Probing the Role of Calcium Signaling in Huntington's Disease (HD)

By

Minahil Tariq

(Registration No: 00000401527)

Department of Sciences

School of Interdisciplinary Engineering & Sciences (SINES)

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(August 2024)

Probing the Role of Calcium Signaling in Huntington's Disease (HD)

By

Minahil Tariq

(Registration No: 00000401527)

A thesis submitted to the National University of Sciences & Technology, Islamabad,

in partial fulfillment of the requirements for the degree of

Master of Science in
Bioinformatics

Supervisor: Dr. Ishrat Jabeen

School of Interdisciplinary Engineering & Sciences (SINES)

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(August 2024)

THESIS ACCEPTANCE CERTIFICATE

Certified that the final copy of MS Thesis written by Miss. Minahil Tariq (Registration No. 00000401527), of SINES (School/College/Institute) has been vetted by the undersigned, found complete in all respects as per NUST Statutes/ Regulations/ Masters Policy, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of Masters degree. It is further certified that necessary amendments as pointed out by GEC members and foreign/ local evaluators of the scholar have also been incorporated in the said thesis.

Signature: _____

Name of Supervisor Dr. Ishrat Jabeen

Date: 12th Sept, 2024

Signature (HOD): _____

Dr Fouzia Malik
HoD Sciences
Professor
SINES NUST Sector H-12
Islamabad

Date: 13-9-2024

Signature (Dean/ Principal) _____

Date: 13/09/2024

AUTHOR'S DECLARATION

I hereby state that my MS thesis titled "**Probing the role of calcium signaling in Huntington's disease (HD)**" is my work and has not been submitted previously by me for taking any degree from the National University of Sciences and Technology, Islamabad, or anywhere else in the country/ world.

At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Name of Student: Minahil Tariq

Date: September, 2024

PLAGIARISM UNDERTAKING

I solemnly declare that the research work presented in the thesis titled "**Probing the role of calcium signaling in Huntington's disease (HD)**" is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero-tolerance policy of the HEC and the National University of Sciences and Technology (NUST), Islamabad towards plagiarism. Therefore, I as an author of the above-titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred to/cited.

I undertake that if I am found guilty of any formal plagiarism in the above-titled thesis even after the award of MS degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and NUST, Islamabad have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized thesis.

Student Signature: _____

Name: Minahil Tariq

DEDICATION

I dedicate this thesis to my exceptional parents, siblings, friends, and teachers whose unconditional love, support, and guidance led me to this world of accomplishment.

ACKNOWLEDGEMENTS

All praise is for **Almighty Allah**, the ultimate source of all knowledge. By His grace, I have reached this stage of knowledge with the ability to contribute something beneficial to His creation. I earnestly thank my supervisor, **Prof. Ishrat Jabeen**, for her keen interest, invaluable guidance, encouragement, and continuous support throughout my research journey. I am extremely grateful for her thought-provoking discussions, sound advice, and valuable suggestions. Her mentorship has enabled me to tackle problems more meaningfully and provided me with the resources to pursue my objectives diligently and sincerely.

I am also thankful to my GEC committee members, **Dr. Masood Ur Rehman Kayani** and **Dr. Ammar Mushtaq**, who guided me throughout my project and offered valuable feedback and suggestions to refine my thesis. Additionally, I acknowledge all the faculty members of SINES for their kind assistance at various phases of this research.

My gratitude extends to my colleagues in the **lab**. Special thanks to **Amna Tariq, Arisha Wasim** and **Muqaddas Chaudary**, for their continuous help and feedback at every stage of the research. I am also grateful to my friends for their unwavering support. Lastly, I am deeply thankful to my parents, **Shabana Tariq** and **Tariq Mehmood,** and siblings for their immense love and support throughout this journey.

Thank you everyone!

Table of Contents

ACKNOWLEDGEMENTS ..... **v**

LIST OF FIGURES ..... **ix**

LIST OF TABLES ..... **xi**

ABSTRACT ..... **xv**

CHAPTER 1: INTRODUCTION **17**

1.1 Huntington's Disease 17

1.1.1 Epidemiology of HD 18

1.1.2 Huntingtin Protein 19

1.1.3 Role of Calcium Signaling 21

1.1.4 Clinical Manifestations 24

1.2 Problem Statement 26

1.3 Objectives 26

CHAPTER 2: LITERATURE REVIEW **27**

2.1 Neuronal Dysfunction in Huntington's disease 27

2.2 Endoplasmic Reticulum Stress in HD 28

2.3 Mitochondrial Dysfunction in HD 29

2.4 Crosstalk between ER and Mitochondria 30

2.5 Gut Microbiome and HD patients 32

2.6 Calcium signaling and striatal MSNs 34

2.7 Diagnosis and therapeutic strategies 35

2.8 Challenges: 37

2.8.1 Genetic and Molecular Complexity 38

2.8.2 Clinical trial challenges: 38

2.8.3 Lack of Disease-Modifying Therapies: 38
2.8.4 Drug Delivery Challenges: 39
CHAPTER 3: MATERIALS AND METHODS 40
3.1 Construction of Biological Regulatory Network: 41
3.2 Dynamic Simulations of BRN: 41
3.3 Collection and Preparation of protein structure: 43
3.3.1 HAP40 Removal: 43
3.3.2 Energy Minimization: 44
3.4 Collection of Ligand Dataset: 44
3.5 Molecular Docking: 45
3.6 Molecular Dynamics Simulations: 46
3.6.1 Preparation of Simulation System: 47
3.6.2 Energy Minimization: 48
3.6.3 Equilibration: 49
3.6.4 Production MD: 50
3.6.5 Analysis: 50
CHAPTER 4: RESULTS 52
4.1 Dynamic Simulations of Biological Regulatory Network: 52
4.1.1 Dynamic simulations of BRN in normal state: 56
4.1.2 Dynamic simulations of BRN in perturbed state: 58
4.2 Molecular Docking: 61
4.3 Molecular Dynamics Simulations of Huntingtin Protein: 65
4.4 Molecular Dynamics Simulations of Protein-Ligand complexes: 69
4.4.1 Complex 1: 70
4.4.2 Complex 2: 73

4.4.3 Complex 3:..... 75

4.4.4 Complex 4:..... 79

CHAPTER 5: DISCUSSION 82

CONCLUSION 84

Future Recommendations 85

REFERENCES..... 86

APPENDIX A 94

LIST OF FIGURES

Page No.

FIGURE 1.1: THE RELATIONSHIP BETWEEN THE NUMBER OF CAG REPEATS AND THE POLYQ LENGTH. 19

FIGURE 1.2: CALCIUM DYSREGULATION PATHWAY IN HUNTINGTON'S DISEASE.. 23

FIGURE 2.1: THE DISRUPTIVE EFFECTS OF MUTANT HUNTINGTIN 28

FIGURE 2.2: MUTANT HUNTINGTIN (MHTT) DISRUPTS CALCIUM BETWEEN THE ER AND MITOCHONDRIA. 31

FIGURE 2.3: THIS SCHEMATIC DIAGRAM SHOWS THE GUT-BRAIN AXIS 33

FIGURE 3.1: OVERALL WORKFLOW OF THE METHODOLOGY. 40

FIGURE 3.2: CRYOEM STRUCTURE OF HUNTINGTIN WITH PDB ID 6X9O 44

FIGURE 3.3: C TERMINAL REGION OF HUNTINGTIN PROTEIN ALONG WITH BINDING REGION. 46

FIGURE 3.4: WORKFLOW OF MOLECULAR DYNAMICS SIMULATIONS. 47

FIGURE 4.1: KNOWLEDGE-DRIVEN HUNTINGTIN (HTT) BIOLOGICAL REGULATORY NETWORK (BRN) 53

FIGURE 4.2: OSCILLATORY DYNAMICS OF BRN SIGNALING IN NORMAL CELLS 57

FIGURE 4.3: OSCILLATORY DYNAMICS OF BRN SIGNALING IN PERTURBED CELLS. 59

FIGURE 4.4: RELATIONSHIP BETWEEN LOG PIC50 VALUES AND GOLD SCORES. 62

FIGURE 4.5: DOCKED COMPLEX OF HTT WITH PAROXETINE 63

FIGURE 4.6: DOCKED COMPLEX OF HTT WITH MANNITOL 64

FIGURE 4.7: DOCKED COMPLEX OF HTT WITH AN UNDEFINED LIGAND 64

FIGURE 4.8: DOCKED COMPLEX OF HTT WITH GABAPENTIN 65

FIGURE 4.9: ROOT MEAN SQUARE DEVIATION OF THE HUNTINGTIN PROTEIN BACKBONE 66

FIGURE 4.10: RMSF OF PROTEIN RESIDUES OVER A 100 NS SIMULATION..... 68

FIGURE 4.11: RMSD GRAPH INDICATING MD SIMULATIONS OF PAROXETINE 71

FIGURE 4.12: RMS FLUCTUATION GRAPH OF PAROXETINE 72

FIGURE 4.13: RMSD GRAPH INDICATING MD SIMULATIONS OF MANNITOL..... 74

FIGURE 4.14: RMS FLUCTUATION GRAPH OF MANNITOL..... 75

FIGURE 4.15: RMSD GRAPH INDICATING MD SIMULATIONS OF UNDEFINED LIGAND 77

FIGURE 4.16: RMS FLUCTUATION GRAPH OF UNDEFINED LIGAND 78

FIGURE 4.17: RMSD GRAPH INDICATING MD SIMULATIONS OF GABAPENTIN..... 80

FIGURE 4.18: RMS FLUCTUATION GRAPH OF GABAPENTIN. 81

LIST OF TABLES

Page No.

TABLE 4.1: FUNCTIONS OF VARIOUS ENTITIES 54

TABLE 4.2: PARAMETER VALUES OF GENES INTERACTING WITH HTT 61

TABLE 4.3: LIGANDS SELECTED FOR MOLECULAR-DYNAMICS SIMULATIONS 69

Abbreviations:

HD = Huntington's Disease

HTT = Huntingtin Protein

mHTT = Mutant Huntingtin

HAP = Huntingtin Associated Protein

HIP = Huntingitn Interacting Protein

CryoEM = Cryogenic Electron Microscopic

PolyQ = Polyglutamine

BRN = Biological Regulatory Network

PI3K = Phosphatidylinositol-3-Kinase

EGF = Epidermal growth factor

IP3R = Inositol Triphosphate Receptor

BCL-xL = B-cell lymphoma-extra Large

BAX, BAK = Bcl-2-associated X protein, Bcl-2 antagonist killer

CASP3 = Caspase 3

MD = Molecular Dynamics

PDB = Protein Data Bank

RMSD = Root Mean Square Deviation

RMSF = Root Mean Square Fluctuation

Amino Acids one letter and 3 letter codes

Amino Acids	3-letter codes	1-letter codes
Alanine	Ala	A
Aspartic acid	Asp	D
Glutamic acid	Glu	E
Histidine	His	H
Isoleucine	Ile	I
Serine	Ser	S
Lysine	Lys	K
Proline	Pro	P
Leucine	Leu	L
Valine	Val	V
Threonine	Thr	T
Tryptophan	Trp	W

ABSTRACT

Huntington's disease (HD) is a neurodegenerative disorder, caused by the expansion of CAG repeats in the Huntingtin gene. Many recent studies have suggested that calcium dyshomeostasis is a one of the major causative factors in HD pathogenesis. It have been advocated that Mutant Huntingtin (mHTT) affects calcium signaling by sensitizing Inositol 1,4,5 triphosphate receptors (IP3R) to activation by Inositol 1,4,5 triphosphate (IP3). Additionally, Polyglutamine expansion in HTT leads to neuronal death of Medium Spiny Neurons (MSNs) in the striatum due to sensitization of IP3R. Although the genetic cause of the disease has been known for a long time, no effective therapeutic strategy has been developed so far due to complexity of the disease mechanistic and regulation. In this research, we constructed a Biological Regulatory Network (BRN) and performed calcium dyshomeostasis target validation of Huntingtin protein. Subsequently, molecular docking and molecular dynamics simulations were performed to develop potential therapeutic interventions against the identified drug targets for Huntington's disease. The Molecular Docking studies are employed on HTT protein with PDB ID 6X9O with binding site residues Asp2737, Glu2738, Asp2758, Lys2759, Glu3106, Glu3107, and Leu3108. Four potential ligands namely, Gabapentin, Mannitol, Paroxetine and an undefined ligand, were selected from the curated dataset of sixty-nine inhibitors, for further MD Simulations. The average RMSD values for these complexes are 0.58, 0.28, 0.29 and 0.29 respectively while the average RMSF values are 0.25, 0.43, 0.25 and 0.27 respectively. These findings revealed that these complexes maintained stable conformations with high flexibility at functional sites that suggests the potential therapeutic efficacy in modulation of the pathological aspects of HD. Among these four ligand complexes, the inhibitor Mannitol showed the highest RMSD value indicating the highest stability compared to the others. Our study was able to validate HTT as potential drug target and comprehend the

molecular mechanism of the disease which can assist in developing novel therapeutic strategies for the treatment of HD.

CHAPTER 1: INTRODUCTION

1.1 Huntington's Disease

Huntington's disease (HD), previously known as Huntington Chorea (HC), was described by George Huntington in 1872. The gene and mutation of this disease was identified by Huntington's disease Collaborative Research Group in 1993 [1]. Huntington's disease is an autosomal dominant neurodegenerative disorder that is typically caused by expanded CAG repeats which results in expanded polyglutamine stretch in the huntingtin protein (HTT). The age onset and severity of the disease occur depending upon the length of CAG repeats in the Huntingtin gene. The age at which symptoms first appear varies greatly, usually starting in early childhood and lasting until >80 years of age. Patients of all ages can have dysfunction of the nervous system which can result in progressive loss of motor control and difficulties with speech, swallowing, communication, etc. The genetic cause of the disease is known for years but there is no effective treatment available to stop the progression of the symptoms [2]. The genetic locus of huntingtin protein is located on the short arm of chromosome 4 and typically damages the nerve cells of the cerebral cortex and basal ganglia. HD is an inherited disease with a 50% chance of passing it to the next generation. In HD, the striatum of the brain is affected, which is basically the most vulnerable part of the brain [3]. Corticostriatal connections are usually damaged in the early stages of disease progression. These connections undergo changes in two stages. The first stage is presymptomatic stage in which there is an increase in glutamate release with the hyperexcitation of Sympathetic Preanglionic Neurons (SPN). In the next stage, the symptomatic stage, the silencing of SPNs occurs [4]. Increase in glutamate release causes persistent NMDAR activation that results in initiation of apoptotic processes in SPNs [5], [6]. On the other hand, SPNs silencing in later stage is due to loss of cortical

inputs [7]. Neuropathologically, HD is characterized by a general decline in brain size, neuronal dysfunction and striatal death where neuronal dysfunction occurs before the symptoms develop. Striatal Medium Spiny Neurons (MSNs) are vulnerable to this disease. These striatal neurons are accountable for carrying the motor information to basal ganglia. In Huntington's Disease, these neurons are preferentially vulnerable to degeneration leading to substantial motor and cognitive deficits [8].

1.1.1 Epidemiology of HD

There are a lot of reviews on the epidemiology of HD suggesting that the prevalence varies worldwide. Estimates on the prevalence of this disease state that there are tenfold or more variations among different parts of the world. A lot of research has been carried out on the incidence of Huntington's disease in Europe, America and Australia but there is very little knowledge on prevalence in Asia. According to an updated review on the prevalence of Huntington's disease in 2024, the prevalence falls between 0.5 and 1.5 per 100,000 people in Asia. Global estimates on HD stated that the prevalence ranges from 4.1 to 8.4 per 100,000 people and 1.63 to 9.95 per 100,000 people in USA and Europe respectively. On the other hand, Finland and Japan have less than 1 per 100,000 of population [9]. According to another epidemiological study, the prevalence of HD in Korea is 2.22 per 100,000 people. Although the prevalence in Korea is very different from Western countries, it is somewhat similar to Asian countries [10]. In 2019, the prevalence of HD in Chile was 0.72 per 100,000 of the population [11]. Across all the studies conducted between 1985 and 2022, the average prevalence was 3.92 per 100,000 people. As compared to Asia and Africa, the prevalence seems higher in Europe and North America.

1.1.2 Huntingtin Protein

The huntingtin protein is encoded by the huntingtin (HTT) gene, which is located on the short arm of chromosome 4. Huntingtin resides mainly in the cytoplasm, but can also be found in the nucleus. Huntingtin may play a significant role in numerous cellular activities happening in the cytoplasm and nucleus. Huntingtin is a large protein with 3142 amino acids and has no sequence similarity with any other protein [12], [13]. HTT possesses a spherical alpha-helical solenoid shape and its structure consists of five different domains [14]. These domains are the N terminal domain, bridge, C terminal domain, HEAT repeats and proline-rich segment. The most important domains are the N and C terminal domains with multiple HEAT repeats. The bridge domain is smaller than other domains and helps in connecting N and C terminals. This domain contains several types of tandem repeats and plays a very important role in stabilizing the HTT structure. In normal person, polyQ length is $\leq$ 35 while in adult or juvenile HD, the length varies from 40 to 120 repeats approximately as described in **Figure 1.1.**

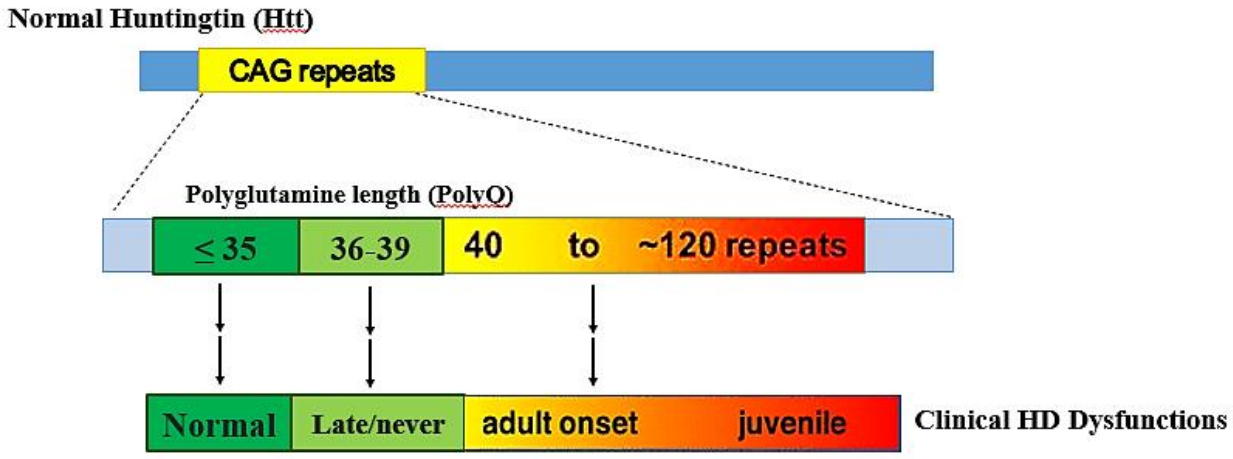


Figure 1.1: The relationship between the number of CAG repeats in the HTT gene and the PolyQ length, with its corresponding impact on the clinical onset of Huntington's disease (HD).

The molecular weight of this protein is approximately 350kDa. The highest levels of HTT are found in neurons of the Central Nervous System (CNS) and testes, expressed mostly in humans and rodents [15]. HTT protein is involved in various biological processes such as apoptosis, proliferation, autophagy, endocytosis, transcriptional regulation and vesicular transport [16]. Among all these functions, the role of huntingtin in transcription and intracellular transport is very significant. The transcriptional factors are transported by huntingtin between the nucleus and cytoplasm for the regulation of transcription [17]. Moreover, Huntingtin directly interacts with these transcriptional factors, spliceosome-related proteins and many other proteins that help in regulation of intracellular transport or endocytosis. These proteins include Huntingtin-associated protein 1 (HAP1), huntingtin-interacting protein 1 and 14 (HIP1 and HIP14), HIP1-related protein (HIP1R) and protein kinase C [18], [19], [20]. Normal huntingtin possesses anti-apoptotic function and prevents neurons from excitotoxicity in vivo [21]. Increasing wild-type huntingtin expression promotes survival and BDNF production, whereas its depletion results in some of the symptoms observed in Huntington's disease (HD) animals. This shows that decreased huntingtin activity may play a role in HD, and wild-type huntingtin may act as a modulator of HD pathophysiology [22]. The expanded PolyQ region results in misfolding and aggregation of HTT which can have harmful effects on neurons. As Huntingtin is located in cytoplasm, the initial events that may lead to HD take place in the cytoplasm. In the early stages of Huntington’s disease, the presence of mutant huntingtin interferes with the normal functioning of mitochondria and the homeostasis of calcium in cells. This disruption hinders the mitochondria’s capacity to regulate calcium levels triggering enzymes such as calpain and increasing the p53 signaling process. Moreover, decreased ATP production due to mitochondrial dysfunction results in cytochrome C release and caspase activation ultimately causing cell death [23]. The abnormal potentiation of N-methyl-D-aspartate

(NMDA) receptors and sensitization of inositol 1,4,5-trisphosphate receptors on the endoplasmic reticulum may contribute to the dysregulation of calcium in medium spiny neurons. These disruptions in calcium regulation are likely pivotal in the development of HD and could account for the oxidative stress, excitotoxicity and energy metabolism dysfunction seen in HD patients [24], [25].

Some of the identified genes, other than the huntingtin gene, that are involved in the development of Huntington's disease include adaptor-related protein complex 1, sigma 1 subunit (*Ap1s1*), and Cd4 Antigen (*Cd4*). The expressions of these genes are significantly downregulated in Huntington's disease [26]. Although the genetic cause of the disease has been known for a long time, no effective treatment to stop the progression of symptoms is available so far due to the complexity of disease mechanistic. To develop an effective therapy, the involvement of both wild-type HTT and mHTT in gene silencing studies has been explored. Depending on the length of the polyQ expansion, mHTT molecules form poisonous aggregates in the central nervous system. For example, mHTT co-aggregates with other proteins that serve critical roles in the cell, resulting in dysfunctional phenotypes. Several researches have been conducted to investigate the effects of HTT gene knockouts and knockdowns on cellular function [27].

