Mutational, Docking and Simulation studies on SERPINI1 gene and its experimental assay with ascorbic acid for its role in Alzheimer's disease

Dissertation submitted in partial fulfilment of the requirement for the degree of

MASTER OF SCIENCE

IN

BIOTECHNOLOGY

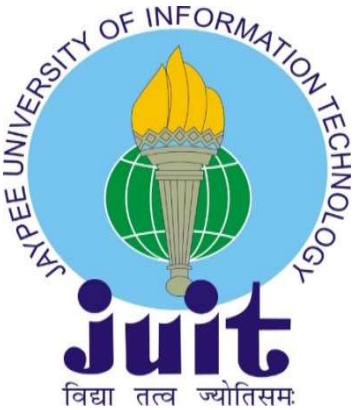
By

Simran Rai (207831)

Under the guidance of

Dr. Tiratha Raj Singh

Dr. Udaybanu M (Co-Guide)



JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY WAKNAGHAT
DEPARMENT OF BIOTECHNOLOGY AND BIOINFORMATICS
HIMACHAL PRADESH-174334

DECLARATION

I hereby declare that the work reported in the M.Sc. dissertation entitle "**Mutational, Docking and Simulation studies on SERPINI1 gene and its experimental assay with ascorbic acid for its role in Alzheimer's disease"** submitted at Jaypee University of Information Technology, Waknaghat, Solan, Himachal Pradesh, India, is an authentic record of my work carried out under the supervision of Dr. Tiratha Raj Singh and Dr. Udaybanu M, Dept. of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat, Solan, Himachal Pradesh-173234, India. I have not submitted this work elsewhere for any other degree or diploma.

Simran Rai, (207831)

Department of Biotechnology and Bioinformatics,

Jaypee University of Information Technology,

Waknaghat, Solan, Himachal Pradesh, India.

Date-

SUPERVISOR'S CERTIFICATE

This is to certify that the work reported in the M.Sc. dissertation entitle "Mutational, Docking and Simulation studies on SERPINI1 gene and its experimental assay with ascorbic acid for its role in Alzheimer's disease**"** submitted by Simran Rai (207831) at Jaypee University of Information Technology, Waknaghat, Solan, Himachal Pradesh, India, is bonafide record of her original work has not been submitted elsewhere for any other degree or diploma.

(Signature of Supervisor)

Dr. Tiratha Raj Singh

Dr Udaybanu M (co-guide)

Associate Professor

Dept. of Biotechnology and Bioinformatics

JUIT, Waknaghat, HP- 173234, India.

Acknowledgement

This thesis would not have been possible without Jaypee University of information technology giving me the opportunity to carry out an extensive project for our dissertation. I would like to thank my dean and all the faculty of my department for giving me this opportunity and helping me choose the right place to carry out my thesis work.

The completion of this study would not have been possible without Dr. Tiratha Raj Singh and Dr. Udaybanu who were my guide and co-guide. I would also like to thank Rohit Shukla, Diksha Manhas, Ms. Somalata Sharma, Ms. Mamta Mishra for helping me out through this entire project. Their expertise assisted me with a great deal with this project. A debt of gratitude is owed to them for giving me a chance to work on an imperative study in a very good environment.

Last but not the least, I would like to thank my family and friends whose constant support helped me reach the completion of this study.

Contents

List Of Figures

List Of Tables

CHAPTER 1

Introduction

==========

A healthy human brain consists of 86 billion neurons. The points of junction where these neurons communicate are termed as synapses[1]. Synapses transmit information from post-synaptic neuron to pre-synaptic neuron in the form of neurotransmitters. Effectual communication can be achieved by neurons through strengthening the molecular structure of synapses [2]. In normal brain, aging leads to weakening of synapses by down regulating the speed of connection within neurons and compromises with memory, decision making skills of an individual. However, in Alzheimer's functioning of neurons can be lost.

Alzheimer's Disease (AD)

AD is a neurodegenerative disorder which impairs memory of an individual and is a protein conformational disease (PCD) caused due to abnormal folding of soluble proteins present in the brain [3], [4]. AD is marked by the presence of amyloid-beta plaques and neurofibrillary tangles in the medial temporal lobe of the brain [5]

History

AD was first reported by Alois Alzheimer while working on a patient named Auguste Deter in Mental Hospital of Frankfurt. Following her death, brain biopsy was conducted in order to identify the relation between the records maintained in successive 4 years and her symptoms. He discovered that the part of the brain which controls thinking, memory, judgment, language was severely damaged. Formation of senile plaques and tangles were observed in neurons and nerve fibers. While presenting in the conference of southwestern Germany, he reported the results of the brain biopsy and symptoms. A local neurological symptom, delusion hallucinations were among the patient symptoms [6]. In this way AD was discovered.

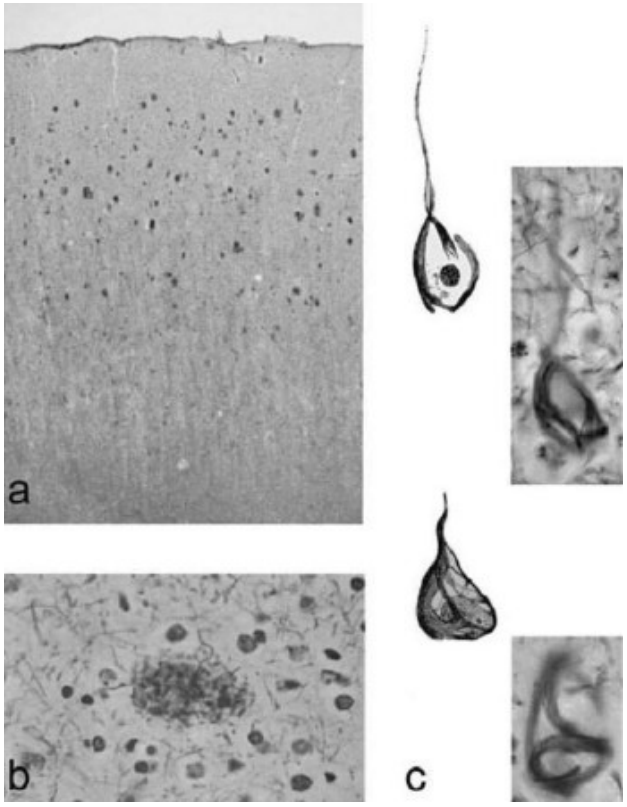


Fig 1: Sample of Auguste Deter brain showing Amyloid plaques (a, b) and neurofibrillary tangles (c) in cerebral cortex [7]

SYMPTOMS

1) Loss of Memory
2) Recognition problems
3) Difficulty in reading, writing and speaking
4) Behavioral changes
5) Confusion of time and place
6) Irritability [7]

Fig 2: Symptoms of AD

Clinical Phases of Alzheimer's disease

1) **Phase 1**:

   Patients appear to be normal and no memory loss occurs at this stage

2) **Phase 2**:

   In this stage patients tend to forget the places where the objects have been kept and may forget the names of people

3) **Phase 3:**

   Patient may find difficulty in reading or might get lost

4) **Phase 4:**

   Patients tend to forget about the events in their own life and face difficulty in performing any complicated task

5) **Phase 5:**

   In this phase patient needs a helping hand in order to survive. They may face difficulty in giving correct responses when asked about their phone number, names of their grandchildren or the place of their graduation.

6) **Phase 6:**

   This phase is the middle stage of dementia. Patients at this stage face difficulty in recognizing their own family members and require constant guidance.

7) **Phase 7:**

   This is the phase of late dementia. At this stage, patients lose their ability to speak and only mumble. They also require assistance in eating and for toileting. [8]

Statistics of AD across the world

It has been estimated that the percentage of population suffering from dementia will increase from 57.4 million (in 2019) to 152.8 million (in 2050). According to the reports of 2019, it was found that more women were suffering from dementia than men[9]

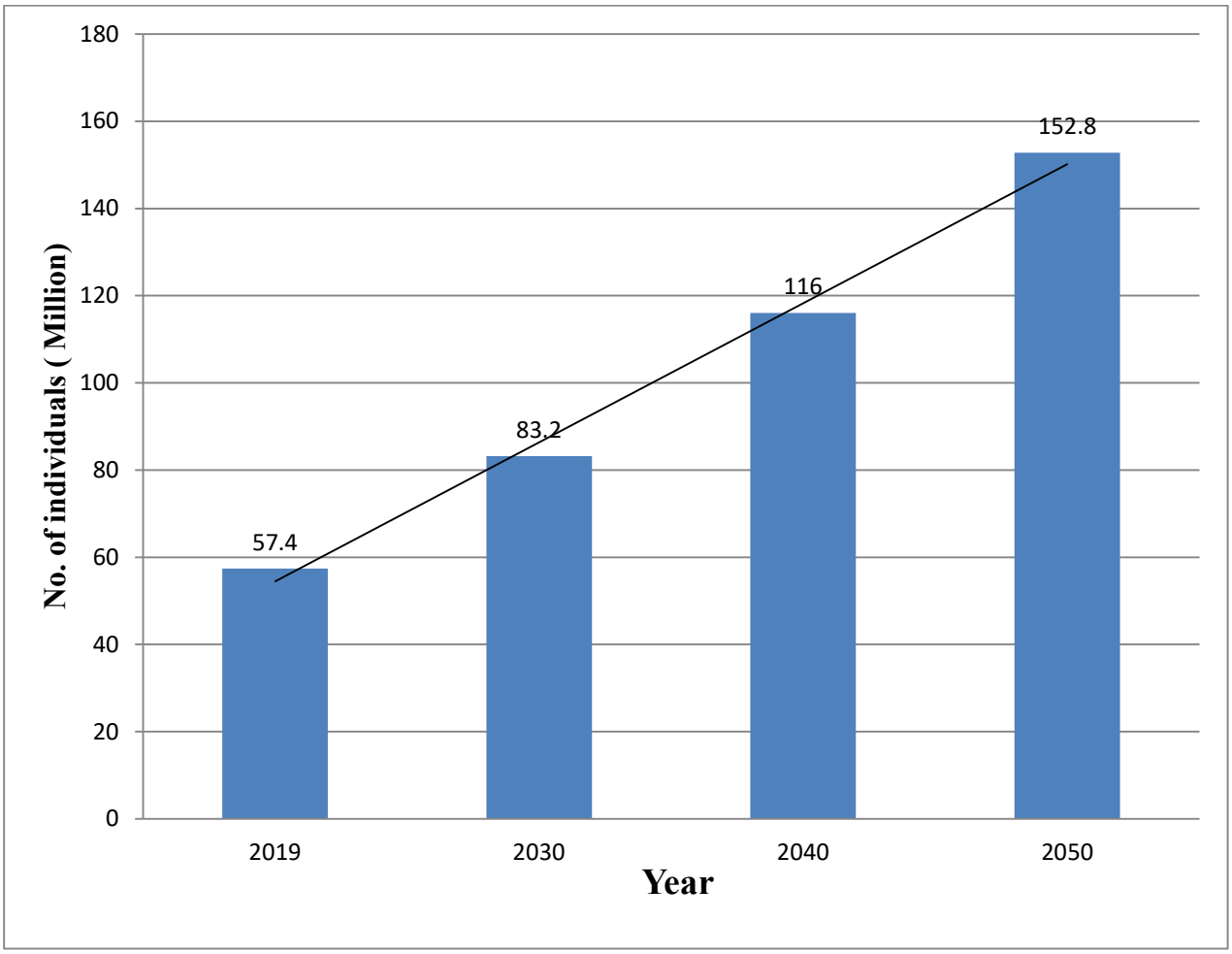


Table1: This graph represents the expected increase in the number of cases from 2019-2050. In 2019, the number of cases around the globe was 57.4 million and it is expected to increase to 83.2 million by 2030, to 116 million by 2040 and 152.8 million by 2050 [9]

On comparison across different regions across the world, the smallest anticipated increase was in high-income Asia Pacific and Western Europe where as the biggest approximate escalation were in North-Africa and Middle east and eastern Sub-Saharan Africa [9].

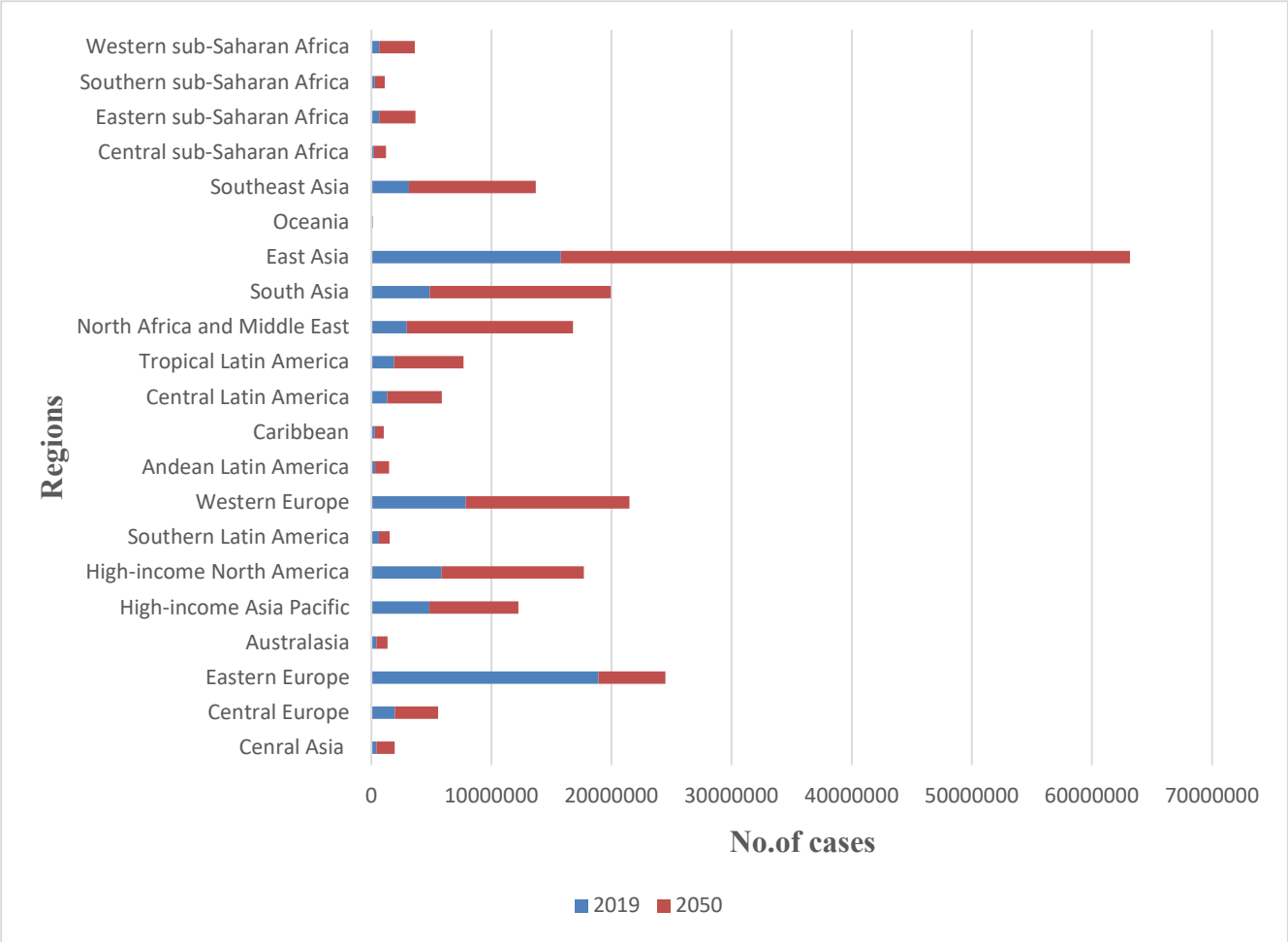


Table 2: It represents the expected increase in number of cases of dementia across different regions of the world

Role of Single nucleotide Polymorphisms (SNPs) in AD

It has been found that SNPs play a key role in variety of human diseases such as autoimmunity, diabetes and AD [10]

SNP

SNP may be defined as substitution of a single base-pair occurring within the coding or non-coding part of the gene. SNPs within the gene can give rise to more than one forms of allelic RNA. The interactions of mRNAs with different bases at SNP locations with cellular components involved in the synthesis , maturation, transport, translation, of mRNA may differ.[11]. It has been found that the SNP frequency differ for coding and non-coding sequences. On comparison of two chromosomes, 3-50 SNPs were found per 10 kilobases. [12]. SNP can be present in either the coding part or the non-coding part of the gene. SNPs within the coding region can further be classified into: Synonymous SNP (sSNP) and non-synonymous SNP (nsSNP). sSNP do not affect the sequence of amino acid, however may affect the rate of translation, alternative splicing and stability of m-RNA. nsSNP may

lead to change in the sequence of amino acid hence altering the structure and function of protein, thereby interfering with various processes. The SNPs within the non-coding region may affect various biological procedures. [13]. SNPs within the genes can affect the structure and function of protein and can lead to Alzheimer's. To identify the deleterious SNP that can probably lead to Alzheimer's various computational tools can be used on the basis of sequence and structure of the gene responsible for AD. Data can be retrieved from dbSNP for SNP of a particular gene. Screening through various sequence and structure-based tools is carried out. In this study, SERPINI1 gene was taken in order to assess the effect of SNPs using computational tools.

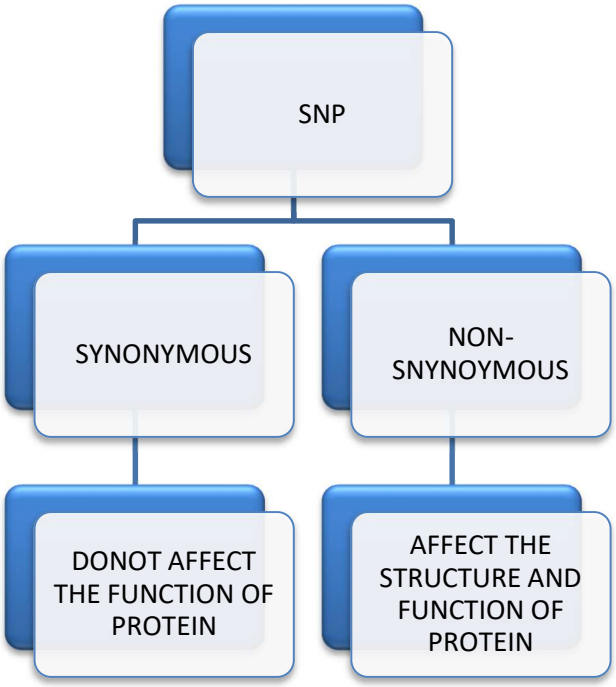


Fig 3: Types of SNP

SERPINI 1 GENE

It is found in nervous system and is a serine protease inhibitor [14]. SERPINI1 gene is located on chromosome 3 constituting of nine exons and encodes for a 410 amino acid protein termed as Neuroserpin with molar –mass of 45 kDa [15] [16] [17]. Neuroserpin has a conserved serpin fold which consists of 3β-sheets and 9α-helices [18] and has a exposed reactive centre loop which follows the process of suicide substrate inhibition [19]. This gene interacts with tissue- type plasminogen activator and has an inhibitory effect on it. This gene plays an important role in regulating how two neurons interact with each other, thus allowing the use and disuse of the neurons thereby increasing the strength of the neurons[20]

SERPINI1 gene and AD

The interaction of Neuroserpin with amyloid-beta plaque changes its confirmation to a less toxic form and hence reduces its toxicity.

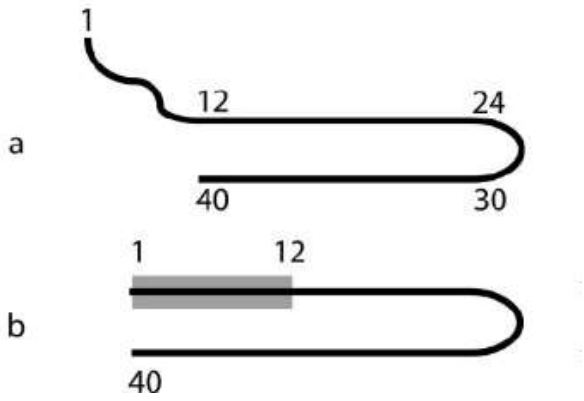


Fig 4: On binding of Neuroserpin (In gray) the confirmation of amyloid-beta plaque changes and is converted to a less toxic form[21]

OBJECTIVES

- To identify the deleterious SNPs using sequence and structure-based tools
- To carry out molecular dynamic simulation of the deleterious SNPs obtained
- To carry out molecular docking of Neuroserpin and Ascorbic acid
- To study the effect of ascorbic acid on animal cell-lines

CHAPTER 2

REVIEW OF LITERATURE

==========

Alzheimer's Disease (AD)

AD is a neurodegenerative disease that affects the brain. Dementia is a term that describes the development and progression of cognitive and functional deterioration that occurs with age and eventually leads to death [22]. Senile plaques and neurofibrillary tangles appear to be the key neuropathological hallmarks of AD [23].

