



IDENTIFYING POTENTIAL BACE-1 ENZYME INHIBITORS /LEAD SMALLMOLECULES: RELEVANCE TO ALZHEIMER'S DISEASES

Deepak.G ^(a)

Dr. A. Dinakara Rao ^(b)

Mohammad Babu Khan ^(c)

^(a)Centre for Bioinformatics Pondicherry University Puducherry-605014

^(b)Associate Professor, Centre for Bioinformatics Pondicherry University Puducherry-605014

^(c) Associate Professor, Centre for Bioinformatics Pondicherry University Puducherry-605014

Abstract

It is unclear what causes Alzheimer's disease. However, genetic data support the amyloid hypothesis, which states that abnormal beta-amyloid protein aggregation initiates the illness process. There is a lengthy pre-clinical stage of Alzheimer's disease. The terrible consequences of amyloidogenic diseases like Alzheimer's. More than 7000 tiny molecules have been screened from ZINC subsets and exposed to 165 compounds and 16. From a research perspective, I have examined numerous databases like the ZINC Subset, Pubchem compound database, and Drug bank database. For docking investigations, these processed molecules were employed. The docking study enabled us to locate a small number of compounds with good inhibitory interactions with the protein, i.e., chemicals that impede the protein's function. These final compounds underwent drug-likeness testing and methods for ADME descriptor validation. These substances or molecules may help develop potent medications to treat Alzheimer's. To determine which protein had the lowest energy conformation and was, therefore, the most stable, a molecular dynamic of the protein was run.

Keywords: Bace-1 Enzyme, ZINC Subset, Beta-Amyloid peptide, Docking, Molecular Dynamic simulation.

1.0 INTRODUCTION

1.1 Role of Beta-Secretase in Alzheimer's Disease:

Amyloid precursor protein (APP), the starting point for amyloid plaques, is one of many proteins associated with the cell membrane, the barrier that encloses the cell. As it is being made inside the cell, APP becomes embedded in the membrane, as a toothpick stuck through the skin of an orange. In several cell compartments, including the outermost cell membrane, specific enzymes snip or cleave, APP into discrete fragments. In 1999 and 2000, scientists identified the enzymes responsible for cleaving APP. These enzymes are called alpha-secretase, beta-secretase, and gamma-secretase. In a breakthrough, scientists then discovered that, depending on which enzyme is involved and the segment of an APP where the cleaving occurs, APP processing could follow one of two pathways that have very different consequences for the cell. In the benign pathway, alpha-secretase cleaves the APP molecule within the portion that has the potential to become a beta-amyloid. This eliminates the production of the beta-amyloid peptide and the potential for plaque buildup. The cleavage releases from the neuron a fragment called sAPP α , which has beneficial properties, such as promoting neuronal growth and survival. The remaining APP fragment, still tethered in the neuron's membrane, is cleaved by gamma-secretase at the end of the beta-amyloid segment. The smaller of the resulting fragments also is released into the space outside the neuron, while the larger fragment remains within the neuron and interacts with factors in the nucleus. In the harmful pathway, beta-secretase first cleaves the APP molecule at one end of the beta-amyloid peptide, releasing sAPP β from the cell. Gamma-secretase cuts the resulting APP fragment, still tethered in the neuron's membrane, at the other end of the beta-amyloid peptide. Following the cleavages at each end, the beta-amyloid peptide is released into the space outside the neuron and begins to stick to other beta-amyloid peptides. These small, soluble aggregates of two, three, four, or even up to a dozen beta-amyloid peptides are called oligomers. Small oligomers may be responsible for reacting with receptors on neighboring cells and synapses, affecting their ability to function. Some oligomers are likely cleared

from the brain. Those that cannot be cleared clump together with more beta-amyloid peptides. As the process continues, oligomers grow larger, becoming entities called protofibrils and fibrils. Eventually, other proteins and cellular material are added, and these increasingly insoluble entities combine to become the well-known plaques characteristic of AD. For many years, scientists thought that plaques might cause all of the damage to neurons that are seen in AD. However, that concept has evolved dramatically in the past few years. Many scientists now think that oligomers may be a significant culprit. Many scientists also think that plaques may be a late-stage attempt by the brain to get this harmful beta-amyloid away from neurons.

1.2 Alzheimer's Diseases (AD):

In 1907, Alois Alzheimer published an account of a 51-year-old female patient, Auguste D., who suffered from intense feelings of jealousy towards her husband, increased memory impairment, disorientation, hallucinations, and often loud and aggressive behavior. After four and a half years of rapidly deteriorating mental illness, Auguste D died in an utterly demented state. Alzheimer's is a brain disease that causes problems with memory, thinking and behavior. It is not a normal part of aging. Alzheimer's gets worse over time. Although Symptoms can vary widely, the first problem many people notice is forgetfulness severe enough to affect their ability to function at home or at work, or to enjoy lifelong hobbies.



Figure 1- Illustration

Other symptoms include confusion, getting lost in familiar places, misplacing things, and trouble with language. Dementia is a general term for memory loss and other intellectual abilities serious enough to interfere with daily life. Alzheimer's is the most common form of dementia.

Alzheimer's disease (AD), also called **Alzheimer's disease**, **Senile Dementia of the Alzheimer Type (SDAT)** or simply **Alzheimer's** is the most common form of dementia. This incurable, degenerative, and terminal disease as first described by German psychiatrist and neuropathologist Alois Alzheimer in 1906 and was named after him. Generally, it is diagnosed in people over 65 years of age, [36] although the less-prevalent early-onset Alzheimer's can occur much earlier. In 2006, there were 26.6 million sufferers worldwide. Alzheimer's is predicted to affect 1 in 85 people globally by 2050. Alzheimer's disease is a progressive gradual decline in the ability to think and remember, as well as to function physically. It is irreversible, and there is no

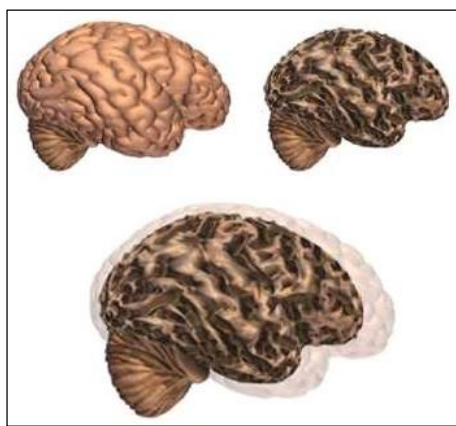


Figure 2- illustration picture

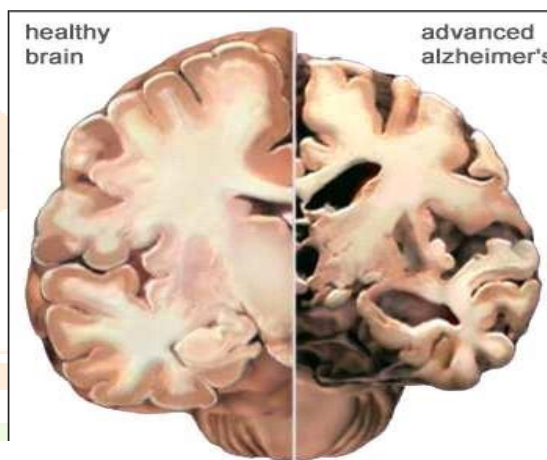
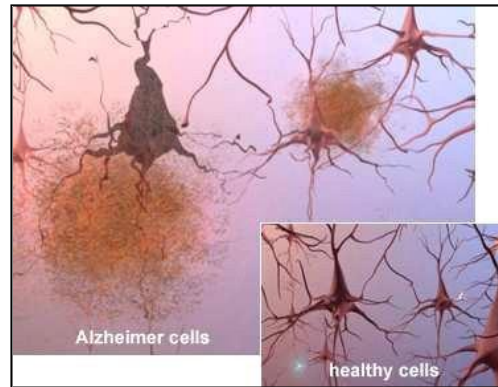


Figure 3- illustration used by author

cure well as to function physically. It is irreversible, and there is no cure. Although the course of Alzheimer's disease is unique for every individual, there are many common symptoms. [38] The earliest observable symptoms are often mistakenly thought to be 'age-related' concerns or manifestations of stress. In the early stages, the most commonly recognized symptom is memory loss, such as difficulty in remembering recently learned facts. When a doctor or physician has been notified, and AD is suspected, the diagnosis is usually confirmed with behavioral assessments and cognitive tests, often followed by a brain scan if available. As the disease advances, symptoms include confusion, irritability and aggression, mood swings, language breakdown, long-term memory loss, and the general withdrawal of the sufferer as their senses decline.

Figure 4 - illustration

Gradually, bodily functions are lost, ultimately leading to death. Individual prognosis is difficult to assess, as the duration of the disease varies. AD develops for an indeterminate period before becoming fully apparent, and it can progress undiagnosed for years. The mean life expectancy following diagnosis is approximately seven years.

