*“Insilco Design of Novel Isoform Selective Histone Deacetylase Inhibitor as a Therapeutic Approach for Alzheimer’s disease Using Multiple Sequence Alignment, Machine Learning, Molecular Docking, ADME, And Mutation Analysis”*

A DISSERTATION

SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF

MASTER OF TECHNOLOGY

IN

**BIOINFORMATICS**

Submitted by:


**Rohan Gupta**

**2K16/BIO/05**


Under the supervision of

**Prof. Pravir Kumar**

Department of Biotechnology

Delhi Technological University

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042


JUNE, 2018

Delhi Technological University

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

## CANDIDATE'S DECLARATION

I, Rohan Gupta, 2K16/BIO/05 student of M.Tech Bioinformatics, hereby declare that the project Dissertation titled "**Insilco Design of Novel Histone Deacetylase Inhibitor as a Therapeutic Approach for Alzheimer's disease Using Multiple Sequence Alignment, Machine Learning, Molecular Docking, ADME, and Mutation Analysis**" which is submitted by me to the department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without paper citation. The work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: Delhi

Date: **ROHAN GUPTA**

Department of Biotechnology
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

## CERTIFICATE

I hereby certify that the Project Dissertation titled "**Insilco Design of Novel Histone Deacetylase Inhibitor as a Therapeutic Approach for Alzheimer's disease Using Multiple Sequence Alignment, Machine Learning, Molecular Docking, ADME, and Mutation Analysis**" which is submitted by Rohan Gupta, 2K16/BIO/05, Department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the student under my supervision. To the best of my knowledge this work has not been submitted in part of full for any Degree or Diploma to this University or elsewhere.

Place: Delhi

Date:

**Prof. PRAVIR KUMAR**
**SUPERVISOR**

ACKNOWLEDGEMENT

At the time of submission of my M.Tech Dissertation, I would first like to thank GOD for giving me patience, strength, capability, and willpower to complete my work. Apart from our efforts, the success of this project depends largely on the encouragement and guidelines of many others. I, therefore, take this opportunity to express my gratitude to the people who have been instrumental in the successful completion of this project.

My initial thank is addressed to my mentor Dr. Pravir Kumar, Professor, Department of Biotechnology, Delhi Technological University, who gave me this opportunity to work in a project under him. It was his enigmatic supervision, constant encouragement and expert guidance which have enabled me to complete this work. I humbly seize this opportunity to express my gratitude to him.

I would also like to extend my sincere gratitude to Professor Jai Gopal Sharma for providing basic infrastructure and facilities.

I would like to extend my sincere gratitude to Dr. Rashmi Ambasta for her keen observation and continuous advice. Her thoughtful inputs on the project issues have been invaluable for the productive progress of the project. I extend my thanks to technical staff Mr. Jitender Singh and Mr. C.B. Singh who had been an aid whenever required. Lastly, I wish to extend my thanks to my family and friends who have supported me through the entire process.

**ROHAN GUPTA**

**2K16/BIO/05**

# ABSTRACT

Histone deacetylases (HDAC) are $Zn2^{+}$ dependent cofactors or class of proteins that play a positive part in cellular transcription and functioning. Overexpression of these proteins are common in the progression of various discrepancies in brain tissues brings about the deregulation of different target proteins engaged with cell development and growth connected with Alzheimer's ailment that causes the shortage in memory and learning capacity. Although various approaches have been applied to control the higher expression of HDACs by repressing them with different compound inhibitors yet constrained effectiveness has been accomplished. In this study, we used ligand-based quantitative structure-activity relationship approach followed by machine learning model generation, molecular docking, and mutation studies in order to predict novel HDAC isoform-selective inhibitor by taking into consideration the previous studied binding calculations and morphological and chemical molecular descriptors. Total 11 novel compounds were selected having quite high binding affinities with different types of HDACs. Out of these compounds, one compound named as ChEMBL1834473 were able to interact at the central $Zn2^{+}$ with both class I and class II HDAC members. Overall chemical bioactivity and binding efficiency of anticipated compound recommended that the proposed compound tend to be a compelling inhibitor for Alzheimer’s disease.

**Keywords**: Alzheimer’s disease, Histone deacetylase, Histone deacetylase inhibitor, Molecular docking, Binding affinity, ChEMBL, SwissDock.

# CONTENTS

CANDIDATE'S DECLARATION ii
CERTIFICATE iii
ACKNOWLEDGEMENT iv
ABSTRACT v
CONTENTS vi
LIST OF FIGURES viii
LIST OF TABLES ix
LIST OF ABBREVIATIONS x
1. INTRODUCTION 1
2. LITERATURE REVIEW 6
   2.1 Alzheimer's disease: An Overview 6
   2.2 Histone Proteins in AD Pathogenesis 7
   2.3 Types of Histone Deacetylase 10
   2.4 Role of Histone Deacetylase Inhibitors in AD 11
3. METHODOLOGY 15
   3.1 Protein-Protein Interaction Network of HDAC And Key Proteins in AD 15
   3.2 Dataset Preparation 15
   3.3 Multiple Sequence Alignment (MSA) 16
   3.4 Pharmacophore Mapping and QSAR Studies 16
   3.5 Machine Learning (ML) Approach in Classification Of Compounds 16
   3.6 Molecular Docking (MD) Technique 17
   3.7 Ammino Acid Substitution Analysis 17
4. RESULTS 18
   4.1 Network Analysis of HDAC and Proteins Involved In AD 18
   4.2 Identity and Similarity between Different HDAC Using MSA 20
   4.3 Identification of Novel Compound Using QSAR Technique Followed By Machine Learning Modelling 23
   4.4 Molecular Docking Analysis 28
   4.5 Mutation Analysis 31
5. DISCUSSION 33

**6. CONCLUSION** **35**

**7. REFERENCES** **36**

## LIST OF FIGURES


| SERIAL NUMBER | NAME OF FIGURE | PAGE NUMBER |
|---|---|---|
| FIGURE 1 | Three Major Pathways Followed By HDAC Inhibitors That Prevents Cell Apoptosis In Neurological Improvement | 3 |
| FIGURE 2 | Classification Of Different Types Of HDAC Inhibitors According To Their Structure Variations | 4 |
| FIGURE 3 | Classical Amyloid Cascade Hypothesis | 7 |
| FIGURE 4 | HDAC Plays An Important Role In Cell Regulation Both In Cytoplasm And Nucleus | 9 |
| FIGURE 5 | Interaction Between Class I HDAC And Proteins That Regulate Brain Functions | 17 |
| FIGURE 6 | Interaction between Class IIa HDAC and Protein Alter Learning Ability | 18 |
| FIGURE 7 | Protein-Protein Interaction between Class IIb and Proteins Causes Memory Impairment and Decreases Learning Ability | 18 |
| FIGURE 8 | Multiple Sequence Alignment of Class I HDAC Isoforms Using ClustalW | 20 |
| FIGURE 9 | Multiple Sequence Alignment of Class II HDAC Isoforms Using ClustalW | 21 |
| FIGURE 10 | Pharmacophore Analysis of Known Inhibitors Using PharmMapper | 23 |
| FIGURE 11 | Graphical Comparison of The Characteristics Features of The Four Models Created Using ML Technique | 25 |
| FIGURE 12 | Graphical Representation of Binding Energy of Novel compounds with Different Types of Histone Deacetylase | 28 |
| FIGURE 13 | Docking Pose of HDAC6 with ChEMBL1834473 | 29 |
| FIGURE 14 | Binding Pocket of HDAC4 with ChEMBL1834473 | 29 |
| FIGURE 15 | Binding Pocket of HDAC1 and HDAC8 with ChEMBL1834473 | 30 |
| FIGURE 16 | Binding Pocket of Mutated (A) HDAC1 and (B) HDAC2 with ChEMBL1834473 | 31 |
| FIGURE 17 | Binding Pocket of Mutated (A) HDAC6 and (B) HDAC10 with ChEMBL1834473 | 31 |

# LIST OF TABLES


| SERIAL NUMBER | NAME OF TABLE | PAGE NUMBER |
|---|---|---|
| TABLE 1 | Classification of Different Types of HDAC's found in Mammals | 2 |
| TABLE 2 | Protein Structure Data with PDB ID, Resolution, and Year of Publication of Individual HDAC | 14 |
| TABLE 3 | Data associated with protein-protein interaction network of Class I, Class IIa, and Class IIb | 19 |
| TABLE 4 | Describes The Active Site Position, Ammino Acid Residues, And Conserved Domain In Different Isoforms Of HDAC | 19 |
| TABLE 5 | Percent Identity Matrix of Class I HDAC | 22 |
| TABLE 6 | Percent Identity Matrix of Class II HDAC | 22 |
| TABLE 7 | Characteristic Features of ML Models Determining The Accuracy of The Models Created | 24 |
| TABLE 8 | Molecular Descriptors of Test Set Compounds | 25 |
| TABLE 9 | Chemical And Physical Properties of Novel Compounds Identified Using DataWarrior | 26 |
| TABLE 10 | Binding Energy of Novel Compounds with Different Isoforms of Histone Deacetylase | 27 |
| TABLE 11 | Binding Energy of ChEMBL1834473 and Currently Used Inhibitors with Different Types of Mutated HDAC | 30 |

# LIST OF ABBREVATIONS

| | |
|---|---|
| **AD** | Alzheimer's Disease |
| **Aβ** | Amyloid β |
| **NFT** | Neurofibrillary Tangles |
| **PS1** | Presenile 1 |
| **PS2** | Presenile 2 |
| **HAT** | Histone Acetyltransferase |
| **HDAC** | Histone Deacetylase |
| **MRFs** | Myogenic Regulatory Factors |
| **SMC3** | Structural Maintenance of Chromosomes protein 3 |
| **STAT1** | Signal Transducer and Activator of Transcription 1 |
| **STAT2** | Signal Transducer and Activator of Transcription 2 |
| **GATA** | GATA Transcription Factors |
| **TRIM29** | Tripartite Motif Containing 29 |
| **MCD** | Malonyl-CoA Decarboxylase |
| **GAPDH** | Glyceraldehyde 3-phosphate Dehydrogenase |
| **BDNF** | Brain-Derived Neurotrophic Factor |
| **GDNF** | Glial Cell Line-Derived Neurotrophic Factor |
| **GSK3β** | Glycogen Synthase Kinase 3 Beta |
| **VPA** | Valporic Acid |
| **APP** | Amyloid Peptide Protein |
| **ZBG** | Zinc Binding Group |
| **PBS** | 4-phenylbutyrate |
| **MEF2** | Myocyte Specific Enhancer Factor 2A |
| **CREB** | cAMP-Response Element-Binding Protein |
| **MeCP2** | Methyl CpG Binding Protein 2 |

# 1. INTRODUCTION

Alzheimer's disease (AD) is a common nervous system disorder that relates to memory deficit and destroys the thinking and learning ability [1]. More than 5 million people in America suffering from the most common form of dementia called as Alzheimer's disease. Among all the cases related to memory impairment and learning deficit, 70%-80% relates to Alzheimer's. The conceivable cause of AD is the loss of synaptic neurons due to deposition of amyloid-β (Aβ) plaques and tau neurofibrillary tangles (NFT) [2]. According to most acceptable theory "amyloid cascade hypothesis" accumulated Aβ peptide leads to cascade of events which results in apoptosis of neuronal cells and initiate the pathogenesis of AD **[**3**]**. According to various evidences genetic changes may occur in the case of familial AD such as overproduction of Aβ peptide in association with presenile 1 and 2 (PS1and PS2), amyloid β precursor protein, and apolipoprotein E **[**4**]**. Epigenetics is the branch of advanced genetics and proteomics that will help in determining the role of histone proteins in the pathogenesis of familial and sporadic AD especially in the case of Sporadic AD. Recent advancement in the field of functional genomics explains the role of histone acetylation and deacetylation in the etiology of AD and other neurodegenerative disorders **[**5**]**. Post-translational modification of histone and non-histone proteins have become the most significant area of research in the pathogenesis of AD [6]. Various studies in the past designated the role of histone acetylation and deacetylation in memory impairment and recover learning and thinking ability [7]. Histone acetyltransferase (HAT) and histone deacetylase (HDAC's) were two key enzymes that catalyze the acetylation and deacetylation activity of histone proteins respectively associated with double helical DNA structure [8][9]. Acetylation by HAT promotes chromatin relaxation by opening the packed histones and DNA [10], thus allowed the binding of specific transcription factors to the respective gene promoter in order to activate transcription and gene expression [11]. In case of deacetylation, nucleosome becomes compact and controlled the binding of the transcription factor to its specific gene promoter leads to controlled gene expression. Any instabilities in the HAT and HDAC's activity leads to abnormal gene expression that may cause vital human disorders such as cancer, neuropathy, diabetes, neoplasm, AD, and Parkinson's disease [12][13].

Till date total 18 human HDAC's have been identified which were divided into two super families and four classes on the basis of structural, functional and phylogenetic analysis [14].

Class I, Class II, Class IV were $Zn^{2+}$ dependent while Class III, which only includes Sirtuins were $NAD^{+}$ dependent [15]. Class I HDAC includes HDAC1, HDAC2, HDAC3, and HDAC8 have sequence similarity to yeast Rpd3 protein while Class II HDAC includes HDAC4, HDAC5, HDAC6, HDAC7, HDAC9, and HDAC10 have sequence similarity to yeast Hda1 protein.

| Class | Type Dependent Cofactor | Type | Subcellular Location | Length of Ammino Acids | Possible Brain Defective Area | Non Histone Substrates | Further Reading |
|---|---|---|---|---|---|---|---|
| I | $Zn^{2+}$ | HDAC1 | Nucleus | 482 | Cortex | p53, MyoD | 17, 18 |
| | $Zn^{2+}$ | HDAC2 | Nucleus | 488 | Cortex | p53, MyoD | 17, 18 |
| | $Zn^{2+}$ | HDAC3 | Nucleus/ Cytosol | 428 | Hippocampus, Cortex, Amigdala | Myocyte Enhancer Factor-2, p65, STAT1, and STAT2 | 19, 20, 21, 22 |
| | $Zn^{2+}$ | HDAC8 | Nucleus/ Cytosol | 377 | Muscle | SMC3 | 23 |
| IIa | $Zn^{2+}$ | HDAC4 | Nucleus/ Cytosol | 1084 | Cortex, Amigdala, Hippocampus, Locus coeruleus | p53,DNAJB8 | 24 |
| | $Zn^{2+}$ | HDAC5 | Nucleus/ Cytosol | 1122 | In Many Parts | GATA1 | 25 |
| | $Zn^{2+}$ | HDAC7 | Nucleus/ Cytosol | 952 | Substantia Nigra, Hippocampus | H1F1α | 26 |
| | $Zn^{2+}$ | HDAC9 | Nucleus/ Cytosol | 1011 | Substantia Nigra | TRIM29 | 27 |
| IIb | $Zn^{2+}$ | HDAC6 | Cytosol | 1215 | Hippocampus | α-Tubulin, HSP90, Cortactin | 28, 29, 30 |
| | $Zn^{2+}$ | HDAC10 | Nucleus/ Cytosol | 669 | Amigdala, Cortex, Hippocampus | Pax3, HSP70 | 31 |
| III | $NAD^{+}$ | Sirt1 | Nucleus | 747 | Cortex, Hypothalamus, Hippocampus | p53, p300 | 32, 33 |
| | $NAD^{+}$ | Sirt2 | Cytosol | 389 | Oligodendrites, Olfactory Neurons | α-Tubulin, HSP90 | 34 |
| | $NAD^{+}$ | Sirt3 | Mitochondria | 399 | Unknown | Ku70 | 35, 36 |
| | $NAD^{+}$ | Sirt4 | Mitochondria | 314 | Unknown | MCD | 37 |
| | $NAD^{+}$ | Sirt5 | Mitochondria | 310 | Cortex | CPS1 | 38 |
| | $NAD^{+}$ | Sirt6 | Nucleus | 355 | Unknown | Unknown | 39 |
| | $NAD^{+}$ | Sirt7 | Nucleolus | 400 | Unknown | P53 | 40 |
| IV | $Zn^{2+}$ | HDAC11 | Nucleus | 347 | In Many Parts | Unknown | 41, 42 |

**Table 1: Classification of different types of HDACs found in mammals.**

Various studies on mouse model provide the strongest evidence in the support of the possible role of HDAC inhibitors in the treatment of neurodegenerative disorders in enhancing learning and thinking ability. Histone deacetylase inhibitors were the effective treatment against neurological disorders via two major neuroprotective mechanisms, including site-specific transcriptional activation of a diseased gene leading to control gene expression and other being modulated histone acetylation homeostasis [43]. Studies such as treatment with

HDAC6 selective inhibitor resulted in alleviated levels of Aβ degrading enzyme neprilysin (NEP) and administration of MS-275 in APP mouse model decreased the Aβ deposition proved the possible role of HDAC inhibitors in the treatment of AD [44][45].

