

# Pravir Kumar

## Kirti MSC Thesis

 Aastha Kaushik

---

### Document Details

Submission ID

trn:oid:::27535:138652615

Submission Date

May 12, 2026, 2:23 PM GMT+5:30

Download Date

May 12, 2026, 2:25 PM GMT+5:30

File Name

Thesis plag check.docx

File Size

2.2 MB

27 Pages

9,037 Words

55,437 Characters

# 3% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

## Filtered from the Report

- ▶ Bibliography
- ▶ Quoted Text
- ▶ Cited Text
- ▶ Small Matches (less than 8 words)

## Match Groups

- 26 Not Cited or Quoted 3%**  
 Matches with neither in-text citation nor quotation marks
- 0 Missing Quotations 0%**  
 Matches that are still very similar to source material
- 0 Missing Citation 0%**  
 Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**  
 Matches with in-text citation present, but no quotation marks

## Top Sources

- 2% Internet sources
- 2% Publications
- 1% Submitted works (Student Papers)

## Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

### Match Groups

- 26 Not Cited or Quoted 3%**  
Matches with neither in-text citation nor quotation marks
- 0 Missing Quotations 0%**  
Matches that are still very similar to source material
- 0 Missing Citation 0%**  
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**  
Matches with in-text citation present, but no quotation marks

### Top Sources

- 2% Internet sources
- 2% Publications
- 1% Submitted works (Student Papers)

### Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

<b>1</b>	Internet	www.mdpi.com	<1%
<b>2</b>	Internet	www.nature.com	<1%
<b>3</b>	Internet	www.pharmacompass.com	<1%
<b>4</b>	Publication	Nicholas S. Caron, E. Ray Dorsey, Michael R. Hayden. "Therapeutic approaches to ...	<1%
<b>5</b>	Internet	www.aladdin-e.com	<1%
<b>6</b>	Internet	dspace.dtu.ac.in:8080	<1%
<b>7</b>	Internet	en.wikipedia.org	<1%
<b>8</b>	Publication	Andrew M Tan, Michal Geva, Y Paul Goldberg, Henk Schuring et al. "Antidopamin...	<1%
<b>9</b>	Internet	www.frontiersin.org	<1%
<b>10</b>	Publication	Jeremy Veenstra-VanderWeele, George M Anderson, Edwin H Cook. "Pharmacoge...	<1%

11	Publication	Borodin, Oleg, Grant D. Smith, Rajdip Bandyopadhyaya, and Oleksiy Byutner. "Mo..."	<1%
12	Student papers	De Montfort University on 2022-04-05	<1%
13	Student papers	Universiti Malaysia Sabah on 2021-12-17	<1%
14	Student papers	University College London on 2021-09-01	<1%
15	Internet	pubmed.ncbi.nlm.nih.gov	<1%
16	Internet	www.aginganddisease.org	<1%
17	Internet	www.benthamscience.com	<1%
18	Publication	Ashok Kumar, Vijay Kumar, Kritanjali Singh, Sukesh Kumar, You-Sam Kim, Yun-Mi ...	<1%
19	Student papers	University of Reading on 2016-12-02	<1%
20	Internet	brieflands.com	<1%

# Virtual identification and screening of novel VMAT2 inhibitors as therapeutic candidates for Huntington's disease

Kirti  
24/MSCBIO/64

## ABSTRACT

The pathology underlying Huntington's disease (HD) results from an unusual polyglutamine tract attributed to a mutation on the short arm of chromosome number 4, in the huntingtin gene, leading to hyperkinetic movement and psychiatric dysfunctions. Presently, HD treatment is palliative in nature and major symptom i.e., chorea is managed by use of Anti-Dopaminergic Medications (ADMs) comprising of Vesicular Monoamine Transporters (VMATs) inhibitors and antipsychotic medications. A subset of VMAT inhibitors specifically Vesicular Monoamine Transporter Type 2 (VMAT2) assist in lowering amount of vesicular monoamines, for instance, dopamine and block its release from pre-synaptic vesicles. As a result, dopamine fails to reach upregulated D2 receptors and therefore reduces chorea. In this study, Valbenazine (VBZ), an FDA approved VMAT2 inhibitor in 2023 for HD mediated chorea is chosen as reference pertaining to its higher efficacy, longer serum shelf-life, safety and reduced psychiatric manifestations as compared to other VMAT2 inhibitors namely Tetrabenazine (TBZ) and Deutetrabenazine (DBZ). The structures similar in conformation to Valbenazine are selected for molecular docking and identifying potential alternatives to VBZ. The potential ligands were also assessed by ADME analysis. The compound CID 163809280 exhibited the strongest interaction with VMAT2, yielding a binding energy of -10.8 kcal/mol suggesting highest favourable binding among the tested ligands.

**Result:** Initial screening of chemical structures retrieved from PubChem yielded about 1,156 structures, which were further filtered on the basis of structural similarity to the lead drug candidate Valbenazine (VBZ). Out of the shortlisted nine compounds, docking analysis revealed that compound CID 163809280 exhibited highest binding energy of -10.8 kcal/mol.

**Conclusion:** ADMET analysis of compound CID 163809280 (binding energy of -10.8 kcal/mol) highlighted its drug-likeness capability and BBB permeability rendering its potential use as a VMAT2 inhibitor. In subsequent studies, experimental determination of binding affinity is recommended to validate these in silico results.

## 1. INTRODUCTION

Huntington's disease (HD), known for its absolute neuropenetrative nature is a widely recognized neurodegenerative disorder. It is the most prevalent autosomal genetic abnormality inherited in dominant pattern and is affecting approximately 10-12 persons per 100,000 globally [1],[2]. George Huntington documented hereditary nature of chorea and concomitant psychiatric and cognitive symptoms that emerge amid age of 30 and 40 years, presently called as Huntington's disease [2],[3]. This monogenic disease occurs due to the pathogenic amplification of cytosine, adenine and guanine (CAG)<sub>n</sub> repeats in exon 1 on the short arm of huntingtin gene (HTT) on chromosome 4p16.3. It later manifests itself with formation of

mutant huntingtin (mHTT) protein containing elongated polyglutamine tract which also serves as a reliable biological predictor in assessing risk and severity of HD. The certainty of disease manifestation depends on penetrance, the normal allele possesses <27 CAG repeats but if the repeats are greater than 40 then the HD will develop with complete penetrance and if ranges from 36-39 repeats there is less penetrance but HD still occurs. Also, earlier onset corresponds to higher disease intensity [3],[4].

Glutamine (Q) encoded by the CAG codon, is produced locally within the lungs, muscles and brain from its precursors glutamate and ammonia via the action of glutamine synthetase enzyme. The HTT gene has glutamine embedded in CAG which by itself is non toxic but when there is polyglutamine expansion the aggregate formation occurs leading to toxicity and secondary issues such as mitochondrial dysfunction (free radicals abundance and oxidative stress markers), inflammatory reactions (imbalanced cytokine and nitric oxide levels), excitotoxicity, nuclear cleavage, transcriptional irregularities and apoptosis. Expanded CAG repeats amount to about 70% of the variation of HD and rest 13% emerges because of polymorphisms in the GRIK2 gene [4].

HD is diagnosed by a positive genetic test or emergence of symptoms pertaining to motor disability which is well stated in Total Motor Score (TMS). The TMS score ranging from 0 which is indicative of no motor disturbances pertaining to HD to maximum of 4, which is suggestive of manifestation of HD. The mutant huntingtin exhibits multi-pronged ramifications such as neuronal dysfunction and apoptosis, some of which are direct effects which includes formation of abnormal protein aggregates due to exon 1 of mHTT fragment. These abnormal protein aggregates have the tendency of causing deleterious effects on axonal transport, proteostasis, gene expression pathways viz., transcription and translation along with major disruptions in functions of mitochondria and synapse. Medium Spiny Neurons (MSNs) are predominantly affected by outcomes of mHTT. The damage to striata occurs in two phases, early degeneration of indirect pathway basal ganglia MSNs thus resulting in hyperkinetic phenotype i.e., chorea and in later phase occurs loss of direct pathway MSNs thereby leading to a hypokinetic/rigid phenotype. HD pathogenesis has been hypothesized by expression of dopamine D2 receptors by indirect instead of direct MSNs, other reasons can be loss of Brain Derived Neurotrophic Factor (BDNF), loss of pyramidal neurons, glutamate induced-neurotoxicity arising from projections at cortico-striatal region and harmful outcomes by translated proteins of repeat associated non ATG sequence [2],[3].

It is marked by chronological deterioration of motor function, behavioral disorder (anxiety, depression, psychosis, anosognia and OCD) and cognitive symptoms culminating to mortality [1,4]. Neurotoxicity of mHTT causes chorea i.e., involuntary muscle movements, incoordination and rigidity, eventually resulting in atrophy of brain particularly at striatum, thalamus, cerebellum, brain stem and cortex. Presently, treatment of HD is palliative in nature and planned to control symptoms as underlying etiological processes aren't fully understood yet [4].

Primarily, the Vesicular Monoamine Transporters (VMATs) facilitate uptake of serotonin (5-hydroxytryptamine or 5-HT), dopamine, epinephrine, histamine and norepinephrine (monoamines) concentrated at axon terminal for their subsequent release in synaptic cleft. VMATs (55kDa) belong to major facilitator superfamily (MFS) and exist in two isoforms namely VMAT1 (solute carrier18A1 or SLC18A1) and VMAT2 (solute carrier18A2 or SLC18A2). VMAT2 is sequestered at vesicular membranes in presynaptic axon terminals and is composed of a cytosolic C-terminus and N-terminus, 12 transmembrane domains. VMAT2 utilizes electrochemical gradient and involves transfer of two protons from lumen to cytoplasm

2

produced by H<sup>+</sup>-ATPase antiporter. Additionally, VMAT2 shields the neurons from intoxicants like methamphetamine and MPP<sup>+</sup> [5],[6].

This study aims to identify novel VMAT2 inhibiting ligands structurally similar to Valbenazine, an FDA -approved and widely researched VMAT2 inhibitor. Furthermore, the resultant ligand profiles were validated both quantitatively and qualitatively by evaluating binding affinity values and by comparing ADMET properties of effective leads.