Hallmarks of AD

Amyloid-beta (Aβ) plaques

Aβ plaques is an insoluble protein (4KDa) formed by the cleavage of amyloid precursor protein (APP). The processing of APP is carried by two different pathways i.e. amyloidogenic (cleavage by β-secretase) and non-amyloidogenic (cleavage by α-secretase). The formation of Aβ is carried out by two membrane bound enzymes**: β-secretase** and **γ-secretase**. Upon cleavage by β-secretase, two fragments are generated sAPPβ and CTFβ, (a 99 amino-acid membrane anchored protein) which is then rapidly cleaved by γ-secretase to produce Aβ. It has been found that 80-90% of Aβ population consist of 40 amino acid fragments i.e. Aβ40 and 5-10% population consist of 42 amino acid fragments i.e. Aβ42. Aβ42 is the key component deposited in the brain of AD patients and is hydrophobic and fibrillogenic in nature [24]. β-secretase which processes 10% of APP is thought to be pace limiting step in this pathway. The remaining APP is cleaved by α-secretase leading to production of 2 fragments sAPPα and a 83 amino-acid CTF α(Membrane tethered α terminal fragment).The cleavage of CTF by γ-secretase leads to the production of a cytosolic component AICD which is involved in signal-transduction.[25] [26]

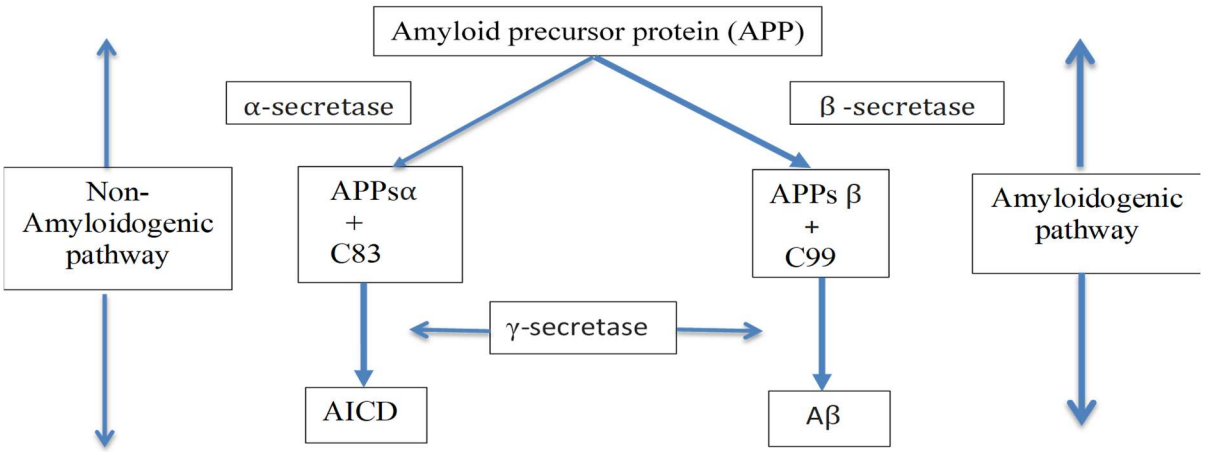


Fig 5: Mechanism for formation of amyloid-beta plaques

Enzymes responsible for Aβ degradation

1) **Neprilysin**: A plasma- membrane anchored metalloprotease involved in breakdown of peptides.[27]
2) **Insulin degrading enzyme (IDE):** It leads to degradation of amyloid-beta plaques[28].

Aβ Toxicity

The aggregation of Aβ is the major cause of neurotoxicity and Aβ oligomers are most potent form and contribute to AD

Consequences of Aβ toxicity

1) **Free radical formation**: The interaction of reactive oxygen species (ROS) with lipids and protein moieties can lead to oxidation of proteins and peroxidation of lipids which compromises with the integrity of the membrane and reduces the activity of enzymes such as glutamine synthetase (GS) and creatine kinase (CK), which are important for neuronal activity.

   [29]. Peroxidation of lipids can lead to suppression of ion-motive ATPase, glial cell Na+ dependent glutamate and signaling pathway disruption which can lead to death of neurons[30] [31] . Aggregation of Aβ leads to oxidation of DNA, thus leading to its damage[32]

2) **Loss of synapses by activation of innate immune system:**

   Aggregation of amyloid plaques can lead to activation of Toll-like receptors such as TLR6, TLR4 which are likely to phagocytose the neuronal cells. Also an inflammatory response is produced which prevents the clearance of amyloid-beta by microglial cells and thus leading to loss of synapses and neuronal death.[33] [34].

3) Pathogenesis of tau

Formation of Neurofibrillary tangles leads to induction of tau protein which are hyper phosphorylated and has an impact on late stage AD-pathogenesis.[35].

Neurofibrillary Tangles (NFTs)

NFTs are key neuropathological feature of Alzheimer's disease. The loss of cytoskeleton microtubules and tubulin associated proteins has been seen in NFT-bearing neurons. The tau proteins is hyperphosphorylated in the brains of AD patients[36]. Hyperphosphorylation of specific amino acids in tau protein leads to detachment of proteins from microtubules in patients with AD disrupting the transport structure and resulting in cell death. Thus hyperphosphorylated tau plays a key role in neurofibrillary alterations and the etiology of Alzheimer's disease [37].

Tau Pathology

Tau protein is encoded by MAPT gene which comprises of an N-terminal region, proline -rich domain, a microtubule-binding domain and a C-terminal. In human brain, six primary tau isoforms are produced through alternate splicing [38] [39]. Tau is a microtubule-associated protein which binds to the surface of microtubule via microtubule binding domain (MBD) and regulates the morphology of microtubules and axoplasmic flow [40] [41]. Hyperphosphorylation prevents the association of tau proteins with the microtubules leading to self-agglomeration of tau proteins and instability of microtubules causing neuronal degeneration [42]. Phosphorylation in AD patients can be carried out by various protein kinases such as proline-directed protein kinases (PDPKs), non-PDPKs and tyrosine kinases which include Src family kinases (SFKs), FYN and ABL family kinases [43] [42]. Phosphatases such as phosphatase1 (PP1), PP2, PP2A, PP2B can cause hyperphosphorylation by repression of these phosphatases.[42]. It has also been found that Aβ can mediate the tau toxicity by guiding toxic tau species to dendritic spines resulting in deterioration of dendrites and disintegration of spines [44] [45]. Mitigation of tau proteins at serine/ threonine residues can be carried out by β-N-acetylglucosamine (GlcNAc) which is termed as O-GlcNAcylation [46]- [47]. In AD patients, Tau O-GlcNAcylation was lowered which lead to increased tau hyperphosphorylation[48]. The metabolism of glucose provides Uridine diphosphate (UDP)-GlcNAc for O-GlcNAcylation, thus any damage in glucose metabolism can lead to reduction in level of tau O-GlcNAcylation and facilitating tau hyperphosphorylation which leads to formation of neurofibrillary tangles [49] [50].

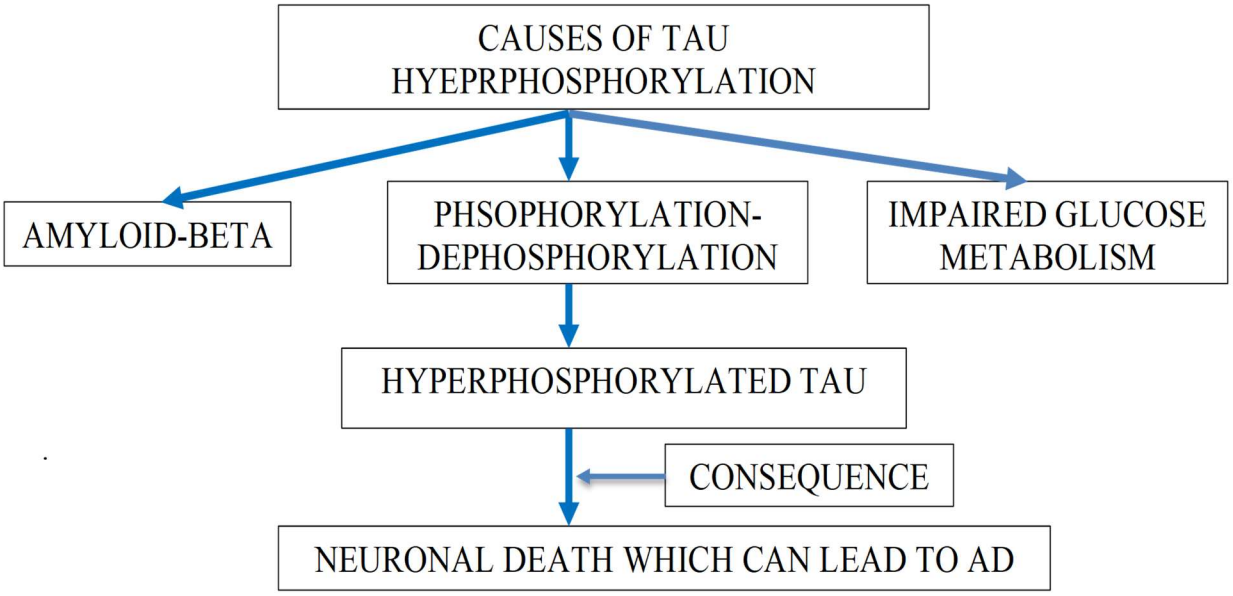


Fig 6: Schematic representation of causes of tau hyperphosphorylation and its consequence

Genetics of AD

It has been found that mutations in three genes i.e., Amyloid precursor protein (APP), Presenilin1 (PSEN1) and Presenilin 2 (PSEN2) can lead to Familial AD which is dominantly inherited and accounts for less than 5% of AD. The late-onset AD is caused due to mutations in the Apolipoprotein E (APOE) gene [51].

Genetics of Early-onset AD

1) **APP gene**

APP gene is located on chromosome 21. It exists in three major isoforms which includes APP695, APP751, and APP770. The isoform APP695 is present in the central nervous system (CNS) whereas the two other major isoforms are found both in peripheral and CNS [52] [53]. This gene is often involved in maintaining the formation of synapses and neuronal plasticity [54]. The synthesis of APP protein occurs in the Endoplasmic reticulum and is modified post-transcriptionally in the Golgi body which is then transported to the surface of cell by means of secretory pathway [55].

APP processing

APP processing occurs via two major pathways: amyloidogenic and non-amyloidogenic. Amyloidogenic pathway involves the cleavage of APP by α-secretase and γ-secretase leading to the formation of non-pathogenic fragments, (sAPPα and a C-terminal fragment). In non-amyloidogenic pathway, APP is processed via two enzymes β-secretase and γ-secretase leading to production of

Aβ fragments of varying length [56] . Aβ fragment of length 40 amino-acids are largely innocuous whereas Aβ fragments of length 42 and 43 amino-acids are self-aggregating in nature and are found to be highly toxic [57]. Aβ40 is the most prevalent species formed in the amyloidogenic pathway in normal conditions with Aβ42 accounting barely for 10% of total Aβ.

Mutations In APP gene and its role in AD

Alternative splicing of 18 exons of APP gene leads to production of proteins of varying length with Aβ peptide encoded by 16^{th} and 17^{th} exon. Till now, 26 harmful mutations have been discovered which are present within the Aβ sequence or regions surrounding it [58].

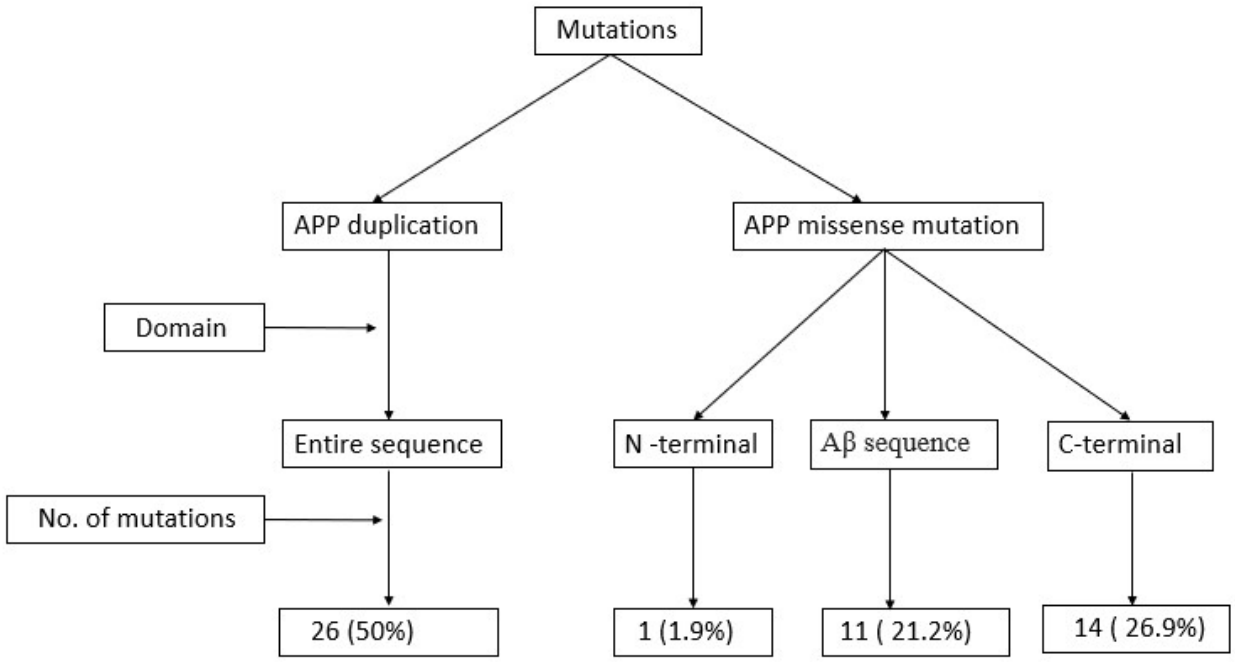


Fig 7: This figure represents the total number of pathogenic mutations within the APP gene

Mutations can occur either in N-terminal region within the β-secretase cleavage site, in C-terminal region near to γ-secretase cleavage site and near the ε-secretase cleavage site, or within the Aβ domain the cleavage site for α-secretase [59].

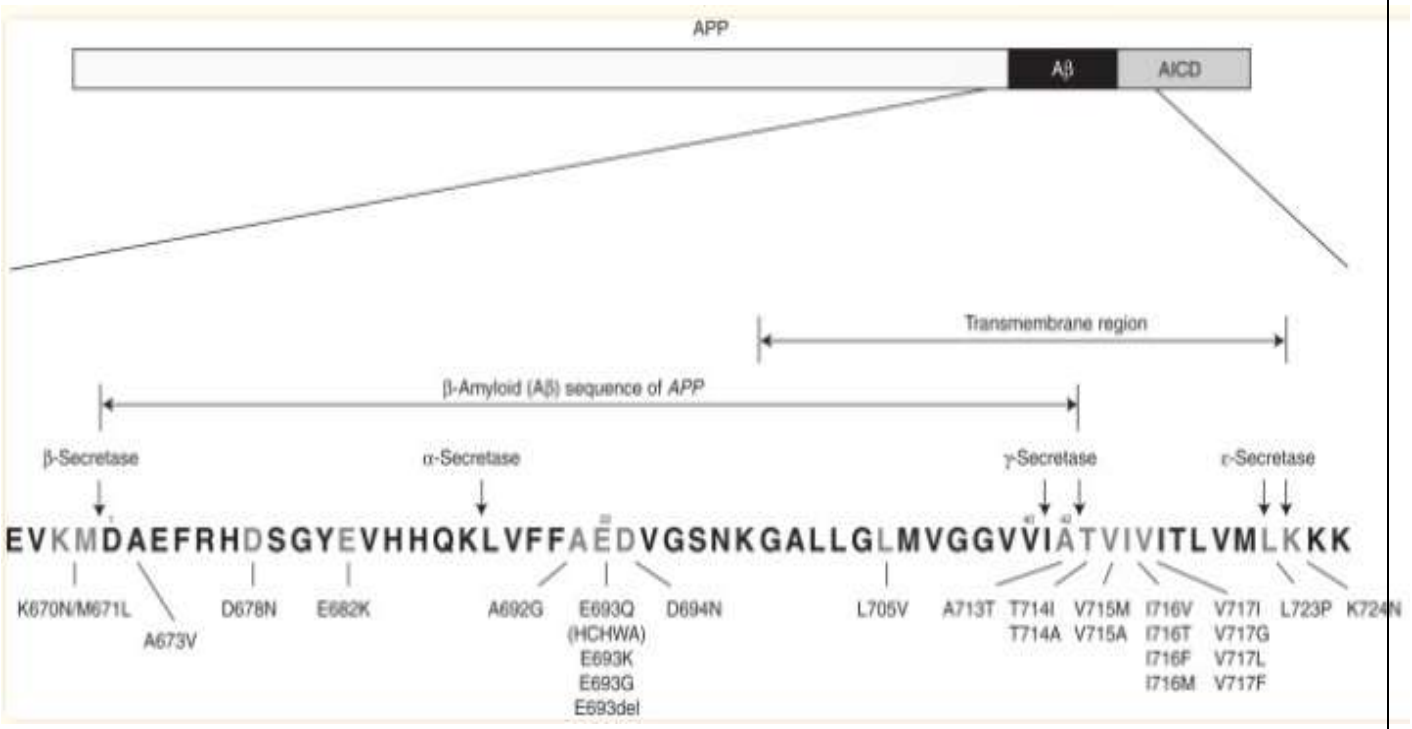


Fig 8: This figure represents the structure of APP gene and regions where the mutations occur [59].

Mutations in N-terminal region

Mutations occurs within the codon 670 and 671 located on exon 16 of APP gene lead to substitution of lysine-to-asparagine in codon 670 and methionine to leucine on codon 671. These mutations led to increase in the levels of Aβ42 and Aβ40 species [60].

Mutations in Aβ-domain

Mutations within this domain can compromise with the activity of α-secretase and hydrophobic character of Aβ species increases, thus favoring the formation of amyloid fibrils. Various harmful mutations have been reported in this region where in the formation of amyloid fibril is favored [61].

Mutations in C-terminal

Mutations within this region can lead to formation of Aβ species of longer length mainly Aβ42 which is facilitated more towards aggregation. Mutations occur at various codons such as codon 717, 723,724 and so on. These mutations affect the activity of the various secretases present within that regions [62].

These mutations led to increase in the level of Aβ42 species prone to aggregation and are the hallmarks of properties associated with AD which includes neuronal death, formation of tangles.

2) **Presenilin1 gene (PSEN1)**

The PSEN1 gene is present on chromosome 14 and encodes for presenilin 1 protein and is involved in break-down of amyloid β precursor[63]. The presenilin 1 protein formed complexes with the γ-secretase and carries out the proteolysis of the proteins wherein the presenilin 1 is often called the proteolytic subunit of γ-secretase. The complex formed leads to breakdown of APP into small peptides which includes sAPP and various other varied length amyloid-beta precursor [64].

Mutations in PSEN1 gene

Mutations within the PSEN1 gene leads to production of amyloid species of longer lengths (Aβ42) [65]. Around 176 mutations have been discovered within the PSEN 1 gene which decreases the activity of γ-secretase and increases the Aβ42/ Aβ40 ratio [66]. Mutations such as N135S in the PSEN1 gene have been associated with loss of memory , Notch signaling and other neurodegenerative changes which lead to AD [67] . PSEN1 gene is often mutated among the three genes responsible for dominant pattern and occurs at a young age i.e. within 30-50 years [68].

3) **Presenilin2 gene (PSEN2)**

This gene is located on chromosome 1,consisting of 17 exons and encodes for presenilin2 protein composed of 448 amino-acids and shares sequence homology with PSEN1 gene[68]. The protein formed consists of nine-transmembrane domains. However, it differs from PSEN1 at N-terminal and the hydrophilic loop where in the hydrophobic region is highly conserved [69]. This gene also complexes with the γ-secretase and is responsible for proteolysis of Aβ [71].

Function:

It complexes with γ-secretase and is responsible for breakdown of APP to produce varying lengths of Aβ. Mutations within the PSEN2 gene can increase the activity of γ-secretase and thus increased production of aggregated Aβ42. Also mutations in PSEN2 gene can lead to apoptosis [70].

Mutations in PSEN2 gene

A point mutation was identified in PSEN2 gene which lead to substitution of asparagine to isoleucine. These mutations were associated with increase in the level of Aβ42 production and initiated the events for AD [72]. Probable damaging mutations in PSEN 2 gene discovered are A85V, T122P/R, K155Efx10,N141I, M239V/I, T430M [73].

