DESIGN AND SYNTHESIS OF H_3 RECEPTOR INVERSE AGONISTS WITH AchE INHIBITOR ACTIVITY AND QSAR STUDY OF H_3 RECEPTOR ANTAGONISTS

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Technology

in

Biotechnology & Medical Engineering

By

AKALABYA BISSOYI

(207BM207)

Department of Biotechnology & Medical Engineering

National Institute of Technology

Rourkela-769008, Orissa, India

2009

DESIGN AND SYNTHESIS OF H_3 RECEPTOR INVERSE AGONISTS WITH AchE INHIBITOR ACTIVITY AND QSAR STUDY OF H_3 RECEPTOR ANTAGONISTS

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Technology

in

Biotechnology & Medical Engineering

By

AKALABYA BISSOYI

(207BM207)

Under the Guidance of

Prof. Gyana R. Satpathy

Department of Biotechnology & Medical Engineering

National Institute of Technology

Rourkela-769008, Orissa, India

2009

Dr. Gyana Ranjan Satpathy
Professor and Head
Department of Biotechnology & Medical Engineering
National Institute of Technology, Rourkela, Orissa, India

Certificate

This is to certify that the thesis entitled **"Design and Synthesis of H_3 receptor inverse agonists with AchE inhibitor activity and QSAR study of H_3 receptor antagonist"** by **Akalabya Bissoyi** submitted to the National Institute of Technology, Rourkela for the Degree of Master of Technology is a record of bonafide research work, carried out by him in the Department of Biotechnology and Medical Engineering under my supervision. I believe that the thesis fulfils part of the requirements for the award of master of Technology. The results embodied in the thesis have not been submitted for the award of any other degree.

Prof. Gyana .R. Satpathy

Department of Biotechnology and Medical Engg.

NIT Rourkela

Acknowledgement

I avail this opportunity to express my indebtedness, deep gratitude and sincere thanks to my guide, **Prof. Gyana R. Satpathy,** Head of the Department, Department of Biotechnology and Medical Engineering for his in depth supervision and guidance, constant encouragement and co-operative attitude for bringing out this thesis work successfully. I enjoyed the freedom he gave me throughout this work that was instrumental in discovering a **scientific "self"** within me.

Finally I extend my sincere thanks to **Mr.Shadrack Jabes B, Mr.Sripad Chandan Patnaik, Mr Jagannath Mallick**, **Miss Sheetal Arora**, **Ms. Deepanwita Das** who have helped me during my dissertation work and have been involved directly or indirectly in my endeavor.

And it goes without saying, that I am indebted to my parents **Mr. Sarat Chandra Bissoyi** and **Mrs. Shanti Bissoyi** and Brother **Sabyasachi Bissoyi**, whose patience, support and endurance made completion of my course a reality

Akalabya Bissoyi

Roll No: 207BM207

M.Tech (Biotechnology)

Dept. of Biotechnology & Medical Engg.

N.I.T. Rourkela, Orissa, India.

CONTENTS

List of Tables

List of Schemes/flowchart

List of Figures

Fig 15: Details of the AchE active site in which the superposition of the crystallographic orientation of donepezil with the top 3 ranked solution suggested by GOLD is shown.

Fig 16: Potential energy graph of protein –ligand complex during 1ns molecular simulation in the active site using GROMACS 3.3.1

Fig 17: Plot showing the RMSD deviation of ligand in the solvated protein during 1ns molecular simulation in the active site using GROMACS 3.3.1

Fig 18: Plot showing the RMSD deviation of l in the solvated protein back bone during 1ns molecular simulation in the active site using GROMACS 3.3.1

ABBREVIATIONS

AchE	Acetyl-cholinesterase inhibitors.
H_3	Histamine-3 receptor.
AD	Alzheimer's disease.
QSAR	Quantitative Structure Activity Relationship.
MDS	Molecular Dynamics Simulation.
CADD	Computer-assisted drug design.
APP	Amyloid Precursor Protein.

ABSTRACT

Currently, acetyl cholinesterase and *N*-methyl-D-aspartate antagonists are commercially available for the treatment of Alzheimer's disease (AD). Approach of using multifunctional inhibitors to reduce the side effects of available drugs is the main objective of this work. Presently, Histamine-3 (H_3) receptor antagonists are used for the treatment of several neurodegenerative disorders such as Epilepsy, Alzheimer's and Parkinson's diseases. Both H_3 and AchE inhibitors cure the symptoms of Alzheimer by enhancing the acetylcholine levels in the brain. But the mechanism of action involved in both the cases is different. Here, we propose histamine-3 antagonist with acetyl cholinesterase (AchE) inhibitor activity as a novel class of drugs which can be used to treat Alzheimer's disease with less adverse peripheral effects caused by excessive AchE inhibitor. Our present study can be divided into two parts. In the first part, homology modeled structure of H_3 active site and available crystal structure of AchE was used to collect the information for pharmacophore identification. The important descriptors were identified based on comparative 2D-QSAR and 3D-QSAR study of 28 druggable compounds for H_3 receptor collected from the literature. In the second part, five hybrid molecules were generated based on the pharmacophore of H_3 receptor and known pharmacophore of AchE inhibitors. All five hybrid molecules were screened through ADME/tox filters. The hybrid molecule was validated through GOLD docking score in both AchE and H_3 receptor. The best hybrid compound (hybrid-3) was then evaluated by molecular dynamics (MD) simulation in water solvent model using 3D model of human H_3 receptor (build based on bovine rhodopsin structure).

Keywords: *Alzheimer's disease, Homology model, QSAR, docking, pharmacophore modeling, Molecular dynamic study, ADME/tox*

CHAPTER 1

INTRODUCTION

INTRODUCTION

Bioinformatics is conceptualizing biology in terms of molecules (in the sense of physical chemistry) and applying "***informatics techniques***" (derived from disciplines such as applied math's, computer science and statistics) to ***understand*** and ***organize*** the ***information*** associated with these molecules, on a ***large scale***. In short, bioinformatics is a management information system for molecular biology and has many ***practical applications***.

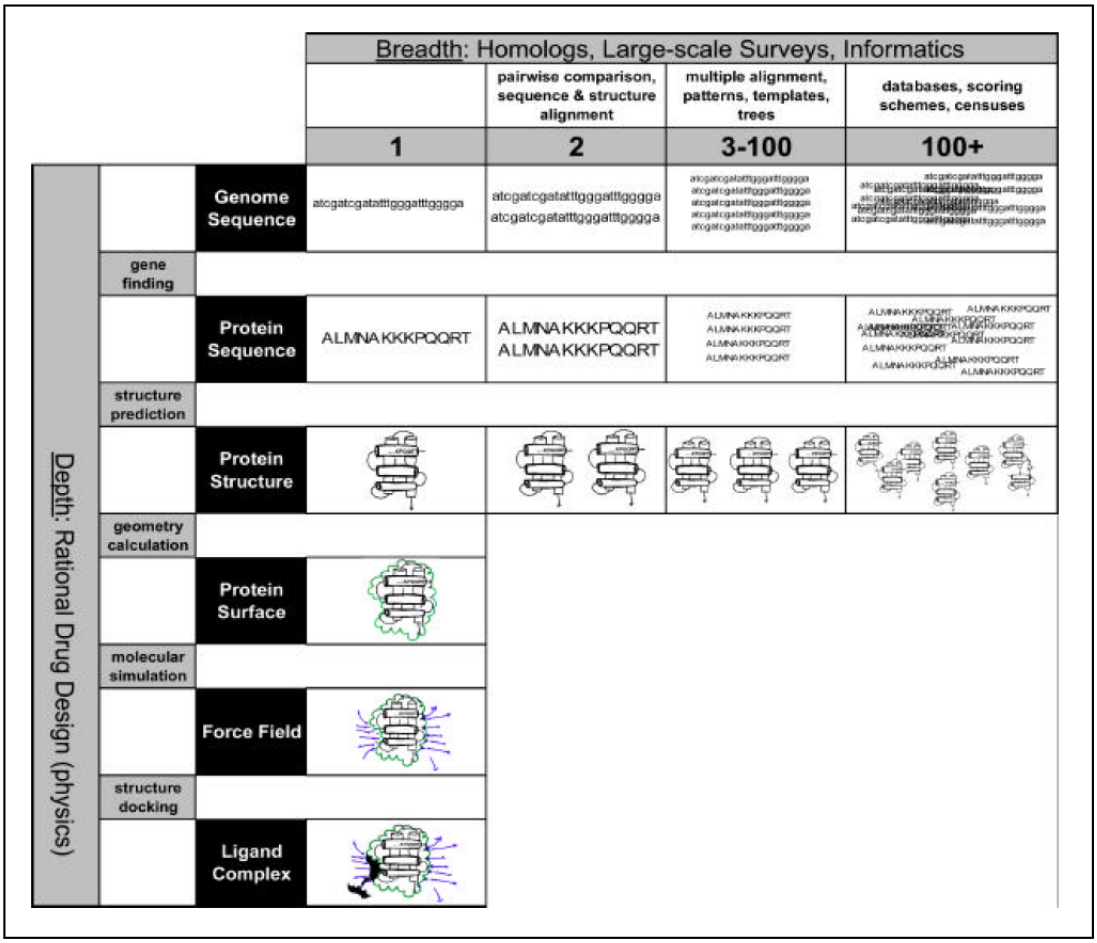


Fig 1 Evolution of Bioinformatics

Fig.1 represents a paradigm shifts during the past couple of decades have taken much of biology away from the laboratory bench and have allowed the integration of other scientific disciplines, specifically computing. The result is an expansion of biological research in breadth and depth.

The vertical axis demonstrates how bioinformatics can aid rational drug design with minimal work in the wet lab. Starting with a single gene sequence, we can determine with strong certainty, the protein sequence. From there, we can determine the structure using structure prediction techniques. With geometry calculations, we can further resolve the protein's surface and through molecular simulation determine the force fields surrounding the molecule. Finally docking algorithms can provide predictions of the ligands that will bind on the protein surface, thus paving the way for the design of a drug specific to that molecule.

The horizontal axis shows how the influxes of biological data and advances in computer technology have broadened the scope of biology. Initially with a pair of proteins, we can make comparisons between the between sequences and structures of evolutionary related proteins. With more data, algorithms for multiple alignments of several proteins become necessary. Using multiple sequences, we can also create phylogenetic trees to trace the evolutionary development of the proteins in question. Finally, with the deluge of data we currently face, we need to construct large databases to store, view and deconstruct the information. Alignments now become more complex, requiring sophisticated scoring schemes and there is enough data to compile a genome census – a genomic equivalent of a population census – providing comprehensive statistical accounting of protein features in genomes.

Applications of Bioinformatics

- Database query tools
- Sequence analysis and molecular Evolution
- Genome mapping and comparison
- Gene identification
- Structure prediction
- Drug design and drug target identification

CADD

Drug design is the approach of finding drugs by design, based on their biological targets. Typically a drug target is a key molecule involved in a particular metabolic or signalling pathway that is specific to a disease condition or pathology, or to the infectivity or survival of a microbial pathogen.

Computer-assisted drug design (CADD), also called computer-assisted molecular design (CAMD), represents more recent applications of computers as tools in the drug design process. In most current applications of CADD, attempts are made to find a ligand (the putative drug) that will interact favorably with a receptor that represents the target site. Binding of ligand to the receptor may include hydrophobic, electrostatic, and hydrogen-bonding interactions. In addition, solvation energies of the ligand and receptor site also are important because partial to complete desolvation must occur prior to binding.

This approach to CADD optimizes the fit of a ligand in a receptor site. However, optimum fit in a target site does not guarantee that the desired activity of the drug will be enhanced or that undesired side effects will be diminished. Moreover, this approach does not consider the pharmacokinetics of the drug.

Based on the information that is available, one can apply either,

- **Ligand-based drug design** is applicable when the structure of the receptor site is unknown, but when a series of compounds have been identified that exert the activity of interest. To be used most effectively, one should have structurally similar compounds with high activity, with no activity, and with a range of intermediate activities.

- **Receptor-based drug design** incorporates a number of molecular modeling techniques, one of which is docking. Docking allows scoring based on force fields, which include both van der Waals and electrostatic interactions. These results illustrate the potential for docking programs to search objectively for ligands than are complementary to receptor

sites, thereby assisting researchers in identifying potential drugs than may be considerably different from existing drugs.

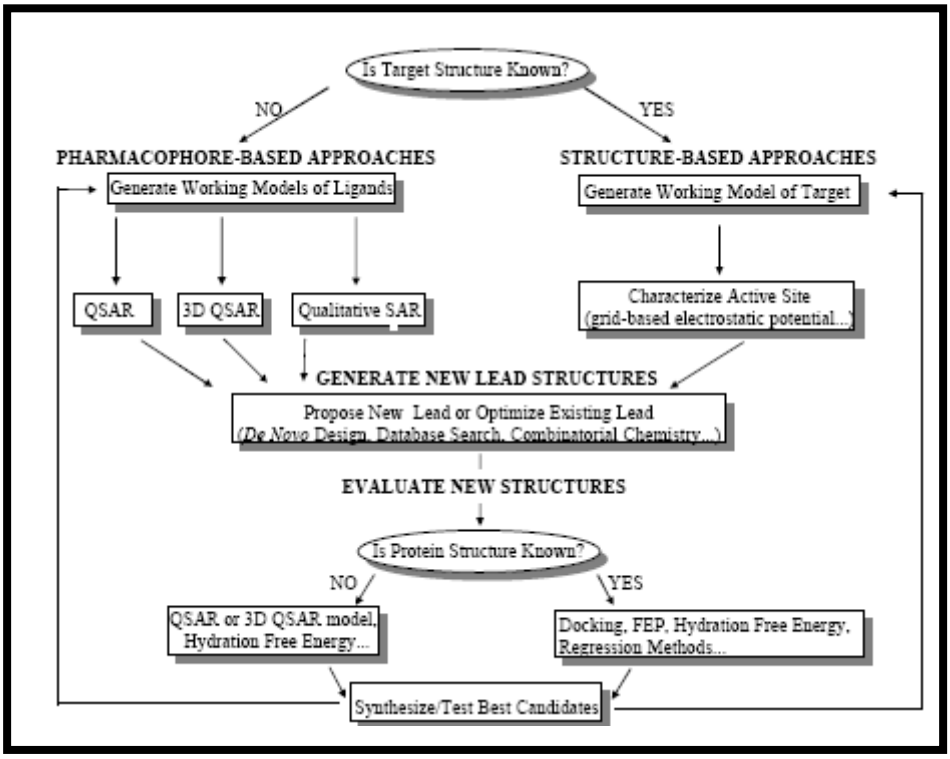


Fig 2 Flowchart representing CADD

Benefits of CADD

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

- **Cost Savings.** Many biopharmaceutical companies now use computational methods and bioinformatics tools to reduce cost burden. 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.
- **Time-to-Market.** The predictive power of CADD can help drug research programs choose only the most promising drug candidates. By focusing drug research on specific lead candidates, biopharmaceutical companies can get drugs to market more quickly.

- **Insight.** One of the non-quantifiable benefits of CADD and the use of bioinformatics tools is the deep insight that researchers acquire about drug-receptor interactions. When we show researchers new molecular models of their putative drug compounds, their protein targets and how the two bind together, they often come up with new ideas on how to modify the drug compounds for improved fit.

CADD and Bioinformatics

Bioinformatics was seen as an emerging field with the potential to significantly improve how drugs are found, brought to clinical trials and eventually released to the marketplace. CADD methods are heavily dependent on bioinformatics tools, applications and databases. As such, there is considerable overlap in CADD research and bioinformatics.

There are several key areas where bioinformatics supports CADD research.

- Virtual High-Throughput Screening (vHTS)
- Sequence Analysis
- Homology Modeling
- Similarity Searches
- Drug Lead Optimization
- Physicochemical Modeling
- Drug Bioavailability and Bioactivity

CADD and bioinformatics together are a powerful combination in drug research and development. An important challenge for going forward is skilled and efficient management of all the bioinformatics resources available to us. Bioinformatics is a field of study which focuses on the analysis of large amounts of biological data. The primary focus is the analysis and interpretation of complex biological molecules, such as DNA and proteins. Because these molecules tend to be very complex, they must be analyzed using computers.

Molecular biologists often use laboratory-based experiments to analyze a molecule of biological interest. Usually, these are molecules of DNA or protein. Both DNA and protein molecules are

chains of smaller molecules attached to each other to build a large string-like molecule. Proteins are the 'machinery' of biology - they do everything from carrying oxygen through the blood (an example of this is Hemoglobin) to movement (an example of this is actin and myosin, which are found in muscle fibers). DNA is the primary data source of life. Cells use DNA molecules to build proteins, which then perform most biological processes.

Biology of bioinformatics

A molecule of DNA is composed of units called *nucleic acids*. These nuceic acids (there are four distinct nucleic acids in DNA) are assembled and stored in the nucleus of a cell. The nucleus is basically a storage house of data for the cell. Anything a cell could possibly want, it can build, using the information stored in its DNA. When a cell wants to build a protein, it finds the appropriate section of DNA, unravels it, and decodes it. This decoding process is how proteins are built. Proteins are composed of units called *amino acids*. There are twenty amino acids that combine to form most proteins. Proteins are the 'machinery' of a cell. They can perform many functions, like transportation, structural support, movement, metabolism, etc. A basic principle of bioinformatics is that everything about these molecules can be inferred from the sequence of building blocks (nucleic acids or amino acids) from which the molecule is composed. For example: In theory, the structure of a protein (its shape) can be determined by analyzing the amino acids that make up the protein. This structure can often be used to deduce its function.

Computational side of bioinformatics

Chemists have developed methods for understanding the shape and behavior of small molecules, using mathematical analysis. They might use computers (or even just a pencil and paper) to study these molecules. As biochemists attempt to study larger and larger molecules, they have found that these methods are no longer feasible. It would simply take too much time to determine the structure of a large molecule. Using the classical methods, and a very fast computer (to perform the mathematics), it would take years, centuries, or even millennia to find the exact structure of a very large protein. Because of this, researchers have developed new computational methods to approximate the structure (or other properties) of a molecule in a much more reasonable amount of time. Protein structure analysis is a very important part of bioinformatics, but it does not encompass the field. There are other many studies being done in bioinformatics. Some of the questions that are addressed by these studies follow:

How does a particular protein bind to another?

Which proteins will be built given a specific strand of DNA?

How can DNA analysis predict genetic disorders and diseases?

How has a biological process or structure changed through evolution?

What diseases is a person especially vulnerable to, given their genetics?

All of these questions can be at least partially answered using an extensive analysis of scientific data with modern bioinformatics algorithms.