## 2. LITERATURE REVIEW

### 2.1 Huntington's Disease: Epidemiology and overview

8 Huntington's disease (HD) is an incremental neurodegenerative disorder and marked by triad of motor, psychiatric and cognitive ramifications. It is generationally passed down in an autosomal dominant mode i.e., if a parent is heterozygous with amplified CAG trinucleotide expansion in exon 1 of HTT gene at chromosome 4p16.3, then the progeny has a 50% chance of receiving amplified CAG expansion trait. The probability of occurrence of HD differs between usual interrupted repeat sequence i.e., [(CAG)<sub>n</sub>-CAA-CAG] spanning over 95% of alleles and rarely continuous repeat of sequence (CAG)<sub>n</sub> exhibited by approximately 1% alleles in symptomatic individuals. The length of repeats of CAG can expand or contract either due to paternal transmission or maternal transmission. Progeny that inherits allele with lesser penetrance (36-39 repeats of CAG) are prone to develop HD but may not manifest symptoms. In contrast, progeny inheriting allele with absolute penetrance (≥40 repeats of CAG) is certain to develop symptomatic HD [7]. In adult HD, CAG repeats vary from 36 to 55. Juvenile HD, characterized by cortical myoclonus and CAG repeats exceeding fifty five, occurs due to early onset of clinical features by age of 20 years, generally paternally transmitted and ranges from 1% to 15% of persons diagnosed with HD. They are known to have Westphal variant HD which is an akinetic rigid syndrome. Similarly, late onset HD in about 4.5% to 11.5% HD patients is also rare in occurrence and manifests after the age of 60 years. Its progression rate is rapid and duration is shorter (8-12 years) in comparison to adult onset HD [8].

Medium Spiny Neurons in basal ganglia (set of subcortical nuclei features connected to each other in centre of hemispheres to cortex thalamus and brainstem; comprises of striatum, globus pallidus, subthalamic nucleus and substantia nigra) brain are most damaged tissue in HD [9]. Other heritable factors may influence variability accounting upto 10 to 20%, for instance, loss of interruption variants, especially CAG-CCG repeat LOI variant which is related with speedy degeneration of motor and cognitive features, PMS1 at chromosome 2 gets affected, MLH1 at chromosome 3, FAN1 at chromosome 11 etc.[10]. Incremental atrophy of caudate nucleus and putamen bilaterally is a major neuropathological hallmark of HD. It is due to targeted deterioration of GABAergic medium spiny neurons and specific stimulatory neurons of cerebral cortex. In affected individuals, region-specific loss of neurons in cortex and basal ganglia may underlie morphological variability in manifestation of symptoms. Inclusion structures in between neurons consisting huntingtin, encoded by HTT, are also a remarkable feature of HD. Many theories have been proposed and adopted regarding HD pathology such as dysfunction of mitochondria, which suggests that mHTT causes mitochondrial malfunction due to compromised calcium regulation and scarce ATP generation. The second theory, namely, proteolysis theory supposes that cytotoxic effect and neuronal degeneration are a result of cleavage of HTT protein by caspases or other proteases

in cytoplasm and fragments are translocated into nucleus. Glutamate induced excitotoxicity of neurons causes apoptosis of neuronal cells [11] [12].

3

One of the most pronounced symptom of HD is chorea i.e, non recurring, non rhythmic involuntary abnormal movements characterized by jerking of face and limbs. These worsen within fifteen years of HD manifestation and cannot be suppressed voluntarily but tames during sleep. The selective deterioration of enkephalin possessing MSNs in indirect pathway of basal ganglia (responsible for locomotion) occurs before the loss of substance P-containing MSNs of direct pathway. The depletion of such neurons is hypothesized as causation of chorea [7]. Consequently, complete degeneration of MSNs (hosts D1 receptors) of direct pathway comprising substantia nigra and internal Globus Pallidus (GP) expresses akinetic symptoms manifested in later stages of HD [13].

The epidemiology of HD varies over generations with CAG length varying among different ethnic population worldwide. Globally, its frequency is 3.92 for every 100,000 individuals [14]. It is mostly prevalent in European population approximately affecting 17 people per 100,000 individuals with allele count 36 or more of CAG repeat tract present in every one individual per 400 persons of European lineage. Averagely, CAG repeat length ranges between 18.4 and 18.7 among people of European lineage which is more as compared to East Asian cohort (17.5 to 17.7). It is less prevalent in East Asian population affecting only one or two people per lakh individuals, specific reports citing 0.65 in Japan and 0.42 in China per lakh. Similar trend is observed in African population where only about two per lakh population is affected. Though, Finland has remarkably low penetration of HD in European population. The rate of HD affected people are exhibiting an increasing trend in regions of North America, Europe and Asia pertaining to expansion in life expectancy. Interestingly, in populations near Lake Maracaibo in Venezuela and Tanzania, migration has resulted in increased prevalence of HD allele via genetic drift i.e., HD carriers creates offsprings with people in an isolated community [7], [9].

**Symptoms:** In order to assess phenotypic manifestations of HD, Unified Huntington Disease Rating Scale (UHDRS), covering six different parameters including cognitive motor, behavioral, independence, function and Total Functional Capacity (TFC). Also, latest being Huntington Disease Integrated Staging System (HD-ISS) which assesses progression of HD via motor diagnosis from infancy. Median age of symptom visibility is about 45 years. [7],[15].

The progressive phases of HD can be broadly stratified into:

**Pre-manifest phase** comprising a pre-symptomatic phase i.e., an asymptomatic period of a person with an HD-associated allele, usually childhood till adolescence but is clinically at risk and a pro-dromal stage in which onset of HD is visibly marked by slight changes in coordinations and movement of eyes, olfactory malfunction, speech or swallowing difficulties, challenges in decision-making, mild involuntary movements accompanied with depression and irritable mood. Pro-dromal stage can't be detected by TFC but can be captured by brain imaging [7].

**Manifest:** with time, symptoms intensify as distinguished by dystonia (abnormal posture), dysarthria, intensified chorea, dysphagia, bradykinesia, rigidity, severe weight loss pertaining to malfunctioned metabolism of cholesterol (essential for transmission function by synapse and neuronal health), gait abnormality, no control over intensity of movement. Psychological manifestations worsen including belligerent behaviour and poor self control. These are comorbid with emergence of motor symptoms [7]. Most prevalent psychiatric outcome is depression which is co-related with dysfunction in prefrontal cortex and basal ganglia. Others include changes in circadian rhythm due to untimely melatonin secretions thus can cause

insomnia. The affected individual survives averagely next 15 to 18 years. The TFC score gives an indication of deterioration of symptoms starts to decline [14], [15].

### 2.1.1 Established Diagnostic Methods

HD can be narrowed down in an individual by looking into family history, neuroimaging with distinguishing clinical findings from a proband and a CAG repeat expansion in HTT's exon1 characterized by a positive genetic test. HD-associated alleles have  $\geq 36$  repeats of CAG and further classified as reduced penetrance HD-associated alleles and complete penetrance HD-associated alleles [16].

Genetic testing: primarily small-read amplicon based HTT sequencing and exome and genome based procedure are adopted in research context. Other techniques such as whole genome sequencing (WGS), sequence-based multiple gene panels and long-read sequencing have also been applied in detecting pathogenic HD-associated alleles [17].

**Clinical Reports:** incremental disability in locomotion with chorea along with affected voluntary actions, psychiatric deterioration including emergence of depression, alterations in personality and declining cognitive ability [17].

**Neuroradiologic imaging:** incremental atrophy of striatum especially caudatum and putamen which can also function as a biomarker, gray and white matter deterioration accompanied with lateral ventricles enlargement. Later stages reveal thinning of cortex region especially motor, prefrontal, occipital, dorsomedial and parietal regions.

Familial predisposition: affected males and females span over multiple generations.

**Molecular biomarker detection:** these can be used to assess physiological changes associated with HD in the central nervous system (CNS). NEFL viz., neuron specific intermediate filament protein, is a substantiated biomarker indicating neurodegeneration. The concentration of NEFL in blood and cerebrospinal fluid (CSF) increases with HD progression. The levels of prodynorphin and proenkephalin are reduced in CSF and indicates condition of striatal MSNs. An astroglial activation marker, GFAP or Glial fibrillar acidic protein in plasma is present in high amounts directing severity of HD. Mutant huntingtin or mHTT assays quantify mHTT levels in brain which pertains to phase of HD [7], [17].

### 2.1.2 Recent pharmacological and non-pharmacological approaches to cater HD Symptoms

**Neuroleptic medications:** Presently, there is no complete cure for HD. Medications target either psychological manifestations or physical symptoms of HD. Anti-psychotics are speculated to inhibit chorea via blocking D2 receptors. For instance, atypical and typical neuroleptics are delivered for taming aggressive behaviour and an irritable mood, amantadine (poor NMDA receptor blocker that lowers dyskinesias in HD by increasing dopamine release without inducing parkinsonism), methylphenidate, atomoxetine, modafinil, bupropion and bromocriptine are administered for curing apathy, carbamazepine and lamotrigine (anti-epileptic drug blocks voltage gated Na<sup>+</sup> channels) are prescribed for stabilizing mood, zopiclone, a sedative drug is used for reducing disturbances related to sleep, for insomnia drug named mirtazapine is advised and lastly few specific inhibitors of serotonin uptake are administered for countering anxiety, depression, irritability and obsessive-compulsive disorder. Venlafaxine, duloxetine and desvenlafaxine are administered to HD patients

exhibiting Parkinsonian symptoms as well. Selisistat, an SirT1 inhibitor which cleaves acetyl groups and mHTT on proteins, improves total motor score on administration [4][7][18].

**Anti-excitotoxic drugs:** riluzole, memantine and tetrabenazine [19].

**Drugs targeting Huntingtin Proteolysis:** minocycline, inhibits caspase-1 and capase-3 expression, resulting in reduced chorea [19].

**Drugs inducing HTT aggregate degradation:** Congo red and Trehalose dye, Compound C2-8 and Rapamycin (inhibits mTOR and facilitates autophagy) [20].

**Therapeutic agents targeting mitochondrial dysfunction:** creatine (anti-oxidant reduces 8-hydroxy-20-dexoguanosine in HD pateints), Coenzyme Q10 cofactor (supplementation showed neuroprotective effect and delayed atrophy), Eicosapentaenoic acid (inhibits caspase and reduces action of c-Jun N-terminal kinase (JNK) pathway) and meclizine drug [21].

**Reagents targeting dysregulation of transcription:** sodium phenylbutyrate, HDACi4b (histone deacetylase inhibitor), suberoylanilide hydroxamic acid, mithramycin and chromomycin (GC rich DNA binding antibiotic) [22].