Genetics of Late-onset AD

Most of the AD cases develop later in life at an age of more than 65 years and is termed as Late-onset AD (LOAD). ApolipoproteinE (ApoE) acts as a strongest genetic factor for LOAD [74].

ApoE

Apolipoprotein is made of 299 amino acids with a molecular mass of 34 kDa. ApoE exists in three isoforms: ApoE3, ApoE2, ApoE4. ApoE3 is the most usual form and differs from other two isoforms by single amino acid change at position 112 (for ApoE2) and 158 (for ApoE4). The allelic form E4 is often involved in development of AD [75]. The difference in the single positions of amino acids in these allelic forms affects the structure in such a way that their binding ability to lipids, receptors and Aβ is modified [76] [13].

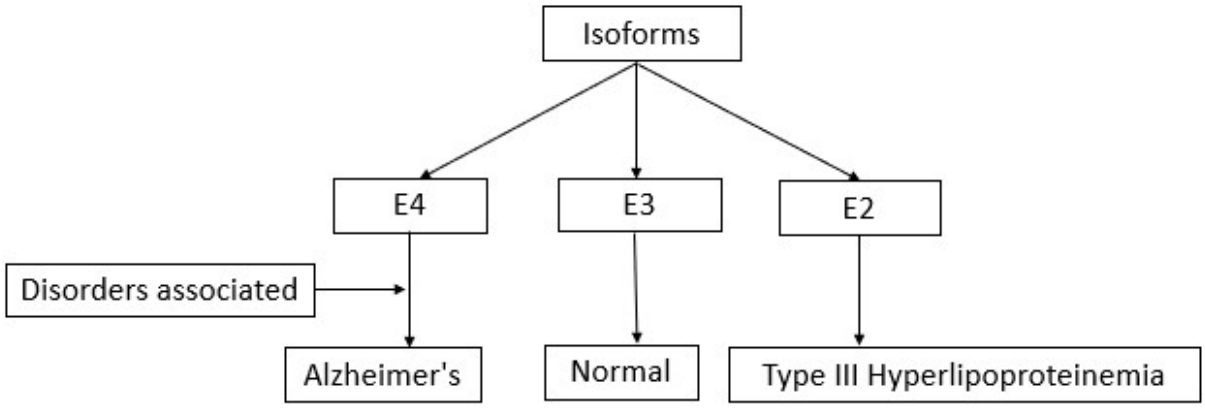


Fig 9: This figure represents the isoforms of lipoprotein and its associated disorders

Functions

It is often involved in transport of lipids and is produced by liver and macrophages. Within the CNS it is often involved in transport of cholesterol with the help of ApoE receptors to neurons [77] [78]. The ApoE is often involved in transport of lipids and is produced at higher concentration in response to injury of nerves and plays an important role in regeneration of neurons by distributing lipids to injured neurons [79].

ApoE4 in AD

The allelic form of apolipoprotein E4 is often involved in the pathogenesis of AD. In the aetiology of AD, ApoE4 confers harmful effects on neuronal activity.

Consequences of formation of allelic form E4 of Apolipoprotein

1) **Hyperphosphorylated tau**

The truncated form of ApoE4 produced in response to any stress or injury may lead to phosphorylation of tau which is neuron dependent and leads to development of neurofibrillary tangles, disruption of cytoskeleton, and impaired functioning of mitochondria [80].

2) **Abnormal lipid metabolism**

Cholesterol is important for formation of synapses, growth of axon, a key event necessary for formation of memory, repair of neurons and learning. ApoE4 carriers may have difficulty in transporting cholesterol thus disrupting the homeostasis of cholesterol and can lead to AD [81] [74].

3) **Impaired Synaptic plasticity and integrity of spines**

The ApoE4 allele contributes to decrease spine density within the hippocampus, thus might affect the integrity of neurons and can increase the risk for dementia. Individuals carrying ApoE4 allele may have impaired functioning of glutamate receptor and trafficking as well as signaling thus may compromise the functions of learning [82].

In this way mutations within any of these genes can increase the risk to AD.

SNP (Single Nucleotide Polymorphism)

Difference in a single genetic code is termed as Single Nucleotide Polymorphism. A point mutation is considered to be SNP if it occurs in more than 1 percent population[83]. Nearly 10 million SNPs are found in the human genome and occur within the DNA contained in the genes. Due to their wide spread frequency, accessibility of analysis, inexpensive genotyping costs and ability to conduct association studies using statistical and bioinformatics techniques, SNPs are regarded as an important biomarkers for illness diagnosis [84]. Sequencing the same genomic area in various groups identifies and characterizes SNPs.

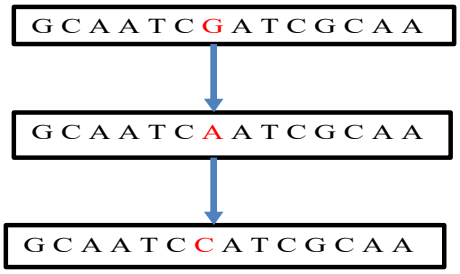


Fig 10: Schematic representation of SNP

SNP classification

SNPs can be found in the gene's coding sequence, non-coding section or within the gene (Intergenic region). SNPs within the coding part are further classified into synonymous and non-synonymous SNPs. Synonymous SNPs have no effect on sequence of amino acid or on the function of protein. Non-synonymous SNPs are further classified into missense and nonsense. For a missense SNP, change in single nucleotide can lead to production of different amino acid hence affecting the function of protein whereas nonsense SNP may be defined as premature termination by stop codon which produces a non-functional protein. SNPs within the coding region of the gene can lead to degradation of m-RNA, may influence gene splicing[85].

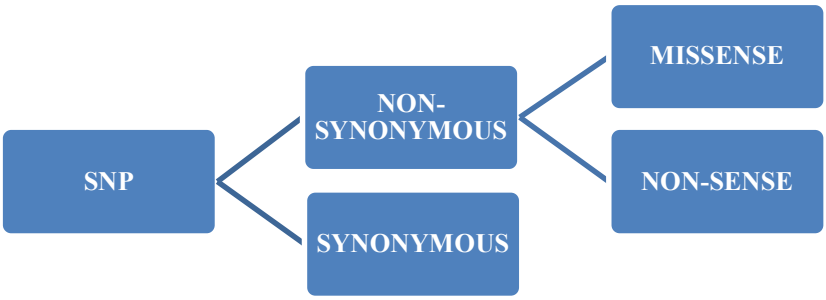


Fig 11: SNP classification

Impact of mutations on structure and functions of gene

SNPs occur after every 100-300 base-pairs and can affect the structure and function of protein. Mutations can affect the expression of gene on the basis of their location.

Table 3: This table represents the location and effect of mutations [86]

Location	Effect
Transcriptionally regulatory elements	Alter the expression of m-RNA
Within genes	Affect RNA-splicing, translation, stability

Within the coding region	Modification in the activity of protein
Premature termination	Can lead to formation of truncated protein and may lead to nonsense phenotype

SERPINI1 GENE

Over 550 genes in human genome encode for proteases, while 150 code for protease inhibitor [87]. 35 serpin coding genes have been found in humans and classified into nine subfamilies i.e. SERPINA, SERPINC, SERPINB, SERPINI [88] [15]. SERPNI1 gene belongs to a family of protease inhibitor which encodes for Neuroserpin and is involved in repression of tissue-type plasminogen activator (tPa) [14]. The protein formed is axonally secreted. It plays an important role in neuronal development and growth of axons. Presence of mutant polymers of Neuroserpin, can result in familial encephalopathy and epilepsy due to mutation in SERPINI1 gene and is expressed in brain on long arm of chromosome 3 with an exon count of 9 (3q26.1) and encodes for a protein having length of 410 amino acids and molar mass of 46,427 Da[89] [90]

```
        10          20          30          40          50
MAFLGLFSLL VLQSMATGAT FPEEAIADLS VNMYNRLRAT GEDENILFSP
        60          70          80          90         100
LSIALAMGMM ELGAQGSTQK EIRHSMGYDS LKNGEEFSFL KEFSNMVTAK
       110         120         130         140         150
ESQYVMKIAN SLFVQNGFHV NEEFLQMMKK YFNAAVNHVD FSQNVAVANY
       160         170         180         190         200
INKWVENNTN NLVKDLVSPR DFDAATYLAL INAVYFKGNW KSQFRPENTR
       210         220         230         240         250
TFSFTKDDES EVQIPMMYQQ GEFYYGEFSD GSNEAGGIYQ VLEIPYEGDE
       260         270         280         290         300
ISMMLVLSRQ EVPLATLEPL VKAQLVEEWA NSVKKQKVEV YLPRFTVEQE
       310         320         330         340         350
IDLKDVLKAL GITEIFIKDA NLTGLSDNKE IFLSKAIHKS FLEVNEEGSE
       360         370         380         390         400
AAAVSGMIAI SRMAVLYPQV IVDHPFFFLI RNRRTGTILF MGRVMHPETM
       410
NTSGHDFEEL
```

Fig 12: Sequence of Neuroserpin[56]

Structure of Neuroserpin

Monomers of human Neuroserpin can exist in three different confirmations i.e., native, latent or cleaved.

1) **Native Neuroserpin**

2) **Cleaved Neuroserpin**
3) **Latent Neuroserpin**

Native Neuroserpin

Neuroserpin has a serpin fold which consists of three β-sheets, nine α-helices and RCL. In order for the substrate to bind, native Neuroserpin has an omega-loop connecting 1 and 2 strands in β- sheet with glycine residues present at 231,236 and 237 positions of the protein. The presence of the glycine residues at these positions provides flexibility to the molecule which is important for substrate binding. The amino acid substitutions at specific base pairs in helix F and strand 1 of β-sheet decreases the firmness of helix F allowing it to fold into latent confirmation [91] [92]. Changes in the F region may affect the inhibitory activity and may lead to polymerization[93].

Cleaved Neuroserpin

In cleaved Neuroserpin, the RCL integrates in β-sheet A and forms a six-stranded β-sheet wherein the RCL remains exposed. The C-terminal portion within the RCL is equilibrated by interaction with β-sheet A [94] [18].

Latent Neuroserpin

The structure is found to be similar to that of PAI-1 (Plasminogen activator inhibitor1) which consists of RCL present within the β-sheet A and the shift to this structure is an auto-regulated [95] [60].

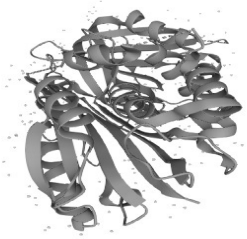


Fig 13: Structure of Neuroserpin (Q99574) [56]

Inhibitory role of Neuroserpin

On studying Neuroserpin it was discovered that it acts as an inhibitor of trypsin-like proteases [14]. Due to presence of arginine and methionine within the P1 and P1' regions of RCL it was believed that it would target serine-like proteases and was found that upon complex formation with tPA it leads to inhibition of tPA and other serine-like proteases[97] [98] [99]. tPA contributes to degeneration of neurons and can lead to injury by the modulation of permeability of neurovascular section and its level decreases with increased level of Neuroserpin. It was found that upon administering Neuroserpin proliferation of cells, plasticity was affected [100] [101] [102].

Mechanisms by which tPA may lead to neuronal death

1) Breakdown of extracellular matrix by the action of plasmin [103] [104] [105]
2) Increase in the concentration of Calcium within the cells can lead to death of neurons via the NMDA receptor [106] [107]
3) Affecting the signaling pathway through LRP receptor [103].

SERPINI1 Gene and AD

It has been found that Neuroserpin can interact with Aβ. It combines in the stoichiometric ratio of 1:1 to form Aβ: Neuroserpin complex. Aβ interacts with β—sheet A of Neuroserpin and a complex is formed which is unable to inhibit tPA. Interaction of Neuroserpin with Aβ occurs via N-terminal fragments and middle parts of Aβ. Upon Aβ-interaction Neuroserpin leads to change in structure of toxic amyloid βeta to less toxic form. It has been found that once the inhibition of tPA has been blocked, tPA may lead to activation of plasmin via plasminogen and plasmin is involved in breakdown of pro-BDNF to m-BDNF thus playing important role in synaptic plasticity[21]. However increase in concentration of Neuroserpin may reduce the levels of plasmin and thus reduces the clearance of Neuroserpin-Aβ complex despite of its role in protecting the CNS by forming complex with Aβ [108].

Neuroserpin in regulation of Aβ- pathology

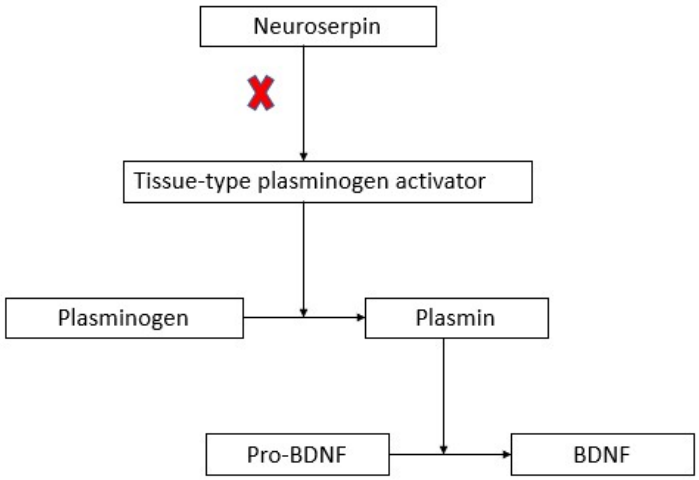


Fig14: This figure explains inhibition of plasmin via tPA inhibition

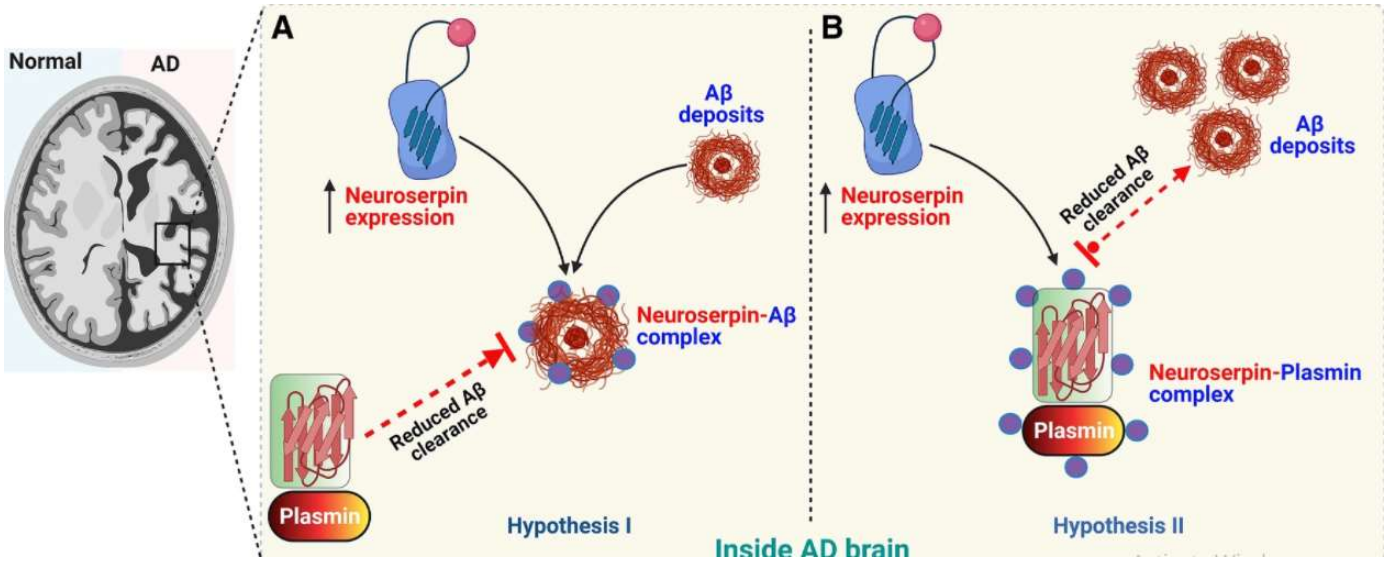


Fig 15: On binding of Neuroserpin to Aβ the complex formed is resistant to plasmin action hence reducing the clearance of this complex (A) whereas when Neuroserpin complexes with plasmin it does not allow the action of plasmin to remove Aβ plaques(B) [109].

These hypotheses led to involvement of Neuroserpin in AD

Ascorbic acid

Ascorbic acid is also known as Vitamin C. It is an electron-donating molecule which is produced by several kinds excluding fishes, birds and higher order primates. Inactive form of i-gulonolactone oxidase developed due to mutations or deletions makes it difficult for these species to synthesize ascorbic acid and in turn have to depend on their diet in order to get Vitamin C [110].

Functions of Vitamin C

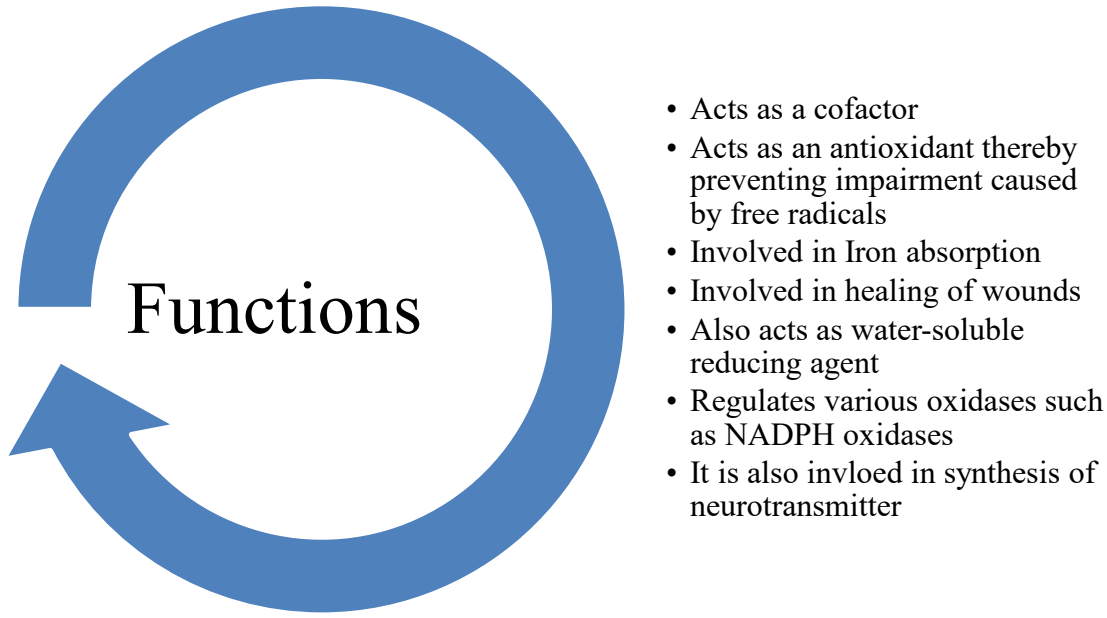


Fig 16: This figure represents the biological functions of Ascorbic acid [111]

Ascorbic acid Inadequacy

Inadequacy of this Vitamin can lead to difficulty in healing of wounds, degeneration of muscles, scurvy, anemia, development of cholesterol plaques , depression, delicate bones [112].

Vitamin C and AD

Aβ accumulation can induce oxidative stress and leads to formation of reactive oxygen species (ROS). Oxidative stress leads to increase in the activity of β-secretase thus leads to formation of Aβ peptide and compromises with the synaptic plasticity [113]. Metals are also involved in affecting the morphology of Aβ. Ascorbic acid provides protection from oxidative stress as it acts as an anti-oxidant

which in turn is responsible for generation of other anti-oxidants such as catalase and vitamin E [114]. It was found that amount of Ascorbic acid was decreased in the AD patients having mutations in APP/PSEN1 gene and thus decreasing the transport of Vitamin C across the brain[115]. Administration of Ascorbic acid reduced the symptoms of AD such as amyloid plaques reduction, which in turn lead to reduction of blood-brain barrier disarrangement and halting of unusual mitochondrial morphology thus reducing the risk of AD [116].

