

**“Screening and designing of KIF5A like motor proteins in
Amyotrophic Lateral Sclerosis (ALS)”**

A DISSERTATION

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

**Masters of Technology
Biomedical Engineering**

Submitted by

ANKITA ARORA

2K17/BME/01



Under the supervision of

Prof. Pravir Kumar

(Professor)

Department of Biotechnology

Delhi Technological University, New Delhi-110042

(Formerly Delhi College of Engineering, University of Delhi)

CANDIDATE'S DECLARATION

The work entitled “**Screening and designing of KIFA5A like motor proteins in Amyotrophic Lateral Sclerosis (ALS)**” was carried out by me in Molecular Neuroscience and Functional Genomics Laboratory under the supervision and guidance of **Prof. Pravir Kumar**, Department of Biotechnology, Delhi Technological University. The extent and sources of information derived from existing literature have been indicated throughout the report at appropriate place. The work is original and has not been submitted in part or full for any other diploma or degree of this organization or any other university.

PLACE:

DATE:

Ankita Arora
(2K17/BME/01)
Department of Biotechnology,
Delhi Technological University

CERTIFICATE

It is to certify that the research work carried out and data presented in this project entitled **“Screening and designing of KIFA5A like motor proteins in Amyotrophic Lateral Sclerosis (ALS)”** submitted for the partial fulfilment of Master of Technology Biomedical Engineering was carried out by **Ms. Ankita Arora** under the guidance and supervision of **Prof. Pravir Kumar, Department of Biotechnology, Delhi Technological University, New Delhi-110042.**

Prof. Pravir Kumar

Supervisor

Delhi Technological University

Prof. Jai Gopal Sharma

Head of Department

Delhi Technological University

ACKNOWLEDGEMENT

I wish to record my deep sense of gratitude and sincere thanks to my research guide **Prof. Pravir Kumar, Department of Biotechnology, Delhi Technological University** for his unstinted inspiration, invaluable guidance, encouragement, keen interest, good wishes and valuable suggestions throughout my entire research tenure.

I express my kind regards and gratitude to **Prof. Jai Gopal Sharma, Head of Department, Department of Biotechnology, Delhi Technological University.**

I would also like to thank **Dr. Rashmi Ambasta** for her support and guidance at various steps during the course of the project that helped in shaping the project and resolving various critical issues faced during the course of project.

I would also like to thanks **Mr. Chhail Bihari and Mr. Jitendre Singh** for providing necessary stuffs and maintaining laboratory in good condition.

I am thankful to my friend for giving me moral boost and making my hopes alive with energy and enthusiasm to carry out the present work and helping hand at every step. My sincere thanks to my parents and family member for encouragement and moral support I received during the tenure of study.

ABSTRACT

KIF5A a motor neuron protein expressed in neuron is responsible for anterograde transportation of organelles, proteins and RNA. Variation within KIF5A leading to disruption of axonal transport serve as a hallmark for various neurodegenerative diseases such as hereditary spastic paraplegia (HSP10), Charcot-Marie-Tooth disease type 2 (CMT2), amyotrophic lateral sclerosis (ALS). Amyotrophic Lateral Sclerosis (ALS) is one of the incurable motor neuron disorders in which progressive loss of upper and lower motor neuron occur, with the incidence of 1-5 per 100,000. Studies have shown KIF5A is a novel ALS gene, an association of rare KIF5A variant with was predominantly due to a mutation in splice site region which result in loss of function of KIF5A protein involved in vesicular transport in mitochondria, Golgi-ER region. Non-synonymous single nucleotide polymorphism (nsSNPs) has potential to alter structure and function of protein thus it is important to differentiate potential damaging and deleterious nsSNPs from neutral. The aim of our study is analyse the functional effect of non-synonymous single nucleotide polymorphism (nsSNPs) leading to dysfunction of KIF5A protein in axonal transport using bioinformatics tools. *In-silico* screening of 512 missense SNPs associated with KIF5A predicted 109 nsSNPs to be damaging in nature. Subsequent analysis of these nsSNPs predicted 5 nsSNPs (A268T, R369W, T644M, R712L and P986L) to be highly deleterious among the entire prediction program. The complete KIF5A protein structure was modeled using *ab-initio* modeling. The study highlighted three possible nsSNPs (T644M, R712L and P986L) to increased stability of mutant protein, thus altering the function of protein. Exact biological mechanism associated with above predicted nsSNPs still needs to validate by *in-vitro* studies. Further we designed novel synthetic compounds to inhibit Pro986Leu variant of KIF5A. A compound library was prepared that consisted of natural compounds retrieved from the ZINC database. The prepared library was then screened against this missense variant KIF5A at specific domains which is involved in ALS and then docking was done. This was completely a new approach to target ALS. The results obtained from this study need to be experimentally validated further so that we can prove our computational work and keep working in that direction with the assurance that our approach is right. The study provided a path to explore association of these predicted nsSNPs in disease susceptibility and to design target dependent drugs for therapeutic application.

TABLE OF CONTENTS

CANDIDATE'S DECLARATION	ii
CERTIFICATE	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	viii
LIST OF TABLES	ix
LIST OF ABBREVIATIONS	x
1. INTRODUCTION	1
2. REVIEW OF LITERATURE	3
2.1 Neurodegenerative disease: An Overview	3
2.2 Amyotrophic Lateral Sclerosis	4
2.3 Kinesin family member 5A (KIF5A)	5
3. METHODOLOGY	8
3.1 Data mining for SNPs	8
3.2 <i>In-silico</i> prediction of deleterious and damaging nsSNP in KIF5A	9
3.3 3D Model Construction	9
3.4 Structure Validation	10
3.5 Stability changes prediction of mutant protein	10
3.6 Active Site Prediction	10
3.7 Virtual Screening	11

3.8 ADME Properties	11
3.9 Molecular Docking	11
4. RESULTS	13
4.1 SNPs id retrieval	13
4.2 Computational analysis of nsSNP based on sequence based prediction tools	13
4.3 Three dimensional structure prediction and validation	15
4.4 Prediction of stability changes of mutant proteins by FOLDX	16
4.5 Active Site Prediction	20
4.6 Virtual Screening	20
4.7 Molecular Docking	21
5. DISCUSSION	27
6. CONCLUSION	30
7. APPENDIX	31
8. REFERENCES	43

LIST OF FIGURES

S.No	Name of Figure	Page No.
Figure 1	Schematic representation of workflow	8
Figure 2	A statistical representation of deleterious/damaging nsSNP predicted by <i>in-silico</i> tools	13
Figure 3	Heat map of KIF5A generated by SNAP	14
Figure 4	3D structure of KIF5A modeled by <i>ab-initio</i> method	15
Figure 5	Ramachandran plot of modeled KIF5A protein	16
Figure 6	Structure of mutants protein generated by WHATIF server	17
Figure 7	The Ramachandran plot of native and mutant protein generated by RAMPAGE	18
Figure 8	Active site residue of KIF5A	20
Figure 9	Natural compounds from different ZINC database	21

LIST OF TABLES

S.No	Table Name	Page No.
Table 1	Functionally important substitution predicted by <i>in-silico</i> screening	14
Table 2	Stability analysis of functionally important nsSNPs	19
Table 3	Chemical name, IUPAC name and fitness score of top ligands docked with KIF5A using GOLD	21
Table 4	ADME and pharmacological parameters prediction of active compounds using QikProp	24

LIST OF ABBREVIATIONS

ADME	Absorption Distribution Metabolism and Elimination
ALS	Amyotrophic Lateral Sclerosis
CMT2	Charcot-Marie-Tooth disease type 2
DDG	Free Energy Change
fALS	Familial Amyotrophic Lateral Sclerosis
GOLD	Genetic Optimization for Ligand docking
HMM	Hidden Markov Model
HOPE	Have yOur Protein Explained
HSP	Hereditary Spastic Paraplegia
HTVS	High Throughput Virtual Screening
I-TASSER	Iterative Threading ASSEmby Refinement
KIF5A	Kinesin Family Member 5A
MND	Motor Neuron Disorder
nsSNP	Non-synonymous Single Nucleotide Polymorphism
PANTHER	Protein Analysis Through Evolutionary Relationships
PDB	Protein Data Bank
PhD-SNP	Predictor of human Deleterious Single Nucleotide Polymorphisms
sALS	Sporadic Amyotrophic Lateral Sclerosis
SIFT	Sorting Tolerant From Intolerant
SNAP	Screening of Non-Acceptable Polymorphism
SNP	Single Nucleotide Polymorphism
SOD1	Superoxide Dismutase 1
SP	Standard Precision
SVM	Support Vector Machines
XP	Extra Precision

1. INTRODUCTION

Neurodegenerative disease is genetic condition that is depicted as escalating neuronal cell death, which is marked by loss of neurons within cerebrum region of brain or spinal cord [1]. Kinesin Family Member 5A (KIF5A) protein is expressed in neurons, functioning as microtubule motor, a component of multi-subunit complex in mitochondrial region in transportation of intracellular proteins and organelles. KIF5A consist of a motor domain at N-terminal motor domain, an alpha helical coiled stalk region and a cargo binding domain at C-terminal and is involved in axonal transport [2,3]. Missense mutation in N-terminal domain leads to monogenic spastic paraplegia and Charcot-Marie-Tooth disease type 2 [4,5]. Additionally, a C-terminal frame-shift mutation is related with a neurodevelopment disorder such as neonatal intractable myoclonus [6]. Kinesin encodes the neuronal kinesin heavy chain (KHC) implicated in the anterograde axonal transport [7]. Recent discovery has revealed KIF5A as a novel gene associated with Amyotrophic lateral sclerosis (ALS) using a large-scale genome-wide association study and exome sequencing [8]. Amyotrophic lateral sclerosis (ALS) is foremost neurodegenerative diseases that arise in two forms of ALS such as familial and sporadic, is major persistent motor neuron diseases, with an occurrence of 1–5 for every 100,000 [9]. Because of the absence of compelling treatment, ALS prompts to death between 2 and 5 years after diagnosis, majorly because of failure of respiratory tracts. However major cases are sporadic (sALS) [10], but only 5–10% are associated with a genetic mutation which is inherited through family [11]. From the year 1993, more than 36 genes mutations have been related with focalization of ALS, and changes in few of these have anticipated disrupting functions of cytoskeletal and intracellular transport system [12,13]. GWAS study has identified missense variant p.Pro986Leu has a significant association with ALS risk and central role of kinesin in axonal transport leads to outcome that mutation in KIF5A would disrupt this process [14]. Disruption in axonal transport and alteration in the cytoskeletal are major hallmarks in ALS patients and majorly leads to degeneration of motor neuron pathogenesis [15]. Missense variant Pro986Leu associated with the C-terminal cargo-binding domain which is associated with ALS is distinct from the HSP and CMT missense mutation in N-terminal motor domain. Missense mutation within the C-terminal domain has significantly observed to affect the binding of microtubule and hydrolysis of ATP, leading to loss of anterograde transport of cargo proteins within the region of dendrites and axon [16]. This loss of function due to missense mutation in C-terminal domain in KIF5A is the major cause of ALS. Thus there is an immediate requirement for the preventive measures that can significantly repair or stabilize the cytoskeleton function loss due to mutation. This provides us an opportunity for development of novel drugs that could possibly treat both familial and sporadic ALS.

Genetic variations are important for evolution whereby the association between the variation and their phenotypic effects would be helpful in understanding the disorders. The polymorphisms may occur either in the coding and non-coding regions of a gene. The coding Single Nucleotide Polymorphisms (SNPs) are either synonymous or non-synonymous [17]. The synonymous SNPs arise due to the degeneracy of the genetic code where the amino acid sequence is not changed but non-synonymous SNPs will change the amino acid that can lead to a change in protein function. It has been estimated that as many as 93% of all human genes contain at least one SNP.

Our study focuses to identify the nsSNPs associated with KIF5A through *in-silico* prediction tools and design novel inhibitor molecules that can be used to target the missense mutation Pro986Leu within the C-terminal domain of KIF5A via *in-silico* approach. Here we have modeled the structure of protein KIF5A using *ab-initio* modeling. Virtual screening against the natural compound's library in Schrodinger suite was performed to identify the novel inhibitor molecule of the protein. Further top compounds from the screening were taken and they were simulated again for the best confirmations. The results that will be obtained from this study need to be experimentally validated so that we prove our computational work and keep working in that direction with the assurance that our approach is right. Therefore, via designing a novel inhibitor using high-throughput *in-silico* screening methods that can be used to target such missense mutation in C-terminal of KIF5A protein can provide a much more promising drug against ALS.

2. REVIEW OF LITERATURE

2.1 NEURODEGENERATIVE DISEASE: AN OVERVIEW

Degeneration of the neurons is one of the main character traits of many incapacitating, incurable neurodegenerative disease that are considerably increased in prevalence such as Alzheimer's or ALS or Parkinson's disease [18]. Due to this reason, there is a pressing need to grow novel and more successful therapeutic strategies to combat these disastrous diseases. A model system from in- vitro cell based to a unicellular organism to the highly complicated animal model system have turned out to be a valuable aspect to explore network system and underlying cause of neurodegenerative diseases, and these advances have now started to give promising restorative therapies [19].

Neurodegenerative diseases appear to a noteworthy risk to the health of humankind. In recent years these age-dependent disorders are progressively prevailing in the elderly population [18]. Instances of neurodegenerative diseases are dementia, Huntington's disease, Amyotrophic Lateral Sclerosis' disease, Parkinson's disease, ataxias, Alzheimer's disease etc. However, these ailments differ in their pathophysiology –with few causing memory damage and loss and others influencing a man's capacity to talk, move and breathe [20]. Efficacious medication and treatment are urgently needed with in-depth knowledge about mechanism and cause of each disease [15].