### Nucleic acid therapeutic approach

It utilizes gapmer anti-sense oligonucleotide (ASO) to reduce huntingtin. The ASO bound to target mRNA can recruit RNaseH at unmodified central area, resulting in degradation of transcript [23]

**FAN1 gene targeted therapy:** FAN1 gene functions to cleave DNA during repair mechanism of crosslinking of the two DNA strands, shows its ability as a gene modifier. It is shown to repair loopouts which arose at CAG repeats [17], [24].

**Zinc Finger Protein:** reduces mutant protein expression by inhibiting transcription of mHTT. ZFPs bind to extended CAG repeats and reduces mHTT level without changing the gene [25].

**CRISPR-Cas9 therapy:** excised CAG repeat to correct HD alleles or inactivate alleles associated with HD or target HTT gene, resulting in reduction of mHTT [25].

**RNA interference (RNAi) and antisense oligonucleotide (ASO)** blocks transcription of mutant huntingtin. An example of ASO, Tominersen, which ligates to wild type HTT and alongwith mHTT mRNA thus initiating degradation. Synthetic peptides and intrabodies, target proline rich domains of HTT. Novel viral vectors use such as, AAV1, AAV5, AAV9, AAV-PHP.B and CREATE. **AAV-encoded miRNA induces constant suppression of mHTT in caudate and putamen** [26].

### Others

**Small molecule administration:** Ubiquilin, lowers mHTT aggregation when over expressed in hippocampus and cortex. Chaperonins, for instance, TRiC (CCT1-CCT8 subunit) reduces mHTT aggregation by lowering inclusions count, number of fibrillar oligomers and mHTT fragments. BN82451 ireduces release of glutamate by blocking Na<sup>+</sup> channels. It exhibited enhanced motor function and lowered brain atrophy. It is under phase II clinical trial [27].

**Focused Ultrasound (FUS):** it is a non invasive and focused brain delivery mechanism. It has assisted in enhancing efficacy of siRNA, adeno-associated virus (AAV) vector-based gene therapy and glial cell-line derived neurotrophic factor (GDNF). It works by transiently affecting internal system of blood-brain barrier (BBB) and its endothelial cells.

Administration of microbubbles and subsequent sonication of target region produces transient

18

4

15

5

openings in endothelial cells of BBB, as a result augmenting permeability and assists drug delivery in a limited time [12], [28].

**Stem Cell Therapy:** new neurons have the capability to replace the degenerating cells at diseased brain regions. C17.2 neuronal stem cells have been utilized in neurodegenerative disorder models and they take up neuronal morphology [24], [28].

**Deep Brain Stimulation (DBS):** at globus pallidus relieves symptoms of chorea. A neurostimulator surgically implanted at midbrain changes electrical and neurological action of neurons near input electrode [28].

In order to cure adverse physical outcomes of HD, a course of anti-dopaminergic drugs like tetrabenazine (TBZ), deutetabenazine (DBZ) and valbenazine (VBZ) are prescribed. Symptom specific drugs include sodium valproate or levetiracetam for tackling hyperkinesia. Chorea specific treatment comprises use of typical neuroleptics like sulpiride and atypical neuroleptics like olanzapine, aripiprazole, risperidone and quetiapine may reduce choreic movements temporarily. Risperidone exhibits higher affinity towards D2, D3 and 5HT<sub>2A</sub> receptor blockade. Use of anti-psychotics is majorly beneficial for psychiatric conditions but lacks consensus towards use for chorea [29].

**VMAT2 inhibitors:** these basically exhausts the vesicular monoamines and blocks their outflow from pre-synaptic vesicles. anti-dopaminergic drugs including valbenazine (in lower dosage) can suppress chorea but should be discontinued as soon as side effects emerges [7].

Also, drugs containing levodopa or its constituents and drugs inhibiting monoamine oxidase must be avoided as they may contribute to enhance chorea. Non-pharmacological therapies, for instance, physical therapy to address imbalance and gait issues, mental counselling, cognitive behavioural therapy, swallow and speech-language therapy are employed. Positive behavioural changes like abstinence from smoking and alcohol, mild physical exercise must be encouraged [30], [31].

6

## 2.2 VMATs biology and pharmacology- role in dopamine homeostasis

The Vesicular Monoamine Transporters (VMATs), constitute a group of transporter proteins that function in transporting monoamines namely serotonin (5-HT), dopamine (DA), norepinephrine (NE) and histamine from cytoplasm to vesicle and are located on the secretory vesicles of neuronal membrane, mast cells, platelets and neuroendocrine cells. They're classified as active transporters of secondary type and their functioning relies on proton-gradient generated by vesicular H<sup>+</sup>-ATPase's proton pump, which pumps proton into the vesicle against their concentration gradient. As per accepted present model, inflow of each monoamine into vesicle from cytosol requires outflow of two protons to cytosol from vesicle membrane. Monoamine-mediated transmission via synapse require VMATs which facilitate concentration, storage, exocytosis (release of monoamines when synaptic vesicle fuses with plasma membrane, regulated by Ca<sup>2+</sup> availability in cytosol) from synaptic vesicles. Action potential produced by neurons mediate exocytosis, and when it develops near the terminal of axon, it causes depolarization of membrane potential that triggers voltage-sensitive Ca<sup>2+</sup> channels and induces inflow of Ca<sup>2+</sup>[32], [33].

Primarily, there exist two isoforms of VMATs based on their distribution in tissues and binding preference towards molecules and inhibitors, namely, Vesicular Monoamine Transporter Type 1 (VMAT1) and Vesicular Monoamine Transporter Type 2 (VMAT2). VMAT1 is sequestered in membrane of neuroendocrine cells especially enterochromaffin and chromaffin granules. VMAT2 is sequestered in membrane of platelets and islet  $\beta$ -cells of

16

10

pancreas. There is 62% sequence identity in both VMAT1 and VMAT2 sequences but difference lies in characteristic specificity towards substrate and pharmacological features. Serotonin has almost identical affinity binding towards both VMAT1 and VMAT2 but monoamines dopamine, norepinephrine, epinephrine have 3 times potent affinity towards VMAT2. Histamine exhibits 30 times higher potency towards VMAT2. Both VMAT1 and VMAT2 are constituents of solute carrier 18 family (SLC18) and their structure basically comprise of 12 transmembranes (TM) orderly arranged as halves exhibiting pseudo-symmetry containing six helices each (N-terminal domain, TMs 1-6 and C-terminal domain, TMs 7-12) with a central binding region for attachment of neurotransmitters, inhibitors and polyamines. Both encode a huge luminal loop (LL) which comprises of multiple N-linked glycosylation regions and between LL1 and LL4 is present a disulfide bridge [32], [34].

### 2.2.1 VMAT2 structure and inhibitor binding

The dysregulation of dopamine leads to either dopaminergic hyperactivity in the form of hyperkinesia (excessive involuntary movement) eg., tremors, myoclonus tics, dystonia and chorea or dopaminergic hypoactivity. VMAT2 has a crucial role in presynaptic dopamine release and its dysregulation generates a hyperdopaminergic state leading to movement disorders. By suppressing the recycling of vesicular dopamine, VMAT2 inhibitors lower synaptic dopamine at striatal terminal, thereby ameliorating hyperdopaminergic drive that causes chorea in HD. This approach restores balance across indirect basal ganglia and direct basal ganglia pathways by reducing D2 receptor overstimulation thus emerging as a promising strategy for targeting movement disorders. Driven by cytosol-directed H<sup>+</sup> gradient, VMAT2 unlocks its cytosolic gate and gears for entry of monoamine. When a substrate binds, it triggers coordinated movement of TMs towards lumen of vesicle. At low-pH (pH 5.5) physiological conditions, residues Glu312 and Asp399 gets protonated in a sequential manner to aid movement of substrate in translocation funnel and release of substrate. Later, VMAT2 is again reset to its cytosol-open configuration [35].

### 2.3 VMAT2 inhibitors- pharmacology, mechanism and limitations

VMAT2 inhibitors (FDA approved) and antipsychotics are subset of antidopaminergic medications (ADMs) commonly employed to control HD motor symptoms specifically chorea and its behavioral manifestations respectively. Managing chorea associated with HD is an attempt to enhance longevity, well-being and elevate productivity of patients suffering from HD. Primarily, VMAT2 inhibitors block the levels of monoamine (particularly dopamine) thus restricting ligation with receptors on post-synaptic neuron. VMAT2 inhibitors function in similar lines with dopamine antagonists and deplete presynaptic dopamine at striatal nerve terminals via cytosolic monoamine oxidase. This blocks dopamine from going towards activated D2 receptors. As a result, restoring balance in direct and indirect pathways and managing motor symptoms in hyperkinetic disease. ADMs influence cognition and function measurements as tested via Total Functional Capacity (TFC) and latest being Huntington's Disease Integrated Staging System. Higher the TFC the better the function and independence [1], [13]. Declining TFC scores are applied in clinical trials to assess deterioration in HD.

FDA approved VMAT2 inhibitor drugs namely, Tetrabenazine (Xenazine), Deutetrabenazine (Austedo) and Valbenazine (Ingrezza) are used to treat hyperkinetic movements viz., chorea in HD. Tetrabenazine (TBZ), a synthetic selective reversible inhibitor of VMAT2 works by inhibiting the dopamine pathway through VMAT2. Deutetrabenazine (DBZ) which is a deuterated form of TBZ also inhibits VMAT2 and has a longer half-life (9-10hours). Both TBZ

and DBZ reduce central monoamines by reversible inhibition of VMAT2. The downside is that TBZ possesses shorter serum half-life and resulted in side effects such as somnolence and suicidal tendency. Currently, the biological implications in CNS by these ADMs is still pursued as an area of research, but a consensus has been achieved of the fact that ADMs modify dopamine signaling in HD pathology. About 30%-50% initial stage HD patients receive at least one ADM prescription which exceeds upto 58.9% in later stages of HD, usually ingesting more than one ADM [1], [2], [3], [36].