Drug Design based on Bioinformatics Tools

Drug design is the approach of finding drugs by design, based on their biological targets. Typically a drug target is a key molecule involved in a particular metabolic or signaling pathway that is specific to a disease condition or pathology, or to the infectivity or survival of a microbial pathogen. Some approaches attempt to stop the functioning of the pathway in the diseased state by causing a key molecule to stop functioning. Drugs may be designed that bind to the active region and inhibit this key molecule. However these drugs would also have to be designed in such a way as not to affect any other important molecules that may be similar in appearance to the key molecules. Sequence homologies are often used to identify such risks. Other approaches may be to enhance the normal pathway by promoting specific molecules in the normal pathways that may have been affected in the diseased state. The structure of the drug molecule that can specifically interact with the biomolecules can be modeled using computational tools. These tools can allow a drug molecule to be constructed within the biomolecule using knowledge of its structure and the nature of its active site. Construction of the drug molecule can be made inside out or outside in depending on whether the core or the R-groups are chosen first. However many of these approaches are plagued by the practical problems of chemical synthesis. Newer approaches have also suggested the use of drug molecules that are large and proteinaceous in nature rather than as small molecules. There have also been suggestions to make these using mRNA. Gene silencing may also have therapeutical applications.The processes of designing a new drug using bioinformatics tools have open a new area of research. However, computational techniques assist

one in searching drug target and in designing drug in silico, but it takes long time and money. In order to design a new drug one need to follow the following path.

Identify Target Disease:

One needs to know all about the disease and existing or traditional remedies. It is also important to look at very similar afflictions and their known treatments. Target identification alone is not sufficient in order to achieve a successful treatment of a disease. A real drug needs to be developed. This drug must influence the target protein in such a way that it does not interfere with normal metabolism. One way to achieve this is to block activity of the protein with a small molecule. Bioinformatics methods have been developed to virtually screen the target for compounds that bind and inhibit the protein. Another possibility is to find other proteins that regulate the activity of the target by binding and forming a complex.

Study Interesting Compounds:

One needs to identify and study the lead compounds that have some activity against a disease. These may be only marginally useful and may have severe side effects. These compounds provide a starting point for refinement of the chemical structures.

Detect the Molecular Bases for Disease:

If it is known that a drug must bind to a particular spot on a particular protein or nucleotide then a drug can be tailor made to bind at that site. This is often modeled computationally using any of several different techniques. Traditionally, the primary way of determining what compounds would be tested computationally was provided by the researchers' understanding of molecular interactions. A second method is the brute force testing of large numbers of compounds from a database of available structures.

Rational drug design techniques:

Unlike the historical method of drug discovery, by trial-and-error testing of chemical substances on cultured cells or animals, and matching the apparent effects to treatments, rational drug design begins with a knowledge of specific chemical responses in the body or target organism, and tailoring combinations of these to fit a treatment profile. Due to the complexity of the drug design process two terms of interest are still serendipity and bounded rationality. Those

challenges are caused by the large chemical space describing potential new drugs without side-effects.

A particular example of rational drug design involves the use of three-dimensional information about biomolecules obtained from such techniques as x-ray crystallography and NMR spectroscopy. This approach to drug discovery is sometimes referred to as structure-based drug design. The first unequivocal example of the application of structure-based drug design leading to an approved drug is the carbonic anhydrase inhibitor dorzolamide which was approved in 1995.

Another important case study in rational drug design is imatinib, a tyrosine kinase inhibitor designed specifically for the bcr-abl fusion protein that is characteristic for Philadelphia chromosome-positive leukemias (chronic myelogenous leukemia and occasionally acute lymphocytic leukemia). Imatinib is substantially different from previous drugs for cancer, as most agents of chemotherapy simply target rapidly dividing cells, not differentiating between cancer cells and other tissues.

The activity of a drug at its binding site is one part of the design. Another to take into account is the molecule's drug likeness, which summarizes the necessary physical properties for effective absorption. One way of estimating drug likeness is Lipinski's Rule of Five.

Refinement of compounds:

Once you got a number of lead compounds have been found, computational and laboratory techniques have been very successful in refining the molecular structures to give a greater drug activity and fewer side effects. This is done both in the laboratory and computationally by examining the molecular structures to determine which aspects are responsible for both the drug activity and the side effects.

Quantitative Structure Activity Relationships (QSAR):

Quantitative structure-activity relationship (QSAR) is the process by which chemical structure is quantitatively correlated with a well defined process, such as biological activity or chemical reactivity. For example, biological activity can be expressed quantitatively as in the concentration of a substance required to give a certain biological response. Additionally, when physiochemical properties or structures are expressed by numbers, one can form a mathematical relationship, or

quantitative structure-activity relationship, between the two. The mathematical expression can then be used to predict the biological response of other chemical structures.

QSAR's most general mathematical form is:

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

This computational technique should be used to detect the functional group in your compound in order to refine your drug. This can be done using QSAR that consists of computing every possible number that can describe a molecule then doing an enormous curve fit to find out which aspects of the molecule correlate well with the drug activity or side effect severity. This information can then be used to suggest new chemical modifications for synthesis and testing.

Solubility of Molecule:

One need to check whether the target molecule is water soluble or readily soluble in fatty tissue will affect what part of the body it becomes concentrated in. The ability to get a drug to the correct part of the body is an important factor in its potency. Ideally there is a continual exchange of information between the researchers doing QSAR studies, synthesis and testing. These techniques are frequently used and often very successful since they do not rely on knowing the biological basis of the disease which can be very difficult to determine.

Drug Testing:

Once a drug has been shown to be effective by an initial assay technique, much more testing must be done before it can be given to human patients. Animal testing is the primary type of testing at this stage. Eventually, the compounds, which are deemed suitable at this stage, are sent on to clinical trials. In the clinical trials, additional side effects may be found and human dosages are determined.

Motivation for the study:

AD (Alzheimer’s disease) is a progressive, degenerative brain disease. The neuropathology of Alzheimer’s disease (AD) is characterized by several features, including extracellular deposition

of amyloid β peptide (Aβ)-containing plaques in the cerebral cortical regions. Aβ (amyloid β) peptides, enzymatically derived from the proteolytic processing of the transmembrane APP (amyloid precursor protein), are the principal cause of neuronal dysfunction and death in the brains of AD patients. Currently, for treat Alzheimer's disease (AD) either acetyl cholinesterase or *N*-methyl-D-aspartate antagonists are available. Acetyl cholinesterase inhibitors have limited efficacy because they often have unpleasant side effects. Here we consider the viability of a single molecule having the actions of both an AchE inhibitor and histamine H_3 receptor antagonist. Both histamine H_3 receptor antagonists and AchE inhibitors synergist the cholinergic neurotransmission in cortex. We have selected histamine H_3 receptor antagonist that it raises acetylcholine levels mostly in brain as its mode of action will primarily be on the central nervous system. And combination of both activities in a single molecule will decrease the unpleasant side effect.

CHAPTER 2

LITERATURE SURVEY

Literature survey

ALZHEIMERS DISEASE

Alzheimer's disease is a progressive, irreversible brain disorder which attacks slowly and steals the mind of its victims. It is a neurodegenerative disease characterized by progressive cognitive deterioration together with declining activities of daily living and neuropsychiatric symptoms or behavioral changes. It is the most common type of dementia (decline in memory and mental abilities

HISTORY:

In 1901, Dr. Alois Alzheimer, a German psychiatrist, interviewed a patient named Auguste D, age 51 suffering from declining mental health. On November 3,1906, he presented Auguste D's case to the 37th Assembly of Southwest German Psychiatrists and described the neurofibrillary tangles and amyloid plaques that have come to be considered the hallmark of the disease. Eventually, the term Alzheimer's disease was adopted formally in the psychiatric and neurological nomenclature to describe individuals of all ages with the characteristic common symptom pattern, disease course and neuropathology.

STATISTICS:

In the United States of America, AD was the 7th leading cause of death in 2004, with 65,829 number of deaths (and rising).At over $100 billion per year, AD is the third most costly disease in the U.S., after heart disease and cancer. There are an estimated 24 million people with dementia worldwide. By 2040, it is projected that this figure will have increased to 81 million. More than 5 million Americans are estimated to have Alzheimer’s disease. It is Projected that 14.3 million Americans will have the disease by mid-century: a 350 percent increase from 2000.The federal government estimates spending approximately $647 million for Alzheimer’s disease research in fiscal year 2005.

PROGRESSION OF DISEASE AND SYMPTOMS:

These images represent a cross-section of the brain as seen from the front. The cross-section on the left represents a brain from a normal individual and the one on the right represents a brain with Alzheimer's disease.

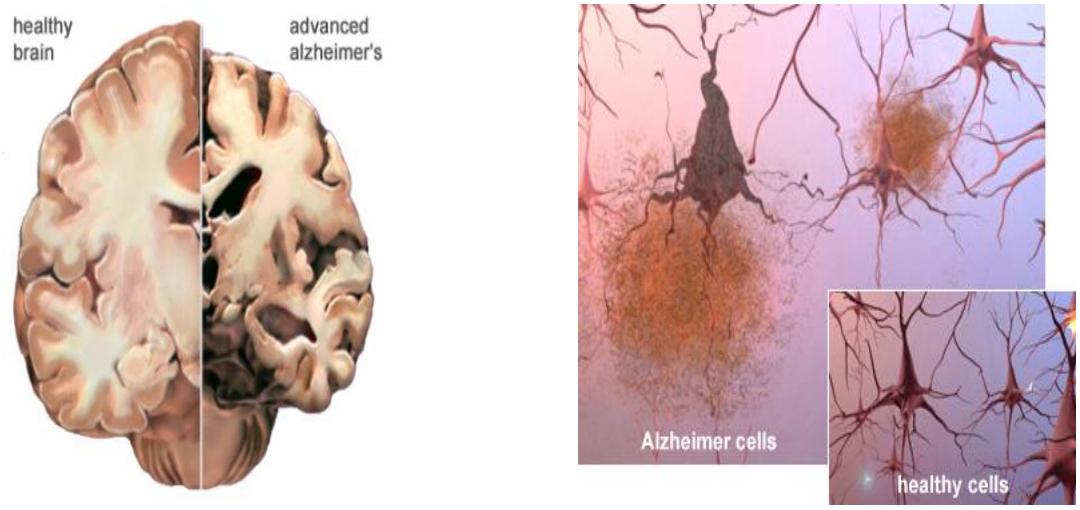


Fig 3 (a) comparative structure of both healthy brain and AD brain, and Fig (b) is the structure of nerve cell with tangles and without tangles.

In Alzheimer's disease, there is an overall shrinkage of brain tissue. The grooves or furrows in the brain, called sulci (plural of sulcus), are noticeably widened and there is shrinkage of the gyri (plural of gyrus), the well-developed folds of the brain's outer layer. In addition, the ventricles, or chambers within the brain that contain cerebrospinal fluid, are noticeably enlarged. In the early stages of Alzheimer's disease, short-term memory begins to decline (see box labeled ‘memory') when the cells in the hippocampus, which is part of the limbic system, degenerate. The ability to perform routine tasks also declines.

As Alzheimer's disease spreads through the cerebral cortex (the outer layer of the brain), judgment declines, emotional outbursts may occur and language is impaired. Progression of the disease leads to the death of more nerve cells and subsequent behavior changes,such as wandering and agitation. The ability to recognize faces and to communicate is completely lost in the final stages. Patients lose bowel and bladder control, and eventually need constant care. This stage of complete dependency may last for years before the patient dies. The average length of time from diagnosis to death is 4 to 8 years, although it can take 20 years or more for the disease to run its course.The most striking early symptom is loss of memory (amnesia), which usually manifests as minor forgetfulness that becomes steadily more pronounced with illness progression, with relative preservation of older memories. As the disorder progresses, cognitive (intellectual) impairment extends to the domains of language (aphasia), skilled movements (apraxia), recognition (agnosia), and those functions (such as decision-making and planning) closely related to the frontal and temporal lobes of the brain as they become disconnected from the limbic system, reflecting extension of the underlying pathological process. This pathological process consists principally of neuronal loss or atrophy, principally in the temporoparietal cortex, but also in the frontal cortex, together with an inflammatory response to the deposition of amyloid plaques and neurofibrillary Tangles One of the hallmarks of Alzheimer's disease is the accumulation of amyloid plaques between nerve cells (neurons) in the brain. Amyloid is a general term for protein fragments that the body produces normally. Beta-amyloid is a fragment of a protein that is snipped from another protein called amyloid precursor protein (APP). In a healthy brain, these protein fragments would be broken down and eliminated. In Alzheimer's disease, the fragments accumulate to form hard, insoluble plaques. Neurofibrillary tangles consist of insoluble twisted fibers that are found inside of the brain's cells. They primarily consist of a protein called tau**,** which forms part of a structure called a microtubule. The microtubule helps transport nutrients and other important substances from one part of the nerve cell to another (the axon is the long thread like extension that conducts nerve impulses away from the body of a nerve cell, and dendrites are any of the shor t branched thread like extensions that conduct nerve impulses towards the nerve cell body. In Alzheimer's disease the tau protein is abnormal and the microtubule structures collapse.

CAUSES/RISK FACTORS:

Three major competing hypotheses exist to explain the cause of the disease.

1)CHOINERGIC HYPOTHESIS: The oldest, on which most currently available drug therapies are based, is known as the "cholinergic hypothesis" and suggests that AD is due to reduced biosynthesis of the neurotransmitter acetylcholine. The medications that treat acetylcholine deficiency have served to only treat symptoms of the disease and have neither halted nor reversed it. The cholinergic hypothesis has not maintained widespread support in the face of this evidence, although cholingeric effects have been proposed to initiate large-scale aggregation leading to generalized neuroinflammation.

2)MISFOLDED AND AGGREGATED PROTEINS: Research after 2000 includes hypotheses centered on the effects of the misfolded and aggregated proteins, amyloid beta and tau. The two positions differ with one stating that the tau protein abnormalities initiate the disease cascade, while the other believes that beta amyloid deposits are the causative factor in the disease .[The tau hypothesis is supported by the long-standing observation that deposition of amyloid plaques do not correlate well with neuron loss;however, a majority of researchers support the alternative hypothesis that amyloid is the primary causative agent.The amyloid hypothesis is initially compelling because the gene for the amyloid beta Precursor APP is located on chromosome 21, and patients with trisomy 21 - better known as Down syndrome - who thus have an extra gene copy almost universally exhibit AD-like disorders by 40 years of age.The traditional formulation of the amyloid hypothesis points to the cytotoxicity of mature aggregated amyloid fibrils, which are believed to be the toxic form of the protein responsible for disrupting the cell's calcium ion homeostasis and thus inducing apoptosis.It should be noted further that ApoE4, the major genetic risk factor for AD, leads to excess amyloid build up in the brain before AD symptoms arise. Thus, beta-amyloid deposition precedes clinical AD.[Another strong support for the amyloid hypothesis, which looks at the beta-amyloid as the common initiating factor for the Alzheimer's disease, is that transgenic mice solely expressing a mutant human APP gene develop first diffuse and then fibrillar beta-amyloid plaques, associated with neuronal and microglial damage.

APP PROCESSING:

APP is a single-transmembrane protein with a 590-680 aa long extracellular amino terminal domain and an approximately 55aa cytoplasmic tail which contains intracellular trafficking signals (Fig 1). mRNA from the APP gene on chromosome 21 undergoes alternative splicing to yield eight possible isoforms, three of which (the 695, 751 and 770 amino acid isoforms) predominate in the brain (2,3). APP695 is the shortest of the three isoforms and is produced mainly in neurons. Alternatively, APP751, which contains a Kunitz-protease inhibitor (KPI) domain, and APP770, which contains both the KPI domain and an MRC-OX2 antigen domain, are found mostly in non-neuronal glial cells. All three isoforms share the same Ab, transmembrane and intracellular domains and are thus all potentially amyloidogenic. The normal function of APP is currently unknown, although in neurons it has been demonstrated to be localized in synapses where is may play a role in neurite extension or memory

Figure 1

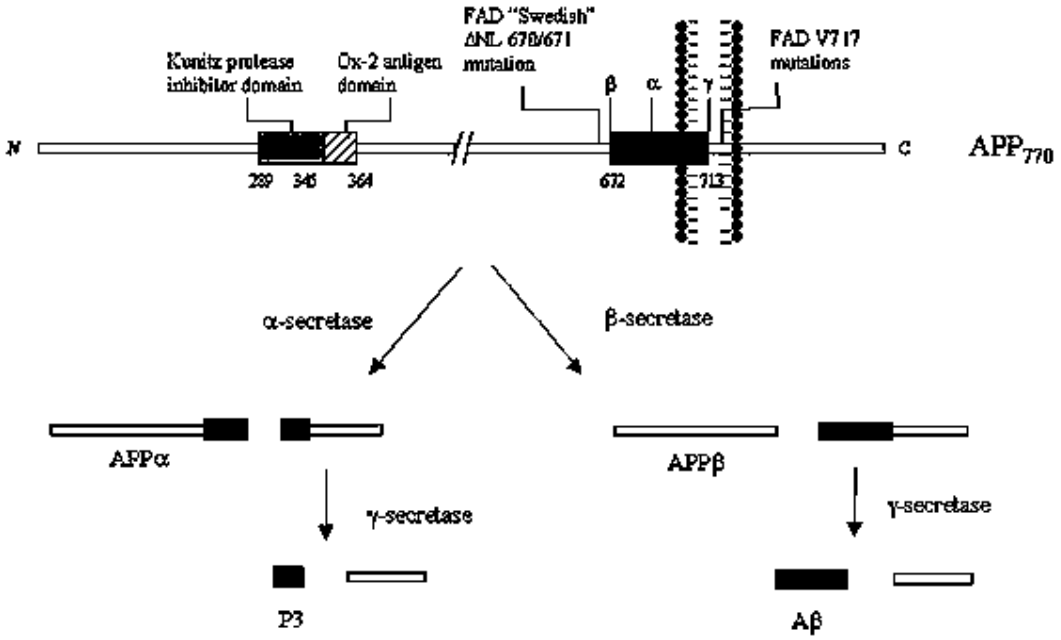


Fig 4. APP and its proteolytic products. Shown is APP770, the largest of the three

predominant isoforms found in the brain. Shorter isoforms are produced by alternative splicing of mRNA to remove the OX2 antigen domain (APP751) or both the OX2 and Kunitz protease inhibitor domains (APP695). APP can undergo proteolytic processing via 2 pathways. Cleavage by a-secretase occurs within the Ab domain and generates the large soluble N- terminal APPa and a non-amyloidogenic C - terminal fragment. A further proteolysis of this fragment by g-secretase generates the non-amyloidogenic peptide p3. Alternatively, cleavage of APP by b-secretase occurs at the beginning of the Ab domain and generates a shorter soluble N-terminus, APPb, as well as an amyloidogenic C-terminal fragment (C99). Further cleavage of this C-terminal fragment by g-secretase generates Ab. Cleavage by g-secretase or multiple g-secretases can result in C-terminal heterogeneity of Ab to generate Ab40 and Ab42.

3**) GENETIC LINKAGE**:

Rare cases of Alzheimer's are caused by dominant genes that run in families. The cases often have an early age of onset. Mutations in presenilin-1 or presenilin-2 genes have been documented in some families. Mutations of presenilin 1 (PS1) lead to the most aggressive form of familial Alzheimer's disease (FAD). Evidence from rodent studies suggests that the FAD mutation of PS1 results in impaired hippocampal-dependent learning which is correlated with reduced adult neurogenesis in the dentate gyrus.

Mutations in the APP gene on chromosome 21 can also cause early onset disease.ThePresenilins have been identified as essential components of the proteolytic processing machinery that produces beta amyloid peptides through cleavage of APP.Alzheimer's disease is definitely linked to the 1st, 14th, and 21st chromosomes, but other linkages are controversial and not yet confirmed. While some genes predisposing to AD have been identified , such as ApoE4 on chromosome 19, sporadic AD also involves other risk and protective genes still awaiting confirmation

RISK FACTORS:

Age : Alzheimer's usually affects people older than 65, but can, rarely, affect those younger than 40. Less than 5 percent of people between 65 and 74 have Alzheimer's. For people 85 and older, that number jumps to nearly 50 percent.