One approach to find out about how these ailment functions, are to build a framework of a model system that summarizes the distinctive feature of the disease. Various significant experimental model organisms such as the fruit fly, mouse, nematode worm, and even baker's yeast have been utilized for a long time to study neurodegenerative diseases and have given in-depth knowledge into the mechanism of disease [21].

Recently gained the capacity to create induced pluripotent stem cells (iPSCs) has made it conceivable to produce patient-specific cell lines in a tissue culture plate thus producing human disease models [22]. Lately, there have been developments that take into consideration cells to be cultured in three dimensions, to differentiate into organoids that further differentiate into human tissues [23]. These organoid frameworks allow cell-cell interactions and modeling of complex cytoskeletal structure and studied in highly specific detail and majorly in contexts of the physiology of tissue than isolated cells in culture plates. Moreover, proofs also suggest that these neurodegenerative diseases are not just diseases

of neuronal death but non-neuronal cells such as glial cells present mainly in brain and central nervous system plays a significant role in the progression of disease [24].

2.2 AMYOTROPHIC LATERAL SCLEROSIS

Amyotrophic lateral sclerosis (ALS) is a lethal upper and lower motor neuron disorder that is, described by a dynamic loss of motor neurons (LMNs) at the spinal or bulbar level. In 1869, French neurologist Jean-Martin Charcot first depicted ALS [25]. In 1939, the disease became notably known in the United States when baseball player Lou Gehrig was diagnosed with the disease. It is otherwise called as Charcot disease, one of the five MNDs that damage motor neurons [11].

ALS is grouped into two forms. The most widely recognized is sporadic (90–95%) which has no observable hereditarily acquired component. The remaining 5–10% of the cases are genetically acquired disease known as familial-type ALS (FALS). The first inception of any indication is normally between the ages of 50 and 65 [26]. The most widely recognized manifestation that shows up in both types of ALS is weakening of muscle, twitching, and cramping, which in long run can lead to muscle impairment [27]. In the most latter stages, symptoms such as dyspnea and dysphagia develop in ALS patients [28].

Moreover, incoming year's incidence of ALS cases over the world will increment to almost 400,000, prevalently because of the aging of the population. This increase is anticipated to place a gigantic burden on global healthcare systems, in particular, because the annual healthcare cost per patient with ALS is among the highest for any neurological disease. Approximately 10% of ALS cases display a family history (FALS), whereas the remaining 90% of ALS cases are sporadic (SALS) in nature [29].

Regardless of almost half a century when disorder was first described, despite extensive clinical research there has been no effective strategy to provide therapeutic aid till date. It primarily damages neurons in motor areas of a spinal cord, brain, and brain stem leading to escalating degeneration and atrophy of skeletal muscles lastly leading to paralysis of these voluntary muscles [30]. The main characteristic feature of ALS observed to be highly complicated, undeniably leading to an improper therapeutic treatment. With just about 5–10% cases are familial (fALS), following law of inheritance with not less than 13 specific gene sites of significant impact known to cause to familial disorder. SOD1 (superoxide dismutase 1) gene mutations are widely studied especially mutations due to functional gain, which causes approximately 20% fALS cases while only 5% sALS cases of disorder [31]. However, the

correct mechanisms which enhance development of disease by SOD1 are still unknown. Nevertheless, first transgenic SOD1^{G93A} mouse model development in the year 1993 which intently copies human fALS pathology was paradigmatic as it has enabled researchers to closely study ALS disease mechanisms in an animal model for every first time [32]. Till today, mutant SOD1 mice stay as most broadly utilized as model for ALS to investigate at a molecular and cellular level of disease progression and to examine the potential adequacy of novel therapeutic molecules [33].

Riluzole is the main as of now treatment choice for ALS and is just mildly effective. In the year 1995, a drug was approved for used as medication in USA however it got approval to be used in Europe in the year 1996. Despite the fact primary large-scale studies exhibiting the experimental viability of riluzole in enhance the survival rate in diseased patients was performed in 1990 [34]. It has shown successive comparative adequacy in this mutant mouse model as seen in human ALS trials, in this way approving preclinical study [35]. Like in human patients it is outstanding that simply that riluzole neglected to have an advantageous impact on disease inception and had just unassumingly enhanced survival in these mice. In pursued years, over 60 compounds have been examined as a conceivable therapeutics for ALS. In most cases, drug compounds that achieved the CTs stage since approval of riluzole have neglected to exhibit human efficacy [36].

2.3 KINESIN FAMILY MEMBER 5A (KIF5A)

KIF5A is an individual from kinesin relative proteins group that is essentially communicated in areas of neurons [37]. As a feature of complex multi-subunit structure, it performs as microtubule motor in the mitochondrial region, transportation of intracellular protein and organelle [7]. Missense mutations at position 12q13.3 in the KIF5A gene most successive cause of hereditary spastic paraplegia (HSP10) which are autosomal dominant in nature, influencing principally respiratory tracts and some of the times additionally the peripheral nervous system [4,38]. Furthermore, the phenotypic range of KIF5A mutations involves additionally peripheral neuropathy (Charcot-Marie-Tooth disease type 2; CMT2) [6] and a highly complicated juvenile neurological disorder with leukoencephalopathy, abnormalities in optical nerve, myoclonus, hypotonia, dysphagia, hearing loss, and early developmental arrest [39,40].

Only 5% of patients show a positive family ancestry (fALS) having a neuronal disease affecting majorly motor neuron, amyotrophic lateral sclerosis (ALS) most often as an autosomal dominant trait for Mendelian inheritance [41]. Ever since the year 1993, in more than 36 mutated genes have been related with the progression of ALS, however changes in few of these have been anticipated to damages

functioning of cytoskeletal and transportation of intracellular proteins and organelles [42]. Datasets containing two extensive examinations based on sequencing of whole genome or association across genome testing proposed likewise relationship between variants in KIF5A and ALS [14,43]. Missense variant within the domain of KIF5A has shown genome wide statistical significance which is associated with ALS risk. Studies using rare variant burden analysis applied to exome sequencing have revealed association between ALS risk and rare KIF5A loss of function variant. Further replication cohort studies has shown Pro986Leu variant and loss of function variants in KIF5A domain was observed to be associated with risk of developing ALS allele [8].

Kinesin protein is motor protein which is microtubule in nature expressing in neuron, performing transportation of intracellular organelles functioning as a cargo by binding to adaptor protein in eukaryotic cells. There occurs three isoform of heavy chain of KIF : KIF5A, KIF5B and KIF5C [3]. These protein dimerize in homo or hetero form via coiled- coiled stalk and via tail domain binding to create a complex within two kinesin light chains [38]. Mutation in KIF5A is known to disrupt axonal transport and directly leads to impairment in the motor neuron pathogenesis [44]. Kinesins proteins are involved in axonal transport that leads to hypothesize that deviation in KIF5A sequence would lead to disruption in axonal transport process. Thus is the one of the significant indicator observed in ALS patients that directly damage motor neuron [7,44]. KIF5 protein is also involved in transportation RNA and RNA-binding protein in the region of dendrite and axons [38] by the help of cargo proteins such as FUS and hnRNPA1 associated with ALS [45–47]. Also, KIF5 involved in VAPB transportation by the help of adaptor protein protruding [48] and mutation in VAPB has been reported to be associated with ALS and muscular atrophy [49]. It is also known for transportation of neurofilament in axonal region and abnormal accumulation of neurofilament serve as a marker for ALS [50], mutation within neurofilament heavy polypeptide (NEFH) were found to be associated with ALS [51].

KIF5 protein is majorly responsible for transportation of mitochondria and any defects in the transportation would significantly decrease the survival and serve as the hallmark for ALS [52]. It is also involved in GABA_A receptors and AMPA-type transport [53,54]. ALS associated genes such as NEK1 and PFN1 play major role in neurite like membrane protrusion formation [48] and its interaction with cytoskeleton enhances the number of protein related to cytoskeleton in KIF5A mutation such as peripherin, PFN1, NEFH and TUBA4A which are implicated in pathogenesis of ALS [12,13].

Studies have shown that variation within the cargo binding C-terminal domain has been related with ALS, while variations in N-terminal domain due to missense mutation lead to CMT2 and hereditary

spastic paraplegia. These variations due to missense mutation within N-terminal domain leads to disruptive anterograde transportation mediated by KIF5A within the axonal and dendritic region of the cargo which further affect binding of microtubule and hydrolysis of ATP. These would further degrade the axonal retrograde transport process as seen in CMT2 and hereditary spastic paraplegia [16]. Lesions within the protein aggregates in cytoplasmic region are seen consistently in motor neuron cell and spread in neurites anterogradely is observed in ALS. Studies in Zebrafish has indicated that Loss of function variants within C-terminal domain or truncation in C-terminal leads to damaging localization of mitochondria in axonal region and distort binding of specific cargo protein [55]. Within the cell body there in aggregation and distortion of cargo binding causes its hoarding which results in neurite deficiency at the terminals. It has been observed that patients with multiple sclerosis has pile up of amyloid protein precursor and phosphorylated neurofilament in the cell bodies of neurons which leads to neurodegeneration due in adequate amount of KIF5A expression and cargo binding [56].

3. METHODOLOGY

3.1 Data mining for SNPs:

Retrieval of protein sequence and SNP of KIF5A was carried out from NCBI (in FASTA format) and dbSNP database respectively for *in-silico* analysis. **Figure 1** provides detailed schematic representation of the study; our search was only limited to non-synonymous missense SNPs and was further subjected to analyze their deleterious and disease causing effects on KIF5A protein.

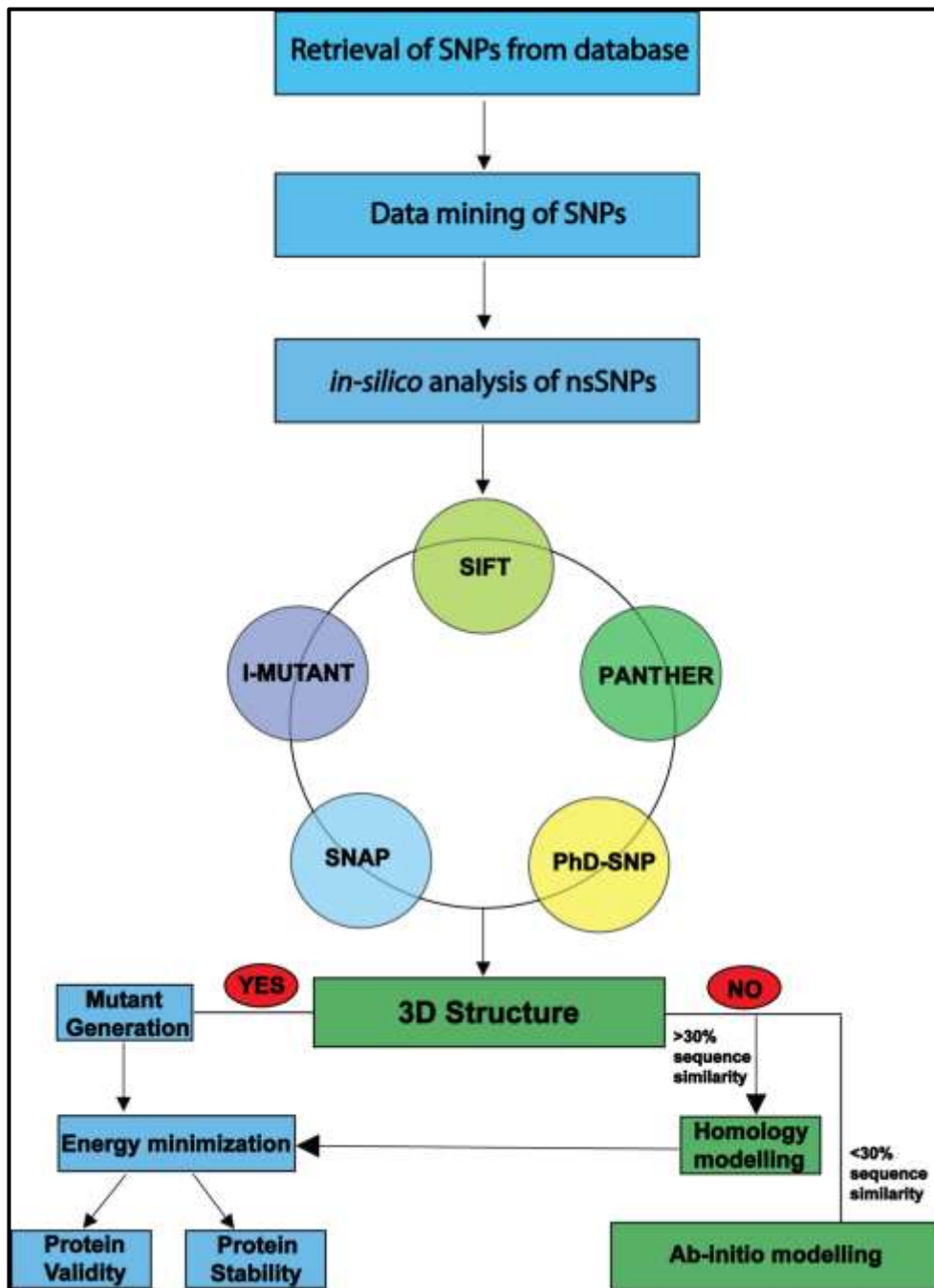


Figure 1: Schematic representation of the workflow

3.2 *In-silico* prediction of deleterious and damaging nsSNPs in KIF5A:

Prediction of highly damaging and deleterious missense SNPs associated with KIF5A gene was first screened based on sequence information by SIFT (Sorting Intolerant from Tolerant), version 2. SIFT evaluates importance of substitution on the basis of sequence homology, physical properties of amino acid and also calculates conservation of evolutionary sequence among the species. It generate the SIFT score (ranges from 0 to 1) to predict nature of substitution, score ≤ 0.05 were classified as damaging and > 0.05 as tolerated. In our study, sequence information was used as query to predict deleterious or damaging nature of missense variant [57]. PANTHER (Protein Analysis Through Evolutionary Relationships) server was used to evaluate the impact of substituted amino acid on the functional and structural aspects of protein based on evolutionary conservation among the lineage by developing HMM. It evaluates time-length of a preserved amino acid in lineage resulting to protein of interest. Longer the preservation time, greater the likelihood of functional impact [58]. PhD-SNP (Predictor of human Deleterious Single Nucleotide Polymorphisms) server predicts the effect of nsSNPs on the basis of support vector machines (SVM) and sequence information [59]. SNAP² evaluates changes caused by nsSNPs on the functionality of the protein by developing neural-network method that uses information such as secondary structure and compares solvent accessibility, conservation of the native and mutated protein. Mutation list and protein list is required to develop network. The software predicts the reliability index and generate a score for each substitution which can be distinguished into neutral (-100, strongly predicted) or effect (+100, strongly predicted) [60]. A support vector machine based (SVM-based) tool I-Mutant2.0 is an SVM-based tool which evaluates changes in the stability of protein upon single-site mutation. It determines disease related nsSNP on the basis of sequence and structure information of protein and predicts sign (+ increase, - decrease) and value of free energy changes (DDG) at a given temperature. It utilizes ProTherm, an experimental protein mutation database [61].

3.3 3D Model Construction:

Three dimensional structure of KIF5A protein was predicted using Iterative Threading ASSEmbly Refinement (I-TASSER). I TASSER utilizes hierarchical approach to determine protein structure and function. Initially, it recognizes the structural template from PDB by utilizing multiple threading approaches, LOMETS, with construction of full-length atomic models through iterative template fragment assembly simulations. Three dimensional model generated by I-TASSER was evaluated in

PDBsum that provides pictorial view of the molecules and construct the structure such as DNA, protein chains, ligands with metal ions and schematic representation of their interactions [62].

3.4 Structure Validation:

The topmost 3D protein model having highest score were selected and further subjected to structural validation by PDBsum and PROCHECK to generate Ramachandran Plot. Ramachandran Plot evaluates dihedral angle of the amino acid residues and based on their phi and psi dihedral angle it determine energetically allowed residues, thus ascertaining structural and functional properties of a protein [63,64]. PDBsum provides pictorial view of the 3D structure of protein and also provides information about angles, helices, motifs, beta sheets and strands. A good protein structure has more than 90% residues in favoured region [65]. PROCHECK evaluates stereo chemical quality of protein by evaluating residue-by-residue geometry and its overall structural geometry [66]. Further quality assessment of model structure was performed by RAMPAGE. It takes energy minimized structure of modeled protein as the input. The server evaluates the protein structure on the basis of dihedral angle and number of residues in favourable, allowed and disallowed region based on Ψ and Φ angles [67].

3.5 Stability changes prediction of mutant protein:

nsSNPs that were predicted to be highly damaging through SIFT, PANTHER, PhD-SNP, SNAP and I-Mutant server were further narrowed down on the basis of minor allele frequency (>5%). nsSNPs with highly deleterious and damaging effects were taken into consideration for further analysis. Then the mutants structure were generated using WHATIF server [68] and further energy minimization of mutants and native structure was performed using YASARA [69]. Structural stability of mutant with respect to wild was calculated using FOLDX. It utilizes atomic description of protein structure to produce quantitative estimation of significant interactions imparting to stability of protein [70].

3.6 Active Site Prediction

Prior to docking prominent binding site prediction of KIF5A protein was identified using MetaPocket 2.0 server. Top 3 major pocket binding pockets were retrieved for analysis of active binding residues and comparison of docking results. MetaPocket uses consensus method to predict ligand binding sites present on the protein surface. It utilizes methods such as LIGSITE^{CS}, PASS, Q-site Finder, SURFNET, Fpocket, GHECOM, ConCavity and POCASA to improve success rate of prediction. The topmost

binding site was selected and active site residues within 10Å from the mutated sites that are present in the server were selected as prominent binding site residue [71].

3.7. Virtual Screening

Compound libraries of around 1 lakhs natural occurring compounds present in the ZINC version 12 database were inputted for screening purpose. ZINC version 12 is a freely available database consisting of curated chemical compounds in ready to dock 3D format that can be used for virtual screening. LigPrep tool was used to prepare ligands that can be used furthered. The set of ligands prepared were docked to binding site of the target protein using Glide module 8 of Schrödinger suite for purpose of virtual screening. Glide utilizes grids for fast scoring and ligands were filtered out based on Lipinski rule of five as well as Veber criteria and reactive functional group, OiKProp considering both ADME and blood- brain- barrier properties. Glide provides a wide range of speed vs accuracy options, high throughput virtual screening (HTVS) mode for efficiently enriching millions of compound libraries, then standard precision (SP) mode for reliably docking tens to hundreds of thousands of ligand with high accuracy, and extra precision (XP) mode where false positive are further eliminated by means of extensive sampling and advanced scoring, which leads to higher enrichment. Glide ranked the ligands in the order of the lowest docking score [72].

3.8. ADME properties

The Absorption Distribution Metabolism and Elimination (ADME) properties were evaluated using QikProp module of Schrodinger suite for computing drug ability and filtering out compounds at an early stage. It consists of various pharmacokinetic properties such as rule of five, blood-brain barrier permeability, octanol/water coefficient, human percent oral absorption etc [73].

3.9. Molecular Docking

Top 10 compounds with the lowest GLIDE score were filtered out and were further used to carry out docking procedure. GOLD (Genetic Optimization for Ligand docking) was utilized to perform docking procedure, it uses genetic algorithm to search range of ligands possessing ligand flexibility, such as full range of flexibility in acyclic ligand and partial cyclic ligand, together with fractional protein flexibility of in neighborhood of the active site of protein, and assures the basic need of displacement of loosely bound water with binding of ligand. Docking with GOLD utilizes ten independent runs were performed for each molecule and generated a simple scoring function to evaluate binding complexes that are

generated. Higher the Gold Fitness score better the solution for docking. Further the topmost compound was analyzed with the known drug for molecular dynamics simulations [74].

4. RESULTS

4.1. SNPs id retrieval:

SNPs in KIF5A were retrieved from dbSNP database in NCBI, which included 511 non-synonymous missense variants, 7 frame-shifts, 308 synonymous and 9 nonsense mutants. Furthermore we selected non-synonymous missense variants for our analysis in order to predict their effects on structure, function and stability of the protein (**Figure 2**).

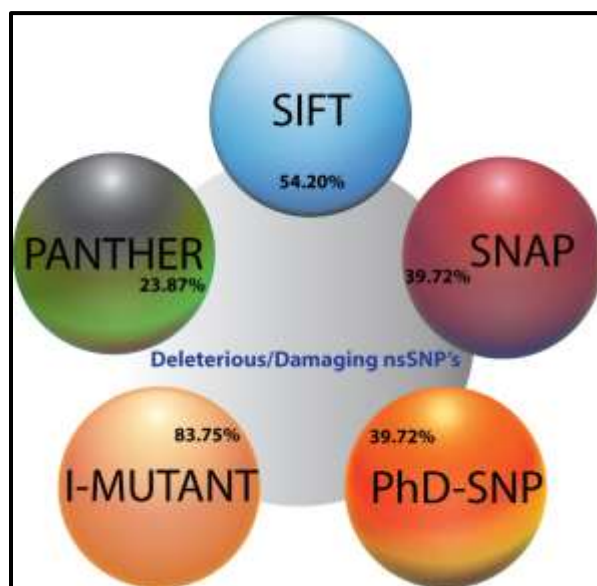


Figure 2: A statistical representation of the deleterious/damaging nsSNP predicted by various *in-silico* tools.

4.2 Computation analysis of nsSNP based on sequence based prediction tools:

The combine result of highly deleterious and damaging SNPs common in all prediction tool (SIFT, PANTHER, PhD-SNP, SNAP and I-Mutant) are provide in **Table 1**. SIFT evaluates all the nsSNPs based on the sequence homology and physical nature of amino acid. From a total of 512 nsSNPs, 277 substitutions were found to be deleterious and were marked as “Affects protein function” (tolerance index 0.00) while the remaining 234 substitutions were tolerated. PANTHER and PhD-SNP predicted 203 nsSNPs to be deleterious in nature based on HMM and SVM based methods respectively. Heatmap of KIF5A protein was generated based on neural-network using SNAP² tool (**Figure 3**), where 203 nsSNPs were predicted to be damaging and 308 to be neutral. I-mutant server predicted change in protein stability due to substitution on the basis of change in DDG (free energy change), among the nsSNPs submitted 428 nsSNPs were predicted to be associated with decrease in the stability while the remaining showed increase in the stability of

protein upon mutation. Additionally, 109 nsSNPs predicted to be damaging in nature, affecting the protein structure were found to be common in all the prediction software including SIFT, PANTHER, PhD-SNP, SNAP2 and I-mutant. Of these nsSNPs, only 5 nsSNPs having rsids rs139015012 (A268T), rs140929639 (R369W), rs200965784 (T644M), rs373969485 (R712L) and rs113247976 (P986L) were chosen based on minor allele frequency (>5%) for further analysis.

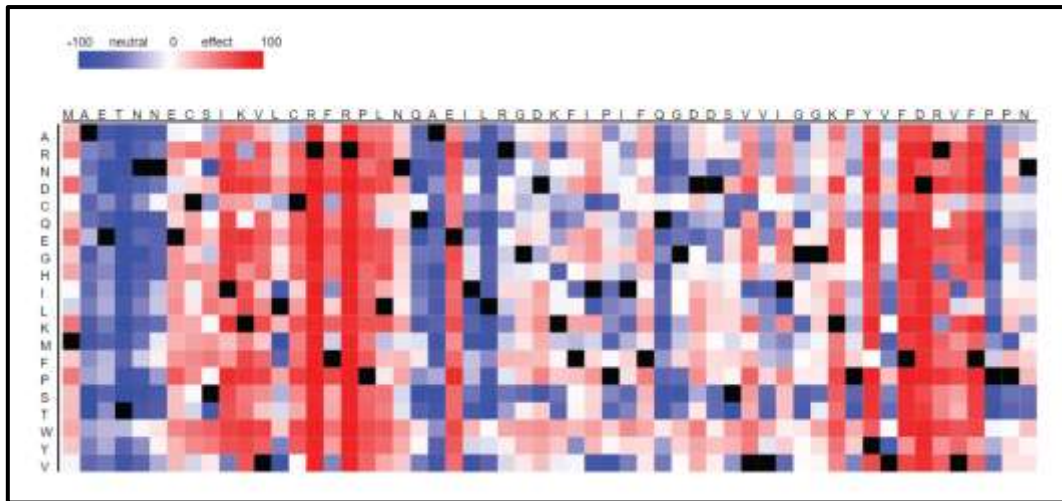


Figure 3: Heatmap of KIF5A generated by SNAP

Variant id	AA variant	SIFT score	SIFT Prediction	PANTHER Prediction	PhD-SNP Prediction	SNAP score	SNAP Prediction	DDG value (Free energy change)	I-Mutant Prediction
rs139015012	A268T	0.01	A	PD	D	4	E	-1.21	De
rs140929639	R369W	0.02	A	PD	D	48	E	-1.03	De
rs200965784	T644M	0.01	A	PD	D	2	E	0.17	De
rs373969485	R712L	0.01	A	PD	D	11	E	-0.32	De
rs113247976	P986L	0	A	PD	D	2	E	-0.79	De

Table 1: Functionally important substitution predicted by *in-silico* screening;

4.3 Three dimensional structure prediction and validation:

3D structure for KIF5A protein was generated by I-Tasser based on the protein sequence information. Model with the highest C-score (-1.2) was selected and considered as a good model (**Figure 4**). To further validate the structure, PDBsum and PROCHECK servers were utilized. Ramachandran Plot (RC plot) was obtained from PROCHECK program, which checks stereo-chemical quality of protein structure and generate a number of postscript plots, analyzing residue-by-residue and overall geometry. The reliability of the predicted model was assured with 70% of residues in most favoured region, 22.5% of residues in allowed region and 4.5% in generously allowed region however only 3% of residues lied in disallowed region (**Figure 5**). Energy minimized structure of modeled protein was further analysed for quality assessment by RAMPAGE (**Figure 7a**).