1.1.3 Role of Calcium Signaling

Calcium signaling controls a wide range of cellular functions, including neuronal excitability, muscle contraction, cell division, proliferation, metabolism, bio-energetics, autophagy, and apoptosis. As a second messenger, calcium is responsible for transmitting signals between the plasma membrane and intracellular machinery. Since calcium is a messenger that affects multiple signal transduction pathways important for cells, their perturbation might eventually lead to cell death [28]. Intracellular calcium concentration is maintained by various calcium pumps and

calcium release channels. Inositol 1,4,5 triphosphate receptors (IP3Rs) are calcium release channels that are mainly required for maintaining ER-Mitochondria contacts. IP3Rs play a major role in the pathogenesis of Huntington's disease. Polyglutamine expansion in HTT leads to neuronal death of MSNs in the striatum due to sensitization of IP3R. Glutamate released from corticostriatal projection neurons stimulates NMDAR and mGluR5 receptors. Activation of NMDAR leads to calcium influx into the cell. Similarly, activation of mGluR5 leads to IP3R1 production and calcium release via IP3. Mutant Huntingtin (mHTT) affects calcium signaling by sensitizing IP3R1 to activation by IP3. This sensitized IP3R1 stimulates the activity of NMDAR. Stimulation of these glutamate receptors results in calcium overload in the cytosol. Excessive calcium is taken into mitochondria via the activity of Mitochondrial Calcium Uniporter (MCU). Consequently, the calcium storage capacity of mitochondria is exceeded leading to MPTP apoptosis. The whole process is shown in **Figure 1..**

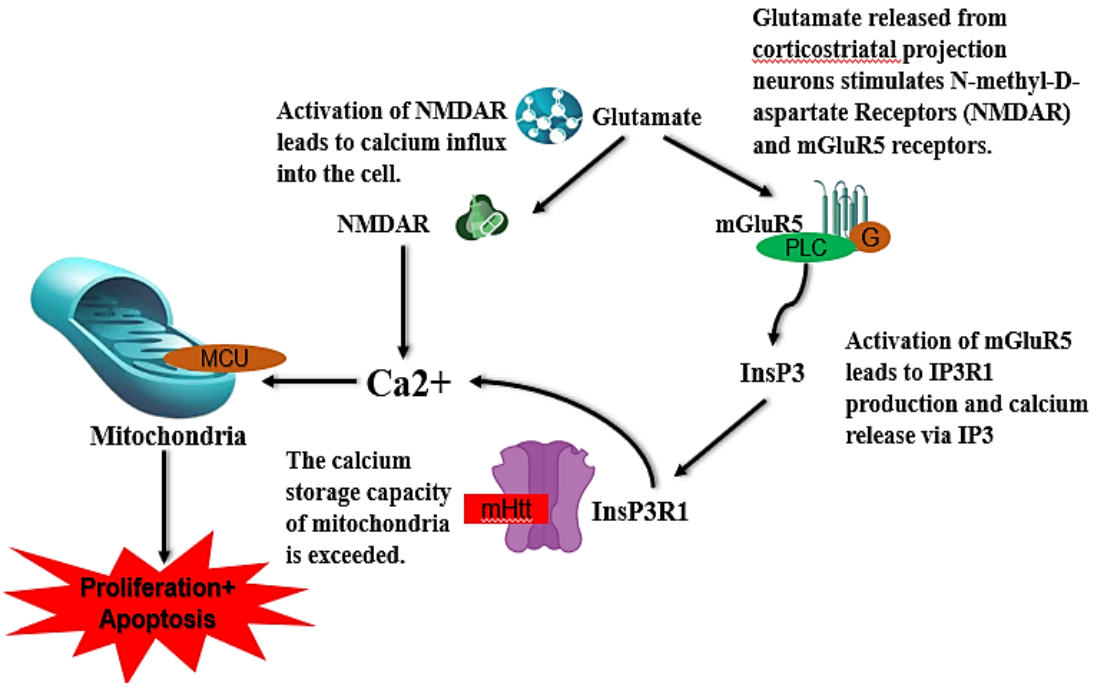


Figure 1.2: Calcium dysregulation pathway in Huntington's Disease. This figure depicts the pathway of calcium influx and signaling in neurons affected by mutant Huntingtin. It shows how glutamate activates NMDAR and mGluR5 receptors, resulting in elevated calcium levels through IP3R1, exceeding mitochondrial storage capacity and promoting cellular events such as proliferation and apoptosis.

mHTT also affects Voltage-gated Calcium Channels (VGCC) and Sarco-endoplasmic Reticulum Calcium Pump (SERCA) causing calcium dyshomeostasis [29]. The activity of the store-operated calcium entry is increased in neuronal cells. TRPC1 channels also play an important role in the SOC pathway in HD neurons. TRPC1-mediated SOC pathway is determined to be a novel therapeutic target for HD and other neurodegenerative illnesses. Moreover, Cadmium is a pathogenic neuromodulator that can induce neurotoxicity and apoptosis in striatal HD cells by dysregulating mitochondrial health and protein degradation pathways [30]. The presence of mutant huntingtin leads to significant alterations in calcium signaling pathways that results in neuronal dysfunction and death. The mechanisms include increased activity of voltage-gated calcium

channels (VGCCs) which cause excessive calcium entry into neurons contributing to excitotoxicity. The interaction of mHTT with IP3Rs and ryanodine receptors (RyRs) enhances their sensitivity that leads to excessive calcium release from the endoplasmic reticulum (ER), depleting ER calcium stores, and triggering stress responses. Additionally, the activity of calcium pumps and exchangers such as plasma membrane Ca^{2+}-ATPase (PMCA) and the sodium/calcium exchanger (NCX) is altered in HD that disrupts the cell's ability to release calcium and leading to sustained elevated intracellular calcium levels. The consequences of calcium dyshomeostasis in HD neurons include excitotoxicity where elevated calcium levels activate various calcium-dependent enzymes leading to the degradation of essential cellular components and neuronal death [31].

1.1.4 Clinical Manifestations

HD, being a devastating neurodegenerative disease, often manifests itself as a triad of motor, cognitive, and psychiatric symptoms. Early symptoms may include mild personality changes, mood swings, impatience, and cognitive difficulties. Moreover, in the early stages, intellectual functions may deteriorate slowly. These symptoms further proceed to motor symptoms which usually entail chorea, a condition characterized by involuntary, irregular, and unpredictable movements. Chorea results in the alteration of muscle control actions, often leading to difficulties in performing daily activities such as walking, talking, and swallowing. This results in the alteration of the muscle control actions. Psychiatric symptoms can be severe which mainly include depression, anxiety, irritability, apathy, and psychosis. Patients frequently experience hypokinesia, which is an impairment in voluntary movement, as well as rigidity, characterized by muscle stiffness and inflexibility. Dystonia, which includes abnormal, persistent muscle contractions that result in repetitive movements or abnormal postures, can also occur. In the final stages, secondary

illnesses such as pneumonia, malnutrition and skin ulcers, etc. are the main cause of death of HD patients rather than the disease itself. The risk of aspiration pneumonia grows as swallowing ability deteriorates, whereas malnutrition is caused by difficulties in eating and maintaining an adequate nutritional intake. Skin ulcers may also develop as a result of prolonged immobility and poor skin care, increasing the challenges that patients face in the late stages of the disease. The major hallmark of HD is cognitive decline with an early decline in short-term memory, further progressing to impairments in verbal fluency, attention and abstract reasoning [32]. HD has a significant influence on individuals and families, affecting not only the patient's physical and mental health but also their social and financial well-being. Due to the progressive nature of the disease, families often face increasing caregiving responsibilities and financial hardships over time. Carers often experience high levels of stress, anxiety, and depression as well [33]. Huntington's disease (HD) is a dominantly inherited progressive neurological disease characterized by chorea, an involuntary brief movement that tends to flow between body regions. HD is typically diagnosed based on clinical findings in the setting of a family history and may be confirmed with genetic testing. Predictive testing is available to family members at risk, but only experienced clinicians should perform the counseling and testing. Multiple areas of the brain degenerate, mainly involving the neurotransmitters dopamine, glutamate, and γ-aminobutyric acid. Although pharmacotherapies theoretically target these neurotransmitters, few well-conducted trials for symptomatic interventions have yielded positive results and current treatments have focused on the motor aspects of HD. Tetrabenazine is a dopamine-depleting agent that may be one of the more effective agents for reducing chorea, although it has a risk of potentially serious adverse effects. Some newer neuroleptic agents, such as olanzapine and aripiprazole, may have adequate efficacy with a more favorable adverse effect profile than older neuroleptic agents for

treating chorea and psychosis. There are no current treatments to change the course of HD, but education and symptomatic therapies can be effective tools for clinicians to use with patients and families affected by HD.

1.2 Problem Statement

The dysregulation of calcium signaling pathways is a pivotal contributor to the pathogenesis of Huntington's disease (HD), yet specific molecular targets within these pathways remain elusive. That’s why identifying key molecular targets involved in calcium signaling in HD and subsequently developing targeted therapeutic interventions to modulate these pathways is necessary for the treatment of Huntington’s disease.

1.3 Objectives

The main objective of this research is:

1. To probe the role of Calcium signaling in Huntington’s disease (HD) using biological regulatory network analysis.
2. Integrated molecular modeling strategies for the design and screening of the potential hits against Huntington’s disease (HD).

CHAPTER 2: LITERATURE REVIEW

2.1 Neuronal Dysfunction in Huntington's disease

HD is caused by selective degeneration of Medium Spiny Neurons (MSNs) in the striatum, particularly in the caudate nucleus and putamen [34]. The neuronal dysfunction in HD is caused by a variety of factors. The primary cause of this degeneration is toxic function of mHTT which is mainly due to abnormal expansion of polyQ tract. Aggregation of mHTT in neuronal nuclei leads to the neuronal dysfunction that ultimately cause cell degeneration and death [35]. Calcium signaling is another factor that plays a crucial role in neuronal function and its dysregulation plays a major role in the pathology of HD. Due to CAG expansion in HTT, calcium homeostasis gets disrupted that leads to excessive release of calcium from the Endoplasmic Reticulum and Mitochondria causing apoptosis [36]. **Figure 2.1** shows mHTT disrupts the activity of several transcription factors, resulting in altered gene expression profiles in vulnerable neurons. For example, mHTT inhibits the action of REST (RE1-Silencing Transcription factor), causing the derepression of neuronal genes that contribute to cell death [37].

Glutamate is the primary neurotransmitter in the brain which is essential for synaptic plasticity and transmission. Dysregulation of glutamate signaling contributes significantly to the synaptic dysfunction. NMDA and AMPA receptors are two major types of glutamate receptors involved in these processes. Excitotoxicity, caused by the over activation of NMDA receptors is another important mechanism in neuronal degeneration. mHTT increases the activation of NMDA receptors, causing excessive Ca^{2+} influx and neuronal excitotoxicity. This leads to the selective susceptibility of MSNs, which is highly sensitive to glutamatergic input [38]. Similarly, when

mHTT interacts with AMPARs, the synaptic transmission and plasticity is affected greatly which results in complete synaptic loss [39].

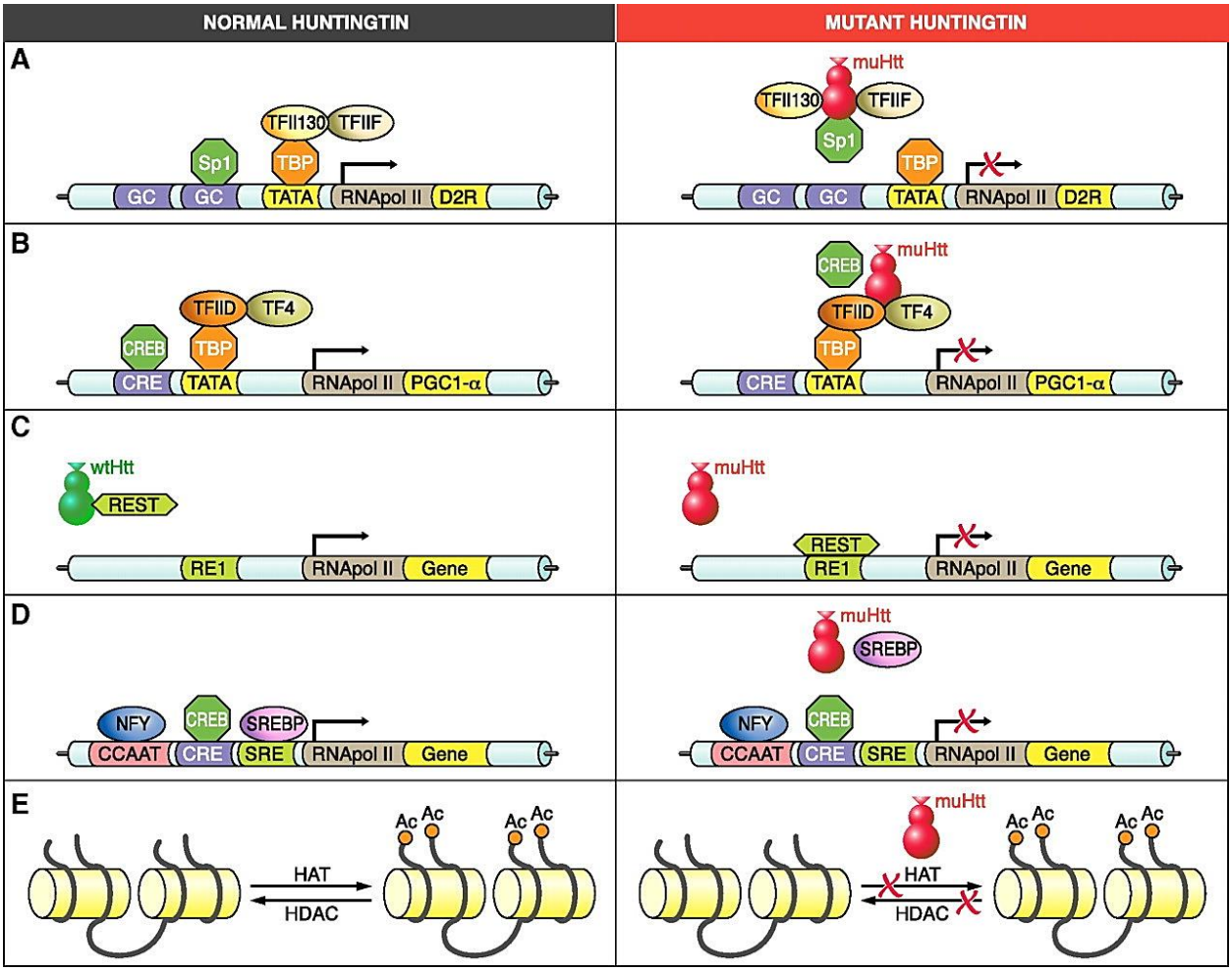


Figure 2.1: The disruptive effects of mutant huntingtin on genes essential for brain function and metabolism by highlighting its intricate interactions with transcriptional regulators in Huntington's disease. Prominent pathways include perturbed histone acetylation affecting overall gene expression, impaired SREBP processing affecting cholesterol production, altered REST/NRSF site impacting BDNF transcription, and interference with Sp1 and CREB activity [37].

2.2 Endoplasmic Reticulum Stress in HD

Endoplasmic Reticulum is a dynamic organelle that plays its role in protein synthesis, lipid synthesis and calcium storage. Disruptions in ER function cause ER stress, which induces the

activation of unfolded protein response (UPR) of ER to restore homeostasis. The UPR is mediated by three key sensor proteins: IRE1, PERK, and ATF6. These mechanisms assist reduce ER stress by improving protein folding, degrading misfolded proteins, and lowering the overall protein synthesis. In HD, these pathways get disrupted leading to neuronal death. Studies have revealed that the activation of Inositol-requiring enzyme 1 (IRE1) causes the formation of a TRAF2-ASK1 complex, which promotes c-Jun N-terminal kinase (JNK) signaling and induces cell death in neurons expressing mHTT. PERK inhibition in striatal neurons expressing mHTT enhances cell survival, suggesting that regulation of PERK activity could be a therapeutic target. MK-28, a Protein Kinase R like Kinase (PERK) activator, has demonstrated promising outcomes in HD models by enhancing motor and cognitive abilities and slowing the progression of the disease. Impaired Activating transcription factor 6 (ATF6) processing in HD causes decreased expression of protective genes, which contributes to neuronal vulnerability. Studies suggest that increasing ATF6 activity can improve cellular susceptibility to ER stress [40].

2.3 Mitochondrial Dysfunction in HD

Mitochondrial dysfunction is very crucial in the pathogenesis of Huntingtin's disease as it plays a major role in regulating intracellular calcium homeostasis and signaling pathways. In HD, mHTT disrupts mitochondrial bioenergetics by impairing oxidative phosphorylation and reducing ATP production. This energy deficit is affected by the Calcium overload in mitochondria, which further results in impairment of mitochondrial function [41]. mHTT directly interacts with the outer membrane of mitochondria i.e. MOM that may trigger calcium release in HD patients. When mHTT interacts with MOM, it induces the opening of permeability transition pore (PTP) of mitochondria which may triggered by calcium resulting in apoptosis. Moreover, recent research suggests that N-terminal fragments of mHTT interact with the inner membrane of

mitochondria, TIM23, leading to suppression of protein import and neuronal death. Neuronal cells majorly depend on the ATP production of mitochondria for the maintenance of synaptic functions. mHTT impairs mitochondrial function, leading to reduced ATP production and increased oxidative stress through N methyl D aspartate receptors. Hence, mitochondrial dysfunction worsens neuronal vulnerability and contributes to the progressive degeneration observed in HD [42].

2.4 Crosstalk between ER and Mitochondria

The physical contact and functional interactions between the ER and mitochondria, particularly at the mitochondria-associated ER membranes (MAMs), are important for the proper maintenance of cellular homeostasis. MAMs regulate calcium signaling, lipid metabolism, and mitochondrial dynamics. In HD, disrupted ER-mitochondrial communication exacerbates mitochondrial dysfunction, leading to increased oxidative stress and neuronal death [43]. The proteins that are involved in ER and Mitochondrial contact and calcium homeostasis are Voltage-Dependent Anion Channel 1 (VDAC1), Glucose-Regulated Protein 75 (Grp75) and Inositol 1,4,5-trisphosphate Receptor (IP3R) as depicted in **Figure 2.2.** VDAC1 is located in the outer membrane of mitochondria and helps in the transportation of calcium between ER and mitochondria. Grp75 acts as a chaperone, bridging IP3R on the ER membrane with VDAC1 on the mitochondrial membrane, forming a complex that helps in the regulation of Ca^{2+} flux [44]. In HD, when the levels of VDAC1 and Grp75 are increased, the calcium transport from ER to Mitochondria is enhanced to counteract the over activation of NMDA receptors [45]. Defective mitochondrial dynamics and protein degradation pathways play a very important roles in cadmium-induced neurotoxicity and cell death in striatal cells of HD patients. The presence of mutant huntingtin affects the vulnerability of neurons to environmental toxins, such as cadmium. Mitochondrial dynamics, which include

mitochondrial fission, fusion, and transport, are crucial for maintaining the mitochondrial function and cellular health. In HD striatal cells, disruption of these mitochondrial dynamics occurs due to cadmium exposure. This disruption is characterized by increased mitochondrial fragmentation, reduced mitochondrial fusion, and impaired mitochondrial transport. Such alterations result in mitochondrial dysfunction, decreased mitochondrial membrane potential, reduced ATP production, and increased oxidative stress. These mitochondrial defects contribute significantly to neuronal cell death.

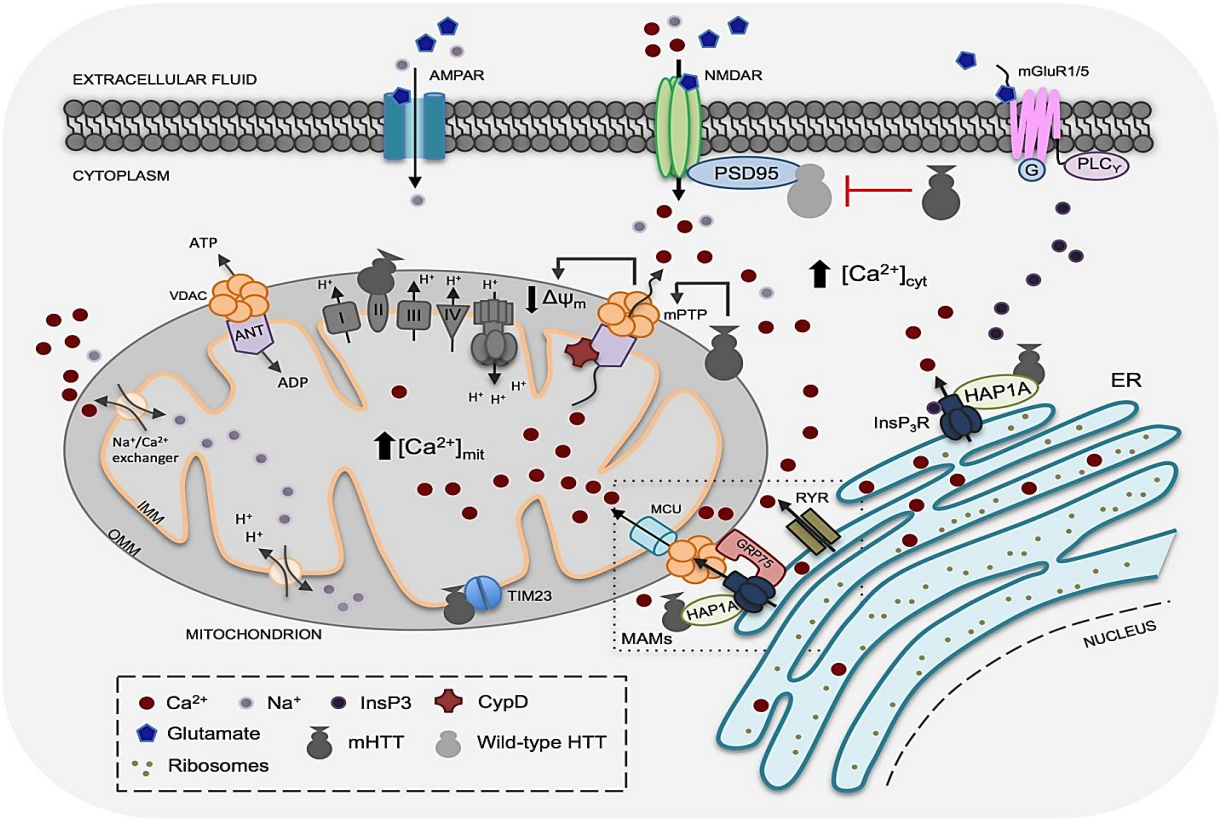


Figure 2.2: Mutant Huntingtin (mHTT) disrupts calcium buffering between the ER and mitochondria, causing calcium overload in mitochondria. This overload is exacerbated by the interaction of various proteins at ER-mitochondria contact sites, leading to mitochondrial dysfunction and cell damage [46].

Protein degradation pathways, specifically the ubiquitin-proteasome system (UPS) and autophagy-lysosome pathway, are also greatly affected in those striatal cells that are exposed to cadmium. The UPS is responsible for degrading misfolded proteins, while the autophagy-lysosome pathway is critical for degrading larger protein aggregates and damaged organelles. Cadmium exposure in HD cells leads to the accumulation of misfolded proteins and defective organelles due to the impaired function of these degradation pathways. This accumulation results in cellular stress and neurotoxicity. Additionally, cadmium exposure also induces ER stress and activates the UPRs in HD cells. Prolonged ER stress contributes to apoptotic cell death. The combined effects of defective mitochondrial dynamics, impaired protein degradation pathways, and ER stress lead to significant neurotoxicity and cell death in HD patients [47]. In short, disruptions in this crosstalk lead to mitochondrial dysfunction, oxidative stress, and ER stress, all of which contribute to neuronal death.

2.5 Gut Microbiome and HD patients

In HD, the composition of the gut microbiota is greatly affected because CNS receives and sends signals to the gut via neural pathways. Studies have shown that HD patients experience a significant decrease in beneficial bacteria and an increase in potentially harmful bacteria. These alterations may lead to an imbalance in gut microbial communities, dysbiosis, which negatively affects health and normal functioning of gut [48]. The gut microbiota plays a very important role in maintaining the metabolism of the body. The overall schematic diagram of gut brain axis is shown in **Figure 2.3.** The short chain fatty acids (SCFAs) like butyrate, propionate and acetate have anti-inflammatory properties and are very crucial for gut health but in HD patients, SCFAs production is reduced due to dysbiosis which leads to metabolic dysfunction and inflammation

[49]. The gut microbiota is involved in the synthesis and regulation of neurotransmitters such as serotonin, dopamine, and gamma-aminobutyric acid (GABA).

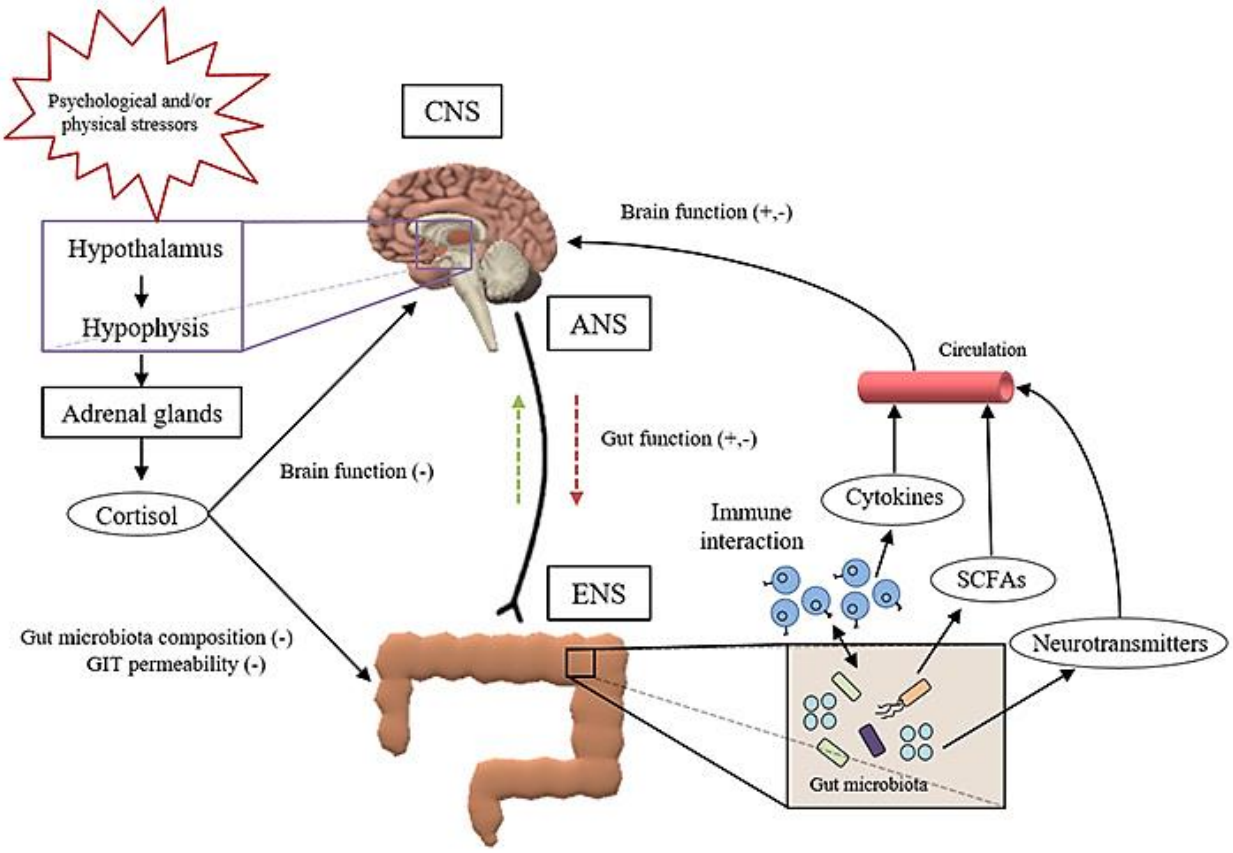


Figure 2.3: This schematic diagram shows the gut-brain axis and the feedback loops that are mediated by hormonal, neural, and immunological pathways connecting the gut microbiota, adrenal glands, and the central nervous system (CNS). The interactions involve the hypothalamus, hypothalamic-pituitary-adrenal (HPA) axis, neurotransmitters, and cytokines, highlighting their impact on brain and gut function [50].