Fewer than three percent of individuals live more than fourteen years after diagnosis. The cause and progression of Alzheimer's disease are not well understood. Research indicates that the disease is associated with plaques and tangles in the brain. Currently used treatments offer a small symptomatic benefit; no treatments to delay or halt the progression of the disease are as yet available. As of 2008, more than 500 clinical trials have been conducted for identification of a possible treatment for AD, but it is unknown if any of the tested intervention strategies will show promising results. A number of non-invasive, life-style habits have been suggested for the prevention of Alzheimer's disease, but there is a lack of adequate evidence for a link between these recommendations and reduced degeneration. Mental stimulation, exercise, and a balanced diet are suggested, as both a possible prevention and a sensible way of managing the disease. Because AD cannot be cured and is degenerative, management of patients is essential. The role of the main caregiver is often taken by the spouse or a close relative. Alzheimer's disease is known for placing a great burden on caregivers; the pressures can be wide-ranging, involving social, psychological, physical, and economic elements of the caregiver's life. In developed countries, AD is one of the most costly diseases to society.

2.0 WHAT IS ALZHEIMER DISEASE?

Alzheimer's disease (AD) is the most common form of dementia among older people. Dementia is a brain disorder that seriously affects a person's ability to carry out daily activities. AD begins slowly. It first involves the parts of the brain that control thought, memory and language. People with AD may have trouble remembering things that happened recently or names of people they know. Over

time, symptoms get worse. People may not recognize family members or have trouble speaking, reading or writing. They may forget how to brush their teeth or members who must care for them AD usually begins after age 60. The risk goes up as you get older. Your risk is also higher if a family member has had the disease No treatment can stop the disease. However, some drugs may help keep symptoms from getting comb their hair. Later on, they may become anxious or aggressive, or wander away from home. Eventually, they need total care. This can cause great stress for family worse for a limited time. National Institute on aging.

3.0 HISTORY AND CURRENT STATUS:

Progressive mental deterioration in old age has been recognized and described throughout history. However, it was not until 1906 that a German physician, Dr. Alois Alzheimer, specifically identified a collection of brain cell abnormalities as a disease. One of Dr. Alzheimer's patients died after years of severe memory problems, confusion and difficulty understanding questions. Upon her death, while performing a brain autopsy, the doctor noted dense deposits surrounding the nerve cells (neuritic plaques). Inside the nerve cells he observed twisted bands of fibers (neurofibrillary tangles). Today, this degenerative brain disorder bears his name, . When found during an autopsy, these plaques and tangles mean a definite diagnosis of Alzheimer's disease (AD) specifically identified a collection of brain cell abnormalities as a disease. One of Dr. Alzheimer's patients died after years of severe memory problems, confusion and difficulty understanding questions. Upon her death, while performing a brain autopsy, the doctor noted dense deposits surrounding the nerve cells (neuritic plaques). Inside the nerve cells he observed twisted bands of fibers (neurofibrillary tangles). Today, this degenerative brain disorder bears his name, and when found during an autopsy, these plaques and tangles mean a definite diagnosis of Alzheimer's disease (AD). Since its discovery more than 100 years ago, there have been many scientific breakthroughs in AD research. In the 1960s, scientists discovered a link between cognitive decline and the number of plaques and tangles in the brain. The medical community then formally recognized Alzheimer's as a disease and not a normal part of aging. In the 1970s, scientists made great strides in understanding the human body, and AD emerged as a significant area of research interest. This increased attention led in the 1990s to important discoveries and a better understanding of complex nerve cells in the brains of AD patients. More research was done on AD susceptibility genes, and several drugs were approved to treat the cognitive symptoms of the disease. Over the last decade, scientists have substantially progressed in understanding potential environmental, genetic and other ,risk factors for

AD, the processes leading to the formation of plaques and tangles in the brain, and the brain regions that are affected. Specific genes related to both the early-onset and late-onset forms of AD have been identified, but genetic risk factors alone do not fully explain its causes, so researchers are actively exploring environment and lifestyle to learn what role they might play in the development of this disease. More effective treatment options have been approved by the Food and Drug Administration (FDA). However, AD is still incurable. The drugs currently in use treat only the symptoms, not the cause of the disorder, and they only slow the progression of cognitive decline .

3.1 CAUSES:

Many changes take place in the brain of a person with AD. The three major characteristics that reflect the pathology, or damage, caused by the disease are: Amyloid plaques. Found in the spaces between neurons, plaques consist of largely insoluble deposits of aggregated protein fragments called beta-amyloid peptides, other proteins, remnants of neurons, degenerating dendrites and axons, glia, and other cellular material. Scientists used to think that plaques caused most of the damage to neurons seen in AD. Now, however, many think that more soluble forms of beta-amyloid, seen earlier in the plaque formation process, may be the major culprits. Neurofibrillary tangles. Found inside neurons, neurofibrillary tangles are abnormal aggregates of a protein called *tau*. Healthy neurons are internally supported in part by structures called microtubules, which help guide nutrients and molecules from the cell body to the end of the axon. *Tau*, which normally has a certain number of phosphate molecules attached to it, binds to microtubules and stabilizes them. In AD, an abnormally high number of additional phosphate molecules attach to *tau*. As a result, *tau* disengages from the microtubules and begins to clump together with other threads of *tau*, eventually forming neurofibrillary tangles. When this happens, the microtubules disintegrate and the neuron's transport system collapses. As with beta-amyloid, some scientists think that early soluble forms of abnormal *tau* may cause the damage to neurons. Loss of connections between cells and cell death. This feature of AD likely results from the accumulation of beta-amyloid and abnormal *tau*. When neurons lose their connections, they cannot function properly and eventually die.

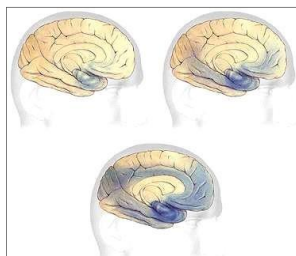


Figure 5

As neuronal death spreads through the brain, affected regions begin to shrink in a process called brain atrophy. By the final stage of AD, damage is widespread and brain tissue has shrunk significantly. Very rarely, people develop AD in their 40s or 50s. In many of these cases, the disease runs in families and is caused by a mutation in one of three genes that a person has inherited from a parent. This form of the disease is called “early-onset” AD. Not all early-onset cases are caused by such mutations. More than 90 percent of AD cases develop in people older than 60, however. The development and pathology of this form of AD, called “late-onset” Alzheimer’s disease, are very similar to those of early-onset AD. We don’t yet completely understand the causes of late-onset AD, but they probably include genetic, environmental, and lifestyle factors. The importance of these factors in increasing or decreasing the risk of developing AD differs from person to person. Scientists hope that what Perhaps the greatest mystery is why AD largely strikes people of advanced age. Why does it take 30 to 50 years or more for people to develop signs of the disease? Research on how the brain changes normally as people age will help provide answers to this important question. They learn about early onset AD also can be applied to the late-onset form of the disease.

4.0 OTHER RISK FACTORS FOR ALZHEIMER’S:

While age and family history comprise the vast majority of Alzheimer causes, researchers have recently made discoveries between the development of Alzheimer’s and other extraneous circumstances. For instance, there appears to be a strong link between future risk of developing Alzheimer’s disease and head injury. Also, the connection between heart and brain has shown significant influence in brain health and Alzheimer’s development. With every heartbeat, your heart sends 20 to 25 percent of the body’s blood to the brain and with that blood the brain uses 20 percent of the food and oxygen it carries. That being the case, when the heart is weakened by heart attack, heart disease, stroke, diabetes or cholesterol, the brain runs a higher risk of being prone to Alzheimer’s.

Other evidence suggests that healthy aging is key to maintaining a healthy brain and fending off

Alzheimer's. One should refrain from smoking and heavy consumption of alcohol, and it's also important to stay socially connected and exercise regularly.

5.0 SYMPTOMS:

Memory loss that disrupts everyday life is not a normal part of aging. The Alzheimer's Association has developed a checklist to help you recognize the difference between regular, age-related memory changes and Alzheimer's disease. There's no clear line that separates regular changes from warning signs. It's always a good idea to check with a doctor if a person's abilities seem to be declining.

Memory loss: Forgetting recently learned information is one of the most common early signs of dementia. A person begins to forget more often and is unable to recall the information later

What's normal? Forgetting names or appointments occasionally.

Difficulty performing familiar tasks

People with dementia often find it hard to plan or complete everyday tasks. Individuals may lose track of the steps involved in preparing a meal, placing a telephone call or playing a game. **What's normal?** Occasionally forgetting why you came into a room or what you planned to say

5.1.1.1.1 Problems with language

People with Alzheimer's disease often forget simple words or substitute unusual words, making their speech or writing hard to understand. They may be unable to find the toothbrush, for example, and instead ask for "that thing for my mouth."