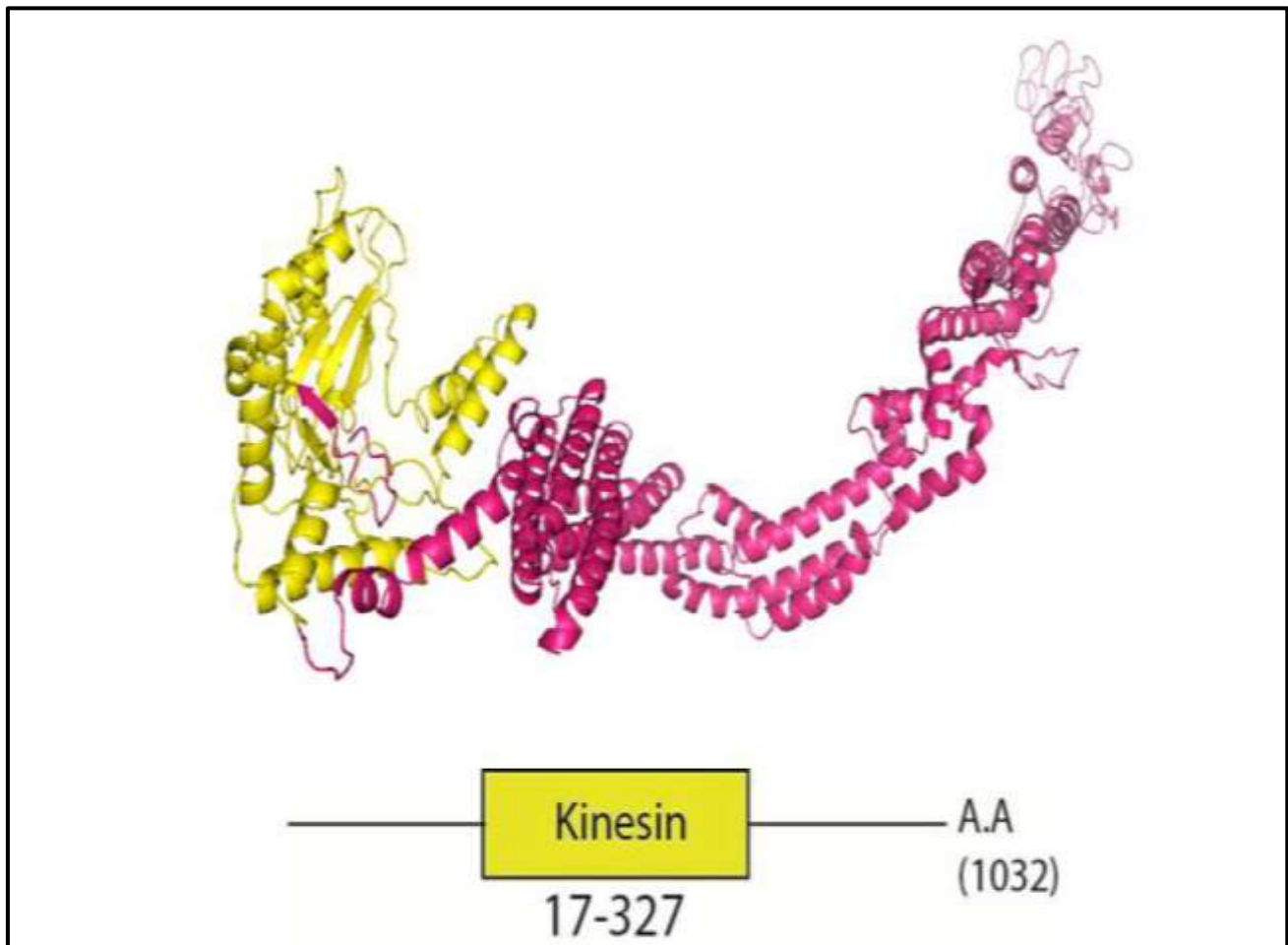


Figure 4: 3D structure of KIF5A modeled using *ab-initio* method

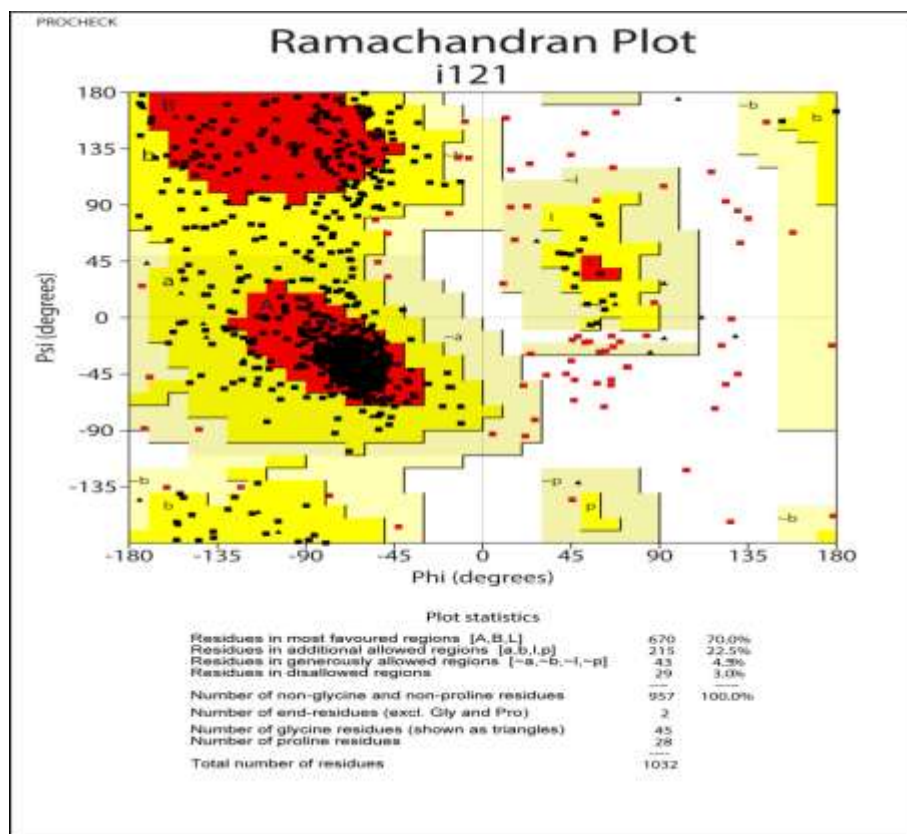


Figure 5: Ramchandran plot of modeled KIF5A protein

4.4 Prediction of stability changes of mutant proteins by FOLDX:

5 nsSNPs (A268T, R369W, T644M, R712L and P96L) cumulatively predicted as highly deleterious and damaging by SIFT, PANTHER, PhD-SNP, SNAP2 and I-mutant were modeled using WHATIF server and were further energy minimized by YASARA (**Figure 6**).

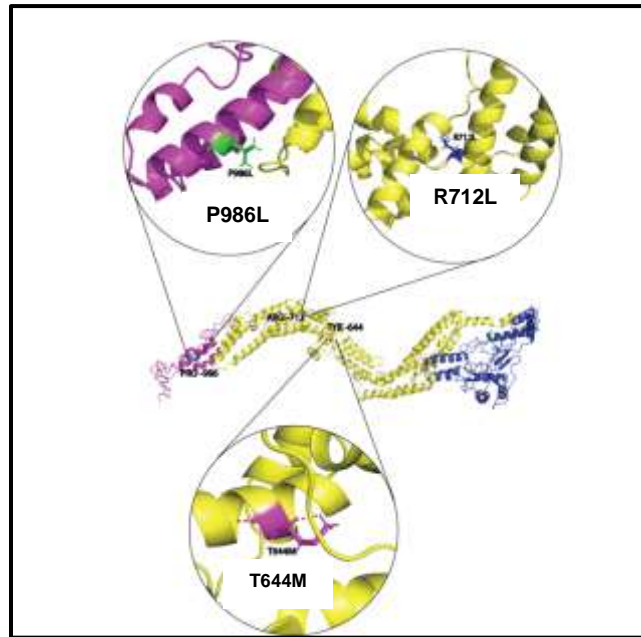


Figure 6: Structure of mutants protein generated by WHATIF server

All three substitution predicted as most deleterious by *in-silico* analysis are highlighted in stick model, T644M in purple, R712L in blue and P986L in green with polar hydrogen bonds surrounding it.

Quality of all the 5 nsSNPs was assessed by RAMPAGE (**Figure 7b-f**) which is indicative of Ramachandran Plot and assured good quality of the mutant protein. Structural stability of KIF5A variants was analyzed in respect to wild by FOLDX which depicts prominent increased stability in 3 variants as compared to wild type (**Table 2**) based on total energy along with other factors, including Van der Waals clashes, Electrostatics and Van der Waals, showing a significant increase in total energy. T644M, R712L and P986L were determined as prominent mutation affecting the protein. These variants were further analyzed by HOPE (Have yOur Protein Explained) which predicted all these 3 nsSNPs to be highly deleterious and damaging in all prediction server used [75]. At 644th position wild residue (Threonine) of KIF5A is highly conserved and any substitution or mutation at this position would damage structure of the protein. HOPE predicts that T644M mutation would leads to the loss of hydrogen bonds since the changes in the exposed residue at this site will disturb the correct folding of the protein. Also 712th position is highly conserved in wild type (Arginine) and substitution (Leucine) would introduce hydrophobicity in the structure, leading to the loss of hydrogen bonds and thus destroying the interaction of the exposed wild-type residue. At 986th position wild-type residue (Proline) of KIF5A is 100% conserved and mutation at this position would damage the protein and leads to disease ALS as it is located in splice variant. HOPE predicts that mutation at this position would create bumps in the structure and would distort the local structure of the protein. These mutations chosen on the basis of FOLDX and HOPE can be further analyzed by molecular dynamics simulation to

understand the internal motions and conformational changes in a definite time scale in mutant with respect to wild-type residue of KIF5A.

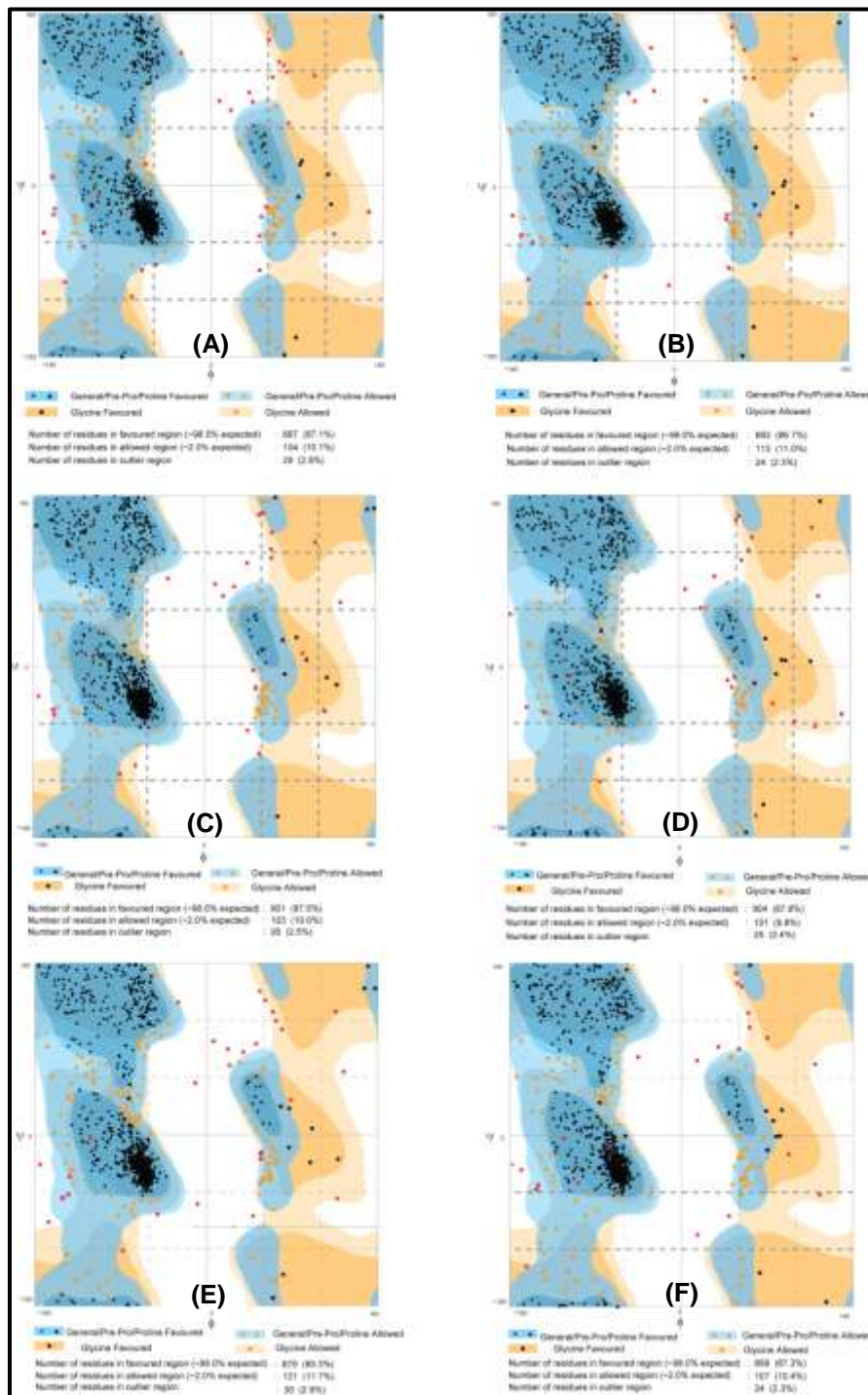


Figure 7: The Ramachandran plot of native and mutant protein generated by Ramapage. (A) The native KIF5A protein showed 87.1% in favoured, 10.1% in allowed and 2.8% in outlier (B) A268T (86.7%, 11% , 2.3%) (C) P986L (87.5%, 10%, 2.5%) (D) R369W (87.8%, 9.8%,2.4%)(E) R712L (85.3%, 11.7%, 2.9%) and (F) T644M (87.3%, 10.4%, 2.3%) residues in favoured, allowed and outlier region.