IN HD, dysbiosis can negatively affect the synthesis and regulation of these neurotransmitters, resulting in imbalanced brain function and behavior. For example, reduced levels of beneficial bacteria that synthesize GABA can contribute to common psychiatric symptoms such as anxiety and depression [51]. Intestinal permeability, known as “leaky gut”, is also associated with neurodegenerative diseases. In this condition harmful substances such as lipopolysaccharides

(LPSs) translocate from gut into bloodstream. When the level of LPSs gets elevated in the blood, it may triggers neuroinflammation and neuronal damage in the brain [52].

2.6 Calcium signaling and striatal MSNs

Medium Spiny Neurons (MSNs), located in the critical region of the brain i.e. striatum, are involved in motor control and cognitive functions. In HD, MSNs are greatly affected and the cAMP-protein kinase A (PKA)-calcium signaling axis in striatal medium spiny neurons gets disrupted. The cAMP-PKA pathway is essential for numerous neuronal functions, such as neurotransmitter release, synaptic plasticity, and gene expression. Due to the presence of mutant huntingtin protein, there is reduced cAMP production and PKA activity, which leads to dysregulation of this pathway. Consequently, important proteins required for sustaining synaptic function and brain health become less phosphorylated causing abnormal neuronal function. Early changes in striatal activity occur before the onset of motor symptoms. These changes include abnormal neuronal activity, indicating early synaptic dysfunction and dysregulation of neural network function. Abnormal glutamate receptor signaling is primarily specific to MSNs. It is suggested in previous studies that transcription factor BCL11B, which is expressed in all MSNs, can cause neurodegeneration in Huntington's disease. MSNs that lack BCL11B showed abnormal responses to glutamate and failed to integrate dopaminergic and glutamatergic simulations. Also, gene enrichment analysis showed that disorder risk genes among BCL11B-regulated pathways, primarily relating to the cAMP-PKA-calcium signaling axis and synaptic signaling were overrepresented. Therefore, it is demonstrated that excitotoxicity and high calcium levels are the primary causes of HD cell death, with cortico-striatal glutamate sensitizing striatal cells to mutant huntingtin. Alternatively, mutant huntingtin makes striatal cells sensitive to glutamate subsequently [53]. The cAMP-PKA-Ca^{2+} signaling axis is not the only synaptic mechanism that is

being affected in HD. Some other processes include neurotransmitter release and synaptic plasticity. For instance, HD is associated with modifications to synaptic vesicle cycling, which involves neurotransmitter release and recycling. By interacting with proteins involved in vesicle fusion and release, mHTT inhibits synaptic vesicle release, which reduces neurotransmitter release and causes impairments in synaptic transmission. Moreover, mHTT decreases the quantity of active synaptic vesicles by affecting the presynaptic terminals. The disturbance of vesicle cycling is a contributing factor to the advancement of HD and synaptic dysfunction. These structural abnormalities correlate with cognitive deficits observed in HD patients, indicating the critical role of synaptic structure in maintaining normal cognitive function [54]. Another common feature of HD is altered synaptic connectivity and structural integrity. Synapses are located in dendritic spines, which show altered morphology and density over time. Research has demonstrated that altered spine morphology and reduced spine density result in reduction of synaptic connection. These structural abnormalities correlate with cognitive deficits observed in HD patients, indicating the critical role of synaptic structure in maintaining normal cognitive function. The contribution of striatal astrocytes in the pathophysiology of HD is also demonstrated in previous studies. It is revealed that HD treatment approaches that target only neurons are likely to fail. It is also concluded that strategies that involve mHTT lowering and astrocytes' function restoration can be more beneficial [55].

2.7 Diagnosis and therapeutic strategies

Huntington disease is devastating to patients and their families — with autosomal dominant inheritance, onset typically in the prime of adult life, progressive course, and a combination of motor, cognitive and behavioral features. The disease is caused by an expanded CAG trinucleotide repeat (of variable length) in HTT, the gene that encodes the protein huntingtin. In mutation

carriers, huntingtin is produced with abnormally long polyglutamine sequences that confer toxic gains of function and predispose the protein to fragmentation, resulting in neuronal dysfunction and death. We describe the relationship between CAG repeat length and clinical phenotype, as well as the concept of genetic modifiers of the disease. We discuss normal huntingtin protein function, evidence for the differential toxicity of mutant huntingtin variants, theories of huntingtin aggregation and the many different mechanisms of Huntington's disease pathogenesis. We describe the genetic and clinical diagnosis of the condition, its clinical assessment and the multidisciplinary management of symptoms, given the absence of effective disease-modifying therapies. We review past and present clinical trials and therapeutic strategies under investigation, including impending trials of targeted huntingtin-lowering drugs and the progress in development of biomarkers that will support the next generation of trials [56]. The accumulation of mutant huntingtin protein affects cellular functioning and causes neuronal cell death. Recent research investigations have also highlighted the significance of calcium signaling in HD, demonstrating that dysregulation of calcium homeostasis may contribute to neurodegeneration. Understanding these molecular and cellular mechanisms is crucial for developing targeted therapeutic strategies aimed at slowing disease progression and alleviating symptoms. Huntington's disease is diagnosed primarily through clinical evaluation, family history, and genetic testing. A thorough neurological examination is essential to identify the typical motor, cognitive, and psychiatric symptoms associated with HD. Genetic testing is the most definitive way for diagnosing HD because it can detect the presence of the mutant HTT gene. This test includes analyzing a blood sample to identify the number of CAG repeats in the gene, with a larger number of repeats indicates the existence of HD [56].

In addition to genetic testing, neuroimaging techniques such as magnetic resonance imaging (MRI) and computed tomography (CT) scans may be employed to assess brain structure and identify characteristic changes associated with HD. Functional imaging, such as positron emission tomography (PET) and functional MRI (fMRI) can provide additional insights into the functional abnormalities in the brain. These imaging techniques are valuable for both diagnosis and tracking disease progression [57]. Therapeutic strategies for HD aim to manage symptoms and improve the quality of life for patients. Currently, there is no cure for HD, and treatment focuses on reducing motor, cognitive, and psychiatric symptoms. Medications such as Ttrabenazine and Deutetrabenazine are typical medications used to treat chorea and other involuntary movements. These drugs deplete dopamine levels in the brain, thereby lowering the severity of motor symptoms [58]. Techniques such as antisense oligonucleotides (ASOs) and RNA interference (RNAi) aim to decrease the production of mHTT by targeting its mRNA. Early clinical trials have demonstrated some success, but delivery optimization and reducing off-target effects remain critical concerns [59]. Gene editing technologies like CRISPR/Cas9 offer the potential to correct the genetic mutation underlying HD. Before this strategy is widely implemented, issues with delivery, specificity, and ethical considerations need to be addressed. Another approach is to improve the delivery of neurotrophic factors which maintain the life and functionality of neurons. These include brain-derived neurotrophic factor (BDNF) and other growth factors that may lessen neuronal loss and improve symptoms [60]. Multidisciplinary care and supportive therapies are essential for managing symptoms and improving the quality of life for HD patients.

2.8 Challenges:

Huntington’s disease is a neurodegenerative disorder that has a significantly high prevalence in some areas. It can greatly affect the routine life of affected individuals. Research in this area can

be helpful in improving the quality of life of patients. HD is a genetic disorder and understanding the genetics of this disease is very important. Patients of all ages can have dysfunction of the nervous system which can result in progressive loss of motor control and difficulties with speech, swallowing, communication, etc. The genetic cause of the disease is known for years but there is no effective treatment available to stop the progression of the symptoms. Despite extensive research, effective disease-modifying treatments remain elusive, presenting significant challenges in managing HD.

2.8.1 Genetic and Molecular Complexity

The major challenge in treating HD stems from its genetic basis and the molecular complexity. The expanded CAG repeats cause HTT gene to become toxically active, which disrupts multiple biological processes including protein homeostasis, mitochondrial function, and synaptic transmission. The widespread impact of mHTT on various cellular pathways complicates the development of targeted therapies

2.8.2 Clinical trial challenges:

HD is a rare disease with a relatively small patient population that makes clinical trials challenging. Due to the progressive nature of HD, patients may experience different stages of disease progression, making the effectiveness of a treatment more difficult. Due to complexity of HD symptoms, it might be challenging to define sensitive and suitable outcome measurements. Given the genetic basis of HD, ethical concerns may arise regarding genetic testing [61].

2.8.3 Lack of Disease-Modifying Therapies:

Currently, no disease-modifying therapies have been approved for HD. The majority of treatments are symptomatic, that treats only the symptoms of the disease, such as cognitive deficits, psychological and motor abnormalities like chorea, and not the disease itself. These treatments do

not alter the disease onset and only provide limited relief. The development of effective therapies is needed to treat the HD pathology more effectively and permanently [62].

2.8.4 Drug Delivery Challenges:

One of the biggest challenges in treating HD is the efficient delivery of therapeutic agents to the brain. Many medications are less effective because of the blood-brain barrier (BBB), which restricts their entry. To overcome this barrier, experts are exploring novel drug delivery systems such as gene therapy vectors and nanoparticle carriers, but these approaches face technical and safety challenges [63]

CHAPTER 3: MATERIALS AND METHODS

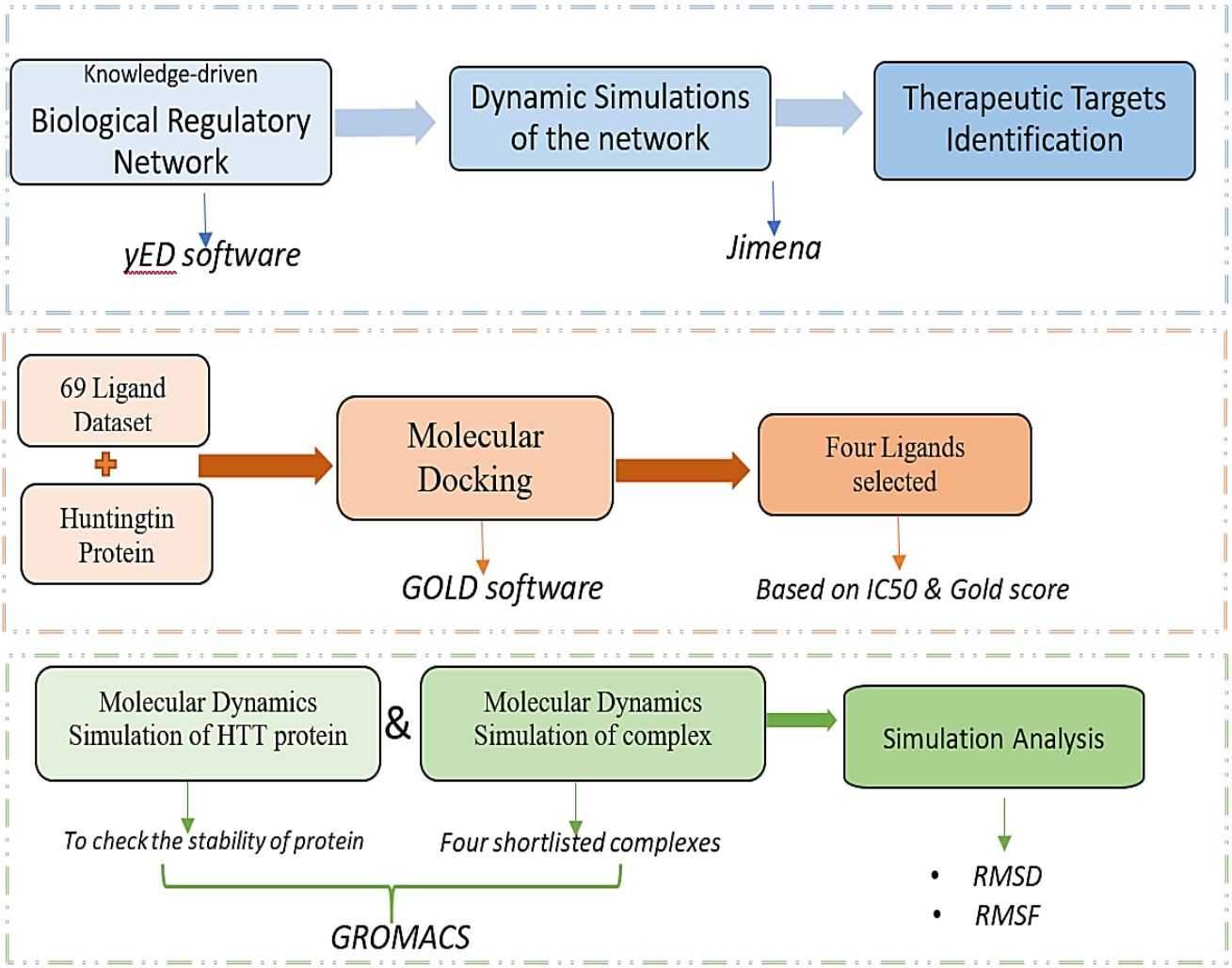


Figure 3.1: Overall Workflow of the methodology including construction of BRN, Molecular Docking and Molecular Dynamics Simulations.

3.1 Construction of Biological Regulatory Network:

To investigate the regulation relationships of the huntingtin protein which is implicated in the pathology of Huntington's disease, a knowledge-driven biological regulatory network (BRN) was constructed. A wide range of proteins namely, PI3K, AKT, EGF, IP3R, RAF1, BCL-XL, BAX, BAK. CASP3, p53 and Glutamate, were added in this network that either activate or inhibit HTT. These proteins were selected after a thorough analysis of the scientific literature and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database entries. This selection was made on the basis of the interactions of these proteins with huntingtin. In the BRN, proteins are represented as nodes and the type of interactions (either positive or negative) are represented as edges that connect these nodes. This network structure provides a clear understanding of how HTT interacts with different cellular pathways and its impacts on cellular functions [64]. An open-source GUI-based graphical software called yED software version 3.23.1 was used to graphically represent this biological regulatory network. This software facilitates an interactive exploration of the network, enhancing the understanding about the role of HTT in cellular mechanisms and identifying the potential therapeutic targets for Huntington's disease [65].

3.2 Dynamic Simulations of BRN:

A dynamic biological regulatory network (BRN) was used to analyze the functional implications of the huntingtin protein in different cellular events. Proteins were classified as either active or inactive using the network model which assigned a value of 1.0 for active states and 0.0 for inactive states. This binary approach helped in the exploration of how variations in HTT activity influence neuronal survival, synaptic function, and neurodegeneration. The network was dynamically simulated under three different cellular conditions: (a) normal cells, (b) cells expressing mutant

HTT and (c) in the presence of HTT inhibitors. A Java-based simulation framework known as Jimena was used for the dynamic simulations. Jimena is particularly designed to model genetic regulatory networks dynamically [66]. A Boolean approach was used to model the activation and inhibition states of the nodes in the network thus reducing complicated biological interactions to binary outcomes denoting the presence (1) or absence (0) of activity. This method allowed for clear, computable insights into how changes in HTT expression or function can results in cellular dysfunction.

The behavior of the network was controlled by ordinary differential equations (ODEs) that describe the rate at which protein activities vary over time, providing a dynamic view of cellular responses to changes in HTT levels [28]. These kinds of simulations are essential for comprehending the intricate biological relationships involved in neurological disorders and for developing therapies against the harmful effects of mutant proteins such as HTT. The SQUAD (Semi-Quantitative simulation of large biological regulatory networks) method was used to run simulations to thoroughly analyze the dynamic behavior of the BRN focused on HTT. This method initially treats the network as a discrete dynamical system and makes use of the binary decision diagram technique. SQUAD makes it easier to convert this discrete model into a continuous dynamical biological regulatory network which enables more complex simulations that demonstrate gradual changes in biological activity. The SQUAD parameters were adjusted to match the sensitivity and response dynamics of the biological components in the network with a steepness of 10 and a decay rate of 1.0 [67]. The simulations ran over a period of 500 time steps and step size (dt) of 0.05. After completion of simulations, the data was visualized in graphs form. These simulations and visualizations are essential for elucidating the roles of HTT and associated proteins in cellular processes. In order to reduce the harmful impacts of Huntington's disease,

potential therapeutic targets can be identified by understanding the dynamic interactions within this network.

3.3 Collection and Preparation of protein structure:

The structural analysis of the huntingtin protein is a crucial component especially for understanding its potential in the context of Huntington's disease. The protein structure, as shown in the **Figure 3.2**, was extracted from Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) database. For the purpose of this analysis, crystal structure of HTT with PDB ID 6X9O was used [68]. This structure was predicted in 2021 with resolution of 2.60 Å. This is a cryoEM structure of huntingtin in complex with HAP40 which provides comprehensive details about the spatial arrangement and conformational dynamics of HTT which are very important for understanding its function and interaction with other cellular components [69].

3.3.1 HAP40 Removal:

The initial step in preparing the HTT structure involved the removal of the HAP40 complex. Since current research is mainly focused on the HTT protein and its potential binding sites for therapeutic agents, it was essential to isolate HTT from HAP40. Therefore, we removed HAP40 from HTT protein. Ultimately, HTT protein is isolated from the complete structure which was used for further study and analysis.

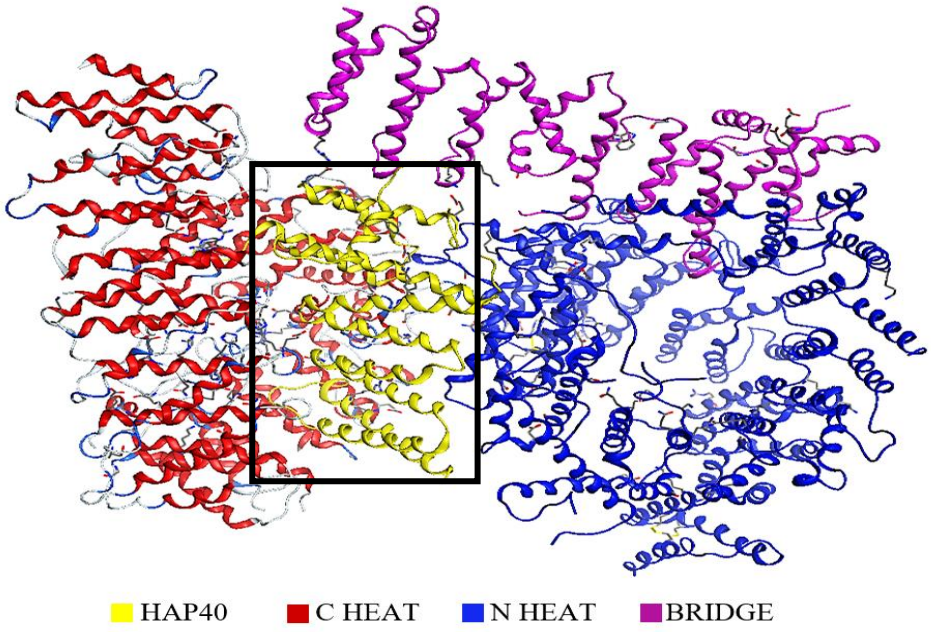


Figure 3.2: CryoEM structure of huntingtin in complex with HAP40 with PDB ID 6X9O. The highlighted region is HAP40 which is being removed from the protein.

3.3.2 Energy Minimization:

The next step involved energy minimization of this updated HTT structure. Energy minimization is crucial to remove any steric clashes or unfavorable geometries and stabilize the protein structure with energetically beneficial configuration, making the docking and simulations more accurate. The energy of this protein structure was minimized using Molecular Operating Environment (MOE) software and Amber99 force field was used [70].

3.4 Collection of Ligand Dataset:

A comprehensive dataset of sixty-nine HTT inhibitors, along with their pIC_{50} values, was collected from existing literature [71] and the ChEMBL database, which is a database of bioactive molecules with drug-like properties [72]. This dataset is give in **Table A.1**. 3D structures of these inhibitors were generated using the Open Babel tool, an open-source software designed to convert molecular data files [73]. This step was very important for further docking studies, that helps in accurate

prediction of the interactions between HTT and its inhibitors by providing precise spatial orientation and 3D confirmations.

3.5 Molecular Docking:

Only the C-terminal region of HTT was selected for docking to specifically target the interactions relevant to Huntington's disease. This region was selected due to its potential therapeutic binding sites. The binding region in C terminal of HTT is highlighted in **Figure 3.3.** The amino acid residues that are present in binding pocket includes Asp2737, Glu2738, Asp2758, Lys2759, Glu3106, Glu3107, and Leu3108 [71]. These residues were chosen because they were involved in ligand binding, according to previous studies. Molecular Docking was employed using Genetic Optimization for Ligand Docking (GOLD) software, for identification of optimal interactions between ligands and protein. The binding cavity was set to 45 Å to accommodate the molecular size of the ligands, with the origin for docking specified at coordinates 143.56, 228.48 and 182.69 for the x, y, and z axes respectively. Ten poses of each ligand were generated of which only the highest pose of each ligand was selected. Higher scores indicate a more favorable ligand-protein interaction. The GOLD score is a measure of the efficacy of docking. This step was crucial to determine the most promising candidates for potential inhibitors of HTT.

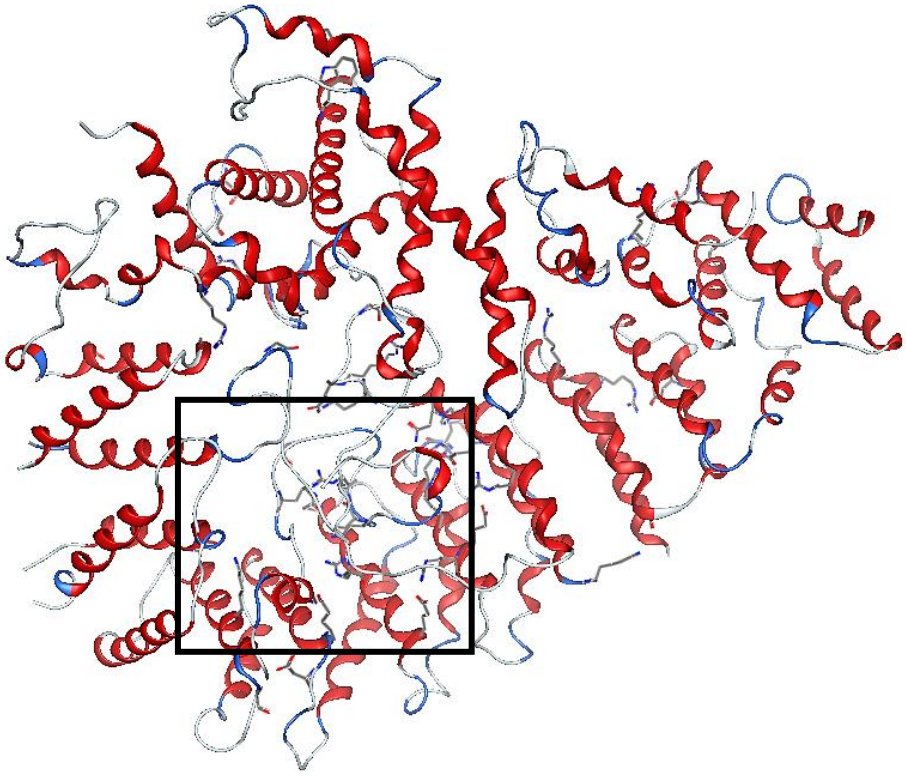


Figure 3.3: C terminal region of Huntingtin protein along with binding region for docking simulations.

3.6 Molecular Dynamics Simulations:

Molecular dynamics (MD) simulations provide insights into the protein's dynamics, its interaction with other molecules, and the effects of its mutation on cellular systems. Through understanding these interactions through MD simulations, we were able to proceed the targeted development of therapeutic interventions [74]. The procedure typically involves initializing the system with prepared structures followed by equilibration under controlled temperature and pressure conditions and then conducting production runs where the system's evolution is observed over time as depicted in **Figure 3.4**. To employ MD simulations, GROMACS, a software that is typically used for MD simulations of biological molecules, is used [75].