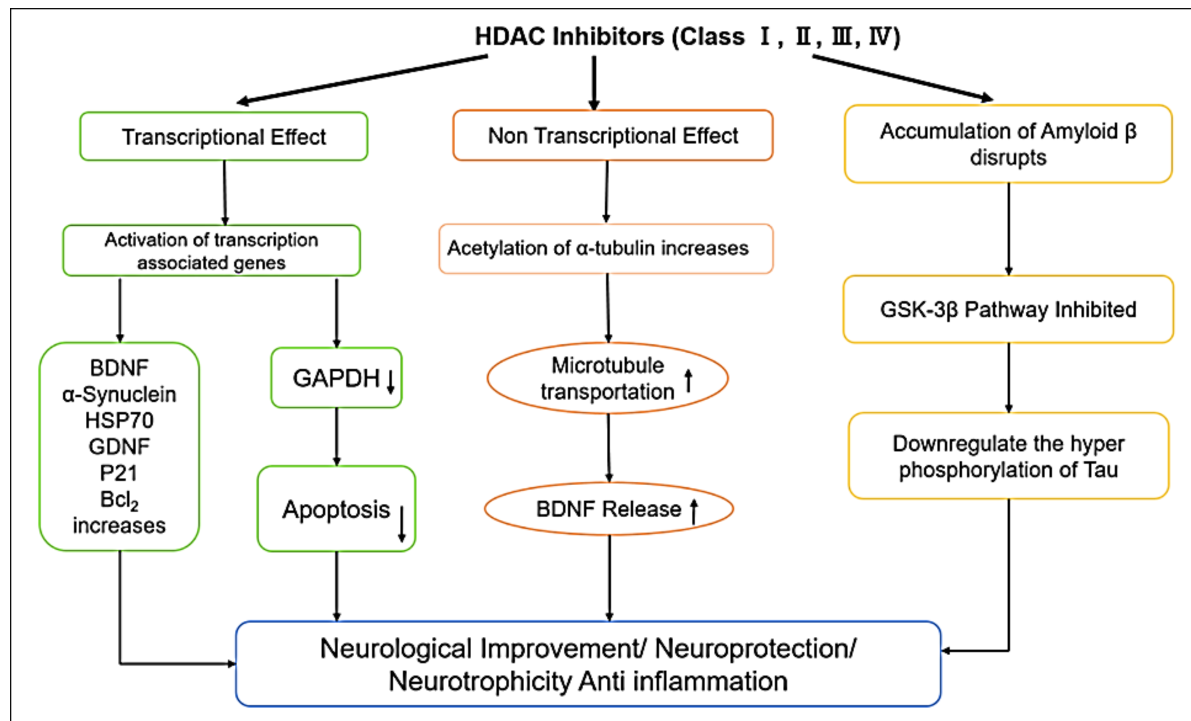


**Figure 1: Three major pathways followed by HDAC inhibitors that prevents cell apoptosis in neurological improvement. Transcriptional effect in which genes associated with neurons gets activated increases the expression of BDNF, GDNF, P21 while decreases the activity of GAPDH due to which cell apoptosis inhibits. Non transcriptional effect increases the acetylation of α tubulin due to which microtubule transportation and expression of BDNF increases significantly. In the third pathway accumulation of amyloid peptide protein declines and inhibition of GSK3β pathway occurs due to which hyper phosphorylation of tau protein decreases which leads to neurological improvement.**

The HDAC2 inhibitor was found to promote the growth dendritic spine density and further degrade the Aβ deposition [46] [47] whereas inhibiting HDAC3 prevents the synaptic plasticity induced by amyloid oligomers [48]. In one study it was found that inhibiting the activity of HDAC6 promotes the clearance of Aβ and tau [49]. It was also known to increase the axonal transport system and restores α-tubulin acetylation [50]. Studies suggested that Valporic acid (VPA) and lithium inhibits the production of Aβ peptide in HEK293 cell line when transfected with Swedish $APP_{751}$ [51]**.** Another study found the role of VPA in

decreasing the Aβ production by inhibiting the GSK-3β- mediated pathway that alleviates the memory linked disorder in AD mouse model [52].

HDAC inhibitors were classified into four groups based on their structure, catalytic activity, and a side chain. These include hydroxamic acids (Vorinostat, Belinostat) [53], benzamide (Entinostat, mocetinostat, SK-7041) [54], carboxylates (Valproic acids) [55], and cyclic tetra peptides (Trapoxin A, Romidepsin) [56]. Among these short-chain carbolic acids were weak inhibitors while cyclic tetra peptides were structurally complex inhibitors and under clinical trials [57]. Though these inhibitors were different, but they share three characteristic features termed as a Zinc-binding group (ZBG) [58] lipophilic linker and hydrophobic cap [59]. ZBG interacts with central $Zn^{2+}$ through coordination bond and formed hydrogen bonds with the residues near $Zn^{2+}$ while hydrophobic cap binds with the residues surrounding the active sites.

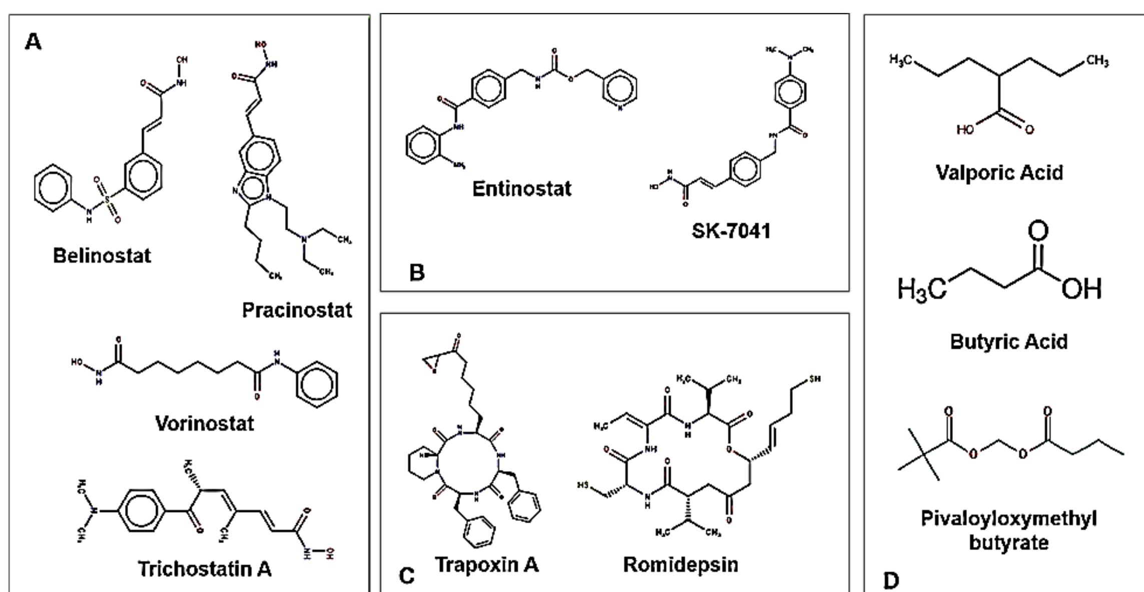


**Figure 2: Classification of different types of HDAC inhibitors according to their structure variations (A) hydroxamic acid inhibitors, (B) benzamide inhibitors, (C) cyclic tetra peptides inhibitors, and (D) carboxylates inhibitors.**

Various studies in the past concluded that histone deacetylase regulates the activity of acetylation process and histone deacetylase inhibitors alleviates the level of histone proteins acetylation and help in memory and learning ability thus prevent neuronal cell death [60]. Three proteins namely Aβ, GSK3β, and tau help play a protective role in diseased condition which regulates their activity on the application of HDAC inhibitors. In mice model increasing expression of HDAC2 but not HDAC1 results in decreased synaptic plasticity,

synaptic numbers, and memory formation ability and thus Vorinostat a pan HDAC2 inhibitor prevents the reduced synaptic plasticity and helps in memory formation [61]. Akhtar et al. suggested that overexpression of HDAC2 decreases the neurotransmission, indicating the possible role of HDAC2 in synaptic plasticity. HDAC3 express strong neurotoxic property which is cellular selective in nature. HDAC3 is phosphorylated during the activation of GSK3β pathway, thus inhibiting the GSK3β pathway prevents from HDAC3 mediated neuronal cell death and neurotoxicity [62]. In another experiment it was found out that 4-phenylbutyrate (PBA) down regulates the phosphorylation of tau protein indirectly via increase in the inactive GSK3β pathway in AD mouse model [63]. PBA promotes the clearance of accumulated intraneuronal Aβ protein and helps in moderation of endoplasmic reticulum (ER) stress in diseased mouse model. Nicotinamide which is known for $NAD^+$ dependent HDAC competitive inhibitor restores the memory impairment and learning deficit in mouse model by selectively decreases phosphorylation of tau protein at Thr231 position due to which concentration of acetylated α-tubulin increases. As compared to normal healthy brain, concentration of HDAC6 found to be significantly increased in cortex and hippocampus region of diseased mouse model [64]. Tubacin a selective HDAC inhibitor inhibits the hyper phosphorylation of tau protein which initiates the misrelated movement of mitochondrial neuron by activating GSK3β pathway. Several studies found the possible function of oxidative stress in the pathology of AD and it was suggested that selective inhibition of HDAC6 protects against neurodegenerative disorders and neuronal cell death induced by oxidative stress [65]. In one study it was stated the abnormal activity of HDAC4 causes neuronal apoptosis while inhibiting the overexpression of HDAC4 activity promotes cell differentiation and cell growth [66].

# 2. REVIEW OF LITERATURE

## 2.1 Alzheimer's disease: An Overview

A neurodegenerative disorder which is characterized by memory impairment and learning inability is called as Alzheimer's disease (AD) which occur in the cerebrum part of the brain and commonly develop with the age. According to well-studied hypothesis the most common basis of AD in humans is the accumulation of β-amyloid protein in the brain due to cleavage of amyloid peptide protein leads to neuronal apoptosis and neuronal cell death. The percentage of diseased population upsurges substantially after the age of 70-75 and it may be shown that this data may rise up to 50% over the population having age above 85. People suffering from heart disease, high blood pressure, and diabetes were at high risk of developing symptoms associated with AD. Other common risk factors that lead to AD includes elevated levels of blood cholesterol, person who is unable to complete eight full years' education, and person who undergone through distressing head injury particularly associated with apoE4 gene [67]. Individuals having in close proximity with AD will exhibit several characteristics features from the early stage of disease and these features will become more dominant as the disease progress over the years. The symptoms associated with AD include memory impairment, having trouble in performing well known errands, problems regarding distance and time judgement, usually lost their belongings or misplace them, and frequent changes in behavior, attitude, and personality. All these features were associated with memory impairment and learning inability in one way or other way. Thus it may be concluded that memory and learning inability due to neuronal cell death were key features found in person having AD [68]. The onset of AD associated with slowly progressive as the age increases and the condition may be further deteriorate when the thinking ability affects along with memory impairment which is considered as second stage of diseased condition. Short term memory problems such as attitude and behavior changes, failing to switch off buttons and doors, withdrawal from the social life occur in the early stages of AD. With the time, symptoms such as difficulty in mental thinking ability and different cerebral functions develops which is considered as second stage of diseased condition.

Till now the major cause of AD associated with genetic phenomenon which is well explained by "Amyloid cascade hypothesis". Years of studies demonstrated that genetic changes in genes such as tau protein, β-peptide, and ApoE leads to neuronal cell death and apoptosis. In

half of the total diseased population mutation linked with key protein found to be major cause of AD which leads to excessive production of amyloid-β (Aβ).

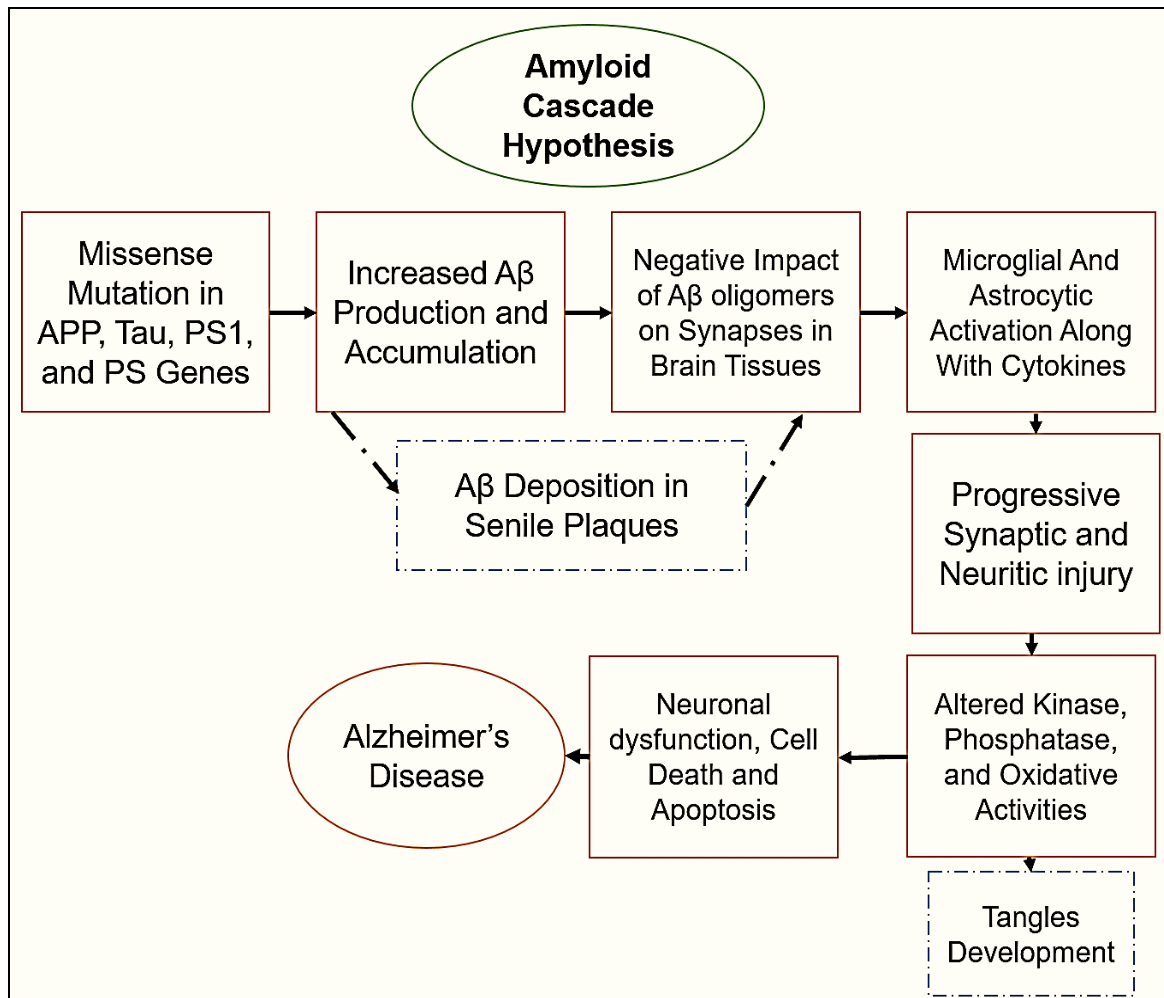


**Figure 3: Classical amyloid cascade hypothesis major cause of AD in which missense mutation in APP and Tau causes increased Aβ production and accumulation in senile plaques which activates microglial and astrocytes in brain tissues leads to neuritic injury, cell death and apoptosis causes AD.**

## 2.2 Histone Proteins in AD Pathogenesis

Chromatin remodelling and alteration of double stranded DNA structure represents the central phenomenon of gene regulation and expression during the brain development process and memory enhancement. Modification of histone proteins at cellular level become the main topic in the treatment of AD. Mainly epigenetic process deals with the modification of histone proteins. In one study, addition of acetyl group to histone known as acetylation is a

key event in cell growth and survival during the pathogenesis of AD found in the postmortem brain tissues [69]. Till now transcriptional mechanism in pathogenesis of AD is not much clear as in other neurodegenerative disorder but several studies have found the crucial link between AD and HDAC. In one study, it was found out that the α-secretase and β-secretase initially cleave the crucial protein APP found in the brain tissues which generate β-amyloid fragments and intracellular tail fragments. Accumulation of β-amyloid fragments in the brain tissues imbalance the gene expression and function leads to cell death and apoptosis. Another study concluded the role of histone deacetylase by activating the GSK-3β pathway leads to hyper phosphorylation of tau protein [70]. Low potassium and glutamate levels in brain trigger the neuronal translocation of HDAC4 and represses MEF2 and CREB associated transcription leads to neuronal cell apoptosis. Recent development found out that HDAC6 facilitate both UPS dependent and autophagy mediated degradation of excessive accumulated proteins.

HDAC1, a class I HDAC, was shown to play a protective role in a p25/cyclin-dependent kinase 5-inducible transgenic mouse model. Activation of p25 gene expression in mouse model leads to postnatal neurodegeneration which is mediated by a direct interaction between p25 and a catalytic region of HDAC1. Reduction in HDAC1 activity trigger aberrant expression of various cell cycle proteins and DNA injury, leading to neuronal death. The in vivo findings of class I HDAC activity on neuronal cell were reviewed in cortical region of brain. In these neurons, the overexpression of HDAC1, but not a catalytically inactive mutant, protected cultured neurons against p25-induced neurotoxicity. Pharmacological inhibition of HAT activity was also found to be protective against p25 induced neurotoxicity and the degree of protection was not increased by HDAC1 coexpression [71]. These studies suggest that the neuroprotective effect of HDAC1 is mediated by histone deacetylation.

Moreover the role of HDAC4 in regulating neurodegeneration is still unknown and require large amount of studies. Initially it was reported that the over expression of HDAC4 in cultured cerebellar granule neurons leads to cell death and cell apoptosis. In one study, it was found out that HDAC4 activity leads to neurological improvement by protecting the neurons against low-potassium-induced cell death. HDAC4 was also was shown to protect cultured cortical neurons against homocysteic acid and 6-hydroxy dopamine induced toxicity. Arguing strongly for a survival-promoting role for HDAC4 in neurons is the finding that the cerebellums of HDAC4 knock-out mice, which die within 2 weeks of birth, display postnatal degeneration of Purkinje neurons. Using HDAC inhibitors and mutant HDAC4 constructs

lacking the deacetylase domain, it was found that HDAC4-mediated neuroprotection does not require deacetylase activity [72]. Rather, neuroprotection by HDAC4 involves an inhibition of cell cycle proteins such as cyclin-dependent kinase-1. A recently published report by confirms a neuroprotective role for HDAC4. It demonstrated that HDAC4 reduced naturally occurring neuronal death in the retina, and that HDAC4 overexpression in a mouse model of retinal degeneration prolonged photoreceptor survival.