ENERGY	A268T	R369W	T644M	R712L	P986L	WILD
Total Energy	615.02	606.91	661.57	648.18	633.26	624.44
Back H bond	-721.78	730.94	-709.29	-723.66	-730.48	-730.35
Side H bond	-334.77	-316.22	-326.58	-315.20	-328.82	-327.20
Vander Waal Energy	- 1124.01	-1126.83	- 1124.92	- 1123.52	-1127.31	-1125.58
Electrostatic Energy	-48.39	-47.83	-48.73	-47.19	-48.32	-54.70
Solvation Polar	1742.76	1735.43	1757.82	1751.18	1755.898	1749.14
Solvation Hydrophobic	- 1386.09	-1393.35	- 1390.40	- 1389.15	-1390.45	-1388.59
Van der Waals clashes	45.44	45.57	52.81	43.73	44.38	38.12
Torsion energy	36.09	36.09	37.25	36.26	34.78	39.59
Backbone Van der Waals	895.60	886.66	899.50	907.80	903.38	892.31
Entropy side chain	675.04	670.75	678.96	671.20	678.50	678.18
Entropy main chain	1741.22	1747.55	1740.65	1754.43	1756.93	1752.05
Water bonds	0.00	0.00	0.00	0.00	0.00	0.00
Helix dipole	-10.14	-13.91	-7.00	-11.24	-11.60	-8.56
Loop entropy	0.00	0.00	0.00	0.00	0.00	0.00
Cis bond	1.12	1.12	1.12	1.12	0.15	2.89
Disulfide	-5.62	-5.63	-5.60	-5.63	-5.63	-5.62
Electrostatic	0.00	0.00	0.00	0.00	0.00	0.00
Partial covalent interactions	0.00	0.00	0.00	0.00	0.00	0.00
Energy ionization	4.13	5.11	5.49	5.84	5.24	5.06
Entropy complex	0.00	0.00	0.00	0.00	0.00	0.00

Table 2: Stability analysis of functionally important nsSNPs

4.5. Active Site Prediction:

Active site was identified using MetaPocket 2.0 which identified the top three binding pocket within the protein. The topmost binding pocket was identified and analyzed further. This helped us to predict the binding pocket within the 10Å region of the residue Pro986 which also included the other active site residues too. Active site identified (**Figure 8**) predicts these residues (SER 977 GLY 979 ALA 980 THR 981 SER 982 SER 983 GLY 984 GLY 985 LEU 986 LEU 987 ALA 988 SER 989 TYR 990 GLN 991 LYS 992) as most prominent site for binding of inhibitor.

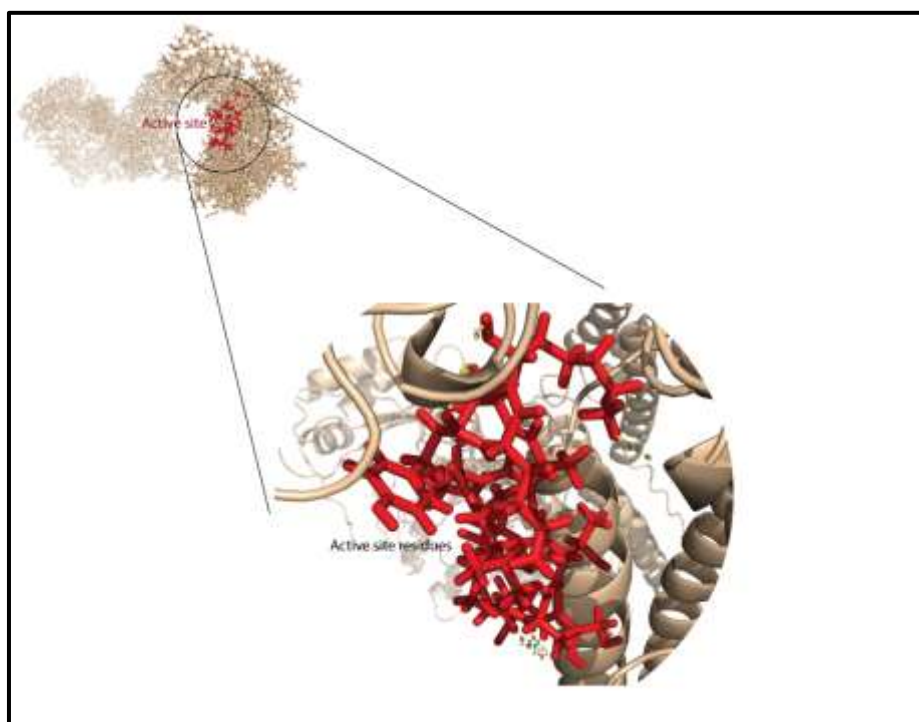


Figure 8: Active site residues (shown in red stick model) for KIF5A

4.6 Virtual Screening

Natural compounds from different company databases available in ZINC were downloaded. ZINC database contain collection of chemical compounds that are available commercially, especially for virtual screening was used. About 167504 natural compounds from 12 different databases were retrieved for virtual screening. Further ligands were prepared in Schrödinger suit using LigPrep application. Energy minimization was performed to generate conformers for each compound which is an essential step for virtual screening.

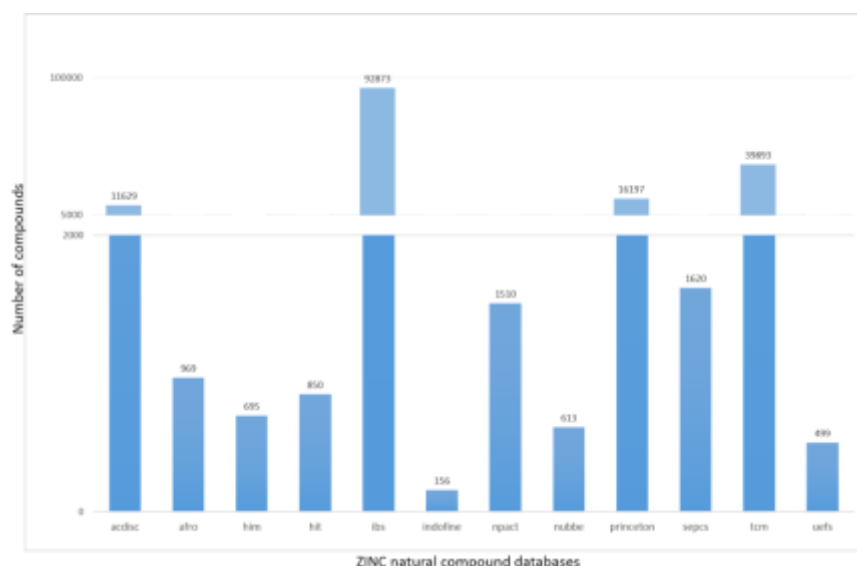

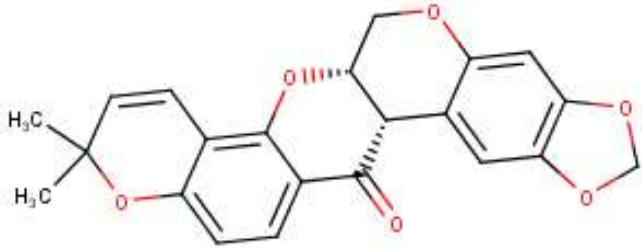
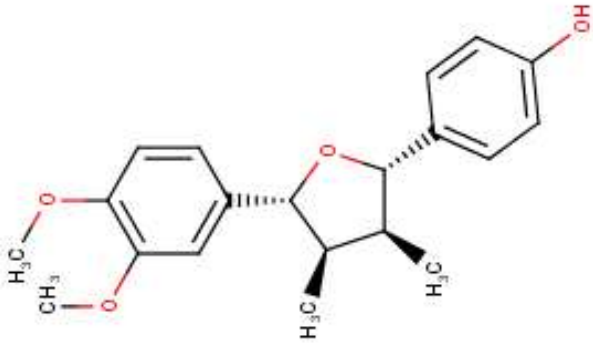
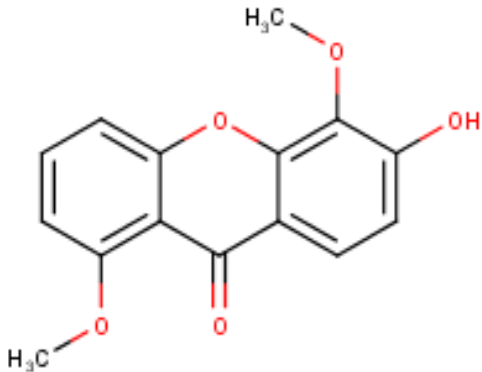


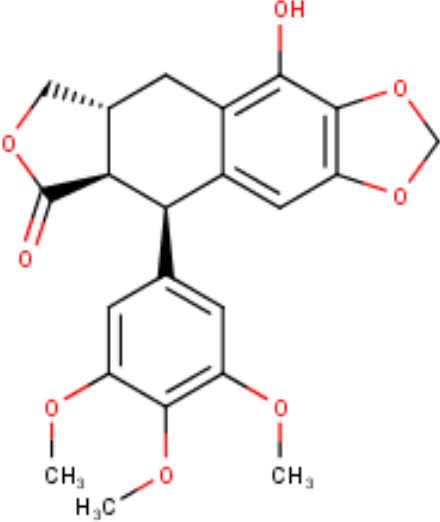

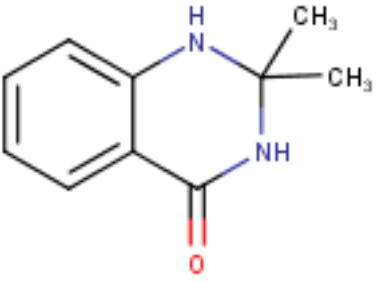
Figure 9: Natural compounds from different ZINC databases

4.7. Molecular Docking:

Compound library was screened for ADME properties in QikProp suite of Schrodinger which predicted approximately 15,000 compounds most likeable to be a drug molecule. These compounds were screened using three algorithm steps within GLIDE software. From each step top 10% compounds were used in the next step to generate list of natural compounds with the increasing rank. Virtual screening with GLIDE program was used for screening library with standard default parameters. Library screening was performed to sort out non-docked compounds from the library. Top 20 compounds with best lowest glide score were further analyzed in detailed docking by utilizing GOLD program. The topmost compound with best GOLD Fitness score was selected for further simulation for validation of the process.

Ligand name	Ligand structure	Glide score	GOLD score
<u>Columbianetin</u>		-6.044	59.02

<p><u>(1R,2S)-1-[2-(1,3-benzodioxol-5-yl)-3-methyl-benzofuran-5-yl]propane-1,2-diol</u></p>		<p>-5.947</p>	<p>52.83</p>
<p><u>Millettone</u></p>		<p>-5.88</p>	<p>50.26</p>
<p><u>4-[(2S,3S,4R,5R)-5-(3,4-dimethoxyphenyl)-3,4-dimethyl-tetrahydrofuran-2-yl]phenol</u></p>		<p>-5.849</p>	<p>50.09</p>
<p><u>6-hydroxy-1,5-dimethoxy-xanthen-9-one</u></p>		<p>-5.806</p>	<p>48.46</p>

<p><u>beta-peltatin</u></p>	 <p>The structure shows a complex polycyclic system. It features a central benzene ring with three methoxy groups (-OCH₃) at the 1, 3, and 5 positions. This ring is attached to a larger ring system that includes a lactone ring and a hydroxyl group (-OH).</p>	<p>-5.806</p>	<p>47.91</p>
<p><u>(3R,4S)-4-hydroxy-3-methoxy-4-(4-methoxyphenyl)-1,3-dihydroquinolin-2-one</u></p>	 <p>The structure shows a quinolinone core. At the 4-position, there is a hydroxyl group (-OH) on a dashed bond and a 4-methoxyphenyl group (-C₆H₄-OCH₃) on a solid wedge. At the 3-position, there is a methoxy group (-OCH₃) on a solid wedge.</p>	<p>-5.774</p>	<p>46.29</p>
<p><u>2,2-Dimethyl-1,2,3-trihydroquinazolin-4-one</u></p>	 <p>The structure shows a quinazolinone core. The nitrogen at position 2 is substituted with two methyl groups (-CH₃).</p>	<p>-5.77</p>	<p>44.54</p>

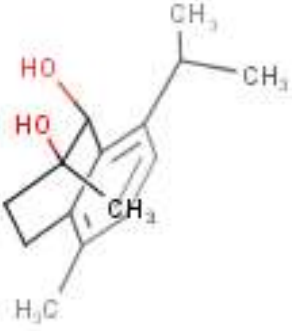
<p><u>(1R,2S)-8-isopropyl-2,5-dimethyl-tetralin-1,2-diol</u></p>		<p>-5.765</p>	<p>44.37</p>
---	---	----------------------	---------------------

Table 3: Chemical structure, IUPAC names and fitness scores of top scored ligands docked with KIF5A using docking program GOLD

Ligand Name	log S for aqueous solubility	Predicted CNS activity	log BB for brain / blood	Apparent MDCK permeability (nm/s)	% Human Oral Absorption	log Kp for skin permeability	log Khsa for serum protein binding
<u>Columbinetin</u>	-3.24	0	-0.231	1076.222	93.349	-2.438	-0.184
<u>(1R,2S)-1-[2-(1,3-benzodioxol-5-yl)-3-methyl-benzofuran-5-yl]propane-1,2-diol</u>	-2.918	0	-0.08	1578.226	100	-1.839	0.054