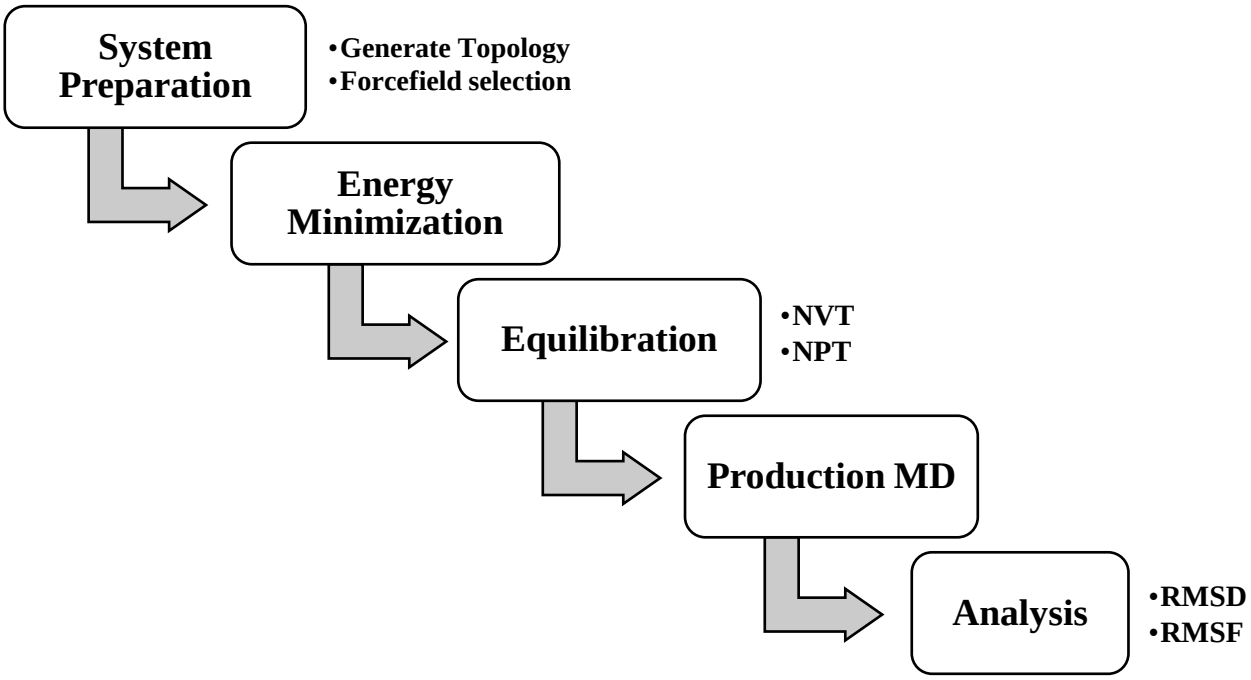


Figure 3.4: Workflow of Molecular Dynamics Simulations.

3.6.1 Preparation of Simulation System:

In molecular dynamics simulations, the main objective of system preparation was to build an accurate and physiologically precise model of the molecular system under study. In order to ensure correct representation of the protein and ligand structures for simulations, preparation of these structures was essential. The preparation included generating topology files, setting up solvent conditions, and ensuring that the system was neutral and energy-minimized. For an accurate simulation of molecular dynamics and interactions under realistic conditions, this setup was crucial.

AMBER03 force field was used to carefully set up MD simulations for the huntingtin protein. This force field is highly regarded for its accuracy in protein dynamics research and was chosen due to its remarkable capacity to capture intricate bio-molecular interactions [76]. Moreover, this force field is also used by other researchers for the simulations of HTT protein [77]. The TIP3P water

model which is well recognized for its realistic description of water molecules crucial to the stability and behavior of biomolecules within the cellular environment was used to accurately replicate physiological conditions. In order to neutralize the system, the simulation setup also involved the addition of counter ions, such as chloride. This step is essential as proteins like huntingtin may carry net charges due to their amino acid composition and maintaining electrical neutrality is essential for avoiding artifactual forces within the simulation [78]. Moreover, to avoid boundary effects that could impact the protein's interactions and dynamics, a cubic box with sufficient space between the protein and the box edges was added. This approach helped to ensure that interactions within the box were accurately represented without undue influence from periodic images of the protein. One of the most important steps in preparing the simulation system was solvating the protein once it was placed inside a cubic simulation box. This step was important because it adds water molecules surrounding the protein to create a physiologically relevant environment that closely resembles the biological settings in which the protein functions naturally. Water was included because it helps proteins fold and function correctly. It also affects the system's dynamic interactions which is important for accurate simulation results [79].

3.6.2 Energy Minimization:

In the energy minimization step of molecular dynamics simulations, the simulation system was stabilized through precise adjustment of atomic positions. The objective was to minimize the potential energy of the system by resolving any steric clashes and unfavorable interactions that might have occurred during system setup. This procedure was essential because it ensures that the simulation starts from a steady and realistic baseline, improving the precision and reliability of the outcomes. The methodology involved using GROMACS utilities: grompp compiled the system with the correct parameters and prepared it for minimization, while mdrun performed the

minimization itself. Without proper energy minimization, simulations might yield false results due to unstable starting conditions, that might affect the interpretation of bio-molecular functions and interactions [80].

3.6.3 Equilibration:

The equilibration phase of molecular dynamics simulations for a protein-ligand complex involved several crucial steps to ensure system stability under simulated physiological conditions. The equilibration phase helped in maintaining the temperature, pressure and density for simulations. During equilibration, position restraints were used to keep the ligand in its proper position and orientation. This helped in creating a position restraint topology specifically for the ligand in order to examine the interactions without significant displacement from the binding site. The system uses temperature coupling groups to maintain the equilibrium of temperature across different components of the system. An improvement over the Berendsen approach, the 'V-rescale' thermostat was used to maintain the proper kinetic ensemble [81]. Equilibration of MD simulations involved two critical steps: NVT equilibration which is basically canonical ensemble and NPT equilibration i.e. isothermal-isobaric ensemble. In NVT run, the system's temperature was stabilized. In order to ensure that the molecular velocities are balanced at the desired temperature, it was a crucial step. During this phase, the volume of the system remained constant while the number of particles and the temperature were controlled [82]. The NPT run was started after temperature equilibration in order to stabilize the system's pressure and density. To better mimic physiological conditions, this phase helped in adjusting the simulation box's volume based on the temperature and external pressure. During this stage, the Parrinello-Rahman barostat is usually employed to precisely control pressure [83]. These steps were very important for preparing the system for production runs, so that the simulated environment closely mimics real-world

conditions, and providing a stable platform for observing the true dynamics of the protein-ligand interaction.

3.6.4 Production MD:

The molecular dynamics simulation proceeded on to the production phase once the equilibration stages were completed. This stage was critical for gathering dynamic data under stabilized conditions. At a particular temperature and pressure, interactions within the system were enabled when the positional restraints previously applied to the ligand were released. The production run typically extended over a specified duration, set to 10 nanoseconds for initial studies, to collect adequate data for analysis. The execution began with preparing the system using the grompp tool from GROMACS which compiled the necessary input files using the latest simulation checkpoint that includes all relevant system states preserved from the equilibration phase. The coordinate file from the NPT phase (npt.gro), the checkpoint file (npt.cpt), and the MD parameter file (md.mdp) were all referenced in the command structure which ensured that all conditions (including pressure and temperature) are accurately carried over into the production run [84]. In addition to collecting trajectories, the production MD run observed how the system changes in simulated settings that offers insights into how biomolecules behave in conditions similar to physiological ones [85].

3.6.5 Analysis:

In the analysis phase of molecular dynamics simulations, especially with periodic boundary conditions, recentering and rewrapping of coordinates was required for maintaining the visual and analytical coherence of the system. When molecules move outside of the simulation box, they may appear fragmented or may even jump across it. The protein was recentered and every molecule was rewrapped using **trjconv** to maintain a compact structure which was crucial for accurate analysis. This step helped the molecules to appear intact and correctly positioned within the visual

field that prevented them from appearing broken or disjointed as a result of crossing the boundary of the simulation box. The trajectory was adjusted to maintain a consistent representation which was helpful in further analyses like RMSD (Root Mean Square Deviation) and RMSF (Root Mean Square Fluctuation) calculations. RMSD evaluates the structural stability by comparing the protein's position over time to its initial state while RMSF analyzes the flexibility of individual residues that provides insights into dynamic regions within the protein [86]. RMSD measurement was crucial for determining how much the protein's structure changed during the simulation from its initial configuration. We selected the "Backbone" atoms of the protein for this analysis because by focusing on the backbone we could get a clear picture of the overall movements and structural changes without having to deal with side-chain fluctuations. The flexibility of individual amino acid residues along the protein chain was analyzed using RMSF. By focusing on the "Backbone" atoms, flexible or dynamic parts of the protein were identified. Higher RMSF values are frequently correlated with functionally significant areas or ligand-binding regions [87].

CHAPTER 4: RESULTS

To unravel the intricate role of calcium signaling within the pathophysiology of Huntington's disease, approach of biological regulatory network analysis combined with advanced molecular modeling techniques was employed for the identification of potential therapeutic targets. The findings shown here demonstrate the effectiveness of combining molecular biology knowledge with computational biology methods. Integrated molecular modelling has made it possible to find and evaluate drugs that help to alter the progression of HD. These findings aids in the advancement of therapeutic approaches to reduce the disruptions in calcium signaling observed in HD, potentially leading to the management and treatment of this disease.

4.1 Dynamic Simulations of Biological Regulatory Network:

In the constructed knowledge-driven BRN, the HTT protein is illustrated as an essential component involved in regulating key cellular processes that intersect with the pathology of Huntington's disease. The complete architecture of this network is shown in the **Figure 4.1**. The simulation was conducted using the SQUAD method, over 50-time units with a time step (dt) of 0.01. The network indicates how HTT regulates calcium signaling, an essential mediator of cellular functions, via its interactions with glutamate and the IP_3R calcium channel. This modulation of calcium dynamics by HTT is critical, as disruptions in calcium homeostasis are a hallmark of Huntington's disease. The key cellular responses of this network are apoptosis and proliferation. The network also reveals the impact of HTT on the PI3K/Akt pathway and subsequently interacting with the apoptosis regulator p53, indicating HTT's involvement in pathways leading to cell survival. EGF activates its receptor EGFR, causing RAF1 to become active, which promotes cell proliferation. HTT interacts with glutamate to affect the IP3R-mediated calcium signaling,

essential for cellular responses. Elevated calcium levels, affected by HTT, can activate apoptotic pathways through BAX/BAK and Caspase-3, or promote survival via BCL-XL depending on the cellular environment. By analyzing these interactions within the BRN, we highlighted the complex role of HTT in both promoting and inhibiting pathways leading to cell proliferation and apoptosis as well as its potential impact on the progression of Huntington's disease.

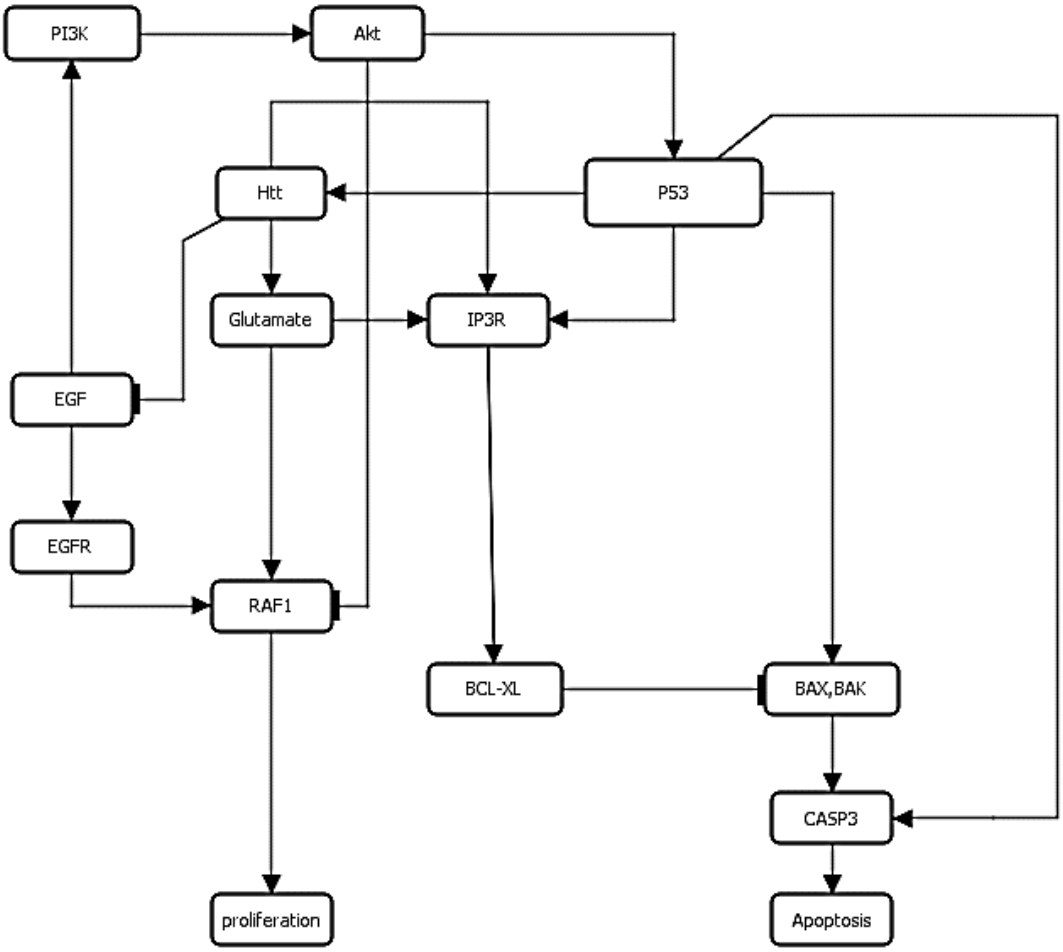


Figure 4.1: Knowledge-driven Huntingtin (HTT) biological regulatory network (BRN). Each node is represented as a box, with arrows indicating activation and lines ending in a flat line denoting inhibition. Key nodes include PI3K, Akt, HTT, glutamate, IP3R, p53, and CASP3. This figure illustrates pathways affecting proliferation and apoptosis, highlighting the complex role of HTT in cellular dynamics. EGF and EGFR are also included to show their role in activating downstream RAF1 leading to cell proliferation.

In this network, the HTT protein is assumed as a stable state (SS), serving a central role in mediating interactions and signaling cascades related to cellular functions. Initially, every node, including HTT, begins with a zero-concentration parameter, which represents the system's still state before any perturbations. The scaled concentration values for simulations were depicted as 0 for inactive state and 1 for hyperactive state. Moreover, values ranging from 0.1-0.4 shows a low activation state and values between 0.5 and 0.9 represent an active state. The simulations were run up to the time where every node in the network reached a stable state that provided insights into the interactions between HTT and other BRN nodes in terms of activation and inhibition. Functions of all the biological entities that are involved in regulating apoptosis and proliferation are elaborated in **Table 4.1.**

Table 4.1: Functions of various entities that are participating in regulation of apoptosis and proliferation in the constructed BRN.

Entity	**Function**	**Reference**
PI3K/Akt	Mediates cell survival and proliferation through activation of p53 and other downstream pathways.	[88]
EGF	Stimulates EGFR, leading to the activation of RAF1 and influencing cell proliferation.	[89]

HTT	Interacts with glutamate and IP3R to regulate calcium signaling, influencing both apoptosis and other cellular signaling pathways.	[90]
IP3R	Activates BCL-XL. Releases calcium from the endoplasmic reticulum, influencing cellular responses such as apoptosis or survival.	[91].
P53	Induces apoptosis in response to stress, influenced by Akt and potentially modulated by HTT in the network.	[92]
BCL-XL	Acts as an anti-apoptotic entity, influenced by other signaling pathways like Akt, contributing to cell survival.	[93]
BAX,BAK	Promote apoptosis, activated in response to apoptotic signals potentially regulated by p53 and calcium levels.	[93]

CASP3	Executes apoptosis, activated by signals from BAX/BAK as part of the apoptotic response.	[94]

4.1.1 Dynamic simulations of BRN in normal state:

This simulation highlights the intricate relationships that exist within the BRN under normal conditions, where HTT plays a pivotal role in regulating the activity of pro-apoptotic and pro-survival pathways. The steady and cyclical behavior of HTT, along with its influence on downstream calcium signaling promotes cellular processes including proliferation while suppressing apoptosis. These results provide valuable insights into the dynamics of cellular signaling pathways, crucial for understanding the normal physiological roles of HTT and its impact on cellular health. The oscillatory behavior of various signaling molecules, including HTT, was carefully analyzed over the 50 time units of the dynamic simulation of the BRN as shown in **Figure 4.2**. The threshold set for the analysis of BRN was 0.5 as the entity having value close to 1 is considered to be in hyperactive state. Therefore, any value that is below 0.5 shows low activation state. HTT exhibited a steady state value of 0.176, suggesting a moderately low activation under normal physiological conditions. The network's ability to regulate the activation of the pathways leading to cell proliferation and apoptosis is dependent on this regulation of HTT. IP3R demonstrated a notably high activation level (0.901) during the simulation, indicating an active calcium signaling mechanism facilitated by the regulatory effect of HTT. BCL-XL, a key anti-apoptotic molecule, shows extremely high activity (0.998), indicating inhibition of apoptosis despite the activation of pro-apoptotic signals. This elevated BCL-XL activity could be a compensatory mechanism enhanced by HTT to maintain cellular integrity and survival. The observed level of apoptosis (0.36) reflects a balance maintained by HTT, where it modifies

apoptotic signaling to prevent unnecessary cell death while still allowing for the removal of potentially harmful cells. The interaction of pro-apoptotic proteins like p53 with HTT and BCL-XL highlights a highly regulated mechanism ensuring integrity and survival of the cell. Akt, a vital kinase in survival pathways, with a low value (0.045), suggests that HTT may regulate cell survival pathways independently of the Akt pathway under normal conditions. RAF1 and proliferation have high values (0.757 and 0.87 respectively), indicating high cell proliferation signaling. The interaction of HTT with these pathways is essential for maintaining controlled cell growth. The tumor suppressor protein P53, with a low activation (0.053), plays a key role in the activation of apoptosis in response to cellular stress or DNA damage.

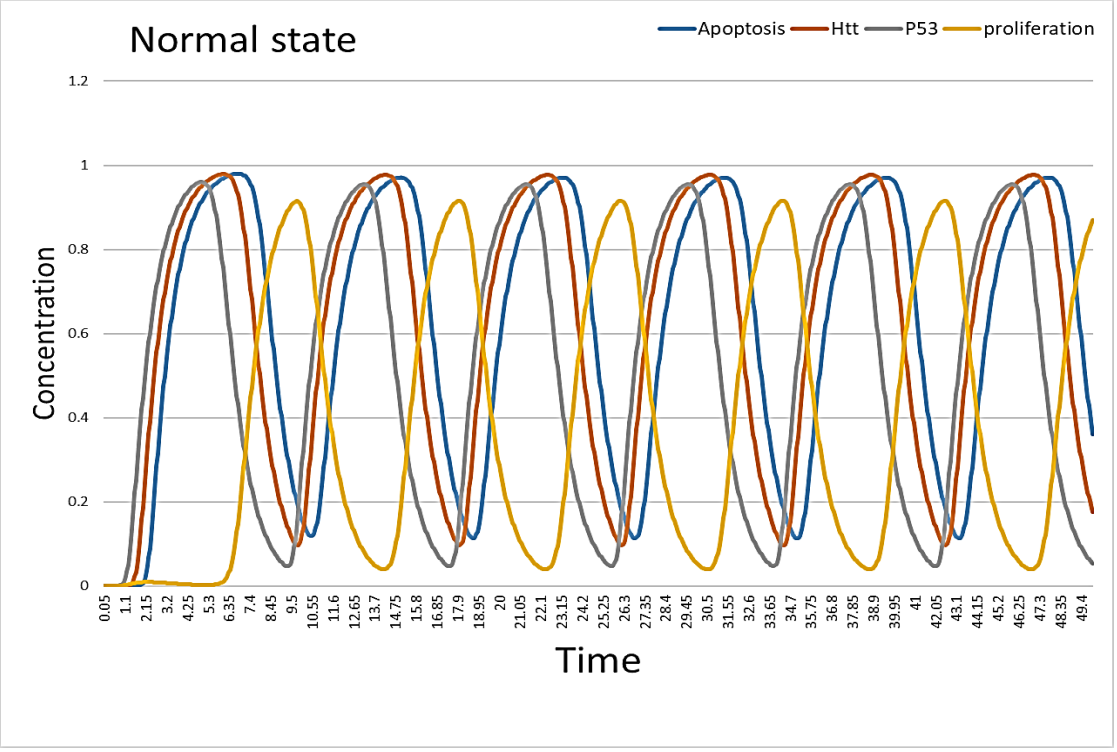


Figure 4.2: Oscillatory Dynamics of BRN Signaling in Normal Cells. The graph represents the fluctuations in activation levels of signaling pathways such as Akt, apoptosis, EGF, Glutamate, HTT, IP3, and p53 as well as cellular proliferation.

Nodes such as HTT and IP3R show closely synchronized oscillations, indicating a direct interaction between HTT modulation and IP3R activity, which affects calcium signaling. The graph illustrates that proliferation has larger amplitude oscillations, indicating strong activation and a crucial function in promoting cell development, whereas apoptosis shows lower amplitude oscillations, indicating controlled activation under normal conditions.

4.1.2 Dynamic simulations of BRN in perturbed state:

The simulation results with HTT perturbation showed significant alterations in the activation levels of various nodes within the BRN, indicating the significant impact of HTT on cellular signaling processes, shown in **Figure 4.3**. HTT exhibits a moderate activation level (0.354), indicating that its disrupted state impacts other pathways. IP3R has a high activation level (0.98), which points to significant calcium signaling, a key mediator of various cellular functions including apoptosis and proliferation. This is complemented by significant activation of Akt (0.905) and PI3K (0.525), which are central to the cell survival pathways, representing the cell's adaptive responses to maintain viability under stress. Apoptosis, BCL-XL and p53 are almost fully activated (0.99), highlighting a strong apoptotic response potentially triggered by HTT perturbation.

Moreover, CASP3, a key executor of apoptosis, also shows high activation (0.928), which aligns with the elevated levels of apoptosis and p53, highlighting the enhanced cell death. HTT has a strong impact on normal cellular processes, as evidenced by the low activation levels of RAF1 and proliferation (0.00042 and 0.00303 respectively), which indicate a significant reduction of cell growth and proliferation. These findings highlight the crucial role of HTT in regulating key cellular processes, where its perturbation leads to a clear activation of apoptotic pathways and suppression of cell proliferation. The high activation of apoptotic and growth factors suggests a cellular

environment striving for balance between survival and programmed cell death under stress conditions induced by HTT perturbation.

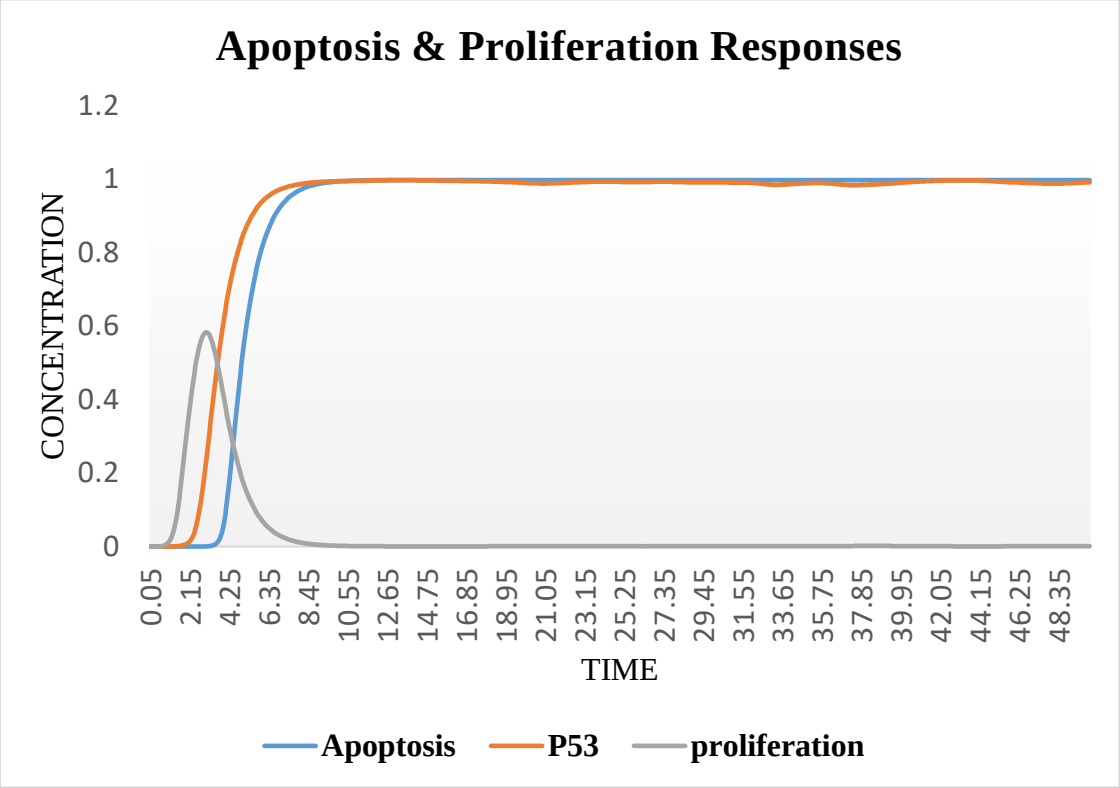


Figure 4.3: Oscillatory Dynamics of BRN Signaling in perturbed cells. This graph illustrates modifications in signaling dynamics that occur in response to cellular perturbations, providing information on the regulatory mechanisms and potential responses that may be targeted in treatment strategies.