### 2.3.1 Tetrabenazine (TBZ) or Xenazine

5 Tetrabenazine ( $C_{19}H_{27}NO_3$  or IUPAC name: 9,10-dimethoxy-3-(2-methylpropyl)-1,3,4,6,7,11b-hexahydrobenzo[a]quinolizin-2-one), is informally known as progenitor molecule from which latest VMAT2 inhibitors are produced. It was first formulated in 1950 for curing psychosis but pertaining to its similarity towards reserpine, it was applied for ameliorating hyperkinetic locomotive disorders. In the year 2008, it got approval of FDA for HD-chorea treatment. TBZ inhibition effect at VMAT2 occurs via two consecutive steps, at lumenal-open configuration of VMAT2 there is low affinity attachment of TBZ which then results in high-affinity TBZ-bound occluded state facing lumen has been validated by cryo-EM experimental data 1. This configuration has rendered TBZ inaccessible from any of the two sides of transmembranes. TBZ rests inside hydrophobic central cavity produced by TMs 1,4,7 and 20 along with large aromatic groups surrounding TBZ through C-domain and smaller non-polar groups at N-domain. It promotes an off-cycle and dead end configuration via luminal entrance. It shows selectivity in inhibiting VMAT2 and not VMAT1. TBZ use in human clinical trials have exhibited a lesser serum half-life and is readily metabolized into its stereoisomers, therefore drug label mentions its use thrice a day. Fluoxetine and Paroxetine induces metabolism of TBZ in liver through CYP2D6 systems. Adverse mental side effects of TBZ have been reported such as increased suicidal tendency, akathisia, depression and drowsiness [37], [38], [39], [40].

### 2.3.2 Deutetrabenazine (dTBBZ) or Austedo

3 Deutetrabenazine [ $C_{19}H_{27}NO_3$  or (3S,11bS)-3-(2-methylpropyl)-9,10-bis(trideuteriomethoxy)-1,3,4,6,7,11b-hexahydrobenzo[a]quinolizin-2-one] is the deuterated isomer of TBZ in which deuterium atoms replace six hydrogen atoms attached to two methoxy groups. This transformation creates strong bonds between heavy hydrogen atoms carbon thus rendering dTBZ a higher resistance to metabolism as compared to TBZ, thus it requires more energy for cleaving carbon-deuterium bonds. It was developed with the aim of reducing side effects like sedation and akathisia which emerged at peak doses and to control dosing for managing chorea[15]. dTBZ also metabolizes into active metabolites similar to TBZ namely (+/-) deuterated  $\beta$ -HTBZ and (+/-)  $\alpha$ -HTBZ. Among the four metabolites, (+)  $\beta$ -HTBZ and (+)  $\alpha$ -HTBZ are involved in VMAT2 inhibition. It was developed as a stable and long lasting alternative of TBZ and its plasma half life ( $t_{1/2}$ ) is about two times greater as compared to TBZ. Its mechanism of action involves controlling dopamine concentration levels in specific cerebral parts by VMAT2 inhibition and decreased re-uptake of monoamines. Its usage is associated with notable side effects such as fatigue, insomnia and somnolence. It is administered once or twice per day which is lesser frequent than TBZ (three times per day) [40], [41], [42].

### 2.3.3 Valbenazine (VBZ) or Ingrezza

8

Valbenazine ( $C_{24}H_{38}N_2O_4$  or ), in August 2023, was approved for amelioration of chorea associated with HD after it demonstrated high efficiency in randomized controlled trials. It is the valine ester derivative of (+)  $\alpha$ -HTBZ, this esterified modification assists in delaying metabolism. Its pharmacological effect is due to its active and selective hydrolysed metabolite viz., [+] - $\alpha$ -dihydrotrabenazine ([+] - $\alpha$ -HTBZ) [metabolised by cytochrome P450 2D6 (CYP3A4/5 enzymes)], and works selectively by regulating monoamine release in CNS impacting motor functions [32], [43]. NBI-136110 is another metabolite of VBZ obtained after its mono-oxidation but has minimal effect. Among the three VMAT2 inhibitors, valbenazine exhibits the longest half-life of 15-22 hours and within 30 mins to 1 hour it attains maximum plasma concentrations ( $C_{max}$ )[44]. Cryo-EM structure of VBZ sequestered to VMAT2 in lumen facing conformation, in which luminal gates stays opened and VBZ attaches to central cavity via abundant overlapping non-polar molecules participating in hydrophobic binding demonstrates stable structure thus explaining longer half-life [43] Residue E312 plays a crucial role in forming salt bridge with VBZ's amine residues (positively charged). A single daily dosage of VBZ has been found to demonstrate enhanced safety, tolerability, efficacy and specifically effective for patients prone to or suffering from psychiatric malfunctions as it tends to produce negligible side effects attributing to its highly selective nature. Even though TBZ exhibits quick and higher efficacy but it comes at the cost of elevated toxicity profile therefore VBZ is prescribed as its a better alternative to TBZ. VBZ has outperformed other VMAT2 inhibitors and is most promising candidate in parameters of efficacy, tolerability, curbing overall motor manifestations and high risk events [45], [46], [47], [48].

## 2.4 Computational Aided Drug Discovery (CADD) Rationale

Procedure of drug discovery is a tedious, time consuming and expensive task. Also, there is huge chemical space that has “drug-like” chemical environment comprising an estimated  $10^{60}$  small molecules which is difficult to explore. US FDA consolidates the drug discovery process into five steps i.e., discovery and development step (optimization of hit to lead, hit molecules are modified via different means to enhance their selectivity and activity to a particular target includes screening large libraries, management, designing chemical structures, identification and optimization of lead hits) the second phase involves pre-clinical trials in which animal testing and models of organ are employed to assess safety and efficacy of designed drugs, third and longest phase is clinical trials on humans (3 phases I, II and III), successful candidates are commercialized and lastly regular post market safety of drug is conducted. All this process takes up years and huge sums of money [49].

Computer-aided drug design (CADD) and Artificial Intelligence (AI) have emerged as a feasible option for drug discovery. CADD virtualizes the process of drug discovery. CADD-leading studies have reported high rates success in identifying bioactive substances as therapeutic targets. Computational framework of CADD comprises performing in silico screening, ligand-receptor docking, quantitative structure-activity relationship (QSAR) modeling and novel designing for predicting binding interactions [50].

### Drug design approaches via CADD:

**1. Ligand-Based Drug Design (LBDD):** it is most frequently used approach that utilizes structural data of molecular structures which are already tested on target of interest. The objective is to determine patterns called as QSAR models which can be extended to obtain quantitative connection between chemical segment and its pharmacological results. It relies on information of known drug molecules to design novel drug compounds. Application of this approach has its limitations such as focusing only on 2D interpretation of molecules,

huge experimental data required to validate results and only work on congeneric ligand libraries [50], [51]

**QSAR Modeling:** it can be used to create predictions on pharmacological action of novel compounds on the basis of structural features, assisting researchers to build knowledgeable modifications to improve potency of a drug along with reducing its side effects. Similarity Ensemble Approach (SEA) is one critical tool employed to measure the precision of k-nearest neighbors (kNN) QSAR models.

**Pharmacophore Modeling:** pharmacophore refers to spatial organization of critical parameters in a molecule essential for its pharmacological function. It enables contemporary drug designing by creating rational designs of new molecular compounds possessing optimized pharmacological features. This information can improve knowledge of receptor-ligand interactions and generate drugs with least side effects [52]

**2. Structure-Based Drug Design (SBDD):** with the advent of availability of 3D structures of molecules SBDD means enables to address small molecule screening and structural analysis. The former comprises techniques of molecular docking, post docking molecular dynamics, AI based SBDD screening, enhanced sampling post docking and free-energy perturbation (FEP). Structural analysis includes Molecular Dynamics, AI based structure predictions etc. 2 Availability of highly consolidated databases namely, DrugBank, ZINC20/22, PubChem etc., have enabled structure-based drug design (SBDD) and exploring pathways and genetic variants [53]

Following is description of key requirements for CADD:

- **Chemical databases, in silico screening and molecular modeling:** objective of de novo drug formulation is to create chemically plausible molecules exhibiting desirable pharmacological profile. Initially, screening of many chemical libraries to its biological leads is performed by identification of lead. Alternatively, molecular modeling of 3D structures of ligand and protein is performed to mimic molecular behaviour 3. Pertaining to its high cost, modern virtual screening was adopted for CADD which is broadly bifurcated into structure and ligand based approach. Currently, chemical libraries are categorized into three classes viz., general (PubChem, ChEMBL), specialized and natural product based (COCONUT 2.0).
- **Generating chemical structures in machine-encodable format:** structures represented in 2-dimensional, 3-dimensional, linear format and connection table format. 2D and 3D format enables spatial connectivity meanwhile, linear formats in form of SMILES assist machine learning (ML) due to its compact alphanumeric strings. Platforms like RDkit enables canonicalization and cleaning of chemical structures thus linking chemical data for virtual analysis.
- **ADMET (Absorption, Distribution, Metabolism, Excretion and Toxicity) Analysis:** poor pharmacokinetic profile of drugs leads to its dismissal, often detected in late clinical trials. To overcome such limitations, determining ADMET profile computationally is preferred. SwissADME is one such platform that enables pharmacological profiling via SMILES of query drug compound.
- **Prediction of target and receptor identification:** various strategies like network-based approaches, comparative genomics and target fishing are adopted for identifying molecular leads to which compound of interest may interact. Comparative genomics enables screening of exhaustive genomic data computationally to localize

essential genes. Network based approaches systematically analyse interactions in biological system for target selection. Target fishing integrates target and ligand focused strategies assisted by ML to predict drug-target interactions.

- **Molecular Docking:** it is critical to structure-based drug design (SBDD) and is used to predict conformation of ligand and its compatibility in active sites of target protein. The docking algorithm accuracy is essential for accurate scoring function. Advanced molecular docking tools enlisted below are crucial [49], [51], [54].

## 2.5 Molecular Docking rationale and methodologies

It is the most exhaustive procedure in CADD and involves identification of most optimal ligand-receptor conformation. Numerous small molecules can be screened against the desired target (protein or nucleic acid) at known binding site and helps rationalizing experimental data to a molecular stage. Molecular docking categories can be bifurcated as protein-protein and protein-ligand, former is applied for predicting interface of two macromolecules and latter is used to visualize localization of small molecule into protein's active site. Generally, docking is a two step process, initially the algorithm creates substitute ligand poses and then a scoring function validates and ranks them according to estimated binding affinity. The conformational site algorithm finds conformational area of the ligand to determine a state which fits at the binding site. Scoring functions are formulated by retrieving structural data from X-ray databases (from PDB and CSD) and later converting atom pair selections into distance-dependent binary comparison potentials by applying Boltzmann law [55].

Essentially, three kinds of scoring functions are present according to equations namely, force field-based (evaluation of energy by a force field), knowledge based (statistical assessment of most frequently observed interactions among ligand and protein) and empirical scoring functions (collection of various terms that represent various intermolecular associations). The docking results depend on scoring functions used which are generally empirical and are effective at evaluating orientation of pose than estimating absolute values of affinity [56].