What's normal? Sometimes having trouble finding the right word.

Disorientation to time and place

People with Alzheimer's disease can become lost in their own neighborhoods, forget where they are and how they got there, and not know how to get back home.

What's normal? Forgetting the day of the week or where you were going.

Poor or decreased judgment

Those with Alzheimer's may dress inappropriately, wearing several layers on a warm day or little clothing in the cold. They may show poor judgment about money, like giving away large sums to telemarketers.

What's normal? Making a questionable or debatable decision from time to time

Problems with abstract thinking

Someone with Alzheimer's disease may have unusual difficulty performing complex mental tasks, like forgetting what numbers are and how they should be used

What's normal? Finding it challenging to balance a checkbook

Misplacing things

A person with Alzheimer's disease may put things in unusual places: an iron in the freezer or a wristwatch in the sugar bowl.

What's normal? Misplacing keys or a wallet temporarily.

Changes in mood or behaviour

Someone with Alzheimer's disease may show rapid mood swings – from calm to tears to anger – for no apparent reason.

What's normal? Occasionally feeling sad or moody

Changes in personality

The personalities of people with dementia can change dramatically. They may become extremely confused, suspicious, fearful or dependent on a family member.

What's normal? People's personalities do change somewhat with age.

Loss of initiative:

A person with Alzheimer's disease may become very passive, sitting in front of the TV for hours, sleeping more than usual or not wanting to do usual activities.

What's normal? Sometimes feeling weary of work or social obligations.

6.0 OBJECTIVE:-

To target Beta-Secretase (The main enzyme that cleaving the, Amyloid Precursor Protein (APP), in order to find new lead molecules/ Test compounds (therapeutic molecules).

7.0 MATERIALS AND METHODS:-

The Designing of the novel inhibitor is carried out by the method of the Structure based drug designing. In these mainly 4 steps are involved. They are

- *Searching of the Template Molecule in the PDB.*
- *Searching of the suitable ligands from the database.*
- *Virtual Screening and Molecular docking*
- *Analysis of the Docked complex and finding out the best Ligand*

The ligand which interacts more with the receptor molecule is known as the best Inhibitor. The molecular Dynamic Simulation has been performed by using the GROMACS software.

8.0 ABOUT DATABASES:

ZINC Database:

A critical barrier to entry into structure-based virtual screening is the lack of a suitable, easy to access database of purchasable compounds. We have therefore prepared a library of 727 842 molecules, each with 3D structure, using catalogs of compounds from vendors (the size of this library continues to grow). The molecules have been assigned biologically relevant protonation states and are annotated with properties such as molecular weight, calculated LogP, and number of rotatable bonds. Each molecule in the library contains vendor and purchasing information and is ready for docking using a number of popular docking programs. Within certain limits, the molecules are prepared in multiple protonation states and multiple tautomeric forms. In one format, multiple conformations are available for the molecules. This database is available for free download (<http://zinc.docking.org>) in several common file formats including SMILES, mol2, 3D SDF, and DOCK flexibase format. A Web-based query tool incorporating a molecular drawing interface enables the database to be searched and browsed and subsets to be created. Users can process their own molecules by uploading them to a server. Our hope is that this database will bring virtual screening libraries to a wide community of structural biologists and medicinal chemists.

PUBCHEM:

PubChem provides information on the biological activities of small molecules. It is a component of NIH's Molecular Libraries Roadmap Initiative. PubChem includes substance information, compound structures, and Bioactivity data in three primary databases, Psubstance, Pcompound, and PC Bioassay, respectively

- Psubstance contains more than 62 million records. We can check the count of substance records as of today.

- Pccompound contains more than 26 million unique structures. We can check the count of compound records as of today.

- PCBioAssay contains more than 2000 BioAssays. Each BioAssay contains a various number of data points. You can check the count of BioAssay records as of today.

Each hit provides information about synonyms, chemical properties, chemical structure including SMILES and InChI strings, bioactivity, and links to structurally related compounds and other NCBI databases like PubMed. In the text search form the database fields can

be searched by adding the field name in square brackets to the search term. A numeric range is represented by two numbers separated by a colon. The search terms and field names are case insensitive. Parentheses and the logical operators AND, OR, and NOT can be used. AND is assumed if no operator is used

DRUGBANK:

The DrugBank database is a unique bioinformatics and cheminformatics resource that combines detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information. The database contains nearly 4800 drug entries including >1,350 FDA-approved small molecule drugs, 123 FDA-approved biotech (protein/peptide) drugs, 71 nutraceuticals and >3,243 experimental drugs. Additionally, more than 2,500 non-redundant protein (i.e. drug target) sequences are linked to these FDA approved drug entries. Each DrugCard entry contains more than 100 data fields with half of the information being devoted to drug/chemical data and the other half devoted to drug target or protein data.

Virtual screening:

(VS) is a computational technique used in drug discovery to search libraries of small molecules in order to identify those structures which are most likely to bind to a drug target, typically a protein receptor or enzyme. Virtual screening has been defined as the "automatically evaluating very large libraries of compounds" using computer programs. As this definition suggests, VS has largely been a numbers game focusing on how the enormous chemical space of over 10^{60} conceivable compounds to a manageable number that can be synthesized, purchased, and tested. Although searching the entire chemical universe may be a theoretically interesting problem, more practical VS scenarios focus on designing and optimizing targeted combinatorial libraries and enriching libraries of available compounds from in-house compound repositories or vendor offerings. As the accuracy of

the method has increased, virtual screening has become an integral part of the drug discovery process.

Methods:

There are two broad categories of screening techniques: ligand-based and structure-based.

Ligand-based :

Given a set of structurally diverse ligands that binds to a receptor, a model of the receptor can be built by exploiting the collective information contained in such set of ligands. These are known as pharmacophore models. A candidate ligand can then be compared to the pharmacophore model to determine whether it is compatible with it and therefore likely to bind. A different strategy is to develop logic-based rules describing features of substructures and chemical properties related to activity using support vector inductive logic programming. The logic-based features provide insights into activity which can be understood by medicinal chemists. Support vector machine integrate the features to yield a quantitative QSAR, which is then used to screen a database of molecules. This approach is well suited to scaffold hopping to identify novel active molecules and is implemented in the package

Another approach to ligand-based virtual screening is to use 2D chemical similarity analysis methods to scan a database of molecules against one or more active ligand structure.

A popular approach to ligand-based virtual screening is based on searching molecules with shape similar to that of known actives, as such molecules will fit the target's binding site and hence will be likely to bind the target. There are a number of prospective applications of this class of techniques in the literature.

Structure-based:

Structure-based virtual screening involves docking of candidate ligands into a protein target followed by applying a scoring function to estimate the likelihood that the ligand will bind to the protein with high affinity.

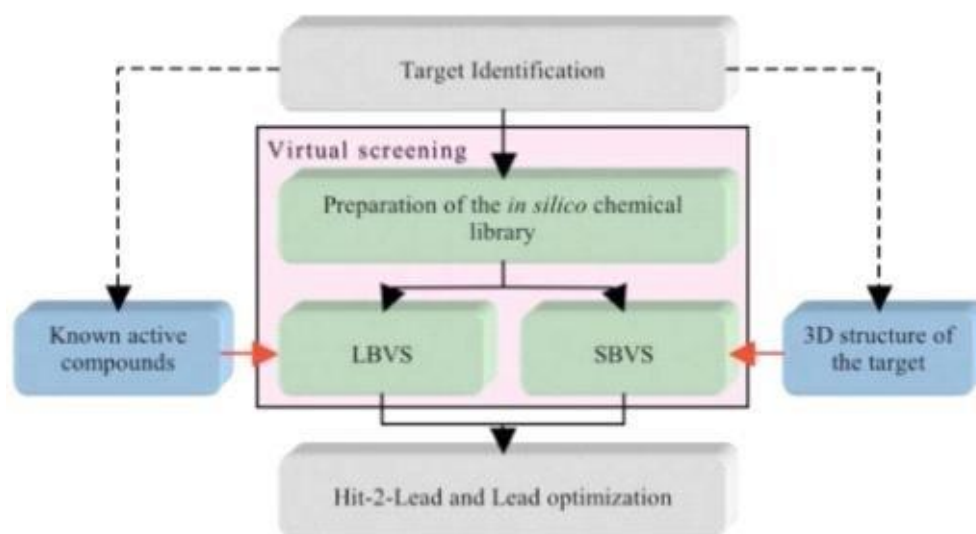


Figure 6- Process

8.1.1.1 Target Tertiary structure Protein:

Molecular Description of protein PDB Id: 3WB4

Classification:

Hydrolase

Structure Weight:

47261.01

Molecule:

Beta-secretase 1

Polymer:

1 Type:

protein

Length: 416

Chains:

A

Fragment:

ACTIVE PROTEASE DOMAIN, UNP residues 43-454

Organism

Homo sapiens

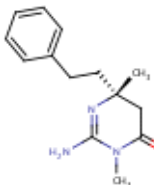
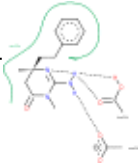
Gene Names

BACE1 BACE KIAA1149



Figure 7: 3D structure view of Bace-1 Protein.