Nodes such as IP3R, P53, and BCL-XL first show a rapid peak, indicating an immediate response to stimuli, mainly affecting the regulation of apoptosis and calcium signaling. These levels then stabilize, indicating ongoing activation to balance cellular stress responses. On the other hand, Akt shows an early peak and then levels off at a lower level, reflecting reduced survival signaling. Apoptosis exhibits an abrupt peak followed by a gradual decrease, indicating that survival signals initially trigger the process but subsequently reduce it. Proliferation remains consistently low, highlighting suppressed growth activity in this environment. This pattern illustrates how cells

dynamically balance survival and apoptosis in response to changes, maintaining stability. The overall values of all the entities before and after the simulations are mentioned in **Table 4.2** for both normal and perturbed state.

Table 4.2: Parameter values of genes interacting with HTT before and after dynamic simulations.

In Normal State			**In Perturbed State**	
Nodes	**Before**	**After**	**Before**	**After**
Akt	0	0.044774602	0	0.905925365
Apoptosis	0	0.360417022	0	0.997001879
BAX,BAK	0	1.71E-05	0	1.40E-05
BCL-XL	0	0.998461765	0	0.999595273
CASP3	0	0.113128303	0	0.928863532
EGF	0	0.450963806	0	0.323654064
EGFR	0	0.232454786	0	0.525968583
Glutamate	0	0.549036194	0	0.677973866
HTT	0	0.175760653	0.5	0.354550846
IP3R	0	0.900810562	0	0.9890607
P53	0	0.053312784	0	0.996155605
PI3K	0	0.232454786	0	0.525968583
RAF1	0	0.757025327	0	0.003039979
proliferation	0	0.870037839	0	4.20E-04

4.2 Molecular Docking:

In the molecular docking study conducted on the HTT protein, specific binding pockets were identified, comprising residues Asp2737, Glu2738, Asp2758, Lys2759, Glu3106, Glu3107, and Leu3108 [95]. These binding pockets are at coordinates 143.56, 228.48, and 182.69 on the x, y, and z axes, respectively. For every ligand, ten different poses were generated and evaluated using

the gold score. The GOLD score is calculated by adding internal and external hydrogen bond energies, Van der Waal energies, and internal torsion using the formula below:

GOLDScore = ΔG (HB_int) + ΔG (HB_ext) + ΔG (VDW_int) + ΔG (VDW_ext) + ΔG (TOR_int)

Where HB_int and HB_ext are internal and external hydrogen bond energies respectively, VDW_int is Van der Waal energies and TOR_int is represented for internal torsion.

For each ligand, only the pose with the highest score was selected for further analysis. These top-scoring poses were then evaluated for their potential inhibitory activity by correlating the gold scores with logarithmic pIC_{50} values, producing an R-square value, 0.9499, that quantitatively assessed the relationship between docking score and biological activity as shown in **Figure 4.4**.

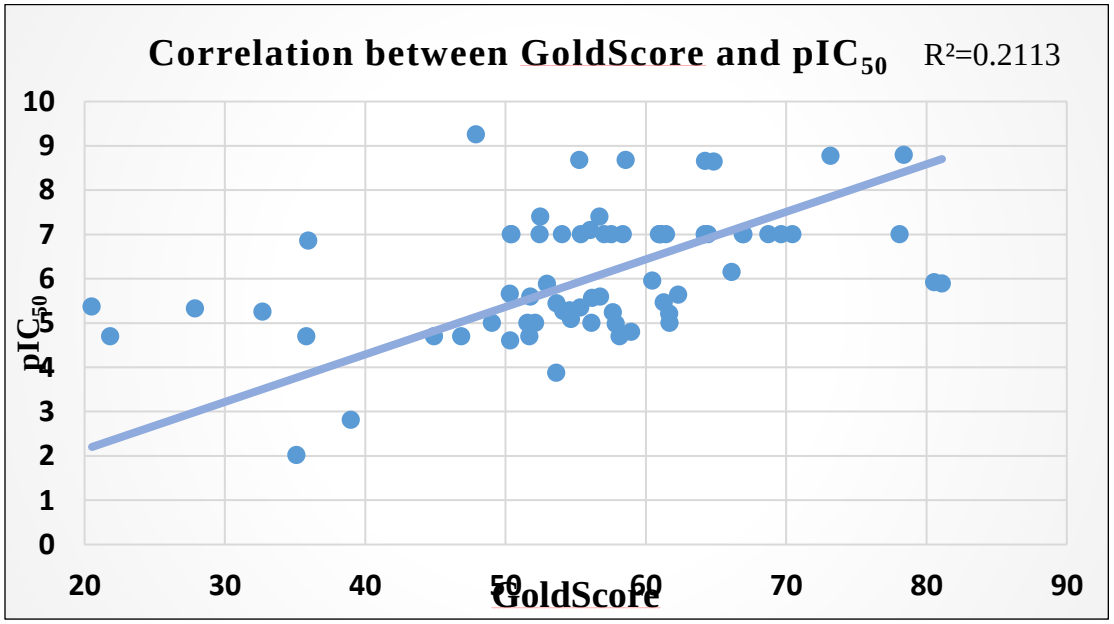


Figure 4.4: Relationship between log pIC50 values and Gold scores for different ligands that are docked with the huntingtin protein. Each data point represents a ligand, plotted according to its Gold score (x-axis) and corresponding log pIC50 value (y-axis). The trend line illustrates the relationship between docking scores and biological activity and the R-squared value indicating the fit of the model.

Based on the graphical analysis of this correlation, four ligands were chosen for subsequent molecular dynamics (MD) simulations. This selection included the ligand with the highest IC_{50} value, the one with the lowest, and two randomly chosen from the set, aiming to provide a comprehensive analysis of their dynamic interactions and stability within the HTT binding site in a simulated biological environment. The interactions of these ligands with protein (docked complexes) are shown in the **Figure 4.5, 4.6, 4.7 and 4.8** for Paroxetine, Mannitol, undefined ligand and Gabapentin respectively.

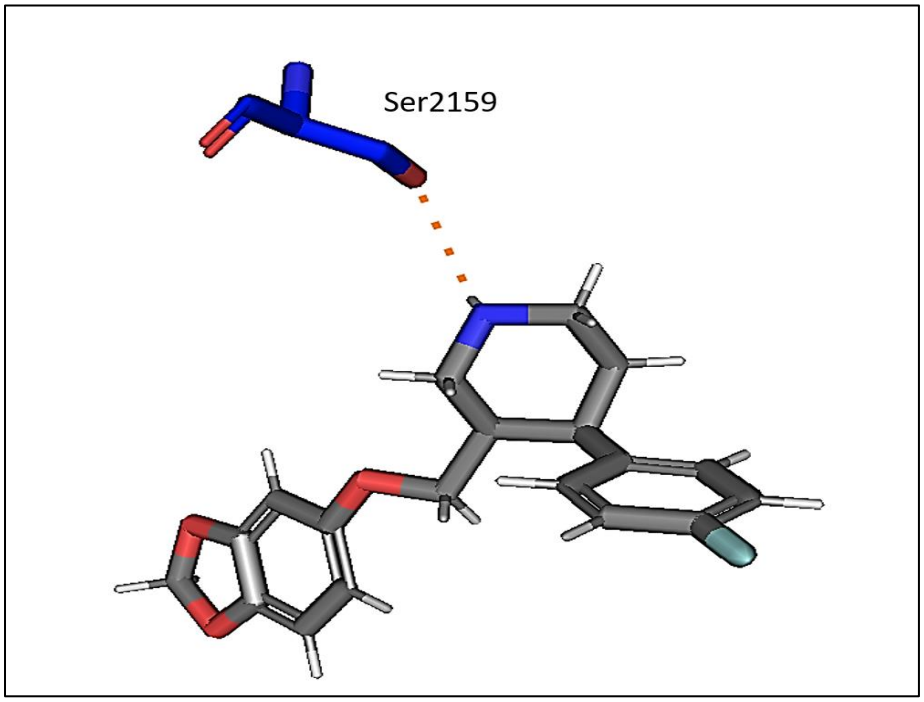


Figure 4.5: Docked Complex of HTT with Paroxetine having Goldscore 47.91.

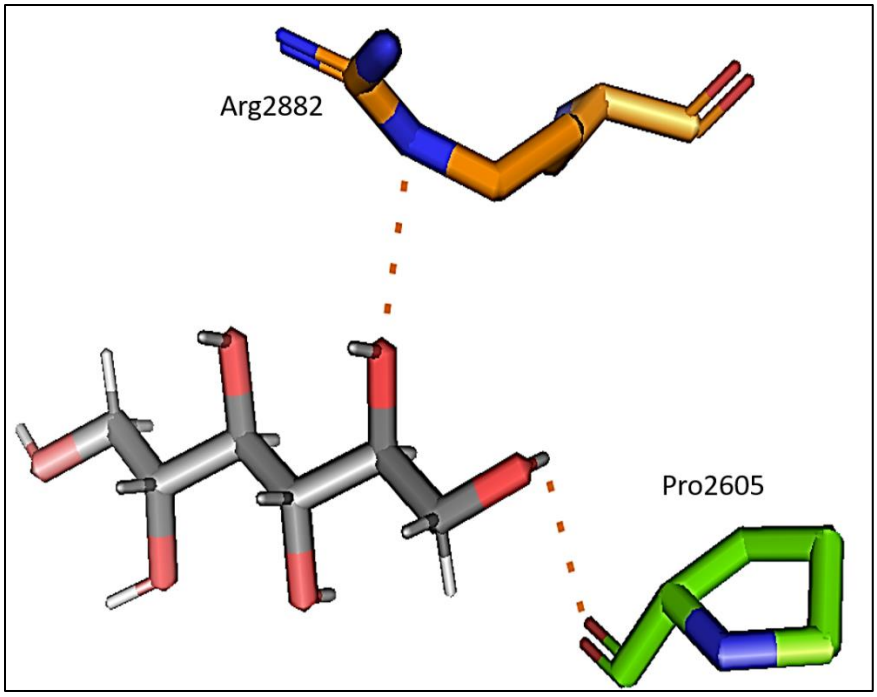


Figure 4.6: Docked Complex of HTT with Mannitol having Goldscore 35.11.

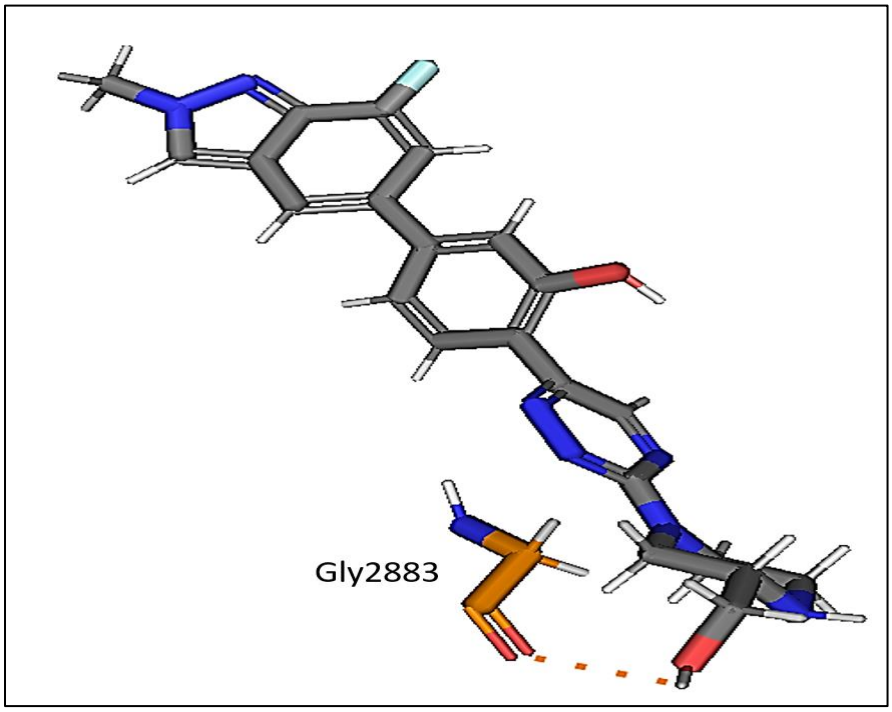


Figure 4.7: Docked Complex of HTT with an undefined ligand having Goldscore 78.41.

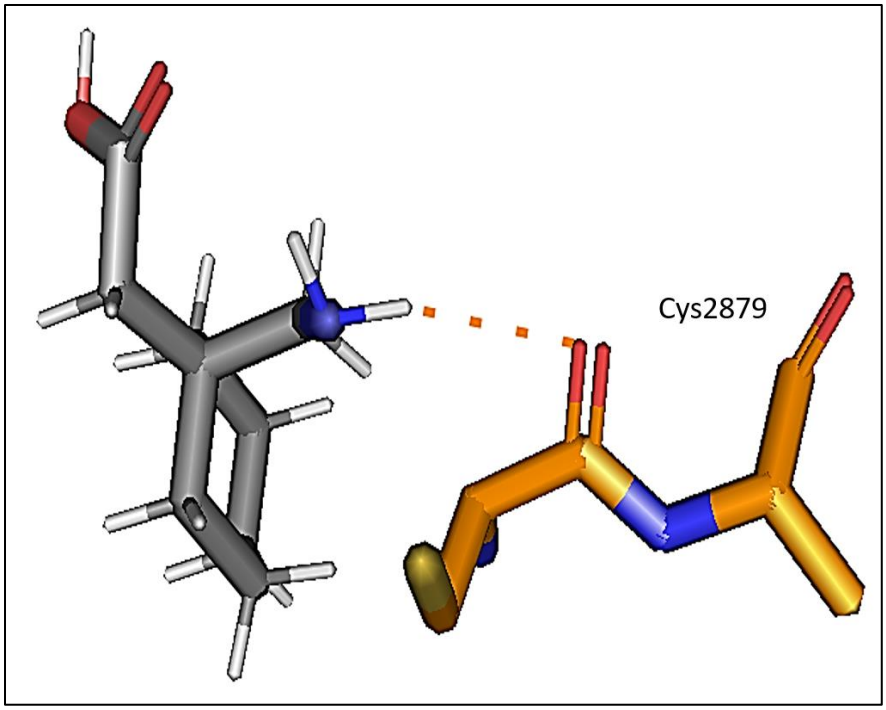


Figure 4.8: Docked Complex of HTT with Gabapentin having Goldscore 35.96.

4.3 Molecular Dynamics Simulations of Huntingtin Protein:

The simulations were performed with the Linux-based GROMACS software, which probes the artificial body environment that allows protein–ligand complexes to stabilize within the system. The overall stability of the structure was evaluated by examining the RMSD and RMSF that helped in understanding the insights into the dynamic behavior of the protein under simulated physiological conditions. There is a notable rise in RMSD during the early stages of the simulation (around the first 20 ns), as seen in **Figure 4.9**, which suggests that the protein is converging from its initial conformation towards a more energetically favorable structure. This initial adjustment phase is common as the system equilibrates. After the initial increase, the graph fluctuates between approximately 0.2 nm and 0.3 nm which appears to stabilize. These fluctuations are normal and represent how dynamic the protein is as it interacts with its environment. The relatively narrow range indicates that the protein has reached a stable conformation that does not deviate

dramatically from this new average structure. The lack of significant increases in RMSD values after the initial period (between 0 and 20 ns) suggests that the protein does not undergo major conformational changes for the remainder of the simulation. This stability represents that the overall structure of a protein is being maintained. This kind of analysis is important for confirming the protein's stability under simulated physiological conditions and that it is suitable for further analysis.

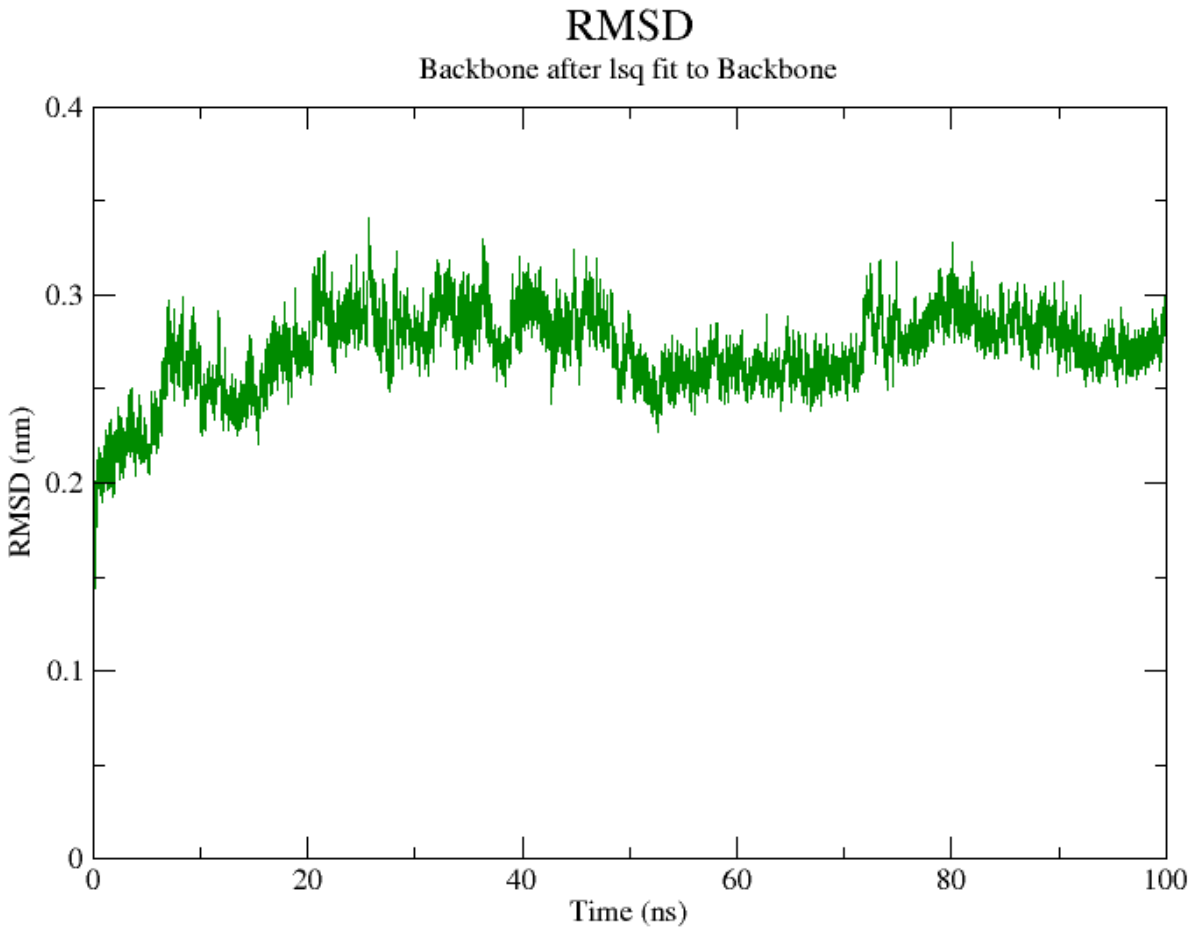


Figure 4.9: Root Mean Square Deviation of the Huntingtin Protein Backbone over 100 ns of Molecular Dynamics Simulation.

The RMSF graph, as shown in **Figure 4.10**, highlights areas of stability and flexibility of amino acid residues and also offers a comprehensive perspective of the protein's dynamic landscape

across its sequence. Peaks in the RMSF graph show residues with high mobility or flexibility during the simulation. These areas may consist of unstructured regions or loops not part of stable secondary structures such as beta sheets or alpha helices. Specific peaks, particularly those reaching as high as 0.3 nm or more, suggest significant movement relative to the protein's average structure. These might be involved in binding processes, conformational changes, or regulatory mechanisms. Areas with lower RMSF values (closer to the baseline) are typically structured regions. These might include well-folded domains that are crucial for maintaining the protein's structural integrity. These areas frequently play important physiological roles potentially encompassing active sites for protein-protein interactions. The pronounced peaks in the region of residues around 2400-2500 indicate high flexibility. Towards the end of the graph (around residue numbers 3100-3200) the dramatic increase in fluctuation suggests a highly dynamic region possibly an exposed loop that might interact with other cellular components or be involved in regulatory processes.

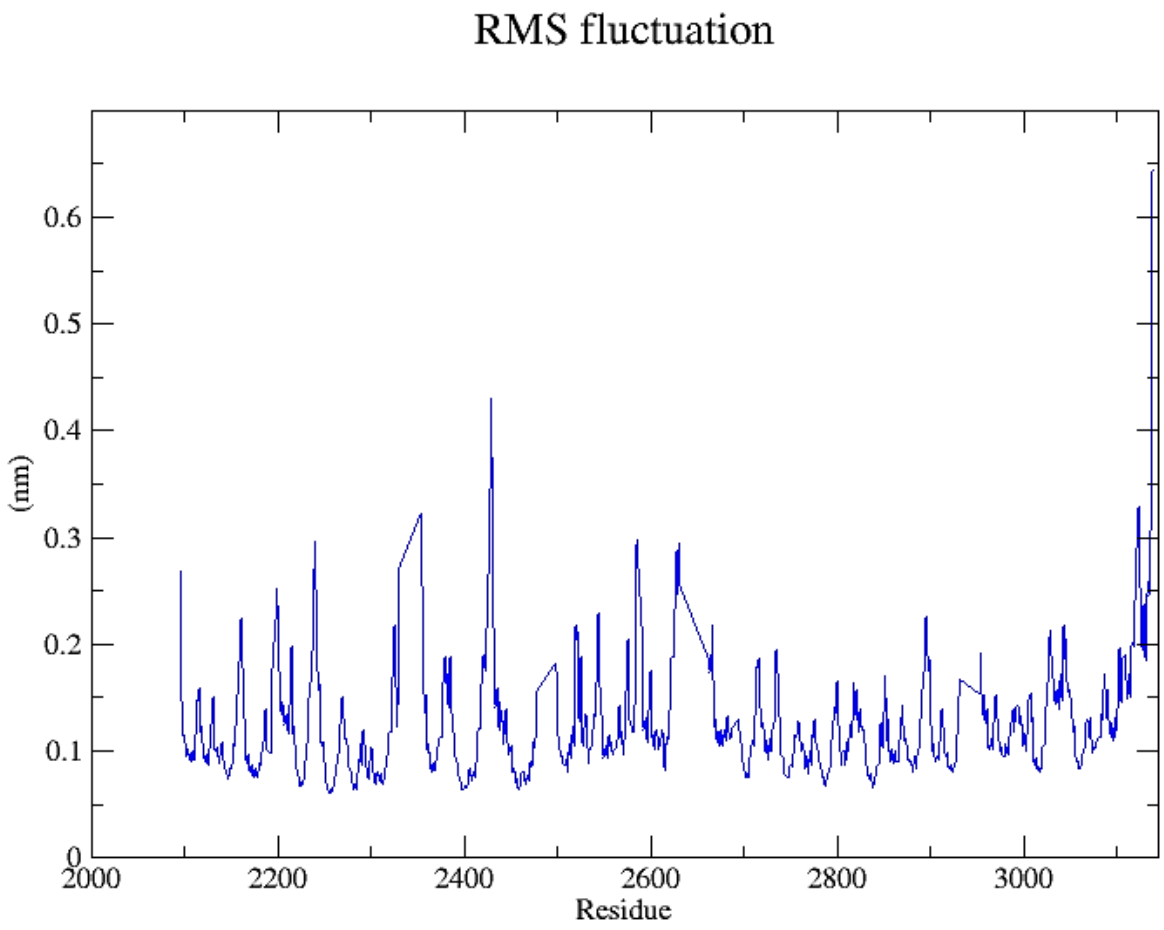
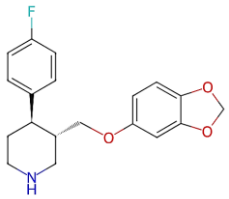
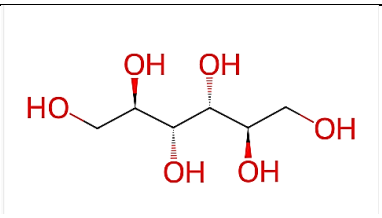


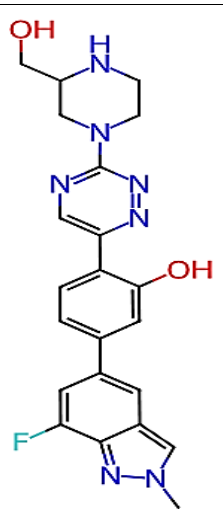
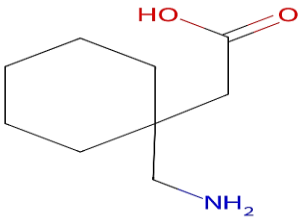
Figure 4.10: RMSF of protein residues over a 100 ns simulation with higher peaks indicating highly unstable residues, demonstrating significant fluctuations during the simulation.