Molecular docking on the basis of degrees of freedom in calculation can be classified as rigid, flexible ligand docking, semi-flexible and ensemble ligand docking. In rigid docking, ligand as well as protein are kept rigid. In flexible ligand docking, ligand is free to navigate through various conformational states while target is kept rigid. Induced fit or semi-flexible approach aims to avoid minor steric clashes to a rigid side chain that could hinder shortlisting of favourable docking poses. Ensemble docking performs docking against protein conformations ensemble, usually derived from MD simulations [56], [57].

### Tools for molecular docking

- **AutoDock Vina:** it is used to predict the binding affinity values and orientation of ligands. It is fast and easy for novice but lags at accuracy when dealing with complex systems.
- **GOLD:** it is suitable for flexible ligands and small hydrophobic ligands but it ranks ligands in large cavities. It is supported by both Windows and Linux. 3
- **PyRx:** openly accessible virtual molecular docking tool that compiles both AutoDock and AutoDock Vina.
- **iGEMDOCK:** it is used for post-screening evaluation and inferring pharmacological results of screened compounds. It generates pose via genetic algorithm and has empirical scoring functions.

- **SwissDock:** it is a freely accessible web-server that predicts interaction of molecules at atomic level. It conducts protein-ligand docking simulations by using CHARMM based scoring functions[58].

#### Basic workflow of molecular docking:

11

- **Protein Preparation:** obtaining pre-processed 3D structure of protein from PDB. After setting parameters, generates side chains, removal of water molecules, stabilization of charges and add missing residue.
- **Active site prediction:** predict the site of concern in receptor. Heteromolecules and water molecules are insignificant, if present.
- **Generating ligands:** either can be drawn using ChemDraw tools or retrieved from databases like ZINC. It is advised to use Lipinski's Rule of Five while selecting ligand.
- **Docking:** of protein and ligand can adopt several approaches, such as Monte Carlo approach in which initial configuration of ligand is synthesized consisting random rotations and conformation at active site, blind docking approach that scans total surface of target protein and detect all probable ligand binding sites etc. After docking, the interactions are assigned by ranking binding affinity values.
- **Visualization:** can be performed in two ways, either search for score functions used by software or validate score's decomposition. Protein-ligand interactions are visually represented by softwares like UCSF Chimera, Discovery Studio, PyMol etc.

Usage of molecular docking procedure which is basically a static method has its obstacles that are enlisted below:

- **Poor processing of solvation effects:** it refers to energy expenditure on displacing molecules of water and system's conformational entropy. It only takes into account the ultimate state of ligand-receptor system under vacuum conditions. Chances of obtaining false positive results of molecular substances chosen by scoring functions are high. Thus, emerges the need of complementing docking results by extensive computational means post docking operations, like molecular dynamics (MD)
- **Fidelity of Predictive models:** accuracy of computational models is a limiting factor in CADD as they are dependent on theoretical models. The theoretical models can lag to detect intricate details of biological systems. Re-calibration of scoring features, inclusion of vast molecular information and repeated validation alongside experimental data can assist in validating theoretical predictions.
- **Quality and quantity of data:** CADD predictions rely on data it is trained upon, so if data is insufficient or of poor quality it may generate inaccurate predictions. Removal of outliers, executing standard protocols (uniform assay conditions), endpoint calculations and consistent supervised data input can minimize inaccuracies.
- **High dependency on computational predictions:** without experimental data, relying solely on computational predictions should be avoided as it can lead to misguided results.
- **Demonstrating flexibility of molecules:** potential drug candidates and their respective target proteins possess high flexibility. Exact depiction of this flexibility is challenging especially in molecular docking and can impact results.

- **Difficulty in interpreting AI Models:** AI and ML models are becoming more complex with time pertaining to its 'black-box' nature and their predictions are also becoming difficult to interpret. It is harder to interpret why a specific drug is predicted active and optimized [57].

## 2.6 ADMET Profiling in virtual screening

12

The ultimate accomplishment of drug design does not only rest on fact that it has ability to bind to the site of target but is also pharmacokinetically stable i.e., its metabolism, distribution in body and solubility are within acceptable range. Drug Metabolism and pharmacokinetic (DMPK) properties can be anticipated prior to pre-clinical stage by CADD tools taking in consideration drug metabolism, its bioavailability and probable drug-drug associations [59]

ADMET profile prediction strategies are categorized into two approaches namely, structure-based means utilizing docking or dynamics and ligand based means utilizing QSAR models which are derived from biological and chemical datasets. SwissADME is a freely accessible platform that provides predictions on drug-likeness, pharmacokinetic features and physicochemical characteristics of a query compound. It requires linear notations of compounds in form of SMILES (molecular structures represented by ASCII characters) and exhibits compact alphanumeric strings of compounds. It aids in determining drug-likeness property of any biological molecule, which are notably advanced oral drug compounds. SwissADME consists of five rule-based filters out of which Lipinski (Pfizer) filter is apex rule for predicting drug-likeness [60], [61]

## 2.7 Research Void and Objectives Undertaken

Major challenges in treating neurodegenerative diseases like HD, is the difficulty to deliver therapeutic molecules across the blood-brain barrier (BBB). BBB is a protective layer that covers brain from impact of harmful compounds but at the same time imposes restriction of numerous drugs. Allele-selective methods of treating HD approaches lack to consider probable side effects from disrupting the biological function of wild-type protein. They do not target whole HD population through SNP related approaches or reducing the transcription of different genes sharing similar nucleotide sequences with few CAG-focusing therapies [62], [63].

Preceding literature evaluation: Many in silico research literature have investigated VMAT2 as target for drugs. Even so, there remains limited literature that specifically focuses on Valbenazine analogues utilizing integrated docking and ADMET filtering, thus retaining a gap that this study aims to address by systematically screening, characterizing and prioritizing novel Valbenazine-like VMAT2 inhibitors for its potential application in Huntington's chorea. Many studies use homology modeling and lacks comprehensive docking-based screening of focused set of VBZ analogue library. The rationale to investigate Valbenazine-like molecules stems after analyzing preceding studies in which structural modifications have led to creation of novel VMAT2 inhibitors.

This study places its focus on structure-based in silico optimization of analogues of Valbenazine as therapeutic candidates for symptomatic treatment of HD chorea. The work is preliminary in its computational methods, including preparation of VMAT2 structure, validation of molecular docking protocol by using co-crystallized VBZ, virtual screening of VBZ analogues and ADMET profiling of hit candidate. The results are solely computational

19

predictions, shortlisted lead as potential VMAT2 inhibitor and can be validated by experimental data.

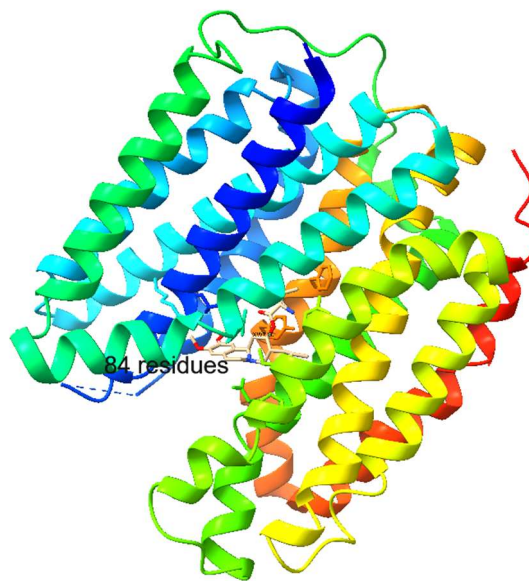
### 3. METHODOLOGY

#### 3.1 Deduction of target protein i.e. human VMAT2 complex structure

6

The three-dimensional (3D) structure of human VMAT2-Valbenazine complex was extracted from Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB, <https://www.rcsb.org/>), using PDB code 9KQ8 (extended PDB ID: pdb\_00009kq8). The protein structure was acquired utilizing cryo-electron microscopy at 3.38 Å spatial resolution and represents VBZ occupied in central cavity of VMAT2 in a lumen-facing conformation. This inhibitor-bound conformation can be used as an appropriate template for structure-based screening of valbenazine-like ligands[64]. The file was downloaded in PDB format for further in silico computations.

13

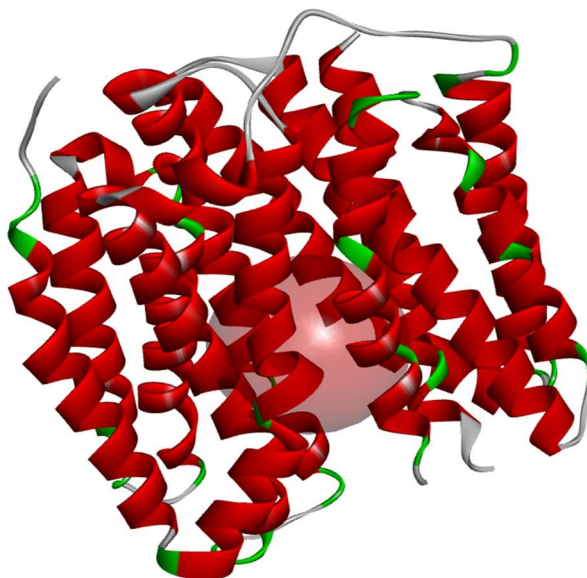


**Fig.1. Human VMAT2-Valbenazine complex in lumen-facing conformation captured by using UCSF Chimera software**

#### 3.2 Preparation of target protein structure

BIOVIA Discovery Studio Visualiser, a publicly accessible molecular modelling application, was used to identify chains (Chain A), non standard residues i.e., co-crystallized ligand (VBZ PDB ID:XW7) and water molecules (if any) in VMAT2-VBZ complex. All crystallographic water molecules were eliminated and to optimize the structure polar hydrogens (at physiological pH) were included which are essential for accurate depiction of hydrogen bonding interactions and electrostatic interactions.

The co-crystallized ligand i.e., valbenazine was initially kept to define the binding site by highlighting SBD sphere and sphere attributes with coordinates  $x=110.4\text{\AA}$ ,  $y=119.7\text{\AA}$  and  $z=116.8\text{\AA}$  were noted. It ensures that docking simulations occur in ambit of known inhibitor binding cavity i.e., at the interface of transmembrane helices. Then, heteroatom (valbenazine) was removed for generating an empty binding site for further docking the protein with ligand library. The final structure was saved in PDBQT format [65].