8.2 Ligand Chemical Component:

Identifier	Formula	Name	View Interactions
<u>0B3</u> <u>Ligand</u>	 $C_{14}H_{19}N_3O$	(6R)-2-amino-3,6-dimethyl-6-(2-phenylethyl)-5,6-dihydropyrimidin-4(3H)-one	

9.0 SEARCHING OF THE SUITABLE LIGANDS MOLECULES FROM THE VARIOUS DATABASES:-

The suitable ligand for the selected receptor was selected by the process 2D- similarity search. In this process we will search for the similar type of ligands from the various small molecule data bases such as PUBCHEM, Ligand Info, Drug Bank, Zink, CHEMBANK, etc. The drug like compound select from this database on the basis of ligand template. Drug Bank database is a unique bioinformatics and cheminformatics resource that combines detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information. The database contains 6826 drug entries including 1431 FDA-approved small molecule drugs, 133 FDA-approved biotech (protein/peptide) drugs, 83 nutraceuticals and 5211 experimental drugs. Additionally, 4435 non- redundant protein (i.e. drug target/enzyme/transporter/carrier) sequences are linked to these drug entries. Each Drug Card entry contains more than 150 data fields with half of the information being devoted to drug/chemical data and the other half devoted to drug target or protein data. The 2d similarity search includes the following steps. They are: About 100 hit obtained from the database which is downloaded individually. These ligands have been made it in to single dataset by using the Discovery Studio 2.1 Suite to make them as a single set of ligands. This set can be used for the future activities such as the ADME Screening and Virtual screening process then for the Molecular docking as well.

9.1.1.1.1.1 Molecular Docking by using the SCRODINGER Bio suite 9.2:-

Molecular Docking is the process in which the intermolecular interaction between 2 molecules can be studied *In-silico*. In this process the Macro molecule is the protein receptor. The micro molecule is the Ligand molecule which can be act as an inhibitor. So, the Docking process involves the following steps:

- *Protein preparation*
- *Ligand preparation*
- *Receptor grid generation*
- *Ligand receptor docking*
- *Analysis of the docked complex and pose viewing*

The Protein-Ligand docking interaction was performed by using the Schrodinger Bio Suite

9.2 This Schrodinger bio Suite includes the various applications for the various Insilco processes. So, the Insilco process of Virtual Screening and molecular docking was performed by following the steps they are:

9.1.1.1.1.2 Protein Preparation by using the Protein preparation Wizard:-

For any In-silico process protein preparation is the important step, because most of the Protein structures that we take as the input are X-ray crystal structures, they are static, and they don't include H atoms. They have implicit water molecules which are helpful for the crystal stability, to study any one the biological process they are not helpful. So, to make the structure dynamic and functional protein has to be prepared. The protein preparation was done by performing the following steps. They are:

9.1.1.1.1.3 Ligand Preparation by using the 'Ligprep':-

Ligprep is the application of the Schrodinger bio suite, which is used to prepare the ligand by using the various force fields and they also prepares the ligands in a such a way that they can be used directly used for the ligand receptor docking. Different ligand compound were taken from the database. The Ligprep involves the following steps. They are:

- The application Ligprep is present under the applications Tab of the Schrodinger MaestroBio suite.

- In this the input of the ligand molecules are given which is the output of the Toxicity prediction (Extensible) program.
- Then the Epik was used to generate the possible ionic states of the all the molecules.
- Then OPLS2005 was taken as the Force field for minimize the structures.

Generate all possible combinations were selected to get the maximum no. of possible no. of ligand states.

- Then Ligprep was performed by starting the Ligprep program. In this way ligands are prepared by the Ligprep Program.

10.0 MOLECULAR DOCKING BY USING GLIDE:-

The next step is the Molecular docking. The docking is the process in which the Interaction between 2 molecules can be studied. As we have already prepared our protein and ligands, now we can perform the Molecular Docking. For the process of molecular docking Glide Application can be used, which is also a part of Schrodinger Bio suite 9.2

The Glide (Grid Based Ligand Docking with Energetic). This programme uses the latest force field such as the OPLS2005 For the calculations. The docking process by using Glide includes the 2 main steps, they are

- Receptor Grid Generation
- Ligand receptor docking

11.0 RECEPTOR GRID GENERATION:-

Receptor grid generation is the first step of the docking process in which the program will generate Grid Box around the active site of the protein molecule. This will help in the docking process to allow the ligand go bind it in to the main site of interaction, where active site of the protein complex is present, and the process of the receptor Grid generation can be performed by the following steps. They are

- *The Receptor Grid generation is present in the Glide Application in the Schrodinger Maestro.*

- *The input for the receptor Grid generation was given by taking the Co-crystallized ligand as the Center of the Grid.*
- *The active site residues are identified from the literature given as the input in constraints as the H-Bond constraints in the Constraints Tab. This active site includes both motif A and motif B.*
- *Then The default parameters are set for the Grid size and place then all others*
- *Then the Receptor Grid was generated by starting the program. The grid file was generated in the form of out.zip. This can be further used in the next ligand receptor docking step.*
- *In this way the Receptor Grid was generated.*

12.0 RECEPTOR LIGAND DOCKING: -

Receptor ligand docking is the process of Molecular docking in which the Protein molecule acts as your receptor and small molecule as Ligand. In the Glide program we can do the Docking in three ways, they are

- HTVS (High throughput virtual Screening)*
- SP (standard Precession)*
- XP (Xtra Precession)*

The HTVS is useful for the process of virtual Screening in which large no. of small molecules or ligands can be screened virtually to identify whether they can be interacted with the Receptor or not, It is a fast method to analyze the molecules and can be useful Screen a large no. of small molecule Libraries. Whereas the SP docking is the stand form of docking, it will take little more longer time to give the accurate interaction result of the docked molecules with the receptors. XP docking is the Précised method of docking where it will give the molecular interactions very precisely and accurately. So, it is a slow method when compared with the other method. These methods are present as a options in the receptor ligand docking menu of the Glide application.

The process of Receptor Ligand docking can be performed by the following the steps given below, they are:

- i. The Receptor ligand docking can be initiated by giving the ligand files (.sdf) in the ligand field and receptor grid file (.zip) as input files.
- ii. Then the HTVS option was selected initially to check whether the ligands have the interaction or not,
- iii. Then all the default parameters are set, the no, and runs per the docking also set as default.
- iv. Then the Ligand receptor docking was performed in HTVS manner.
- v. The output files are generated in the form of the out.maez form, then they are saved and scores viewed in the project table and analyzed and saved for the future purpose.
- vi. Then the SP docking was also performed for that ligand which has scored better in HTVS docking, to check their interactions in a precise manner.
- vii. Then SP docking results and scores were saved and analyzed.
- viii. Then XP docking was also performed for those molecules who have the better score and interactions in the SP docking, by giving the same parameters
- ix. Then those scores and results are also saved and analyzed for the future purpose.

In this way the Receptor ligand docking was done by the analysis of these results and scores, interactions we can conclude that which ligand can be said as the best possible drug like compound for the taken Receptors. The Results are analyzed and they are given in the Results and discussion, the results are discussed and brief conclusions are also given.

13.0 MD SIMULATION OF THE DOCKING PROTEIN :-

GROMACS-4.5.4 molecular dynamics package and 43a1 all atom force field was used to analyze the stability of models. The protein models were solvated with the Monte Carlo simulated TIP3P water using Cubic box with the size of 1.2 nm. Periodic boundary conditions were applied in all directions and the system was neutralized by adding sodium and chloride counter ions, replacing the water molecules. Subsequently, a maximum of 50,000 steps of energy minimization was carried out for the constructed models by the steepest descent algorithm with a tolerance of 1000 kJ/mol/nm. A twin range cutoff was used for long-range interactions using PME method: 1.0 nm for Vander Waals interactions and electrostatic interactions. These minimized and solvated systems were used as reasonable structures in terms of geometry and solvent orientation for further simulations. All bond

angles were constrained with LINCS algorithm while geometry of water molecules was constrained with SATTLE algorithm. Weak coupling method V-rescale was used for regulating the temperature while Parrinello-Rahman is used for setting pressure of the system. Equilibration MD for both temperature (300 K) and pressure (1 atm) was carried out for 100 ps. We observed that the temperature, pressure, density and total energy of the system were well equilibrated. These pre-equilibrated systems were subsequently used in the 3000 ps (3 ns) production MDS with a time-step of 2 fs. Structural coordinates were saved at every 2 ps and analyzed using GROMACS analysis tools. The lowest potential energy conformations were selected from the 5 MDS trajectory and further refined by energy minimization.