Heredity: Your risk of developing Alzheimer's appears to be slightly higher if a first- degree relative — parent, sister or brother — has the disease. Although the genetic mechanisms of Alzheimer's among families remain largely unexplained, researchers have identified a few genetic mutations that greatly increase risk in some familiesfew genetic mutations that greatly increase risk in some families.

Sex: Women are more likely than men are to develop the disease, in part because they live longer.

Heredity: The same factors that put you at risk of heart disease, such as high blood pressure and high cholesterol, may also increase the likelihood that you'll develop Alzheimer's disease . Poorly controlled diabetes is another risk factor. And keeping your body fit isn't your only concern — you've got to exercise your mind as well. Some studies have suggested that remaining mentally active throughout your life, especially in your years, reduces the risk of Alzheimer's disease.

Education levels: Studies have found an association between less education and the risk of Alzheimer's. Some researchers theorize that the more you use your brain, the more synapses you create , which provides a greater reserve as you age. It remains unclear, however, whether less education and less mental activity create a risk of Alzheimer's or if it's simply harder to detect Alzheimer's in people who exercise their minds frequently or who have more education.

Toxicity: One long-standing theory is that overexposure to certain trace metals or chemicals may cause Alzheimer's. For a time, aluminum seemed a likely candidate, because some people with Alzheimer's have deposits of aluminum in their brains. After many years of studies, however, no one has been able to link aluminum exposure directly to Alzheimer's. At

this point, there's no evidence that any particular substance increases a person's risk of Alzheimer's.

Head injury: The observation that some ex-boxers eventually develop dementia suggests that serious traumatic injury to the head (for example, a concussion with a prolonged loss of consciousness) may be a risk factor for Alzheimer's. Several studies indicate a definite link between the two, but others show no link.

Hormone replacement therapy:

The exact role hormone replacement therapy may play in the development of dementia isn't yet clear. Throughout the 1980s and '90s, evidence seemed to show that estrogen supplements given after menopause could reduce the risk of dementia. But results from the large-scale Women's Health Initiative Memory Study indicated an increased risk of dementia for women taking estrogen after age 65. The verdict is not yet in on whether estrogen affects the risk of dementia if given at an earlier age.

TESTS/DIAGONOSIS:

No single test can detect Alzheimer's. Instead, the disease is diagnosed by symptoms, findings on neurologic examination, and results from diagnostic tests. These tests help exclude other conditions that might cause the signs and symptoms. A diagnosis of Alzheimer's may be " probable ," meaning that other causes of the symptoms have been ruled out and the most likely cause is Alzheimer's disease. First, the patient will have a complete physical exam, along with a detailed history of symptoms and medical history, including medications. Examination by neurology specialists will help identify signs of Parkinson's disease , strokes , tumors and other medical conditions that ma impair memory and thinking, as well as physical function. Tests may include:

Mental status and neuropsychological assessments:

To determine which thinking and memory functions may be affected and to what degree, the patient will be asked questions to measure cognitive functions for attention, learning, recall, language and visuospatial abilities. The tests are compared to the tests of other patients of

similar age and education. The patient and people familiar with the patient will be interviewed about the patient's emotional state and day-to-day routines . They will also be asked about possible alcohol or drug abuse, head trauma and other causes for memory loss. Family members or close friends can provide valuable information about how the patient's behavior and personality have changed.

Psychiatric assessments:

In addition, the patient may have a psychiatric assessment to uncover possible depression or other mental illness.

Blood tests:

The patient's blood will be checked for infections or conditions such as vitamin deficiency, anemia, medication levels, disorders of the thyroid, kidneys or liver, and other factors that can cause memory loss.

Brain imaging:

Internal images of the brain help detect strokes, tumors or other conditions that may have affected the brain. Brain images can show changes to structures I in the brain that are associated with memory, such as the hippocampus. Various brain imaging techniques are

Computed tomography (CT scan):

In this test, an X-ray machine rapidly rotates around the brain while taking a series of thin X-ray beams that produce two-dimensional images.

Magnetic resonance imaging (MRI):

This test uses powerful magnets and radio waves to produce a detailed, three-dimensional view of the brain. Most patients are asked to undergo an MRI scan. Depending on the individual, another scan technique may be performed.

Positive emission tomography (PET) or Single-photon emission computerized tomography (SPECT):

These two fairly new techniques may be needed for clinical-related or research-related study. For both tests, a small amount of radioactive material is injected into the patient and emission detectors are placed on the brain. PET provides visual images of brain activity. SPECT is used to measure blood flow to various regions of the brain.

Other tests:

Other tests that sometimes provide important diagnostic information include the Electroencephalogram (EEG), electromyogram (EMG), urine tests , and tests on cerebrospinal fluid (CSF) obtained by a lumbar puncture. Physicians discuss with the patient and family which tests are most appropriate to establish the correct diagnosis

RISK REDUCERS:

Intellectual stimulation (e.g., playing chess or doing crosswords)Regular physical exercise Regular social interaction A Mediterranean diet with fruits and vegetables and low in saturated fat, supplemented in particular with: B vitamins Omega-3 fatty acids, especially Docosahexaenoic acid Fruit and vegetable juice High doses of the antioxidant Vitamin E (in combination with vitamin C) seem to reduce Alzheimer's risk in cross sectional studies but not in a randomized trial and so are not currently a recommended preventive measure because of observed increases in overall mortality Cholesterol-lowering drugs (statins) reduce Alzheimer's risk in observational studies but so far not in randomized controlled trials Female Hormone replacement therapy is no longer thought to prevent dementia based on data from the Women's Health Initiative Long - term usage of non-steroidal anti-inflammatory drugs used to reduce joint inflammation and pain, are associated with a reduced likelihood of developing AD, according to some observational studies. The risks appear to outweigh the drugs' benefit as a method of primary prevention

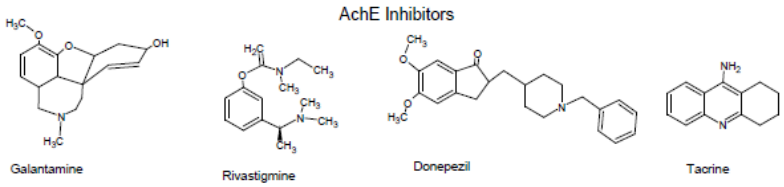
Alzheimer's disease Medications

Currently, there is no cure for Alzheimer's disease (AD). However, there are medications that can help control its symptoms. In addition, treatments are also available to help manage agitation, depression or psychotic symptoms (hallucinations or delusions) which may occur as the disease progresses.

1. **Acetyl-cholinesterase inhibitors (AchE inhibitors)**

Cholinesterase inhibitors slow the metabolic breakdown of acetylcholine, and make more of this chemical available for communication between cells. This helps in delaying the progression of cognitive impairment and can show efficacy for some patients in the early to middle stages of AD. All the four treatments are approved for mild to moderate symptoms of AD. In 2006, the FDA for the treatment of severe AD symptoms approved one drug Aricept®. The four FDA-approved cholinesterase inhibitors are:

- **Razadyne®** (galantamine)
- **Exelon®** (rivastigmine)
- **Aricept®** (donepezil)
- **Cognex®** (tacrine)

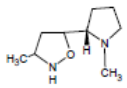


2. **Nicotinic acetylcholine receptors (nAChRs)**

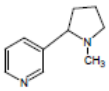
Neuronal nicotinic acetylcholine receptors (nAChRs) are found in the central and peripheral nervous systems that regulate number of diseases, including Alzheimer's discussed and investigated.(21)The nicotinic acetylcholine receptors (nAChRs) structurally it is pentameric and(22,23) Nicotine/acetylcholine binds to the receptor and depolarization occur, nAChR subtypes are found in different locations of the central and peripheral nervous system and the α4β2 and α4β4 subtypes appear to play a role in neurodegeneration(specially Alzheimer's

disease (AD and Attention Deficit Hyperactivity Disorder (ADHD).) Thus, the therapeutic potential of nAChR agonists, such as nicotine, in Alzheimer's, Parkinson's disease, Tourette's syndrome, and pain is being recognized Various analogues of nicotine(pyridine and pyrrolidine) were synthesized And one class agonists has entered clinical trials to treat the Alzheimer disease from the Abbott Laboratories(ABT 418) (.23)

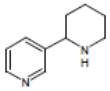
nAchRs agonists



ABT-418



Nicotine

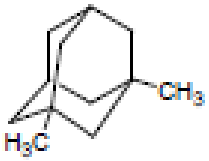


Anabasine

3. **N-methyl,D-aspartate inhibitors (NMDA)**

Persistent activation of central nervous system's NMDA receptors by the excitatory amino acid glutamate has been hypothesized to contribute to the symptomatology of AD. Thus inhibiting this receptor might improve symptoms in AD patients. Namenda® **(memantine)** was the first NMDA inhibitor to be approved by FDA .It protects the brain nerve cells by inhibiting the release of glutamate (neurotransmitter).Namenda binds to NMDA receptor and decrease the calcium to flow into the cell, which in turn decease the cell degeneration. Namenda causes minor side effect like dizziness, confusion, headache and constipation that are well-tolerated.

NMDA antagonist



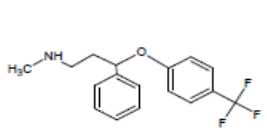
Mimantine

4. **Treatment of anxiety, depression & psychosis**

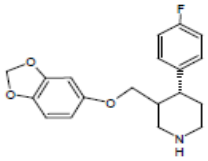
Often, in case of mild to severe Alzheimer's disease person often experiences depression, agitation, paranoid thoughts, delusions and hallucinations. The individual may be unable to communicate, be frustrated by his or her limitations, misunderstand what is happening or simply forget how to respond appropriately. Some common drugs currently in use are fluoxetine,

paroxetine, buspirone to treat depression and anxiety if the later condition becomes so severe that the patient becomes danger for him/herself.

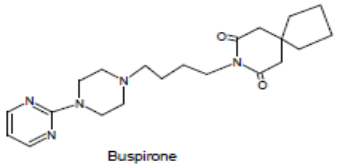
Anxiety,depression drugs



Fluoxetine



Paraxetine

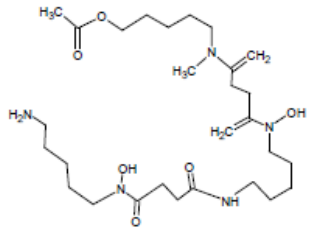


Buspirone

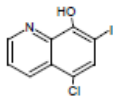
5. A-beta deposit antagonists (metal chelators)

Increasing evidence shows that several metal species such as aluminum, iron, zinc, and copper induce A-βeta aggregation and neurotoxicity in the AD brain by producing reactive oxygen species. *In vitro* studies have been done, and metal binding ligands have also been employed. Metal chelators like Desferrioxamine, MPAC and clioquinol, have shown efficacy in vitro cell and animal models of AD patients. AD patients have elevated levels of copper and zinc in the neocortex. The transition metals are particularly concentrated in neuritic plaques and potentiate A-βeta aggregation and neurotoxicity *in vitro*. Clioquinol chelates with copper and zinc in postemortem AD brains and solubilizes A-βeta. Thus, A-βeta accumulation in the brain may be significantly reduced by treatment with Clioquinol as a therapeutic agent. Crystal structure analysis confirms the coordination chemistry behind clioquinol's possible role as metal chelator.

Metal chelators



Desferrioxamine

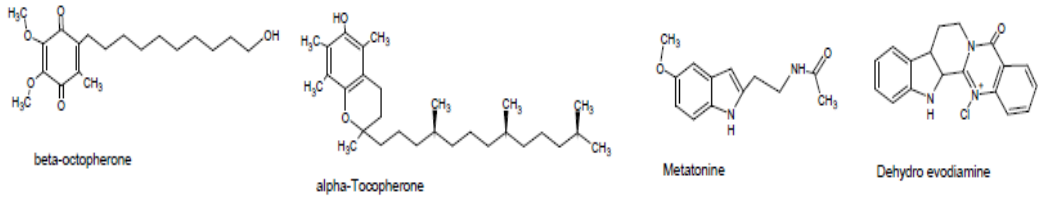


Clioquinol

6. **Antioxidants**

From *in-vitro* study suggests that due to metal binding generation of reactive oxygen species which is responsible for neuodegeneration. Thus, antioxidant plays important role in reduce oxidative injury and prove beneficial in retarding or preventing the onset and progression of AD in patients. An extract from Antioxidant like *Ginkgo* biloba **(Egb761)**, melatonin, idebenone, dehydroevodiamine hydrochloride, manganese porphyrin, salen and vitamin E**,** was examined to assess efficacy and safety in patients with AD showed improvement on the Alzheimer's Disease. Melatonin also has antiamyloidogenic activities. *Dehydroevodiamine Hydrochloride* (DHED) extracted from *Evodia rutaecarpa*. Results showed that DHED also protects neurons against glutamate and hydrogen peroxide. DHED decreases reactive oxygen species production and cell death suggesting that DHED might be useful in treatment of AD, vascular dementia and stroke. DHED is currently under clinical studies so that they can be in future used as a potential candidate for AD treatment.

Antioxidants



7. **Monoamine Oxidase Inhibitors (MAO-inhibitors)**

Selegiline

Research shows that when certain MAO inhibitors such as selegiline are used, they delayed the deterioration related to the Alzheimer's disease. The effect was also seen with vitamin E or a combination of both. These drugs, although, doesn't have significant effect on cognitive ability.

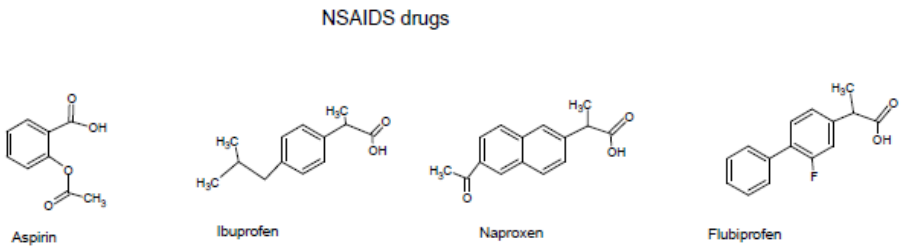
Rasagiline and TVP1022

N-propargyl-1(R)-aminoindan, rasagiline & its optical isomer TVP1022 are selective irreversible inhibitors for MAO. They are structurely very similar to selegiline. Both compounds have

similar neuroprotective activities with neuronal cell cultures, which is associated to the propargylamine functionality. However, rasagiline inhibits MAO-B to a much greater extent.

8. **Non-steroidal Anti-inflammatory Drugs (NSAIDs)**

Inflammation around around the Abeta plaques causes the destruction of neuron is thought to be a major factor in the pathogenesis of AD. NSAIDs decreases the inflammation by inhibiting the cyclooxygenase-1 and cyclooxygenase-2 (COX-1 and COX 2), which are responsible for the oxidation of arachidonic acid to prostaglandins. Individuals study shows that regular use of using conventional NSAIDs decreased incidence of AD and Act as neuroprotective..



Role of histamine-3 receptor in prospect of Alzheimer's

The histamine receptor are broadly classify in to four receptor subtypes H1, H2,H3, and H4, that mediate the diverse biological effects (Brown et al., 2001). Like all histamine receptors the H_3 receptor is a G-protein coupled receptor. H3 receptors are widely expressed in the mammalian brain, particularly in areas involved in cognitive processes and arousal, such as the cerebral cortex, hippocampus, basal ganglia, and hypothalamus (Martinez-Mir et al., 1990; Pollard et al.,1993). Activation of H3 autoreceptors results in the inhibition of histamine synthesis and heteroreceptors leads to the inhibition of release of other neurotransmitters such as acetylcholine, noradrenaline, dopamine, and 5-HT from nonhistaminergic neurons (Blandina et al., 1996; Schlicker and Kathmann, 1998; Brown et al., 2001).Conversely, blockade of H3 receptors with selective antagonists can increase the release of neurotransmitters involved in cognitive processes (Fox et al., 2005). Selective H3 antagonists have been shown to improve performance in a diverse range of rodent cognition paradigms, including object recognition. These observations provide the first evidence of a regulatory role of histamine H3 receptors on

cortical acetylcholine release in vivo. Moreover, they suggest a role for histamine in learning and memory and may have implications for the treatment of degenerative disorders associated with impaired cholinergic function.The presynaptic inhibitor-histamine H_3 receptor, when activated decreases the release of acetylcholine from cholinergic neurons. This has been shown both in the gastrointestinal and central nervous system. Consequently, when a histamine h3 receptor antagonist is present, an increase release of acetylcholine is observed.thus; both mechanisms contribute to the same neurochemical end result (increased synaptic levels of acetylcholine) in different ways. The histamine h3 receptor antagonists increase the amount of acetylcholine molecules entering the synaptic space, and AchE inhibitors entering the synaptic space , and AchE inhibitors prolong their survival time and hence the likelihood of their interacting with a postsynaptic cholinergic neurotransmitter and eliciting a biological effect.

In order to achieve a specific, cognition-enhancing level of cholinergic neurotransmission, a lower dose of an AchE inhibitor-histamine H3 receptor antagonist compound will likely be required than of either an AchE inhibitor or histamine H3 receptor antagonist alone, a combination molecule will have a higher potentcy for increasing acetylcholine levels than either an AchE inhibitor alone or histamine h3 receptor antagonist alone. It is likely reduced compound requirements for efficacy will then improve the side effect profile.

We describe herein the molecular modeling efforts that lead to the identification of an AchE inhibitor-histamine H3 receptor antagonist. The use of available crystal structure information of AchE receptor, homology model structure of H3 receptor based on bacterial rhodopsin crystal structure, pharmacophore modeling, QSAR, rigid docking and molecular dynamic simulation.

CHAPTER 3

TOOLS FOR STUDY

TOOLS FOR STUDY:

3.1Sequence alignments Tools

Sequence alignments provide a powerful way to compare novel sequences with previously characterized genes. Both functional and evolutionary information can be inferred from well designed queries and alignments. BLAST (Basic Local Alignment Search Tool), provides a method for rapid searching of nucleotide and protein databases. Since the BLAST algorithm detects local as well as global alignments, regions of similarity embedded in otherwise unrelated proteins can be detected. Both types of similarity may provide important clues to the function of uncharacterized proteins.

3.2. Homology modelling

3.2.1 On Line Homology Modelling Softwares

Swiss model (www.expasy.ch/swissmod/ SWISS-MODEL.html) is a fully automated protein structure Homology Modeling server. It has a first approach mode that helps performs Homology Modeling. The user has to enter his / her email id and input the protein sequence in Fasta format. It allows the user to choose the BLAST limit for template selection. It can search the pdb file from the pdb database with the user providing the name of the pdb file or the user can upload his / her own pdb file. The output file is a pdb file that is returned to the user's email address. The result can be forwarded by Swiss Model to PHD Secondary structure prediction at Columbia University and Fold Recognition Server (3D-pssm) of the ICRF. Swiss Model however does not accept the sequences for homology modelling when similarity is less than 25%[23].