**Fig.2. Prepared target protein active site for docking of ligand library**

14

### 3.3 Ligand library selection and preparation

- Reference ligand retrieval: The 3D conformation of the reference ligand Valbenazine (Compound CID 24795069) was downloaded from the freely accessible website of PubChem Database.
- Refining Valbenazine analogs: Each compound exhibiting structural similarity to Valbenazine were browsed using the ‘Similar Structures Search’ and Tanimoto threshold was set to 90% that yielded 1,156 structures. The results were further refined by applying filters ranging close to the chemical and physical features of valbenazine, for instance, molecular weight, complexity, H-bond donor count etc. Nine final potential compounds were obtained which were saved in 3D Structure Data File (SDF) format which assists in collectively retaining data consisting of atom types and coordinates [66].

### 3.4 Molecular docking operations

- Docking software used: Virtual molecular screening is performed in order to dock libraries encompassing small molecules to a macromolecule to yield candidate compounds with preferred biological mechanisms. Multi-OS compatible PyRx Python Prescription 0.8 (<https://pyrx.sourceforge.io/>), an open available and user friendly virtual screening software, has altogether integrated AutoDock Vina, AutoDock4, 3D Viewer (Mayavi) and Open Babel was employed for molecular docking operations. The prepared protein target VMAT2 structure was loaded in PyRx and modified to macromolecule. Each ligand was loaded individually via the Open Babel tab.

- Energy minimization: After specifying the macromolecule and ligands the AutoDock Wizard was selected and energy minimization of the ligands were performed for generating conformations with relatively stable energy and reduce steric hindrance. The energy minimized ligands were then converted to AutoDock Ligands (pdbqt) format. The grid box was generated with coordinates  $x=110.4\text{\AA}$ ,  $y=119.7\text{\AA}$  and  $z=116.8\text{\AA}$  and dimensions of  $25\text{\AA} \times 25\text{\AA} \times 25\text{\AA}$  such that it encompassed central cavity and key residues (Glu312, Tyr341, Tyr433 etc.) participating in inhibitor binding [67].
- Docking simulations and output consolidation: VMAT2 was modeled as a rigid receptor and ligands were automatically assigned rotatable bonds during PDBQT conversion, thus treated as flexible molecules. AutoDock Vina predicts binding affinity in kcal/mol and yields multiple binding conformations for every individual ligand. AutoDock vina was allowed to run which yielded virtual screening results containing binding affinity values (in kcal/mol) and Root Mean Square Deviation (RMSD) values which denotes resemblance of a predicted docked ligand pose to actual ligand position (lower the score similar the ligand pose;  $<2\text{\AA}$  generally considered ideal). The results were extracted into a separate CSV (Comma Separated Value) file. The primary results from virtual screening consist of most suitable predicted binding modes (of RMSD value  $0\text{\AA}$ ) and corresponding binding affinity values [67], [68].

15

## 4. RESULTS

### 4.1 Authentication of docking protocol

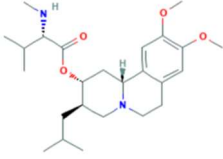
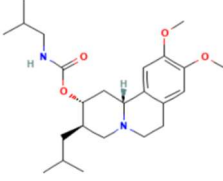
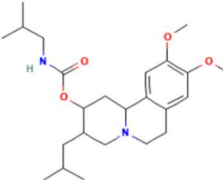
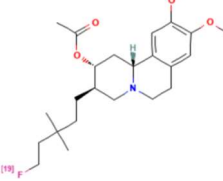
Re-docking of Valbenazine to the prepared target protein structure i.e., VMAT2 (PDB ID: 9KQ8) was performed to validate the docking methodology. The highest-ranked pose was precisely similar to the experimentally determined binding orientation with calculated RMSD value of  $0\text{\AA}$  and retained key residues such as V232, L37, Y433, I308 and E312 which contributes in forming a salt bridge with VBZ's protonated charged amine. This comparison confirmed that the calculated grid parameters and docking protocol was suitable to move forward for docking with VBZ analogues.

### 4.2 Analysis of binding affinity

Binding affinity refers to measurement of how strong are interactions between protein and ligands and also provides insight onto stability of receptor-ligand complex. A negative value indicates that ligand is predicted to dock with protein macromolecule. The higher the negative value, the better the predicted docking between protein and ligands. The binding affinity was  $-9.2\text{kcal/mol}$  for the reference drug and out of the nine ligands the highest binding affinity was noted in compound CID 163809280 with the value of  $-10.8\text{kcal/mol}$ . Dissociation constant ( $K_d$ ) is the ligand concentration at which 50% of the protein binding sites are filled. It was calculated using equation  $\Delta G = RT \ln K_d$  where,  $\Delta G$  is Gibbs free energy,  $R$  is universal gas constant,  $T$  is absolute temperature (in Kelvin),  $\ln$  is natural logarithm and  $K_d$  is dissociation constant expressed in Molarity (M) [69]. Lower value of  $K_d$  indicates stronger binding of ligand to receptor and the compound CID 163809280 exhibits lowest  $K_d$  value of  $1.19 \times 10^{-8}\text{M}$  indicating most stable binding among the nine VBZ analogs and provides approximate measure of inhibition potency (Table 1).

**Table 1: Docking Details of The Ligand Library**

S. No.	Compound CID	Chemical Structure	Binding Affinity (kcal/mol)	Dissociation constant ( $K_d$ ) in M	Interacting Ligands
1.	24795069 (Reference)		-9.2	$1.77 \times 10^{-7}$	Glu312, Ala337, Ile308, Phe334, Val232, Tyr433, Tyr341, Leu228, Ser196, Leu37, Asn34, Asp22
2.	163809280		-10.8	$1.19 \times 10^{-8}$	Phe334, Ile308, Ala337, Val232, Tyr341, Phe429, Leu228, Cys430, Ser200, Asp426, Leu37, Asp33, Ser196, Asn34
3.	156193539		-9.4	$1.27 \times 10^{-7}$	Asp33, Ser196, Leu37, Asn34, Leu228, Val232, Tyr341, Ile308, Phe334, Ala337
4.	57730406		-8.5	$5.79 \times 10^{-7}$	Tyr341, Tyr433, Val232, Asn305, Leu228, Lys138, Asn34, Leu37, Leu30, Asp33, Gln142, Ser196
5.	140537505		-9.2	$1.77 \times 10^{-7}$	Ser196, Asn34, Asp33, Leu37, Lys138, Leu228, Tyr433, Phe334, Tyr341, Ala337, Ile308
6.	140537506		-9.2	$1.77 \times 10^{-7}$	Tyr341, Val232, Tyr433, Asn305, Leu228, Lys138, Asn34, Leu37, Leu30, Asp33, Gln142, Ser196, Asn305

7.	148738001		-9.0	$2.49 \times 10^{-7}$	Phe334, Ile308, Ala337, Val232, Tyr433, Tyr341, Leu228, Leu37, Asn34, Ser196, Asp33
8.	156174442		-9.0	$2.49 \times 10^{-7}$	Asp33, Lys138, Ser196, Asn34, Leu37, Glu312, Ile308, Val232, Phe334, Ala337, Tyr433, Tyr341, Asp426, Cys430
9.	156174445		-9.0	$2.49 \times 10^{-7}$	Val232, Tyr433, Asn34, Phe429, Leu228, Lys138, Leu37, Ala139, Pro236
10.	57730405		-8.4	$6.85 \times 10^{-7}$	Asp33, Ser196, Leu37, Asn34, Tyr433, Ile308, Tyr341, Asn305

17

### 4.3 Visualization of virtual screening results

BIOVIA Discovery Studio Visualizer version 25.1.0.24284 is a software suite comprising tools for 2D or 3D visualization of protein-ligand interactions and was utilized for viewing molecular docking results. The interactions occurring were generated in 2D format along with the interacting residues for the re-docked reference compound and docked ligand library comprising of top hit compound. Crucial interacting residues were examined in each 2D ligand-receptor interaction, for instance, hydrogen bonding interactions with Tyr433 and Tyr341, salt bridge presence on side chain of Glu312 and aromatic interactions Phe334. The 2-Dimensional (2D) presentation of interactions between reference compound Valbenazine and VMAT2 is exhibited in Fig1, similarly rest of 2D interactions are arranged in descending order from best docked to least favourable docked pose