13.1 Result and Discussion: -

Energetically favorable conformation is very important for docking studies and other analysis because all the force fields utilize potential energy function to calculate the interaction energy. Molecular dynamic simulation was performed to analyze the stability of the homology models of the proteins and lowest energy conformation for each protein was chosen for further docking studies. Energy Graph is constantly all the simulate time.

The MD analysis concerned the stability of simulated Bace-1 protein (Figure:3) during the simulation by monitoring the RMSD profiles as computed by the atomic displacements from MD trajectories. Figure 2A shows that RMSD value 2.5 to 3.0 nm constant between during simulation, and then remains quite stable in the following simulation.

RMSF: Residues 246 to 278 and 310-318 this Amino Acid more fluctuate towards ligand , which suggested that the Binding affinity of the protein has increase towards ligand making protein docking ligand complex stable as the shown in RMSF graph.

Gromacs Energies

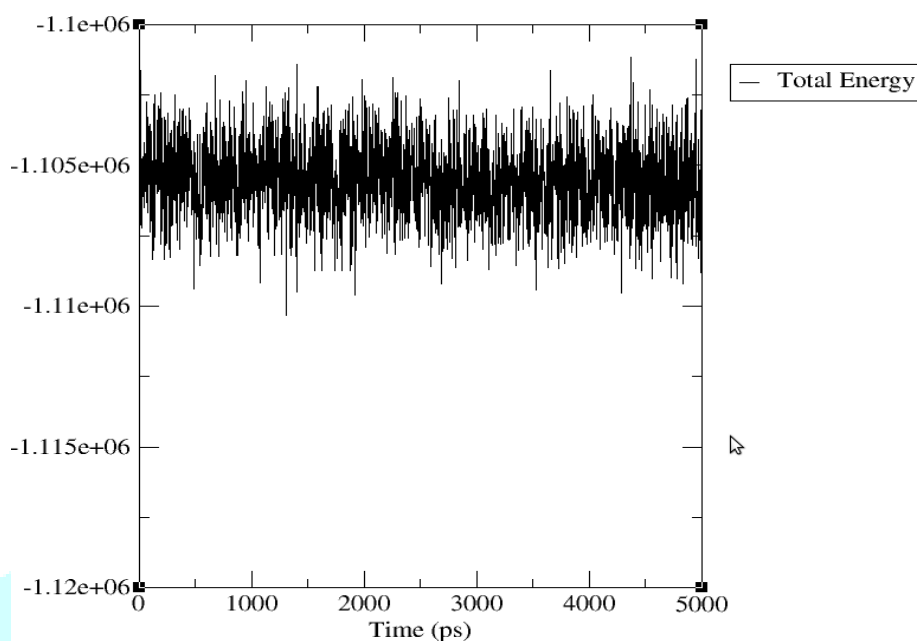


Figure 8 Molecular dynamics-based analyses for the model refinement. Potential energy shows the energetically stable conformation for further docking studies. Lowest energy conformation obtained at 5 ns.

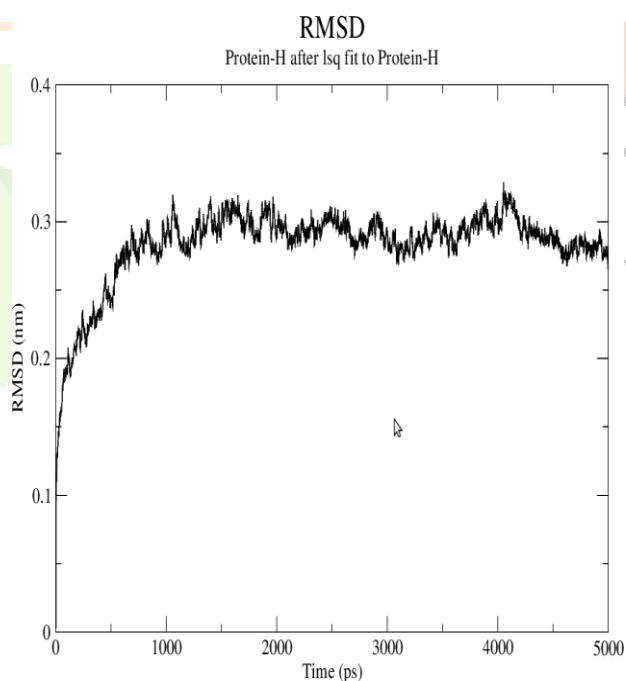


Figure 9 RMSD of the backbone atoms with respect to initial structure shows stable nature of model after initial equilibration time

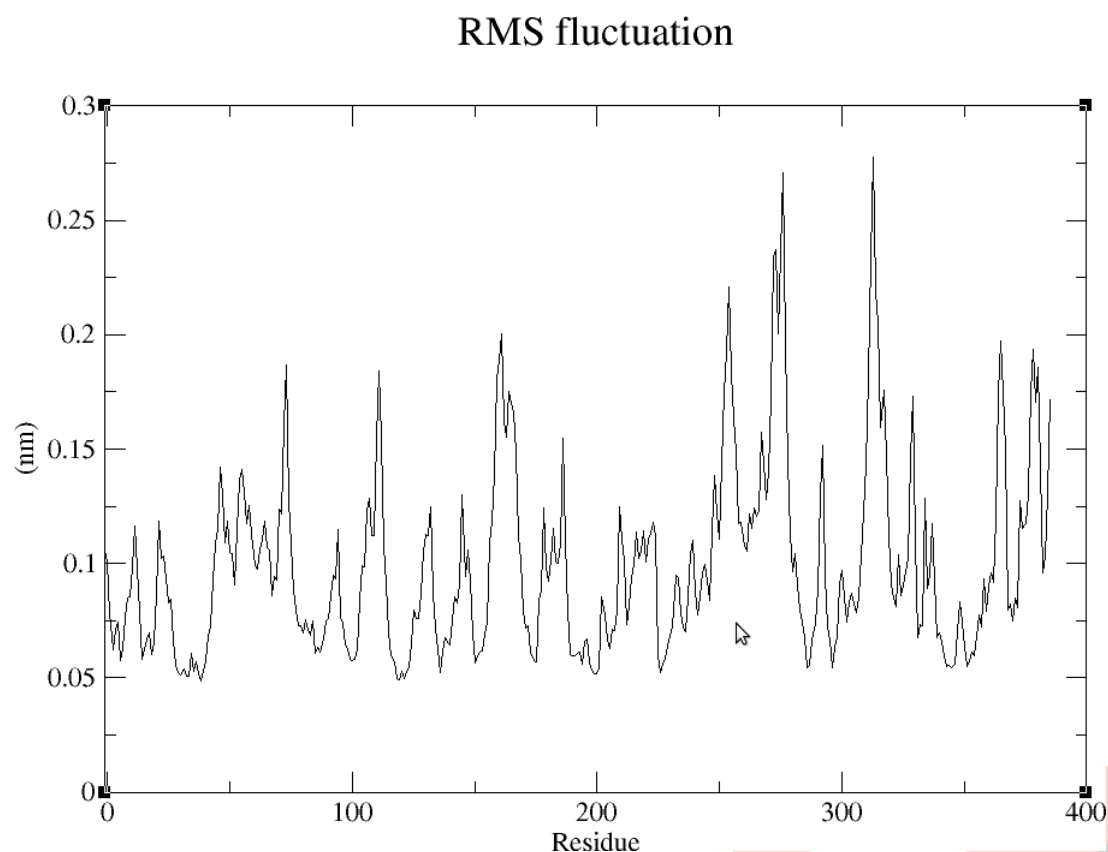


Figure 10 RMSF graph showing fluctuation in protein with respect to time.

14.0 RESULTS OF THE MOLECULAR DOCKING PROCESS BY USING THE GLIDE:-

Test compounds which are obtained from the database screening used to form a ligand library. This library is taken and the Molecular docking with the Receptor protein which is the Bace-1 of human was performed by using the Glide program which is one of the Application in the Schrodinger Bio suite 9.2. **All ligands were taken and performed the three methods of docking viz. 1>HTVS, 2>SP and 3>XP.**

The results of the Docking studies can be studied by the analysis of the Number of Hydrogen bonds the ligand is forming with the receptor active site residues and the No. of Hydrophobic interactions also plays an important role in the Complex between protein and ligand.