Geno3D (http://geno3d-pbil.ibcp.fr) performs Comparative protein structure Modeling by spatial restraints (distances and dihedral) satisfaction. Geno3D is most frequently used for Homology or Comparative protein structure Modeling.Geno3d accepts input similar to Fasta format but only the one letter code has to be used. The result is obtained in the pdb format that can be viewed in any Molecular Modeling software.Geno3d offers many other features, it allows the user to select PDB entries as templates for Molecular Modeling after a 3 step iterative PSI BLAST. It presents the output for each template, along with the secondary structure prediction, displays percent of agreement in secondary structure and repartition of information from template on query

sequence. The output link is sent to the user's email address. It also notifies the user when it's server begins the Homology Modeling. It has an option where the user can decide how many models to generate. The main idea behind having more than one model generated is that the user may have a better flexibility and understanding. It also returns a superimposed pdb file which has the models superimposed on each other. This is one of the good points in Geno3d as it allows us to compare the various models generated in one window. All the results obtained can be downloaded as a archive.tar.Z that can be opened in WinZip in windows and in UNIX or Linux platforms. So the user does not have to save results in webpage effect or in a document file. It also displays the Ramachandran plot in the result[24].

CPHmodels Automated neural-network based protein modeling server (Http://www.cbs.dtu.dk/services/CPHmodels/). CPHmodels is a collection of databases and methods developed to predict protein structure. It performs prediction of protein structure using Comparative Modeling. It does not accept more than 900 amino acids in the input sequence. The sequences are kept confidential and are deleted after processing. This program did not give me appropriate results. The error it displayed was similar to the one displayed by Swiss Model.

3.2.2 Offline Homology Modelling Software:

MODELLER is used for homology or comparative modeling of protein three-dimensional structures. It is built in FORTRAN. It will runs on python script file commands. Modeller is most frequently used for homology or comparative protein structure modeling. Modeller helps determine the spatial restraints from the templates. It generates a number of 3D models of the sequence you submit satisfying the template restraints. MODELLER automatically calculate a full-atom model. MODELLER models protein 3D structure keeping in the constraints of spatial restraints. The restraints can be derived from a number of different sources. These include NMR experiments (NMR refinement),cross-linking experiments, fluorescence spectroscopy, rules of secondary structure packing (combinatorial modeling),image reconstruction in electron microscopy, homologous structures (comparative modeling),site-directed mutagenesis, residue-residue and atom-atom potentials of mean force, etc. Modeller is not an automated homology modelling tool.

It is a very specific program. Any error in the format of the sequence alignment prevents the modeller from performing Homology Modeling. The program is very specific about the extension names of the file formats used for Homology Modeling. It is a very reliable program and it allows the user to specify what he wants in the end result. Modeller runs on platforms like Win XP, Linux, Sun Solaris and Macintosh.

Deep View - Swiss-PdbViewer is an application that provides a user friendly interface allowing to analyze several proteins at the same time. The proteins can be superimposed in order to deduce structural alignments and compare their active sites or any other relevant parts. Amino acid mutations, H-bonds, angles and distances between atoms are easy to obtain thanks to the intuitive graphic and menu interface. DeepView - Swiss-PdbViewer has been developped by Nicolas Guex (GlaxoSmithKline R&D). Swiss-PdbViewer is tightly linked to SWISS-MODEL, an automated homology modeling server developed within the Swiss Institute of Bioinformatics (SIB) at the Structural Bioinformatics Group at the Biozentrum in Basel[27].

3.2.3 Structure Analysis and Verification Server

PROCHEK Checks the stereochemical quality of a protein structure by analyzing residue-by-residue geometry and overall structure geometry. It is tell about: Covalent geometry, Planarity, Dihedral angles, Chirality, Non-bonded interactions[28].

WHAT_CHEK derived from a subset of protein verification tools from the WHATIF program; this does extensive checking of many sterochemical parameters of the residues in the model[29].

DOPE: The DOPE model score is designed for selecting the best structure from a collection of models built by MODELLER. DOPE uses the standard MODELLER energy function.

ERRAT is a protein structure verification algorithm that is especially well-suited for evaluating the progress of crystallographic model building and refinement. The program works by analyzing the statistics of non-bonded interactions between different atom types. A **single** output plot is produced that gives the value of the error function *vs.* position of a 9-residue sliding window. By comparision with statistics from highly refined structures, the error values have been calibrated to give **confidence** limits. ERRAT will give an “overall quality factor” and if it is

a high 90% range protein structure is good. This is extremely useful in making decisions about reliability[30].

VERIFY_3D determines the compatibility of an atomic model (3D) with its own amino acid sequence (1D) by assigned a structural class based on its location and environment (alpha, beta, loop, polar, nonpolar etc) and comparing the results to good structures.Then a database generated from vetted good structures is used to obtain a score for each of the 20 amino acids in this structural class. For each residue, the scores of a sliding 21-residue window (from -10 to +10) are added and plotted[31].

PROVE Calculates the volumes of atoms in macromolecules using an algorithm which treats the atoms like hard spheres and calculates a statistical Z-score deviation for the model from highly resolved (2.0 Å or better) and refined (R-factor of 0.2 or better) PDB-deposited structures[32].

3.6 Docking Tools

Autodock 4.0

The introduction of AutoDock 4 comprises three major improvements:

1. The docking results are more accurate and reliable.
2. It can optionally model flexibility in the target macromolecule.
3. It enables AutoDock's use in evaluating protein-protein interactions.

AutoDock 4 offers many new features and improvements over previous versions. The most significant is that it models flexible side chains in the protein. We can get both the 3D structure and the inhibition constants.

AutoDock4 scoring functions are van der Waals forces, Hydrogen Bonding, Electrostatics, Desolvation, Torsional.

Binding energy=Intermolecular energy+ Torsional energy

$\Delta G_{bind} = \Delta G_{vdw} + \Delta G_{ele.} + \Delta G_{H\text{-}bond} + \Delta G$ desolv $+\Delta G_{tors}$

Here ΔG=change in free energy

The aim of this part is to re-dock the ligand present in the crystal structure of a protein using an automated docking suite called 'AutoDock'. The GUI for AutoDock is AutoDockTools (ADT), which was used to perform the entire docking task. More information is available on the AutoDock suite homepage http://www.scripps.edu/mb/olson/doc/autodock/.After preparing the protein and ligand files through chimera, all files must be transferred to the Autodock directory. Further modification to the protein and ligand, something like fixing the torsion residues ec are made and the files are saved in PDBQT format. The grid box is setted on the protein. After Autodock completes, the docking results are saved in a file named "dlg" in the directory. The conformation with the lowest docking energy is ranked best by AutoDock. One can see the clusters of docked conformations based on Binding energies by Opening and scrolling down the "dlg" file until we find 'CLUSTERING HISTOGRAM'. Make a note of the cluster rank, lowest docked energy, number of conf. in the cluster.AutoDock is used to perform computational molecular docking of small molecules to proteins, DNA, RNA and other important macromolecules, by treating the ligand and selected parts of the target as conformationally flexible. It uses a scoring function based on the AMBER force field, and estimates the free energy of binding of a ligand to its target. Novel hybrid global-local evolutionary algorithms are used to search the phase space of the ligand-macromolecule system.

3.7 GOLD

GOLD (Genetic Optimisation for Ligand Docking) is a genetic algorithm for docking flexible ligands into protein binding sites. GOLD provides all the functionality required for docking ligands into protein binding sites from prepared input files. GOLD offers a choice of scoring functions, GoldScore, ChemScore, Astex Statistical Potential (ASP) and User Defined Score which allows users to modify an existing function or implement their own scoring function.

The GOLD fitness function (Goldscore) is made up of four components:

o protein-ligand hydrogen bond energy (external H-bond)

o protein-ligand van der Waals (vdw) energy (external vdw)

o ligand internal vdw energy (internal vdw)

o ligand torsional strain energy (internal torsion)

oOptionally, a fifth component, ligand intramolecular hydrogen bond energy (internal H-bond), may be added. output files will contain a single internal energy term S(int) which is the sum of the internal torsion and internal vdw terms.

S(int)= internal torsion+ internal vdw

The fitness score is taken as the negative of the sum of the component energy terms, so that larger fitness scores are better. If any constraints have been specified, then an additional constraint scoring contribution S(con) will be made to the final fitness score. Similarly, when docking covalently bound ligands a covalent term S(cov) will be present.

The ChemScore function was trained by regression against measured affinity data. ChemScore estimates the total free energy change that occurs on ligand binding as:

?Gbind = ?G0 + ?GH-bond + ?Gmetal + ?Glipo +?Grot

The final ChemScore value is obtained by adding in a clash penalty and internal torsion terms, which militate against close contacts in docking and poor internal conformations. Covalent and constraint scores may also be included.

Chemscore = ?Gbinding + Pclash + cinternalPinternal + (ccovalentPcovalent +Pconstraint)

3.8 QSAR :

The Quantitative Structure Activity Relationship (QSAR) paradigm is based on the assumption that there is an underlying relationship between the molecular structure and biological activity. On this assumption QSAR attempts to establish a correlation between various molecular properties of a set of molecules with their experimentally known biological activity.

There are two main objectives for the development of QSAR:

1) Development of predictive and robust QSAR, with a specified chemical domain, for prediction of activity of untested molecules.

2) It acts as an informative tool by extracting significant patterns in descriptors related to the measured biological activity leading to understanding of mechanisms of given biological activity. This could help in suggesting design of novel molecules with improved activity profile.

QSAR's most general mathematical form is: Activity = f (physiochemical and/or structural properties)

Data Requirement and Handling: Biological Activity

For QSAR analysis, a dataset of a series of synthesized molecules tested for its desired biological activity is required. For a QSAR to bevalid and reliable, the activity of all of the chemicals covered must be elicited by a common mechanism. The quality of the model is totally dependent on the quality of the experimental data used for building the model.

Biological activity can be of two types:

1) Continuous Response: MEC, IC50, ED50, % inhibition

2) Categorical Response: Active/Inactive

In order to have confidence in QSAR analysis, biological data of at least 20 molecules is recommended:

1) Preferably tested in the same lab and by the same biological assay method.

2) With wide range and uniform distribution of the activity data.

3) Activity should well-defined in terms of either real number (continuous response, and cannot be e.g. >1000 or <1000) or in a particular class (categorical response).

3.8.1 Molecular Descriptors

Molecular descriptors can be defined as a numerical representation of chemical information encoded within a molecular structure via mathematical procedure. Type of QSAR is based on the dimensionality of molecular descriptors used:

1. 0D- These are descriptors derived from molecular formula e.g. molecular weight, number and type of atoms etc.
2. 1D- A substructure list representation of a molecule can be considered as a one-dimensional (1D) molecular representation and consists of a list of molecular fragments (e.g. functional groups, rings, bonds, substituent etc.).
3. 2D- A molecular graph contains topological or two dimensional (2D) information. It describes how the atoms are bonded in a molecule, both the type of bonding and the interaction of particular atoms (e.g. total path count, molecular connectivity indices etc.).
4. 3D- These are calculated starting from a geometrical or 3D representation of a molecule. These descriptors include molecular surface, molecular volume and other geometrical properties. There are different types of 3D descriptors e.g. electronic, steric, shape etc.

5. 4D-In addition to the 3D descriptors the 4th dimension is generally in terms of different conformations or any other experimental condition.

3.8.2 Selection of training and test set:

QSAR models are used increasingly to screen chemical databases and/or virtual chemical libraries for potentially bioactive molecules. These developments emphasize the importance of rigorous model validation to ensure that the models have both the ability to explain the variance in the biological activity (internal validation) and also the acceptable predictive power (external validation).For model validation the dataset is required to be divided into training set (for building the QSAR model) and test set (for examining its predictive ability). For any QSAR model, it is of crucial importance that the training set selected to calibrate the model exhibits a well balanced distribution and contains representative molecules.

Following are the methods for division of the dataset into training and test set:

1) **Manual Selection:** This is done by visualizing the variation in the chemical and biological space of the given dataset.

2) **Random Selection:** This method creates training and test set by random distribution.

3) **Sphere Exclusion Method:** This is a rational method for creation of training and test set. It ensures that the points in the both the sets are uniformly distributed w.r.t. chemical and biological space.

4) **Others:**

a) Experimental Design: full factorial, fractional factorial etc.

b) Onion Design

c) Cluster Analysis

d) Principal Component Analysis

e) Self Organizing Maps (SOM)

3.8.3 Variable selection methods:

There are a hundreds of molecular descriptors available for building a QSAR model. Not all of the molecular descriptors are important in determining the biological activity, and hence to find the optimal subset of the descriptors which plays an important role in determining activity, a

variable selection method is required. The variable selection method could be divided mainly into two categories:

1) Systematic variable selection: These methods add and/or delete a descriptor in steps one-by-one in a model.

a) Stepwise forward

b) Stepwise forward-backward

c) Stepwise backward

2) Stochastic variable selection: These methods are based on simulation of various physical or biological processes. These methods creates model starting from randomly generated model(s) and later modifying these model(s) by using different process operator(s) (e.g. perturbation, crossover etc.) to get better model(s).

a) Simulated Annealing

b) Genetic/Evolutionary Algorithms

c) Modified Particle Swarm Optimization

d) Artificial Ant Colony System

3.8.4 Statistical Methods:

• A suitable statistical method coupled with a variable selection method allows analyses of this data in order to establish a QSAR model with the subset of descriptors that are most statistically significant in determining the biological activity.

• The statistical methods can be broadly divided into two : linear and non-linear methods. In statistics a correlation is established between dependent variable(s) (biological activity) and independent variable(s) (molecular descriptors).The linear

method fits a line between the selected descriptors and activity as compared to non-linear method which fits a curve between

the selected descriptors and activity.

• The statistical method to build QSAR model is decided based on the type of biological activity data. Following are few commonly used statistical methods:

Categorical Dependent Variable

a) Discriminant analysis

b) Logistic regression

c) k-Nearest Neighbor classification

d) Decision Trees

e) SIMCA

Continuous Dependent Variable

a) Multiple Regression

b) Principal Component Regression

c) Continuum Regression

d) Partial Least Squares Regression

e) Canonical Correlation Analysis

f) k-Nearest Neighbor method

g) Neural Networks

Commonly used:

3.8.5 Interpretation of Model:

• Multiple regression is widely used method for building QSAR model. It is simple to interpret a regression model, in which contribution of each descriptor could be seen by the magnitude and sign of its regression coefficient.

• A descriptor coefficient magnitude shows its relative contribution w.r.t other descriptors and sign indicates whether it is directly (+) or inversely (-) proportional to the activity.

3.9 Molecular Dynamics Simulation

Although normally represented as static structures, molecules such as lysozyme are in fact dynamic. Most experimental properties, for example, measure a time average or an ensemble average over the range of possible configurations the molecule can adopt. One way to investigate the range of accessible configurations is to simulate the motions or dynamics of a molecule numerically. This can be done by computing a trajectory, a series of molecular configurations as a function of time, by the simultaneous integration of Newton's equations of motion

$$\frac{dr_i(t)}{dt} = v_i(t) \qquad \text{(eqn. 1)}$$

and

$$\frac{dv_i(t)}{dt} = \frac{F_i(t)}{m_i} \qquad \text{(eqn. 2)}$$

for all atoms (i = 1, 2 ,...,N) of the molecular system. The atomic coordinates, r, and the velocity, v, of atom, i, with mass, mi, thus become functions of time. The force Fi exerted on atom i by the other atoms in the system is given by the negative gradient of the potential energy function V which in turn depends on the coordinates of all N atoms in the system:

$$F_i(t) = -\frac{\delta V(r_1(t), r_2(t), \ldots, r_N(t))}{\delta r_i(t)} \qquad \text{(eqn. 3)}$$

For small time steps δt, eqn. (2) can be approximated by

$$V_i(t+\Delta t/2) = v_i(t-\Delta t/2) + \frac{F_i(t) \cdot \Delta t}{m_i} \qquad \text{(eqn. 4)}$$

and eqn. (1) likewise by

$$r_i(t+\Delta t) = r_i(t) + v_i(t+\Delta t/2)\Delta t \qquad \text{(eqn. 5)}$$

Thus a 100 ps (10^{-10} seconds) molecular dynamics simulation involves 10^5 to 10^4 integration steps. Even using the fastest computers only very rapid molecular processes can be simulated at an atomic level. As with any aspect of modelling, the accuracy of the predicted dynamics will depend on the validity of the underlying assumptions of the model. In this case the model is essentially defined by the force field that is used. For this exercise we will be using the GROMOS96 empirical force field.

This technique is commonly referred to as Molecular Dynamics (MD). A detailed explanation of the concepts behind this numerical technique can be found at http://www.fisica.uniud.it/~ercolessi/md/md.

Factors that govern the outcome of MD simulations are:

- choice of the degrees of freedom
- force field parameters
- treatment of non-bonded interactions
- salvation effects
- boundary conditions
- treatment of temperature and pressure
- integration time step
- starting configuration

As with any aspect of modeling, the accuracy of the predicted dynamics will depend on the validity of the underlying assumptions of the model. In this case this is essentially defined by the model for the intermolecular interactions (or potential energy) used. That model is a mathematical function (force field) that describes how the value for the potential energy depends on the spatial arrangement of all the atoms.

3.9.1 Gromacs, the MD Package

The following is designed to acquaint you with the general features of the molecular dynamics software package Gromacs. Gromacs is a widely used molecular dynamics simulation package developed at the University of Groningen. Information on Gromacs can be found at http://www.gromacs.org/.

To run a simulation several things are needed:

- a file containing the coordinates for all atoms
- information on the interactions (bond angles, charges, Van der Waals)
- Parameters to control the simulation.

The .pdb or .gro file contains the coordinates for all atoms and is the input structure file for MD simulation. The interactions are listed in the topology (.top) file and the input parameters are put into a .mdp file..