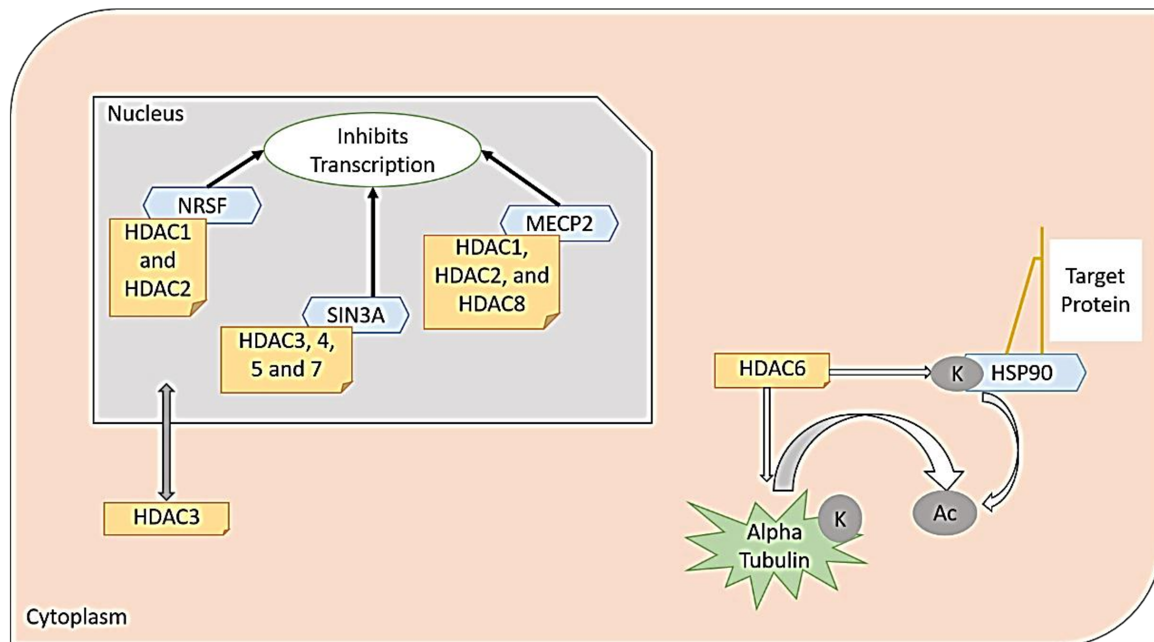


**Figure 4: HDAC plays an important role in cell regulation both in cytoplasm and nucleus. In Nucleus HDAC 1 and HDAC2 regulates the activity of NRSF gene, HDAC3, HDAC4, HDAC5, and HDAC7 regulates gene SIN3A while HDAC1, HDAC2, and HDAC8 inhibits MECP2 activity which causes transcription and leads to cell death and apoptosis. In cytoplasm HDAC6 acetylates alpha tubulin and HSP90 to inhibit transcription associated genes.**

## 2.3 Types of Histone Deacetylase

Mammals express 18 HDAC proteins, which have been grouped into four classes based on their homology to yeast deacetylase proteins. Class I HDACs (HDAC1, HDAC2, HDAC3 and HDAC8) are homologues of the yeast HDAC RPD3 protein. These HDACs are expressed ubiquitously, localized predominantly in the nucleus (with the exception of HDAC3, which can also be found in the cytoplasm) and possess high enzymatic activity. HDAC1 and HDAC2 are structurally very similar and within cells, are found complexed with corepressors such as the mammalian paired amphipathic helix protein Sin3 and the protein

CoREST, as well as with the polycomb-repressive complex 2 (PRC2) and the nucleosome remodeling and histone deacetylation (NuRD) complex [73].

Class II HDACs are homologous to the yeast HDAC HDA1 and are further divided into class IIa (HDAC4, -5, -7 and -9) and class IIb (HDAC6 and -10) HDACs. Class IIa HDACs are characterized by large N-terminal extensions with conserved binding sites for the transcription factor myocyte-specific enhancer factor 2A (MEF2) and 14-3-3 protein eta, and can shuttle between the nucleus and cytoplasm in a phosphorylation-dependent manner. Phosphorylation is mediated by kinases such as calcium/calmodulin-dependent protein kinase (CaMK) and protein kinase D at conserved serine residues in the N-terminal region of these proteins. Class IIa HDACs also display a restricted expression pattern. For example, HDAC4, -5 and -9 are highly expressed in the brain, heart and skeletal muscle, while HDAC7 is abundant in endothelial cells and thymocytes [74]. Class IIb HDACs lack the Nterminal extension but possess two non-identical catalytic domains in tandem. HDAC6 is localized exclusively in the cytoplasm where it is physically associated with tubulin. Deacetylation of tubulin as well as other cytoskeletal and transmembrane proteins by HDAC6 has been reported [75]. Relatively little is known about the intracellular localization, expression pattern or functions of HDAC10. HDAC11 shares sequence conservation with both RPD3 and HDA1 and is therefore placed in class IV. Its expression is highest in the brain, heart, testis and kidney. Little is known about the protein's enzymatic activity, its substrates or the proteins with which it associates. Classes I, II, and IV HDACs are all zinc-dependent enzymes, sometimes referred to as "classical HDACs". In contrast, class III HDACs require a nicotinamide adenine dinucleotide ion ($NAD^+$) for their catalytic activity and are homologous to the NAD dependent HDAC SIR2 yeast protein [76]. These proteins, called sirtuin, share no sequence or structural similarity with the classical HDACs and deacetylate proteins by a mechanism that is distinct from the acetyl-lysine hydrolysis mechanism utilized by classical HDACs [77].

## 2.4 Role of Histone Deacetylase Inhibitors in AD

With a single exception, as discussed later, small-molecule HDAC inhibitors investigated for the potential treatment of neurodegenerative disorders do not affect sirtuin (class III HDAC) activities. These compounds, which considerably differ in potency and HDAC isoenzyme selectivity, belong to four different classes, namely the short chain fatty acids, hydroxamic acids, benzamide and cyclic tetra peptides [78]. The fatty acid group of HDAC inhibitors

comprises the compounds SB, phenyl butyrate (PB) and Valproic acid (VPA). In one study it was linked that the previously reported anti proliferative action of SB to histone hyper acetylation suggesting an HDAC inhibitory function [79].

Today, the SB derivative PB is an orphan drug which achieved FDA approval for the treatment of urea cycle disorders. VPA, a commonly used anticonvulsant and mood stabilizer, achieved FDA approval in 1987 and its HDAC inhibitory function was discovered in 2001 [80]. Using HDACs isolated from rat liver, SB and VPA inhibit total HDAC activity incompletely and at comparatively high millimolar concentrations, which is in line with a pronounced class I selectivity shown for VPA. This is somewhat in contrast to the observation that VPA and SB alter histone acetylation levels already at submillimolar doses in single cell cultures, which suggests that mechanism(s) other than direct interference with the catalytic activity of HDAC isoenzymes are involved [81]. Indeed, Kramer and co-workers have shown that VPA and SB (but not TSA and SNDX-275, reduce HDAC2 (class I) protein levels. For VPA, this effect has been shown to be based on proteasomal degradation of HDAC2 but not of other class I HDAC isoenzymes [82].

Transcriptome-wide studies mainly performed using neoplastic cells have shown that HDAC inhibitors can affect the transcription levels of 7 – 10% of all genes. However, there is increasing evidence that these changes are not solely based on histone hyper acetylation due to the observation that HATs and HDACs are not just for histones. Numerous non-histone protein targets of HDACs including transcription factors and regulators, signal transduction mediators, DNA repair enzymes, nuclear import regulators, chaperone proteins, structural proteins, inflammation mediators and viral proteins have been identified which are likely to contribute to the hitherto reported effects induced by HDAC inhibitors [83]. An additional mechanism of action is given by the finding that HDAC inhibitors counteract gene silencing by DNA methylation mediated by methyl-CpGbinding protein 2 (MeCP2). MeCP2 binds tightly to chromatin in a methylation-dependent manner and associates with a corepressor complex containing HDAC1 and HDAC2, suggesting that the fundamental mechanisms of epigenetic gene regulation, DNA methylation and histone acetylation, are linked by MeCP2 [84]. Jones and co-workers demonstrated that gene silencing conferred by MeCP2 and methylated DNA can be relieved by HDAC inhibition using the pan-HDAC inhibitor TSA. A striking observation is that the consequences of HDAC inhibition are not limited to changes in protein acetylation but may also bring about changes in the state of DNA methylation [85].

The neurotoxic p25 derives from proteolysis of p35, and over activates the tau kinase cyclin-dependent kinase 5. Transgenic mice expressing high levels of p25 exhibit hyper phosphorylation of tau as seen in AD, and neurodegeneration [86]. By employing a tetracycline inducible mouse model (CK-p25) expressing p25 under control of a neuron-specific promoter, in one study it was demonstrate that recovery of learning and memory is associated with chromatin remodeling. In CK-p25 mice, p25 causes severe cognitive defects (fear conditioning, spatial learning) associated with neuronal loss and tangle formation (cortex, hippocampus) similar to AD pathology [87]. Very similar to the RSTS mouse model, application of an HDAC inhibitor (SB, 1200 mg/kg/day, i.p.) mitigated learning and memory defects in CK-p25 mice. These effects were not restricted to diseased CK-p25 animals. Application of either SB (100 ng) or TSA (50 ng) into the brain ventricles of wild-type mice following fear conditioning significantly improved their long-term memory abilities when tested 24 h later [88]. An intriguing observation is that environmental enrichment induces histone H3/H4 hyper acetylation in wild-type mice and normalizes memory performance in diseased CK-p25 mice [89].

Gene profiling studies conducted using RNA isolated from the brains of patients with Huntington's disease and from different Huntington's disease mouse models have revealed alterations in the expression of a large number of genes. In the mouse models, many of these alterations occur before symptoms become obvious, suggesting that they play a causal role in disease pathogenesis [90]. In contrast to Huntington's disease and other neurodegenerative disorders in which transcriptional dysregulation and alterations in acetylation patterns are well described, evidence for a causal role of transcriptional dysfunction in the pathogenesis of the more prevalent neurodegenerative disorders (e.g., AD, PD, amyotrophic lateral sclerosis and stroke-induced brain injury) is somewhat limited. Despite the lack of mechanistic rationale, however, HDAC inhibitors have been used in experimental models of a large number of neurodegenerative disorders, and with great success. Indeed, their effectiveness might well be the best evidence for transcriptional dysfunction in most neurodegenerative diseases.

The utility of HDAC inhibitors as potential therapeutic agents for neurodegenerative disorders was first described in a Drosophila model of Polyglutamine expansion disease. The best studied Polyglutamine disorder is Huntington's disease, which is caused by glutamine expansion within the coding region of the huntingtin gene (HTT). Mutant Htt has been shown to disrupt transcription through different mechanisms. For example, it inhibits the activity of

CREB-binding protein, a histone acetyl transferase with global effects on transcription, through direct interaction and sequestration [91]. Mutant Htt also interacts with transcription factors such as Sp1, which regulates the transcription of a large number of genes. Indeed, known targets of Sp1 display decreased expression in human Huntington's disease and in mouse models thereof. In addition to these mechanisms involving protein–protein interaction, polyQ-expanded Htt can alter post-transcriptional modifications of histones resulting in the condensation of chromatin to a more repressed conformation [92]. Vorinostat and butyrate were shown to arrest ongoing progressive neuronal degeneration and reduce lethality induced by the polyglutamine domain of Htt (polyQ-Htt) or just a polyglutamine peptide (Q48). Even when administered to flies already exhibiting neurodegeneration, HDAC inhibitors reduced further neurodegeneration. A similar reduction in the extent and rate of neurodegeneration was observed when a partial loss of function mutant form of Drosophila Sin3A, a corepressor protein that is part of active HDAC complexes, was overexpressed in the flies [93]. Because this genetic manipulation could be expected to reduce HDAC activity, it suggested that the HDAC inhibitors act by inhibiting the activity of HDACs as opposed to affecting cellular processes other than the deacetylase pathways. Drosophila expresses five different classical HDACs (Rpd3, Hdac3, and HDAC4, HDAC6 and histone deacetylase X, an orthologue of HDAC11) and it was not known which HDACs were responsible for the neuroprotective effects of vorinostat and butyrate. In a more recent study, the same group reported that polyQ-Htt-overexpressing flies with a reduced level of Rpd3 suffered less neurodegeneration and lived longer than those with normal Rpd3 expression [94]. Reduced expression of each of the other four HDACs using multiple heterozygous mutations and/or short hairpin-type RNA constructs had no beneficial effect on neurodegeneration or life span. These observations suggest that pharmacological inhibitors that selectively target human orthologues of Rpd3 would be excellent therapeutic agents for neurodegenerative disorders [95].

# 3. METHODOLOGY

## 3.1 Protein-Protein Interaction Network of HDAC and Key Proteins in AD

The numerous cellular associations and interactions that happen between proteins are at the center of cell behavior and their systematic depiction gives setting in molecular biological science. The STRING database (http://string-db.org) expects to give a basic evaluation and assimilation of protein–protein interactions, including direct (physical) and in addition aberrant (useful) affiliations [96]. The newfangled form 10.0 of STRING covers in excess of 2000 creatures, which has required novel, adaptable calculations for exchanging cooperation data between living beings. The protein for protein-protein interaction were selected on the basis of their biological function in the pathogenesis of AD. It will gives insight of possible effect of HDAC on key regulatory proteins of neuronal survival and cell apoptosis.

## 3.2 Dataset Preparation

For dataset collection high-resolution co-crystallized class I and class II protein structures were extracted from RCSB PDB database [98]. These structures additionally went through the dock prep section of UCSF Chimera 1.10 [99] on account of including charges and missing residues. Active sites and conserved domain information of these structures were acquired from PDBSum [100].

| Protein Name | PDB ID | Resolution | Year |
|---|---|---|---|
| **HDAC1** | 5ICN | 3.3 | 2016 |
| **HDAC2** | 5IWG | 1.66 | 2016 |
| **HDAC3** | 4A69 | 2.06 | 2012 |
| **HDAC8** | 5FCW | 1.91 | 2016 |
| **HDAC4** | 2VQV | 3.3 | 2008 |
| **HDAC5** | 5UWI | 2.14 | 2017 |
| **HDAC6** | 5EEN | 1.8 | 2016 |
| **HDAC7** | 3ZNS | 2.4 | 2013 |
| **HDAC10** | 5TD7 | 2.85 | 2017 |

**Table 2: Protein Structure Data with PDB ID, Resolution, and Year of Publication of Individual HDAC.**

## 3.3 Multiple Sequence Alignment (MSA)

Binding and catalytic mode information of class I and class II HDAC has been studied from different scientific sources. To distinguish between the structural similarity and identity between different class protein sequences were obtained from NCBI protein database [101] and for MSA ClustalW [102] tool was used with a gap open penalty of 10 and gap extension penalty of 0.05. BLOSUM [103] was used as a weight matrix in multiple sequence alignment.

## 3.4 Pharmacophore Mapping and QSAR Studies

For Pharmacophore mapping and Quantitative Structural Activity Relationship (QSAR) studies seven known HDAC inhibitors were identified through various literature sources. Pharmacophore mapping of known inhibitors was performed using PharmMapper [104] with default parameters. Fit score and z' score was carefully calculated and role playing binding sites such as hydrophobic sites, hydrogen donor sites, and hydrogen acceptor sites were characterized. For performing QSAR studies of seven known inhibitors, compounds were subjected to EMBL-EBI ChEMBL database [105], ChemAxon's Chemicalize [106]. While 2D/3D properties, molecular descriptors, and bioactivities of compounds were extricated from Marvin Sketch [106], Marvin View [106], and Molinspiration [107] [108].

## 3.5 Machine Learning (ML) Approach in Classification of Compounds

For ML model creation, the training set was obtained from Binding Database [109] which were classified into inhibitors and non-inhibitors based upon IC50 value. Compounds having the highest inhibitory effect that is IC50 $< 2500$ nM were classified as inhibitor/actives while the compounds having weak inhibitory potential that is IC50 $> 2500$ nM were classified as non-inhibitor/decoys. For the preparation of test set known selective and non-selective synthetic molecules were obtained from ChEMBL database. Molecular descriptors of both training set as well as test set were computed through ACD ChemSketch [110] and DataWarrior [111]. Total 19 attributes were calculated on the basis of which test set data were classified as inhibitor or non-inhibitor. Machine learning models were applied to test set data in order to compute their binding effect with the help of Weka 3.8 [112]. Total four models were created which are as follow random forest model, deep learning model, logistic model, and k-star model. After classification, the data sorted out to reduce the number of

compounds in order to obtain novel inhibitor having high binding energy with both class I and class II HDAC.

## 3.6 Molecular Docking (MD) Technique

Molecular docking studies were performed to check the binding relation between the ligand and receptor. Molecular docking studies of the top-ranked compounds with both class I and class II HDAC were performed with SwissDock [113] default parameters. Total binding energy and binding pose of ligand to specific HDAC were examined carefully. Finally, ADMET analysis of compounds was predicted using AdmetSAR [114]. Toxic effects of compounds were computed using DataWarrior and Molinspiration.

## 3.7 Ammino Acid Substitution Analysis

Permanent alteration in the nucleotide sequence of genome termed as a mutation in protein structure causes the dysfunction or dysregulation of protein activity and structure. The catalytic sites or binding sites of different Histone Deacetylase possess missense mutation as well as silent coding mutation. The data of ammino acid mutation in HDAC were retrieved from the COSMIC database [115]. UCSF Chimera was used to carry out the ammino acid substitution using the command swapaa (aa name): (position) (chain) or it can be done with the help of option under tools named as morph conformations. The modified structures was saved into Pdb file separately to carry out the molecular docking studies with SwissDock.