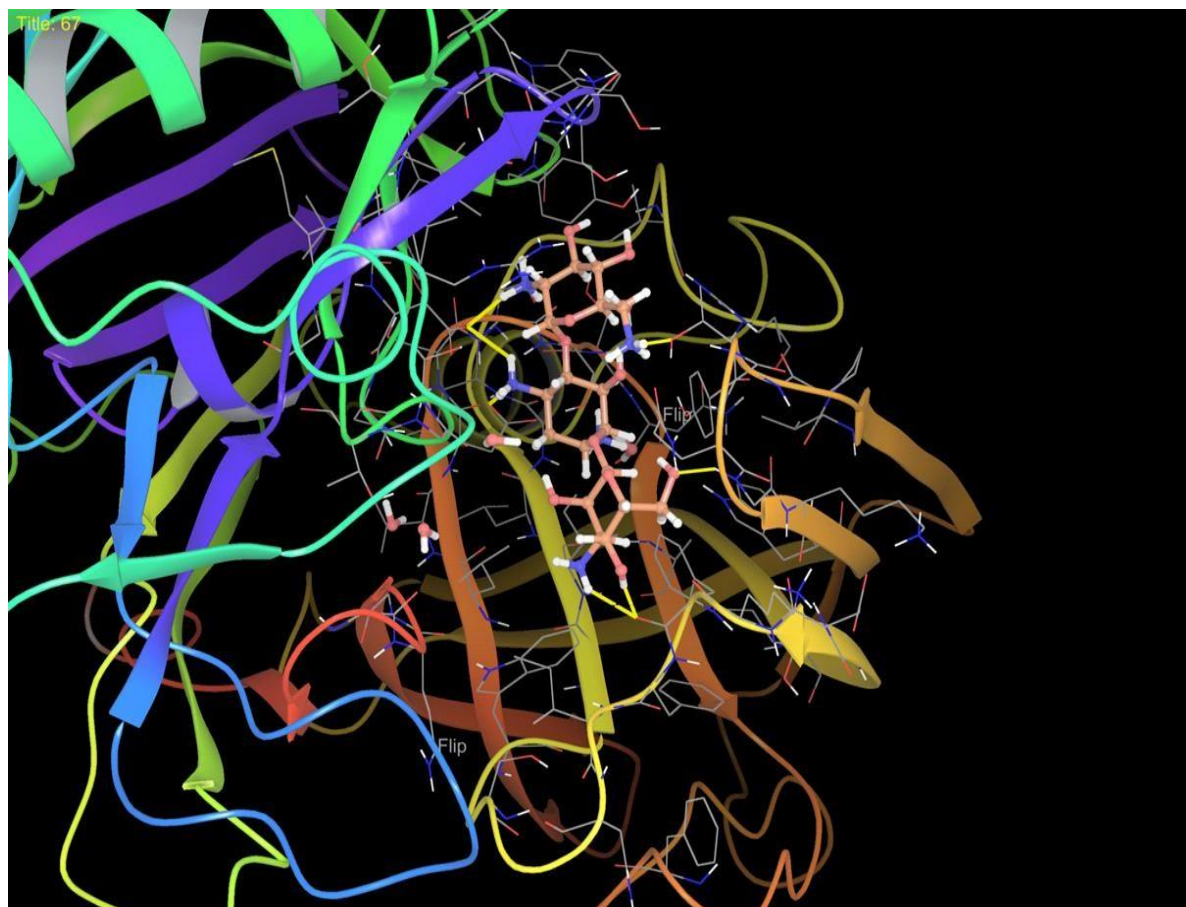
The Strength of the interaction can also be analyzed by the Scoring of the each possible interaction. That score is called as the glide score or G-score. The more the negative score the better the interaction. The ligand or compound giving highest negative score is choose as a best ligand and further analysis is performed by taking it with the receptor protein. The binding site residue suggest by CAST-P server and active site obtain from the literature is similar with the interaction site observe in the docked complex of protein and ligand molecule.

The various methods of docking and the interactions were tabulated in the given tables below:

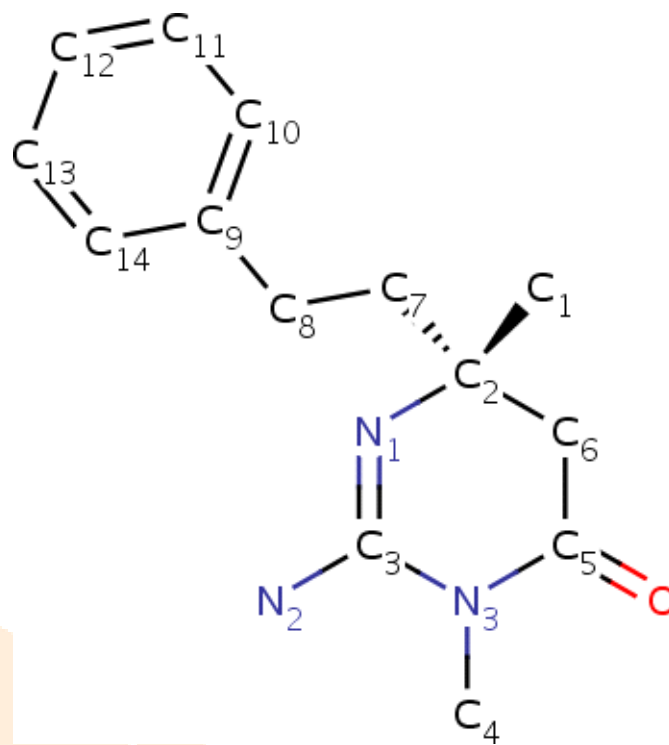
Table 1 - Interaction analysis of compounds with 3WB4 protein(Beta-Secretase (BACE1) were independently Docked using Schrodinger. Docking Score and Glide Energy residues participating in the H- bonding are shown in the Table.

No. of Test compounds	Residues of Base -1 involved in interaction withTest compounds		Docking Score	Glide Energy
	Hydrogen Bond	Hydrophobic Interaction		
Testcompound 2	4-HB	G230, K107, D228	-13.399282	-45.932744
Testcompound 28	5-HB	D32,F108,T72,232	-12.524351	-45.927282
Testcompound 144	7-HB	D32,228,K107,T231,Y198,R235	-12.437461	-63.557447
Testcompound 40	7-HB	K107,G230,D28,228,I226,T231,R235	-11.79099	-55.419461
Testcompound 158	7-HB	Q73,K107,D228,R128,G34	-11.683985	-68.292271

No. of Test compounds	Residues of Base -1 involved in interaction with Test compounds		DockingScore	Glide Energy
	Hydrogen Bond	Hydrophobic Interaction		
Test compound 6	4-HB	Y198,R128,Q73,F108	-11.645063	-63.87108
Test compound 16	5-HB	K107,G230,D228,R235	-11.275358	-50.153616
Test compound 113	7-HB	D32,228,R235,T72,G230,Y71	-11.222617	-53.412012
Test compound 67	7-HB	D32,228,T72,G74,F108	-11.06299	-57.354266
Test compound 86	9-HB	D32,228T72,R235,K107,F108	-10.828531	-56.443688
Test compound 15	4-HB	D32,G230,T72,F109	-10.773839	-58.475864
Test compound 145	6-HB	D32,K107,F108,Y198,T231	-10.75888	-48.446711
Test compound 73	7-HB	G11,34,230,D32,K107,F109,	-10.706107	-66.531727
Test compound 5	8-HB	G34,Q73,K107,F108,109,D228	-10.693217	-48.343251
Test compound 41	8-HB	D32,228,Q73,G74,R235,T329	-10.569875	-53.066374
Test compound 42	6-HB	G11,D32,228,T73,Y198,R235	-10.545335	-52.066374



14.1.1.1.1.1 Figure 10- Illustration Shows the Ligand Interaction(Hydrogen bond Interaction shows in Yellow Colour)PDB Id : 3WB4 (Beta- Secretase1)



Ligand , OB3 : (6R)-2-amino-3,6-dimethyl-6-(2-phenylethyl)- 5,6-dihydropyrimidin-4(3H)-one
Fig6 : 2-d diagram of ligand compound .