<u>Millette</u>	-3.643	0	- 0.606	512.036	94.693	-2.896	-0.13
<u>4-[(2S,3S,4R,5R)-5-(3,4-dimethoxyphenyl)-3,4-dimethyl-tetrahydrofuran-2-yl]phenol</u>	-3.274	0	- 0.255	1131.112	100	-1.76	-0.052
<u>6-hydroxy-1,5-dimethoxy-xanthene-9-one</u>	-4.136	-1	- 0.781	647.694	96.949	-1.9	-0.171
<u>beta-peltatin</u>	-2.132	0	- 0.047	2120.414	100	-1.64	-0.685
<u>(3R,4S)-4-hydroxy-3-methoxy-4-(4-methoxy</u>	-5.558	0	- 0.639	520.142	100	-2.023	0.461

<u>phenyl</u> <u>-1,3-</u> <u>dihydro</u> <u>quinolin</u> <u>-2-one</u>							
<u>2,2-</u> <u>Dimeth</u> <u>yl-</u> <u>1,2,3-</u> <u>trihydro</u> <u>quinazo</u> <u>lin-4-</u> <u>one</u>	-3.499	-1	- 0.277	2057.25	100	-1.478	-0.477
<u>(1R,2S)-</u> <u>8-</u> <u>isoprop</u> <u>yl-2,5-</u> <u>dimeth</u> <u>yl-</u> <u>tetralin-</u> <u>1,2-diol</u>	-2.798	1	0.395	1031.715	100	-3.746	0.287

Table 4: ADME and pharmacological parameters prediction for active compounds using QikProp

5. DISCUSSION

KIF5A is a motor protein involved in the transportation of the intracellular organelles, RNA and protein in microtubule based manner [76]. It mediates the transport of RNA and RNA binding protein within the dendritic and axonal region. KIF5A is involved in anterograde transportation of mitochondria with the help of adaptor protein, important or essential for maintaining viability and functioning of neurons. Dysregulation of mitochondrial transportation is associated with various neurodegenerative diseases [77,78]. It has been previously reported that missense substitution in KIF5A motor domain and alpha-helical coiled-domain is associated with hereditary spastic paraplegia and Charcot-Marie-Tooth disease type 2 (CMT2) [79,80]. Recently, mutation within the C-terminal region of KIF5A affecting splicing in exon 27 was associated with amyotrophic lateral sclerosis (ALS) [8]. KIF5A plays a central role in axonal transportation which has provided way to introspect that mutation within KIF5A can lead to disruption of axonal transport which contribute to pathogenesis of motor neuron degeneration [76]. Mutation in KIF5A would lead to accumulation of neurofilament and impaired transportation of mitochondria which can serve as the hallmark for neurodegenerative disease [46, 81–83]. Studies conducted in zebrafish have shown loss-of-function within the C-terminal leads to truncation and disruption of axonal localization within mitochondria [55]. In neuronal cell body decreased expression of KIF5A and binding of cargo results in accumulation of amyloid precursor proteins and phosphorylated neurofilament and serve as the hallmark for neurodegeneration[56].

In-silico analysis would provide us key to predict the effect of single nucleotide variation on structure and function of the protein thus serving an alleyway to study its role in neuroedegeneration. In our study we used sequence based prediction tools to deduce the effect of nsSNPs on KIF5A structure and function.

High-throughput computational screening of nsSNPs associated with KIF5A protein predicted 5 nsSNPs: rs139015012 (A268T), rs140929639 (R369W), rs200965784 (T644M), rs373969485 (R712L) and rs113247976 (P986L) to be highly deleterious and damaging to structure and function of KIF5A protein. Furthermore variants T644M, R712L and P986L were predicted to increase the stability of the mutants rendering changes in the structure of the protein. A recent study has revealed that missense variant P986L located in the C-terminal domain of KIF5A can lead to amyotrophic lateral sclerosis (ALS) [8]. Variation within C-terminal cargo binding domain at 986th position has been associated with abnormal splicing at exon 27 implicating

KIF5A as a novel gene associated with ALS phenotype and pathogenesis [84]. This *in-silico* based study needs to be validated *in-vitro* to determine the role and exact pathway through which these nsSNPs with increased protein stability would alter the functioning of protein and would lead to disease progression. This study, thus, paves the gateway for studying the role of these nsSNPs in KIF5A protein thus leading to neurodegeneration and developing a mechanism for drug targeting and biomarkers for therapeutic application.

Specificity and accuracy are the major criterion for the drug discovery technology. Drug molecule should never hamper the hemostasis of the host system. Our initial work in computation screening of nsSNP of KIF5A depicted three missense variant (T644M, R712L, P986L) were observed to have high chances of affecting protein function. Of these Pro986Leu missense variant was found to be most damaging to KIF5A function, which was further analyzed through simulations.

The protein structure was unavailable and *ab-initio* procedure was employed to predict model of KIF5A. Model structure with best confidence score (C-value = -1.40) was selected for further structural and functional analysis. We performed both PROCHECK and VERIFY3D softwares to evaluate the quality of the modeled structures and Ramachandran Plot obtained from PROCHECK ensures the stereo chemical quality of the modeled protein structure, which analyses residue-by-residue and overall geometry, assured reliability of structure having 70% residues in allowed region and additional 22.5% allowed residues. VERIFY3D provides “3D-1D” profile on the basis of local environment of each residue such as side-chain fraction, local secondary structure which provides the reliability of dimensional structures. Further MD simulation of the modeled structure was performed for 30 ns to validate the structure. It was observed that there were no structural discrepancies in radius of gyration, hydrogen bonds, rmsd and SASA values. Data retrieved provide a stable modeled structure with low energy state that was used further for docking.

Binding pocket was defined around 10 Å region of Pro986, which included all other residue also. Multiple sequence alignment of the target protein depicts that the active site is highly conserved among the species in the course of evolution. The top best 10 ligands were chosen based on the GLIDE and GOLD score. Virtual screening using GLIDE program was performed using standard default setting. Library screening filtered out docked ligands from the non-docked ligands from compound library used in virtual screening. After filtering non-docked ligands further remaining compounds were used in detailed docking. The approach used evaluate protein docking with primary score function (GLIDE

energy) and then re-scoring output list with secondary score function (GOLD Fitness Score). Lead molecule that would be identified through computational screening would require to test experimentally in order to confirm the proposed LEAD molecule. The significance of this work is providing a relatively inexpensive approach to screen.

6. CONCLUSION

In conclusion, computational analysis of screening non-synonymous missense variation revealed three SNPs to be highly deleterious and damaging altering structure and function of KIF5A. These nsSNP can be mapped into the KIF5A that would help to reveal their mechanism to disrupt axonal transportation and lead to neurodegeneration.

This study integrates structural analysis, virtual screening, molecular docking and MD simulations to design a novel synthetic inhibitor compound to inhibit missense variant of KIF5A. In this work we have proposed probable chemical molecules which could be tested to devise drug molecules. In absence of crystal structures, we used *ab-initio* modeling to predict the three dimensional structure with the good reliability score. Virtual screening was carried out using GLIDE program, followed by GOLD scoring algorithm and ADME properties. The scope of this work is to use data for cost-effective experimental screening. The proposed lead molecule could provide a potential drug to treat ALS.

APPENDIX

S.No	dbSNP rs#	WposM	SIFT PREDICTION	PANTHER	PHD-SNP	SNAP Prediction	Imutant
1	rs758595412	M1L					
2	rs780433858	A2V					
3	rs1244142117	E3Q					
4	rs755532095	N5D					
5	rs1434215781	N6D					
6	rs940856567	C8Y					
7	rs1376496207	S9N					
8	rs748864821	K11T					
9	rs748864821	K11R					
10	rs770302674	P18S					
11	rs1371053668	A22T					
12	rs745558435	A22G					
13	rs771847226	L25Q					
14	rs1161092929	D28E					
15	rs573410126	F30L					
16	rs199955108	I31T					
17	rs540463538	I33F					
18	rs1207026111	I42V					
19	rs149569914	I42M					
20	rs769763596	R51C					
21	rs773336059	R51H					
22	rs536777412	F53L					
23	rs1169287923	P55S					
24	rs1014548294	P55Q					
25	rs1162282839	N56H					
26	rs774510700	N56T					
27	rs759785671	T57M					
28	rs1348634786	T58S					
29	rs975867260	H64R					
30	rs756639633	A65V					
31	rs1230201278	M68T					
32	rs1288040135	Q69H					
33	rs1060502525	G77D					
34	rs1299838179	N79S					
35	rs1060502524	S90L					
36	rs1208551351	H94R					
37	rs1237275454	M96V					
38	rs1237275454	M96L					
39	rs1485959906	M96I					
40	rs1193791841	K99N					
41	rs1247952950	L100P					

42	rs1171944114	D102N					
43	rs777103564	I109M					
44	rs1161615188	R111Q					
45	rs762474618	I112V					
46	rs144277716	R114Q					
47	rs1332933023	H119Y					
48	rs887626771	Y121C					
49	rs1434891074	E125D					
50	rs758987045	K142R					
51	rs978901860	R144C					
52	rs1060502522	D145H					
53	rs1207892747	T152S					
54	rs1343723233	T152K					
55	rs1273893076	V156M					
56	rs1469562271	E158K					
57	rs1191334617	D159G					
58	rs748551786	R162W					
59	rs770143931	R162Q					
60	rs1429751650	V166I					
61	rs748426915	C169F					
62	rs1159429693	R172C					
63	rs1392363068	S176G					
64	rs749561961	S176R					
65	rs1319232269	P177L					
66	rs1299152245	D185N					
67	rs147510678	G187A					
68	rs140144799	S189L					
69	rs769315791	R191C					
70	rs1488871976	R191H					
71	rs1193407286	H192R					
72	rs1265956693	V193M					
73	rs773071687	A194V					
74	rs762585533	V195F					
75	rs879254292	E200K					
76	rs1057524193	S202R					
77	rs1057519195	S202T					
78	rs387907287	R204Q					
79	rs1162007628	I208V					
80	rs769491011	M218V					
81	rs1409047729	E219A					
82	rs772697257	T220M					
83	rs911601615	S225C					
84	rs940334329	L228V					
85	rs690016545	D232N					
86	rs387907289	G235E					

87	rs151129834	V248M					
88	rs375693647	D250E					
89	rs387907285	E251K					
90	rs1191735053	A252T					
91	rs121434441	N256S					
92	rs947028044	K257M					
93	rs1174199852	V265A					
94	rs1131692233	S267P					
95	rs139015012	A268T					
96	rs772172027	E271D					
97	rs373795817	T273S					
98	rs121434443	Y276C					
99	rs1291404967	V277F					
100	rs121434442	R280C					
101	rs387907288	R280H					
102	rs1057523746	K283E					
103	rs1224640834	R286K					
104	rs761812789	S291C					
105	rs761812789	S291F					
106	rs1416085161	R297W					
107	rs1404057685	R297Q					
108	rs1303126235	T298M					
109	rs1290797018	M300T					
110	rs1215911052	S305A					
111	rs1305787310	P306A					
112	rs1295922331	S308G					
113	rs1381157336	S308I					
114	rs919318252	N310S					
115	rs267603608	D311N					
116	rs1048845476	D311E					
117	rs1285398673	A312S					
118	rs199886915	S316T					
119	rs1012819766	R323W					
120	rs1468981472	T330S					
121	rs1385454359	V333I					
122	rs1187290769	L335F					
123	rs1399092435	E336K					
124	rs1298908633	Y346C					
125	rs748119521	E347Q					
126	rs1352725704	E351K					
127	rs200763210	E351V					
128	rs1465327244	T353A					
129	rs1465327244	T353S					
130	rs777886455	T353R					
131	rs1331463458	K354R					

132	rs749301835	A355S					
133	rs771092084	E358D					
134	rs774586838	T359M					
135	rs886049700	I360T					
136	rs121434444	A361V					
137	rs371335708	E364K					
138	rs1434261778	A365S					
139	rs1201387610	E366K					
140	rs764640324	L367P					
141	rs749955896	S368I					
142	rs140929639	R369W					
143	rs1390179670	R369Q					
144	rs1390179670	R369L					
145	rs1333285940	W370C					
146	rs986406337	R371C					
147	rs751273893	R371H					
148	rs754551006	N372K					
149	rs1367412468	E374G					
150	rs767095969	N375D					
151	rs866804200	E380D					
152	rs752522963	R381C					
153	rs755855553	R381H					
154	rs1290297748	A383V					
155	rs143326964	G384R					
156	rs753854351	E385K					
157	rs757142042	E385A					
158	rs1052752751	E386Q					
159	rs1425088200	L389V					
160	rs541181624	E392K					
161	rs541181624	E392Q					
162	rs75907338	E392D					
163	rs776826605	E395D					
164	rs908820800	V399M					
165	rs1021367307	N400S					
166	rs1287914572	N402T					
167	rs1313023330	I405V					
168	rs773924317	V406M					
169	rs1216923500	V407M					
170	rs767181570	R408C					
171	rs1203347152	R408H					
172	rs148888970	A410T					
173	rs541896344	A410V					
174	rs1399145820	E413G					
175	rs758500897	R414W					
176	rs1170006745	R414Q					