4.4 Molecular Dynamics Simulations of Protein-Ligand complexes:

The top poses of the selected protein–ligand complexes were used to perform the molecular dynamics simulations. Four different protein-ligand complexes were selected for the MD simulations based on their docking scores and logpIC_{50} values. The ligands that were selected for further simulations are CHEMBL4875017, CHEMBL940, CHEMBL689, and CHEMBL490 listed in **Table 4.3**. These ligands demonstrate different docking efficacy and potency, as indicated by their Gold score and logpIC_{50} values. These complexes were then subjected to rigorous MD simulations using GROMACS software. The aim was to investigate dynamic interactions and explore the stability and flexibility of the protein conformations when bound to these ligands. To evaluate the structural dynamics and integrity of these complexes, we focused on RMSD and RMSF of these complexes.

Table 4.3: Ligands selected for Molecular-Dynamics Simulations of Protein-Ligand complexes based on their gold score and Logp IC50 values.

Chembl ID	Name	Structure	pIC50	Goldscore
CHEMBL490	PAROXETINE		9.25	47.91
CHEMBL689	MANNITOL		2.02	35.12

CHEMBL4875017	Unknown		8.79	78.41
CHEMBL940	GABAPENTIN		6.85	35.96

4.4.1 Complex 1:

The RMSD values as in **Figure 4.11** show an initial rise and stabilize around 0.3 nm. This shows that the protein-ligand complex reached an equilibrated state early in the simulation. The fluctuations within a relatively small range indicate that the ligand-protein complex has a stable structure with moderate flexibility. The continuous pattern around a central value indicates that major conformational changes are not significant, and the complex does not experience significant unfolding or large-scale structural rearrangement. Towards the end of the simulation, the values increased significantly from 0.11 to 0.37, indicating conformational adjustment or increased mobility in the protein-ligand complex with the average RMSD value of **0.58**. Such changes could

be important for understanding the dynamic nature of the ligand binding and its effects on the protein structure.

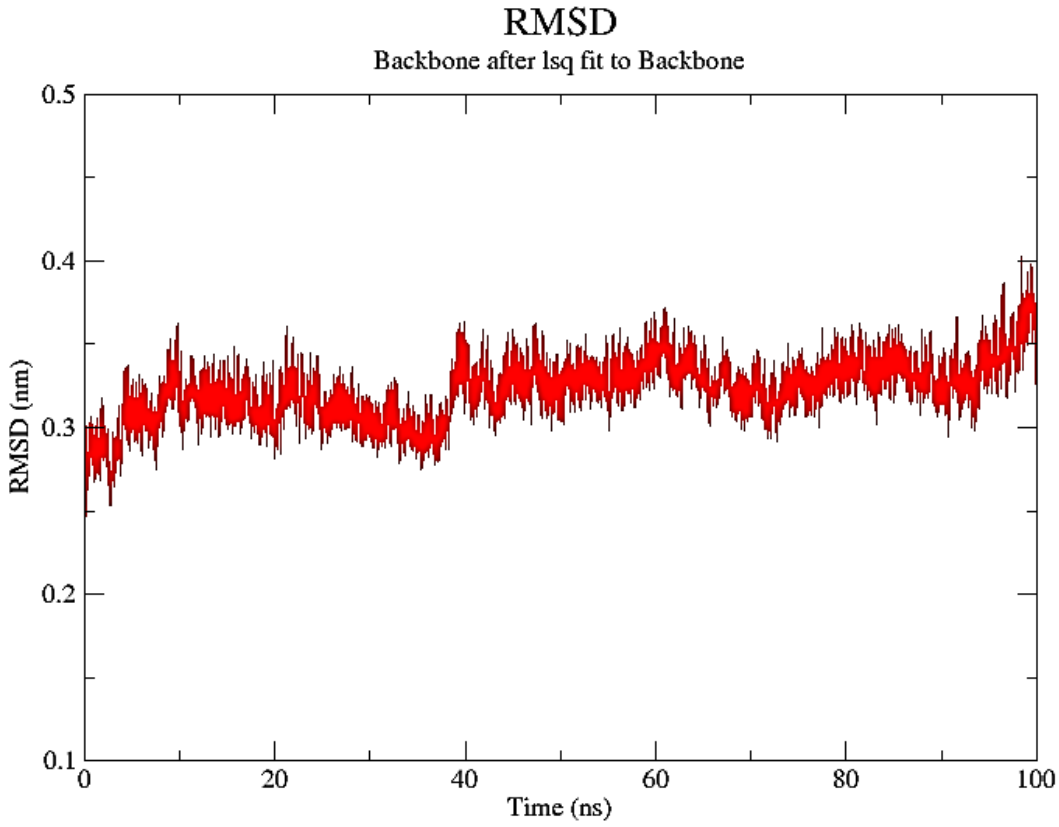


Figure 4.11: RMSD graph indicating MD simulations of protein-ligand complex, with the ligand (Paroxetine) having highest pIC_{50} value.

The RMSF graph, shown in **Figure 4.12**, displays residues with high flexibility, indicating areas of the protein that vary significantly during the simulation. Notably, three peaks are witnessed, indicating residues with significant fluctuations, Pro 2200, His 2380, and Trp 2421, each with RMS fluctuations close to 0.4 nm. These high RMSF values indicate that these residues are highly flexible, which could point to conformational transition sites, loops, or surface-exposed areas. On the other hand, other regions with lower RMSF values indicate more stable and less flexible parts of the protein, which could correspond to secondary structural elements such as alpha-helices or beta-sheets. The calculated average RMSF value is **0.25**. Overall, the graph reveals the dynamic

behavior of the protein that highlights the areas of both high flexibility and stability which can provide information about the protein's functional mechanisms.

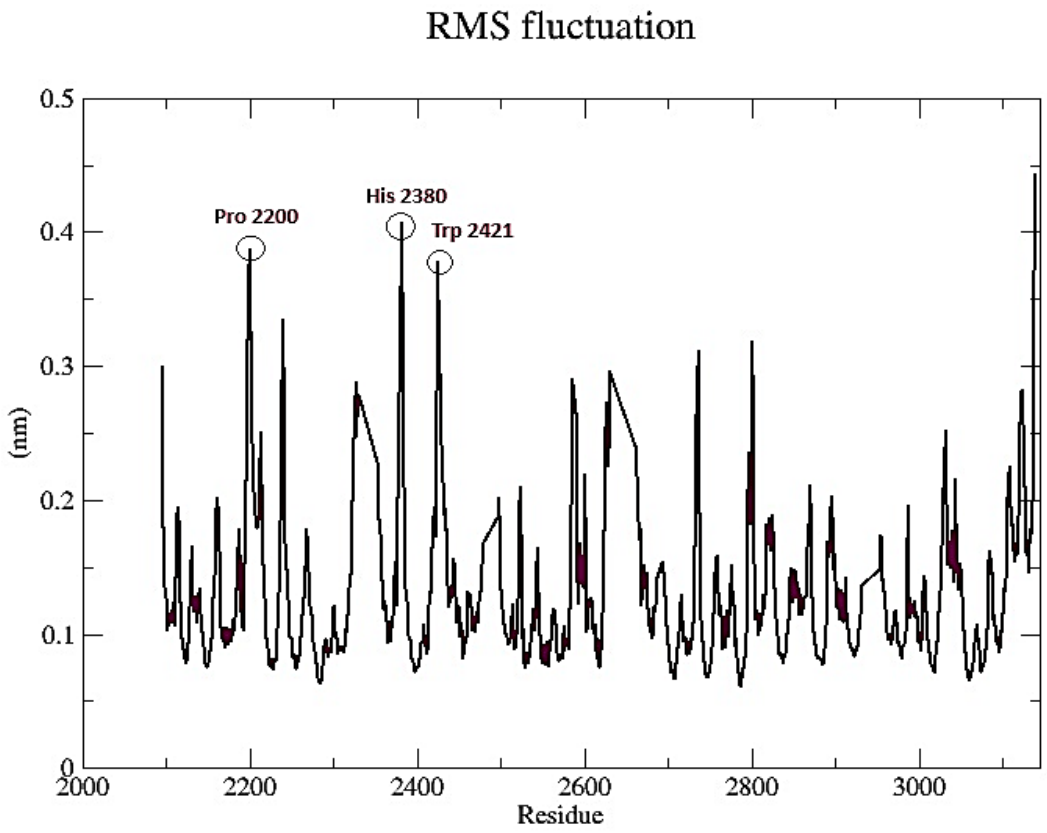


Figure 4.12: RMS fluctuation graph illustrating protein residue flexibility, with higher peaks at Pro 2200, His 2380, and Trp 2421 indicating regions of high conformational mobility within the protein.

4.4.2 Complex 2:

The RMSD values fluctuate between 0.15 and 0.35 nm as shown in **Figure 4.13**. This range shows moderate conformational changes as the protein-ligand complex explores various configurations under physiological conditions simulated during the MD. Between 20 and 40 ns, there is a clear stabilization phase in which RMSD values hover around 0.25 nm, indicating an interval of relative structural stability. After 40 ns, the RMSD becomes more variable, particularly around 60 ns and at the end of the simulation period (80-100 ns). These peaks indicate transient conformations or dynamic interactions within the protein-ligand complex, which could be caused by ligand binding effects or local folding/unfolding events inside the protein itself. The average RMSD value (**0.28**) indicates that the protein-ligand complex achieved a dynamic equilibrium with moderate fluctuations in the backbone structure, highlighting the stability of the protein under the simulated conditions and the effective binding of the ligand.

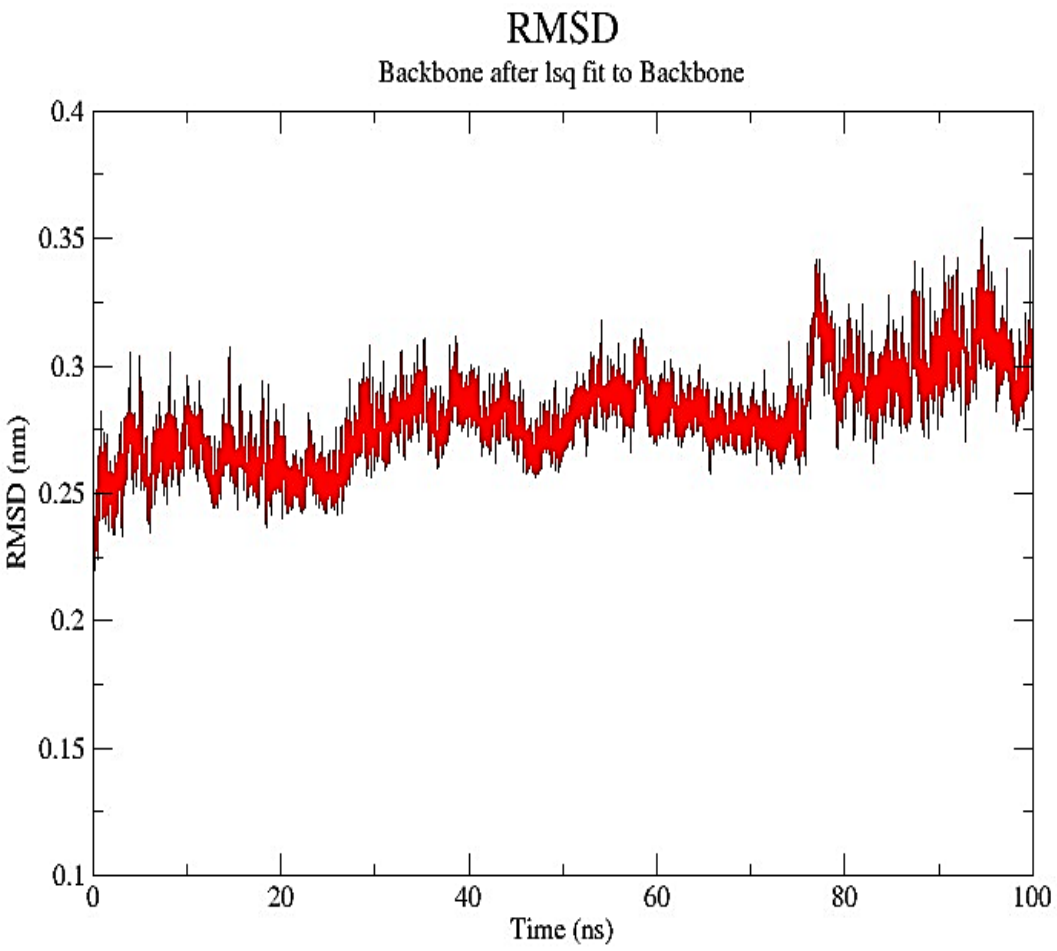


Figure 4.13: RMSD graph indicating MD simulations of protein-ligand complex, with the ligand (Mannitol) having least pIC_{50} value.

The RMSF graph, shown in **Figure 4.14**, typically shows fluctuations ranging from about 0.1 nm to around 0.6 nm, showing varying degrees of mobility across various regions of the protein.

A prominent peak at residue Leu 2328 extends up to roughly 0.75 nm, which is significantly greater than the average variation seen in other regions. This shows that Leu 2328 is extremely flexible or exposed to solvent, and may play an important role in the protein's dynamics or interaction with the ligand. The average RMSF value that was calculated is **0.43**.

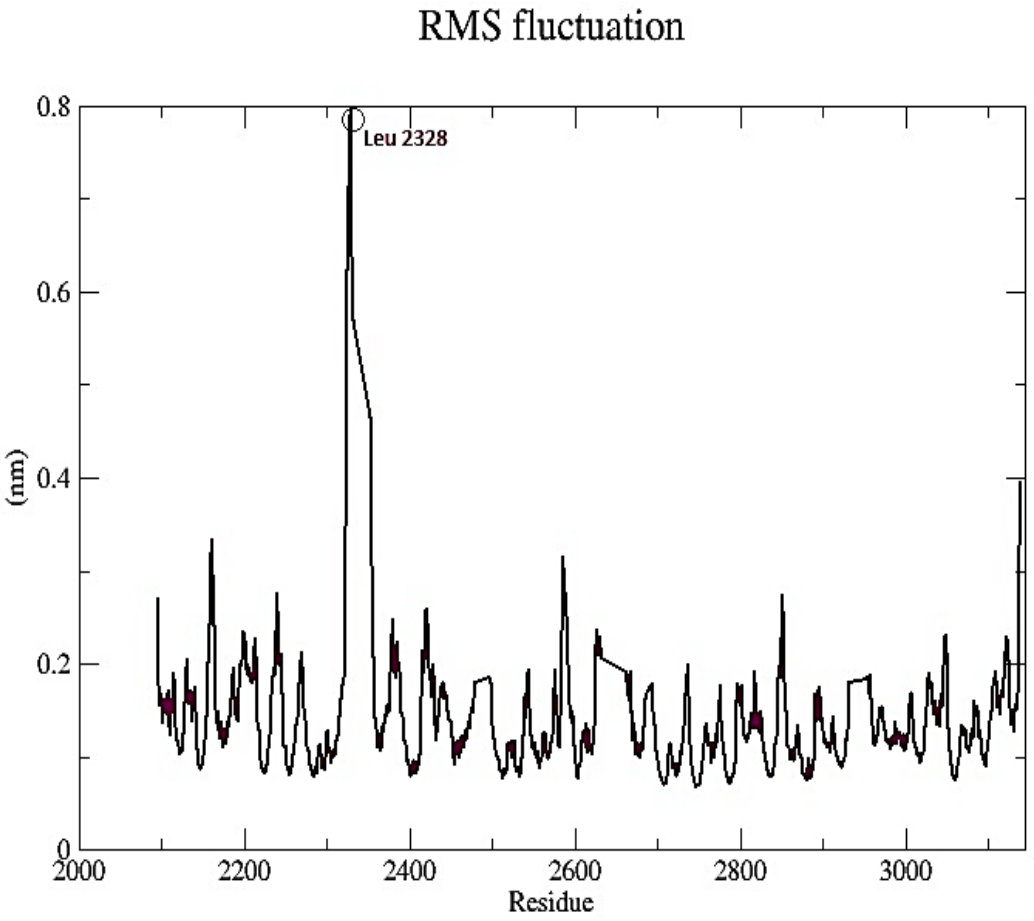


Figure 4.14: RMS fluctuation graph illustrating protein residue flexibility, with higher peak at Leu 2328 indicating region of high conformational flexibility within the protein.

4.4.3 Complex 3:

The stability observed in the graph represents how the complex behaves over time, which is important for understanding the dynamics of the protein-ligand interaction during the simulation. The RMSD values decrease and stabilize around 0.25 nm by about 20 nanoseconds as seen in

Figure 4.15. This lower and more stable RMSD indicates that the complex has established a relatively stable conformation within the simulated environment, potentially representing a more biologically relevant state. From 20 to 60 nanoseconds, the RMSD values fluctuate consistently within a relatively narrow range. These fluctuations reflect the dynamic nature of the protein-ligand interaction, in which the complex constantly adapts but remains within a restricted structural region, implying a balance of flexibility and stability. Towards the end of the simulation (60 to 100 nanoseconds), the RMSD values become slightly more variable. The average RMSD value for this complex is **0.29**. This indicates that the complex is exploring different conformational regions or responding to simulation settings. Overall, the RMSD graph suggests that the protein-ligand complex achieves a reasonable level of stability after an initial period of significant structural adjustment.

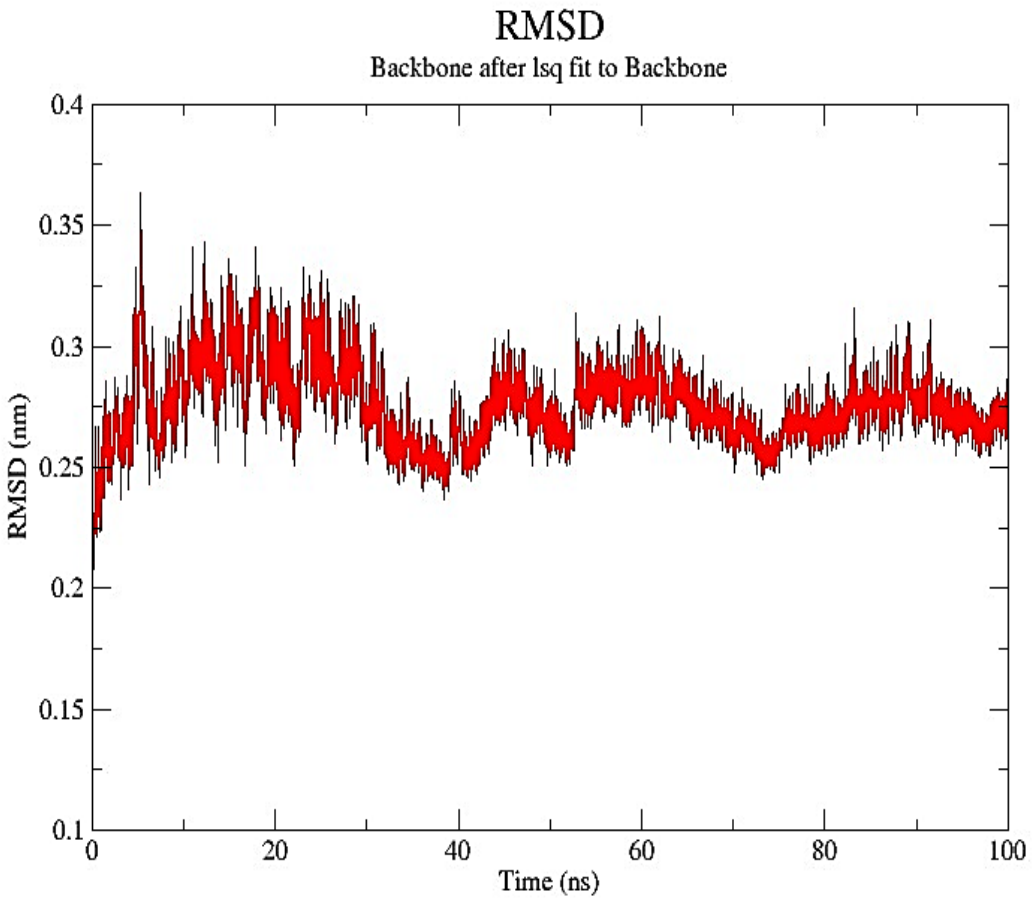


Figure 4.15: RMSD graph indicating MD simulations of protein-ligand complex (undefined ligand with CHEMBL ID 4875017).

The RMSF graph points out several residues with significant changes, including Val 2095, Leu 2196, Thr 2631, and Thr 2800. These peaks, particularly those reaching 0.4 nm or higher, show areas of considerable mobility within the protein structure. This extreme flexibility can be due to the protein's functional features such as ligand binding, enzyme activity, or interaction with other molecules. The residues with the highest RMSF (such as Thr 2631 and Thr 2800), which can be seen in **Figure 4.16**, may be crucial for protein function because they may be engaged in binding sites that can be helpful to support biological processes such as catalysis, substrate binding or structural transformations. The average value for this RMSF graph is **0.25**.

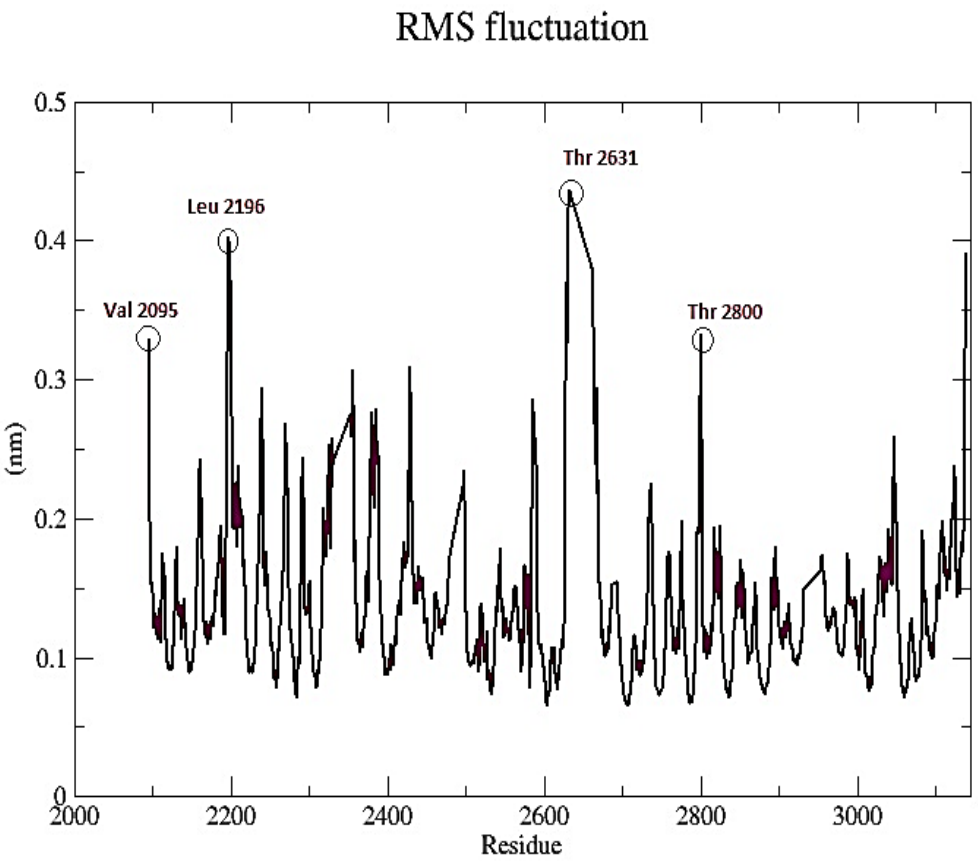


Figure 4.16: RMS fluctuation graph illustrating protein residue flexibility, with higher peak at Val 2095, Leu 2196, Thr 2631 and Thr 2800 indicating region of high conformational flexibility within the protein.