# 4 RESULTS

## 4.1 Network Analysis of HDAC and Proteins Involved In AD

Protein-protein interaction (PPI) networks enable to establish mutual effect of two or more highly similar class of protein. In AD pathogenesis many protein including α-synuclein (SNCA), Apolipoprotein E (ApoE), Amyloid β precursor protein (APP), Beta-secretase 1 (BACE), Presenilin-1 (PS-1), Tumor necrosis factor (TNFα), p35, Mitogen-activated protein kinase 1 (MAPK1), microtubule-associated protein tau (MAPT), and Glycogen synthase kinase 3 beta (GSK3β). Several studies in the past suggested the association of these proteins with different isoenzymes of HDAC during the pathogenesis of AD or neuronal cell demise. Impairment between the bindings of these proteins with HDAC's leads to deregulate their function which ultimately cause AD.

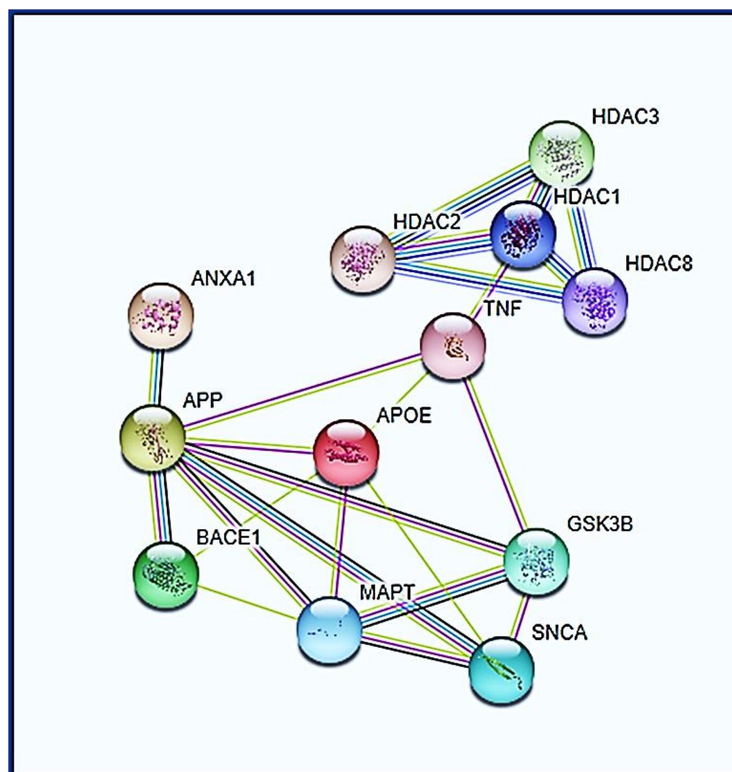


**Figure 5: Interaction Between Class I HDAC And Proteins That Regulate Brain Functions. HDAC1 Is Directly Associated With TNF and Causes Its Acetylation Which Results in Accumulation of APP, Overexpression of ApoE, and Inhibits GSK3β Pathway.**

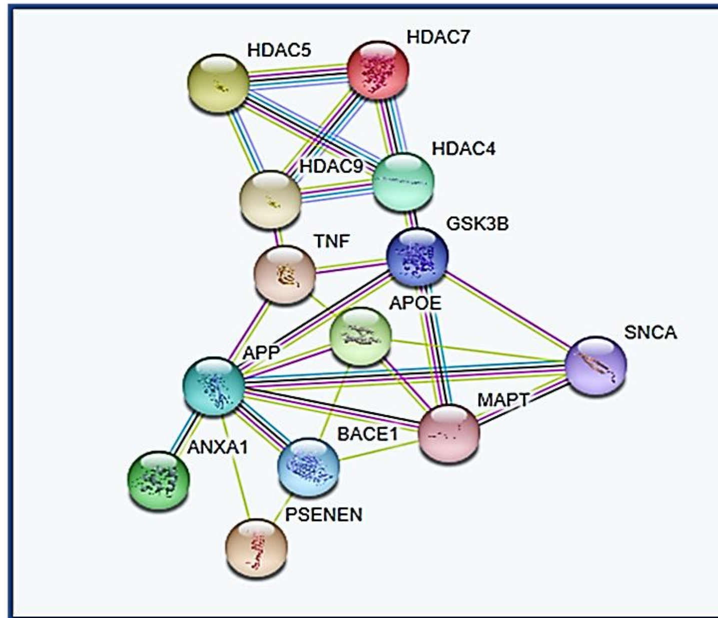


**Figure 6: Interaction between Class IIa HDAC and Protein Alter Learning Ability. HDAC4 Inhibits GSK3β Pathway Directly and HDAC9 Causes Hyper Phosphorylation of TNF. HDAC7 And HDAC5 Does Not Directly Linked With Neurodegeneration Rather They Interact With Either HDAC9 Or HDAC4 Causes Neuronal Cell Apoptosis.**

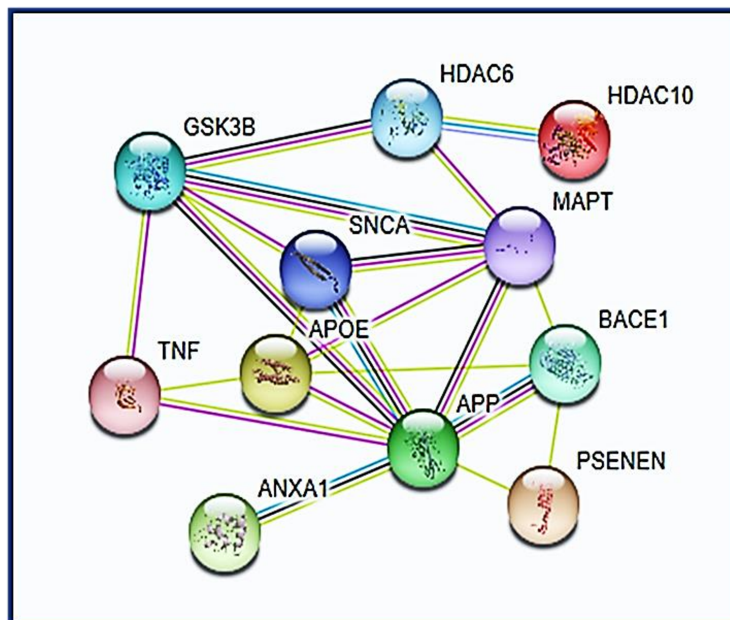


**Figure 7: Protein-Protein Interaction between Class IIb and Proteins Causes Memory Impairment and Decreases Learning Ability. HDAC6 Inhibits GSK3β Pathway And Causes Deacetylation Of MAPT Which Hyperphosphorylated Tau Protein Leading Neuronal Cell Death.**

| | Number of Nodes | Number of Edges | Average Node Degree | Local Clustering Coefficient | Expected Number of Edges | PPI p-Value |
|---|---|---|---|---|---|---|
| Class I | 12 | 23 | 3.83 | 0.755 | 9 | 0.00013 |
| Class IIa | 13 | 26 | 4 | 0.684 | 7 | 2.63-08 |
| Class IIb | 11 | 21 | 3.82 | 0.681 | 5 | 1.91E-08 |

**Table 3: Data associated with protein-protein interaction network of Class I, Class IIa, and Class IIb.**

## 4.2 Identity and Similarity between Different HDAC Using MSA

| Type | Active Site | Ammino Acids Residues | Conserved Domain |
|---|---|---|---|
| **Hdac1** | 141, 142, 149, 150, 176, 178, 264, 261, 301, 303 | H, H, G, F, D, H, D, L, G, Y | 140-303 |
| **Hdac2** | 14, 142, 150, 151, 177, 179, 265, 272, 302, 304 | H, H, G, F, D, H, D, L, G, Y | 141-304 |
| **Hdac3** | 134, 135, 143, 144, 170, 172, 259, 266, 296, 298 | H, H, G, F, D, H, D, L, G, Y | 134-298 |
| **Hdac4** | 802, 803, 811, 812, 840, 842, 934, 974 | H, H, G, F, D, H, D, G | 802-974 |
| **Hdac5** | 832, 833, 841, 842, 870, 872, 954, 1004 | H, H, G, F, D, H, D, G | 832-1004 |
| **Hdac6** | 215, 216, 224, 225, 253, 255, 346, 384 | H, H, G, Y, D, H, D, G | 610-780 |
| **Hdac7** | 542, 626, 669, 670, 678, 679, 707, 709,738, 801, 809, 810, 841 | P, D, H, H, G, F, D, H, F, D, P, L, G | 542-841 |
| **Hdac8** | 100, 101, 141, 142, 143, 151, 152, 178, 180, 208, 267, 274, 304, 306 | Y, D, W, H, H, G, F, D, H, F, D, M, G, Y | 100-306 |
| **Hdac9** | 782, 783, 791, 792, 820, 822, 914, 954 | H, H, G, F, D, H, D, G | 782-954 |
| **Hdac10** | 134, 135, 143, 144, 172, 174, 265, 303 | H, H, G, F, D, H, D, G | 134-303 |
| **Hdac11** | 41, 140, 141, 142, 143, 151, 152, 153, 181, 183, 257, 261, 301, 302, 304 | K, G, F, H, H, G, F, C, D, H, N, D, S, G, Y | 41-304 |

**Table 4: The Following Table Describes The Active Site Position, Ammino Acid Residues, And Conserved Domain In Different Isoforms Of HDAC. According To Table Histidine, Glycine, Tyrosine, Aspartic Acid, and Leucine Were Most Important Ammino Acid That Take Part In Binding With Small Biomolecules or Inhibitors.**

Binding site analysis of class I and class II showed that residues ranging from position 140 to position 250 considered as the active site for class I with ammino acids such as Histidine, Glycine, and Leucine as active ammino acids. While residues ranging from position 800 to position 900 were considered as the active site for class IIa and residues ranging from

position 150 to position 300 were considered as the active site for class IIb. Studies of catalytic sites explained that Histidine, Glycine, Aspartic acid, and Phenylalanine were important catalytic ammino acids for HDAC class II. $Zn^{2+}$ has centrally formed a coordination bond with one Histidine and two Aspartic acid, respectively for both classes I as well as class II HDAC.

```
Q9BY41.2        MEEPEEPADSGQSLVPVYIYSPEYVSM--CDSLAKIPKRASMVHSLIEAYALHKQMRIVK   58
O15379.2        -----------MAKTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFK   49
NP_004955.2     -----MAQTQGTRRKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYR   55
Q92769.2        ----MAYSQGGGKKKVCYYYDGDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYR   56
                           * *. :  .     .    *:*  :.*.*:  *.*:::* : :
Q9BY41.2        PKVASMEEMATFHTDAYLQHLQKVSQEGDDDHPD-SIEYGLGYDCPATEGIFDYAAAIGG  117
O15379.2        PYQASQHDMCRFHSEDYIDFLQRVSPTNMQGFTKSLNAFNVGDDCPVFPGLFEFCSRYTG  109
NP_004955.2     PHKANAEEMTKYHSDDYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSTG  115
Q92769.2        PHKATAEEMTKYHSDEYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSTG  116
                *  *. .:*  :*:: *:..*: :   . . . .    :.:* ***.  *:*::.    *
Q9BY41.2        ATITAAQCLIDGMCKVAINWSGGWHHAKKDEASGFCYLNDAVLGILRLRRKFERILYVDL  177
O15379.2        ASLQGATQLNNKICDIAINWAGGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDI  169
NP_004955.2     GSVASAVKLNKQQTDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDI  175
Q92769.2        GSVAGAVKLNRQQTDMAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDI  176
                .:: .*  *     .:*:**:** ***** *******:** *:.**.* : . *:**:*:
Q9BY41.2        DLHHGDGVEDAFSFTSKVMTVSLHKFSPGFFPGTGDVSDVGLGKGRYYSVNVPIQDGIQD  237
O15379.2        DIHHGDGVQEAFYLTDRVMTVSFHKYGNYFFPGTGDMYEVGAESGRYYCLNVPLRDGIDD  229
NP_004955.2     DIHHGDGVEEAFYTTDRVMTVSFHKYGEY-FPGTGDLRDIGAGKGKYYAVNYPLRDGIDD  234
Q92769.2        DIHHGDGVEEAFYTTDRVMTVSFHKYGEY-FPGTGDLRDIGAGKGKYYAVNFPMRDGIDD  235
                *:******::**  *.:*****:**:.   ******: ::*  .*:**.:* *::***:*
Q9BY41.2        EKYYQICESVLKEVYQAFNPKAVVLQLGADTIAGDPMCSFNMTPVGIGKCLKYILQWQLA  297
O15379.2        QSYKHLFQPVINQVVDFYQPTCIVLQCGADSLGCDRLGCFNLSIRGHGECVEYVKSFNIP  289
NP_004955.2     ESYEAIFKPVMSKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLP  294
Q92769.2        ESYGQIFKPIISKVMEMYQPSAVVLQCGADSLSGDRLGCFNLTVKGHAKCVEVVKTFNLP  295
                :.*  : : ::.:* : ::*..:*** *:*::. * : .**:  * .:*::  * .:*:: :
Q9BY41.2        TLILGGGGYNLANTARCWTYLTGVILGKTLSSEIPDHEFFTAYGPDYVLEITPSCRPD-R  356
O15379.2        LLVLGGGGYTVRNVARCWTYETSLLVEEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQ  349
NP_004955.2     MLMLGGGGYTIRNVARCWTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPS-NMTNQ  353
Q92769.2        LLMLGGGGYTIRNVARCWTYETAVALDCEIPNELPYNDYFEYFGPDFKLHISPS-NMTNQ  354
                 *:******.: *.****** *.: :  : .*:* ::*  :.**: *.   * .  :
Q9BY41.2        NEPHRIQQILNYIKGNLKHVV---------------------------------------  377
O15379.2        NSRQYLDQIRQTIFENLKMLNHAPSVQIHDVPADLLTYDRTDEADAEERGPEENYS----  405
NP_004955.2     NTNEYLEKIKQRLFENLRMLPHAPGVQMQAIPEDAIPEESGDEDEDD---PDKRISICSS  410
Q92769.2        NTPEYMEKIKQRLFENLRMLPHAPGVQMQAIPEDAVHEDSGDEDGED---PDKRISIRAS  411
                *  . :::* : :  **: :
Q9BY41.2        ------------------------------------------------------------  377
O15379.2        --RPEAPNEFYDGDHDNDK--------------------------ESDVEI--------  428
NP_004955.2     DKRIACEEEFSDSEEEGEGGRKNSSNFKKAKRVKTEDEKEKDPE-EKKEVTEEEKTKEE-  468
Q92769.2        DKRIACDEEFSDSEDEGEGGRRNVADHKKGAKKARIEEDKKETEDKKTDVKEEDKSKDNS  471

Q9BY41.2        -----------------  377
O15379.2        -----------------  428
NP_004955.2     --KPEAKGVKEEVKLA-  482
Q92769.2        GEKTDTKGTKSEQLSNP  488
```

**Figure 8: Multiple Sequence Alignment of Class I HDAC Isoforms Using ClustalW where * Marks the Identity between The Isoforms and ˚ Donates the Percent Similarity Between Different HDAC.**

Final results of multiple sequence alignment showed that HDAC1 and HDAC2 share 85.06% sequence similarity using the BLOSUM similarity matrix which is highest as compared to

other HDAC's in class I. The percentage of sequence similarity for HDAC1 reduces to 40.54% and 59.1% respectively with HDAC8 and HDAC3. Similarly, in class II the sequence similarity between HDAC5 and HDAC9 stands more at 61.11%. Among class IIb the sequence similarity between HDAC6 and HDAC10 remains at 41.58%. The results explained the similarity and identity between different isoforms of HDAC. Studies showed that class IIa has another putative zinc binding site with Histidine and Cysteine take part in binding with central $Zn^{2+}$ atoms. The catalytic residues in HDAC form a tunnel-shaped pocket which allows the lipophilic part of the inhibitor to form a hydrogen bond with the selective receptor.