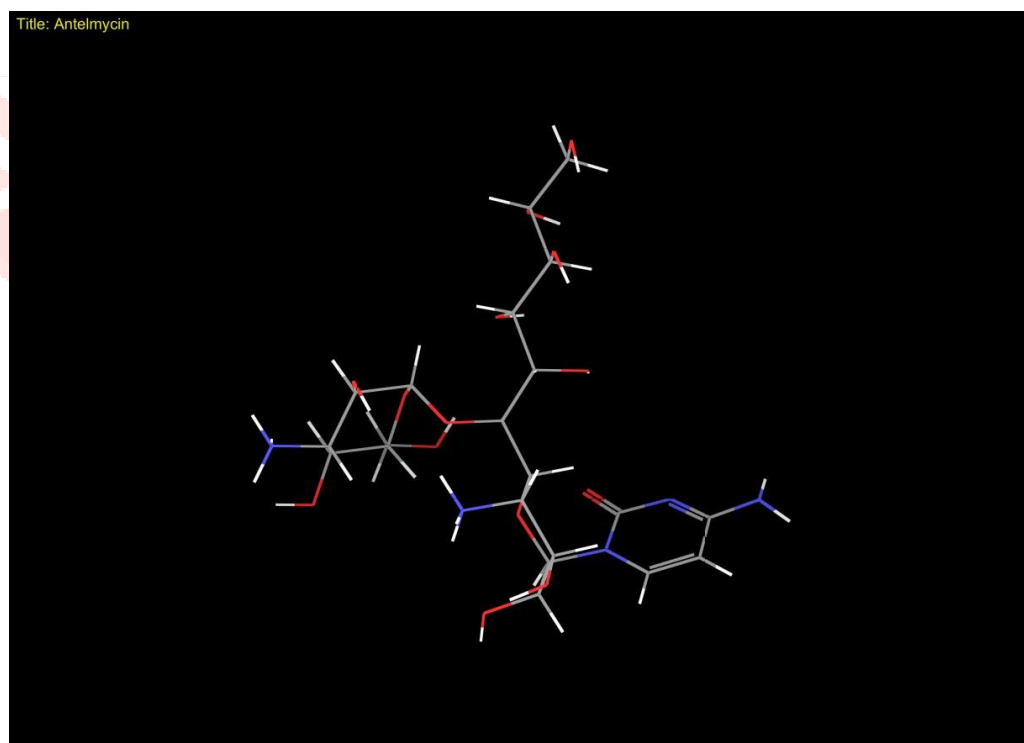


Figure 11 - - Ligand picture in 3-D mode.

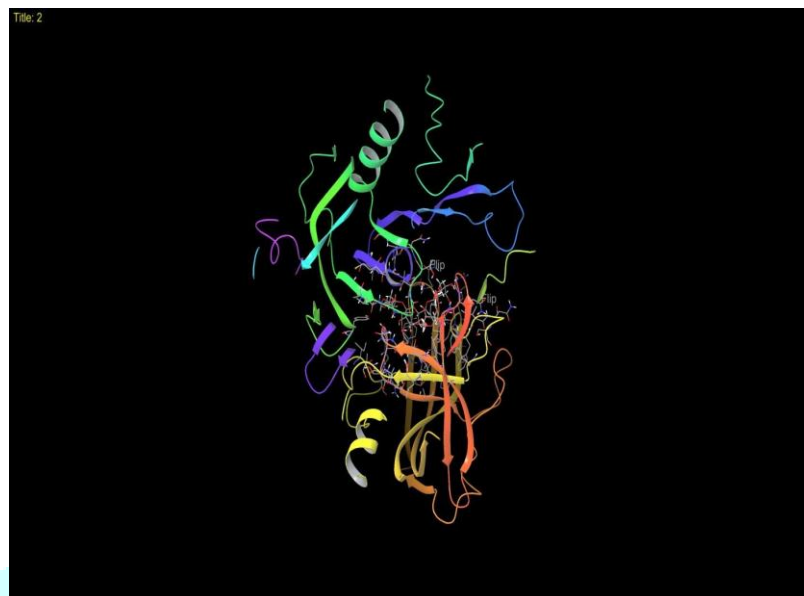


Figure 12 -Shown the Ligand Interaction a Docking pose.

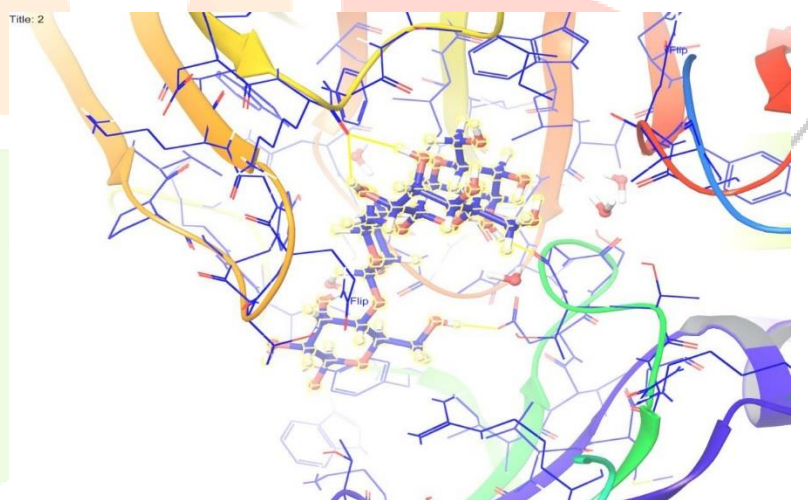
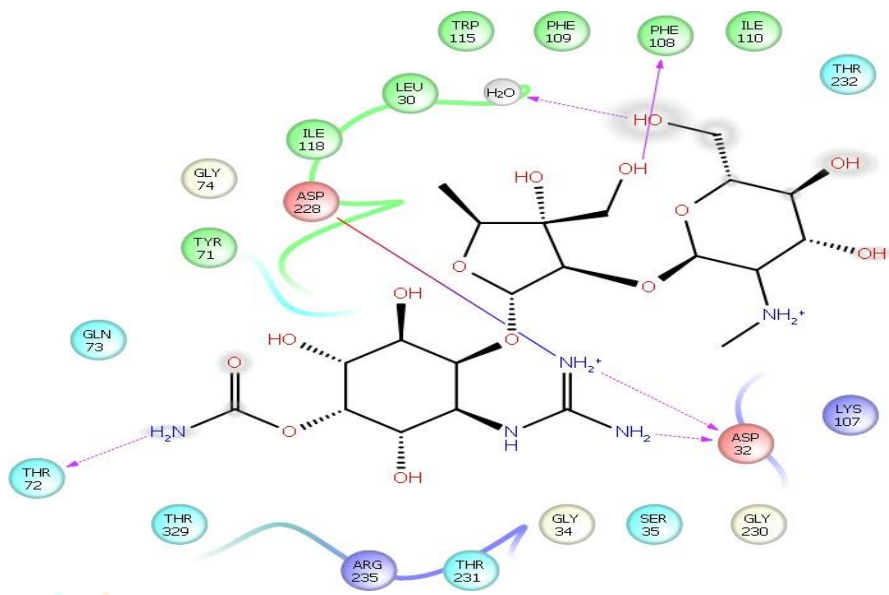


Figure 13 Shown the Ligand Interaction a Docking pose.



1.1.1.1 Figure 14 Ligand Interaction shows with different Amino Acid Residues position

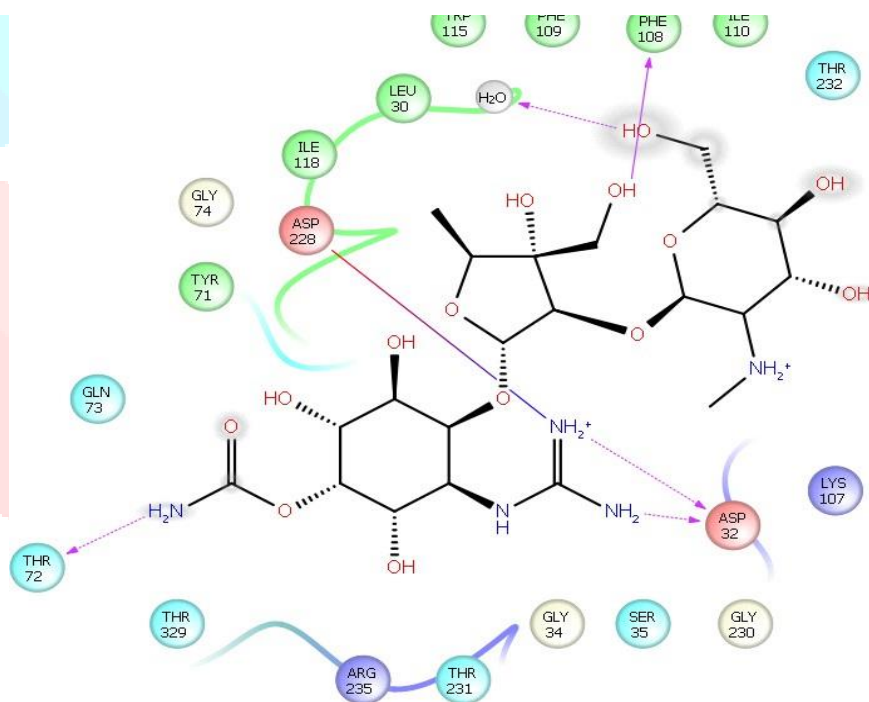
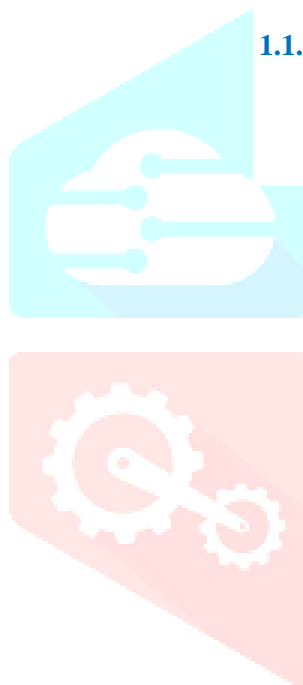