Mtt assay

This assay was first reported by Mosmann. It is a calorimetric assay used to measure the viability of cell, its proliferation and loss of viable cells. It is based on the principle that the yellow-colored tetrazolium salt (3-(4,5-deimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) often called as MTT is reduced to purple colored formazan by the cells which are metabolically active thus leading to differentiation between viable and non-viable cells. The NADPH-dependent oxidoreductase enzymes are responsible for this reduction. The quantification can be carried out by taking the absorbance at 500-600nm with the help of spectrophotometer. Dark colored solution represents higher viable and active cells

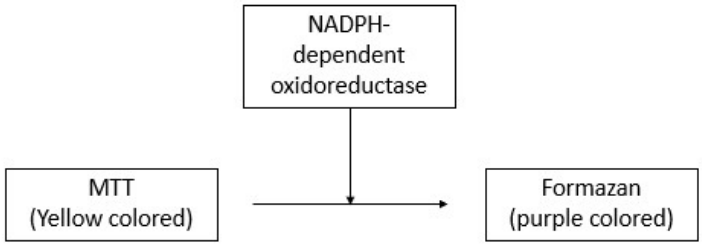


Fig17: MTT reaction

CHAPTER 3

METHODOLOGY

==========

3.1 Data Retrieval

Data for SNPs was collected from dbSNP (https://www.ncbi.nlm.nih.gov/snp). National centre of Biotechnology Information created the dbSNP database in order to address the large-scale sampling design, for mapping of genes and evolutionary biology [117]. Data for SNP was collected from dbSNP in order to screen those mutations. For SERPINI1 gene, a total of 33,072 SNPs was found out of which 337 missense SNPs were taken for further screening.[75]. The sequence of this gene was obtained from UniProt with UniProt ID Q99574. Structure was retrieved from Protein Data bank (PDB)

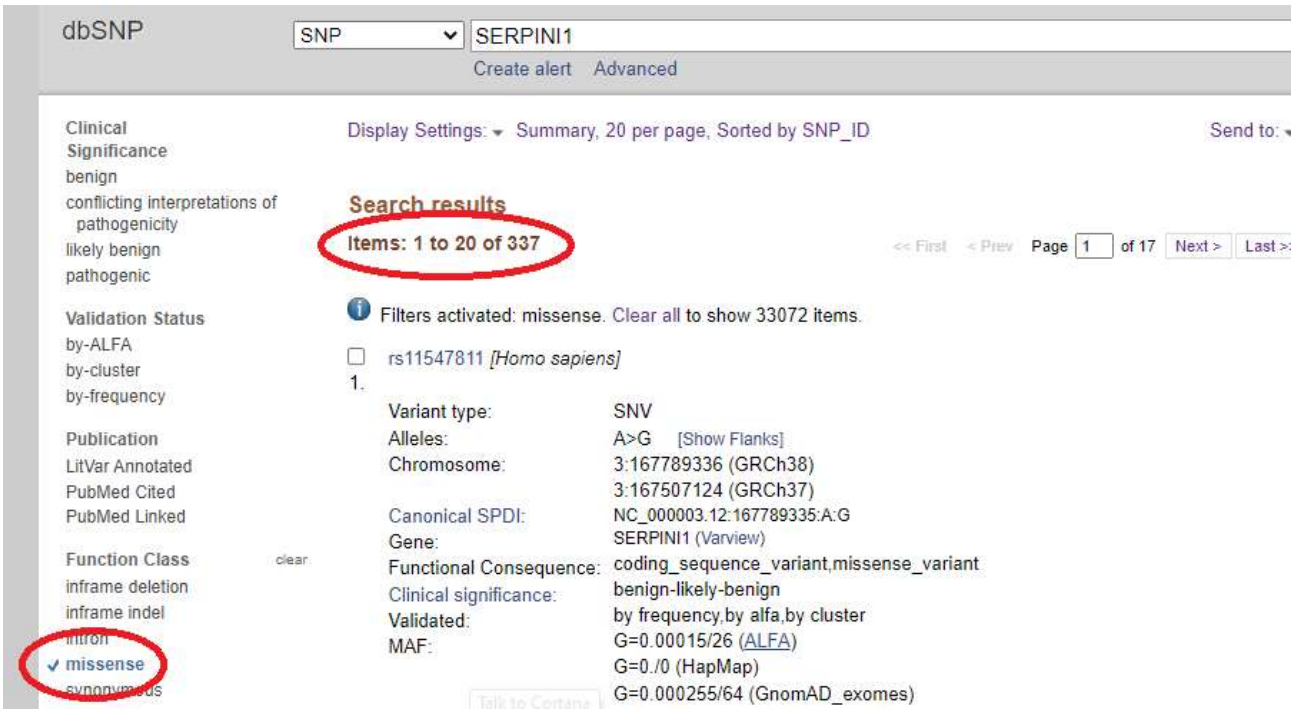


Fig 18: 337 missense SNPs were obtained from dbSNP database

3.2 Deleterious SNP identification using Sequence Based Tools

The changes in amino acid sequence and structure of protein were detected using various sequence-based tools such as PROVEAN, PANTHER, Poly-Phen2, Meta-SNP, PredictSNP1, SusPect, PMut, SNAP2, SNPs&GO.

1) PROVEAN (http://provean.jcvi.org/index.php)

Protein Variation Effect Analyzer is a tool that predicts if change in the sequence of amino acid has any effect on the protein function. PROVEAN is a novel tool that uses alignment-based scoring technique.[119]. PROVEAN can predict substitutions in single as well as multiple amino acids by using similar-scoring technique. It collects information about homologous sequences from NCBI and an alignment score is designated to both query and mutant and a score is specified which is the difference between the mean alignment score for query and mutant protein. The SNPs in this tool is categorized as deleterious and neutral. The threshold value was set to be -2.5 in which variants having a score of -2.5 or less -2.5 were considered to be deleterious [120].

2) PANTHER (http://www.pantherdb.org/tools)

Protein Analysis Through Evolutionary Relationship. PANTHER usually works on the principle of how long an amino acid has been preserved in a lineage which leads to formation of protein. More the time of preservation more chances of functional impact on the protein. This mode of determination is known as PANTHER-PSEP (position-specific evolutionary preservation). PANTHER-PSEP plays an important role in identification of SNPs involved in genetic variation which can lead to human disease[121]. Using curated databases of the protein families, PANTHER identifies the function of the proteins, a product of a gene[122]. PANTHER often asses the function of protein using an evolutionary model which predicts the consequences of genetic diversity. [123]. PANTHER currently consist of trees of 5000 protein families which is further classified into 30000 [124].

3) PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/)

It predicts the effect of amino acid substitutions on the functioning and structure of protein. The identification of probably damaging SNPs are characterized on basis of sequence number, phylogeny and structural features which characterize the substitutions [125]

4) Meta-SNP (http://snps.biofold.org/meta-snp)

Meta-predictor of Disease-causing variants. Meta-SNP requires an input of protein which can either be FASTA format or the sequence of protein. It also requires the list of mutations which are separated by commas. Scoring System can be classified as:

- For PANTHER: If the score is greater than 0.5 the mutation is predicted to be deleterious
- For PhD-SNP: If the score is greater than 0.5, the mutation is found to be deleterious
- For SIFT: A positive value with a score greater than 0.05, the mutation is found to be neutral
- For SNAP: The score greater than 0.5 is found to have deleterious effect and leads to diseased condition
- Meta-SNP: If the score is greater than 0.5 it is predicted to have deleterious effect [126]

5) Predict SNP1 (http://loschmidt.chemi.muni.cz/predictsnp)

Predict SNP runs multiple programs used to identify the effect of mutation on the function of protein. Combination of results from Predict SNP, MAPP, PhD-SNP, Polypehn-1, PANTHER are studied [127]. The scoring system is used to identify whether the mutation is deleterious or neutral. If the value lies between (-1,0), the consequences of mutations are said to be neutral. If the value lies between (0,1) the mutations are found to be deleterious. The larger the gap between the Predict SNP score and zero, the more harmful the mutations are.[128].

6) SusPect http://www.sbg.bio.ic.ac.uk/suspect/about.html)

SuSPect predicts the phenotypic implications of missense mutations based on sequencing, structure, and systems biology factors. A score of 0-100 is depicted for diseased and neutral conditions Color coding depicts the nature of mutants (blue=neutral, red=disease). On clicking the score parameter, details of the variant such as its projected solvent accessibility , degree of conservation at that residue [129].

7) PMut (http://mmb2.pcb.ub.es:8080/PMut/)

PMut predicts the deleterious nature of single nucleotide variations using neural networks. PMUT operates on two levels:

- It obtains data from local databases for mutational hotspots
- It evaluates a given SNP in a protein [130]

8) SNAP2 (https://rostlab.org/owiki/index.php/Snap2)

Evolutionary information extracted from multiple sequence alignment is crucial input signal for the prediction. Distinction between synonymous and non-synonymous is made by taking various sequences and its features. Score prediction is made in the range of -100 to +100 (-100 represents strong neutrality and +100 represents the diseased condition[88].

9) SNPs&GO (https://snps.biofold.org/snps-and-go/snps-and-go.html)

It uses a SVM-based classifier in order to identify the deleterious SNPs [122].

3.3 Structure-based tools to identify deleterious SNPs

Six Structure based tools were used to identify which missense SNPs were found to deleterious

1) I-Mutant

It predicts the stability of protein upon single nucleotide variation. I-Mutant completes following tasks:

- How single point mutation affects the direction of protein stability when the tertiary structure is known.
- $\Delta\Delta$G can be calculated from the given tertiary structure
- To determine the direction of protein stability from the sequence of protein.
- $\Delta\Delta$G can be calculated using the sequence of protein

For $\Delta\Delta$G more than 0 indicates the stability of protein whereas the value of $\Delta\Delta$G in negative indicates that mutation is destabilizing. Once $\Delta\Delta$G value is calculated, RI can also be predicted [132].

2) DUET

DUET is a bioinformatics tools that predicts the stability of protein due to missense SNPs. DUET predicts the effect of missense mutations by combining the results of SDM and mCSM. SDM calculates the free energy change by comparison between the folded and unfolded state of both wild-type and mutants. mCSM depicts the properties of wild-type residue in the form of graph [133].

3) CUPSAT

In order to predict $\Delta\Delta$G value, CUPSAT utilizes torsional angle and specific atom potentials wherein the output gives information about any change in the stability of protein and features such as secondary structure, torsional strain in the mutated region[134].

4) MUpro

A machine learning program used to identify the effect of substitution on the stability of protein. The advantage of this tool is that even if the structure of protein is not known, the sequence can be used to predict the consequences of SNPs. The accuracy obtained from both sequence and structure-based information is 84.2% and 84.5%.

5) DynaMut

It is a tool used to critically evaluate the changes in dynamics of protein in order to properly comprehend the molecular repercussions of mutation[135].

6) Align-GVGD

It is a tool that uses Multiple Sequence Alignment in order to describe the biochemical properties.

Method of disease prediction:

- The difference in the biochemical property at each alignment place is translated into Granthan Variation score (GV)
- Then a Grantham difference score is generated on the basis of difference in the properties to that of substituted amino acids
- A spectrum is formed which gives idea whether the mutation will affect the biological function of protein. A range of spectrum from (C0,C15,C25,C35,C45,C55,C65) is given with C65 being the most deleterious whereas C0 the least likely to effect the biological function[136].

3.4 Conserved regions identification of SERPINI1 gene

Identification the conserved regions is governed by a bioinformatics tool called **ConSurf** which identifies the conserved regions within the nucleic acid or amino acids based on phylogenetic relationships between homologous sequences. This leads to identification of regions which are crucial for structure and functioning of protein. In order to build an evolutionary relationship, a multiple sequence alignment is carried out for the query sequence and a phylogenetic tree is

constructed. This information is then used to identify the rate of evolution of each position within the sequence [137]. Conservation scores are calculated in order to identify the rate of evolution. The sites which evolve slowly are termed as conserved regions whereas some evolve at a faster rate and are termed as variable regions. The evolutionary rate is calculated by two methods i.e. Bayesian and Maximum Likelihood method. Lower the conservation score, highly conserved the region is and higher the conservation score, the sequence is less likely to be conserved [138]

3.5 Analysis of interactions of SERPINI1 gene using STRING

A database which is used to predict the interaction of the query protein with other proteins. This database contains information about 5090 organisms and 24.6 million proteins. Physical interaction is not required for the two proteins to be associated with each other, instead they can share some functions and contribute to the same pathway. However, the interactions must be precise enough in order to develop a functional map. The association of two proteins is expressed in the form of interaction scores which gives a confidence score whether the association is true or not[139].

3.6 Molecular dynamic simulation (MDS)

MDS was carried out for monomeric structure of SERPINI1 gene using GROMACS version 2019.4. Development of protein framework was done using gromos54a7 force field. Gmxeditconf tool was used to develop the cubic simulation box. Gmx solvate tool was used to carry out solvation using point-charge water model. Electro-neutralization of the complex was carried out using gmxgenion tool. Optimization of structure by minimizing the energy of the molecule and removal of steric clashes was carried out. Equilibrium of the system was set by heating the system up to 300K and maintaining the pressure and density of the system. The structures obtained from NPT equilibration phase were subjected to a simulation time of 100 ns [122].

3.7 Orientation analysis

GROMACS analysis tool was used to in order to determine the trajectory of the molecules. Root mean square deviation (RMSD) and root-mean square fluctuations were determined for the wild and mutant proteins using gmxrms and gmxrmsf tools. Inter-particle hydrogen bonding, solvent-accessible surface area (SASA) and radius of gyration of the protein was calculated. Investigation of secondary structure was done with the help of do_dssp tool [140].

3.8 Molecular docking

A computational tool which is used for understanding the interactions between protein and ligand. Here the protein taken was Neuroserpin and the ligand was ascorbic acid. Docking helps in understanding the confirmation and the binding free energy of the interacting molecules [141]. In order to calculate the ligand binding confirmation auto dock makes use of semi-empirical free energy force. Preparation of ligand and protein was carried out. Hydrogen atoms and atomic charge was added to the structure [142].

3.9 MTT assay

Procedure:

1) **Chemicals and equipments**

Gibco Dulbecco's Modified Eagle's Media (DMEM), 0.05% Trypsin-EDTA and Penstrep (Penicillin Streptomycin), MTT, Dimethyl sulphoxide, Phosphate saline buffer, Fetal bovine serum (FBS), Ascorbic Acid and were purchased from Thermo-Fisher Scientific, India .

2) **Cell-culture and Viability test**

Neuro 2A (N2A) cell line was procured from National Centre of Cell Sciences, Pune, India. N2A cell line is derived from the neural crest cells of mouse, has neuronal and amoeboid stem cell morphology that has been extensively used to study neuronal differentiation, axonal growth and signaling pathways. Cell lines were maintained in DMEM with 10% (v/v) FBS and 1% Penstrep antibiotic at 5% CO_2 concentration and 37℃ temperature in an incubator.

3) **MTT assay**

The MTT assay is used to measure cellular metabolic activity as an indicator of cell viability, proliferation and cytotoxicity. This colorimetric assay is based on the reduction of a yellow tetrazolium salt (3-(4,5-dimethylthiazol-2-yl)-2,5- diphenyltetrazolium bromide or MTT) to purple formazan crystals by metabolically active cells. The viable cells contain NAD(P)H-dependent oxidoreductase enzymes which reduce the MTT to formazan. The insoluble formazan crystals are dissolved using a solubilization solution and the resulting-colored solution is quantified by measuring absorbance at 500-600 nanometers using a multi-well spectrophotometer. It is a quantitative assay that allows rapid and convenient handling of a high number of samples. Approximately 1×10^5 mL^{-1} cells in their exponential growth phase were seeded in a flat-bottomed 96-well plate and were incubated for 24 hours at 37°C in a CO_2 incubator. After 24 hours of incubation Ascorbic acid at varied concentrations of 25, 50,75,100 ug/ml was added and cells were incubated for 24 hours at in a 5% CO_2. After 24 hours of incubation,10 µL of MTT reagent was added to each well and was further incubated for 4 hours. Cell viability was measured after

recording the absorbance at 570 nm with a Multiskan Thermo Fisher Microplate Spectrophotometer reader

CHAPTER 4

RESULTS AND DISCUSSION

==========

1) SNPs for SERPINI1 GENE in dbSNP

A total of 33,072 SNPs were obtained from dbSNP, out of which 337 missense SNPs were selected for screening in order to identify the deleterious SNPs associated with disease. Missense SNPs are the SNPs present in coding-regions which effect the structure and function of protein

Table 4: No and types of SNP retrieved from dbSNP

Types of SNP	No of SNPs
1) Inframe del	2
2) Inframe deletion	2
3) Intron	31830
4) Missense	337
5) Synonymous	125

2) Screening using sequence-based tools

337 missense SNPs were taken and screened using sequence-based tools: PROVEAN, PANTHER, Poly-Phen2, Meta-SNP, PredictSNP1, SusPect, PMut, SNAP2, SNPs&GO. The SNPs were screened in order to identify the deleterious nature of SNPs

Table 5: This table represents the total number of SNPs in various sequence-based servers

TOOLS	TOTAL NO. OF DELETERTIOUS SNPs
Provean	102
Panther	129
Poly-Phen2	142
Meta-SNP	110
Predict SNP1	98
SusPect	66
PMut	130
Snap2	101
SNPs&GO	101

Parameters which identified the SNPs to be deleterious:

1) **Provean:** BLAST is performed by clustering of 30 closely related sequences which is used to generate the alignment score and an average is set in order to develop the final threshold score for Provean. The threshold score is -2.5 and a score equal to or less than -2.5 is found to be deleterious.
2) **Panther:** On the basis of the score of position-specific evolutionary preservation Panther identifies whether the mutation is damaging or not. Three different parameters are set on the basis of which the SNPs are classified as deleterious i.e.

- **Probably damaging:** If preservation time is more than 450 million years with a false positive rate of 0.2
- **Possibly damaging:** If the preservation time is between 450my to 250my with a false positive rate of 0.4
- **Probably benign:** If the preservation time is less than 200 my.[143]

3) **Poly-phen2:** It determines the probability of a deleterious mutation using the model of 5%/10% False positive rate (FPR) for HumDiv model and FPR score for HumVar model.

- Mutations are considered to be probably damaging if the probability score are below the first FPR value.
- Mutations are considered probably damaging if the probability score is below the second FPR value
- Mutations are said to be benign if FPR value is above the second higher value[144]

4) **Meta-SNP:** A threshold score is calculated which identifies the SNP to be deleterious or not. A threshold score greater than 0.5 confirms a SNP to be deleterious [145]
5) **Predict-SNP1:** A confidence score was set for the SNP to be deleterious. If the score was 0.5 then the mutation was deleterious[128]
6) **PMut:** It classifies the SNP to be deleterious or neutral on the basis of a prediction score set between 0 to 1. Mutations in the scoring parameter of 0 to 0.5 are classified as neutral and in the range of 0.5 to 1 are considered deleterious.[130]
7) **SusPect:** It produces a color-coding system (Blue represents neutrality and red represents disease causing SNP) in order to identify the deleterious SNPs and a scoring system table from a score of 0-100 is also produced wherein the value of 50 is considered as threshold value above which a variant is associated to cause disease
8) **Snap2:** A scoring system is developed which classifies the SNP to be deleterious. Scoring from -100 to +100 is produced wherein all the predictions above 0 are deleterious and all the predictions below or equal to 0 are neutral
9) **SNPs&GO:** Reliability index is used for prediction of disease associated SNPs. If the probability is greater than 0.5 then the SNP is found to be deleterious[146].

On the basis of these scoring parameters various structure-based tools predict a SNP to be deleterious or neutral

Table 6: This table represents the SNPs which were found to be deleterious in all nine sequence-based servers

ID	Mutation	PROV-EAN	PANTHER	Poly-Phen2	Meta-SNP	Predict SNP1	SusPect	Pmut	SNAP2	SNPs&GO
rs121909051	S49P	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease

rs121909053	G392E	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs121909054	G392R	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs759200368	N152K	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1466251791	L257P	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1577418477	S52R	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs140384336	L307S	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs201722989	G63R	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs761819286	N110K	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs767376971	G58R	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs769089864	N45D	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs772860274	A352V	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs779089806	F132L	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1252549703	P293L	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1373128074	T199S	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1439828732	S252G	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1448356396	Y185C	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1711681549	G311E	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1727417409	P50Q	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1727419831	A64P	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1727465744	A109T	Delete-	Probably	Probably	Disease	Disease	Disease	Disease	Effect	Disease

		rious	damaging	damaging						
rs1727656927	L267P	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1727659223	P293S	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1560020646	I380T	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1727575153	Y225C	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs61761892	T199I	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease
rs1712452015	R381G	Delete-rious	Probably damaging	Probably damaging	Disease	Disease	Disease	Disease	Effect	Disease

RESULT: Out of 337 SNPs, only 23 were found to be deleterious in all nine-sequence based servers. These SNPs were then further screened using structure-based tools.

3) Screening using structure-based tools

The 23 SNPs were further screened using structure-based tools. Prediction was carried out using 6 structure-based tools i.e. I-Mutant, MUpro, DynaMut, Align-GVGD, CUPSAT, DUET. Out of 23 SNPs only 7 SNPs were found to be deleterious in all six servers.