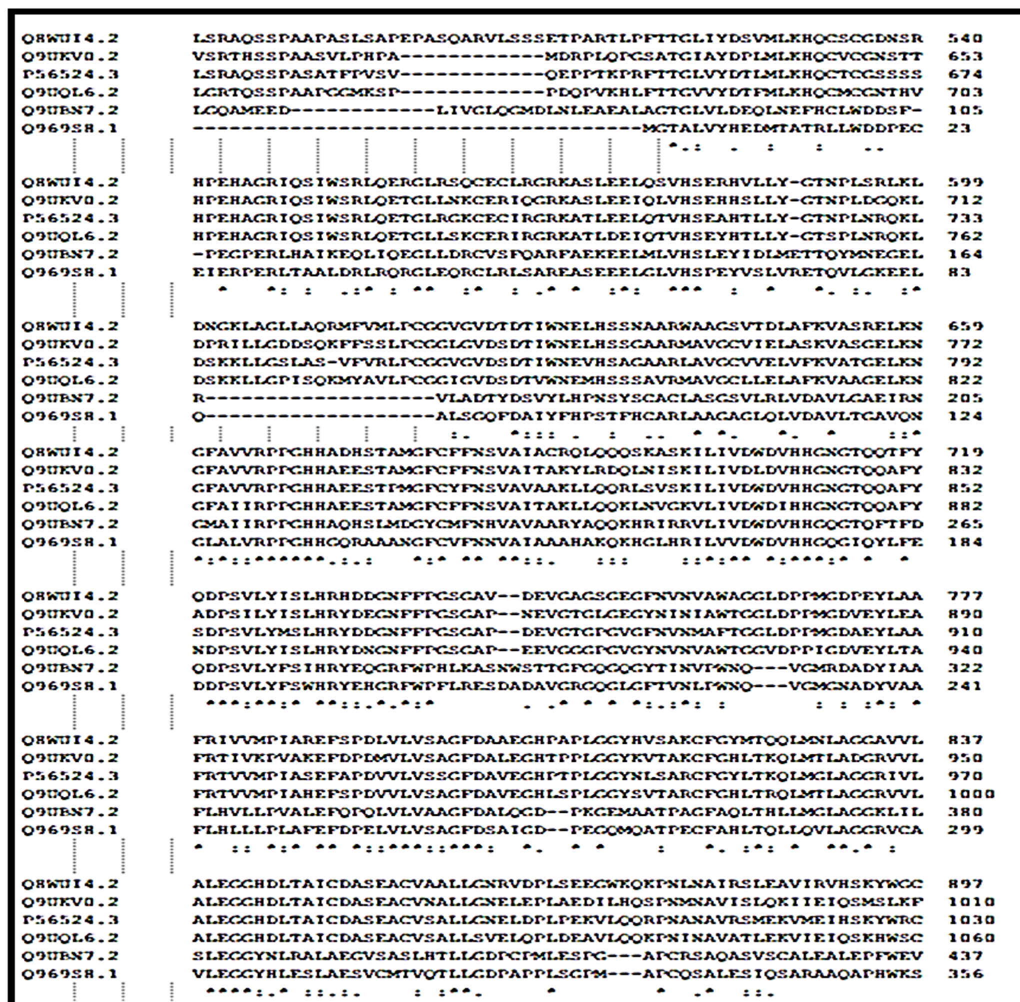

**Figure 9: Multiple Sequence Alignment of Class II HDAC Isoforms Using ClustalW.**

|  | Q9BY41.2 | O15379.2 | NP_004955.2 | Q92769.2 |
|---|---|---|---|---|
| **Q9BY41.2** | 100 | 40.98 | 40.54 | 40.7 |
| **O15379.2** | 40.98 | 100 | 59.1 | 59.1 |
| **NP_004955.2** | 40.54 | 59.1 | 100 | 85.06 |
| **Q92769.2** | 40.7 | 59.1 | 85.06 | 100 |

**Table 5: Percent Identity Matrix of Class I HDAC.**

|  | Q8WUI4.2 | Q9UKV0.2 | P56524.3 | Q9UQL6.2 | Q9UBN7.2 | Q969S8.1 |
|---|---|---|---|---|---|---|
| **Q8WUI4.2** | 100 | 48.24 | 53.37 | 48.29 | 33.77 | 36.75 |
| **Q9UKV0.2** | 48.24 | 100 | 59.69 | 61.11 | 36.66 | 38.42 |
| **P56524.3** | 53.37 | 59.69 | 100 | 61.6 | 36.12 | 36.59 |
| **Q9UQL6.2** | 48.29 | 61.11 | 61.6 | 100 | 35.45 | 34.41 |
| **Q9UBN7.2** | 33.77 | 36.66 | 36.12 | 35.45 | 100 | 41.58 |
| **Q969S8.1** | 36.75 | 38.42 | 36.59 | 34.41 | 41.58 | 100 |

**Table 6: Percent Identity Matrix of Class II HDAC.**

## 4.3 Identification of Novel Compound Using QSAR Technique Followed By Machine Learning Modelling

Pharm mapping studies of six known inhibitors with HDAC8 describe the pharmacophore features which were the important characteristics of a compound to be HDAC inhibitor. Studies explained that there should be 4 hydrophobic sites, 2 hydrogen donor sites, and 1 hydrogen acceptor site present in the compound. Moreover, the fit score should be around 3 and the normalized fit score should be around 0.5 for HDAC inhibitor. Z- score is a score produced from the particle's fit score and a library score grid ascertained in advance . It joins the fit score and its comparing vector in the score grid together and standardizes it to a vector with a mean of zero and a standard deviation of one. Contrasted with the fit score z'- score apply the pharmacophore coordinating strategy as well as think of some as measurement factor lying behind, say, typical dissemination a haphazardly given particle's fit score may take after. By and large, expansive positive z'- score ought to shows high essentialness of the objective to a question compound, also, substantial negative z'- score demonstrates the objective may not be sufficiently huge. By doing this, we trust the z'- score will signifies the unadulterated fit score with more factual significance and certainty contrasting.

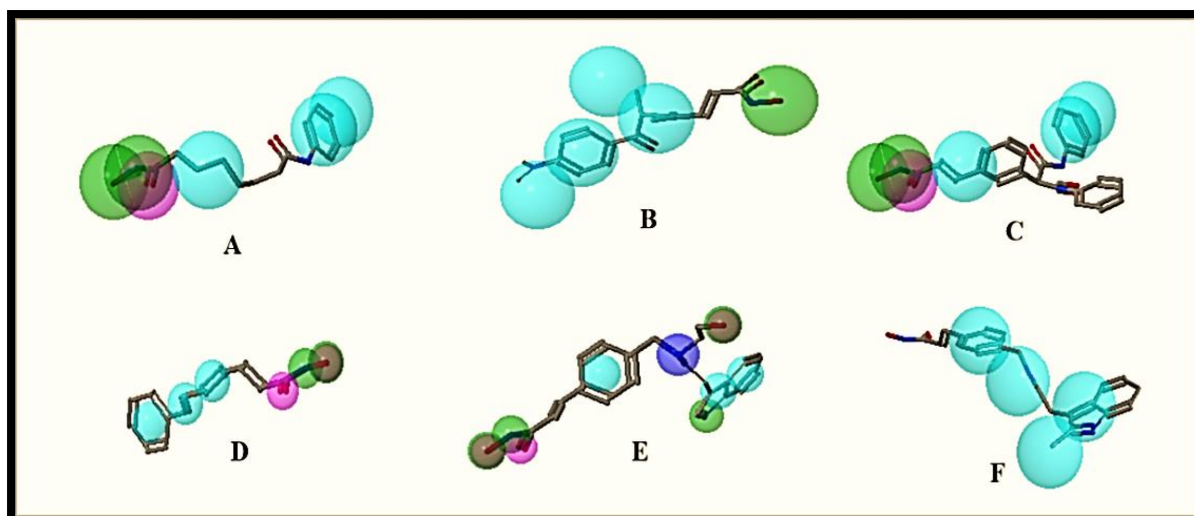


**Figure 10: Pharmacophore Analysis of Known Inhibitors Using PharmMapper. (A) Chembl98 (B) Chembl99 (C) Chembl116620 (D) Chembl159249 (E) Chembl356066 (F) Chembl483254. Blue Color Represents The Hydrophobic Residues, Green Color Represents The Hydrogen Donor Sites, And Pink Color Represents The Hydrogen Acceptor Site Of Receptor.**

In order to create machine learning models, the training set and test set should be carefully examined and for this purpose, molecular descriptor studies of the compounds were carried out. Training compounds were collected from the binding database based upon IC50 values. Total 3234 compounds were used as training data in Weka 3.8 while 5064 compounds were collected as a test set data from various literature sources. The compounds were treated into DataWarrior to calculate total 19 descriptors in order to identify the class of compounds. Total 4 model names as a random forest model, deep learning model, KStar model, and logistic model were created to classify the test set data. Our generated model predicted that 2347 compounds were classified as inhibitors while the remaining compounds were classified as non-inhibitors. To check the reliability and predictability of generating models several statistical parameters such as root mean square error, relative squared error, root relative squared error, and precision was calculated carefully. The compounds classified as inhibitors were further sorted out on the basis of Lipinski's rule of 5 and other characteristics features on the basis of which a small chemical compound classified as drug molecule.

Random forest model is widely accepted and used model because of its simplicity level and reliability. It is based on supervised learning method. The mechanism behind the working of model is to create a forest that is n ot organized and based on bagging phenomenon.it is used

in both classification and regression problems. Deep learning metods are different in nature as they based on learning data and can be supervised method or unsupervised methods. The basic feature of deep learning methods are their similarity to biological nervous system representation. Kstar is a type of ML method based on the input training data and gives output of test data based on previously feed data structure. It is based on distance between the different input attributes and thus more predictale than other similar type of algorithms.

Logistic model is also a type of regression based algorithm and depends on the calculation of attributes value. Logistic model is widely used to calculate the score among traumiatic and injury patients. It is widely accepted mocdel in many cases such as in polling.

| Characteristics | Random Forest | Logistic Model | K-Star Model | Deep Learning Method | Average |
|---|---|---|---|---|---|
| **Accuracy** | 95.82 | 97.98 | 86.95 | 97.4 | 94.5375 |
| **ROC Area** | 0.993 | 0.989 | 0.934 | 0.992 | 0.977 |
| **PRC Area** | 0.993 | 0.985 | 0.938 | 0.992 | 0.977 |
| **Kappa Statistic** | 0.9045 | 0.9543 | 0.7483 | 0.941 | 0.887025 |
| **TP Rate** | 0.958 | 0.98 | 0.89 | 0.974 | 0.9505 |
| **FP Rate** | 0.066 | 0.032 | 0.149 | 0.039 | 0.0715 |
| **Precision** | 0.958 | 0.98 | 0.889 | 0.974 | 0.95025 |
| **Mean Absolute Error** | 0.0951 | 0.0596 | 0.1169 | 0.0662 | 0.08445 |
| **Root Mean Squared Error** | 0.1841 | 0.1476 | 0.3136 | 0.1574 | 0.200675 |
| **Relative Absolute Error** | 21.44% | 13.44% | 26.35% | 14.93% | 19.04% |
| **Root Relative Squared Error** | 39.09% | 31.34% | 66.60% | 33.42% | 42.61% |

**Table 7: Characteristic Features of ML Models Determining The Accuracy of The Models Created.**

From the table above it is clearly identified that accuracy of logistic model method is highest that is 95.6% followed by deep learning method, random forest, and kstar method which is 97.4%, 95.8%, and 86.9% respectively. From the results it may be concluded that overall accuracy of our generated system is more than 94%. thus our prediction is reliable and proceeds for further tests.

Accuracy in Weka is defined as the overall correctly identified instance during the algorithm. In other words it is the ratio of correct predictions out of total predictions made by a particular model. Root mean squared error is defined as the average error made on the test set during the output variable. Area under ROC curve denotes the plot against TPR and FPR or we can say that it is the graph between sensitivity v/s specificity.

**Figure 11: Graphical Comparison of The Characteristics Features of The Four Models Created Using ML Technique.**

| | 1834473 | 2312466 | 256322 | 271308 | 3353068 | 3648287 | 3648298 | 3655942 | 3655956 | 474746 | 490722 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Deriding Energy | 61.07 | 102.46 | 72.6 | 65.33 | 97.56 | 112.96 | 67.81 | 54.83 | 69.07 | 76.23 | 69.56 |
| Van Der Waals Volume | 262.13 | 374.5 | 315.3 | 285 | 313.76 | 419.31 | 317.53 | 282.7 | 300.48 | 351.7 | 336.6 |
| MMFF94 Energy | 119.11 | 151.44 | 138.8 | 93.76 | 258.54 | 205.62 | 177.1 | 137.04 | 153.58 | 167.1 | 136.5 |
| Platt Index | 68 | 102 | 72 | 68 | 84 | 96 | 74 | 68 | 74 | 80 | 80 |
| Randic Index | 11.15 | 16.17 | 12.17 | 11.72 | 13.15 | 15.18 | 12.71 | 11.78 | 12.13 | 13.62 | 13.56 |
| Balban Index | 1.32 | 0.99 | 1.23 | 1.23 | 1.25 | 1.57 | 1.49 | 1.43 | 1.45 | 1.45 | 1.43 |
| Harary Index | 77.59 | 130.82 | 86.99 | 81.32 | 100.73 | 128.7 | 93.15 | 84.69 | 91.83 | 104.6 | 103.4 |
| Hyper Weiner Index | 6075 | 21975 | 8151 | 7342 | 9505 | 13669 | 8045 | 5362 | 5688 | 9668 | 9986 |
| Szeged Index | 2081 | 5958 | 2564 | 2436 | 3308 | 4056 | 2592 | 2139 | 2282 | 3033 | 3082 |
| Wiener Index | 1399 | 3887 | 1771 | 1612 | 2073 | 3048 | 1858 | 1404 | 1508 | 2213 | 2254 |
| Wiener polarity | 32 | 51 | 35 | 33 | 43 | 47 | 37 | 35 | 40 | 41 | 39 |

**Table 8: Molecular Descriptors of Test Set Compounds**

| | 1834473 | 2312466 | 256322 | 271308 | 3353068 | 3648287 | 3648298 | 3655942 | 3655956 | 474746 | 490722 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Total Molweight** | 348.773 | 436.47 | 352.4 | 319.3 | 362.38 | 454.59 | 351.40 | 320.35 | 334.37 | 379.4 | 379.4 |
| **clogP** | 2.241 | 2.0433 | 2.629 | 2.001 | 2.1436 | 2.7118 | 2.1121 | 2.2175 | 2.1407 | 2.994 | 2.046 |
| **clogS** | -3.989 | -3.913 | -3.817 | -3.854 | -2.593 | -2.824 | -2.568 | -3.814 | -3.914 | -3.206 | -2.82 |
| **H-Acceptors** | 8 | 9 | 5 | 6 | 7 | 7 | 6 | 6 | 6 | 6 | 7 |
| **H-Donors** | 3 | 3 | 3 | 3 | 2 | 2 | 3 | 3 | 3 | 3 | 3 |
| **Total Surface Area** | 246.2 | 335.78 | 276.4 | 253.2 | 272.13 | 351.39 | 279.03 | 249.82 | 259.99 | 303.7 | 296.7 |
| **Relative PSA** | 0.46019 | 0.3069 | 0.298 | 0.288 | 0.2646 | 0.2325 | 0.2456 | 0.2842 | 0.2730 | 0.225 | 0.274 |
| **Polar Surface Area** | 141.16 | 123.12 | 108.2 | 92.93 | 87.46 | 106.67 | 87.38 | 87.14 | 87.14 | 87.38 | 104.4 |
| **Shape Index** | 0.69565 | 0.6060 | 0.68 | 0.666 | 0.6296 | 0.5 | 0.6153 | 0.5416 | 0.52 | 0.571 | 0.571 |
| **Metal-Atoms** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Rotatable Bonds** | 4 | 6 | 6 | 5 | 3 | 9 | 7 | 5 | 5 | 8 | 8 |
| **Amides** | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| **GPCR ligand** | -0.34 | 0.12 | -0.07 | 0.05 | 0.28 | 0.3 | 0.25 | 0.01 | 0.21 | 0.11 | 0.33 |
| **Ion channel modulator** | -0.94 | 0.27 | -0.28 | -0.11 | -0.07 | -28 | -0.18 | -0.37 | -0.04 | -0.24 | -0.11 |
| **Kinase inhibitor** | 0.47 | 0.7 | 0.28 | 0.39 | 0.38 | 0.18 | 0.23 | 0.35 | 0.38 | 0.05 | 0.11 |
| **Nuclear receptor ligand** | -0.61 | -0.45 | -0.33 | -0.29 | -55 | -0.03 | 0.07 | -0.27 | 0.09 | 0.05 | 0.17 |
| **Protease inhibitor** | 0.02 | -0.2 | 0 | -0.14 | 0.52 | 0.39 | 0.34 | 0.32 | 0.58 | 0.29 | 0.39 |
| **Enzyme inhibitor** | 0.29 | 0.12 | 0.12 | 0.17 | 0.46 | 0.59 | 0.51 | 0.32 | 0.45 | 0.4 | 0.53 |
| **LE from clogP** | 0.51592 | 0.3612 | 0.470 | 0.497 | 0.4404 | 0.3672 | 0.4577 | 0.4946 | 0.4757 | 0.417 | 0.425 |
| **LLE from clogP** | 6.4086 | 6.6464 | 5.950 | 6.697 | 6.5253 | 5.8549 | 6.5632 | 6.4366 | 6.5287 | 5.529 | 6.642 |
| **LELP from clogP** | 4.3437 | 5.6562 | 5.585 | 4.024 | 4.8666 | 7.3837 | 4.6141 | 4.4827 | 4.4998 | 7.169 | 4.807 |
| **Mutagenic** | none | none | none | none | None | none | None | none | none | None | None |
| **Tumorigenic** | none | none | none | none | None | none | None | none | none | None | None |
| **Reproductive Effective** | none | none | none | none | None | none | None | none | none | None | None |
| **Irritant** | none | none | none | none | None | none | None | none | none | None | None |

**Table 9: Chemical And Physical Properties of Novel Compounds Identified Using DataWarrior.**

Pharmacokinetics and pharmacodynamics properties of predicted compounds were calculated using different standard descriptors to check the efficiency and efficacy of compound as a drug molecule. The properties such as protease inhibitor, enzyme inhibitor, kinase inhibitor, ion channel modulator, GPCR ligand, and nuclear receptor ligand were calculated along with Mutagenicity, Tumorigenic effect, and reproductive effect with the help of Molinspiration.

The results indicated that all the compounds did not show mutagenic, tumorigenic, and reproductive effect.

