
[7] Cohen Claude N. "Guidebook on molecular modeling in drug design" (NewDelhi: Elsevier India Pvt.Ltd,2006),68-71,139-158,273-279.
[8] V.Kothekar. "Essentials of Drug Designing" (Delhi: Dhruv publications first edition,2005).
[9] http://ihumans.com/news-comments_archive/DrugDesignApproaches.html
[10] http://en.wilkipedia.org/wiki/drud_design
[11] Jiang Tao, Zhang Michael. "Current topics in computational biology" (Delhi: Gopaljee printers, 2004), 502-520.
[12] Bohm H.J (1992). "Computer-Aided Mol Design",593-606.
[13] http://biotech.icmb.utexas.edu/pages/bioinfo.html
[14] Madsen (2002). "Textbook of Drug Design and Discovery", 415-428.
[15] Jorgensen WL (1991). "Rusting of the lock and key model for protein-ligand binding",254-267,954–955.
[16] Moon and Howe (1991). "Computer design of Bioactive molecules: A method for receptor based de-novo ligand design",314-328.
[17] Bohm, H.J (1993). "Ligand design in 3D QSAR in drug design",386-405
[18] Edwards BS, (2006). "Integration of virtual and physical screening",377–385.
[19] Lengauer T, Rarey M (1996). "Computational methods for molecular docking",402–406.
[20] N. Claude (1996). "Guidebook on Molecular Modeling in Drug Design".
[21] Kitchen DB, Bajorath J (2004). "Docking and scoring in virtual screening for drug discover",935–949.
[22] Shoichet BK (2004). "Molecular docking using shape descriptors",380–397.
[23] Shelley M, Francis P, Shenkin PS (March 2004). "A new approach for rapid, accurate docking and scoring method",1739–1749.
[24] Feig M, Onufriev A (2004). "Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures",265–284.
[25] Suresh PS, Singh VP (January 2008). "An in silico [correction of insilico] approach to bioremediation: laccase as a case study",845–849.
[26] http://en.wilkipedia.org/wiki/AIDS
[27] Hooper E (1999). "The River:A Journey to the Source of HIV and AIDS".
[28] http://www.mcld.co.uk/hiv/?q=HIV%20genome
[29] Davies DR (1990). "The structure and function of the aspartic proteinases",189–215.
[30] Brik A, Wong CH (January 2003). "HIV-1 protease: mechanism and drug discovery",5–14.
[31] Spinelli.S (1991). "The three-dimensional structure of the aspartyl protease from the HIV-1",1391-1396.
[32] Harrison RW, (2007). "Atomic resolution crystal structures of HIV-1 protease",232-242.
[33] Guss DA (1994). "The acquired immuno deficiency syndrome: an overview for the emergency physician,Part 2",491–497.
[34] http://www.ncsl.org/programs/health/transhiv.html
[35] http://www.brown.edu/Courses/bio_160/Projects1999/hiv/pathog.html
[36] Taylor D.J, Stout G.W. "Biological science" (USA: Cambridge university press, 1997),22.
[37] Nupur Mehrotra, Rajeev Soni. "Pharmacogenomics:A Step Towards Personalized Medicine",2005, vol III.
[38] Roitt Ivan. "Microbiology" (Gower medical publishing Ltd,2001),240-242.
[39] Ananthanarayan R, Panicar Jayaram. "Microbiology" (Chennai Orient Longman Pvt Ltd,2005),582-598.
[40] N.G.Rick. "Drugs-From discovery to approval" (New Jersey: john Wiley and sons,2002),10-30,44-66.