177	rs755243063	E419K					
178	rs748402153	R422C					
179	rs1272478065	R422H					
180	rs140917012	R423C					
181	rs778095192	R423H					
182	rs749775296	L424R					
183	rs1192349908	Y425N					
184	rs1201421601	L428F					
185	rs370132723	L428R					
186	rs1217850974	D430N					
187	rs1240356713	D430E					
188	rs1267848296	I435M					
189	rs776102034	N436T					
190	rs1469089525	Q437R					
191	rs1379332888	I442L					
192	rs773133711	I442T					
193	rs886042524	S458A					
194	rs1445439522	S458F					
195	rs573756569	R460Q					
196	rs962405110	E464K					
197	rs1271167685	V466F					
198	rs771021589	R468W					
199	rs774379377	R468Q					
200	rs376249839	E469D					
201	rs1245070281	H472Y					
202	rs1318552404	L473V					
203	rs1183553852	Q474K					
204	rs1373971092	Q474H					
205	rs1323875105	D478N					
206	rs867837134	A479V					
207	rs754236824	A480T					
208	rs757571901	V484M					
209	rs577815919	L488R					
210	rs1467600087	D499N					
211	rs758937106	Q500R					
212	rs1167631801	V505G					
213	rs1002863375	S509R					
214	rs1338931080	Q513R					
215	rs1034361893	Q521P					
216	rs1034361893	Q521R					
217	rs747632443	K522T					
218	rs755635476	Q534K					
219	rs1055742325	R535W					
220	rs777398624	R535Q					
221	rs753612831	Q537K					

222	rs1386445853	V539A					
223	rs1386445853	V539G					
224	rs778864713	R544Q					
225	rs199608047	R546Q					
226	rs771938131	I547M					
227	rs779997056	A548T					
228	rs147493358	L551V					
229	rs1411015114	K556Q					
230	rs142701108	E560K					
231	rs769941982	V563I					
232	rs1385718036	I564T					
233	rs1161389904	V565A					
234	rs763336788	G568R					
235	rs766817381	G568E					
236	rs1299010073	K571M					
237	rs770034039	P573A					
238	rs773435627	E575D					
239	rs754373609	S577G					
240	rs1249966359	S577N					
241	rs555833807	S577R					
242	rs760135493	A579T					
243	rs1242381298	E583D					
244	rs768069998	R588Q					
245	rs1417164425	L589P					
246	rs761362390	I591V					
247	rs1307954078	V598L					
248	rs1471037370	K599R					
249	rs1288743658	R604W					
250	rs931757596	R604Q					
251	rs1348214633	R606W					
252	rs750134009	R606Q					
253	rs978303061	N610S					
254	rs978303061	N610I					
255	rs987234168	V613M					
256	rs757985743	E614D					
257	rs758289742	H616Y					
258	rs202045039	R617C					
259	rs751295639	R617H					
260	rs1217911242	K618R					
261	rs1487870979	M619L					
262	rs748248329	R624W					
263	rs1174310373	R624Q					
264	rs750826157	Q630E					
265	rs1343792922	S634T					
266	rs1485918899	H636R					

267	rs754718000	E637K					
268	rs767277745	R641H					
269	rs1422332328	S642L					
270	rs200965784	T644M					
271	rs779228716	M647K					
272	rs1316855645	S649N					
273	rs374729476	V650M					
274	rs374729476	V650L					
275	rs775994678	E651D					
276	rs1438614778	K653R					
277	rs1236467846	K654R					
278	rs768984357	R655W					
279	rs772780431	R655Q					
280	rs770636366	H656R					
281	rs1427243293	E658K					
282	rs1477671960	Y661C					
283	rs1396694334	S663F					
284	rs1426873247	S665R					
285	rs1398974670	D666N					
286	rs767365871	L668V					
287	rs752474331	L668P					
288	rs142495781	A669T					
289	rs753975521	A669D					
290	rs753975521	A669V					
291	rs1279648981	A673T					
292	rs751973359	E675D					
293	rs1191059574	E679K					
294	rs1261934561	E679D					
295	rs1212119444	A681T					
296	rs1177916529	K685E					
297	rs745954376	K685R					
298	rs538986257	E686K					
299	rs1322426342	P687A					
300	rs1391420451	D688N					
301	rs1465463508	T689A					
302	rs1235560158	Q690E					
303	rs1159132449	Q690R					
304	rs756599532	D693E					
305	rs753047724	K697N					
306	rs971784374	A698T					
307	rs1386216646	Q702H					
308	rs1379699288	S705N					
309	rs756646293	R707W					
310	rs1360689553	R707Q					
311	rs778204772	A709T					

312	rs368671947	A709V					
313	rs779614400	R712W					
314	rs373969485	R712Q					
315	rs373969485	R712P					
316	rs373969485	R712L					
317	rs1437177704	L714V					
318	rs377539747	R716W					
319	rs554894381	R716Q					
320	rs769871370	R718W					
321	rs1162438120	R718Q					
322	rs773330335	D719E					
323	rs1457616326	E720K					
324	rs879254299	I721S					
325	rs759690535	E723K					
326	rs759690535	E723Q					
327	rs1331338616	E723G					
328	rs879254315	Q725P					
329	rs948053969	Q725H					
330	rs753083346	K726N					
331	rs756519830	T727N					
332	rs756519830	T727I					
333	rs757751684	I728T					
334	rs370904481	E730K					
335	rs751019108	E730G					
336	rs754614296	L731F					
337	rs1243609134	K732E					
338	rs754561927	L734V					
339	rs780535025	Q736E					
340	rs752462267	K737R					
341	rs755813386	Q739H					
342	rs1369177203	E743K					
343	rs754593021	K744N					
344	rs749023785	A747D					
345	rs1218452123	E750K					
346	rs1269224635	L752V					
347	rs778958309	K753R					
348	rs387907286	E755K					
349	rs1288959239	H757Y					
350	rs140281678	E758K					
351	rs1269746317	T761P					
352	rs1271325672	T761S					
353	rs768803434	K762R					
354	rs749616160	Q764K					
355	rs1160784315	E765K					
356	rs1359073475	E765V					

357	rs765493045	T767R					
358	rs765493045	T767I					
359	rs773433247	L769M					
360	rs201321066	Y770H					
361	rs766924686	E771K					
362	rs537001094	R772Q					
363	rs760213269	H773R					
364	rs1289214173	E774D					
365	rs1318769789	K777R					
366	rs1461718357	K781R					
367	rs1242302845	G782C					
368	rs1236823586	T786S					
369	rs1457728758	V787I					
370	rs781124099	R789W					
371	rs368469872	R789Q					
372	rs372633156	R798C					
373	rs940071825	R798H					
374	rs749574366	D804N					
375	rs775039195	T806M					
376	rs1402124748	T807I					
377	rs1276356855	R808Q					
378	rs1425429900	K811R					
379	rs772785144	S812G					
380	rs376501264	S812T					
381	rs766050696	E818K					
382	rs766216492	S820G					
383	rs751295549	G821R					
384	rs373328181	H824Y					
385	rs777903778	Q826H					
386	rs1363391519	K829N					
387	rs193920755	S831F					
388	rs754171273	F832V					
389	rs919116365	K842E					
390	rs953001626	K845T					
391	rs779426030	Q846R					
392	rs1003993907	L847M					
393	rs757582532	V848A					
394	rs765503816	R849H					
395	rs780457980	D853N					
396	rs755458436	K860Q					
397	rs899351451	R864Q					
398	rs1295436882	T868M					
399	rs748759066	A874G					
400	rs748759066	A874V					
401	rs1256355579	A878T					

402	rs1484153242	K883R					
403	rs1017745641	A886T					
404	rs1017745641	A886S					
405	rs1184691547	M887L					
406	rs774040770	M887I					
407	rs745502722	K888N					
408	rs767911747	R891H					
409	rs267603612	R892W					
410	rs1467847947	R892Q					
411	rs1192823802	Q894L					
412	rs375737596	R899H					
413	rs1198903062	A903T					
414	rs1337848884	A903V					
415	rs141590802	V904I					
416	rs758637368	R905C					
417	rs766671105	R905H					
418	rs766671105	R905L					
419	rs1217692160	K907M					
420	rs748632605	S908T					
421	rs267603613	S909L					
422	rs1333544617	G910S					
423	rs1476843909	G910D					
424	rs199953180	K911T					
425	rs199953180	K911R					
426	rs201098122	R912W					
427	rs771687270	H914Y					
428	rs1375277418	Q917K					
429	rs1419915559	I918V					
430	rs1252859027	P921T					
431	rs763041798	V922I					
432	rs1231571339	R923W					
433	rs770895618	R923Q					
434	rs1163611256	P928S					
435	rs752993142	P928Q					
436	rs1420407259	A929T					
437	rs1329021125	A929V					
438	rs999227495	S931P					
439	rs1284934647	T933S					
440	rs1381639610	N934K					
441	rs764612223	P935S					
442	rs1272664645	P935L					
443	rs754276707	Y936D					
444	rs779422769	R939W					
445	rs1188692048	R939Q					
446	rs1431696242	S940N					

447	rs201359295	Y946C					
448	rs150672943	T947A					
449	rs1478755471	T947S					
450	rs1478755471	T947I					
451	rs138919700	N948T					
452	rs138919700	N948S					
453	rs749151018	L950P					
454	rs566471141	Q952E					
455	rs1292729014	Y958H					
456	rs1342455145	L959Q					
457	rs746095110	A961T					
458	rs1220546540	P963T					
459	rs1268871610	P963H					
460	rs772241902	S965F					
461	rs775853828	T966S					
462	rs149433823	T966I					
463	rs1461615572	S967L					
464	rs764332673	D968H					
465	rs777259435	M969L					
466	rs1484472471	Y970C					
467	rs1260487453	F971Y					
468	rs769120712	N973D					
469	rs776798327	N973K					
470	rs1172218291	C975Y					
471	rs146406013	T976A					
472	rs139801016	T976I					
473	rs1159045681	S977N					
474	rs1159045681	S977T					
475	rs200008143	S978T					
476	rs774024961	S978R					
477	rs1156869738	G979E					
478	rs866481110	A980P					
479	rs759169064	A980G					
480	rs759169064	A980V					
481	rs1387856537	T981K					
482	rs1211468940	S982F					
483	rs1328438334	G984S					
484	rs371548640	G985S					
485	rs371548640	G985C					
486	rs113247976	P986R					
487	rs113247976	P986L					
488	rs1057520171	Q991K					
489	rs1259619958	Q991P					
490	rs757241534	N994H					
491	rs146202502	N997I					

492	rs1060502523	R1007K					
493	rs762906034	R1007S					
494	rs1349876246	D1009E					
495	rs751496558	P1011L					
496	rs781330396	G1013A					
497	rs931443243	Y1014C					
498	rs1048929484	A1016S					
499	rs1485170378	E1017G					
500	rs1186511989	K1021T					
501	rs1186511989	K1021R					
502	rs752920933	F1023L					
503	rs756348570	F1023C					
504	rs1441630174	P1024S					
505	rs1172190035	P1024H					
506	rs778286513	Q1027H					
507	rs894996728	E1028D					
508	rs1276293459	T1029A					
509	rs1206507768	A1030E					
510	rs749631031	A1031T					
511	rs749631031	A1031P					

Appendix Table 1: List of all missense SNPs in KIF5A and their prediction by *in-silico* tools

REFERENCES

- [1] M. Simone, A. Trabacca, E. Panzeri, L. Losito, A. Citterio, and M. T. Bassi, “KIF5A and ALS2 Variants in a Family With Hereditary Spastic Paraplegia and Amyotrophic Lateral Sclerosis,” *Front. Neurol.*, 2018.
- [2] L. S. B. Goldstein, “Kinesin molecular motors: Transport pathways, receptors, and human disease,” *Proc. Natl. Acad. Sci.*, 2002.
- [3] H. Miki, Y. Okada, and N. Hirokawa, “Analysis of the kinesin superfamily: Insights into structure and function,” *Trends in Cell Biology*. 2005.
- [4] E. Reid *et al.*, “A Kinesin Heavy Chain (KIF5A) Mutation in Hereditary Spastic Paraplegia (SPG10),” *Am. J. Hum. Genet.*, 2002.
- [5] C. Goizet *et al.*, “Complicated forms of autosomal dominant hereditary spastic paraplegia are frequent in SPG10,” *Hum. Mutat.*, 2009.
- [6] Y. T. Liu *et al.*, “Extended phenotypic spectrum of KIF5A mutations,” *Neurology*, 2014.
- [7] N. Hirokawa, Y. Noda, Y. Tanaka, and S. Niwa, “Kinesin superfamily motor proteins and intracellular transport,” *Nature Reviews Molecular Cell Biology*. 2009.
- [8] A. Nicolas *et al.*, “Genome-wide Analyses Identify KIF5A as a Novel ALS Gene,” *Neuron*, 2018.
- [9] O. Hardiman *et al.*, “Amyotrophic lateral sclerosis,” *Nat. Rev. Dis. Prim.*, vol. 3, p. 17071, Oct. 2017.
- [10] W. Van Rheenen *et al.*, “Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis,” *Nat. Genet.*, 2016.
- [11] R. Chia, A. Chiò, and B. J. Traynor, “Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications,” *The Lancet Neurology*. 2018.
- [12] B. N. Smith *et al.*, “Exome-wide rare variant analysis identifies TUBA4A mutations associated with familial ALS,” *Neuron*, 2014.
- [13] C. H. Wu *et al.*, “Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis,” *Nature*, 2012.
- [14] R. L. McLaughlin *et al.*, “Genetic correlation between amyotrophic lateral sclerosis and schizophrenia,” *Nat. Commun.*, 2017.
- [15] S. Roy, V. M. Y. Lee, and J. Q. Trojanowski, “Axonal Transport and Neurodegenerative Diseases,” *Encycl. Neurosci.*, vol. 1762, pp. 1199–1203, 2010.