Table 7: This table represents the mutations which were destabilizing in all six structure-based servers

Variant ID	**Mutation**	**I-Mutant**	**DUET**	**CUPSAT**	**MUpro**	**DynaMut**	**Align-GVGD**
rs121909053	G392E	DECREASE	DESTABILIZING	DESTABILIZING	DECREASE	DESTABILIZING	Class C65
rs121909054	G392R	DECREASE	DESTABILIZING	DESTABILIZING	DECREASE	DESTABILIZING	Class C65
rs1466251791	L257P	DECREASE	DESTABILIZING	DESTABILIZING	DECREASE	DESTABILIZING	Class C65
rs140384336	L307S	DECREASE	DESTABILIZING	DESTABILIZING	DECREASE	DESTABILIZING	Class C65

rs1373128074	T199S	DECREASE	DESTABILIZING	DESTABILIZING	DECREASE	DESTABILIZING	Class C65
rs1727465744	A109T	DECREASE	DESTABILIZING	DESTABILIZING	DECREASE	DESTABILIZING	Class C65
rs1712452015	R381G	DECREASE	DESTABILIZING	DESTABILIZING	DECREASE	DESTABILIZING	Class C65

Parameters which identified the SNP to be destabilizing

1) **I-Mutant**: It predicts the effect of single nucleotide variation on the stability of protein and mutations are classified as:

- Neutral mutation: A score between $-0.5 \leq DDG \leq 0.5$
- Large decrease: A score of ≤ -0.5
- Large Increase: A score of >0.5[147]

2) **MUpro**: A confidence score is used to identify the stability of protein. Confidence score less than zero decreases the stability of protein and a confidence score more than 0 leads to increase in the stability of protein
3) **DynaMut**: The value of difference in binding affinity of mutant and wild type protein complex is defined as $\Delta\Delta G$. If this value is less than 0 then the mutation is thought to be destabilizing and if the value is above zero the mutation is stabilizing.[135]
4) **Align-GVGD:** Spectrum is developed with a range of (C0, C15, C25, C35, C45, C55, C65) where C65 is found to most deleterious and C0 is least deleterious
5) **CUPSAT:** $\Delta\Delta G$ values predicts the stability of protein
6) **DUET:** Change in folding free energy upon mutation is calculated. A Positive value indicates the mutation to be destabilizing and negative value indicates the mutation to be stabilizing

4) Identification of conserved residues using ConSurf

The seven SNPs obtained from structure-based tools were screened for evolutionary regions and it was discovered that all seven SNPs were highly conserved.

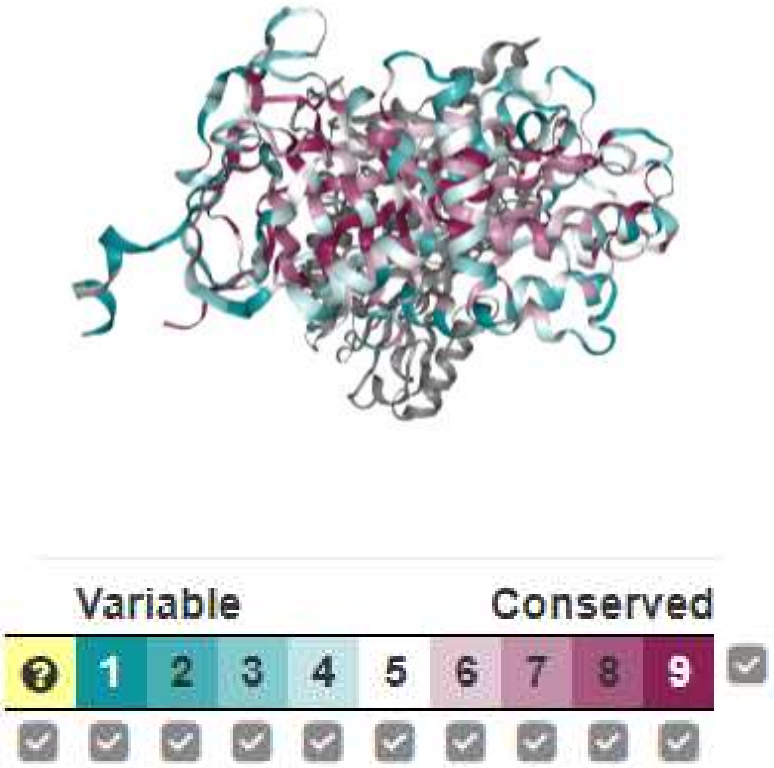


Fig 19: This figure represents the conservation analysis of chain A of SERPINI1 gene

In this figure, the score of 1 denotes the least conserved residue whereas the score of 9 denotes the most conserved residue. ConSurf developed a scoring table for the 7 SNPs which were selected in structure-based tools and it was found that all SNPs were evolutionary conserved. Conservation was found at these residues

POSITION	RESIDUE	CONSERVATION SCORE
392	G	9
392	G	9
257	L	9
307	L	9
199	T	9
109	A	9
58	G	7

5) Multiple sequence alignment to identify the conservation of deleterious SNPs

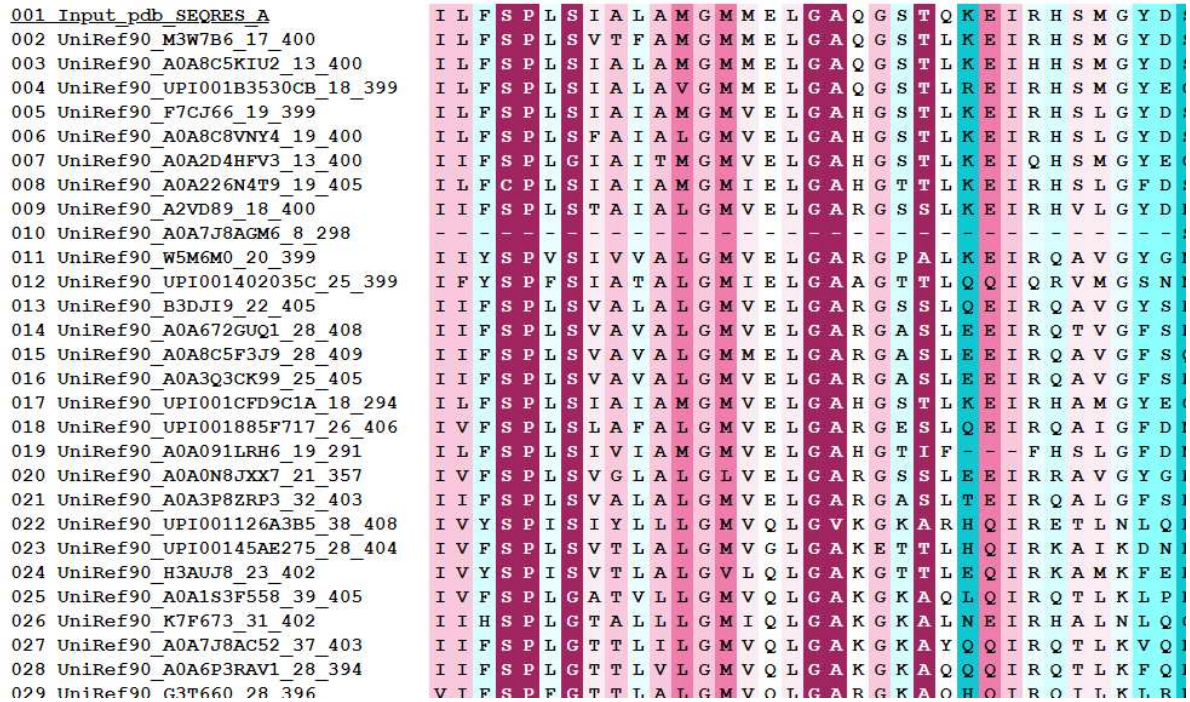


Fig 20: This figure represents the results of Multiple sequence alignment

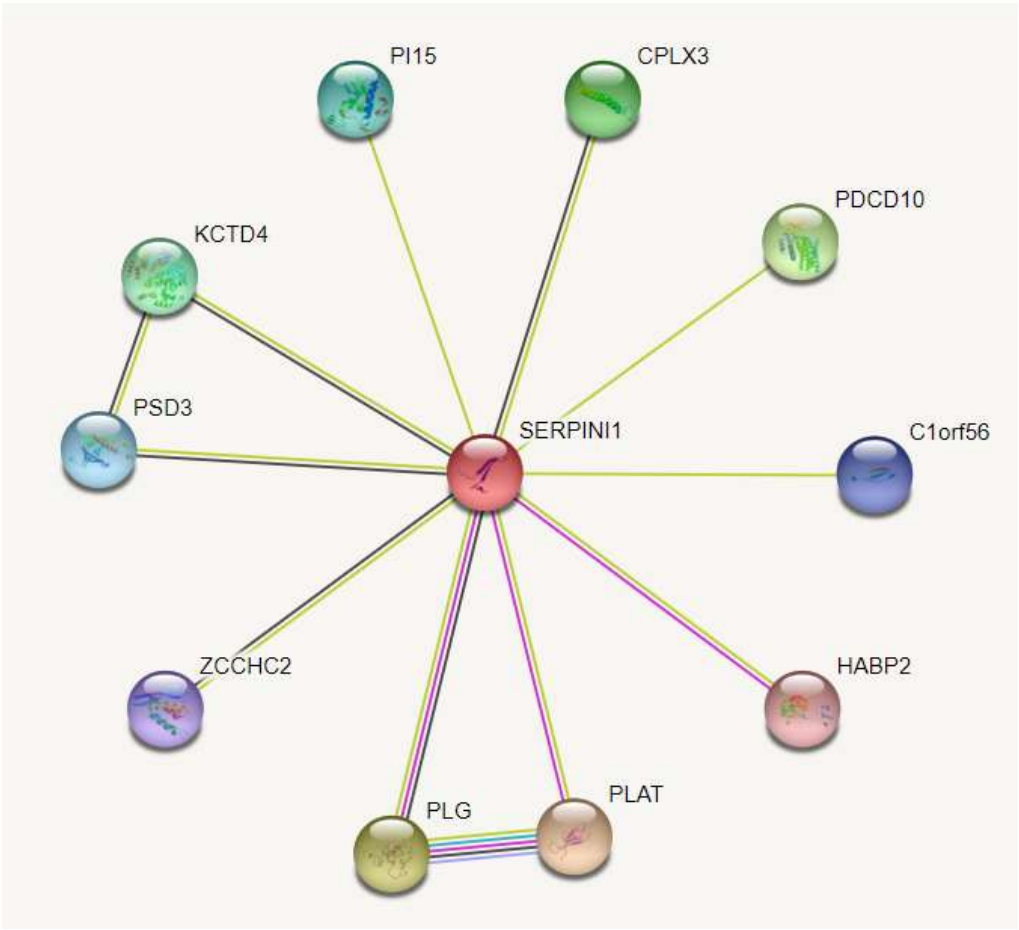
Multiple sequence alignment is carried out in order to check the conservation of residues. Gaps or – indicates insertion or deletion at that point. Here the regions S, P, S G, A represent the conserved regions.

6) Interaction analysis of SERPINI1 gene using STRING database

Information in the form of nodes and edges is represented where in the network nodes represent the proteins and edges represent its association with other proteins.

No of Nodes	11
No of edges	12

Table 8: This table represents total number of nodes and edges found upon analysis



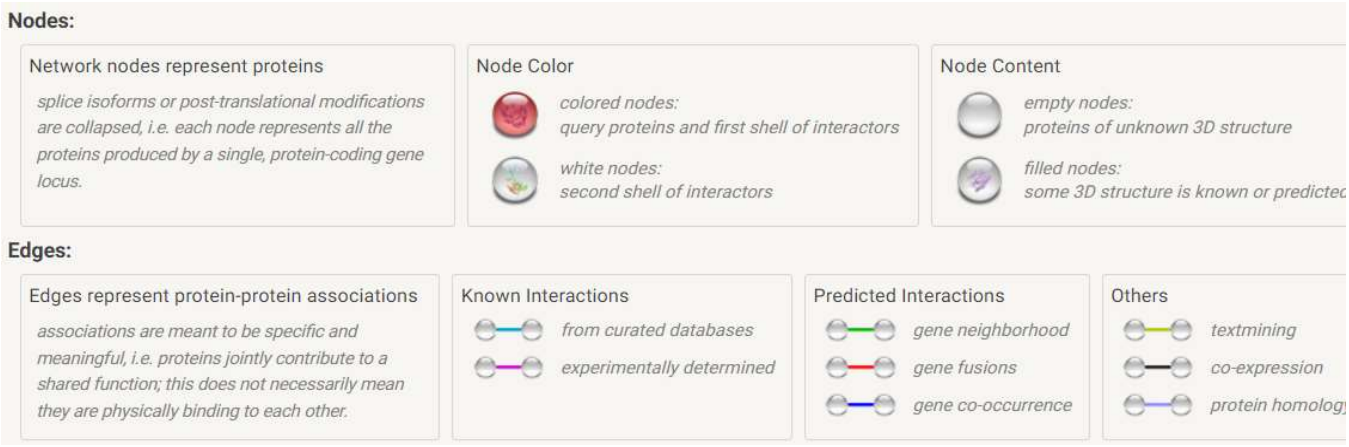


Fig 21: This figure represents the interaction network of SERPINI1 gene using STRING database

Upon analysis it was found that 10 protein-coding genes were in association with SERPINI 1 gene and were its predicted functional partners depending upon the scoring system.

Table 9: This table gives information about all the functional partners of SERPINI1 gene

Functional Partner	Score	Description
PLAT	0.965	Encodes for tissue type plasminogen activator and is involved in conversion of inactive plasminogen to plasmin thereby playing an important role in degradation, neuronal migration
PLG	0.783	It encodes for plasminogen which in turn is further converted to plasmin upon the action of various plasminogen activators which dissolves the fibrin involved in blood clots. It is also involved in development of embryo, tissue remodeling, tumor invasion
PDCD10	0.720	Encodes for programmed cell death protein 10 which is involved in proliferation of cell, regulates apoptotic pathways, cell-migration and assembly of Golgi body, required for normal angiogenesis, vasculogenesis
CPLX3	0.692	It is involved in the regulation of step involved in synaptic vesicle exocytosis
KCTD4	0.667	Potassium channel tetramerization domain containing 4
PI15	0.653	Peptidase inhibitor 15 is a serine protease inhibitor and weakly inhibits trypsin
PSD3	0.638	It encodes for Pleckstrin and Sec7 domain-containing protein 3
C1orf56	0.626	Chromosome 1 open reading frame 56 encodes for MENT protein which is involved in cellular proliferation
ZCCHC2	0.608	Zinc finger cchc domain-containing protein 2

HABP2	0.605	Encodes for Hyaluronan-binding protein 2 which is involved in activation of coagulation factor VII. It also functions as tumor suppressor

STRING database also gives idea about the various biological processes SERPINI1 gene is involved in with information about various associated proteins

Table 10: This table represents the processes in which SERPINI1 gene is involved in and its interaction partners

GO accession	Biological Process	Gene count	Strength	False discovery rate	Matching proteins in network
GO:0042730	Fibrinolysis	6	2.45	2.02E-09	PLAT, SERPINE1, HRG, PLG, PLAUR, ANXA2
GO:1903034	Regulation of response to wounding	8	1.64	4.49E-08	PLAT, SERPINE1, HRG, LRP1, PLG, PLAUR, ANXA2, STK24
GO:0045861	Negative regulation of proteolysis	8	1.32	2.54E-06	PLAT, SERPINE1, HRG, LRP1, PI15, SERPINI1, PLAUR, ANXA2
GO:0010466	Negative regulation of peptidase activity	7	1.41	7.70E-06	SERPINE1, HRG, LRP1, PI15, SERPINI1, PLAUR, ANXA2
GO:0051918	Negative regulation of fibrinolysis	3	2.45	0.00028	SERPINE1, HRG, PLG
GO:0001568	Blood vessel	7	1.12	0.00046	SERPINE1,

	development				HRG, LRP1, PLG, ANXA2, CCM2, PDCD10
GO:0080134	Regulation of response to stress	10	0.81	0.0005	PLAT, SERPINE1, HRG, LRP1, PLG, STK25, PLAUR, ANXA2, STK24, PDCD10
GO:1903036	Positive regulation of response to wounding	4	1.73	0.00059	SERPINE1, HRG, LRP1, PLG
GO:0008631	Intrinsic apoptotic signaling pathway in response to oxidative stress	3	2.22	0.00067	STK25, STK24, PDCD10

7) Molecular dynamics simulation

Simulation was carried out in order to determine the confirmational difference between the wild-type and mutant proteins. MDS was carried out for wild-type protein, G392E, G392R, L257P, L307S, T199S, A109T, R381G for 100ns using GROMACS.

7.1 Analysis of Stability

In order to determine the stability of wild-type and mutant proteins RMSD was carried out. The stability can be predicted by deviations of RMSD produced through MDS. The protein is found to be more stable if the RMSD deviations are small. To identify the discrepancy in the structural confirmation, RMSD of backbone of protein atoms was plotted against time The average RMSD value of wild type protein was 0.426. Also, the values for mutants were as follows: A109T (0.295), T199S (0.343), L257P (0.335), L307S (0.446), R381G 0.2922, G392E (0.414), G392R (0.5133). It was found that mutants G392R and L307S had higher average values and thus are less stable than wild type. Mutant G392E had no effect on

the stability of protein. Mutants such as A109T, T199S, L257P, R381G were more stable than the wild-type since their average values are less than that of wild-type.

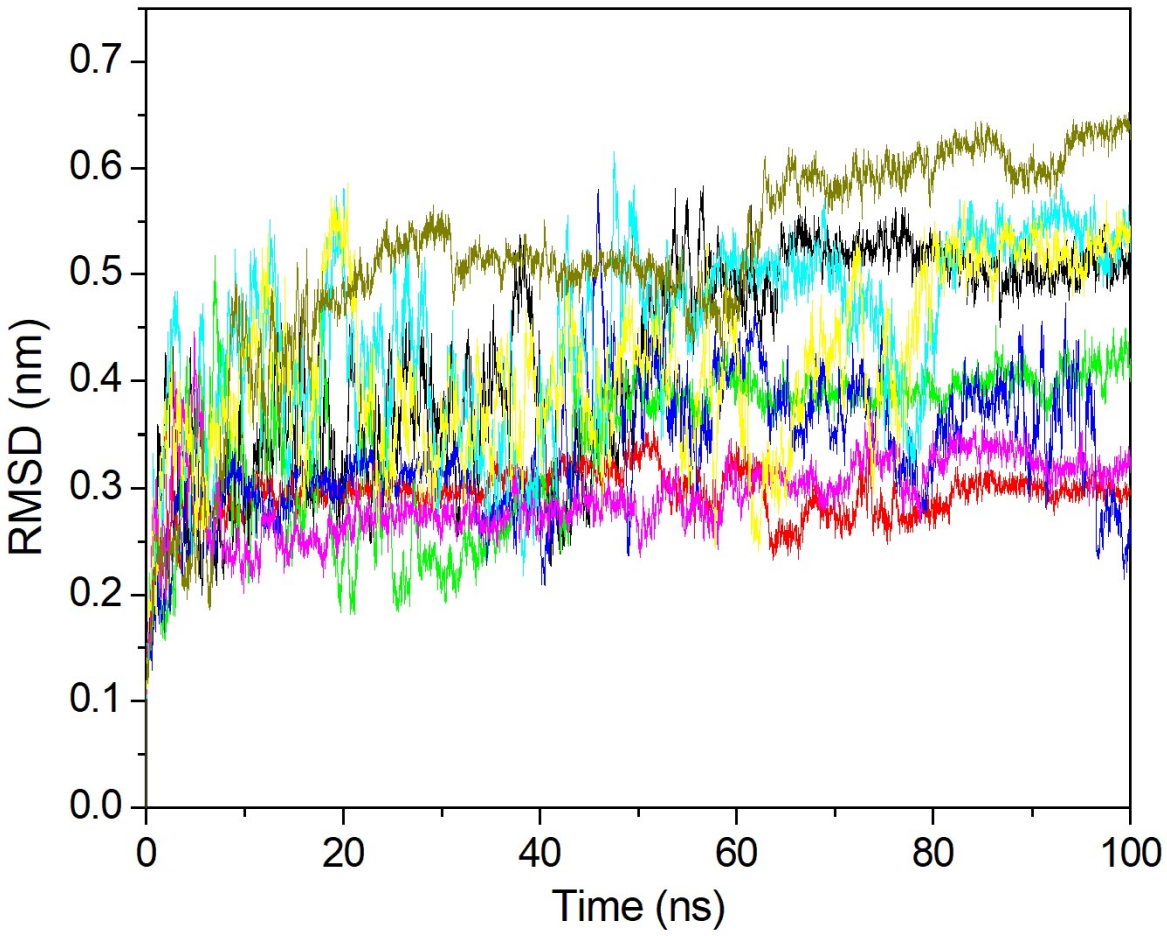


Fig 22: MDS simulation for wild type and mutants (RMSD)

7.2 Flexibility Analysis

Flexibility of the molecules can be calculated using RMSF values. More the RMSF value more is the flexibility of the wild-type and mutants within MDS. For wild type protein the average value was 0. 135nm.The average values for mutants were found to be A109T (0.122), T199S (0.143), L257P (0.127), L307S (0.135), R381G (0.112), G392E (0.143), G392R (0.130). It was found that mutants T199S and G392E were more flexible than wild-type since their average values are more than wild-type.