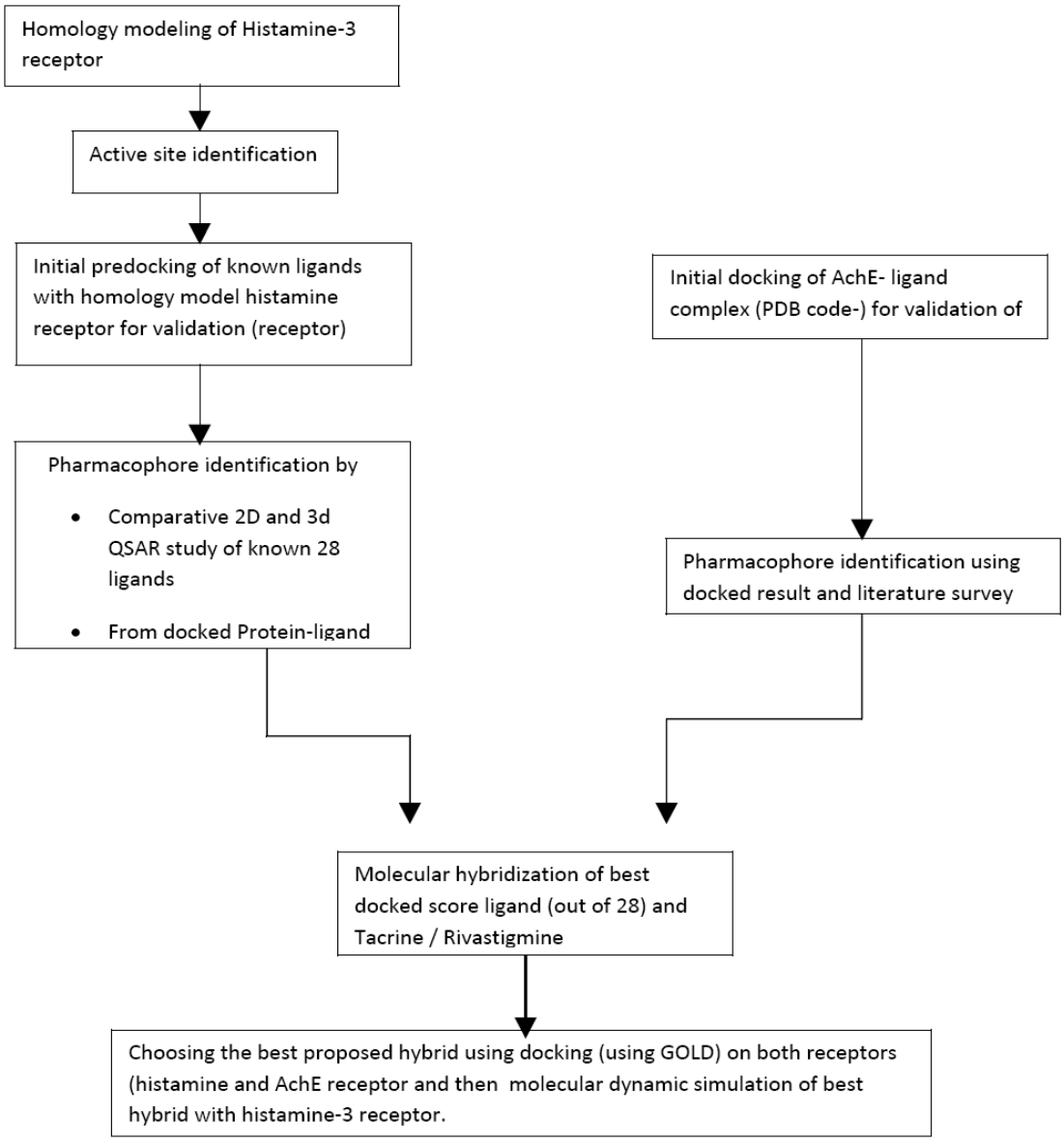
The actual steps in an MD simulation.

- Conversion of the pdb structure file to a Gromacs structure file, with the simultaneous generation of a descriptive topology file.
- Energy minimization of the structure to release strain.
- Running full simulations.
- Analyzing results.

CHAPTER 4

MATERIALS AND METHODS

4. Materials and Methods:



Scheme 1**:** flowchart for designing the dual acting hybrid molecule.

4.1 Sequence alignment and Homology Modeling:

The transmembrane portion of the H_3 receptor was build by homology modeling techniques based on the 2.8 angstrom resolution crystal structure of Bovine Rhodopsin (pdb 1F88) which is the most accurate rhodopsin structure available. The primary sequence of the H_3 receptor was aligned with bovine rhodopsin, the CLUSTALW program was used from its web site at http://www2.ebi.ac.uk/CLUSTALW based upon highly conserved amino acid residues in the seven helices. A model of the human histamine H3 receptor was generated based on the crystal structure 1HZX of bovine rhodopsin.The initial sequence-structure alignment was based on multiple sequence alignments, the prediction of secondary structure, transmembrane helices and highly conserved residues identified.

All homology models were constructed with the SWISS-MODEL server (http://expasy.org/swissmod/SWISSMODEL.HTML). And amino acid side chain conformations were added using program SCWRL3.0. . The model was validated for the stereochemical qualities were checked with (http://nihserver.mbi.ucla.edu/SAVS/). Finally, the structural properties of the target protein were validated by using the Ramachandran plot score.

Molecular Redocking studies H3 and AchE receptors with known active molecules:

Docking study on H_3 receptor

Preparation of Ligands and Protein:

Since ligands are not peptides, Gasteiger charge was assigned and then non-polar hydrogen's were merged. The rigid roots were defined automatically rather manually for each compound considered. The amide bonds were made non-rotatable. The homology model structure of was used for docking study The Kollman charges were added to each atoms of the modeled protein.

Grid Generation:

The Grid box was centered on the Asp 114 of the human histamine-3 receptor. The binding site includes the catalytic center (Asp 114 and Glu 206) and several subsites as (Ser-79,Lec-82,Val-83,Gly-84,Phe-86,Cys-87,Ile-88,Pro-89,Leu-90,Tyr-91,Trp-100,Leu-106,Cys-107,Lys-108,Leu-109,Val-112,Val-113,Asp-114,Tyr-115,Leu-116,Leu-117,Cys-118,Thr-119,Ser-120,Trp-

160,Tyr-167,Gly-168,Ile-171,Glu-175,Phe-192,Phe-198,Glu-206,Trp-371,Tyr-374,Met-378,Tyr-394,Phe-398,Leu-401,Ser-405).The spacing between the Grid points was 0.375 angstroms.

Docking:

4.1.2 With autodock4

Algorithm- Lamarckian genetic algorithm,Population size-150,Number of operations-250000,Rate of Gene mutation -0.02,Rate of Crossover 0.8

The parameters were set using the software Autodock Tools available at (http://mgltools.scripps.edu/downloads) which is made to associate with Autodock 4.0 Binaries downloaded from (http://autodock.scripps.edu/downloads/autodock-registration). The Calculations of Autogrid and Autodock were performed on Linux operating system having system Properties (Intel(R) Pentium(R) D CPU 2.80GHz, 2.0 GB of RAM).

4.1.3 With GOLD

Algorithm- Lamarckian genetic algorithm, Population size-100, Selection pressure-1.1, Number of operations-100000, Number of islands-5, Niche size-2, Crossover frequency-95, Mutation frequency-95, Migration frequency-10.

The Calculation was performed on windows operating system having system Properties (Intel(R) Pentium(R) D CPU 2.80GHz, 2.0 GB of RAM).

4.2 QSAR study of Histamine-3 receptor

4.2.1 Dataset:

Dataset of 28 non-peptide Inhibitor molecules collected from the literature [30] , was considered in this study (figure 5). All the molecules studied had the same parent skeleton.). IC50 is the concentration of the compound leading to 50% inhibitory effect. The logarithm transformation of this parameter has been used as biological end points (log IC50) in the QSAR studies so as to move the data to a nearly normal distribution.

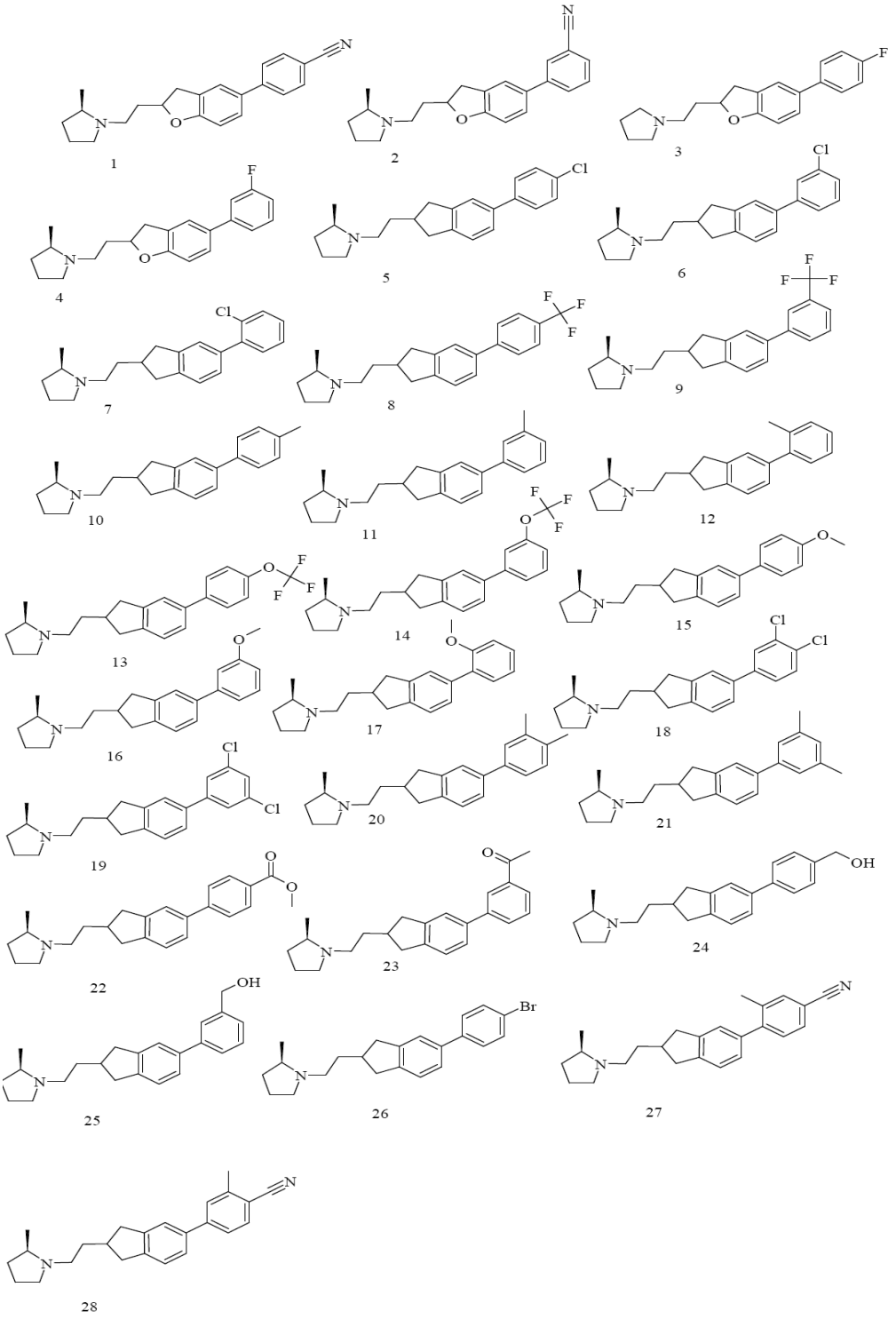


Fig 6: The chemical structures of 28 Inhibitors(data taken from Gfesser et al[49])

4.2.3 QSAR studies:

2d QSAR STUDY

Generation of molecular descriptors:

The selected 28 molecules were built and energy minimized using PRODRG software [25] and saved as the mol2 (output) files. The output files were loaded into **Vlife MDS QSAR** module for evaluation of several molecular descriptors along with the facility to build the QSAR equation and use it for predicting the activity of test/new molecules. These features are managed through an MS-Excel type worksheet. It can calculate all physiochemical descriptors such as Individual, Chi, Chiv, Path count, Chi Chain, Chiv chain, chain path count, Cluster, Path cluster, Kappa, Element count, Estate numbers, Estate contributions, Information theory index. Deselect Dipole Moment, Electro Static, Distance Based Topological Indices, Semi Empirical and Hydrophobicity base logP descriptors (as these are 3D descriptors) by the pointer

Set Data to complete the selection of training and test set

The worksheet shows the divisions of training data set (23molecules) and test data set (5 molecules).

Selection of data based on

1. The max of the test should be less than max of train set
2. The min of the test should be greater than min of train set

Regression methods used:

Multiple regressions

Multiple regressions are the standard method for multivariate data analysis it is also called as ordinary least squares regression (OLS). This method of regression estimates the values of the regression coefficients by applying least squares curve fitting method

Regression method- Multiple

Selected variable selection method- Forward-backward, parameter setting- Cross correlation limits-1, Number variable in final equation-4, Term selection criteria-r^2, Ftest-in-4.0.

Regression method- Multiple

Selected variable selection method- Genetic algorithm, Cross correlation limit-1.0, Population-10

Convergence criteria-0.01, Crossover probability-0.9, Number of generations-1000, Convergence length-3, Term selection criteria-r^2, Seed-0

Regression method- Multiple

Selected variable selection method- Simulated annealing, Maximum temperature-100, Minimum temperature-0.01, Decrease tempratureby-10.0, Iteration at given temperature-5, Terms in model-4, Perturbation limits-1.0, Seed-0, Term selection criteria-r^2

Partial least squares regression method

Partial least squares regression is an extension of the multiple linear regression model. In its simplest form, a linear model specifies the relationship between a dependent variable Y,and a set of predictor variable, the X's, so that

$Y=b0 + b1X1 +b2X2 +\ldots+bpXp$

In this equation b0 is the regression coefficient for the intercept and the bi values are the regression coefficients (for variables 1 through p) computed from the data.\

Regression method- PLS

Selected variable selection method- Forward-backward, parameter setting- Cross correlation limits-1, Number variable in final equation-4, Term selection criteria-r^2, Ftest-in-4.0.

Regression method- PLS

Selected variable selection method- Genetic algorithm, Cross correlation limit-1.0, Population-10 Convergence criteria-0.01, Crossover probability-0.9, Number of generations-1000, Convergence length-3, Term selection criteria-r^2, Seed-0.

Regression method- PLS

Selected variable selection method- Simulated annealing, Maximum temperature-100, Minimum temperature-0.01, Decrease tempratureby-10.0, Iteration at given temperature-5, Terms in model-4, Perturbation limits-1.0, Seed-0, Term selection criteria-r^2

Principle component regression (PCR) method

Principle components analysis provides a method for finding structure in such data sets .put simply, rotates the data into a new set of axes, such that the first few axes reflect most of the variations within the data. By plotting the data on these axes, we can spot major underlying

structures automatically. The value of each point, when rotated to a given axis, is called the principle component values.

Regression method- PCR

Selected variable selection method- Forward-backward, parameter setting- Cross correlation limits-1, Number variable in final equation-4, Term selection criteria-r^2, Ftest-in-4.0.

Regression method- PCR

Selected variable selection method- Genetic algorithm, Cross correlation limit-1.0, Population-10 Convergence criteria-0.01, Crossover probability-0.9,Number of generations-1000, Convergence length-3, Term selection criteria-r^2, Seed-0.

Regression method- PCR

Selected variable selection method- Simulated annealing, Maximum temperature-100, Minimum temperature-0.01, Decrease tempratureby-10.0, Iteration at given temperature-5, Terms in model-4, Perturbation limits-1.0, Seed-0, Term selection criteria-r^2

4.3.3 3D QSAR study

k- nearest neighbor QSAR

In k-nearest neighbor algorithm, for classifying a new pattern (molecule),

The system finds the k nearest neighbors among the training set, and used the categories of the k-nearest neighbors to weight the category candidates. The nearness is measured by an appropriate distance metric (e.g. a molecular similarity measure, calculated using descriptors of molecular structures).

Method selected- k-nearest neighbor.

Selected variable selection method- Forward-backward

Parameter setting- Cross correlation limits-1, Number variable in final equation-4, Term selection criteria-q^2, Ftest-in-4.0

Method selected- k-nearest neighbor.

Selected variable selection method- Simulated annealing.

Parameter setting- Maximum temperature-100, Minimum temperature-0.01, Decrease tempratureby-10.0, Iteration at given temperature-5, Terms in model-4, Perturbation limits-1.0

Seed-0, Term selection criteria-r^2

4.3.4 Docking study on AchE receptor

Preparation of Ligands and Protein:

Since ligand (tacrin, rivastigmine) are not peptides, Gasteiger charge was assigned and then non-polar hydrogen's were merged. The rigid roots were defined automatically rather manually for each compound considered. The amide bonds were made non-rotatable. The PDB structure-1EVE of was used for docking study The Kollman charges were added to each atoms of the given protein.

Grid Generation:

The Grid box was centered on the Ser 199, His 439 and Glu 326 of the AchE. The binding site includes several subsites as (TYR70, TRP84, GLY118, TYR 121, TYR130, GLU199, SER200, TRP279, LEU282, ILE287, PHE288, PHE290, GLU327, PHE331, TYR334, HIS440). The spacing between the Grid points was 0.375 angstroms.

4.4 Molecular Dynamics Simulation Studies (best hybrid molecule)

Gromacs is a very powerful molecular simulation package. The best hybrid molecule (results obtained from both GOLD and Autodock 4.0) was further analyzed for the stability of its interaction with the protein (histamine -3 receptor) using molecular dynamics simulation studies in Gromacs. The ligand-receptor complexes resulting from the docking calculation were placed in the box of water.

Preparation of Ligand and Protein:

A GMX topology file was prepared for the best docked molecule in Dundee PRODRG server (http://davapc1.bioch.dundee.ac.uk/programs/prodrg/). After putting drg.pdb coordinates into the empty text box on the webpage, Check the following options*:(1).Chirality-Yes (2).Full charges-Yes (3).Energy Minimization- No Click "Run PRODRG".*

DRGGMX.ITP file is used for building the topology for the drug. In addition, DRGFIN.GRO file is needed for building the coordinate file (*.GRO), or DRGPOH.PDB file can be used. The DRGGMX.ITP file was renamed to drg.itp. Our protein PDB file was too crude to use with pdb2gmx. So it was minimized in Chimera for 100 steps.

Simulation:

The topology file for the protein molecule was generated. The dodecahedron water box was set having diameter 0.65.The dodecahedron water box saves around 30% of computational time in comparison to cubic box. The system was found to have a non-zero total charge. So it was neutralized by adding Na+ ions in the water box. Na+ ions were added by simply replacing the water molecules in the water box.

The energy minimization of protein molecule was done. "Steep" algorithm was used for energy minimization and no constraints were set during the process. The other parameters set for minimization as: emtol 2000, emstep 0.01, nstcomm 1, ns_type grid, rlist 1, coulomb type PME, rcoulomb 1.0, rvdw 1.4. No temperature coupling, pressure coupling was set and no velocity was generated during the process. The system was minimized to Fmax = 2000 in 1500 steps and the converged result came as:

Steepest Descents converged to Fmax < 2000 in 832 steps

Potential Energy = -1.5722345e+06

Maximum force = 1.6236754e+03 on atom 6800

Norm of force = 1.3176494e+04

A position restrained dynamics simulation was run to "soak" the water and the drug into the drug-enzyme complex. In this run, the atom positions of the protein are restrained to restrict their movement in the simulation (i.e. the atom positions are restrained not fixed!). The water and the drug are permitted to relax about the protein. The relaxation time of water is 10 ps. Therefore, a total of 30 ps dynamics run was used to perform the soak under coulomb type, PME which stands for "Particle Mesh Ewald" electrostatics. PME is the best method for computing long range electrostatics (gives more reliable energy estimates). [6, 7] The all bonds option under

constraints applies the Linear Constraint algorithm [8] for fixing all bond lengths (important to use this option when dt > 0.001 ps). Berenson's temperature and pressure coupling methods were used [9]. The reference temperature was set at 300 K and the velocity was also generated.

The molecular dynamics simulation parameters were set for the Gromacs 87 force field, which we are using. For this run, we use the energygrps parameter to establish the groups for the energy output (the md.edr file). This will be important for use in for example linear interaction energy computations later on.