4.4.4 Complex 4:

The graph, as shown in **Figure 4.17**, illustrates that the protein-ligand complex remains relatively stable during the simulation with expected levels of fluctuation for dynamic biological molecules, having average value of **0.29**. The constant distribution of RMSD values, notably between 0.25 and 0.35 nm, indicates that the complex retains structural integrity and allows for essential flexibility. Such stability and flexibility are important for the ligand's functional behavior within the complex as they indicate that the ligand is adequately positioned for interaction with the protein's active sites. This analysis is essential for understanding the complex's behavior in a simulated biological environment, as it can provide insights useful for drug development and other biochemical applications. Towards the end of the simulation (80 to 100 nanoseconds), there is a minor rise in both the frequency and amplitude of fluctuations, with RMSD values approaching 0.35 nm. This tendency could indicate that the complex is increasing its mobility or exploring new conformational space.

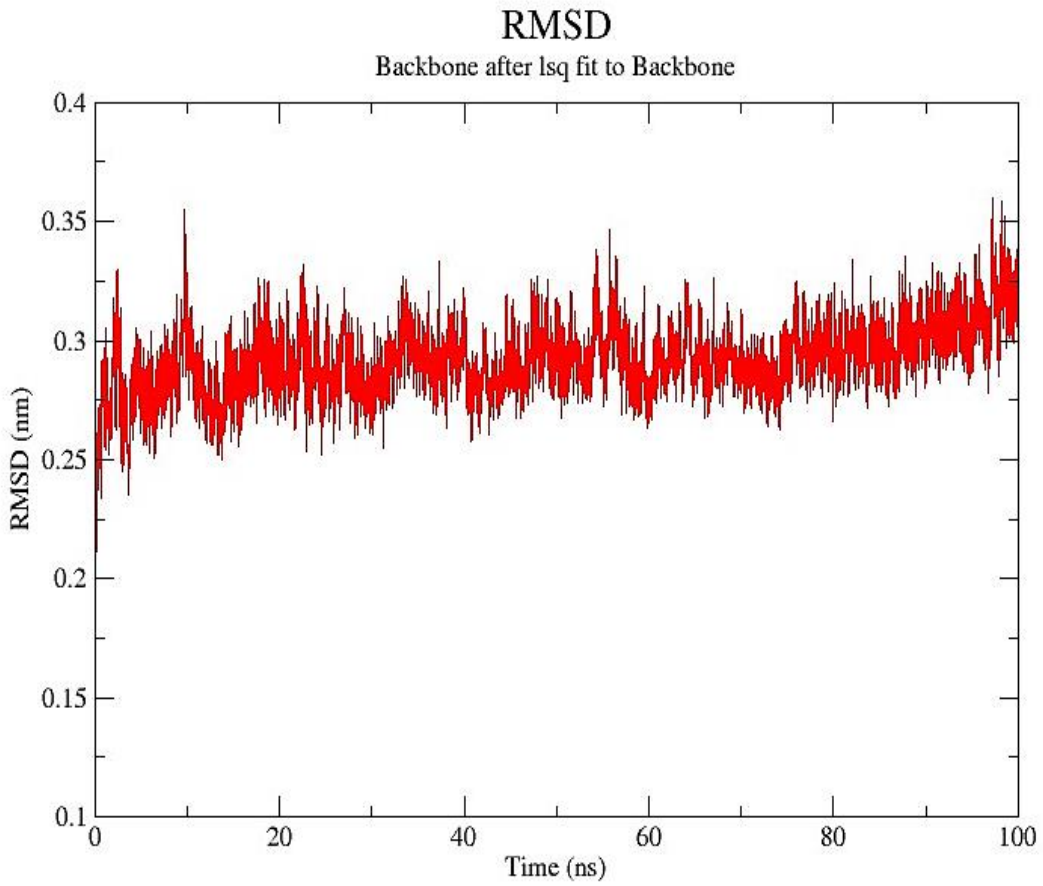


Figure 4.17: RMSD graph indicating MD simulations of protein-ligand complex, with the ligand (Gabapentin).

The graph shown in **Figure 4.18** focuses on residues with high variations, including Ala 2198, His 2240, Asp 2429, Ser 2586, and Ile 2630. Peaks in RMSF at these places, when fluctuations approach or surpass 0.4 nm, indicate high mobility. This could indicate flexible regions within the protein that may be engaged in important dynamic processes. The regions of lower fluctuation, generally around or below 0.2 nm, represent structural rigidity, which could be critical for maintaining the overall integrity and stable framework of the protein structure. The calculated average RMSF value for this protein-ligand complex is **0.27**.

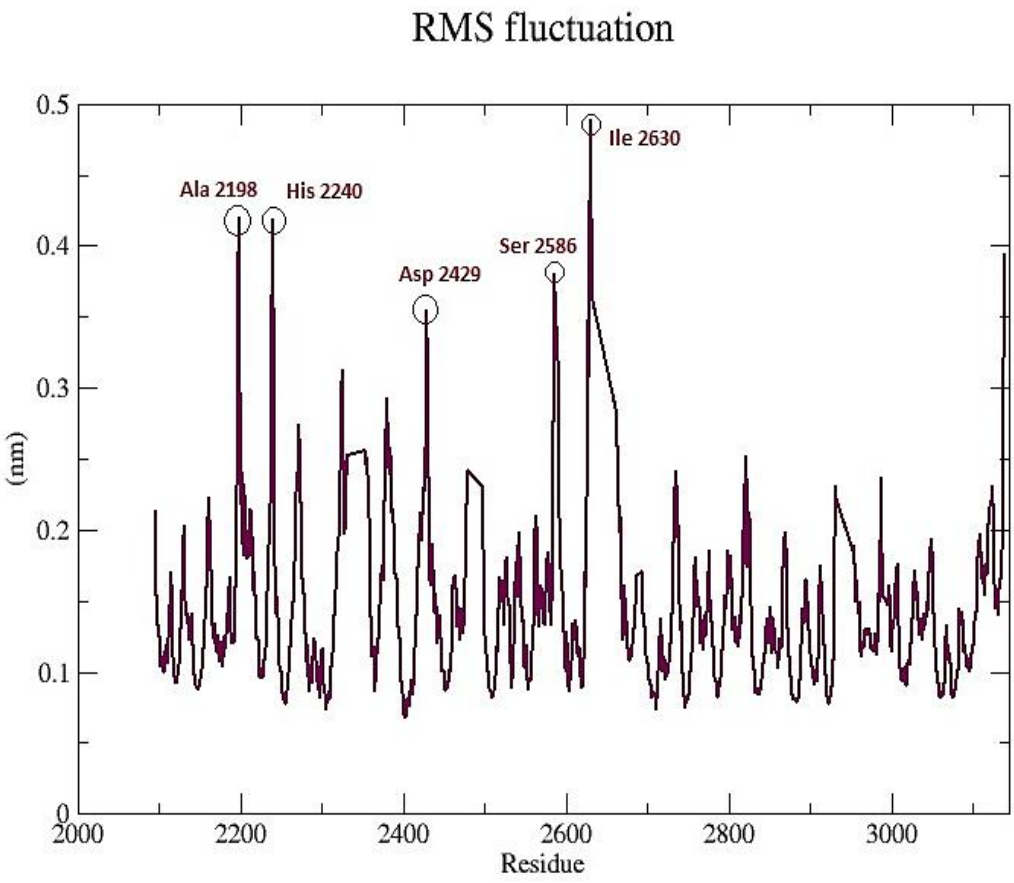


Figure 4.18: RMS fluctuation graph illustrating protein residue flexibility, with higher peak at Ala 2198, His 2240, Asp2429, Ser2586 and Ile 2630 indicating region of high conformational flexibility within the protein.

CHAPTER 5: DISCUSSION

Huntington's disease is a progressive neurodegenerative disorder characterized by motor, cognitive, and psychiatric symptoms that is primarily caused by an expansion of CAG repeats in the Huntingtin gene. This genetic mutation causes an aberrant increase of the polyglutamine stretch in the huntingtin protein, which has a significant impact on its function and contributes to disease pathology. The HTT gene lies on the short arm of chromosome 4, and the length of the CAG repeat expansion correlates with disease severity and onset. The HTT protein is essential for several cellular processes, including neuronal excitability, synaptic plasticity, and mitochondrial function. Glutamate signaling and ER-mitochondrial communication are both disrupted in HD, deteriorating synaptic dysfunction and mitochondrial degradation. This causes increased oxidative stress and neuronal death, particularly in medium spiny neurons (MSNs). Furthermore, HD affects the gut-brain axis, altering the gut microbiota composition, which might have an impact on neuronal health as well as disease progression. In our research, we constructed a knowledge-driven BRN of HTT protein that provides a comprehensive overview of the complex interactions and regulatory mechanisms involving HTT. Important pathways and molecular interactions that are perturbed in HD were identified by simulating Biological Regulatory Network that was helpful in the identification of targets for therapeutic intervention. This study highlights HTT protein as a potential drug target because of its interaction with various proteins and complex cellular pathways that are also involved in apoptosis and proliferation. We also performed molecular docking and dynamic simulations to investigate the interactions of potential inhibitors with the HTT protein. Using the crystal structure of HTT (PDB ID 6X9O), especially focusing on its C-terminal region, we targeted the interactions that are most relevant to the pathology of this disease. To identify specific binding sites on HTT, including residues Asp2737, Glu2738, Asp2758, Lys2759,

Glu3106, Glu3107, and Leu3108, GOLD software was used. A dataset of sixty-nine HTT inhibitors, with associated PIC_{50} values obtained from the literature and the ChEMBL database, guided our selection of four prominent ligands for further research. To examine the dynamic interactions between these ligands and the HTT protein, rigorous MD simulations were performed using GROMACS software. This strategy was used to illustrate the stability and flexibility of protein conformations while bound to these ligands, which is critical for understanding the molecular mechanisms. To analyze the results of simulations, RMSD and RMSF analysis were employed to evaluate the structural dynamics and integrity of the protein-ligand complexes. The RMSD data showed that the complexes quickly reached a state of relative stability, with only minor fluctuations during the simulation period. This shows a strong and stable association between the HTT protein and the selected ligands. Moreover, RMSF values identified those regions of the protein that showed significant flexibility during the interaction. Additionally, the identified flexible regions correspond to the binding sites which means that when the entire complex remains stable, these active sites have the ability to preserve the required mobility for functional activity and possibly help in the therapeutic efficiency of the ligand. Our findings highlight the use of molecular docking and MD simulations for the investigation and validation of HTT interactions with potential inhibitors. The identified ligands not only bind with high affinity, but also preserved HTT's structural integrity and flexibility leading to understand the potential pathways for therapeutic intervention.

CONCLUSION

This research examined the molecular mechanistic of Huntington's disease, focusing mainly on the structural and functional dynamics of the huntingtin protein as disrupted by pathological CAG repeat expansions. By combining a biological regulatory network with modeling strategies such as molecular docking and MD simulations, we were able to understand the complex pathogenic interactions in HD. Using the GOLD software, we identified those HTT binding sites that are important for interaction with potential inhibitors which allowed us to choose four promising ligands from a large dataset namely Paroxetine, Mannitol, Gabapentine and an undefined ligand with CHEMBL ID 4875017. This selection was made on the basis of correlation graph between gold score and pIC_{50} values of these ligands. The inhibitor, Paroxetene, was considered best of all because of its highest pIC_{50} value. Afterwards, MD simulations using GROMACS provided valuable insights into the stability and flexibility of these protein-ligand complexes over time. The stability observed emphasizes a strong interaction of the residues Val 2095, Leu 2196, Ala 2198, Pro 2200, His 2240, Leu 2328, His 2380, Trp 2421, Asp2429, Ser2586, Ile 2630, Thr 2631 and Thr 2800 at the binding sites that suggests potential therapeutic efficacy in treating HD symptoms, while the observed flexibility at certain protein regions highlights the preserved functional capacity of HTT essential for its biological roles. The average RMSD values for the complexes with Paroxetene, Mannitol, Gabbapentin and an undefined ligand (CHEMBL4875017) are 0.58, 0.28, 0.29 and 0.29 respectively and the average RMSF values for these ligand complexes are 0.25, 0.43, 0.27 and 0.25 respectively. The RMSD and RMSF analyses highlighted the conformational integrity and functional dynamics of these complexes, indicating the therapeutic potential of these ligands. Our findings not only increased our understanding of the

molecular basis of the disease but also proposes a novel framework for the development of targeted therapies.

Future Recommendations

Future research should focus on in vivo validation of these ligands for the evaluation of their efficacy and safety in clinical models of Huntington's disease. Moreover, investigation of additional genetic and environmental interactions that may influence treatment outcomes will be helpful for the development of personalized therapeutic strategies to combat this fatal disorder.

REFERENCES

[1] A. Aubeeluck and E. Wilson, "Huntington's disease. Part 1: essential background and management," *Br. J. Nurs. Mark Allen Publ.*, vol. 17, no. 3, pp. 146–151, Feb. 2008, doi: 10.12968/bjon.2008.17.3.28402.

[2] C. A. Ross and M. R. Hayden, "Huntington's disease," in *Analysis of Triplet Repeat Disorders*, Garland Science, 1998.

[3] "Parameters Affecting the Pathogenicity of Huntingtin Protein - ProQuest." Accessed: Jun. 28, 2024. [Online]. Available: https://www.proquest.com/openview/6d2f30f457fae2a2adb3de19bcee1050/1?cbl=18750&pq-origsite=gscholar&parentSessionId=Vxu%2F89d1RMvuSDfg7B1FCvNgPCLx%2BSt6rZvA8b1E%2FDs%3D

[4] "Frontiers | Cortical and Striatal Circuits in Huntington's Disease." Accessed: Jun. 28, 2024. [Online]. Available: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2020.00082/full

[5] "Balance between synaptic versus extrasynaptic NMDA receptor activity influences inclusions and neurotoxicity of mutant huntingtin | Nature Medicine." Accessed: Jun. 28, 2024. [Online]. Available: https://www.nature.com/articles/nm.2056

[6] M. DiFiglia, "Excitotoxic injury of the neostriatum: a model for Huntington's disease," *Trends Neurosci.*, vol. 13, no. 7, pp. 286–289, Jul. 1990, doi: 10.1016/0166-2236(90)90111-m.

[7] Y. P. Deng, T. Wong, C. Bricker-Anthony, B. Deng, and A. Reiner, "Loss of corticostriatal and thalamostriatal synaptic terminals precedes striatal projection neuron pathology in heterozygous Q140 Huntington's disease mice," *Neurobiol. Dis.*, vol. 60, pp. 89–107, Dec. 2013, doi: 10.1016/j.nbd.2013.08.009.

[8] M. E. Ehrlich, "Huntington's Disease and the Striatal Medium Spiny Neuron: Cell-Autonomous and Non-Cell-Autonomous Mechanisms of Disease," *Neurotherapeutics*, vol. 9, no. 2, pp. 270–284, Apr. 2012, doi: 10.1007/s13311-012-0112-2.

[9] B. Papanna, C. Lazzari, and M. Rabottini, "Huntington's disease prevalence in Asia: a systematic review and meta-analysis," *Riv. Psichiatr.*, vol. 59, no. 1, pp. 4–12, Jan. 2024.

[10] R. Kim *et al.*, "Analysis of *HTT* CAG repeat expansion among healthy individuals and patients with chorea in Korea," *Parkinsonism Relat. Disord.*, vol. 118, p. 105930, Jan. 2024, doi: 10.1016/j.parkreldis.2023.105930.

[11] "Huntington's Disease in Chile: Epidemiological and Genetic Aspects | Neuroepidemiology | Karger Publishers." Accessed: Jun. 28, 2024. [Online]. Available: https://karger.com/ned/article-abstract/57/3/176/837737/Huntington-s-Disease-in-Chile-Epidemiological-and?redirectedFrom=fulltext

[12] P. Harjes and E. E. Wanker, "The hunt for huntingtin function: interaction partners tell many different stories," *Trends Biochem. Sci.*, vol. 28, no. 8, pp. 425–433, Aug. 2003, doi: 10.1016/S0968-0004(03)00168-3.

[13] "Huntingtin Is Present in the Nucleus, Interacts with the Transcriptional Corepressor C-terminal Binding Protein, and Represses Transcription* - Journal of Biological Chemistry."

Accessed: Jun. 28, 2024. [Online]. Available: https://www.jbc.org/article/S0021-9258(19)82307-4/fulltext

[14] “Huntingtin’s spherical solenoid structure enables polyglutamine tract-dependent modulation of its structure and function | eLife.” Accessed: Jun. 28, 2024. [Online]. Available: https://elifesciences.org/articles/11184
[15] “Normal huntingtin function: an alternative approach to Huntington’s disease | Nature Reviews Neuroscience.” Accessed: Jun. 28, 2024. [Online]. Available: https://www.nature.com/articles/nrn1806
[16] Q. Guo *et al.*, “The cryo-electron microscopy structure of huntingtin,” *Nature*, vol. 555, no. 7694, pp. 117–120, Mar. 2018, doi: 10.1038/nature25502.
[17] “Huntingtin Interacts with a Family of WW Domain Proteins | Human Molecular Genetics | Oxford Academic.” Accessed: Jun. 28, 2024. [Online]. Available: https://academic.oup.com/hmg/article-abstract/7/9/1463/582598
[18] “A huntingtin-associated protein enriched in brain with implications for pathology | Nature.” Accessed: Jun. 28, 2024. [Online]. Available: https://www.nature.com/articles/378398a0
[19] M. A. Kalchman *et al.*, “HIP1, a human homologue of S. cerevisiae Sla2p, interacts with membrane-associated huntingtin in the brain,” *Nat. Genet.*, vol. 16, no. 1, pp. 44–53, May 1997, doi: 10.1038/ng0597-44.
[20] R. R. Singaraja *et al.*, “HIP14, a novel ankyrin domain-containing protein, links huntingtin to intracellular trafficking and endocytosis,” *Hum. Mol. Genet.*, vol. 11, no. 23, pp. 2815–2828, Nov. 2002, doi: 10.1093/hmg/11.23.2815.
[21] E. Cattaneo, C. Zuccato, and M. Tartari, “Normal huntingtin function: an alternative approach to Huntington’s disease,” *Nat. Rev. Neurosci.*, vol. 6, no. 12, pp. 919–930, Dec. 2005, doi: 10.1038/nrn1806.
[22] “The biological function of the Huntingtin protein and its relevance to Huntington’s Disease pathology - PMC.” Accessed: Jun. 28, 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3237673/
[23] M. Borrell-Pagès, D. Zala, S. Humbert, and F. Saudou, “Huntington’s disease: from huntingtin function and dysfunction to therapeutic strategies,” *Cell. Mol. Life Sci.*, vol. 63, no. 22, pp. 2642–2660, Nov. 2006, doi: 10.1007/s00018-006-6242-0.
[24] “Disturbed Ca2+ signaling and apoptosis of medium spiny neurons in Huntington’s disease | PNAS.” Accessed: Jun. 28, 2024. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.0409402102
[25] M. M. Zeron *et al.*, “Increased Sensitivity to N-Methyl-D-Aspartate Receptor-Mediated Excitotoxicity in a Mouse Model of Huntington’s Disease,” *Neuron*, vol. 33, no. 6, pp. 849–860, Mar. 2002, doi: 10.1016/S0896-6273(02)00615-3.
[26] “Expression analysis of novel striatal-enriched genes in Huntington disease | Human Molecular Genetics | Oxford Academic.” Accessed: Jun. 28, 2024. [Online]. Available: https://academic.oup.com/hmg/article/19/4/609/608845
[27] “Recent approaches on Huntington’s disease (Review).” Accessed: Jun. 28, 2024. [Online]. Available: https://www.spandidos-publications.com/10.3892/br.2022.1587
[28] H. Ismatullah, I. Jabeen, and M. T. Saeed, “Biological Regulatory Network (BRN) Analysis and Molecular Docking Simulations to Probe the Modulation of IP3R Mediated Ca2+ Signaling in Cancer,” *Genes*, vol. 12, no. 1, Art. no. 1, Jan. 2021, doi: 10.3390/genes12010034.
[29] K. Y.a, V. V.a, S. A.v, and K. E.v, “Huntington’s disease: calcium dyshomeostasis and pathology models,” *Acta Naturae Англоязычная Версия*, vol. 9, no. 2 (33), Art. no. 2 (33), 2017.

[30] M. Czeredys, “Dysregulation of Neuronal Calcium Signaling via Store-Operated Channels in Huntington’s Disease,” *Front. Cell Dev. Biol.*, vol. 8, Dec. 2020, doi: 10.3389/fcell.2020.611735.
[31] M. Giacomello, J. C. Oliveros, J. R. Naranjo, and E. Carafoli, “Neuronal Ca(2+) dyshomeostasis in Huntington disease,” *Prion*, vol. 7, no. 1, pp. 76–84, 2013, doi: 10.4161/pri.23581.
[32] K. M. Shannon, “Chapter 1 - Huntington’s disease – clinical signs, symptoms, presymptomatic diagnosis, and diagnosis,” in *Handbook of Clinical Neurology*, vol. 100, W. J. Weiner and E. Tolosa, Eds., in Hyperkinetic Movement Disorders, vol. 100. , Elsevier, 2011, pp. 3–13. doi: 10.1016/B978-0-444-52014-2.00001-X.
[33] R. A. Roos, “Huntington’s disease: a clinical review,” *Orphanet J. Rare Dis.*, vol. 5, no. 1, p. 40, Dec. 2010, doi: 10.1186/1750-1172-5-40.
[34] S. M. de la Monte, J.-P. Vonsattel, and E. P. Richardson Jr., “Morphometric Demonstration of Atrophic Changes in the Cerebral Cortex, White Matter, and Neostriatum in Huntington’s Disease,” *J. Neuropathol. Exp. Neurol.*, vol. 47, no. 5, pp. 516–525, Sep. 1988, doi: 10.1097/00005072-198809000-00003.
[35] S. W. Davies *et al.*, “Formation of Neuronal Intranuclear Inclusions Underlies the Neurological Dysfunction in Mice Transgenic for the HD Mutation,” *Cell*, vol. 90, no. 3, pp. 537–548, Aug. 1997, doi: 10.1016/S0092-8674(00)80513-9.
[36] I. Bezprozvanny and M. R. Hayden, “Deranged neuronal calcium signaling and Huntington disease,” *Biochem. Biophys. Res. Commun.*, vol. 322, no. 4, pp. 1310–1317, Oct. 2004, doi: 10.1016/j.bbrc.2004.08.035.
[37] C. Zuccato, M. Valenza, and E. Cattaneo, “Molecular Mechanisms and Potential Therapeutical Targets in Huntington’s Disease,” *Physiol. Rev.*, vol. 90, no. 3, pp. 905–981, Jul. 2010, doi: 10.1152/physrev.00041.2009.
[38] A. J. Milnerwood and L. A. Raymond, “Early synaptic pathophysiology in neurodegeneration: insights from Huntington’s disease,” *Trends Neurosci.*, vol. 33, no. 11, pp. 513–523, Nov. 2010, doi: 10.1016/j.tins.2010.08.002.
[39] A. J. Milnerwood and L. A. Raymond, “Corticostriatal synaptic function in mouse models of Huntington’s disease: early effects of huntingtin repeat length and protein load,” *J. Physiol.*, vol. 585, no. Pt 3, pp. 817–831, Dec. 2007, doi: 10.1113/jphysiol.2007.142448.
[40] “IJMS | Free Full-Text | Impact of ER Stress and ER-Mitochondrial Crosstalk in Huntington’s Disease.” Accessed: Jun. 28, 2024. [Online]. Available: https://www.mdpi.com/1422-0067/23/2/780
[41] V. Costa and L. Scorrano, “Shaping the role of mitochondria in the pathogenesis of Huntington’s disease,” *EMBO J.*, vol. 31, no. 8, pp. 1853–1864, Apr. 2012, doi: 10.1038/emboj.2012.65.
[42] C. Carmo, L. Naia, C. Lopes, and A. C. Rego, “Mitochondrial Dysfunction in Huntington’s Disease,” in *Polyglutamine Disorders*, C. Nóbrega and L. Pereira de Almeida, Eds., Cham: Springer International Publishing, 2018, pp. 59–83. doi: 10.1007/978-3-319-71779-1_3.
[43] S. Maity, P. Komal, V. Kumar, A. Saxena, A. Tungekar, and V. Chandrasekar, “Impact of ER Stress and ER-Mitochondrial Crosstalk in Huntington’s Disease,” *Int. J. Mol. Sci.*, vol. 23, no. 2, Art. no. 2, Jan. 2022, doi: 10.3390/ijms23020780.
[44] G. Szabadkai *et al.*, “Chaperone-mediated coupling of endoplasmic reticulum and mitochondrial Ca2+ channels,” *J. Cell Biol.*, vol. 175, no. 6, pp. 901–911, Dec. 2006, doi: 10.1083/jcb.200608073.