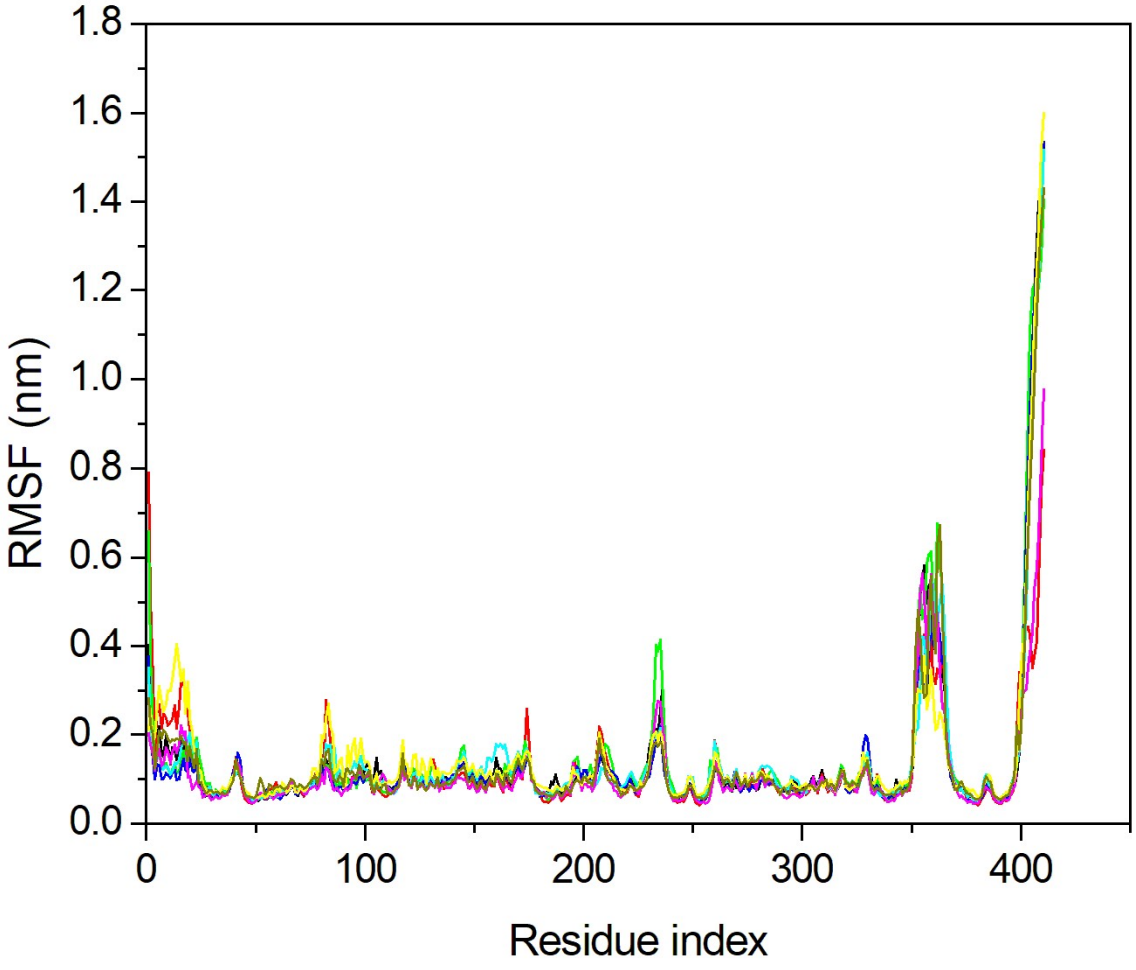


Fig 23: MDS for wild and mutants (RMSF)

7.3 Analysis of compactness

The distance between the two atoms of a protein from the axis of rotation is termed as radius of gyration (Rg). More the value of Rg, less compact and flexible the protein is. For wild type protein the average value was 2.304 whereas for mutants the values were A109T (2.25), T199S (2.28), L257P (2.261), L307S (2.312), R381G (2.28), G392E (2.296, G392R (2.281) nm. Since mutant L307S had more Rg value hence it is highly compact and flexible.

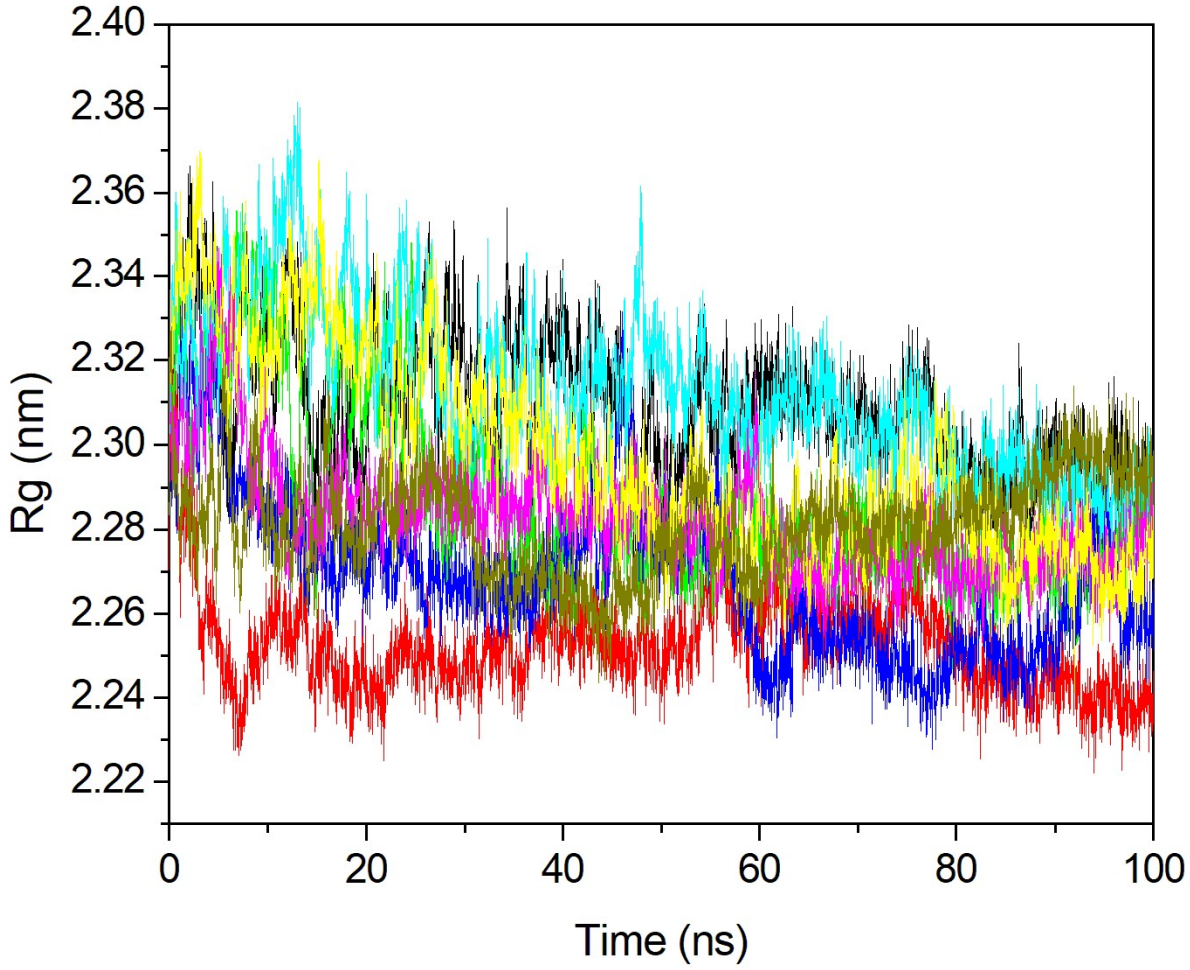


Fig 24: MDS for wild and mutant type (Rg)

7.4 Hydrogen-bond analysis

Hydrogen bonds are a key factor involved in stability of the protein. More the number of hydrogen bonds more stable the protein is and vice-versa. For wild type protein the average number of hydrogen bonds were found to be 327.86. For mutants the number of hydrogen bonds were as follows: A109T (330.78), T199S (319.93) L257P (329.91), L307S (329.62), R381G (325.12), G392E (325.152), G392R (327.68). Mutants R381G and G392R were found to be less stable than wild-type

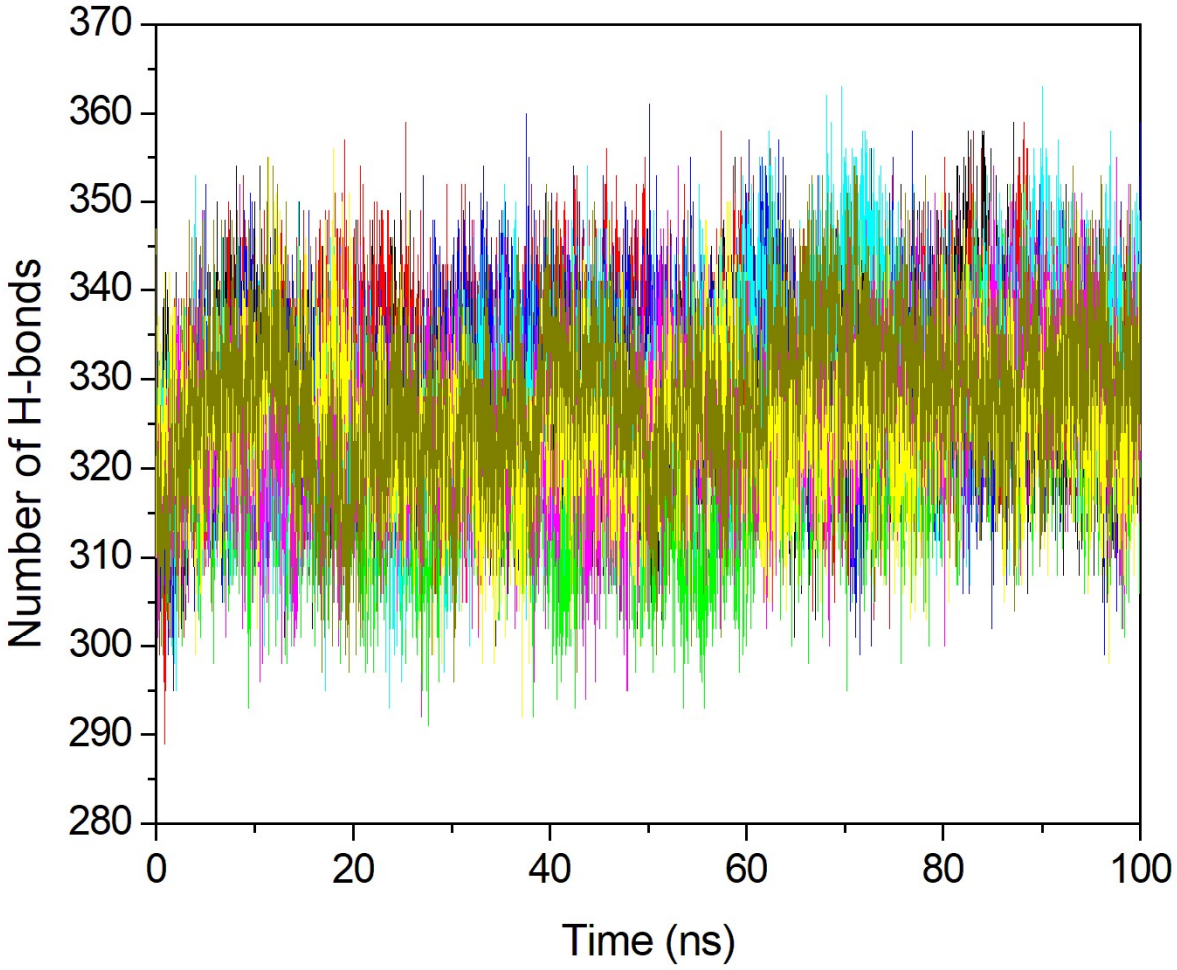


Fig 25: MDS for wild type and mutants (Hydrogen-bond)

7.5 Solvent accessible surface area analysis (SASA)

This factor explains the accessibility of residue of protein to water solvent. More the SASA value more stable the protein is and vice-versa. The average value of wild-type protein is 197.01 and that of the mutants were A109T (197.01), T199S (195.22), L257P (192.07), L307S (195.31), R381G (194.06), G392E (196.68), G392R (191.99).

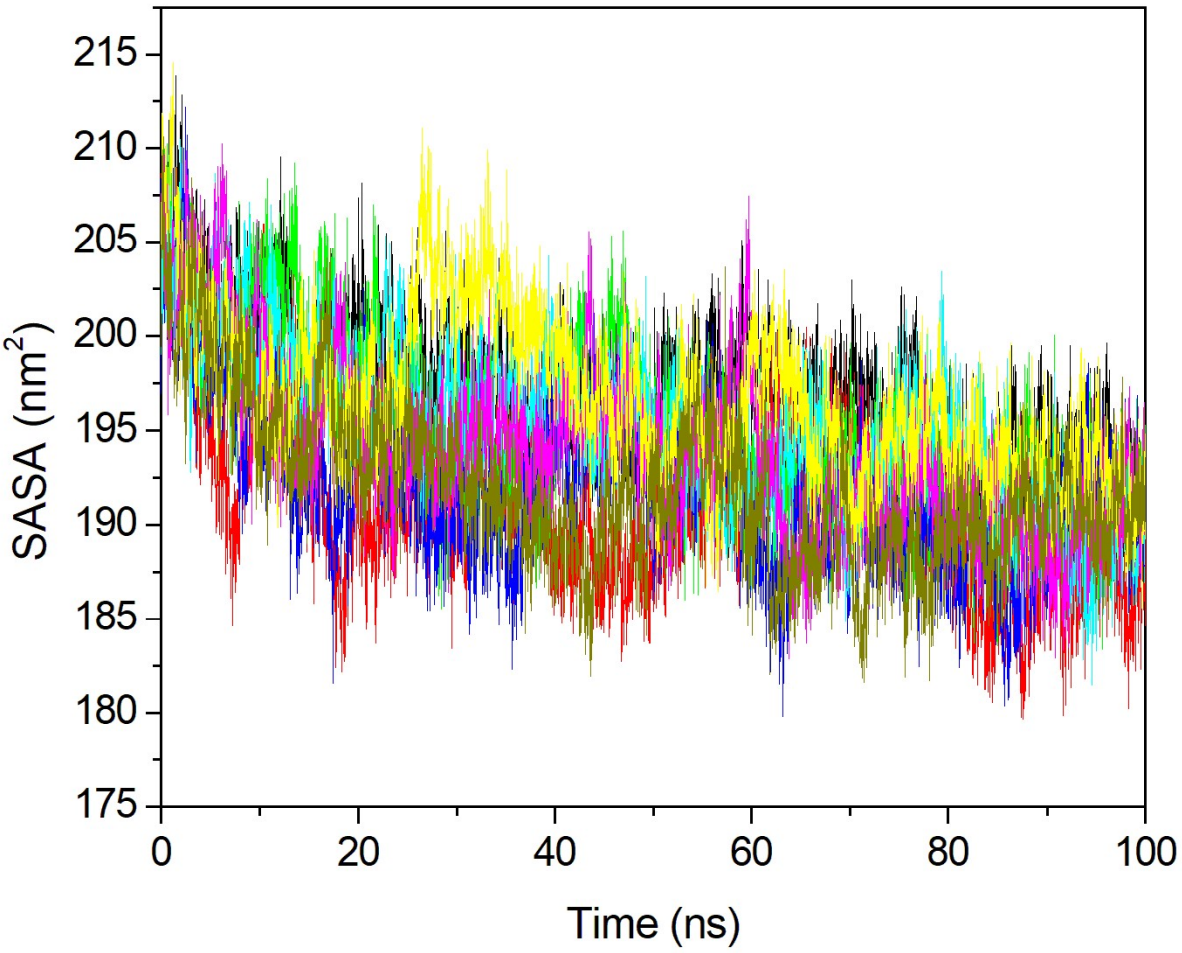


Fig 26: RMSD for wild and mutant type protein (SASA)

Table 11: This table represents the values of different parameters obtained from simulation

Mutations	RMSD	RMSF	gyration	Hydrogen bond	SASA
Wild-type	0.426	0.135	2.304	327.86	197.01
A109T	0.295	0.122	2.25	330.78	191.22
T199S	0.343	0.143	2.28	319.93	195.25
L257P	0.335	0.127	2.261	329.91	192.07
L307S	0.446	0.135	2.312	329.62	195.31
R381G	0.292	0.112	2.28	325.12	194.06
G392E	0.414	0.143	2.296	325.152	196.68
G392R	0.513	0.130	2.281	327.68	191.99

The flexibility of structure is reduced and rigidity is increased in the following three mutations (A109T, L257P, R381G) hence these mutations were found to be lethal

8) Molecular docking

Upon docking it was found that the binding energy of the complex was -3.89 kcal/mol and inhibition constant was 1.41mM. Also, it was found that ascorbic acid binds at Glutamic acid 101, Lysine107, Valine97, Methionine 106 of Neuroserpin indicating the interactions among them

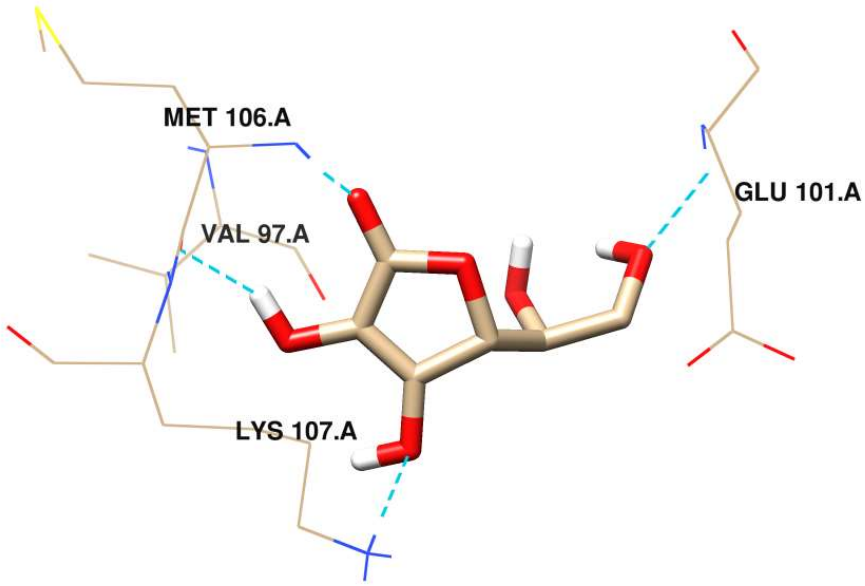


Fig 27: Molecular docking of Neuroserpin (protein) and ascorbic acid (ligand)

9) MTT Assay

From our results (Fig 28 and table 12), after treatment of ascorbic acid, significant viability enhancement effects were observed in Neuro 2A cell lines at 25, 50, 75, 100 µg/ml with proliferation rate of 60.5%, 82.9%, 95.3% and 109%.