The parameters which were set for the final md simulation are as:

Nsteps-= 5000000(number of steps (1ns), Coulomb type-= PME, Vdwtype= Cut-off, Tcoupl= berendsen, Pcoupl = berendsen, pcoupltype = isotropic, constraints= all-bonds, constraint-algorithm = shake

CHAPTER 5

RESULTS

5. RESULTS

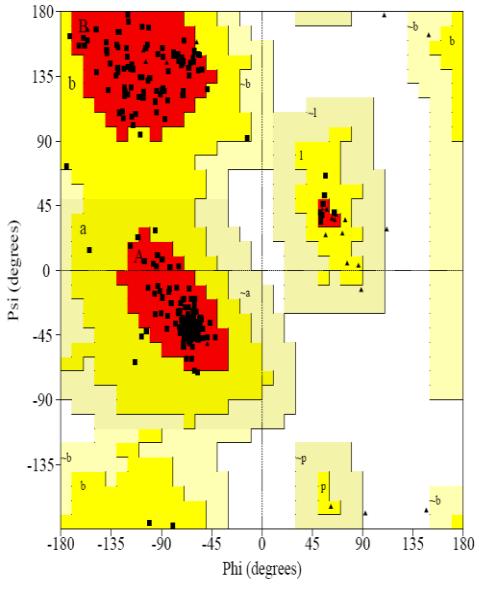
5.1 Sequence alignments and modeling H_3 receptor

Sequence alignment was done with the help of CLUSTALW software using default parameters. Because of an overall 60% sequence identity between the template and the target, as illustrated in Fig. 1, the generation of homology model of the human histamine h3 receptor was done using online Swiss model server (www.expasy.ch/swissmod/ SWISS-MODEL.html) based on the backbone coordinates of the crystal structure 1HZX of bovine rhodopsin. The missing side chain was added using the program SCWRL3.0.

5.2 Structure Analysis and Verification

PROCHEK was used to check the stereochemical quality of the protein structure generated, by analyzing residue-by-residue geometry and overall structural geometry. It tells about: Covalent geometry, Planarity, Dihedral angles, Chirality, Non-bonded interactions. The Ramchandran plot (Fig.7) was generated by PROCHEK of the protein model.

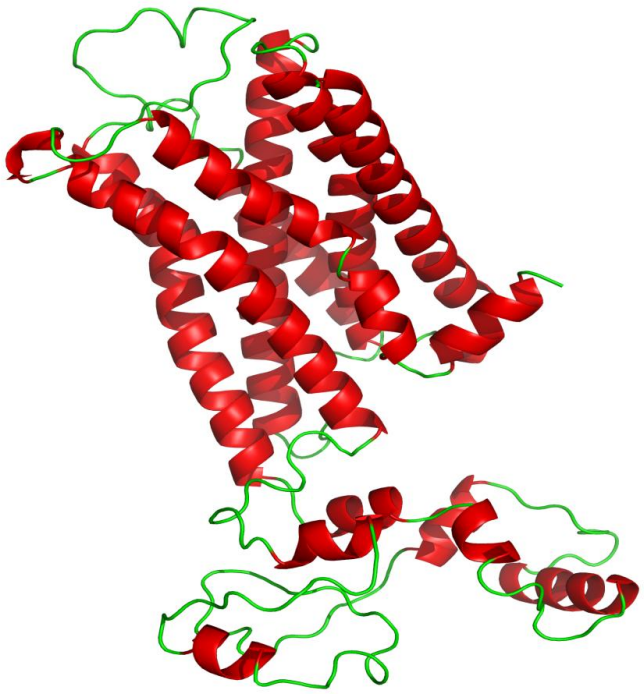
Fig 7 Ramchandran plot of human histamine-3 receptor protein from PROCHECK.



Plot statistics

Residues in most favored regions	[A,B,L]	306 90.3%
Residues in additional allowed regions	[a,b,l,p]	24 7.1%
Residues in generously allowed regions	[~a,~b,~l,~p]	6 1.8%
Residues in disallowed regions		3 0.9%
Number of non-glycine and non-proline residues		339 100.0%
Number of end-residues (excl. Gly and Pro)		2
Number of glycine residues (shown as triangles)		32
Number of proline residues		23
Total number of residues		396

Residue in disallow region are HIS417, THR-149, SER-330



.

Fig 8*:* Homology model structure of histamine-3 receptor protein (cartoon view)

5.3 Redocking study

All 28 molecules were docked into the homology modeled protein of histamine receptor. The docking results were ranked according to the decreasing docking energies of the 100 conformers for each of the Ligands. It was found that most of the ligands with lowest docking energies interacted quite well with the receptor in the pocket. The molecules 11, 18, 23, 25 and 26 had the dock score of less than -8.00 kcal/mol. The docking study was done using Autodock4 software. The AchE receptor's pdb structure 1EVE and 1GQR were considered for docking along with the ligand molecules Rivastigmine and Tacrine for docking studies. Out of the two PDB structures, 1EVE showed good docking score that was less than -11.30 kcal/mol. The ligand Rivastigmine interacted well with the receptor binding pocket. Therefore, for our further studies we considered the PDB structure 1EVE for further investigations.

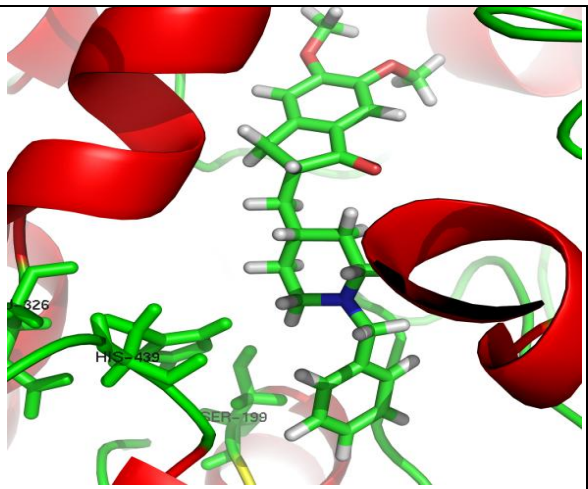


Figure1.(a)

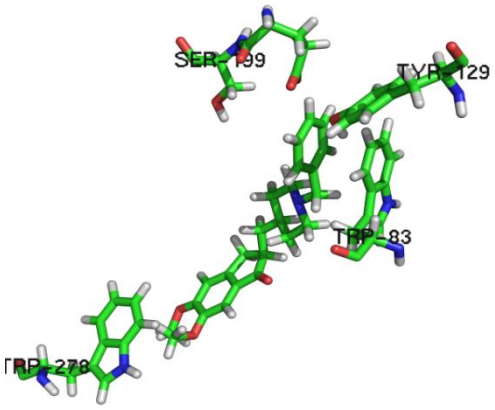


Figure1.(b)

Fig 9*(a)* Rivastigmine within the binding pocket.figure.1 (b) Residue interacting with the AchE receptor

Table: (1): Binding free energies for the 28 molecules (Unmodified compounds) of both H_3

S.No	Compounds	Estimated Free Energy of Binding (kcal/mole)	Estimated inhibition (µm)	Final intermolecular energy(Vdr +H-bond +Dissolved energy) in kcal/mole	Final total internal energy (kcal/mole)	Torsional free energy (kcal/mole)	Unbound system energy (kcal/mole)
1	Compound1	-7.84	1.75	-8.45	+0.12	+1.10	-0.62
2	Compound2	-7.71	2.23	-8.62	-0.61	+1.10	-0.42
3	Compound3	-7.31	4.41	-8.16	-0.65	+1.10	-0.41
4	Compound4	-6.37	3.83	-8.22	-0.67	+1.10	-0.41
5	Compound5	-7.85	1.75	-8.68	-0.69	+1.10	-0.42
6	Compound6	-7.08	6.48	-7.96	-0.68	+1.10	-0.44

7	Compound7	-6.20	28.59	-6.91	0.73	+1.10	-0.35
8	Compound8	-7.18	5.41	-8.94	-0.59	+1.92	-0.43
9	Compound9	-5.64	72.90	-7.10	-0.89	+1.92	-0.43
10	Compound10	-6.88	9.03	-8.09	-0.80	+1.37	-0.64
11	Compound11	-8.15	1.07	-9.04	-0.64	+1.10	-0.44
12	Compound12	-6.17	30.26	-7.15	-0.50	+1.10	-0.38
13	Compound13	-6.14	31.36	-7.95	-0.85	+2.20	-0.46
14	Compound14	-7.14	5.86	-9.17	0.64	+2.20	-047
15	Compound15	-7.57	2.81	-8.68	-0.72	+1.37	-0.46
16	Compound16	-7.14	3.71	-8.82	-0.43	+1.37	-0.47
17	Compound17	-7.57	2.84	-8.85	-0.63	+1.37	-0.54
18	Compound18	-8.54	0.547	-9.45	-0.61	+1.65	-0.43
19	Compound19	-7.85	1.70	-8.48	-0.12	+165	-0.43
20	Compound20	-6.88	8.99	-7.91	-0.53	+1.10	-0.45
21	Compound21	-7.06	6.68	-7.87	-0.74	+1.10	-0.45
22	Compound22	-7.60	2.07	-8.97	-0.80	+1.65	-0.53
23	Compound23	-8.43	0.66	-9.60	-0.72	+1.37	-0.52
24	Compound24	-7.45	3.46	-8.96	-0.57	+1.65	-0.43
25	Compound25	-8.01	1.34	-9.48	--0.61	+1.65	-0.43
26	Compound26	-8.00	1.37	-8.88	-0.65	+1.10	-0.43

27	Compound27	-7.27	4.71	-8.28	-0.44	+1.10	-0.36
28	Compound28	-6.44	19.01	-7.08	-0.84	+1.10	-0.38

Table: (2): Binding free energies for the pharmaceutical compound Rivastigmine with PDB structures (1EVE, 1GQR)

S.No	Compound Name	Estimated Free Energy of Binding (kcal/mole)	Estimated inhibition	Final intermolecular energy(Vdr +H-bond +Dissolved energy) in kcal/mole	Final total internal energy (kcal/mole)	Torsional free energy (kcal/mole)	Unbound system energy (kcal/mole)
1	1EVE-RIVASTGMINE	-11.35	4.75nm	-12.98	-0.70	+1.65	-0.68
2	1GQR-RIVASTGMINE	-9.63	86.82nm	-10.18	-0.20	-1.37	-0.01

Table: (3) Comparison of the MR,SA-MR,GA-MR,PLS,SA-PLS,GA-PLS,PCR,SA-PCR,GA-PCR models for the aryl benzo data set using the selected data set**.**

Two-dimensional QSAR

parameter/molecule	actual	MR	SA-MR	GA-MR	PLS	SA-PLS	GA-PLS
r^2		0.7602	0.7637	0.5977	0.8222	0.4986	0.7148
q^2		0.4869	0.4224	0.2859	-0.2345	-0.2630	0.2570
F-Test		20.0721	10.9900	9.4098	29.2872	9.9458	25.0679
r^2 se		0.2933	0.2850	0.3517	0.2338	0.3827	0.2886
q^2 se		0.4290	0.4455	0.4686	0.6161	0.6074	0.4659
pred_r^2		0.3206	0.6649	0.7019	0.8041	0.4175	0.7465
pred_r^2 se		0.5046	0.4602	0.4340	0.3519	0.6067	0.4002
SELECTED DESCRIPTORS		Polar surface area excluding PandS, SssOE-index, T_2_C_4.	T_2_2_5, T_C_N_7, SssOHcount , T_C_F_3, SsCH3E-index	Polar surface area excluding PandS, T_C_N_6, T_O_O_0.	Polar surface area excluding PandS, T_2_2_5, SssOE-index, T_O_O_7, T_O_O_2.	ChiV5, Polar surface area excluding PandS, T_2_2_0.	T_T_N_3, CHIv0, T_2_2_4, SsOHE-index, SsCH3E-index.
Test data							
Molecule-2	9.569	9.3298	9.0263	9.0230	9.1737	8.8160	9.1278
Molecule-4	8.658	8.2758	8.3146	8.3474	8.3263	8.1845	8.4352
Molecule-8	8.244	8.3478	8.3650	8.3474	8.3263	8.1427	8.2833
Molecule-14	7.959	8.2537	8.3273	8.2184	8.2390	8.5731	8.2339
Molecule-27	8.699	9.5460	9.0197	9.0230	9.1737	9.0109	8.9878

parameter/molecule	actual	PCR(Forward)	SA-PCR	GA-PCR
r^2		0.7646	0.7017	0.3254
q^2		0.3449	0.2556	-0.0750
F-Test		20.5721	10.5836	10.1298
r^2 se		0.2690	0.3112	0.4332
q^2 se		0.4488	0.4916	0.5469
pred_r^2		0.8450	0.6729	0.0858
pred_r^2 se		0.3130	0.4546	0.7601
SELECTED DESCRIPTORS		Polar surface area excluding PandS, T_2_2_5, T_O_O_2, SssOcount.	SaaCHcount SsOHcount, T_T_N_4, T_2_2_6, SsCH3E-index,	ChiV1, Polar surface area excluding PandS, T_2_2_0.
Test data				
Molecule-2	9.569	9.2038	9.0370	8.5937
Molecule-4	8.658	8.3500	8.3048	8.1442
Molecule-8	8.244	8.3500	8.3049	8.4579
Molecule-14	7.959	8.1336	8.3049	8.6562
Molecule-27	8.699	9.2038	9.0093	8.8020

28 molecules which were taken from the literature were divided into two parts, one is test set and others into training set in 1:4 ratio and the test set had maximum and minimum IC_{50} value less than the respective IC_{50} values of training set. Different 2D QSAR methods are used such as SW, GA, MR using some sets of training and test set. This QSAR showed different q^2 values which ranged from 0.4986 and 0.8222. Different descriptors having different q^2 were obtained. To make a common set out of these different descriptors with different q^2 values all the 20 different descriptors were put into the neural network. The q^2 value thus generated which is more then 0.89(i.e. around 90% in prediction of IC_{50} value).

QSAR MODEL USING NEURAL NETWORK

Method -Back propagation

Training set size=23, Test set size=5

selected descriptors- ChiV1, Polar surface area excluding PandS, SaaCHcount, T_T_N_4,T_2_2_6, T_2_2_5, T_O_O_2,SssOcount, T_2_C_4, T_2_2_5,T_C_N_7,SssOHcount,T_C_F_3,SsCH3E-index, T_C_N_6, T_2_2_5, SssOE-index,T_O_O_7,T_O_O_2,T_O_O_0, , T_2_2_0, T_T_N_3,CHIv0,T_2_2_4,SsOHE-index.

Statics- N=23, degree of freedom=1, r^2 =0.9595, F Test =1.1278, pred_r^2 =0.7382, pred_r^2se =0.3132

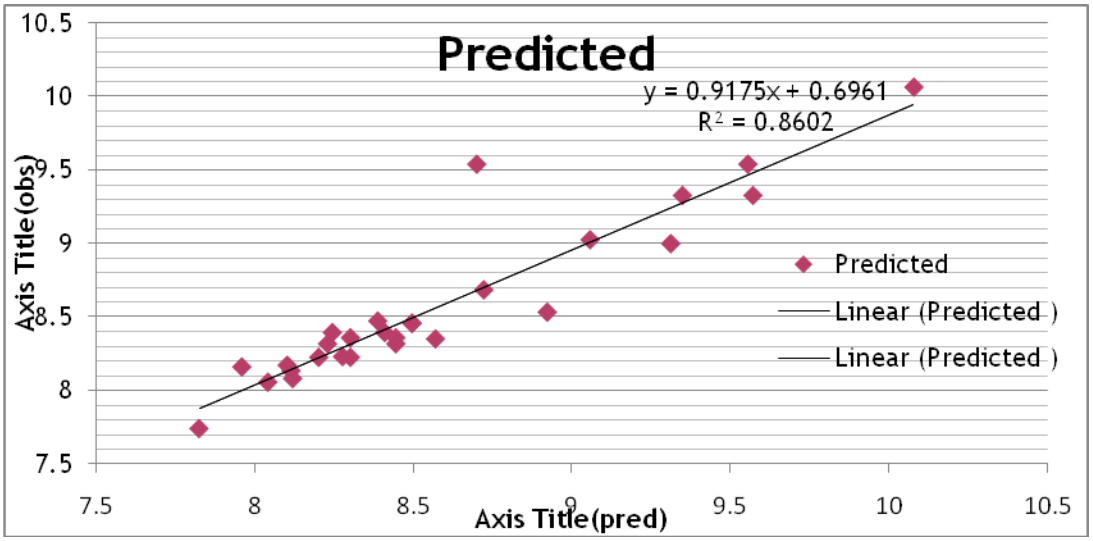


Fig 10 Plot of the observed vs. calculated pk_i values the binding affinities to the H_3 receptors

The relationship between different types of descriptors with activity is shown in graph.

Table: (4) selected descriptors with their function

ChiV0	atomic valence connectivity index (order 0)
ChiV1	atomic valence connectivity index (order 1)
SaaCHcount	the total number of carbon atoms connected with a hydrogen along with two aromatic bonds
SsOHcount	the total number of –OH group connected with one single bond.
SssOcount	total number of oxygen connected with two single bonds.
SsCH3E-index	Electro topological state indices for number of -CH3 group connected with one single bond.
SsOHE-index	Electro topological state indices for number of –OH group connected with one single bond.
SssOE-index	Electro topological state indices for number of oxygen atom connected with two single bonds
Polar Surface Area Excluding PandS	total polar surface area excluding phosphorous and sulphur.
T_2_2_0	the count of number of double bounded atoms (i.e. any double bonded atom, T_2) separated from any other double bonded atom by 0 bonds in a molecule.
T_2_2_4	number of double bounded atoms (i.e. any double bonded atom, T_2) separated from any other double bonded atom by 4 bonds in a molecule.
T_2_2_5	number of double bounded atoms (i.e. any double bonded atom, T_2) separated from any other double bonded atom by 5 bonds in a molecule.
T_2_2_6	This is the count of number of double bounded atoms (i.e. any double bonded atom, T_2) separated from any other double bonded atom by 6 bonds in a molecule.
T_2_C_4	number of double bounded atoms (i.e. any double bonded atom, T_2) separated from carbon atom by 4 bonds in a molecule
T_C_N_7	of number of Carbon atoms (single double or triple bonded) separated from any nitrogen atom (single or double bonded) by 7 bond

	distance in a molecule.
T_C_N_6	number of Carbon atoms (single double or triple bonded) separated from any nitrogen atom (single or double bonded) by 6 bond distance in a molecule.
T_C_F_3	number of Carbon atoms (single double or triple bonded) separated from any nitrogen atom (single or double bonded) by 3 bond distance in a molecule.
T_C_N_6	number of Carbon atoms (single double or triple bonded) separated from any nitrogen atom (single or double bonded) by 6 bond distance in a molecule.
T_O_O_2	the count of number of oxygen atoms (single double or triple bonded) separated from any oxygen atom (single or double bonded) by 2 bond distance in a molecule.
T_O_O_0	number of oxygen atoms (single double or triple bonded) separated from any oxygen atom (single or double bonded) by 0 bond distance in a molecule
T_O_O_7	number of oxygen atoms (single double or triple bonded) separated from any oxygen atom (single or double bonded) by 7 bond distance in a molecule

Interpretation and comparison of SW-KNNMFA and SA-KNNMFA 3d QSAR model

In our QSAR study it can be seen that the KNNMFA models obtained by using two variable selection methods shows that steric potential(1 out 3 in SW and 3 out of three in SA) and the descriptors S_275 common in both generated models.