## 4.4 Molecular Docking Analysis

Comparison between the full fitness of 11 novel compounds to 7 known inhibitors explained that the three compounds such as ChEMBL1834473, ChEMBL271308, and ChEMBL490722 have a higher binding affinity to both class I and class II HDAC due to which they have strong potential to be used as isoform-selective HDAC inhibitors.

| CHEMBLID | HDAC1 | HDAC2 | HDAC3 | HDAC4 | HDAC5 | HDAC6 | HDAC7 | HDAC8 | HDAC10 |
|---|---|---|---|---|---|---|---|---|---|
| CHEMBL1834473 | -2088.59 | -1944.2 | -1678.4 | -1749.1 | -1572.6 | -1424.5 | -1517.5 | -1855.6 | -2986.2 |
| CHEMBL490722 | -2087.795 | -1930.08 | -1669.18 | -1742.56 | -1562.7 | -1414.11 | -1516.32 | -1839.5 | -2950.6 |
| CHEMBL271308 | -2090.9 | -1959.2 | -1684.7 | -1734 | -1566.4 | -1400.4 | -1513.5 | -1863.9 | -2977.7 |
| CHEMBL2312466 | -2087.54 | -1954.1 | -1683.7 | -1736.5 | -1564.7 | -1383.6 | -1514.7 | -1843.9 | -2956.8 |
| CHEMBL256322 | -2037.16 | -1860.7 | -1637.2 | -1694.1 | -1518 | -1363.5 | -1463.5 | -1792.6 | -2922.1 |
| CHEMBL3353068 | -1990.64 | -1836.9 | -1579 | -1637.1 | -1470.5 | -1321.7 | -1422.2 | -1770.7 | -2881.7 |
| CHEMBL3648287 | -2028.31 | -1882.6 | -1611.6 | -1672.2 | -1520.9 | -1368.8 | -1459.7 | -1790.3 | -2895 |
| CHEMBL3648298 | -2055.88 | -1903.1 | -1623.84 | -1716.4 | -1529 | -1385.1 | -1477.2 | -1802.4 | -2934.1 |
| CHEMBL3655942 | -2075.97 | -1920.6 | -1668 | -1742.9 | -1559.2 | -1412.2 | -1506.6 | -1839.3 | -2947.7 |
| CHEMBL3655956 | -2060.5 | -1913.3 | -1649.5 | -1716.6 | -1549.1 | -1401.4 | -1495.8 | -1844.7 | -2938.5 |
| CHEMBL474746 | -2057.79 | -1897.6 | -1650.8 | -1700.9 | -1537.8 | -1390 | -1488.3 | -1814.3 | -2923.1 |

**Table 10: Binding Energy of Novel Compounds with Different Isoforms of Histone Deacetylase.**

Among the known inhibitors entinostat exhibit the highest binding affinity to all forms of HDAC. The three novel compounds have the maximum binding energy or a full fitness as compared to high energy compounds among the known HDAC inhibitors. Table above describes the full fitness or binding affinity of novel compounds with different isoform of HDAC. From the table, it is also observed that the compounds have a highest binding affinity with HDAC10 due to its structural and topographical features.

From the study of Binding pocket and binding site analysis, it is cleared that ChEMBL271308 can be used to inhibit HDAC class I and class IIb but it cannot be used for HDAC class IIa due as the compound is unable to bind with central $Zn^{2+}$ atom which is

characteristics feature as discussed earlier. Also, ChEMBL490722 can be used only with HDAC class IIb due to its specific binding with $Zn^{2+}$ central atom present in the HDAC structure. In class I except HDAC8 other HDAC's like HDAC1, HDAC2, HDAC3 binds with the novel compound to its specific position and in class IIa only HDAC7 shows the potential to binds with the compound with the central $Zn^{2+}$ atom. Docking results of HDAC's with the novel compounds also reflect that HDAC7 have two central $Zn^{2+}$ atom that allows more space for the compound to bind with it specifically.

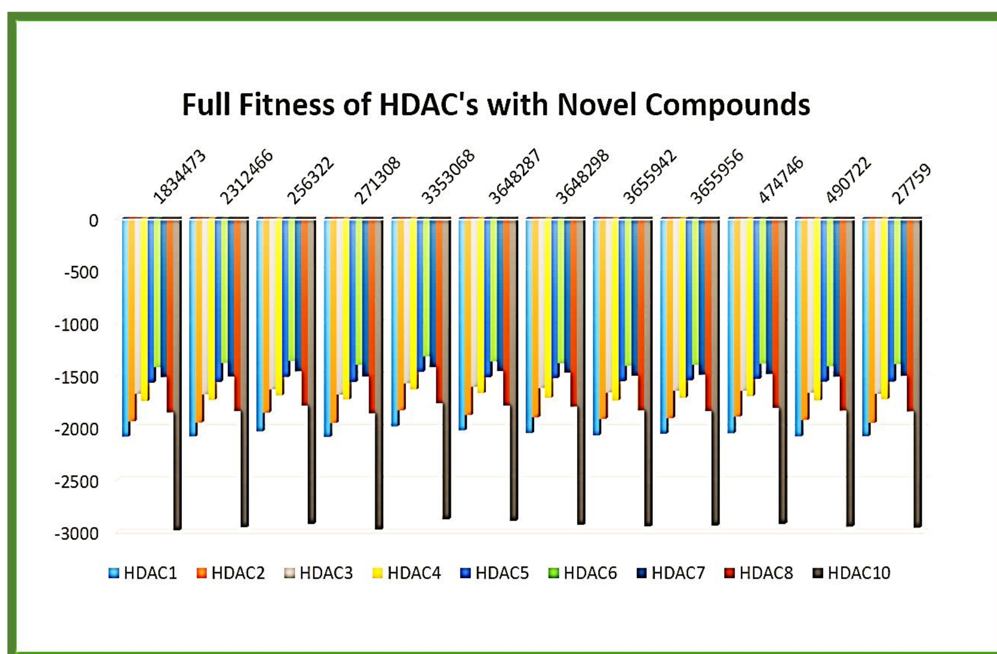


**Figure 12: Graphical Representation of Binding Energy of Novel compounds with Different Types of Histone Deacetylase.**

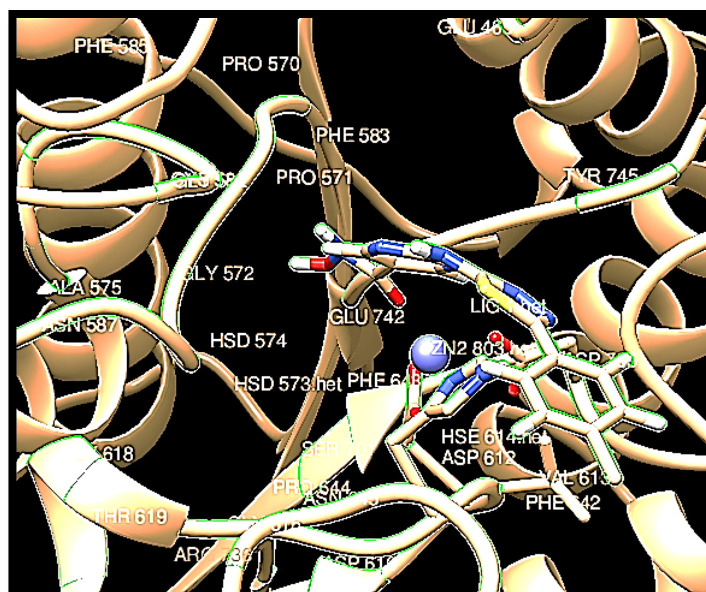


**Figure 13: Docking Pose of HDAC6 with ChEMBL1834473 in which The Ligand is Associated with Central $Zn^{2+}$ Atom and the Residues Take Part in Binding of Receptor with Ligand were GLU742, ASP612, and VAL613.**

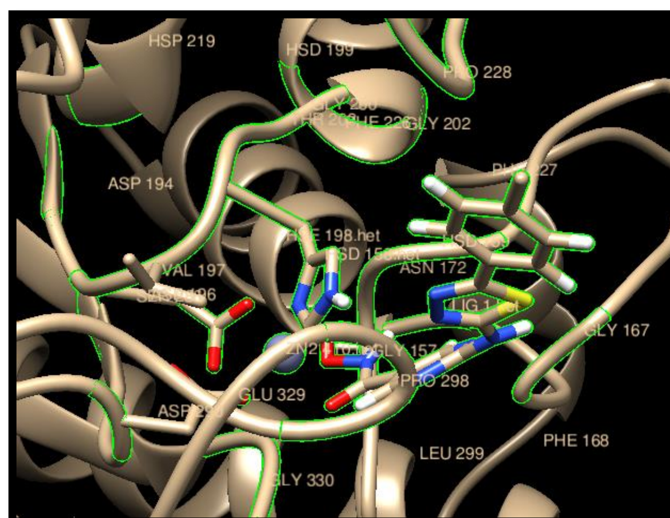


**Figure 14: Binding Pocket of HDAC4 with ChEMBL1834473**

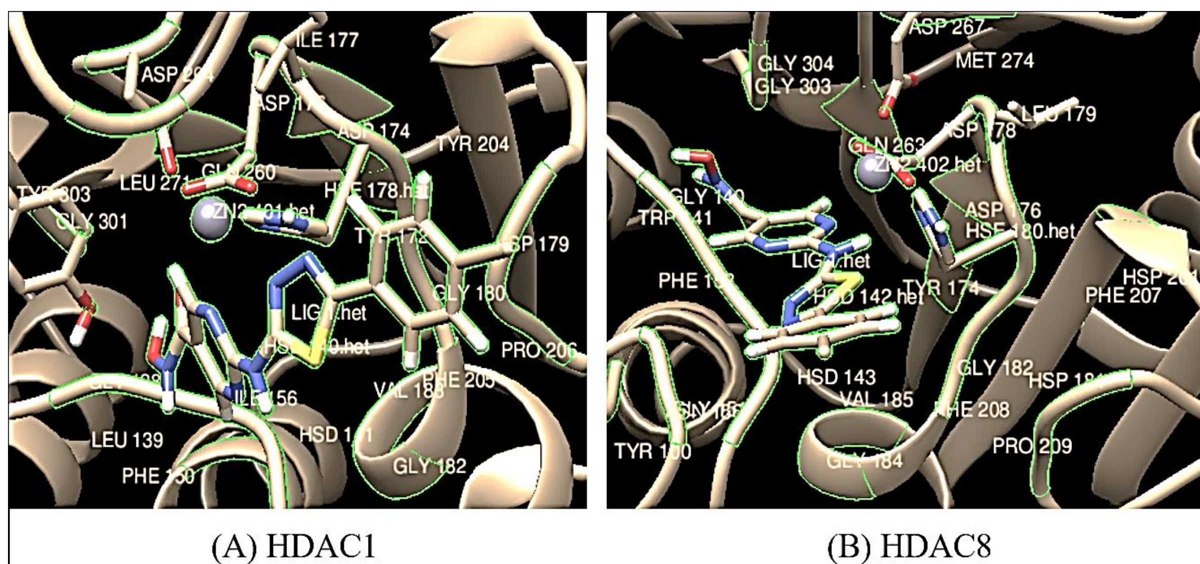


**Figure 15: Binding Pocket of HDAC1 and HDAC8 with ChEMBL1834473**

Another compound named as ChEMBL1834473 possess accurate results which have the ability to binds with class I, class IIa, and class IIb specifically and selectively with the $Zn^{2+}$ atom. Thus from the studies, it can be said that ChEMBL1834473 can be used as isoform-selective inhibitor of HDAC as a neuroprotective agent.

## 4.5 Mutation Analysis

Data obtained from COSMIC database shows critical ammino acid substitution in the ligand-binding pocket of different types of HDAC's. HDAC3, HDAC5, and HDAC7 are the proteins that do not undergo any mutation in its active site whether missense or silent coding mutation.

| Compound ID | HDAC1 | HDAC2 | HDAC4 | HDAC6 | HDAC8 | HDAC10 |
|---|---|---|---|---|---|---|
| **ChEMBL1834473** | **-2078.9** | **-1932.9** | **-1717.5** | **-1382.8** | **-1840.9** | **-2957.9** |
| 98 | -2059.6 | -1917 | -1707.3 | -1370.8 | -1810.1 | -2945.8 |
| 99 | -2018.2 | -1863 | -1663.7 | -1231.6 | -1775.4 | -2895.4 |
| 27759 | -2039.1 | -1927 | -1711.6 | -1352.6 | -1817.2 | -2938.1 |
| 116620 | -2039.1 | -1884.1 | -1683.4 | -1352.3 | -1787.5 | -2890.3 |
| 159249 | -2026.3 | -1879.4 | -1677.4 | -1345.5 | -1780.6 | -2921.4 |
| 356066 | -2002 | -1850.5 | -1647.3 | -1312.1 | -1752 | -2875.3 |
| 483254 | -2025.2 | -1867.5 | -1666.2 | -1336.5 | -1776.1 | -2882.6 |

**Table 11: Binding Energy of ChEMBL1834473 and Currently Used Inhibitors with Different Types of Mutated HDAC.**

Hdac4 shows the maximum number of mutations that is 4 in its ligand binding domain on position such as 167, 168, 198, and 330. HDAC2, HDAC6, HDAC8, and HDAC10 possess one mutated site in its active site while HDAC1 possess two mutated ammino acid residues at active position.

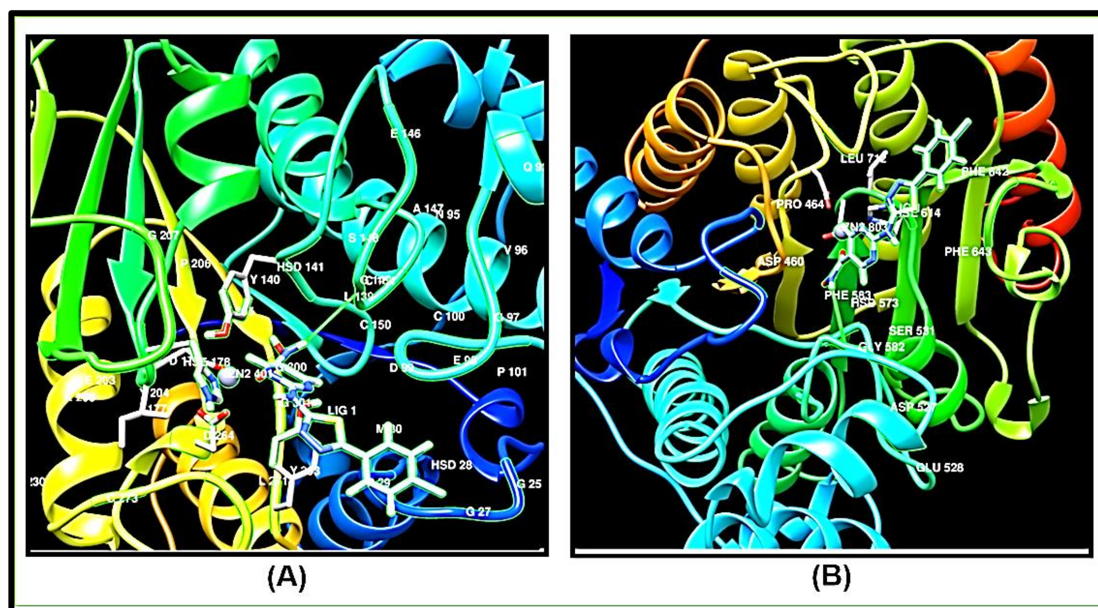


**Figure 16: Binding Pocket of Mutated (A) HDAC1 and (B) HDAC2 with ChEMBL1834473.**

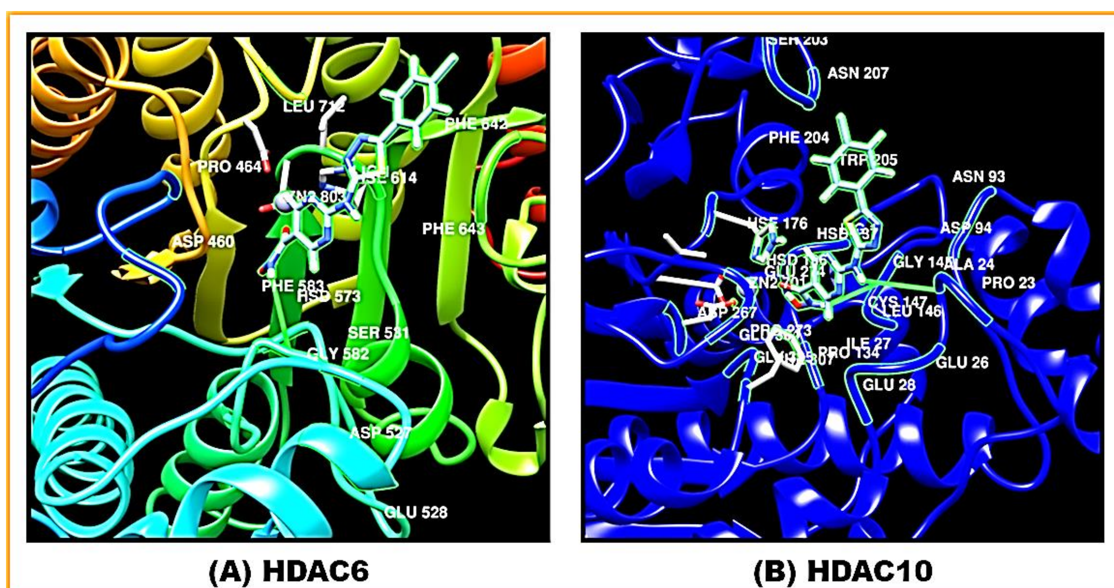


**Figure 17: Binding Pocket of Mutated (A) HDAC6 and (B) HDAC10 with ChEMBL1834473.**

Molecular docking of these mutated structures was carried out with ChEMBL1834473 and 7 approved inhibitors in order to validate whether the novel drug-like compound has the potential to carry out binding process in single nucleotide polymorphism state. Results show that the novel drug-like compound has the higher binding energy of full fitness as compared to the binding energy of seven known inhibitors of HDAC class I and class II. Studies found out that in HDAC4, HDAC8 and HDAC10 mutation causes distortion in the structure due to which binding of several original inhibitors does not takes place in its active site and thus it is unable to facilitate the binding of compound to central $Zn^{2+}$ atom. From the molecular docking results, it can be predicted that the proposed compound has the tendency to bind with the central $Zn^{2+}$ atom in its mutated structure.