Table 12: The effect of ascorbic acid on Neuro2A cells

Concentration	% Cell Proliferation
25	60.511
50	82.95
75	95.31
100	109.80

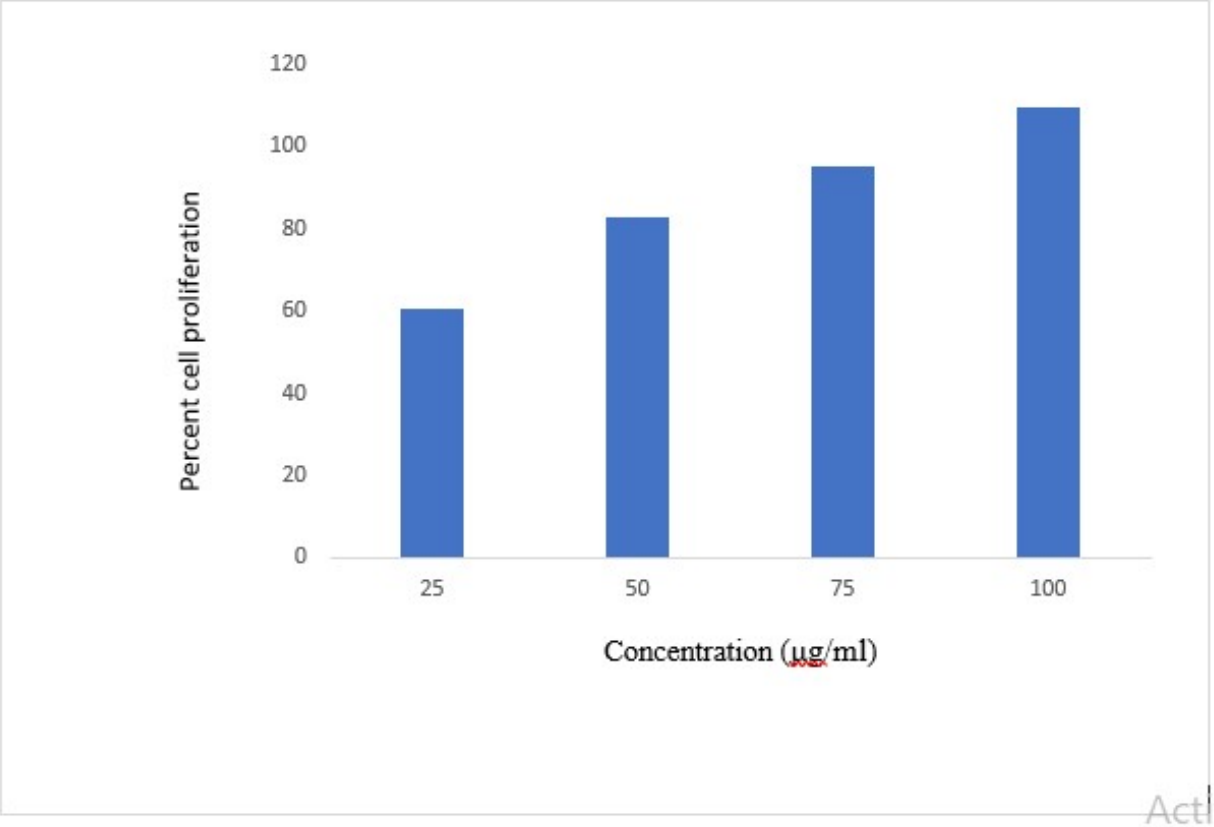


Fig 28: This graph represents the rate of proliferation at varied concentrations

Our in vitro studies demonstrated that ascorbic acid significantly enhanced neuro cells proliferation. Ascorbic acid was involved in promoting neuronal growth. Thus, it could be concluded that ascorbic acid could interact with SERPINI1 gene and downregulate AD.

CONCLUSION

Out of 337 missense SNPs, 23 were found to be deleterious in all nine sequence-based servers.

These 23 SNPs were further screened using structure-based tools out of which seven SNPs were found to be destabilizing Further simulation was carried out and three mutations (A109T, L257P, R381G) were found to be lethal. Molecular docking revealed binding of Ascorbic acid to Neuroserpin. MTT assay was carried out in order to determine the cell-viability and proliferative effects were seen on Neuro 2A cells.

REFERENCES

[1] M. J. Caire, V. Reddy, and M. Varacallo, "Physiology, Synapse," in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2022. Accessed: May 12, 2022. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK526047/

[2] W. C. Abraham, "How long will long-term potentiation last?," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 358, no. 1432, pp. 735–744, Apr. 2003, doi: 10.1098/rstb.2002.1222.

[3] "Protein Misfolding in Conformational Disorders: Rescue of Folding...: Ingenta Connect." https://www.ingentaconnect.com/content/ben/mrmc/2008/00000008/00000009/art00005 (accessed May 14, 2022).

[4] S. S. Adav and S. K. Sze, "Insight of brain degenerative protein modifications in the pathology of neurodegeneration and dementia by proteomic profiling," *Mol. Brain*, vol. 9, no. 1, p. 92, Nov. 2016, doi: 10.1186/s13041-016-0272-9.

[5] V. J. De-Paula, M. Radanovic, B. S. Diniz, and O. V. Forlenza, "Alzheimer's Disease," in *Protein Aggregation and Fibrillogenesis in Cerebral and Systemic Amyloid Disease*, J. R. Harris, Ed. Dordrecht: Springer Netherlands, 2012, pp. 329–352. doi: 10.1007/978-94-007-5416-4_14.

[6] H. D. Yang, D. H. Kim, S. B. Lee, and L. D. Young, "History of Alzheimer's Disease," *Dement. Neurocognitive Disord.*, vol. 15, no. 4, pp. 115–121, Dec. 2016, doi: 10.12779/dnd.2016.15.4.115.

[7] "Alzheimer's disease: Symptoms, stages, causes, and treatments," Sep. 01, 2020. https://www.medicalnewstoday.com/articles/159442 (accessed May 14, 2022).

[8] B. Reisberg, S. H. Ferris, M. J. de Leon, and T. Crook, "The Global Deterioration Scale for assessment of primary degenerative dementia," *Am. J. Psychiatry*, vol. 139, no. 9, pp. 1136–1139, Sep. 1982, doi: 10.1176/ajp.139.9.1136.

[9] E. Nichols *et al.*, "Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019," *Lancet Public Health*, vol. 7, no. 2, pp. e105–e125, Feb. 2022, doi: 10.1016/S2468-2667(21)00249-8.

[10] H. J. Tey and C. H. Ng, "Computational analysis of functional SNPs in Alzheimer's disease-associated endocytosis genes," *PeerJ*, vol. 7, p. e7667, Sep. 2019, doi: 10.7717/peerj.7667.

[11] "The metabolic and molecular bases of inherited disease seventh edition: Edited by C R Sriver, A L Beaudet, W S Sly and D Valle. P 4605. McGraw-Hill, New York. 1995. ISBN 0-07-909826-6 - Vella - 1996 - Biochemical Education - Wiley Online Library." https://onlinelibrary.wiley.com/doi/abs/10.1016/S0307-4412%2896%2980019-7 (accessed May 19, 2022).

[12] A. Chakravarti, “Population genetics--making sense out of sequence,” *Nat. Genet.*, vol. 21, no. 1 Suppl, pp. 56–60, Jan. 1999, doi: 10.1038/4482.

[13] Z. E. Sauna and C. Kimchi-Sarfaty, “Understanding the contribution of synonymous mutations to human disease,” *Nat. Rev. Genet.*, vol. 12, no. 10, pp. 683–691, Aug. 2011, doi: 10.1038/nrg3051.

[14] T. Osterwalder, J. Contartese, E. T. Stoeckli, T. B. Kuhn, and P. Sonderegger, “Neuroserpin, an axonally secreted serine protease inhibitor.,” *EMBO J.*, vol. 15, no. 12, pp. 2944–2953, Jun. 1996.

[15] “The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature - PubMed.” https://pubmed.ncbi.nlm.nih.gov/11435447/ (accessed May 19, 2022).

[16] S. P. Schrimpf *et al.*, “Human neuroserpin (PI12): cDNA cloning and chromosomal localization to 3q26,” *Genomics*, vol. 40, no. 1, pp. 55–62, Feb. 1997, doi: 10.1006/geno.1996.4514.

[17] M. Yazaki *et al.*, “Biochemical characterization of a neuroserpin variant associated with hereditary dementia,” *Am. J. Pathol.*, vol. 158, no. 1, pp. 227–233, Jan. 2001, doi: 10.1016/S0002-9440(10)63961-2.

[18] S. Ricagno, S. Caccia, G. Sorrentino, G. Antonini, and M. Bolognesi, “Human Neuroserpin: Structure and Time-Dependent Inhibition,” *J. Mol. Biol.*, vol. 388, no. 1, pp. 109–121, Apr. 2009, doi: 10.1016/j.jmb.2009.02.056.

[19] S. T. Olson and P. G. W. Gettins, “Chapter 5 - Regulation of Proteases by Protein Inhibitors of the Serpin Superfamily,” in *Progress in Molecular Biology and Translational Science*, vol. 99, E. Di Cera, Ed. Academic Press, 2011, pp. 185–240. doi: 10.1016/B978-0-12-385504-6.00005-1.

[20] PubChem, “SERPINI1 - serpin family I member 1 (human).” https://pubchem.ncbi.nlm.nih.gov/gene/SERPINI1/human (accessed May 20, 2022).

[21] K. J. Kinghorn *et al.*, “Neuroserpin Binds Aβ and Is a Neuroprotective Component of Amyloid Plaques in Alzheimer Disease *,” *J. Biol. Chem.*, vol. 281, no. 39, pp. 29268–29277, Sep. 2006, doi: 10.1074/jbc.M600690200.

[22] A. Association, “2019 Alzheimer’s disease facts and figures,” *Alzheimers Dement.*, vol. 15, no. 3, pp. 321–387, 2019, doi: 10.1016/j.jalz.2019.01.010.

[23] “Alzheimer’s Disease - an overview | ScienceDirect Topics.” https://www.sciencedirect.com/topics/psychology/alzheimers-disease (accessed May 20, 2022).

[24] “Alzheimer’s disease: genes, proteins, and therapy - PubMed.” https://pubmed.ncbi.nlm.nih.gov/11274343/ (accessed May 22, 2022).

[25] X. Cao and T. C. Südhof, “A transcriptionally [correction of transcriptively] active complex of APP with Fe65 and histone acetyltransferase Tip60,” *Science*, vol. 293, no. 5527, pp. 115–120, Jul. 2001, doi: 10.1126/science.1058783.

[26] J. Hardy and D. J. Selkoe, “The amyloid hypothesis of Alzheimer’s disease: progress and problems on the road to therapeutics,” *Science*, vol. 297, no. 5580, pp. 353–356, Jul. 2002, doi: 10.1126/science.1072994.

[27] W. Q. Qiu *et al.*, “Insulin-degrading enzyme regulates extracellular levels of amyloid beta-protein by degradation,” *J. Biol. Chem.*, vol. 273, no. 49, pp. 32730–32738, Dec. 1998, doi: 10.1074/jbc.273.49.32730.

[28] “Qiu WQ, Folstein MF. Insulin, insulin-degrading enzyme and amyloid-beta peptide in Alzheimer’s disease: review and hypothesis. Neurobiol Aging. 2006;27:190–198 - Google Search.” https://www.google.com/search?q=Qiu+WQ%2C+Folstein+MF.+Insulin%2C+insulin-degrading+enzyme+and+amyloid-beta+peptide+in+Alzheimer%E2%80%99s+disease%3A+review+and+hypothesis.+Neurobiol+Aging.+2006%3B27%3A190%E2%80%93198&rlz=1C1CHBF_enIN760IN760&oq=Qiu+WQ%2C+Folstein+MF.+Insulin%2C+insulin-degrading+enzyme+and+amyloid-beta+peptide+in+Alzheimer%E2%80%99s+disease%3A+review+and+hypothesis.+Neurobiol+Aging.+2006%3B27%3A190%E2%80%93198&aqs=chrome..69i57.996j0j7&sourceid=chrome&ie=UTF-8 (accessed May 22, 2022).

[29] “Enhancement of β-Amyloid Peptide Aβ(1–40)-Mediated Neurotoxicity by Glutamine Synthetase - Aksenov - 1995 - Journal of Neurochemistry - Wiley Online Library.” https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1471-4159.1995.65041899.x (accessed May 22, 2022).

[30] “Amyloid beta-peptide impairs ion-motive ATPase activities: evidence for a role in loss of neuronal Ca2+ homeostasis and cell death - PubMed.” https://pubmed.ncbi.nlm.nih.gov/7666206/ (accessed May 22, 2022).

[31] X. Sun, W.-D. Chen, and Y.-D. Wang, “β-Amyloid: the key peptide in the pathogenesis of Alzheimer’s disease,” *Front. Pharmacol.*, vol. 6, p. 221, Sep. 2015, doi: 10.3389/fphar.2015.00221.

[32] S. Varadarajan, S. Yatin, M. Aksenova, and D. A. Butterfield, “Review: Alzheimer’s amyloid beta-peptide-associated free radical oxidative stress and neurotoxicity,” *J. Struct. Biol.*, vol. 130, no. 2–3, pp. 184–208, Jun. 2000, doi: 10.1006/jsbi.2000.4274.

[33] U. Neniskyte, J. J. Neher, and G. C. Brown, “Neuronal Death Induced by Nanomolar Amyloid β Is Mediated by Primary Phagocytosis of Neurons by Microglia *,” *J. Biol. Chem.*, vol. 286, no. 46, pp. 39904–39913, Nov. 2011, doi: 10.1074/jbc.M111.267583.

[34] “TLR2 Is a Primary Receptor for Alzheimer’s Amyloid β Peptide To Trigger Neuroinflammatory Activation | The Journal of Immunology.” https://www.jimmunol.org/content/188/3/1098 (accessed May 22, 2022).

[35] “Innate immunity in Alzheimer’s disease | Nature Immunology.” https://www.nature.com/articles/ni.3102 (accessed May 22, 2022).

[36] A. Metaxas and S. J. Kempf, “Neurofibrillary tangles in Alzheimer’s disease: elucidation of the molecular mechanism by immunohistochemistry and tau protein phospho-proteomics,” *Neural Regen. Res.*, vol. 11, no. 10, pp. 1579–1581, Oct. 2016, doi: 10.4103/1673-5374.193234.

[37] W. Noble, A. M. Pooler, and D. P. Hanger, “Advances in tau-based drug discovery,” *Expert Opin. Drug Discov.*, vol. 6, no. 8, pp. 797–810, Aug. 2011, doi: 10.1517/17460441.2011.586690.

[38] M. Morris, S. Maeda, K. Vossel, and L. Mucke, “The Many Faces of Tau,” *Neuron*, vol. 70, no. 3, pp. 410–426, May 2011, doi: 10.1016/j.neuron.2011.04.009.

[39] “Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer’s disease: Neuron.” https://www.cell.com/neuron/pdf/0896-6273(89)90210-9.pdf?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2F0896627389902109%3Fshowall%3Dtrue (accessed May 24, 2022).

[40] “Differential Regulation of Dynein and Kinesin Motor Proteins by Tau.” https://www.science.org/doi/10.1126/science.1152993 (accessed May 24, 2022).

[41] “A Spatial Gradient of Tau Protein Phosphorylation in Nascent Axons | Journal of Neuroscience.” https://www.jneurosci.org/content/16/18/5727 (accessed May 24, 2022).

[42] “Tau phosphorylation: the therapeutic challenge for neurodegenerative disease: Trends in Molecular Medicine.” https://www.cell.com/trends/molecular-medicine/fulltext/S1471-4914(09)00033-1?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS1471491409000331%3Fshowall%3Dtrue (accessed May 24, 2022).

[43] S.-H. Chung, “Aberrant phosphorylation in the pathogenesis of Alzheimer’s disease,” *BMB Rep.*, vol. 42, no. 8, pp. 467–474, 2009, doi: 10.5483/BMBRep.2009.42.8.467.

[44] “Alzheimer’s disease-type neuronal tau hyperphosphorylation induced by Aβ oligomers - ScienceDirect.” https://www.sciencedirect.com/science/article/abs/pii/S019745800700108X?via%3Dihub (accessed May 24, 2022).

[45] “Aβ Oligomers Cause Localized Ca2+ Elevation, Missorting of Endogenous Tau into Dendrites, Tau Phosphorylation, and Destruction of Microtubules and Spines | Journal of Neuroscience.” https://www.jneurosci.org/content/30/36/11938 (accessed May 24, 2022).

[46] C. S. Arnold, G. V. Johnson, R. N. Cole, D. L. Dong, M. Lee, and G. W. Hart, “The microtubule-associated protein tau is extensively modified with O-linked N-acetylglucosamine,” *J. Biol. Chem.*, vol. 271, no. 46, pp. 28741–28744, Nov. 1996, doi: 10.1074/jbc.271.46.28741.

[47] X. Li, F. Lu, J.-Z. Wang, and C.-X. Gong, “Concurrent alterations of O-GlcNAcylation and phosphorylation of tau in mouse brains during fasting,” *Eur. J. Neurosci.*, vol. 23, no. 8, pp. 2078–2086, Apr. 2006, doi: 10.1111/j.1460-9568.2006.04735.x.

[48] “Overexpression of Dyrk1A contributes to neurofibrillary degeneration in Down syndrome - Liu - 2008 - The FASEB Journal - Wiley Online Library.” https://faseb.onlinelibrary.wiley.com/doi/full/10.1096/fj.07-104539 (accessed May 24, 2022).

[49] C.-X. Gong, F. Liu, I. Grundke-Iqbal, and K. Iqbal, “Impaired brain glucose metabolism leads to Alzheimer neurofibrillary degeneration through a decrease in tau O-GlcNAcylation,” *J. Alzheimers Dis. JAD*, vol. 9, no. 1, pp. 1–12, Mar. 2006, doi: 10.3233/jad-2006-9101.

[50] C.-X. Gong and K. Iqbal, “Hyperphosphorylation of Microtubule-Associated Protein Tau: A Promising Therapeutic Target for Alzheimer Disease,” *Curr. Med. Chem.*, vol. 15, no. 23, pp. 2321–2328, 2008.

[51] R. E. Tanzi, “The Genetics of Alzheimer Disease,” *Cold Spring Harb. Perspect. Med.*, vol. 2, no. 10, p. a006296, Oct. 2012, doi: 10.1101/cshperspect.a006296.

[52] D. Goldgaber, M. I. Lerman, O. W. McBride, U. Saffiotti, and D. C. Gajdusek, “Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer’s disease,” *Science*, vol. 235, no. 4791, pp. 877–880, Feb. 1987, doi: 10.1126/science.3810169.

[53] J. Kang *et al.*, “The precursor of Alzheimer’s disease amyloid A4 protein resembles a cell-surface receptor,” *Nature*, vol. 325, no. 6106, pp. 733–736, Feb. 1987, doi: 10.1038/325733a0.

[54] “Roles of amyloid precursor protein and its fragments in regulating neural activity, plasticity and memory - PubMed.” https://pubmed.ncbi.nlm.nih.gov/12927332/ (accessed Jun. 07, 2022).

[55] L. M. Bekris, C.-E. Yu, T. D. Bird, and D. W. Tsuang, “Genetics of Alzheimer Disease,” *J. Geriatr. Psychiatry Neurol.*, vol. 23, no. 4, pp. 213–227, Dec. 2010, doi: 10.1177/0891988710383571.

[56] C. Van Cauwenberghe, C. Van Broeckhoven, and K. Sleegers, “The genetic landscape of Alzheimer disease: clinical implications and perspectives,” *Genet. Med.*, vol. 18, no. 5, Art. no. 5, May 2016, doi: 10.1038/gim.2015.117.

[57] B. D *et al.*, "Assembly and aggregation properties of synthetic Alzheimer's A4/beta amyloid peptide analogs," *J. Biol. Chem.*, vol. 267, no. 1, Jan. 1992, Accessed: Jun. 07, 2022. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/1730616/

[58] "Genomic organization of the human amyloid beta-protein precursor gene - PubMed." https://pubmed.ncbi.nlm.nih.gov/2110105/ (accessed Jun. 09, 2022).

[59] J. TCW and A. M. Goate, "Genetics of β-Amyloid Precursor Protein in Alzheimer's Disease," *Cold Spring Harb. Perspect. Med.*, vol. 7, no. 6, p. a024539, Jun. 2017, doi: 10.1101/cshperspect.a024539.

[60] "A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid - PubMed." https://pubmed.ncbi.nlm.nih.gov/1302033/ (accessed Jun. 09, 2022).

[61] "Peptides homologous to the amyloid protein of Alzheimer's disease containing a glutamine for glutamic acid substitution have accelerated amyloid fibril formation - PubMed." https://pubmed.ncbi.nlm.nih.gov/1681804/ (accessed Jun. 09, 2022).

[62] "Amyloid, the presenilins and Alzheimer's disease - PubMed." https://pubmed.ncbi.nlm.nih.gov/9106355/ (accessed Jun. 09, 2022).

[63] "Presenilins: members of the gamma-secretase quartets, but part-time soloists too - PubMed." https://pubmed.ncbi.nlm.nih.gov/18697993/ (accessed Jun. 09, 2022).

[64] "PSEN1 gene: MedlinePlus Genetics." https://medlineplus.gov/genetics/gene/psen1/ (accessed Jun. 09, 2022).

[65] R. N. Martins *et al.*, "High levels of amyloid-β Protein from S182 (Glu246) familial alzheimer's cells," *NeuroReport*, vol. 7, no. 1, pp. 217–220, Dec. 1995.

[66] M. Citron *et al.*, "Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid beta-protein in both transfected cells and transgenic mice," *Nat. Med.*, vol. 3, no. 1, pp. 67–72, Jan. 1997, doi: 10.1038/nm0197-67.

[67] T. Moehlmann *et al.*, "Presenilin-1 mutations of leucine 166 equally affect the generation of the Notch and APP intracellular domains independent of their effect on Abeta 42 production," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 12, pp. 8025–8030, Jun. 2002, doi: 10.1073/pnas.112686799.

[68] R. Cacace, K. Sleegers, and C. Van Broeckhoven, "Molecular genetics of early-onset Alzheimer's disease revisited," *Alzheimers Dement. J. Alzheimers Assoc.*, vol. 12, no. 6, pp. 733–748, Jun. 2016, doi: 10.1016/j.jalz.2016.01.012.