Descriptors range for SA-KNNMFA

Electrorostatic –E_ 655(30.0000, 30.0000), E_584 (10.0000, 10.00000)

Steric potential -S_725 (7.7104, 30.0000)

Descriptors range for SW-KNNMFA

Steric potential -S_725 (7.7104, 30.0000), S_795 (-0.3099, 30.0000), S_675 (-0.0113, 30.0000)

Positive range indicates that positive electrostatic potential is favorable for increase in the activity and hence a less electronegative substituent group is preferred in that region. Negative range indicates that negative steric potential is favorable for increase in the activity and hence less bulky substituent group is preferred in that region. Positive range indicates that positive steric potential is favorable for increase in the activity and hence more bulky substituent group is preferred in that region.

Table: (5) Comparison of k-Nearest Neighbor Method for the histamine-3 data set using the optimal test set of 5 molecules.

parameter/molecule	actual	SW kNN-MFA	SA kNN-MFA
q^2		0.6154	0.5223
q^2 _se		0.3364	0.3749
pred_r^2		0.6240	0.1125
pred_r^2 se		0.8810	0.6513
descriptors		S_725,S_675, S_795.	S_725,E_655, E_584

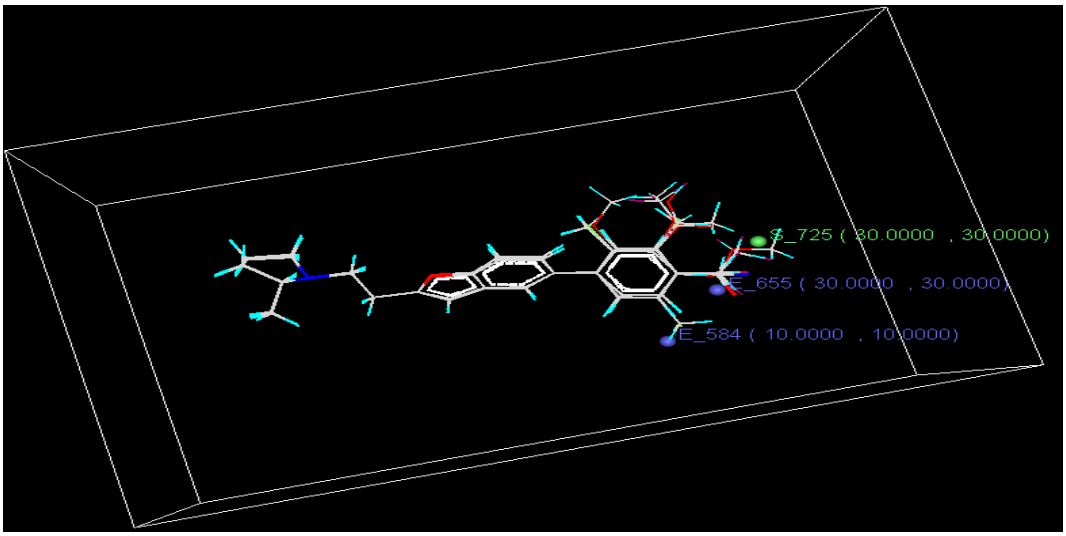


Fig 13 . Distribution of chosen points in the SA kNN-MFA for the aryl benzo data set with the selected test set molecule

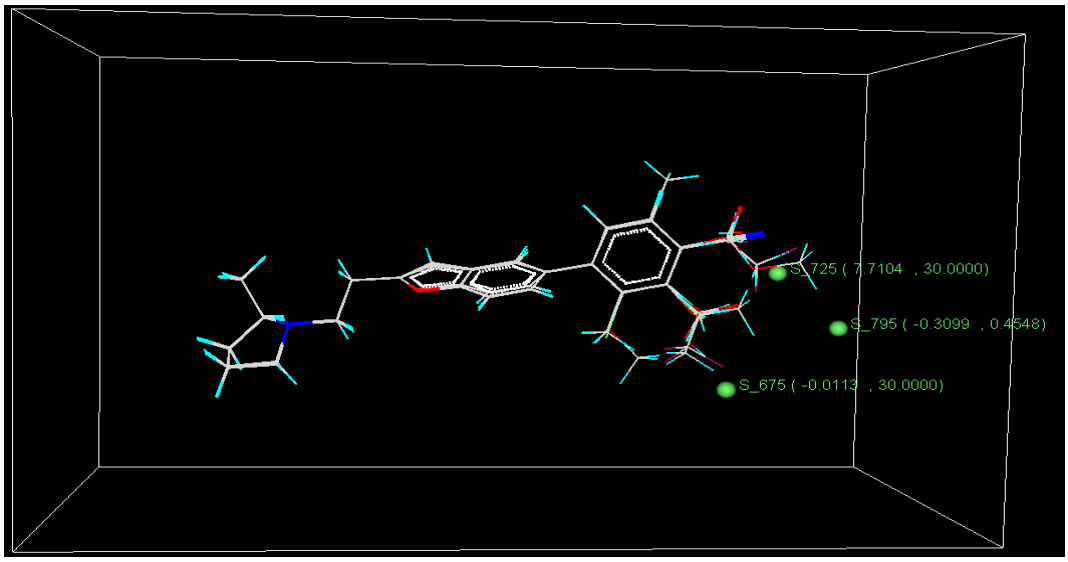


Fig 14 . Distribution of chosen points in the SA kNN-MFA for the aryl benzo data set with the selected test set molecule

Table: (6) 3d-QSAR data k-Nearest Neighbor Method.

SA kNN-MFA model			SW kNN-MFA model		
Test data	Actual	Predicted	Test data	Actual	Predicted
Molecule-1	9.347	9.6590	Molecule-1	9.347	9.5690
Molecule-4	8.658	7.9590	Molecule-4	8.658	8.3010
Molecule-10	8.102	8.7210	Molecule-10	8.102	9.5690

Molecule-21	8.444	8.2290	Molecule-21	8.444	8.2290
Molecule-28	9.553	8.6990	Molecule-28	9.553	8.6990
Training data			Training data		
Molecule-2	9.569	8.9210	Molecule-2	9.569	9.0560
Molecule-3	8.495	8.2010	Molecule-3	8.495	8.2010
Molecule-5	8.201	8.1190	Molecule-5	8.201	8.1140
Molecule-6	8.301	8.2760	Molecule-6	8.301	8.3010
Molecule-7	8.276	8.3870	Molecule-7	8.276	8.3010
Molecule-8	8.244	8.4090	Molecule-8	8.244	8.4090
Molecule-9	8.409	8.2440	Molecule-9	8.409	8.2440
Molecule-11	8.569	8.2290	Molecule-11	8.569	8.2290
Molecule-12	8.387	8.2760	Molecule-12	8.387	8.3010
Molecule-13	8.119	8.2010	Molecule-13	8.119	8.1140
Molecule-14	7.959	8.3010	Molecule-15	8.041	8.3010
Molecule-15	8.041	8.4410	Molecule-16	8.921	8.4090
Molecule-16	8.921	8.4090	Molecule-17	7.824	8.444
Molecule-17	7.824	8.2760	Molecule-18	8.444	8.3010
Molecule-18	8.444	8.4950	Molecule-19	8.301	8.1140
Molecule-19	8.301	8.3870	Molecule-20	8.229	8.3010
Molecule-20	8.229	8.5690	Molecule-22	8.721	8.9210
Molecule-22	8.721	8.6990	Molecule-23	10.076	9.3100
Molecule-23	10.076	9.3100	Molecule-24	9.056	9.5690
Molecule-24	9.056	9.5690	Molecule-25	9.310	8.9210
Molecule-25	9.310	8.4090	Molecule-26	8.114	8.2010
Molecule-26	8.114	8.1190	Molecule-27	9.5530	8.4090
Molecule-27	8.6990	8.7210	Molecule-2	9.569	9.0560

Pharmacophore study of AchE

Using crystal structure of rivastigmine in the active site of AchE. After redocking study revels two set of key interaction between the protein and the ligand. One set of interaction between the trp83 (located at the base of the active site) and quaternary amine of rivastigmine and second interaction of trp278 located at the opening of the cavity.

Pharmacophore study of histamine-3 antagonist

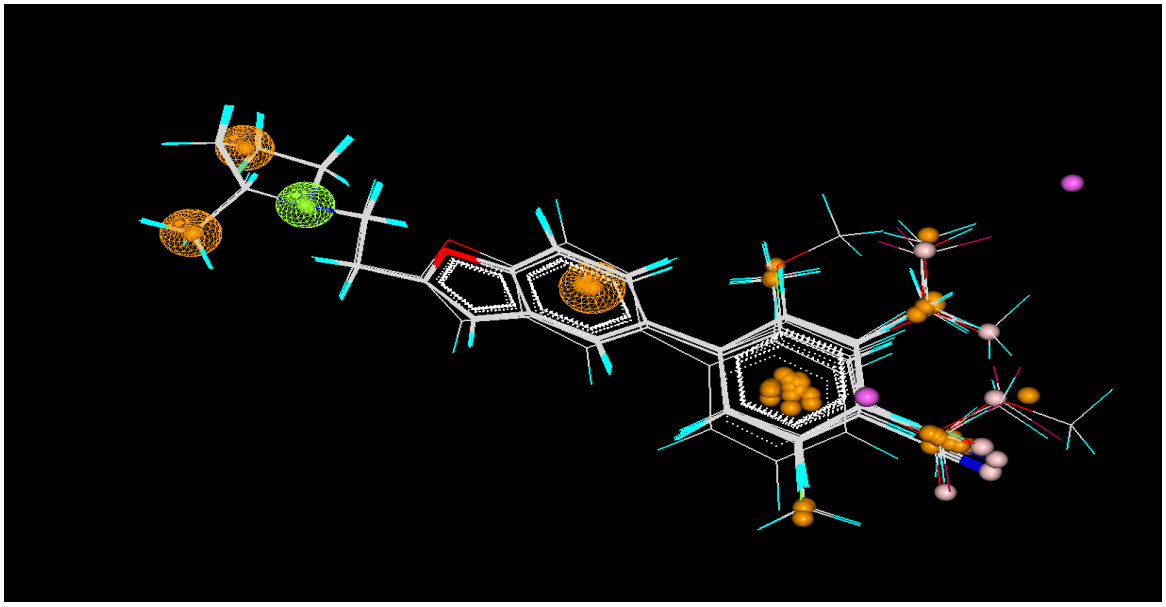
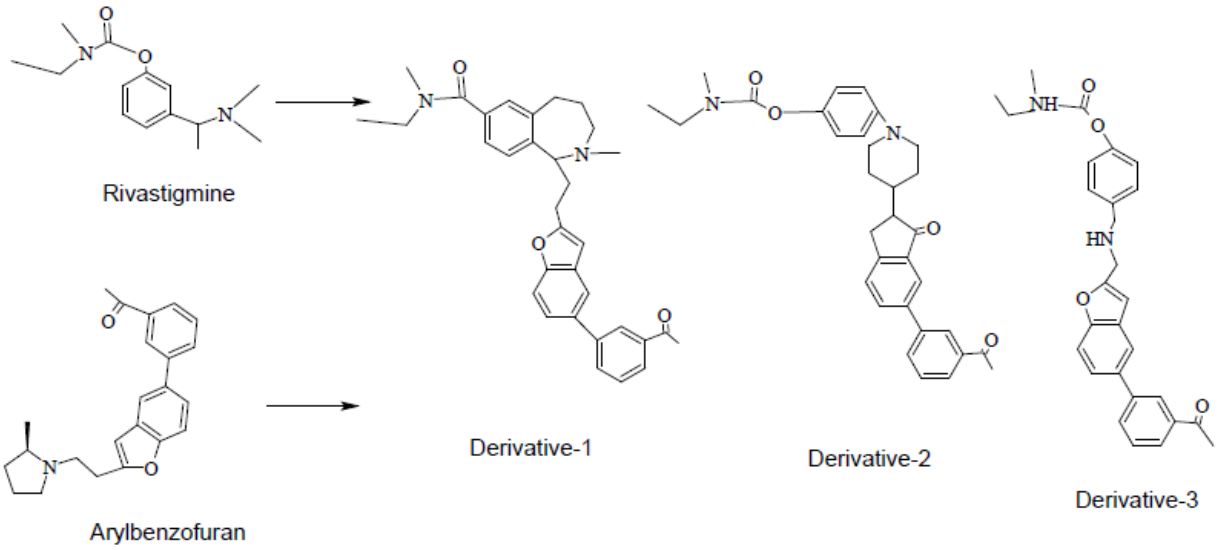


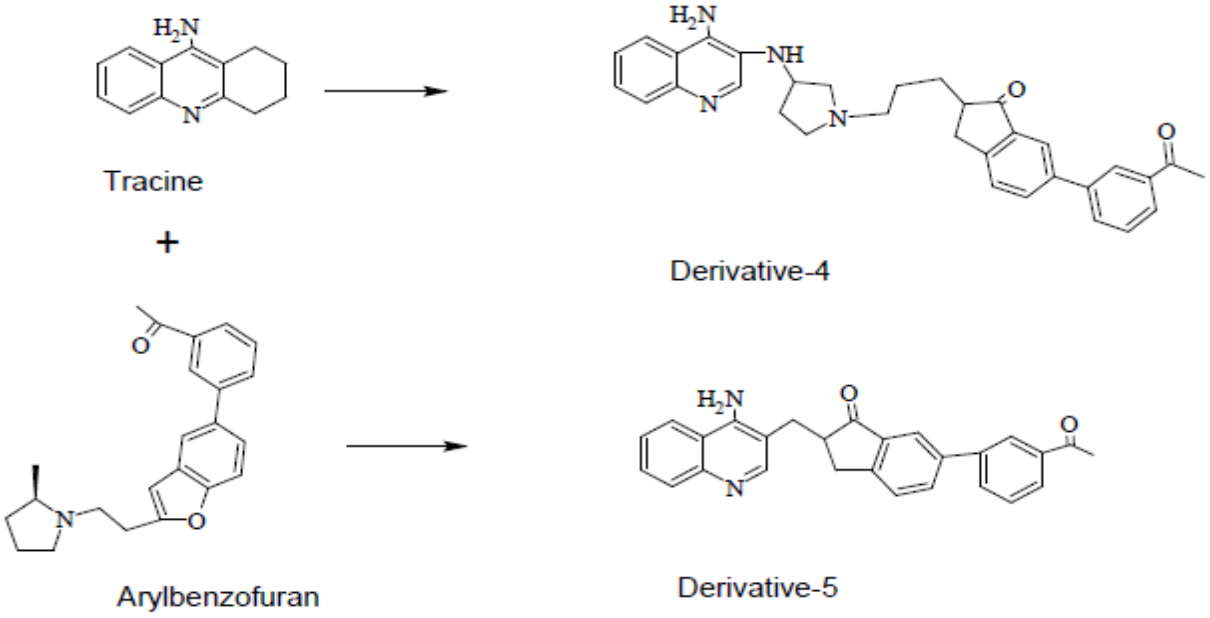
Fig 15 . Distribution of chosen points in the SA kNN-MFA for the aryl benzo data set with the selected test set molecule

5.5 Molecular hybrid

One of the quaternary amine out of two quaternary amine well fits with it. Quaternary amine of AchE inhibitor and h3 antagonist are well fit.So combining two class of molecule at that quaternary amine portion without changing the pharmacophore of other five hybrid design. Molecular hybridization of histamine receptor antagonists and AchE inhibitors.5 proposals of molecular hybridization involving Tacrine and Rivastigmine for AchE inhibitor activity and H_3 antagonist with best IC_{50} value. The entire 5 hybrids molecule passed through Lipinski five rules and none of them violated this rule. Toxicity predictions are done through ADME/TOX filter. Our main goal was to design an H3 receptor antagonist capable of inhibiting AchE. Here we use the crystal structure data of AchE (1EVE). And use the knowledge of the 2d QSAR, 3D QSAR and 3d pharmacophore of H3 receptor



(a)



(b)

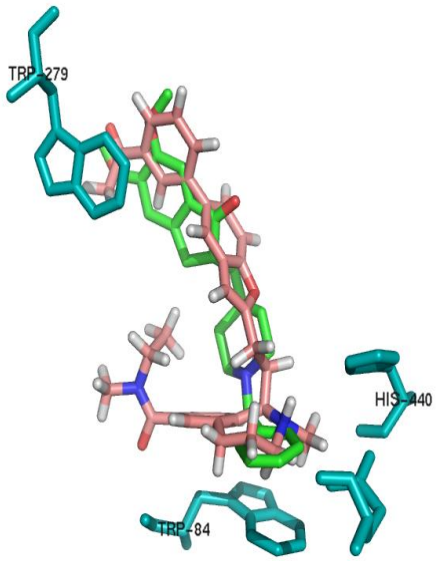
Scheme 2. (a) Derivatives 1, 2, 3 molecular hybridization of rivastigmine with arylbenzofuran ;(b) derivatives 4, 5 molecular hybridization of tacrine with arylbenzofuran

Validation of five hybrids

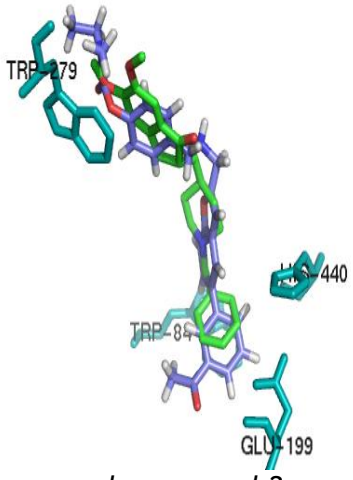
Out of all five hybrid the binding mode of hybrid-3,2,4 very much overlaid with the crystal structure binding mode of decamethonium 1eve.good superposition between the donepezil structure oriented with gold and the same molecule in the crystallographic orientation suggest the method used appropriate.

Table: (7) gold dock score of proposed compound with histamine-3 receptor**.**

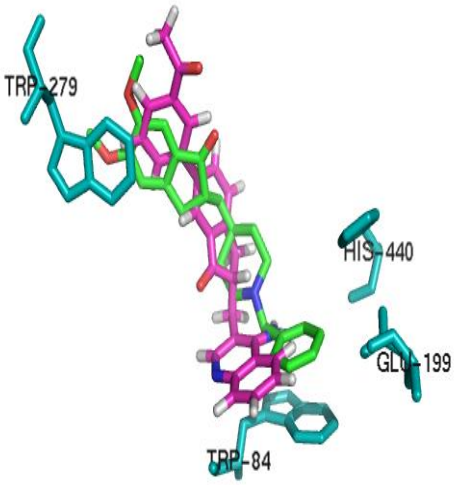
Molecule name	Fitness	S(hb_ext)	S(vdw_ext)	S(hb_int)	S(int)
Arylbenzofuran	42.51	1.92	42.20	0.00	-17.44
Proposal-1	13.40	1.36	45.35	0.00	-50.32
Proposal-2	57.07	0.00	51.97	0.00	-14.39
Proposal-3	76.62	11.51	63.14	0.00	-23.89
Proposal-4	45.92	6.88	41.86	0.00	-18.52
Proposal-5	57.35	0.35	55.41	0.00	-19.19



Proposed compound-3



proposed compound-2



Proposed compound-4

Fig 15. Details of the AchE active site in which the superposition of the crystallographic orientation of donepezil with the top 3 ranked solution suggested by GOLD is shown.