- [16] B. Ebbing *et al.*, “Effect of spastic paraplegia mutations in KIF5A kinesin on transport activity,” *Hum. Mol. Genet.*, 2008.
- [17] P. Kalaiarasan, B. Kumar, R. Chopra, V. Gupta, N. Subbarao, and R. N. K. Bamezai, “In silico screening, genotyping, molecular dynamics simulation and activity studies of SNPs in Pyruvate Kinase M2,” *PLoS One*, 2015.
- [18] M.-T. Heemels, “Neurodegenerative diseases,” *Nature*, vol. 539, p. 179, Nov. 2016.
- [19] L. A. Becker *et al.*, “Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice,” *Nature*, 2017.
- [20] R. Hussain, H. Zubair, S. Pursell, and M. Shahab, “Neurodegenerative Diseases: Regenerative Mechanisms and Novel Therapeutic Approaches,” *Brain Sci.*, vol. 8, no. 9, p. 177, Sep. 2018.
- [21] J. Couthouis *et al.*, “A yeast functional screen predicts new candidate ALS disease genes,” *Proc. Natl. Acad. Sci. U. S. A.*, 2011.
- [22] R. M. Marton and S. P. Paşca, “Neural Differentiation in the Third Dimension: Generating a Human Midbrain,” *Cell Stem Cell*. 2016.
- [23] H. Y. *et al.*, “Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model,” *Genes Dev.*, 2010.
- [24] H. Ilieva, M. Polymenidou, and D. W. Cleveland, “Non-cell autonomous toxicity in neurodegenerative disorders: ALS and beyond,” *Journal of Cell Biology*. 2009.
- [25] R. G. Miller, J. D. Mitchell, and D. H. Moore, “Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND),” *Cochrane Database Syst. Rev.*, no. 3, 2012.
- [26] X. W. Su, J. R. Broach, J. R. Connor, G. S. Gerhard, and Z. Simmons, “Genetic heterogeneity of amyotrophic lateral sclerosis: Implications for clinical practice and research,” *Muscle and Nerve*, 2014.
- [27] M. De Carvalho, T. Matias, F. Coelho, T. Evangelista, A. Pinto, and M. L. Sales Luís, “Motor neuron disease presenting with respiratory failure,” in *Journal of the Neurological Sciences*, 1996.
- [28] R. Chen, F. Grand’Maison, M. J. Strong, D. A. Ramsay, and C. F. Bolton, “Motor neuron disease presenting as acute respiratory failure: A clinical and pathological study,” *J. Neurol. Neurosurg. Psychiatry*, 1996.
- [29] G. Logroscino and M. Piccininni, “Amyotrophic Lateral Sclerosis Descriptive Epidemiology: The Origin of Geographic Difference,” *Neuroepidemiology*, vol. 52, no. 1–2, pp. 93–103, 2019.

- [30] J. G. Nicholls and J. F. R. Paton, "Brainstem: neural networks vital for life," *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 364, no. 1529, pp. 2447–2451, Sep. 2009.
- [31] D. R. Rosen *et al.*, "Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis," *Nature*, vol. 362, no. 6415, pp. 59–62, 1993.
- [32] A. Acevedo-Arozena *et al.*, "A comprehensive assessment of the SOD1G93A low-copy transgenic mouse, which models human amyotrophic lateral sclerosis," *Dis. Model. Mech.*, vol. 4, no. 5, pp. 686–700, Sep. 2011.
- [33] E. Tokuda and Y. Furukawa, "Copper Homeostasis as a Therapeutic Target in Amyotrophic Lateral Sclerosis with SOD1 Mutations," *Int. J. Mol. Sci.*, vol. 17, no. 5, p. 636, Apr. 2016.
- [34] G. Bensimon, L. Lacomblez, and V. Meininger, "A Controlled Trial of Riluzole in Amyotrophic Lateral Sclerosis," *N. Engl. J. Med.*, vol. 330, no. 9, pp. 585–591, Mar. 1994.
- [35] H. Mitsumoto, B. R. Brooks, and V. Silani, "Clinical trials in amyotrophic lateral sclerosis: Why so many negative trials and how can trials be improved?," *The Lancet Neurology*. 2014.
- [36] A. DeLoach, M. Cozart, A. Kiaei, and M. Kiaei, "A retrospective review of the progress in amyotrophic lateral sclerosis drug discovery over the last decade and a look at the latest strategies," *Expert Opin. Drug Discov.*, 2015.
- [37] L. Fagerberg *et al.*, "Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics," *Mol. Cell. Proteomics*, 2013.
- [38] Y. Kanai, Y. Okada, Y. Tanaka, A. Harada, S. Terada, and N. Hirokawa, "KIF5C, a Novel Neuronal Kinesin Enriched in Motor Neurons," *J. Neurosci.*, 2018.
- [39] J. Duis *et al.*, "KIF5A mutations cause an infantile onset phenotype including severe myoclonus with evidence of mitochondrial dysfunction," *Ann. Neurol.*, 2016.
- [40] M. Rydzanicz *et al.*, "KIF5A de novo mutation associated with myoclonic seizures and neonatal onset progressive leukoencephalopathy," *Clin. Genet.*, 2017.
- [41] Y. Yang *et al.*, "The gene encoding alsin, a protein with three guanine-nucleotide exchange factor domains, is mutated in a form of recessive amyotrophic lateral sclerosis," *Nat. Genet.*, 2001.
- [42] I. Puls *et al.*, "Mutant dynactin in motor neuron disease," *Nat. Genet.*, 2003.
- [43] K. P. Kenna *et al.*, "NEK1 variants confer susceptibility to amyotrophic lateral sclerosis," *Nat. Genet.*, 2016.

- [44] S. Millecamps and J. P. Julien, "Axonal transport deficits and neurodegenerative diseases," *Nature Reviews Neuroscience*. 2013.
- [45] C. Vance *et al.*, "Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6," *Science* (80-.), 2009.
- [46] W. Guo *et al.*, "HDAC6 inhibition reverses axonal transport defects in motor neurons derived from FUS-ALS patients," *Nat. Commun.*, 2017.
- [47] H. J. Kim *et al.*, "Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS," *Nature*, 2013.
- [48] F. Matsuzaki, M. Shirane, M. Matsumoto, and K. I. Nakayama, "Protrudin serves as an adaptor molecule that connects KIF5 and its cargoes in vesicular transport during process formation," *Mol. Biol. Cell*, 2011.
- [49] A. L. Nishimura *et al.*, "A Mutation in the Vesicle-Trafficking Protein VAPB Causes Late-Onset Spinal Muscular Atrophy and Amyotrophic Lateral Sclerosis," *Am. J. Hum. Genet.*, 2004.
- [50] W. Y. Wang *et al.*, "Interaction of FUS and HDAC1 regulates DNA damage response and repair in neurons," *Nat. Neurosci.*, 2013.
- [51] C. H. Xia *et al.*, "Abnormal neurofilament transport caused by targeted disruption of neuronal kinesin heavy chain KIF5A," *J. Cell Biol.*, 2003.
- [52] K. N. Karle, D. Möckel, E. Reid, and L. Schöls, "Axonal transport deficit in a KIF5A -/- mouse model," *Neurogenetics*, 2012.
- [53] K. Nakajima, X. Yin, Y. Takei, D. H. Seog, N. Homma, and N. Hirokawa, "Molecular Motor KIF5A Is Essential for GABAA Receptor Transport, and KIF5A Deletion Causes Epilepsy," *Neuron*, 2012.
- [54] F. F. Heisler *et al.*, "GRIP1 interlinks N-cadherin and AMPA receptors at vesicles to promote combined cargo transport into dendrites," *Proc. Natl. Acad. Sci.*, 2014.
- [55] P. D. Campbell, K. Shen, M. R. Sapio, T. D. Glenn, W. S. Talbot, and F. L. Marlow, "Unique Function of Kinesin Kif5A in Localization of Mitochondria in Axons," *J. Neurosci.*, 2014.
- [56] K. Hares, J. Redondo, K. Kemp, C. Rice, N. Scolding, and A. Wilkins, "Axonal motor protein KIF5A and associated cargo deficits in multiple sclerosis lesional and normal-appearing white matter," *Neuropathol. Appl. Neurobiol.*, 2017.
- [57] N. L. Sim, P. Kumar, J. Hu, S. Henikoff, G. Schneider, and P. C. Ng, "SIFT web server: Predicting effects of amino acid substitutions on proteins," *Nucleic Acids Res.*, 2012.

- [58] H. Mi, S. Poudel, A. Muruganujan, J. T. Casagrande, and P. D. Thomas, "PANTHER version 10: Expanded protein families and functions, and analysis tools," *Nucleic Acids Res.*, 2016.
- [59] E. Capriotti, R. Calabrese, and R. Casadio, "Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information," *Bioinformatics*, 2006.
- [60] Y. Bromberg and B. Rost, "SNAP: Predict effect of non-synonymous polymorphisms on function," *Nucleic Acids Res.*, 2007.
- [61] E. Capriotti, P. Fariselli, and R. Casadio, "I-Mutant2.0: Predicting stability changes upon mutation from the protein sequence or structure," *Nucleic Acids Res.*, 2005.
- [62] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang, "The I-TASSER suite: Protein structure and function prediction," *Nature Methods*. 2014.
- [63] F. Carrascoza, S. Zaric, and R. Silaghi-Dumitrescu, "Computational study of protein secondary structure elements: Ramachandran plots revisited," *J. Mol. Graph. Model.*, 2014.
- [64] K. Gopalakrishnan, G. Sowmiya, S. S. Sheik, and K. Sekar, "Ramachandran plot on the web (2.0).," *Protein Pept. Lett.*, 2007.
- [65] R. A. Laskowski, J. Jabłońska, L. Pravda, R. S. Vařeková, and J. M. Thornton, "PDBsum: Structural summaries of PDB entries," *Protein Sci.*, 2018.
- [66] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton, "PROCHECK: a program to check the stereochemical quality of protein structures," *J. Appl. Crystallogr.*, 1993.
- [67] S. C. Lovell *et al.*, "Structure validation by C α geometry: phi,psi and C beta deviation," *Proteins-Structure Funct. Genet.*, 2003.
- [68] G. Vriend, "WHAT IF: a molecular modeling and drug design program.," *J. Mol. Graph.*, 1990.
- [69] E. Krieger *et al.*, "Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8," *Proteins: Structure, Function and Bioinformatics*. 2009.
- [70] J. Schymkowitz, J. Borg, F. Stricher, R. Nys, F. Rousseau, and L. Serrano, "The FoldX web server: An online force field," *Nucleic Acids Res.*, 2005.
- [71] B. Huang, "MetaPocket: A Meta Approach to Improve Protein Ligand Binding Site Prediction," *Omi. A J. Integr. Biol.*, 2009.
- [72] R. A. Friesner *et al.*, "Extra precision glide: Docking and scoring incorporating a

- model of hydrophobic enclosure for protein-ligand complexes,” *J. Med. Chem.*, 2006.
- [73] G. Jones, P. Willett, R. C. Glen, A. R. Leach, and R. Taylor, “Development and validation of a genetic algorithm for flexible docking,” *J. Mol. Biol.*, 1997.
- [74] G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, and W. Sherman, “Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments,” *J. Comput. Aided. Mol. Des.*, 2013.
- [75] H. Venselaar, T. A. H. te Beek, R. K. P. Kuipers, M. L. Hekkelman, and G. Vriend, “Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces,” *BMC Bioinformatics*, 2010.
- [76] N. Hirokawa, S. Niwa, and Y. Tanaka, “Molecular motors in neurons: Transport mechanisms and roles in brain function, development, and disease,” *Neuron*. 2010.
- [77] T. L. Schwarz, “Mitochondrial trafficking in neurons,” *Cold Spring Harb. Perspect. Med.*, 2013.
- [78] Z. H. Sheng and Q. Cai, “Mitochondrial transport in neurons: Impact on synaptic homeostasis and neurodegeneration,” *Nature Reviews Neuroscience*. 2012.
- [79] S. Kaji *et al.*, “Late-onset spastic paraplegia type 10 (SPG10) family presenting with bulbar symptoms and fasciculations mimicking amyotrophic lateral sclerosis,” *J. Neurol. Sci.*, 2016.
- [80] C. O. Guinto *et al.*, “A novel mutation in KIF5A in a Malian family with spastic paraplegia and sensory loss,” *Ann. Clin. Transl. Neurol.*, 2017.
- [81] A. Al-Chalabi *et al.*, “Deletions of the heavy neurofilament subunit tail in amyotrophic lateral sclerosis,” *Hum. Mol. Genet.*, 1999.
- [82] G. M. Palomo and G. Manfredi, “Exploring new pathways of neurodegeneration in ALS: The role of mitochondria quality control,” *Brain Research*. 2015.
- [83] E. F. Smith, P. J. Shaw, and K. J. De Vos, “The role of mitochondria in amyotrophic lateral sclerosis,” *Neuroscience Letters*, 2017.
- [84] D. Brenner *et al.*, “Hot-spot KIF5A mutations cause familial ALS,” *Brain*, 2018.