Figure 15 Ligand Interaction shows with different-2 Amino Acid Residues

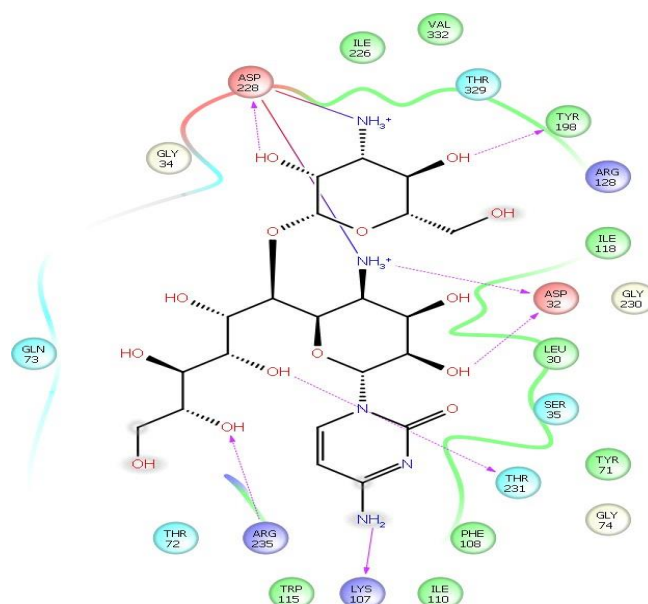


Figure 16 Ligand Interaction shows with different Amino Acid Residues

15.0 CONCLUSION & FUTURE ASPECTS:

15.1 Domain of the Study

The scope for the future study about the project can be studied in more than one ways. Most of the Bace-1 enzyme are involved in the various disease cycles of numerous Amyloidogenesis. Resulting Bace-1 inhibitor may eventually lead to development of drug that would be useful in Alzheimer's Diseases research, the protein-Ligand interaction can also be studied based on this Bace-1 enzyme *In-silico studies*.

Once the new inhibitors or potent lead molecules are found for these Bace-1 the full length MD simulations for longer time scale can be performed for the better analysis and which helps us to understand the stability and sustainability of the protein- inhibitor complexes. This *In-silico* study has scope for the immediate future.

By performing virtual screening a wide range of 3D descriptor can be generated from the dock complex. This descriptor can then be used, in conjunction with docking score, to improve discrimination between true and false positive. The identified lead compound may serve as a base for future studied by doing QSAR studied of identified test compounds.

16.0 CONCLUSIONS

The process of the inhibitor test compound finding through structure-based drug designing (SBDD) of the Bace-1 enzyme have revealed that the new possible inhibitors have been found. Docking studies shows that the ligand binds active site which consist both sub sites for binding of cofactor and substrate in the binding site of the protein structure is a unique one. This is due to the uniqueness of the Active site and binding site present under one domain in the protein structure. Even though 8 to 12 lead molecule Test compounds showed the good interaction with the protein molecule. The complex formed between them is stronger. That can be analyzed through their Docking Scores and the intermolecular interactions between the ligand protein receptor complex and highly stable as it has the high hydrogen bonds and more hydrophobic interactions.

The Molecular Dynamic Simulation (MDS) shows that the 5ns and the potential energy is also decreased. This is the sign of the stable complex at constant temperature and pressure. The other parameters are also supporting the stability phenomenon. So, from the result of the MDs we can say that the protein is stable. Furthermore, if simulation is extended more than 5 ns, more stable structure can be obtained.

Since ligand compound obtain showing good interaction by making more than 6 hydrogen bonds at different site in and there is also present of hydrophobic interaction observe in the docked ligand complex. This shows that protein ligand compound is stable.

Hence this test compound can be served as a lead compound for identification of same type of inhibitor from different database. By further modification and by adding different functional group this drug like compound can be used to treat of Alzheimer's disease in which activity of this enzyme in not clearly known.

17.0 REFERENCES:

- F.K. Brown (1998). "Chapter 35. Chemoinformatics: What is it and how does it Impact Drug Discovery". *Annual Reports in Med. Chem.* 33: 375.
- Brown, Frank (2005). "Editorial Opinion: Chemoinformatics – a ten-year update". *Current Opinion in Drug Discovery & Development* 8 (3): 296–302.
- HANN, M., GREEN, R. Chemoinformatics– a new name for an old problem? *Curr. Opin. Chem. Biol.* 1999, 3, 379–383.
- G. Paris (August 1999 Meeting of the American Chemical Society), quoted by W. Warr at <http://www.warr.com/warrzone.htm>.
- Gasteiger J.(Editor), Engel T.(Editor): *Chemoinformatics : A Textbook*. John Wiley & Sons, 2004, ISBN 3-527-30681-1.
- A.R. Leach, V.J. Gillet: *An Introduction to Chemoinformatics*. Springer, 2003, ISBN 1-4020-1347-7.
- RITCHIE, T. Chemoinformatics; manipulating chemical information to facilitate decision-making in drug discovery. *Drug Discovery Today* 2001, 6(16), 813–814.
- JOHNSON, D.E., BLOWER, P.E., Jr, MYATT, G.J., WOLFGANG, G.H. Chem-tox informatics: data mining using a medicinal chemistry building block approach. *Curr. Opin. Drug Disc. Dev.* 2001, 4(1), 92–101.
- OPREA, T.I., GOTTFRIES, J. Chemography: the art of navigating in chemical space. *J. Comb. Chem.* 2001, 3, 157–166.
- OPREA, T.I. Chemical space navigation in lead discovery. *Curr. Opin. Chem. Biol.* 2002, 6(3), 384–389.
- LEACH, A.R., GILLET, V.J. *An Introduction to Chemo informatics*. Kluwer Academic Publishers, Dordrecht, Netherlands, 2003, 259.
- Drews J, *Drug discovery: a historical perspective*, Science, 287 5463: pp1,960-1,964, 2000
- www.drugbank.ca/ accessed on 29th October 2022.
- "RSC acquires Chem Spider". Royal Society of Chemistry 11 May 2009. <http://www.rsc.org/AboutUs/News/PressReleases/2009/ChemSpider.asp>. Retrieved 2009-05-11
- Welcome ChemMantis to ChemZoo and a Call for Contributions from the Community, 2008-10-23, A. Williams, blog post *Chemical Entities of Biological Interest*
- International Union of Pure and Applied Chemistry. "molecular entity". *Compendium of Chemical Terminology* Internet edition.
- Broach JR, Thorner J. High-throughput screening for drug discovery. *Nature*. 1996;384(Suppl 7 Nov 1996):14-16.
- Brown FK. Chemoinformatics: What is it and how does it impact drug discovery? *Annu Rep Med Chem*. 1998;33:375-384.
- Gasteiger J, Engel T, editors. *Chemoinformatics: A Textbook*. Weinheim, Germany: Wiley-VCH; 2003. 649p.
- Plunkett MJ, Ellman JA. Combinatorial chemistry and new drugs. *Sci Am*. 1997 April;68-73.
- Russo E. Chemistry plans a structural overhaul. *Nature*. 2002;419(6903):4-7.

- ALLEN, F.H., MOTHERWELL, W.D. Applications of the Cambridge structural database inorganic chemistry and crystal chemistry. *Acta Crystallogr, Sect B* 2002, 58(Pt 3 Pt 1), 407–422.
- ALLEN, F.H., DAVIES, J.E., GALLOY, J.J., JOHNSON, O., KENNARD, O., MACREA, C.F., MITCHELL, E.M., MITCHELL, G.F., SMITH, J.M., WATSON, D.G. The developments of versions 3 and 4 of the Cambridge database system. *J. Chem. Inf. Comput.Sci.* 1991, 31, 187–204.
- ABOLA, E.E., BERNSTEIN, F.C., KOETZLE, T.F. The protein data bank. In *The Role of Data in Scientific Progress*, GLAESER, P.S. (ed.) Elsevier, New York, 1985.
- C.D., ELLIS, R.A., GRAESSER, S., MARSHALL, G.R. Display and manipulation in threedimensions.
- In *Pertinent Concepts in Computer Graphics*, FAIMAN, M. NIEVERGELT, J. (Eds). University of Illinois Press, Chicago, IL, 1969, 104–153.
- FUJITA, T., IWASA, J., HANSCH, C. A new substituent constant, σ , coefficients. *J. Am. Chem. Soc.* 1964, 86(December 5), 5175–5180.
- HANSCH, C., LEO, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*. Wiley & Sons, New York, 1979. http://www.udel.edu/chemo/Images/chemometrics_scheme.jpg