# 5 DISCUSSION

Ligand-based drug designing, follows machine learning and molecular docking analysis approach based on the binding or full fitness energy of ligand to the receptor is considered as an effective tool in the drug detection procedure. Here we collected the number of novel compounds for the class I and class II HDAC by creating four machine learning models, for example, logistic model, k-star model, random forest model, and deep learning model based on the Pharmacophore and 2D/3D properties of Trichostatin A, dacinostat, entinostat, Vorinostat, CG-1521, SK683, and Panobinostat. The average accuracy (94.53), precision (0.950), TP rate (0.950), FP rate (0.071), and area under ROC curve of our generated models was quite high as compared to previously created model for selecting a novel isoform-selective HDAC inhibitor. Consequently, 11 hits were identified through their 2D/3D and ADME properties. The protocol applied during the study was quite reliable and efficient as the whole study based on 19 different molecular descriptors of known inhibitors and novel compounds. Also, the data for testing the inhibitor potential includes more than 6000 novel compounds.

Through molecular docking studies, the novel compounds with binding energy or full fitness for both class I and class II HDAC's were selected for comprehensive analysis. The predicted compounds were having the characteristics feature that is a central $Zn^{2+}$ atom in its binding region. Out of these 11 predicted compounds, ChEMBL1834473, ChEMBL490722, and ChEMBL271308 possess the greater binding energy for the class I and class II HDAC. Due to insufficient binding of ChEMBL490722 and ChEMBL271308, these compounds were discarded. ChEMBL1834473 involved in hydrogen bonding with central $Zn^{2+}$ atom and possess the basic pharmacophore sites and hence considered as more potent isoform-selective HDAC inhibitor. Based on the analysis of the catalytic sites it was found out that ChEMBL1834473 forms a hydrogen bonding with ALA142, GLY149, ASN154, ILE156, TYR172, GLY180, GLY182, VAL183, PRO206, and TYR303 residues of HDAC1; GLY157, ASN172, VAL197, PRO298, GLU329, and GLY330 of HDAC4; ASP612, VAL63, HSE614, PHE642, PRO644, SER701, and GLU742 of HDAC6; and HSD143, TYR174, ASP176 GLY184, and VAL185 of HDAC 8. These bonding were carried out because of the conformational modification in the linker section of protein organization.

Moreover to validate the inhibition effect of the compound when protein undergoes mutation, molecular docking of ChEMBL1834473 was carried out with the mutated structure of both

class I and class II HDACs. Single ammino acid replacement or missense mutation took place in the active sites of HDACs except for HDAC3, HDAC5, and HDAC7. Docking energy analysis shows that the compound has quite comparable energy values as compared to the binding energy of 7 known inhibitors of HDAC. However, the proposed compound was able to eliminate the discrepancies caused by the mutation which prevents the binding of known inhibitors to protein structure.

Hence with respect to these results, our proposed novel compound may prove to be an effective therapeutic methodology to Alzheimer's disease. Further studies were required to demarcate the inhibitory potential and effect on disease pathway through in-vitro and in-vivo assays.

# 6 CONCLUSION

We carried out pharmacophoric properties analysis and machine learning approach to generate 4 models of around 6000 novel chemical compounds and classified them as inhibitor or non-inhibitor based upon 19 molecular descriptors. The generated models demonstrated the improved accuracy and predictability of screened hits. Total 11 compounds were selected as HDAC isoform-selective inhibitors based on 2D/3D and ADME properties. The selected compounds were tested for their binding potential through molecular docking method. Docking result analysis stated that three compounds have the higher potential to work as a drug for the proposed target that is HDAC class I and class II. Mutation studies of different types of HDAC also prove the predicted compounds as an effective drug for different HDACs. Thus on the basis of in-silico binding analysis approach, these novel compounds may prove as more potential drugs against class I and class II HDAC as compared to previously known inhibitors to correct the disequilibrium between acetylation and deacetylation activity of HDAC causes life threaten disease named as Alzheimer’s.

# 7 REFERENCES


1. K. Xu, X. Dai, H. Huang and Z. Jiang, "Targeting HDACs: A Promising Therapy for Alzheimer's disease", Oxidative Medicine and Cellular Longevity, vol. 2011, no. 143269, pp. 1-5, 2011.
2. L. Zuo, B. Hemmelgarn, C. Chuang and T. Best, "The Role of Oxidative Stress-Induced Epigenetic Alterations in Amyloid-βProduction in Alzheimer’s Disease", Oxidative Medicine and Cellular Longevity, vol. 2015, no. 604658, pp. 1-13, 2015.
3. J. Hardy, "Has the Amyloid Cascade Hypothesis for Alzheimers Disease been Proved?", Current Alzheimer Research, vol. 3, no. 1, pp. 71-73, 2006.
4. E. Bagyinszky, Y. Youn, S. An and S. Kim, "The genetics of Alzheimer's disease", Clinical Interventions in Aging, p. 535, 2014.
5. A. Peserico and C. Simone, "Physical and Functional HAT/HDAC Interplay Regulates Protein Acetylation Balance", Journal of Biomedicine and Biotechnology, vol. 2011, pp. 1-10, 2011.
6. X. Li, X. Bao and R. Wang, "Neurogenesis-based epigenetic therapeutics for Alzheimer's disease (Review)", Molecular Medicine Reports, vol. 14, no. 2, pp. 1043-1053, 2016.
7. R. Liu, L. Zou and Q. Lü, "Liquiritigenin inhibits Aβ25–35-induced neurotoxicity and secretion of Aβ1–40 in rat hippocampal neurons", Acta Pharmacologica Sinica, vol. 30, no. 7, pp. 899-906, 2009.
8. G. Legube and D. Trouche, "Regulating histone acetyltransferases and deacetylases", EMBO reports, vol. 4, no. 10, pp. 944-947, 2003.
9. T. Richmond and C. Davey, "The structure of DNA in the nucleosome core", Nature, vol. 423, no. 6936, pp. 145-150, 2003.
10. D. Lee, J. Hayes, D. Pruss and A. Wolffe, "A positive role for histone acetylation in transcription factor access to nucleosomal DNA", Cell, vol. 72, no. 1, pp. 73-84, 1993.
11. V. Mutskov, D. Gerber, D. Angelov, J. Ausio, J. Workman and S. Dimitrov, "Persistent Interactions of Core Histone Tails with Nucleosomal DNA following Acetylation and Transcription Factor Binding", Molecular and Cellular Biology, vol. 18, no. 11, pp. 6293-6304, 1998.

12. S. Sidoli, L. Cheng and O. Jensen, "Proteomics in chromatin biology and epigenetics: Elucidation of post-translational modifications of histone proteins by mass spectrometry", Journal of Proteomics, vol. 75, no. 12, pp. 3419-3433, 2012.
13. S. Berger, "The complex language of chromatin regulation during transcription", Nature, vol. 447, no. 7143, pp. 407-412, 2007.
14. I. Gregoretti, Y. Lee and H. Goodson, "Molecular Evolution of the Histone Deacetylase Family: Functional Implications of Phylogenetic Analysis", Journal of Molecular Biology, vol. 338, no. 1, pp. 17-31, 2004.
15. A. RUIJTER, A. GENNIP, H. CARON, S. KEMP and A. KUILENBURG, "Histone deacetylases (HDACs): characterization of the classical HDAC family", Biochemical Journal, vol. 370, no. 3, pp. 737-749, 2003.
16. E. Seto and M. Yoshida, "Erasers of Histone Acetylation: The Histone Deacetylase Enzymes", Cold Spring Harbor Perspectives in Biology, vol. 6, no. 4, pp. a018713-a018713, 2014.
17. R. Broide, J. Redwine, N. Aftahi, W. Young, F. Bloom and C. Winrow, "Distribution of histone deacetylases 1–11 in the rat brain", Journal of Molecular Neuroscience, vol. 31, no. 1, pp. 47-58, 2007.
18. C. Janssen, A. Sarlette, R. Dengler, K. Krampfl and S. Petri, "Distribution of histone deacetylases in post-mortem brain and spinal cord of ALS patients and controls", Aktuelle Neurologie, vol. 35, no. 01, 2008.
19. S. Gregoire, L. Xiao, J. Nie, X. Zhang, M. Xu, J. Li, J. Wong, E. Seto and X. Yang, "Histone Deacetylase 3 Interacts with and Deacetylates Myocyte Enhancer Factor 2", Molecular and Cellular Biology, vol. 27, no. 4, pp. 1280-1295, 2006.
20. N. Blanco-García, E. Asensio-Juan, X. de la Cruz and M. Martínez-Balbás, "Autoacetylation Regulates P/CAF Nuclear Localization", Journal of Biological Chemistry, vol. 284, no. 3, pp. 1343-1352, 2008.
21. O. Kramer, S. Knauer, G. Greiner, E. Jandt, S. Reichardt, K. Guhrs, R. Stauber, F. Bohmer and T. Heinzel, "A phosphorylation-acetylation switch regulates STAT1 signaling", Genes & Development, vol. 23, no. 2, pp. 223-235, 2009.
22. Z. Yuan, "Stat3 Dimerization Regulated by Reversible Acetylation of a Single Lysine Residue", Science, vol. 307, no. 5707, pp. 269-273, 2005.
23. M. Deardorff, M. Bando, R. Nakato, E. Watrin, T. Itoh, M. Minamino, K. Saitoh, M. Komata, Y. Katou, D. Clark, K. Cole, E. De Baere, C. Decroos, N. Di Donato, S. Ernst, L. Francey, Y. Gyftodimou, K. Hirashima, M. Hullings, Y. Ishikawa, C. Jaulin,

M. Kaur, T. Kiyono, P. Lombardi, L. Magnaghi-Jaulin, G. Mortier, N. Nozaki, M. Petersen, H. Seimiya, V. Siu, Y. Suzuki, K. Takagaki, J. Wilde, P. Willems, C. Prigent, G. Gillessen-Kaesbach, D. Christianson, F. Kaiser, L. Jackson, T. Hirota, I. Krantz and K. Shirahige, "HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle", Nature, vol. 489, no. 7415, pp. 313-317, 2012.

24. V. Basile, R. Mantovani and C. Imbriano, "DNA Damage Promotes Histone Deacetylase 4 Nuclear Localization and Repression of G2/M Promoters, via p53 C-terminal Lysines", Journal of Biological Chemistry, vol. 281, no. 4, pp. 2347-2357, 2005.
25. M. Glozak, N. Sengupta, X. Zhang and E. Seto, "Acetylation and deacetylation of non-histone proteins", Gene, vol. 363, pp. 15-23, 2005.
26. J. Jeong, M. Bae, M. Ahn, S. Kim, T. Sohn, M. Bae, M. Yoo, E. Song, K. Lee and K. Kim, "Regulation and Destabilization of HIF-1α by ARD1-Mediated Acetylation", Cell, vol. 111, no. 5, pp. 709-720, 2002.
27. Z. Yuan, L. Peng, R. Radhakrishnan and E. Seto, "Histone Deacetylase 9 (HDAC9) Regulates the Functions of the ATDC (TRIM29) Protein", Journal of Biological Chemistry, vol. 285, no. 50, pp. 39329-39338, 2010.
28. A. Verdel, S. Curtet, M. Brocard, S. Rousseaux, C. Lemercier, M. Yoshida and S. Khochbin, "Active maintenance of mHDA2/mHDAC6 histone-deacetylase in the cytoplasm", Current Biology, vol. 10, no. 12, pp. 747-749, 2000.
29. V. Kekatpure, A. Dannenberg and K. Subbaramaiah, "HDAC6 Modulates Hsp90 Chaperone Activity and Regulates Activation of Aryl Hydrocarbon Receptor Signaling", Journal of Biological Chemistry, vol. 284, no. 12, pp. 7436-7445, 2009.
30. X. Zhang, Z. Yuan, Y. Zhang, S. Yong, A. Salas-Burgos, J. Koomen, N. Olashaw, J. Parsons, X. Yang, S. Dent, T. Yao, W. Lane and E. Seto, "HDAC6 Modulates Cell Motility by Altering the Acetylation Level of Cortactin", Molecular Cell, vol. 27, no. 2, pp. 197-213, 2007.
31. I. Lai, T. Lin, Y. Yao, C. Lin, M. Hsieh and W. Yang, "Histone Deacetylase 10 Relieves Repression on the Melanogenic Program by Maintaining the Deacetylation Status of Repressors", Journal of Biological Chemistry, vol. 285, no. 10, pp. 7187-7196, 2009.
32. H. Cheng, R. Mostoslavsky, S. Saito, J. Manis, Y. Gu, P. Patel, R. Bronson, E. Appella, F. Alt and K. Chua, "Developmental defects and p53 hyperacetylation in

Sir2 homolog (SIRT1)-deficient mice", Proceedings of the National Academy of Sciences, vol. 100, no. 19, pp. 10794-10799, 2003.

33. T. Bouras, M. Fu, A. Sauve, F. Wang, A. Quong, N. Perkins, R. Hay, W. Gu and R. Pestell, "SIRT1 Deacetylation and Repression of p300 Involves Lysine Residues 1020/1024 within the Cell Cycle Regulatory Domain 1", Journal of Biological Chemistry, vol. 280, no. 11, pp. 10264-10276, 2005.
34. B. North, B. Marshall, M. Borra, J. Denu and E. Verdin, "The Human Sir2 Ortholog, SIRT2, Is an $NAD^+$-Dependent Tubulin Deacetylase", Molecular Cell, vol. 11, no. 2, pp. 437-444, 2003.
35. N. R. Sundaresan, S. A. Samant, V. B. Pillai, S. B. Rajamohan, and M. P. Gupta, “SIRT3 Is a Stress-Responsive Deacetylase in Cardiomyocytes That Protects Cells from Stress-Mediated Cell Death by Deacetylation of Ku70,” Molecular and Cellular Biology, vol. 28, no. 20, pp. 6384–6401, 2008.
36. Y. Yang, H. Cimen, M.-J. Han, T. Shi, J.-H. Deng, H. Koc, O. M. Palacios, L. Montier, Y. Bai, Q. Tong, and E. C. Koc, “NAD -dependent Deacetylase SIRT3 Regulates Mitochondrial Protein Synthesis by Deacetylation of the Ribosomal Protein MRPL10,” Journal of Biological Chemistry, vol. 285, no. 10, pp. 7417–7429, 2009.
37. G. Laurent, N. J. German, A. K. Saha, V. C. De Boer, M. Davies, T. R. Koves, N. Dephoure, F. Fischer, G. Boanca, B. Vaitheesvaran, S. B. Lovitch, A. H. Sharpe, I. J. Kurland, C. Steegborn, S. P. Gygi, D. M. Muoio, N. B. Ruderman, and M. C. Haigis, “SIRT4 Coordinates the Balance between Lipid Synthesis and Catabolism by Repressing Malonyl CoA Decarboxylase,” Molecular Cell, vol. 50, no. 5, pp. 686–698, 2013.
38. T. Nakagawa, D. J. Lomb, M. C. Haigis, and L. Guarente, “SIRT5 Deacetylates Carbamoyl Phosphate Synthetase 1 and Regulates the Urea Cycle,” Cell, vol. 137, no. 3, pp. 560–570, 2009.
39. A. Kaidi, B. T. Weinert, C. Choudhary, and S. P. Jackson, “Human SIRT6 Promotes DNA End Resection Through CtIP Deacetylation,” Science, vol. 329, no. 5997, pp. 1348–1353, 2010.
40. O. Vakhrusheva, C. Smolka, P. Gajawada, S. Kostin, T. Boettger, T. Kubin, T. Braun, and E. Bober, “Sirt7 Increases Stress Resistance of Cardiomyocytes and Prevents Apoptosis and Inflammatory Cardiomyopathy in Mice,” Circulation Research, vol. 102, no. 6, pp. 703–710, 2008.

41. L. Gao, M. A. Cueto, F. Asselbergs, and P. Atadja, "Cloning and Functional Characterization of HDAC11, a Novel Member of the Human Histone Deacetylase Family," Journal of Biological Chemistry, vol. 277, no. 28, pp. 25748–25755, 2002.
42. A. Villagra, F. Cheng, H.-W. Wang, I. Suarez, M. Glozak, M. Maurin, D. Nguyen, K. L. Wright, P. W. Atadja, K. Bhalla, J. Pinilla-Ibarz, E. Seto, and E. M. Sotomayor, "Erratum: The histone deacetylase HDAC11 regulates the expression of interleukin 10 and immune tolerance," Nature Immunology, vol. 10, no. 6, pp. 665–665, 2009.
43. G. Donmez, "The neurobiology of sirtuins and their role in neurodegeneration," Trends in Pharmacological Sciences, vol. 33, no. 9, pp. 494–501, 2012.
44. F. Ajamian, A. Salminen, and M. Reeben, "Selective regulation of class I and class II histone deacetylases expression by inhibitors of histone deacetylases in cultured mouse neural cells," Neuroscience Letters, vol. 365, no. 1, pp. 64–68, 2004.
45. E. Hahnen, J. Hauke, C. Tränkle, I. Y. Eyüpoglu, B. Wirth, and I. Blümcke, "Histone deacetylase inhibitors: possible implications for neurodegenerative disorders," Expert Opinion on Investigational Drugs, vol. 17, no. 2, pp. 169–184, 2008.
46. C. Kerridge, N. D. Belyaev, N. N. Nalivaeva, and A. J. Turner, "The Aβ-clearance protein transthyretin, like neprilysin, is epigenetically regulated by the amyloid precursor protein intracellular domain," Journal of Neurochemistry, vol. 130, no. 3, pp. 419–431, Dec. 2014.
47. Y. M. Sung, T. Lee, H. Yoon, A. M. Dibattista, J. M. Song, Y. Sohn, E. I. Moffat, R. S. Turner, M. Jung, J. Kim, and H.-S. Hoe, "Mercaptoacetamide-based class II HDAC inhibitor lowers Aβ levels and improves learning and memory in a mouse model of Alzheimer's disease," Experimental Neurology, vol. 239, pp. 192–201, 2013.
48. K. Krishna, T. Behnisch, and S. Sajikumar, "Inhibition of Histone Deacetylase 3 Restores Amyloid-β Oligomer-Induced Plasticity Deficit in Hippocampal CA1 Pyramidal Neurons," Journal of Alzheimers Disease, vol. 51, no. 3, pp. 783–791, 2016.
49. C. Cook, T. F. Gendron, K. Scheffel, Y. Carlomagno, J. Dunmore, M. Deture, and L. Petrucelli, "Loss of HDAC6, a novel CHIP substrate, alleviates abnormal tau accumulation," Human Molecular Genetics, vol. 21, no. 13, pp. 2936–2945, May 2012.
50. N. Govindarajan, P. Rao, S. Burkhardt, F. Sananbenesi, O. M. Schlüter, F. Bradke, J. Lu, and A. Fischer, "Reducing HDAC6 ameliorates cognitive deficits in a mouse

model for Alzheimers disease,” EMBO Molecular Medicine, vol. 5, no. 1, pp. 52–63, 2012.