[45] M. Brondani *et al.*, "Mitochondrial dysfunction, oxidative stress, ER stress and mitochondria-ER crosstalk alterations in a chemical rat model of Huntington's disease: Potential benefits of bezafibrate," *Toxicol. Lett.*, vol. 381, pp. 48–59, May 2023, doi: 10.1016/j.toxlet.2023.04.011.
[46] L. Naia, I. L. Ferreira, E. Ferreiro, and A. C. Rego, "Mitochondrial Ca2+ handling in Huntington's and Alzheimer's diseases – Role of ER-mitochondria crosstalk," *Biochem. Biophys. Res. Commun.*, vol. 483, no. 4, pp. 1069–1077, Feb. 2017, doi: 10.1016/j.bbrc.2016.07.122.
[47] P. J. Kamitsuka, M. M. Ghanem, R. Ziar, S. E. McDonald, M. G. Thomas, and G. F. Kwakye, "Defective Mitochondrial Dynamics and Protein Degradation Pathways Underlie Cadmium-Induced Neurotoxicity and Cell Death in Huntington's Disease Striatal Cells," *Int. J. Mol. Sci.*, vol. 24, no. 8, p. 7178, Apr. 2023, doi: 10.3390/ijms24087178.
[48] "The microbiota–gut–brain axis in Huntington's disease: pathogenic mechanisms and therapeutic targets - Ekwudo - The FEBS Journal - Wiley Online Library." Accessed: Jun. 29, 2024. [Online]. Available: https://febs.onlinelibrary.wiley.com/doi/full/10.1111/febs.17102
[49] R. M. Stilling, T. G. Dinan, and J. F. Cryan, "Microbial genes, brain & behaviour - epigenetic regulation of the gut-brain axis," *Genes Brain Behav.*, vol. 13, no. 1, pp. 69–86, Jan. 2014, doi: 10.1111/gbb.12109.
[50] P. Alonso-García, R. Martín, and E. Martínez-Pinilla, "Gut microbial imbalance and neurodegenerative proteinopathies: from molecular mechanisms to prospects of clinical applications," *Explor. Neuroprotective Ther.*, vol. 1, no. 1, Art. no. 1, Aug. 2021, doi: 10.37349/ent.2021.00005.
[51] T. R. Sampson *et al.*, "Gut Microbiota Regulate Motor Deficits and Neuroinflammation in a Model of Parkinson's Disease," *Cell*, vol. 167, no. 6, pp. 1469-1480.e12, Dec. 2016, doi: 10.1016/j.cell.2016.11.018.
[52] L. Vitetta, S. Coulson, A. W. Linnane, and H. Butt, "The Gastrointestinal Microbiome and Musculoskeletal Diseases: A Beneficial Role for Probiotics and Prebiotics," *Pathogens*, vol. 2, no. 4, Art. no. 4, Dec. 2013, doi: 10.3390/pathogens2040606.
[53] M. Fjodorova, Z. Noakes, D. C. De La Fuente, A. C. Errington, and M. Li, "Dysfunction of cAMP–Protein Kinase A–Calcium Signaling Axis in Striatal Medium Spiny Neurons: A Role in Schizophrenia and Huntington's Disease Neuropathology," *Biol. Psychiatry Glob. Open Sci.*, vol. 3, no. 3, pp. 418–429, Jul. 2023, doi: 10.1016/j.bpsgos.2022.03.010.
[54] E. T. Koch, M. D. Sepers, J. Cheng, and L. A. Raymond, "Early Changes in Striatal Activity and Motor Kinematics in a Huntington's Disease Mouse Model," *Mov. Disord. Off. J. Mov. Disord. Soc.*, vol. 37, no. 10, pp. 2021–2032, Oct. 2022, doi: 10.1002/mds.29168.
[55] B. S. Khakh and S. A. Goldman, "Astrocytic contributions to Huntington's disease pathophysiology," *Ann. N. Y. Acad. Sci.*, vol. 1522, no. 1, pp. 42–59, Apr. 2023, doi: 10.1111/nyas.14977.
[56] G. P. Bates *et al.*, "Huntington disease," *Nat. Rev. Dis. Primer*, vol. 1, no. 1, pp. 1–21, Apr. 2015, doi: 10.1038/nrdp.2015.5.
[57] S. J. Tabrizi *et al.*, "Predictors of phenotypic progression and disease onset in premanifest and early-stage Huntington's disease in the TRACK-HD study: analysis of 36-month observational data," *Lancet Neurol.*, vol. 12, no. 7, pp. 637–649, Jul. 2013, doi: 10.1016/S1474-4422(13)70088-7.
[58] S. Frank, "Treatment of Huntington's Disease," *Neurotherapeutics*, vol. 11, no. 1, pp. 153–160, Jan. 2014, doi: 10.1007/s13311-013-0244-z.

[59] D. K. Stiles *et al.*, "Widespread suppression of huntingtin with convection-enhanced delivery of siRNA," *Exp. Neurol.*, vol. 233, no. 1, pp. 463–471, Jan. 2012, doi: 10.1016/j.expneurol.2011.11.020.
[60] J. B. Carroll *et al.*, "Potent and selective antisense oligonucleotides targeting single-nucleotide polymorphisms in the Huntington disease gene / allele-specific silencing of mutant huntingtin," *Mol. Ther. J. Am. Soc. Gene Ther.*, vol. 19, no. 12, pp. 2178–2185, Dec. 2011, doi: 10.1038/mt.2011.201.
[61] A. Jiang, R. R. Handley, K. Lehnert, and R. G. Snell, "From Pathogenesis to Therapeutics: A Review of 150 Years of Huntington's Disease Research," *Int. J. Mol. Sci.*, vol. 24, no. 16, Art. no. 16, Jan. 2023, doi: 10.3390/ijms241613021.
[62] X. García-González *et al.*, "Pharmacogenetics in the Treatment of Huntington's Disease: Review and Future Perspectives," *J. Pers. Med.*, vol. 13, no. 3, p. 385, Feb. 2023, doi: 10.3390/jpm13030385.
[63] "IJMS | Free Full-Text | New Avenues for the Treatment of Huntington's Disease." Accessed: Jul. 02, 2024. [Online]. Available: https://www.mdpi.com/1422-0067/22/16/8363
[64] M. Kanehisa and S. Goto, "KEGG: kyoto encyclopedia of genes and genomes," *Nucleic Acids Res.*, vol. 28, no. 1, pp. 27–30, Jan. 2000, doi: 10.1093/nar/28.1.27.
[65] "[PDF] Model Driven ActiveRecord with yEd | Semantic Scholar." Accessed: Jul. 03, 2024. [Online]. Available: https://www.semanticscholar.org/paper/Model-Driven-ActiveRecord-with-yEd-Sedlmeier-Gogolla/294c04471f4a65fb6a76c81472940d13e6f15d8b
[66] "Jimena: efficient computing and system state identification for genetic regulatory networks | BMC Bioinformatics." Accessed: Jul. 03, 2024. [Online]. Available: https://link.springer.com/article/10.1186/1471-2105-14-306
[67] A. Di Cara, A. Garg, G. De Micheli, I. Xenarios, and L. Mendoza, "Dynamic simulation of regulatory networks using SQUAD," *BMC Bioinformatics*, vol. 8, no. 1, p. 462, Nov. 2007, doi: 10.1186/1471-2105-8-462.
[68] R. P. D. Bank, "RCSB PDB - 6X9O: High resolution cryoEM structure of huntingtin in complex with HAP40." Accessed: Aug. 13, 2024. [Online]. Available: https://www.rcsb.org/structure/6x9o
[69] R. J. Harding *et al.*, "Huntingtin structure is orchestrated by HAP40 and shows a polyglutamine expansion-specific interaction with exon 1," *Commun. Biol.*, vol. 4, no. 1, pp. 1–16, Dec. 2021, doi: 10.1038/s42003-021-02895-4.
[70] H. Sanchez-Faddeev *et al.*, "Using Multiobjective Optimization and Energy Minimization to Design an Isoform-Selective Ligand of the 14-3-3 Protein," in *Leveraging Applications of Formal Methods, Verification and Validation. Applications and Case Studies*, T. Margaria and B. Steffen, Eds., Berlin, Heidelberg: Springer, 2012, pp. 12–24. doi: 10.1007/978-3-642-34032-1_3.
[71] "Prediction of a new surface binding pocket and evaluation of inhibitors against huntingtin interacting protein 14: an insight using docking studies - PubMed." Accessed: Jul. 09, 2024. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/21360185/
[72] "ChEMBL: a large-scale bioactivity database for drug discovery | Nucleic Acids Research | Oxford Academic." Accessed: Jul. 09, 2024. [Online]. Available: https://academic.oup.com/nar/article/40/D1/D1100/2903401
[73] "Open Babel: An open chemical toolbox | Journal of Cheminformatics." Accessed: Jul. 09, 2024. [Online]. Available: https://link.springer.com/article/10.1186/1758-2946-3-33

[74] M. Karplus and G. A. Petsko, “Molecular dynamics simulations in biology,” *Nature*, vol. 347, no. 6294, pp. 631–639, Oct. 1990, doi: 10.1038/347631a0.
[75] “GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers - ScienceDirect.” Accessed: Jul. 11, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352711015000059
[76] D. A. Case *et al.*, “The Amber biomolecular simulation programs,” *J. Comput. Chem.*, vol. 26, no. 16, pp. 1668–1688, Dec. 2005, doi: 10.1002/jcc.20290.
[77] “The Predicted Structure of the Headpiece of the Huntingtin Protein and Its Implications on Huntingtin Aggregation - ScienceDirect.” Accessed: Jul. 11, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022283609000801?pes=vor
[78] A. Tuladhar *et al.*, “Ions Tune Interfacial Water Structure and Modulate Hydrophobic Interactions at Silica Surfaces,” *J. Am. Chem. Soc.*, vol. 142, no. 15, pp. 6991–7000, Apr. 2020, doi: 10.1021/jacs.9b13273.
[79] K. Takemura and A. Kitao, “Effects of water model and simulation box size on protein diffusional motions,” *J. Phys. Chem. B*, vol. 111, no. 41, pp. 11870–11872, Oct. 2007, doi: 10.1021/jp0756247.
[80] M. Levitt, “Protein folding by restrained energy minimization and molecular dynamics,” *J. Mol. Biol.*, vol. 170, no. 3, pp. 723–764, Nov. 1983, doi: 10.1016/s0022-2836(83)80129-6.
[81] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, “Molecular dynamics with coupling to an external bath,” *J. Chem. Phys.*, vol. 81, no. 8, pp. 3684–3690, Oct. 1984, doi: 10.1063/1.448118.
[82] W. Dawson and F. Gygi, “Equilibration and analysis of first-principles molecular dynamics simulations of water,” *J. Chem. Phys.*, vol. 148, no. 12, p. 124501, Mar. 2018, doi: 10.1063/1.5018116.
[83] M. Parrinello and A. Rahman, “Polymorphic transitions in single crystals: A new molecular dynamics method,” *J. Appl. Phys.*, vol. 52, no. 12, pp. 7182–7190, Dec. 1981, doi: 10.1063/1.328693.
[84] M. Khaled, B. Strodel, and A. Sayyed-Ahmad, “Comparative molecular dynamics simulations of pathogenic and non-pathogenic huntingtin protein monomers and dimers,” *Front. Mol. Biosci.*, vol. 10, Apr. 2023, doi: 10.3389/fmolb.2023.1143353.
[85] E. Lindahl, B. Hess, and D. van der Spoel, “GROMACS 3.0: a package for molecular simulation and trajectory analysis,” *Mol. Model. Annu.*, vol. 7, no. 8, pp. 306–317, Aug. 2001, doi: 10.1007/s008940100045.
[86] L. Martínez, “Automatic Identification of Mobile and Rigid Substructures in Molecular Dynamics Simulations and Fractional Structural Fluctuation Analysis,” *PLoS ONE*, vol. 10, no. 3, p. e0119264, Mar. 2015, doi: 10.1371/journal.pone.0119264.
[87] M. Karplus and J. Kuriyan, “Molecular dynamics and protein function,” *Proc. Natl. Acad. Sci.*, vol. 102, no. 19, pp. 6679–6685, May 2005, doi: 10.1073/pnas.0408930102.
[88] S. N. Rai *et al.*, “The Role of PI3K/Akt and ERK in Neurodegenerative Disorders,” *Neurotox. Res.*, vol. 35, no. 3, pp. 775–795, Apr. 2019, doi: 10.1007/s12640-019-0003-y.
[89] “Kinase Signalling in Huntington’s Disease - IOS Press.” Accessed: Jul. 23, 2024. [Online]. Available: https://content.iospress.com/articles/journal-of-huntingtons-disease/jhd140106
[90] I. Bezprozvanny, “Calcium signaling and neurodegenerative diseases,” *Trends Mol. Med.*, vol. 15, no. 3, pp. 89–100, Mar. 2009, doi: 10.1016/j.molmed.2009.01.001.

[91] E. Kania, G. Roest, T. Vervliet, J. B. Parys, and G. Bultynck, “IP3 Receptor-Mediated Calcium Signaling and Its Role in Autophagy in Cancer,” *Front. Oncol.*, vol. 7, Jul. 2017, doi: 10.3389/fonc.2017.00140.
[92] K. H. Vousden and C. Prives, “Blinded by the Light: The Growing Complexity of p53,” *Cell*, vol. 137, no. 3, pp. 413–431, May 2009, doi: 10.1016/j.cell.2009.04.037.
[93] J. E. Chipuk, T. Moldoveanu, F. Llambi, M. J. Parsons, and D. R. Green, “The BCL-2 Family Reunion,” *Mol. Cell*, vol. 37, no. 3, pp. 299–310, Feb. 2010, doi: 10.1016/j.molcel.2010.01.025.
[94] “Old, new and emerging functions of caspases | Cell Death & Differentiation.” Accessed: Jul. 23, 2024. [Online]. Available: https://www.nature.com/articles/cdd2014216
[95] G. Babbi, C. Savojardo, P. L. Martelli, and R. Casadio, “Huntingtin: A Protein with a Peculiar Solvent Accessible Surface,” *Int. J. Mol. Sci.*, vol. 22, no. 6, p. 2878, Mar. 2021, doi: 10.3390/ijms22062878.

APPENDIX A

Table A.1: Curated Dataset of sixty nine ligands, along with their SMILES and LogpIC50 values, extracted from CHEMBL database and Literature.

Index	ChEMBL id	SMILES	IC50(logp)
1	CHEMBL2333942	COC(=O)c1ccc(C(=O)Nc2nc3ccc(C(=O)Nc4ccccc4)cc3s2)cc1	5.92
2	CHEMBL445226	CCN1CCN(c2ncnc3ccc(Cl)cc23)CC1	5.44
3	CHEMBL449571	CCCNc1ncnc2ccc(Br)cc12	5.08
4	CHEMBL153483	Clc1ccc(CCNc2ncnc3ccccc23)cc1	5.34
5	CHEMBL153483	Clc1ccc(CCNc2ncnc3ccccc23)cc1	4.97
6	CHEMBL4863336	Oc1cc(-c2c[nH]nc2F)ccc1-c1cnc(N2CCN[C@@H](C3CC3)C2)nn1	8.67
7	CHEMBL4749870	CN(c1cnc(-c2ccc(-n3nccn3)cc2O)nn1)C1CC(C)(C)NC(C)(C)C1	7
8	CHEMBL4761459	Oc1cc(-c2cn[nH]c2)ccc1-c1cc2sc(N3CCC34CCNCC4)nc2nn1	7
9	CHEMBL500972	Oc1ccc(CCNc2ncnc3ccc(Br)cc23)cc1	5.95
10	CHEMBL4850295	CC1(C)CC(n2nnc3cc(-c4ccc(-n5nccn5)cc4O)nnc32)CC(C)(C)N1	7
11	CHEMBL4797633	CN(c1nc2sc(-c3ncc(-c4cn[nH]c4)cc3O)nc2s1)C1CC(C)(C)NC(C)(C)C1.Cl	7
12	CHEMBL4754827	Cc1cn2nc(-c3cn4cc(C5CCNCC5)nc4s3)cc(C)c2n1.Cl	7
13	CHEMBL4743319	Cn1cc(-c2ccc(-c3cnc(NC4CC(C)(C)NC(C)(C)C4)nn3)c(O)c2)cn1	7
14	CHEMBL4867927	CCC1CN(c2ncc(-c3ccc(-c4cc(F)c5nc(C)cn5c4)cc3O)nn2)CCN1	8.65
15	CHEMBL4878942	Cn1nc2cc(-c3ccc(-c4cnc(N5CCN[C@@H](C6CC6)C5)nn4)c(O)c3)cnc2n1	8.76
16	CHEMBL494838	CCN1CCN(c2ncnc3ccc(Br)cc23)CC1	5.65
17	CHEMBL494325	C=CCNc1ncnc2ccc(Br)cc12	5.27
18	CHEMBL494326	Brc1ccc2ncnc(NCc3ccccc3)c2c1	5.21
19	CHEMBL493349	c1ccc2c(N3CCN(Cc4ccc5c(c4)OCO5)CC3)ncnc2c1	5
20	CHEMBL493350	Oc1ccc(CCNc2ncnc3ccccc23)cc1	4.80
21	CHEMBL4847156	CC1CN(c2ncc(-c3ccc(-c4cc(F)c5nn(C)cc5c4)cc3O)nn2)CCN1	8.67
22	CHEMBL4776322	CN(c1ncc(-c2ccc(-c3ncco3)cc2O)nn1)C1CC(C)(C)NC(C)(C)C1	7
23	CHEMBL493153	Oc1ccc(CCNc2ncnc3ccc(Cl)cc23)cc1	6.14
24	CHEMBL149800	Clc1ccc(CCNc2ncnc3ccc(Cl)cc23)cc1	5.46
25	CHEMBL4777225	Cl.Cn1cc2cc(-c3nn4cc(C5CCNCC5)nc4s3)cc(C#N)c2n1	7
26	CHEMBL4875017	Cn1cc2cc(-c3ccc(-c4cnc(N5CCNC(CO)C5)nn4)c(O)c3)cc(F)c2n1	8.79

27	CHEMBL5195936	CC1(C)CC(Oc2ccc(-c3ccc(-c4ccnc(O)c4)cc3O)nn2)CC(C)(C)N1	7.39
28	CHEMBL5175104	CCN1CCC(c2cc(C)c3nc(-c4ccc5nn(C)cc5c4)cc(=O)n3c2)CC1	7.09
29	CHEMBL495005	CN1CCN(c2ncnc3ccc(Br)cc23)CC1	5.88
30	CHEMBL509988	Brc1ccc2ncnc(N3CCNCC3)c2c1	5.56
31	CHEMBL451222	CCN1CCN(c2ncnc3ccccc23)CC1	5
32	CHEMBL489053	CN1CCN(c2ncnc3ccc(F)cc23)CC1	5
33	CHEMBL4778330	CN(c1nc2sc(-c3cnc(-c4cn[nH]c4)c4[nH]ccc34)nc2s1)C1CCNCC1.Cl	7
34	CHEMBL4799758	CC(C)(C)NC1CCN(c2ncc(-c3ccc(-c4c[nH]nc4C#N)cc3O)nn2)C1	7
35	CHEMBL4875823	CC1(C)C=C(c2csc3cc(-c4ccc(-c5cn[nH]c5)cc4O)nnc23)CC(C)(C)N1	7
36	CHEMBL4862205	CC1(C)CC(n2nnc3cc(-c4ccc(-c5ncco5)cc4O)nnc32)CC(C)(C)N1	7
37	CHEMBL4875396	COc1ncnc(-c2ccc(-c3cc4nnn(C5CC(C)(C)NC(C)(C)C5)c4nn3)c(O)c2)n1	7
38	CHEMBL4859053	Cc1cn2cc(-c3ccc(-c4cnc(N5CCN[C@@H](C(C)C)C5)nn4)c(O)c3)ncc2n1	8.63
39	CHEMBL4794728	Cc1cn2nc(-c3cc(F)c4nc(OC5CCN(C)CC5)sc4c3)cc(C)c2n1	7
40	CHEMBL4799079	CN(c1nc2nnc(-c3ccc(-c4c[nH]nc4F)cc3O)cc2s1)C1CCNC1	7
41	CHEMBL444229	CCN1CCN(c2ncnc3ccc(F)cc23)CC1	5.27
42	CHEMBL4761220	CN(c1nc2sc(-c3ncc(-c4cn[nH]c4)nn3)nc2s1)C1CC(C)(C)NC(C)(C)C1.Cl	7
43	CHEMBL442675	Fc1ccc2ncnc(N3CCN(Cc4ccc5c(c4)OCO5)CC3)c2c1	5
44	CHEMBL493324	CN1CCN(c2ncnc3ccc(Cl)cc23)CC1	5
45	CHEMBL4796814	C[C@H]1C[C@H](Oc2nc3sc(-c4ncc(-c5nnn(C)n5)c5cc[nH]c45)nc3s2)CCN1.Cl	7
46	CHEMBL4757451	Cc1nc2ccc(-c3cc(F)c4nc(C5CCNCC5)sc4c3)cc2o1	7
47	CHEMBL4785183	CN1CC2(C1)CN(c1nc3nnc(-c4ccc(-c5c[nH]nc5F)cc4O)cc3s1)C2	7
48	CHEMBL4798090	CC1(C)CC(Nc2cnc(-c3ccc(-c4cn[nH]c4)cc3O)nn2)C(F)C(C)(C)N1	7
49	CHEMBL692	OCC(O)CO	4.7
50	CHEMBL660	NC12CC3CC(CC(C3)C1)C2	5.32
51	CHEMBL61350	NCCSSCCN	2.81
52	CHEMBL940	NCC1(CC(=O)O)CCCCC1	6.85
53	CHEMBL54	O=C(CCCN1CCC(O)(c2ccc(Cl)cc2)CC1)c1ccc(F)cc1	7.39
54	CHEMBL807	CC12CC3CC(C)(C1)CC(N)(C3)C2	5.25
55	CHEMBL229128	Cc1ccccc1OCC(O)CO	4.69
56	CHEMBL13	COCCc1ccc(OCC(O)CNC(C)C)cc1	5.58
57	CHEMBL1232863	C(C(C(C(C(=O)CO)O)O)O)O	4.7
58	CHEMBL689	C(C(C(C(C(CO)O)O)O)O)O	2.02

59	CHEMBL44884	CCC(CO)NCCNC(CC)CO	4.60
60	CHEMBL1160714	CC(C)NCC(COC1=CC=C(C=C1)O)O	4.7
61	CHEMBL490	C1CNCC(C1C2=CC=C(C=C2)F)COC3=CC4=C(C=C3)OCO4	9.25
62	CHEMBL1201066	CN(C)CC(C1=CC=C(C=C1)OC)C2(CCCCC2)O.Cl	5.59
63	CHEMBL452861	CC(C)NCC(COC1=CC=CC2=CC=CC=C21)O	4.69
64	CHEMBL1200979	CC(C)(CO)C(C(=O)NCCCO)O	4.69
65	CHEMBL413	CC1CCC2CC(C(=CC=CC=CC(CC(C(=O)C(C(C(=CC(C(=O)CC(OC(=O)C3CCCCN3C(=O)C(=O)C1(O2)O)C(C)CC4CCC(C(C4)OC)O)C)C)O)OC)C)C)C)OC	5.36
66	CHEMBL45	COc1ccc2[nH]cc(CCNC(C)=O)c2c1	3.87
67	CHEMBL429694	Nc1c(/N=N/c2ccc(-c3ccc(/N=N/c4cc(S(=O)(=O)[O-])c5ccccc5c4N)cc3)cc2)cc(S(=O)(=O)[O-])c2ccccc12.[Na+].[Na+]	5.88
68	CHEMBL374478	CC1C=CC=C(C(=O)NC2=C(C(=C3C(=C2O)C(=C(C4=C3C(=O)C(O4)(OC=CC(C(C(C(C(C(C1O)C)O)C)OC(=O)C)C)OC)C)C)O)O)C=NN5CCN(CC5)C)C	5.63
69	CHEMBL51483	Cc1cc2c(C(C)C)c(O)c(O)c(C=O)c2c(O)c1-c1c(C)cc2c(C(C)C)c(O)c(O)c(C=O)c2c1O	5.23