[69] "PSEN2 presenilin 2 [Homo sapiens (human)] - Gene - NCBI." https://www.ncbi.nlm.nih.gov/gene/5664 (accessed Jun. 09, 2022).

[70] Y. Cai, S. S. A. An, and S. Kim, "Mutations in presenilin 2 and its implications in Alzheimer's disease and other dementia-associated disorders," *Clin. Interv. Aging*, vol. 10, pp. 1163–1172, Jan. 2015, doi: 10.2147/cia.s85808.

[71] "Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease - PubMed." https://pubmed.ncbi.nlm.nih.gov/8705854/ (accessed Jun. 09, 2022).

[72] "A familial Alzheimer's disease locus on chromosome 1 - PubMed." https://pubmed.ncbi.nlm.nih.gov/7638621/ (accessed Jun. 09, 2022).

[73] S. Jayadev *et al.*, "Alzheimer's disease phenotypes and genotypes associated with mutations in presenilin 2," *Brain J. Neurol.*, vol. 133, pp. 1143–54, Apr. 2010, doi: 10.1093/brain/awq033.

[74] C.-C. Liu, T. Kanekiyo, H. Xu, and G. Bu, "Apolipoprotein E and Alzheimer disease: risk, mechanisms, and therapy," *Nat. Rev. Neurol.*, vol. 9, no. 2, pp. 106–118, Feb. 2013, doi: 10.1038/nrneurol.2012.263.

[75] Y. Huang and R. W. Mahley, "Apolipoprotein E: Structure and Function in Lipid Metabolism, Neurobiology, and Alzheimer's Diseases," *Neurobiol. Dis.*, vol. 72PA, pp. 3–12, Dec. 2014, doi: 10.1016/j.nbd.2014.08.025.

[76] "Structural differences between apoE3 and apoE4 may be useful in developing therapeutic agents for Alzheimer's disease | PNAS." https://www.pnas.org/doi/10.1073/pnas.1207022109 (accessed Jun. 09, 2022).

[77] R. W. Mahley and S. C. Rall, "Apolipoprotein E: far more than a lipid transport protein," *Annu. Rev. Genomics Hum. Genet.*, vol. 1, pp. 507–537, 2000, doi: 10.1146/annurev.genom.1.1.507.

[78] G. Bu, "Apolipoprotein E and its receptors in Alzheimer's disease: pathways, pathogenesis and therapy," *Nat. Rev. Neurosci.*, vol. 10, no. 5, pp. 333–344, May 2009, doi: 10.1038/nrn2620.

[79] Y. Huang and R. W. Mahley, "Apolipoprotein E: Structure and Function in Lipid Metabolism, Neurobiology, and Alzheimer's Diseases," *Neurobiol. Dis.*, vol. 72PA, pp. 3–12, Dec. 2014, doi: 10.1016/j.nbd.2014.08.025.

[80] W. J. Brecht *et al.*, "Neuron-specific apolipoprotein e4 proteolysis is associated with increased tau phosphorylation in brains of transgenic mice," *J. Neurosci. Off. J. Soc. Neurosci.*, vol. 24, no. 10, pp. 2527–2534, Mar. 2004, doi: 10.1523/JNEUROSCI.4315-03.2004.

[81] D. H. Mauch *et al.*, "CNS synaptogenesis promoted by glia-derived cholesterol," *Science*, vol. 294, no. 5545, pp. 1354–1357, Nov. 2001, doi: 10.1126/science.294.5545.1354.

[82] Y. Ji, Y. Gong, W. Gan, T. Beach, D. M. Holtzman, and T. Wisniewski, "Apolipoprotein E isoform-specific regulation of dendritic spine morphology in apolipoprotein E transgenic mice and Alzheimer's

disease patients," *Neuroscience*, vol. 122, no. 2, pp. 305–315, 2003, doi: 10.1016/j.neuroscience.2003.08.007.

[83] "Large-Scale Identification, Mapping, and Genotyping of Single-Nucleotide Polymorphisms in the Human Genome." https://www.science.org/doi/10.1126/science.280.5366.1077 (accessed May 24, 2022).

[84] S. Srinivasan, J. A. Clements, and J. Batra, "Single nucleotide polymorphisms in clinics: Fantasy or reality for cancer?," *Crit. Rev. Clin. Lab. Sci.*, vol. 53, no. 1, pp. 29–39, Jan. 2016, doi: 10.3109/10408363.2015.1075469.

[85] W. Sukhumsirichart, *Polymorphisms*. IntechOpen, 2018. doi: 10.5772/intechopen.76728.

[86] F. Robert and J. Pelletier, "Exploring the Impact of Single-Nucleotide Polymorphisms on Translation," *Front. Genet.*, vol. 9, p. 507, Oct. 2018, doi: 10.3389/fgene.2018.00507.

[87] N. D. Rawlings, D. P. Tolle, and A. J. Barrett, "MEROPS: the peptidase database," *Nucleic Acids Res.*, vol. 32, no. suppl_1, pp. D160–D164, Jan. 2004, doi: 10.1093/nar/gkh071.

[88] "Phylogeny of the Serpin Superfamily: Implications of Patterns of Amino Acid Conservation for Structure and Function." https://genome.cshlp.org/content/10/12/1845 (accessed May 26, 2022).

[89] "SERPINI1 serpin family I member 1 [Homo sapiens (human)] - Gene - NCBI." https://www.ncbi.nlm.nih.gov/gene/5274 (accessed May 27, 2022).

[90] "SERPINI1 - Neuroserpin precursor - Homo sapiens (Human) - SERPINI1 gene & protein." https://www.uniprot.org/uniprot/Q99574 (accessed May 27, 2022).

[91] "The N terminus of the serpin, tengpin, functions to trap the metastable native state | EMBO reports." https://www.embopress.org/doi/full/10.1038/sj.embor.7400986 (accessed May 26, 2022).

[92] L. Yang, J. A. Irving, W. Dai, M.-I. Aguilar, and S. P. Bottomley, "Probing the folding pathway of a consensus serpin using single tryptophan mutants," *Sci. Rep.*, vol. 8, no. 1, Art. no. 1, Feb. 2018, doi: 10.1038/s41598-018-19567-9.

[93] "A hydrophobic patch surrounding Trp154 in human neuroserpin controls the helix F dynamics with implications in inhibition and aggregation | Scientific Reports." https://www.nature.com/articles/srep42987 (accessed May 26, 2022).

[94] "Crystal structure of neuroserpin: a neuronal serpin involved in a conformational disease - Briand - 2001 - FEBS Letters - Wiley Online Library." https://febs.onlinelibrary.wiley.com/doi/full/10.1016/S0014-5793%2801%2902764-8 (accessed May 26, 2022).

[95] "Structures of Active and Latent PAI-1: A Possible Stabilizing Role for Chloride Ions | Biochemistry." https://pubs.acs.org/doi/10.1021/bi000290w (accessed May 26, 2022).

[96] S. Takehara *et al.*, “The 2.1-Å Crystal Structure of Native Neuroserpin Reveals Unique Structural Elements That Contribute to Conformational Instability,” *J. Mol. Biol.*, vol. 388, no. 1, pp. 11–20, Apr. 2009, doi: 10.1016/j.jmb.2009.03.007.

[97] “Expression of Neuroserpin, an Inhibitor of Tissue Plasminogen Activator, in the Developing and Adult Nervous System of the Mouse - PMC.” https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6573583/ (accessed May 27, 2022).

[98] “Neuroserpin, a brain-associated inhibitor of tissue plasminogen activator is localized primarily in neurons. Implications for the regulation of motor learning and neuronal survival - PubMed.” https://pubmed.ncbi.nlm.nih.gov/9407089/ (accessed May 27, 2022).

[99] “The Axonally Secreted Serine Proteinase Inhibitor, Neuroserpin, Inhibits Plasminogen Activators and Plasmin but Not Thrombin* - Journal of Biological Chemistry.” https://www.jbc.org/article/S0021-9258(19)84189-3/fulltext (accessed May 27, 2022).

[100] “Tissue plasminogen activator (tPA) increase neuronal damage after focal cerebral ischemia in wild-type and tPA-deficient mice | Nature Medicine.” https://www.nature.com/articles/nm0298-228 (accessed May 27, 2022).

[101] M. Yepes *et al.*, “Regulation of seizure spreading by neuroserpin and tissue-type plasminogen activator is plasminogen-independent,” *J. Clin. Invest.*, vol. 109, no. 12, pp. 1571–1578, Jun. 2002, doi: 10.1172/JCI14308.

[102] L. Fredriksson *et al.*, “Identification of a neurovascular signaling pathway regulating seizures in mice,” *Ann. Clin. Transl. Neurol.*, vol. 2, no. 7, pp. 722–738, 2015, doi: 10.1002/acn3.209.

[103] X. Wang *et al.*, “Lipoprotein receptor–mediated induction of matrix metalloproteinase by tissue plasminogen activator,” *Nat. Med.*, vol. 9, no. 10, Art. no. 10, Oct. 2003, doi: 10.1038/nm926.

[104] T. Sumii and E. H. Lo, “Involvement of Matrix Metalloproteinase in Thrombolysis-Associated Hemorrhagic Transformation After Embolic Focal Ischemia in Rats,” *Stroke*, vol. 33, no. 3, pp. 831–836, Mar. 2002, doi: 10.1161/hs0302.104542.

[105] Z.-L. Chen and S. Strickland, “Neuronal Death in the Hippocampus Is Promoted by Plasmin-Catalyzed Degradation of Laminin,” *Cell*, vol. 91, no. 7, pp. 917–925, Dec. 1997, doi: 10.1016/S0092-8674(00)80483-3.

[106] “Immunotherapy blocking the tissue plasminogen activator-dependent activation of N-methyl-d-aspartate glutamate receptors improves hemorrhagic stroke outcome - ScienceDirect.” https://www.sciencedirect.com/science/article/abs/pii/S0028390812005709?via%3Dihub (accessed May 27, 2022).

[107] K. Benchenane *et al.*, "Anti-NR1 N-terminal-domain vaccination unmasks the crucial action of tPA on NMDA-receptor-mediated toxicity and spatial memory," *J. Cell Sci.*, vol. 120, no. 4, pp. 578–585, Feb. 2007, doi: 10.1242/jcs.03354.

[108] S. Fabbro and N. W. Seeds, "Plasminogen activator activity is inhibited while neuroserpin is up-regulated in the Alzheimer disease brain," *J. Neurochem.*, vol. 109, no. 2, pp. 303–315, 2009, doi: 10.1111/j.1471-4159.2009.05894.x.

[109] "Neuroserpin, a crucial regulator for axogenesis, synaptic modelling and cell–cell interactions in the pathophysiology of neurological disease | SpringerLink." https://link.springer.com/article/10.1007/s00018-022-04185-6 (accessed Jun. 13, 2022).

[110] J. Lykkesfeldt, A. J. Michels, and B. Frei, "Vitamin C1," *Adv. Nutr.*, vol. 5, no. 1, pp. 16–18, Jan. 2014, doi: 10.3945/an.113.005157.

[111] E. J. J. Bruno, T. N. P. Ziegenfuss, and J. Landis, "Vitamin C: Research Update," *Curr. Sports Med. Rep.*, vol. 5, no. 4, pp. 177–181, Aug. 2006, doi: 10.1097/01.CSMR.0000306503.32987.1e.

[112] S. Chambial, S. Dwivedi, K. K. Shukla, P. J. John, and P. Sharma, "Vitamin C in Disease Prevention and Cure: An Overview," *Indian J. Clin. Biochem.*, vol. 28, no. 4, pp. 314–328, Oct. 2013, doi: 10.1007/s12291-013-0375-3.

[113] A. Covarrubias-Pinto, A. I. Acuña, F. A. Beltrán, L. Torres-Díaz, and M. A. Castro, "Old Things New View: Ascorbic Acid Protects the Brain in Neurodegenerative Disorders," *Int. J. Mol. Sci.*, vol. 16, no. 12, pp. 28194–28217, Nov. 2015, doi: 10.3390/ijms161226095.

[114] "Mechanisms involved in the modulation of astroglial resistance to oxidative stress induced by activated microglia: antioxidative systems, peroxide elimination, radical generation, lipid peroxidation - PubMed." https://pubmed.ncbi.nlm.nih.gov/19763738/ (accessed Jun. 16, 2022).

[115] J. Kocot, D. Luchowska-Kocot, M. Kiełczykowska, I. Musik, and J. Kurzepa, "Does Vitamin C Influence Neurodegenerative Diseases and Psychiatric Disorders?," *Nutrients*, vol. 9, no. 7, p. 659, Jun. 2017, doi: 10.3390/nu9070659.

[116] "High-dose of vitamin C supplementation reduces amyloid plaque burden and ameliorates pathological changes in the brain of 5XFAD mice - PubMed." https://pubmed.ncbi.nlm.nih.gov/24577081/ (accessed Jun. 16, 2022).

[117] S. T. Sherry *et al.*, "dbSNP: the NCBI database of genetic variation," *Nucleic Acids Res.*, vol. 29, no. 1, pp. 308–311, Jan. 2001, doi: 10.1093/nar/29.1.308.

[118] "SERPINI1 - SNP - NCBI." https://www.ncbi.nlm.nih.gov/snp (accessed May 27, 2022).

[119] Y. Choi, G. E. Sims, S. Murphy, J. R. Miller, and A. P. Chan, "Predicting the functional effect of amino acid substitutions and indels," *PloS One*, vol. 7, no. 10, p. e46688, 2012, doi: 10.1371/journal.pone.0046688.

[120] L. Sandell and N. P. Sharp, "Fitness Effects of Mutations: An Assessment of PROVEAN Predictions Using Mutation Accumulation Data," *Genome Biol. Evol.*, vol. 14, no. 1, p. evac004, Jan. 2022, doi: 10.1093/gbe/evac004.

[121] H. Tang and P. D. Thomas, "PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation," *Bioinforma. Oxf. Engl.*, vol. 32, no. 14, pp. 2230–2232, Jul. 2016, doi: 10.1093/bioinformatics/btw222.

[122] "Structural and functional analysis of disease-associated mutations in GOT1 gene: An in silico study - PubMed." https://pubmed.ncbi.nlm.nih.gov/34352456/ (accessed May 27, 2022).

[123] L. R. Brunham, R. R. Singaraja, T. D. Pape, A. Kejariwal, P. D. Thomas, and M. R. Hayden, "Accurate Prediction of the Functional Significance of Single Nucleotide Polymorphisms and Mutations in the ABCA1 Gene," *PLoS Genet.*, vol. 1, no. 6, p. e83, Dec. 2005, doi: 10.1371/journal.pgen.0010083.

[124] P. D. Thomas *et al.*, "Applications for protein sequence–function evolution data: mRNA/protein expression analysis and coding SNP scoring tools," *Nucleic Acids Res.*, vol. 34, no. suppl_2, pp. W645–W650, Jul. 2006, doi: 10.1093/nar/gkl229.

[125] I. Adzhubei, D. M. Jordan, and S. R. Sunyaev, "Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2," *Curr. Protoc. Hum. Genet. Editor. Board Jonathan Haines Al*, vol. 0 7, p. Unit7.20, Jan. 2013, doi: 10.1002/0471142905.hg0720s76.

[126] "Meta-SNP - Meta-predictor of disease causing variants." https://snps.biofold.org/meta-snp/pages/help.html (accessed May 27, 2022).

[127] "PredictSNP: Robust and Accurate Consensus Classifier for Prediction of Disease-Related Mutations | PLOS Computational Biology." https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003440 (accessed May 27, 2022).

[128] B. Yoganarasimha, V. Chandramohan, T. P. Krishna Murthy, B. S. Gangadharappa, G. B. Siddaiah, and M. Hanumanthappa, "Prediction of deleterious single nucleotide polymorphisms and their effect on the sequence and structure of the human CCND1 gene," *J. Taibah Univ. Med. Sci.*, vol. 12, no. 3, pp. 221–228, Sep. 2016, doi: 10.1016/j.jtumed.2016.07.009.

[129] C. M. Yates, I. Filippis, L. A. Kelley, and M. J. E. Sternberg, "SuSPect: Enhanced Prediction of Single Amino Acid Variant (SAV) Phenotype Using Network Features," *J. Mol. Biol.*, vol. 426, no. 14, pp. 2692–2701, Jul. 2014, doi: 10.1016/j.jmb.2014.04.026.

[130] C. Ferrer-Costa, J. L. Gelpí, L. Zamakola, I. Parraga, X. de la Cruz, and M. Orozco, "PMUT: a web-based tool for the annotation of pathological mutations on proteins," *Bioinformatics*, vol. 21, no. 14, pp. 3176–3178, Jul. 2005, doi: 10.1093/bioinformatics/bti486.

[131] "Snap2 - Rost Lab Open." https://rostlab.org/owiki/index.php/Snap2 (accessed May 28, 2022).

[132] E. Capriotti, P. Fariselli, and R. Casadio, "I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure," *Nucleic Acids Res.*, vol. 33, no. Web Server issue, pp. W306–W310, Jul. 2005, doi: 10.1093/nar/gki375.

[133] D. E. V. Pires, D. B. Ascher, and T. L. Blundell, "DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach," *Nucleic Acids Res.*, vol. 42, no. Web Server issue, pp. W314–W319, Jul. 2014, doi: 10.1093/nar/gku411.

[134] V. Parthiban, M. M. Gromiha, and D. Schomburg, "CUPSAT: prediction of protein stability upon point mutations," *Nucleic Acids Res.*, vol. 34, no. suppl_2, pp. W239–W242, Jul. 2006, doi: 10.1093/nar/gkl190.

[135] C. H. Rodrigues, D. E. Pires, and D. B. Ascher, "DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability," *Nucleic Acids Res.*, vol. 46, no. W1, pp. W350–W355, Jul. 2018, doi: 10.1093/nar/gky300.

[136] "Align-GVGD | NGRL Manchester." http://www.ngrl.org.uk/Manchester/page/align-gvgd.html (accessed May 28, 2022).

[137] H. Ashkenazy *et al.*, "ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules," *Nucleic Acids Res.*, vol. 44, no. W1, pp. W344–W350, Jul. 2016, doi: 10.1093/nar/gkw408.

[138] "The ConSurf Server." https://consurf.tau.ac.il/overview.php#_Toc311131288 (accessed May 29, 2022).

[139] D. Szklarczyk *et al.*, "STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic Acids Res.*, vol. 47, no. Database issue, pp. D607–D613, Jan. 2019, doi: 10.1093/nar/gky1131.

[140] "Screening of natural compounds from Cyperus rotundus Linn against SARS-CoV-2 main protease (Mpro): An integrated computational approach. - Abstract - Europe PMC." https://europepmc.org/article/med/34090015 (accessed Jun. 14, 2022).

[141] S. Forli, R. Huey, M. E. Pique, M. Sanner, D. S. Goodsell, and A. J. Olson, "Computational protein-ligand docking and virtual drug screening with the AutoDock suite," *Nat. Protoc.*, vol. 11, no. 5, pp. 905–919, May 2016, doi: 10.1038/nprot.2016.051.

[142] "Identification of potential inhibitors of Fasciola gigantica thioredoxin1: computational screening, molecular dynamics simulation, and binding free energy studies: Journal of Biomolecular Structure and Dynamics: Vol 36, No 8." https://www.tandfonline.com/doi/abs/10.1080/07391102.2017.1344141?journalCode=tbsd20 (accessed Jun. 16, 2022).

[143] Z. Wang, C. Huang, H. Lv, M. Zhang, and X. Li, "In silico analysis and high-risk pathogenic phenotype predictions of non-synonymous single nucleotide polymorphisms in human Crystallin beta A4 gene associated with congenital cataract," *PLoS ONE*, vol. 15, no. 1, p. e0227859, Jan. 2020, doi: 10.1371/journal.pone.0227859.

[144] "overview [PolyPhen-2 Wiki]." http://genetics.bwh.harvard.edu/pph2/dokuwiki/overview (accessed May 29, 2022).

[145] E. Capriotti, R. B. Altman, and Y. Bromberg, "Collective judgment predicts disease-associated single nucleotide variants," *BMC Genomics*, vol. 14, no. Suppl 3, p. S2, May 2013, doi: 10.1186/1471-2164-14-S3-S2.

[146] E. Capriotti, R. Calabrese, P. Fariselli, P. L. Martelli, R. B. Altman, and R. Casadio, "WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation," *BMC Genomics*, vol. 14, no. Suppl 3, p. S6, May 2013, doi: 10.1186/1471-2164-14-S3-S6.

[147] Z. Mahmud, S. U. F. Malik, J. Ahmed, and A. K. Azad, "Computational Analysis of Damaging Single-Nucleotide Polymorphisms and Their Structural and Functional Impact on the Insulin Receptor," *BioMed Res. Int.*, vol. 2016, 2016, doi: 10.1155/2016/2023803.