51. H. Qing, G. He, P. Ly, C. Fox, M. Staufenbiel, F. Cai, Z. Zhang, S. Wei, X. Sun, C. Chen, W. Zhou, K. Wang and W. Song, "Valproic acid inhibits Ab production, neuritic plaque formation, and behavioral deficits in Alzheimer's disease mouse models", The Journal of Cell Biology, vol. 183, no. 3, pp. i8-i8, 2008.
52. D.-M. Chuang, Z. Wang, and C.-T. Chiu, “GSK-3 as a Target for Lithium-Induced Neuroprotection against Excitotoxicity in Neuronal Cultures and Animal Models of Ischemic Stroke,” Frontiers in Molecular Neuroscience, vol. 4, 2011.
53. K. Ververis and T. C. Karagiannis, “Overview of the Classical Histone Deacetylase Enzymes and Histone Deacetylase Inhibitors,” ISRN Cell Biology, vol. 2012, pp. 1–12, 2012.
54. S. C. Pandey, “Histone Deacetylase Inhibitors,” Encyclopedia of Psychopharmacology, pp. 749–751, 2015.
55. C. Monneret, “Histone Deacetylase Inhibitors,” ChemInform, vol. 36, no. 19, Oct. 2005.
56. P. A. Marks, “Histone Deacetylase Inhibitors: Development as Anticancer Agents,” Apoptosis and Cancer Therapy, pp. 1014–1032.
57. P. W. Finn, “Histone Deacetylase Inhibitors,” Drug Design of Zinc-Enzyme Inhibitors, pp. 859–879.
58. X. Hou, J. Du, R. Liu, Y. Zhou, M. Li, W. Xu, and H. Fang, “Enhancing the Sensitivity of Pharmacophore-Based Virtual Screening by Incorporating Customized ZBG Features: A Case Study Using Histone Deacetylase 8.,” Journal of chemical information and modeling, vol. 55, no. 4, pp. 861–871, Apr. 2015.
59. K. Chen, L. Xu, and O. Wiest, “Computational Exploration of Zinc Binding Groups for HDAC Inhibition,” The Journal of Organic Chemistry, vol. 78, no. 10, pp. 5051–5055, 2013.
60. M. Mahgoub and L. Monteggia, "A role for histone deacetylases in the cellular and behavioral mechanisms underlying learning and memory", Learning & Memory, vol. 21, no. 10, pp. 564-568, 2014.
61. S. Ahmad Ganai, M. Ramadoss and V. Mahadevan, "Histone Deacetylase (HDAC) Inhibitors - emerging roles in neuronal memory, learning, synaptic plasticity and neural regeneration", Current Neuropharmacology, vol. 14, no. 1, pp. 55-71, 2016.

62. K. Saijo, J. Collier, A. Li, J. Katzenellenbogen and C. Glass, "An ADIOL-ERβ-CtBP Transrepression Pathway Negatively Regulates Microglia-Mediated Inflammation", Cell, vol. 145, no. 4, pp. 584-595, 2011.
63. A. Ricobaraza, M. Cuadrado-Tejedor, A. Pérez-Mediavilla, D. Frechilla, J. Del Río and A. García-Osta, "Phenylbutyrate Ameliorates Cognitive Deficit and Reduces Tau Pathology in an Alzheimer's Disease Mouse Model", Neuropsychopharmacology, vol. 34, no. 7, pp. 1721-1732, 2009.
64. J. Tseng, L. Xie, S. Song, Y. Xie, L. Allen, D. Ajit, J. Hong, X. Chen, R. Meeker and T. Cohen, "The Deacetylase HDAC6 Mediates Endogenous Neuritic Tau Pathology", Cell Reports, vol. 20, no. 9, pp. 2169-2183, 2017.
65. S. M. de la Monte, "Brain Insulin Resistance and Deficiency as Therapeutic Targets in Alzheimers Disease", Current Alzheimer Research, vol. 9, no. 1, pp. 35-66, 2012.
66. N. Majdzadeh, L. Wang, B. Morrison, R. Bassel-Duby, E. Olson and S. D'Mello, "HDAC4 inhibits cell-cycle progression and protects neurons from cell death", Developmental Neurobiology, vol. 68, no. 8, pp. 1076-1092, 2008.
67. W. Strittmatter, A. Saunders, D. Schmechel, M. Pericak-Vance, J. Enghild, G. Salvesen and A. Roses, "Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease.", Proceedings of the National Academy of Sciences, vol. 90, no. 5, pp. 1977-1981, 1993.
68. W. XIA, "Etiological factors of Alzheimer disease and recent advances of its treatment", Neural Regeneration Research, vol. 2, no. 2, pp. 107-111, 2007.
69. F. Coppedè, "Advances in the Genetics and Epigenetics of Neurodegenerative Diseases", Epigenetics of Degenerative Diseases, vol. 1, no. 1, 2013.
70. P. Calissano, C. Matrone and G. Amadoro, "Apoptosis and in vitro Alzheimer's disease neuronal models", Communicative & Integrative Biology, vol. 2, no. 2, pp. 163-169, 2009.
71. F. Bardai, V. Price, M. Zaayman, L. Wang and S. D'Mello, "Histone Deacetylase-1 (HDAC1) Is a Molecular Switch between Neuronal Survival and Death", Journal of Biological Chemistry, vol. 287, no. 42, pp. 35444-35453, 2012.
72. K. Ververis and T. Karagiannis, "Overview of the Classical Histone Deacetylase Enzymes and Histone Deacetylase Inhibitors", ISRN Cell Biology, vol. 2012, pp. 1-12, 2012.

73. X. Yang and E. Seto, "The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men", Nature Reviews Molecular Cell Biology, vol. 9, no. 3, pp. 206-218, 2008.
74. L. Gan and L. Mucke, "Paths of Convergence: Sirtuins in Aging and Neurodegeneration", Neuron, vol. 58, no. 1, pp. 10-14, 2008.
75. D. Taylor, M. Maxwell, R. Luthi-Carter and A. Kazantsev, "Biological and Potential Therapeutic Roles of Sirtuin Deacetylases", Cellular and Molecular Life Sciences, vol. 65, no. 24, pp. 4000-4018, 2008.
76. L. Saunders and E. Verdin, "Sirtuins: critical regulators at the crossroads between cancer and aging", Oncogene, vol. 26, no. 37, pp. 5489-5504, 2007.
77. M. Haberland, R. Montgomery and E. Olson, "The many roles of histone deacetylases in development and physiology: implications for disease and therapy", Nature Reviews Genetics, vol. 10, no. 1, pp. 32-42, 2009.
78. E. Hahnen, J. Hauke, C. Tränkle, I. Eyüpoglu, B. Wirth and I. Blümcke, "Histone deacetylase inhibitors: possible implications for neurodegenerative disorders", Expert Opinion on Investigational Drugs, vol. 17, no. 2, pp. 169-184, 2008.
79. J. Cummings, "Disease modification and Neuroprotection in neurodegenerative disorders", Translational Neurodegeneration, vol. 6, no. 1, 2017.
80. C. Phiel, F. Zhang, E. Huang, M. Guenther, M. Lazar and P. Klein, "Histone Deacetylase Is a Direct Target of Valproic Acid, a Potent Anticonvulsant, Mood Stabilizer, and Teratogen", Journal of Biological Chemistry, vol. 276, no. 39, pp. 36734-36741, 2001.
81. M. Gottlicher, "Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells", The EMBO Journal, vol. 20, no. 24, pp. 6969-6978, 2001.
82. O. Kramer, "The histone deacetylase inhibitor valproic acid selectively induces proteasomal degradation of HDAC2", The EMBO Journal, vol. 22, no. 13, pp. 3411-3420, 2003.
83. A. Avila, B. Burnett, A. Taye, F. Gabanella, M. Knight, P. Hartenstein, Z. Cizman, N. Di Prospero, L. Pellizzoni, K. Fischbeck and C. Sumner, "Trichostatin A increases SMN expression and survival in a mouse model of spinal muscular atrophy", Journal of Clinical Investigation, vol. 117, no. 3, pp. 659-671, 2007.
84. F. Yu, "Histone deacetylase-independent transcriptional repression by methyl-CpG-binding protein 2", Nucleic Acids Research, vol. 28, no. 10, pp. 2201-2206, 2000.

85. P. Jones, G. Jan Veenstra, P. Wade, D. Vermaak, S. Kass, N. Landsberger, J. Strouboulis and A. Wolffe, "Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription", Nature Genetics, vol. 19, no. 2, pp. 187-191, 1998.
86. J. Cruz, H. Tseng, J. Goldman, H. Shih and L. Tsai, "Aberrant Cdk5 Activation by p25 Triggers Pathological Events Leading to Neurodegeneration and Neurofibrillary Tangles", Neuron, vol. 40, no. 3, pp. 471-483, 2003.
87. J. Gräff, D. Rei, J. Guan, W. Wang, J. Seo, K. Hennig, T. Nieland, D. Fass, P. Kao, M. Kahn, S. Su, A. Samiei, N. Joseph, S. Haggarty, I. Delalle and L. Tsai, "An epigenetic blockade of cognitive functions in the neurodegenerating brain", Nature, vol. 483, no. 7388, pp. 222-226, 2012.
88. R. Luthi-Carter, "Dysregulation of gene expression in the R6/2 model of polyglutamine disease: parallel changes in muscle and brain", Human Molecular Genetics, vol. 11, no. 17, pp. 1911-1926, 2002.
89. E. Korzus, M. Rosenfeld and M. Mayford, "CBP Histone Acetyltransferase Activity Is a Critical Component of Memory Consolidation", Neuron, vol. 42, no. 6, pp. 961-972, 2004.
90. C. Brochier, M. Gaillard, E. Diguet, N. Caudy, C. Dossat, B. Ségurens, P. Wincker, E. Roze, J. Caboche, P. Hantraye, E. Brouillet, J. Elalouf and M. de Chaldée, "Quantitative gene expression profiling of mouse brain regions reveals differential transcripts conserved in human and affected in disease models", Physiological Genomics, vol. 33, no. 2, pp. 170-179, 2008.
91. R. Butler and G. Bates, "Histone deacetylase inhibitors as therapeutics for polyglutamine disorders", Nature Reviews Neuroscience, vol. 7, no. 10, pp. 784-796, 2006.
92. J. Pallos, L. Bodai, T. Lukacsovich, J. Purcell, J. Steffan, L. Thompson and J. Marsh, "Inhibition of specific HDACs and sirtuins suppresses pathogenesis in a Drosophila model of Huntington's disease", Human Molecular Genetics, vol. 17, no. 23, pp. 3767-3775, 2008.
93. N. Yuniarti, B. Juliandi, T. Sanosaka and K. Nakashima, "Mid-gestational exposure to histone deacetylase inhibitor suberoylanilide hydroxamic acid influence cortical interneuron and astrocyte in mouse brain", Indonesian Journal of Biotechnology, vol. 22, no. 1, p. 32, 2018.

94. M. Minamiyama, "Sodium butyrate ameliorates phenotypic expression in a transgenic mouse model of spinal and bulbar muscular atrophy", Human Molecular Genetics, vol. 13, no. 11, pp. 1183-1192, 2004.
95. E. Thomas, G. Coppola, P. Desplats, B. Tang, E. Soragni, R. Burnett, F. Gao, K. Fitzgerald, J. Borok, D. Herman, D. Geschwind and J. Gottesfeld, "The HDAC inhibitor 4b ameliorates the disease phenotype and transcriptional abnormalities in Huntington's disease transgenic mice", Proceedings of the National Academy of Sciences, vol. 105, no. 40, pp. 15564-15569, 2008.
96. D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. Tsafou, M. Kuhn, P. Bork, L. Jensen and C. von Mering, "STRING v10: protein–protein interaction networks, integrated over the tree of life", Nucleic Acids Research, vol. 43, no. 1, pp. D447-D452, 2014.
97. H. M. Berman, “The Protein Data Bank,” Nucleic Acids Research, vol. 28, no. 1, pp. 235–242, Jan. 2000.
98. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, “UCSF Chimera-A visualization system for exploratory research and analysis,” Journal of Computational Chemistry, vol. 25, no. 13, pp. 1605–1612, 2004.
99. R. A. Laskowski, “Enhancing the functional annotation of PDB structures in PDBsum using key figures extracted from the literature,” Bioinformatics, vol. 23, no. 14, pp. 1824–1827, 2007.
100. M. Larkin, G. Blackshields, N. Brown, R. Chenna, P. Mcgettigan, H. Mcwilliam, F. Valentin, I. Wallace, A. Wilm, R. Lopez, J. Thompson, T. Gibson, and D. Higgins, “Clustal W and Clustal X version 2.0,” Bioinformatics, vol. 23, no. 21, pp. 2947–2948, Oct. 2007.
101. L. Patthy, “BLOSUM (BLOSUM Matrix),” Dictionary of Bioinformatics and Computational Biology, 2004.
102. X. Liu, S. Ouyang, B. Yu, Y. Liu, K. Huang, J. Gong, S. Zheng, Z. Li, H. Li, and H. Jiang, “PharmMapper server: a web server for potential drug target identification using pharmacophore mapping approach,” Nucleic Acids Research, vol. 38, pp. W609–W614, 2010.
103. A. P. Bento, A. Gaulton, A. Hersey, L. J. Bellis, J. Chambers, M. Davies, F. A. Krüger, Y. Light, L. Mak, S. Mcglinchey, M. Nowotka, G. Papadatos, R. Santos, and

J. P. Overington, "The ChEMBL bioactivity database: an update," Nucleic Acids Research, vol. 42, no. D1, pp. 1083–1090, Jul. 2013.

104. M. Swain, "chemicalize.orgchemicalize.orgby ChemAxon Ltd.," Journal of Chemical Information and Modeling, vol. 52, no. 2, pp. 613–615, Jul. 2012.

105. A. Jarrahpour, M. Motamedifar, M. Zarei, M. H. Youssoufi, M. Mimouni, Z. H. Chohan, and T. B. Hadda, "Petra, Osiris, and Molinspiration Together as a Guide in Drug Design: Predictions and Correlation Structure/Antibacterial Activity Relationships of New N-Sulfonyl Monocyclic β-Lactams," Phosphorus, Sulfur, and Silicon and the Related Elements, vol. 185, no. 2, pp. 491–497, 2010.

106. "Molinspiration cheminformatics," Choice Reviews Online, vol. 43, no. 11, Jan. 2006.

107. T. Liu, Y. Lin, X. Wen, R. N. Jorissen, and M. K. Gilson, "BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities," Nucleic Acids Research, vol. 35, no. Database, pp. D198–D201, Mar. 2007.

108. ACD/ChemSketch Freeware, Advanced Chemistry Development, Inc., Toronto, ON, Canada, www.acdlabs.com, 2014.

109. T. Sander, J. Freyss, M. V. Korff, and C. Rufener, "DataWarrior: An Open-Source Program for Chemistry Aware Data Visualization and Analysis," Journal of Chemical Information and Modeling, vol. 55, no. 2, pp. 460–473, Feb. 2015.

110. I. H. Witten, E. Frank, and M. A. Hall, "Introduction to Weka," Data Mining: Practical Machine Learning Tools and Techniques, vol. 4, pp. 403–406, 2011.

111. A. Grosdidier, V. Zoete, and O. Michielin, "SwissDock, a protein-small molecule docking web service based on EADock DSS," Nucleic Acids Research, vol. 39, no. Web Server Issue, pp. W270–W277, May 2011.

112. F. Cheng, W. Li, Y. Zhou, J. Shen, Z. Wu, G. Liu, P. W. Lee, and Y. Tang, "admetSAR: A Comprehensive Source and Free Tool for Assessment of Chemical ADMET Properties," Journal of Chemical Information and Modeling, vol. 52, no. 11, pp. 3099–3105, 2012.

113. M. Griffith and O. L. Griffith, "COSMIC (Catalogue of Somatic Mutations in Cancer)," Dictionary of Bioinformatics and Computational Biology, 2004.

114. H. Park, S. Kim, Y. E. Kim, and S.-J. Lim, "A Structure-Based Virtual Screening Approach toward the Discovery of Histone Deacetylase Inhibitors: Identification of Promising Zinc-Chelating Groups," ChemMedChem, vol. 5, no. 4, pp. 591–597, Jun. 2010.

115. X. H. Liu, H. Y. Song, J. X. Zhang, B. C. Han, X. N. Wei, X. H. Ma, W. K. Cui, and Y. Z. Chen, “Identifying Novel Type ZBGs and Nonhydroxamate HDAC Inhibitors Through a SVM Based Virtual Screening Approach,” Molecular Informatics, vol. 29, no. 5, pp. 407–420, 2010.