Table: (8) Gold dock score of proposed compound with AchE receptor**.**

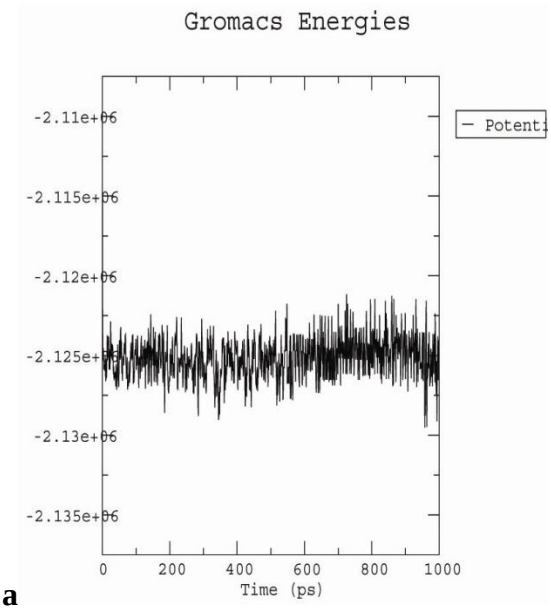
Molecule name	Fitness	S(hb_ext)	S(vdw_ext)	S(hb_int)	S(int)
Donepezil	**56.95**	**0.30**	**49.79**	**0.00**	**-11.82**
Proposal-1	**59.14**	**2.19**	**61.93**	**0.00**	**-28.20**
Proposal-2	**61.14**	**1.25**	**51.45**	**0.00**	**-10.85**
Proposal-3	**81.75**	**11.51**	**65.42**	**0.00**	**-19.71**
Proposal-4	**68.71**	**10.15**	**53.67**	**0.00**	**-15.24**
Proposal-5	**69.24**	**6.01**	**59.58**	**0.00**	**-18.68**

The conformation of highest score obtained with gold for this proposal-3(score=81) and it almost align with the rivastigmine Donepezil in pdb and for histamine-3 receptor proposed compound shows highest score.

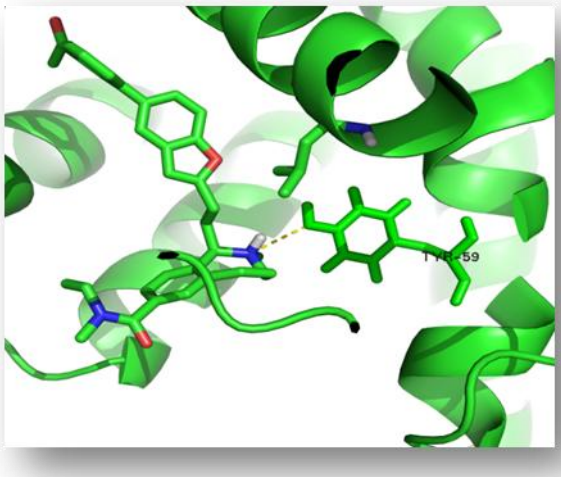
Molecular dynamic simulation MD simulation

The ligand –receptor complex resulting from docking calculation were placed in a box of water using algorithms was carries out for 1ns after initially equilibrated water molecules for 50 ns. An average structure was energy minimized under conjugated gradient and periodic boundary condition. The dynamic behavior and structural change of the receptor was analyzed by calculating the RMSD value for structural movement and change in the elements of secondary structure of the receptor model during the MD simulation.

The structure change of h3 receptor model was evaluated during 1ns MD simulation by use of GROMACS 3.3.1It can be seen from **Fig 17** that the potential energy of the h3 receptor model Ligands reaches to the plateau (2.0 A^0) with in first 200ps.



a



b

Fig 16 (a). Potential energy graph of protein –ligand complex during 1ns molecular simulation in the active site using GROMACS 3.3.1; (**b)** Avg structure during last 50 ps simulation in solvated condition.

Potential energy of protein-ligand complex remain in range between -2.12e+06 to -2.13e+06KJ/mol and the histamine-3 receptor backbone reaches a constant level after 250 ps at

2.0 A^0,but suddenly increases after 460 ps at 2.5 A^0 and then remain same for 1ns simulation time period .

The H-bond between drug molecule and enzyme were analyzed and after calculating the average of all H-bond candidates total 3 H-bonds were observed between ligand and receptor molecule.

The RMSD plots of protein backbone and the drug were obtained separately (Fig18 & Fig17 respectively). The backbone RMSD indicates that the rigid protein structure equilibrates rather quickly in this simulation (after 20 ps). The drug does not equilibrate until after 30 ps. The RMSD for the drug is more variable indicative of its mobility within the binding pocket.

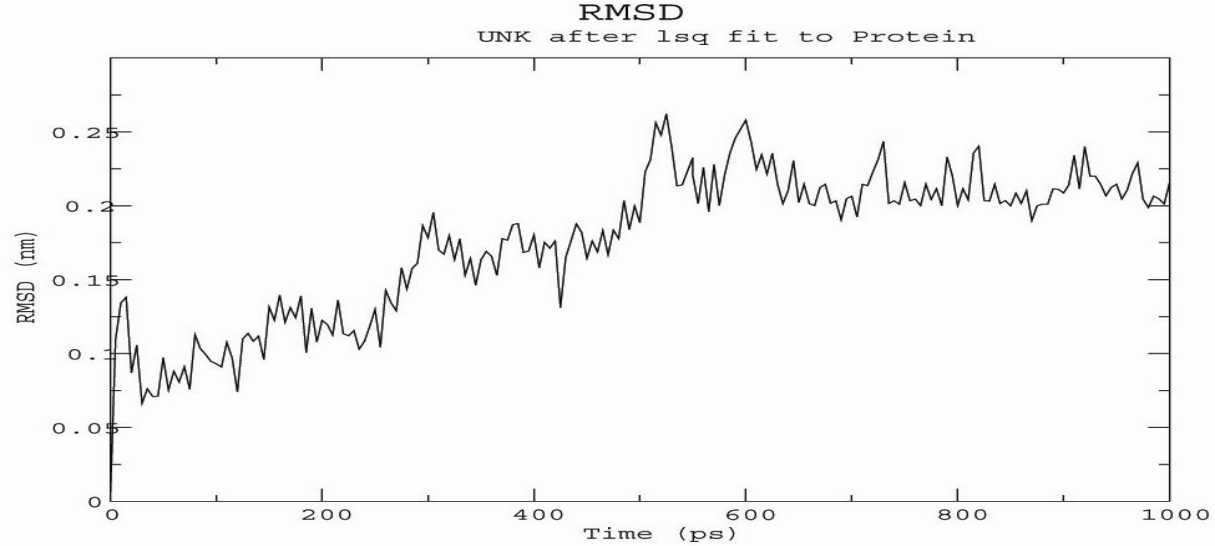


Fig 17. Plot showing the RMSD deviation of ligand in the solvated protein during 1ns molecular simulation in the active site using GROMACS 3.3.1

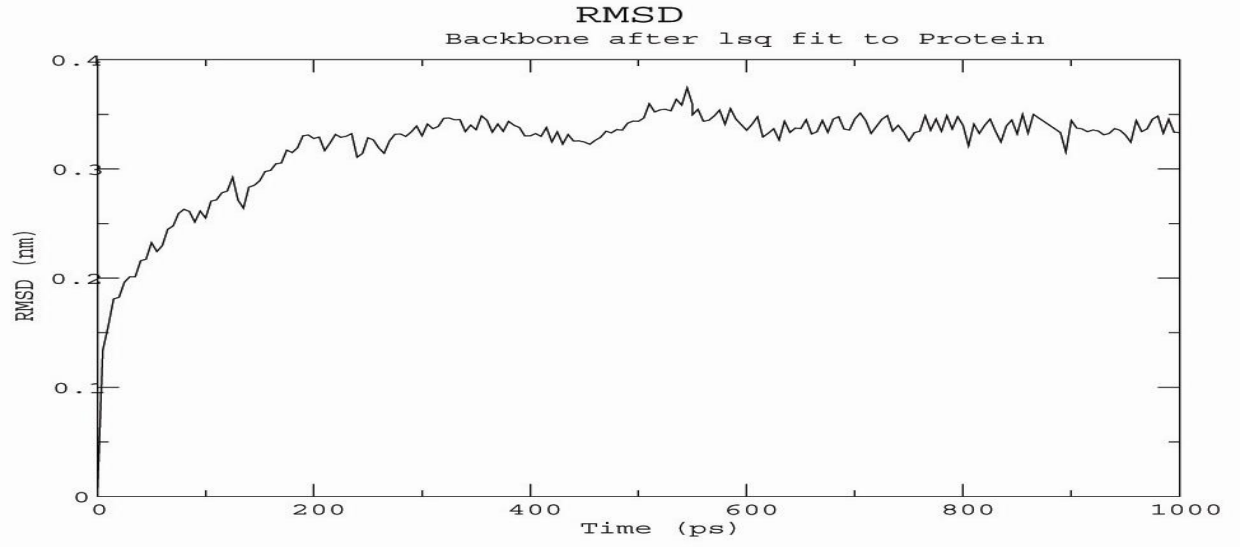


Fig 18. Plot showing the RMSD deviation of l in the solvated protein back bone during 1ns molecular simulation in the active site using GROMACS 3.3.1

The RMSD close to 3.5A^0 (for backbone) and 2.5 A^0 (for drug molecule) and fairly low potential energy close to -2.12e+06 to -2.13e+06 KJ/mol shows high stability of protein-ligand complex shows the likeliness of ligand molecule to be drug like candidates

CHAPTER 5

CONCLUSION

CONCLUSION

In this work, combined 2D QSAR study was carried out using various statistical models. Using all the models, 20 best descriptors were selected. Further, these descriptors were used in neural network back propagation model as a training set to conclude to a given result(Table 3). The value of cross validated squared correlation coefficient "$q^{2"}$ generated more than 0.89 (i.e. around 90% accurate in prediction of IC_{50} value) suggests a good internal productivity of the equation. The equation generated describes the positive contribution of Chiv 0, SssOE-index, SssCH3-index, Polar Surface Area excluding P and S whereas Chiv1, SssO count, SaaCHcount descriptors contributed negative to the inhibitory activity. From 3D QSAR study using SW kNN-MFA and SA kNN-MFA shows that steric potential (1 out of 3 descriptors in SA, 3 out of 3 descriptors in SW are steric potential descriptors) plays major role in determining biological activity. The descriptor S_725 (7.7104, 30.0000) was common in both generated model. Statistically, SW kNN-MFA model is comparatively better as compared to SA kNN-MFA with respect to q^2=0.6154. In the second part of this work, five novel molecular hybrids of the pharmaceuticals Tacrine, rivastigmine and aryl benzofuran, used in the Alzheimer's disease (AD), was designed and evaluated for further investigation and experimental validation. Based on this work, the hybrids proposed have shown very close orientations to the original pharmaceuticals. Our results suggest that the proposal-3, with highest synthetic viability, has more interactions with the both AchE and histamine-3 receptor. During molecular dynamic studies done with Histamine-3 receptor using GROMACS identified the amino acid residues responsible for the formation of the hydrogen bonding with **TYR-59** (tyrosine residue number 59). We propose that this molecule is an interesting pharmaceutical candidate for preparation and further investigation of various wet lab studies.

BIBLIOGRAPHY

1. D. Weininger, (1988) , J. Chem. Inf. Comput. Sci. 28: 31-36.
2. Michael M. C. Lai (1990) , Annu. Rev. Microbiol. 44:303-33
3. Appel R.D, Bairoch A, Hochstrasser D.F, (1994), Trends Biochem. Sci. 19:258- 260
4. Hugh G. Griffin and Annette M. Griffin, (1998). Software Programs For Bioinformatics And Biotechnology Motccutnr Biotchnologv Volume 10,
5. S.Dastmalchi, M.H.Mivehroud, T.Ghafourian, H.Hamzeiy (2007) , J.Molecular Graphics and Modeling, 26 (2008) 834-844.
6. F .U.Axe, S.D.Bembenek ,S. Szalama ,2006, J.Molecular Graphics and Modeling, 24 (2006) 456-464.
7. S.D.Bembenek, J.M.Keith, M.A.Letavic, R.Apodaca, A .J.Barbier,2008 , Bioorganic & medicinal chemistry, Bioorganic & Medicinal chemistry 16 (2008) 2968-2973.
8. C.S. Silva, V.L.Campo, I.Carvalho,2005, J.Molecular Graphics and Modeling, 26 (2006) 169-175.
9. B.Schlegel, C.Laggener, R. Meier,(2007), J. Computer Aided Mol Des ,21:437-453.
10. Volker Thiel (2003), J Gen Virol 84 2305-2315; DOI 10.1099/Vir.0.19424-0.
11. Poutanen, S.M., et Al. (2003), Engl. J. Med. In Press.
12. Rota, P.A., Et Al. (2003), Science. DOI:10.1126/Science.1085952.
13. Marra, M.A., Et Al. (2003), Science. DOI:10.1126/Science.1085953.
14. Kathryn V. Holmes (2003), J. Clin. Invest. 111:1605-1609
15. Xiang Xu1 Et Al, (2003), Nucleic Acids Research, Vol. 31, No. 24.
16. Volker Thiel, (2003), J Gen Virol 84 2305-2315; DOI 10.1099/Vir.0.19424-0.
17. Yang.H, (2003), Proc.Natl.Acad.Sci.Usa 100: 13190-13195
18. Richard A. Lewis ,(2001–2003).,Computer-Aided Drug Design Chemical Modelling: Applications and Theory, Volume 3
19. Guan Y, Zheng BJ, He YQ, Liu XL, Zhuang ZX, Cheung CL, Luo SW, Li PH, Zhang LJ, Guan YJ, Butt KM, Wong KL, Chan KW, Lim W, Shortridge KF, Yuen KY, Peiris JS, Poon LL ,(2003),Science, 302:276-278.
20. Stadler K, Masignani V, Eickmann M, Becker S, Abrignani S, Klenk HD, Rappuoli R ,(2003) Nat Rev Microbiol, 1:209-218

21. Martina BE, Haagmans BL, Kuiken T, Fouchier RA, Rimmelzwaan GF, Van Amerongen G, Peiris JS, Lim W, Osterhaus AD,(2003) Nature 2003, 425:915
22. Fouchier RA, Kuiken T, Schutten M, van Amerongen G, van Doornum GJ, van den Hoogen BG, Peiris M, Lim W, Stohr K, Osterhaus AD,(2003) Nature 2003, 423:240
23. Schwede T, Kopp J, Guex N, and Peitsch MC (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research* 31, 3381-3385.
24. G.N. Ramachandran, C. Ramakrishnan & V. Sasisekharan (1963): *Stereochemistry of polypeptide chain configurations.* In: *J. Mol. Biol.* vol. 7, p. 95-99.
25. Schuettelkopf and D. M. F. Van Aalten (2004), Acta Crystallogr. D60, 1355--1363.
26. Caren J Stark And Cd Atreya (2005), Virology Journal 2:35
27. Zihe Rao Neuman B W et al (2005), J. Virol. 79 :9665–76
28. David A Groneberg, Groneberg Et Al, (2005), Licensee Biomed Central Ltd,DOI:10.1186/1465-9921-6-8.
29. C. Alland1, F. Moreews, D. Boens, M. Carpentier, S. Chiusa, M. Lonquety, N. Renault, Y. Wong, H. Cantalloube, J. Chomilier, J. Hochez, J. Pothier, B.O. Villoutreix, J.-F. Zagury, P. Tuffery., (2005), Nucleic Rpbs: A Web Resource For Structural Bioinformatics, Acids Res., 33:W44-9.
30. Keng-Chang Tsai, (2006), J. Med. Chem. 49, 3485-3495
31. D.-Q. Wei1 (2006), Anti-SARS Drug Screening By Molecular Docking Amino Acids 31: 73–80
32. ACDLABS 11.0, (2006), Www.Acdlabs.Com,
33. G. Fear Et Al. (2007), Pharmacology & Therapeutics Volume 113, Issue 2, February 2007, Pages 354-368.
34. Jing-Fang Wang,(2007), IEEE, 1-4244-1120-3/07/
35. Dariusz Plewczynski (2007), J. Phys.: Condens. Matter 19 ,285207 (18pp)
36. Virupakshaiah Dbm ,(2007),Proceedings Of World Academy Of Science, Engineering And Technology Volume 24 October ISSN 1307-6884
37. M.Lesk, (2005), Introduction to Bioinformatics, Second edition.
38. Jose Luis Medina-Franco (2005), Journal of Computer-Aided Molecular Design 19: 229–242, DOI 10.1007/s10822-005-4789-8.

39. Yi Li and Juha, Rinne, O. and Mosconi, L. and Pirraglia, E. and Rusinek, H. and DeSanti, S. and Kemppainen, N. and Någren, K. and Kim, B. and Tsui, W. and Leon, M. (2008). Eur J Nucl Med Mol Imaging, 35,2169–218.
40. Buccafusco, J. J. Powers,J.C. Hernandez, Prendergast , M. A. Terry, A. V. Jonnala, R. R. (2007). MHP-133,. Neurochem Res 32:1224–1237.
41. Surya Prakash B.N Gupta (2008), E-Journal of Chemistry, Vol. 5, No. 1, Pp. 185-186. Youdim, M. B. H. and Buccafusco, J. J. (2005). J Neural Transm, 112: 519–537.
42. Bondi, M. W. & Jak, A. J. & Wood, L. D. & Jacobson, M. W. & Delis, D. C. & Salmon, D. P. (2008) Neuropsychological Contributions to the Early Identification of Alzheimer's disease. Neuropsychol Rev 18, 73–90
43. Yi Li and Juha, Rinne, O. and Mosconi, L. and Pirraglia, E. and Rusinek, H. and DeSanti, S. and Kemppainen, N. and Någren, K. and Kim, B. and Tsui, W. and Leon, M. (2008). Eur J Nucl Med Mol Imaging, 35,2169–2181.
44. Gustafson, L . Brun, A. Englund, E. Hagnell, O. Nilsson, K. Stensmyr, M . Öhlin, A. Abrahamson, M (1998). Hum Genet, 102,253–257.
45. Bondi, M. W. & Jak, A. J. & Wood, L. D. & Jacobson, M. W. & Delis, D. C. & Salmon, D. P. (2008). Neuropsychol Rev 18, 73–90.
46. Frigerio, C. S. Piscopo, P. (2005).J Neurol, 252, 1033–1036.
47. Lei, F. Birgit K. Jochen, L. Joerg, H. Yihua, Z. and Michael, D. (2008) Bioorganic & Medicinal Chemistry Letters ,18, 2905–2909.
48. Suh, W. H. Suslick, K. S. And Suh, Y.H. (2005). Curr. Med. Chem. – Central Nervous System Agents, 5, 259-269.
49. Gfesser G.A. , R.Faghih, Y.L.Bennami,M.P. Curtis,T.A. Esbenshade,AA. Hncock,M.D Cowart.(2005).Bioorg. Med. Chem. Lett. 15 (2005),8127-8134

.