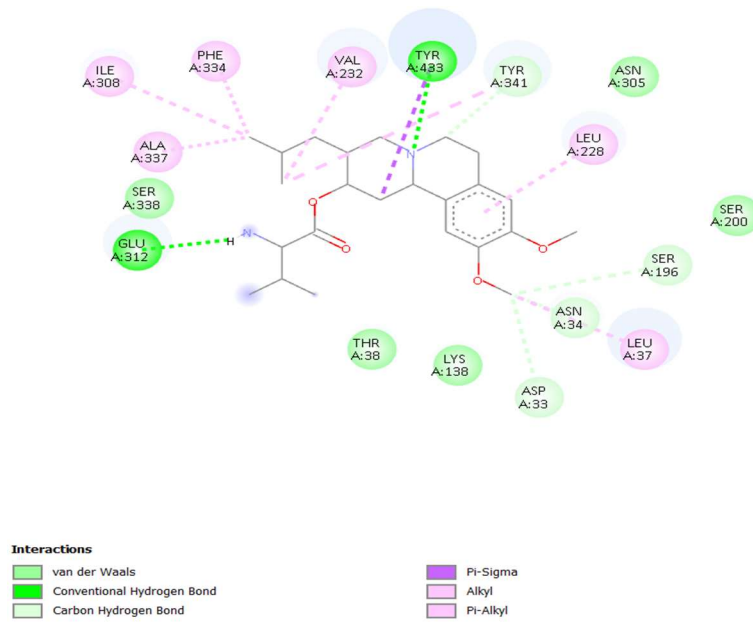


Fig. 3. 2D presentation of interactions between compound CID 24795069 and VMAT2

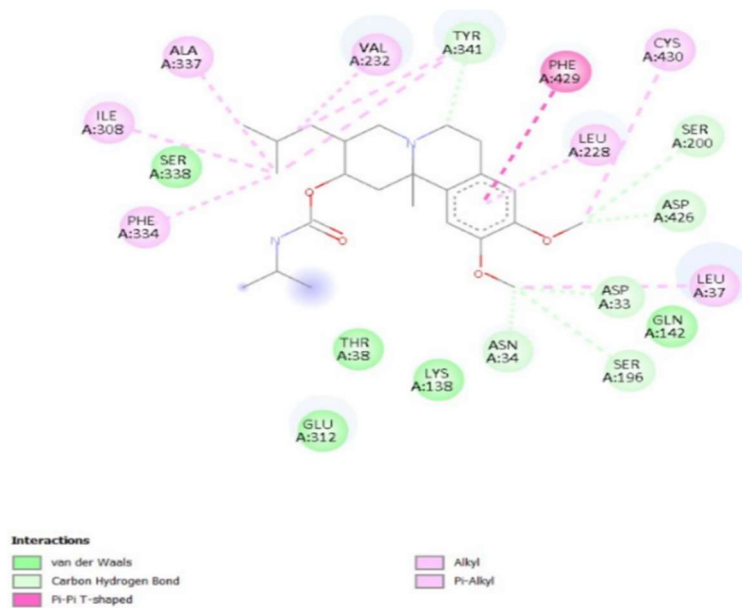


Fig. 4. 2D presentation of interactions between compound CID 163809280 and VMAT2

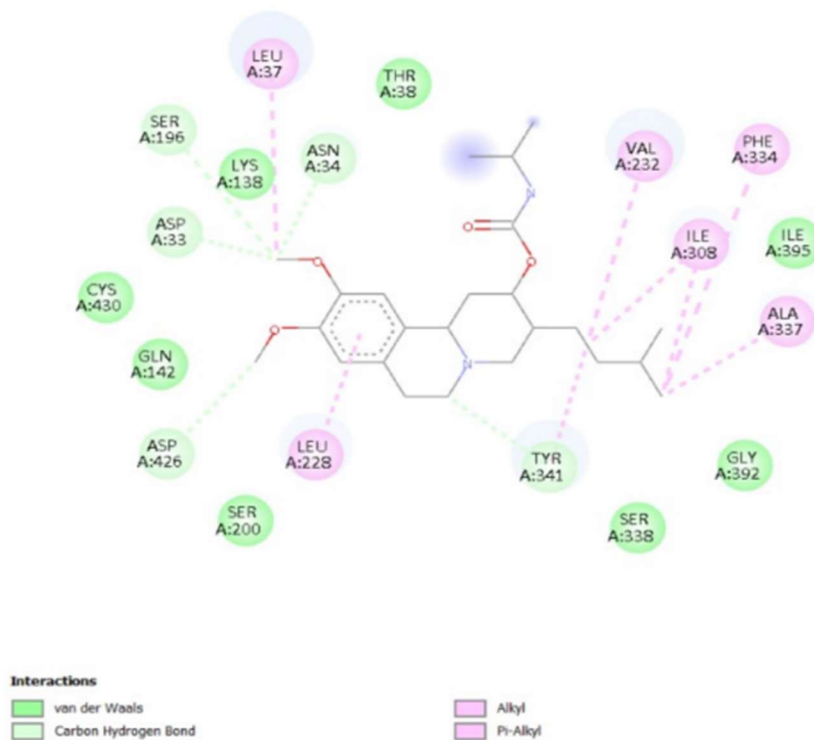


Fig. 5. 2D presentation of interactions between compound CID 156193539 and VMAT2

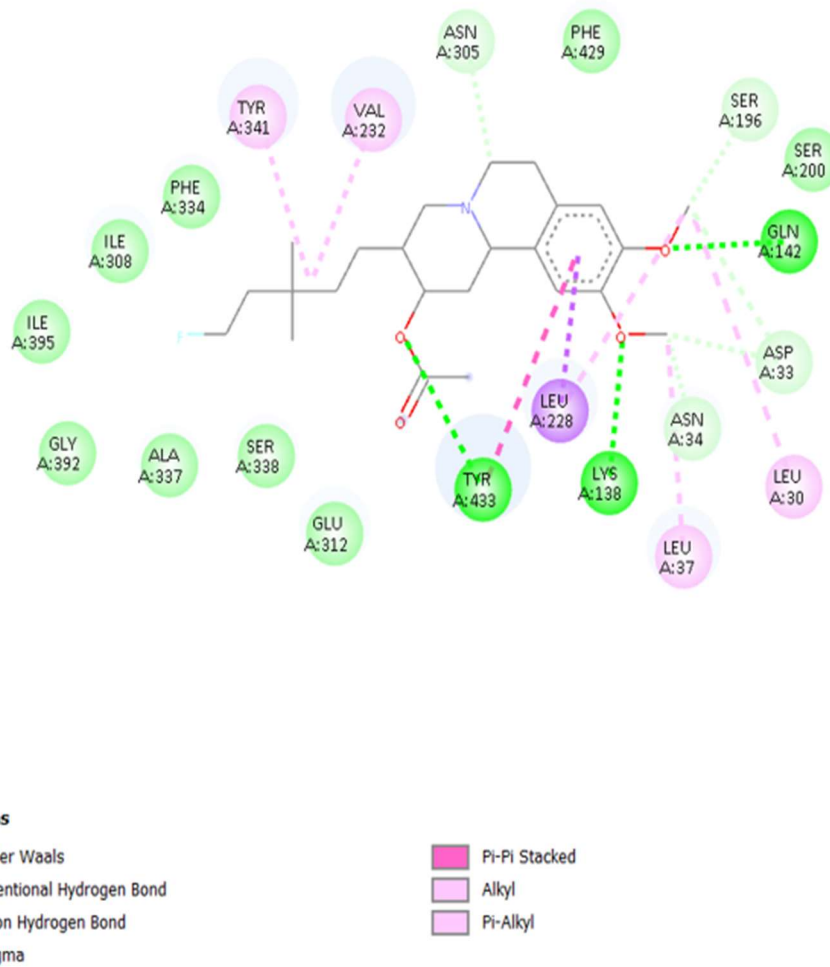


Fig. 6. 2D presentation of interactions between compound CID 57730406 and VMAT2

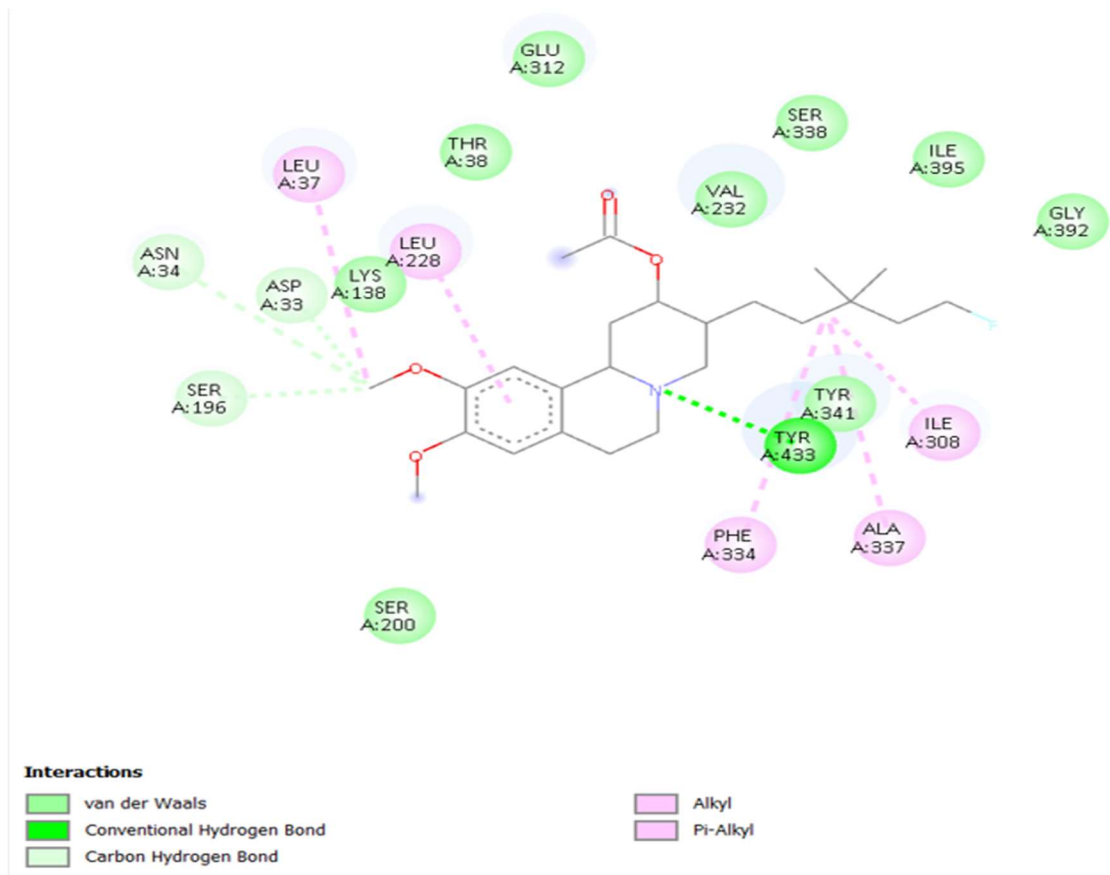


Fig. 7. 2D presentation of interactions between compound CID 140537505 and VMAT2

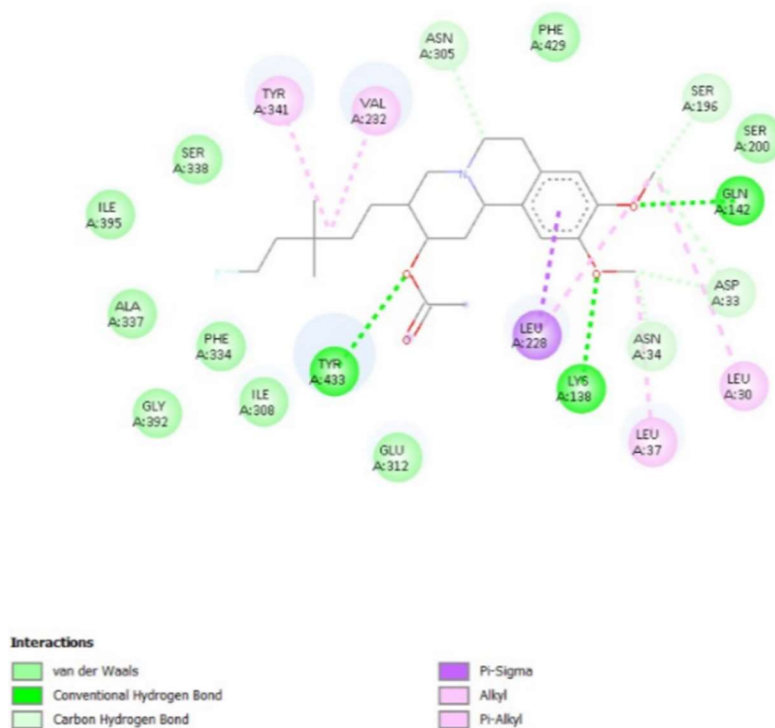


Fig. 8. 2D presentation of interactions between compound CID 140537506 and VMAT2

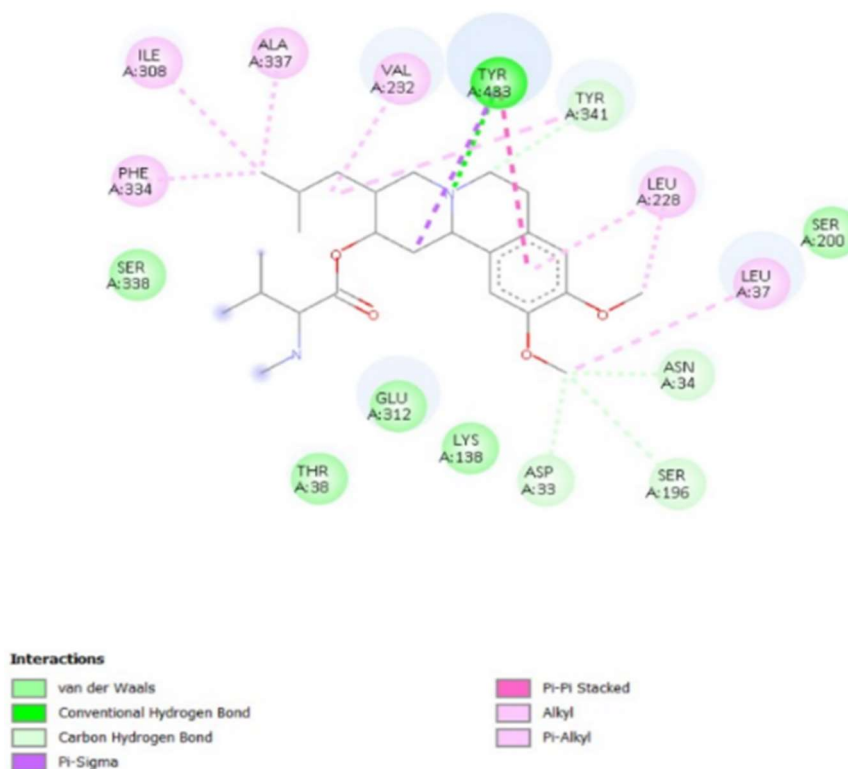


Fig. 9. 2D presentation of interactions between compound CID 148738001 and VMAT2

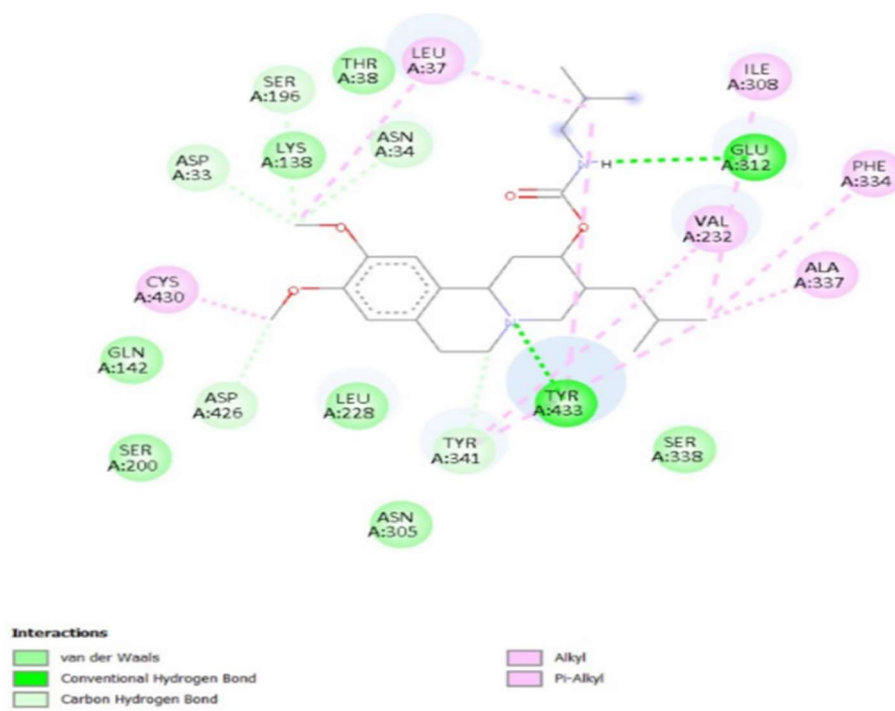


Fig. 10. 2D presentation of interactions between compound CID 15617444 and VMAT2

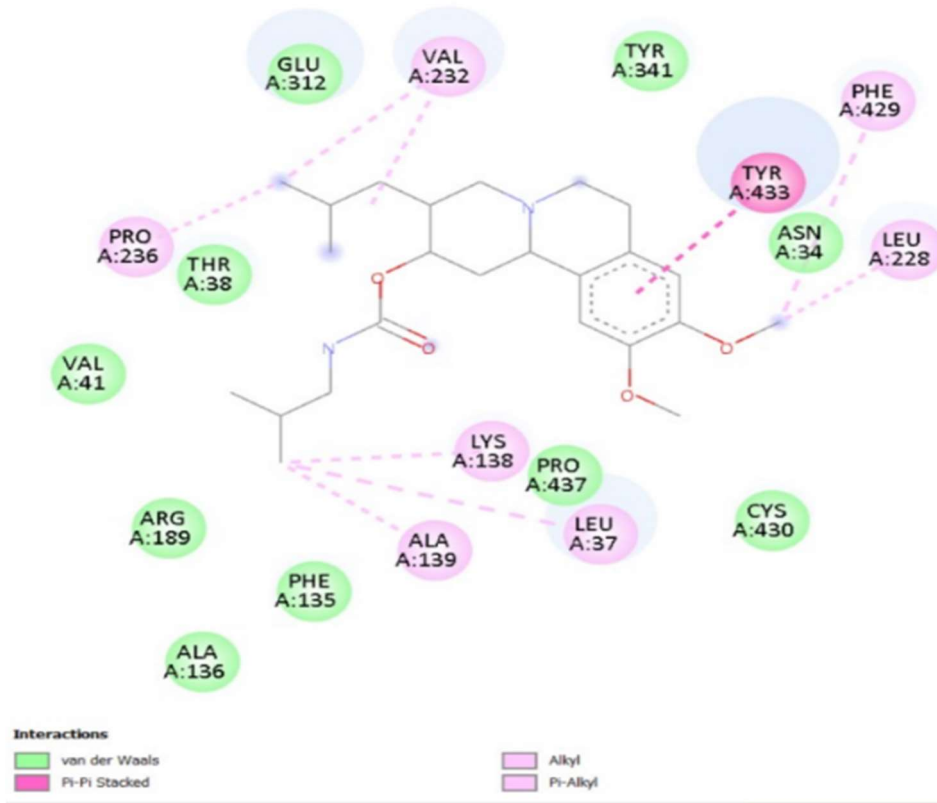


Fig. 11. 2D presentation of interactions between compound CID 156174445 and VMAT2

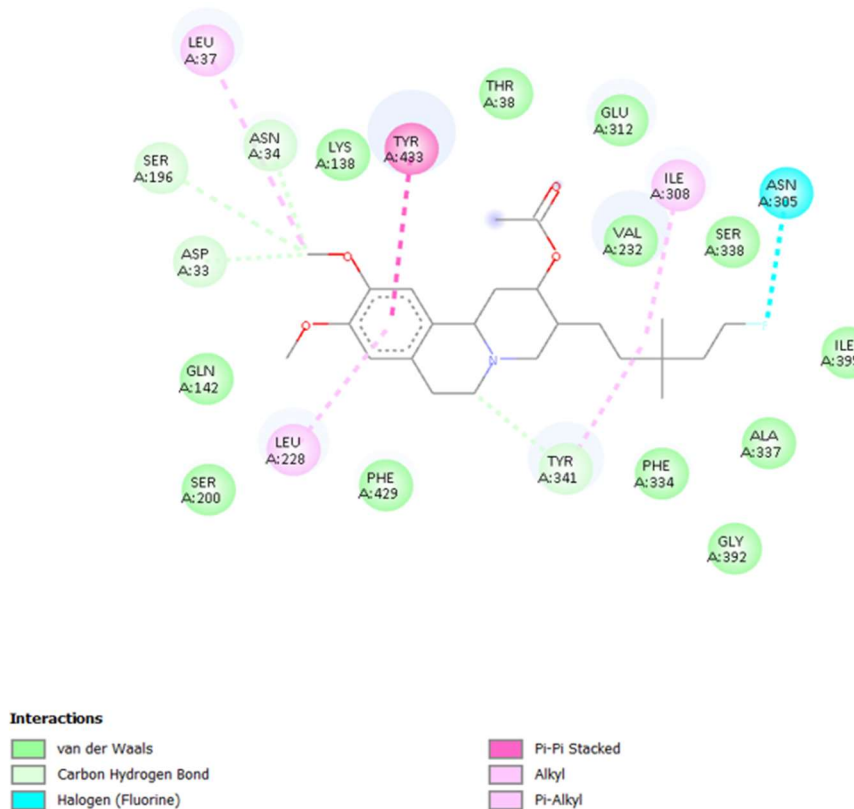


Fig. 12. 2D presentation of interactions between compound CID 57730405 and VMAT2

22

#### 4.4 ADMET Analysis

The acronym ADMET refers to four fundamental pharmacokinetic processes viz., Absorption, Distribution, Metabolism, Excretion and Toxicity. The potential molecules are assessed on the basis of aforementioned parameters and in addition they must exhibit low toxicity. SwissADME was utilized for obtaining physicochemical and pharmacokinetic properties of desired molecules in accordance with the Rule-of-five. It is a non exhaustive tool that can compute multiple molecules and yield results per single molecule via interactive graphs. The SMILES of ligands were used as input and were assessed on various parameters notably BBB permeability, Lipinski rule of five, brenk, PAINS, solubility etc. (Table 2). All the resultant ligands obeyed Lipinski's rule without any violations thus indicating good oral drug-like properties [70]. The webserver ProTox 3.0 integrates chemical similarity screening, fragment-based toxicity patterns, pharmacophore modelling and machine learning to predict multiple outcomes of drug's toxicity. The lead compound, CID 163809280, exhibited acceptable toxicity profile with predicted LD<sub>50</sub> value of 75mg/kg and classified under Class 3 on scale of toxicity. It indicated that improved binding affinity of CID 163809280 doesn't come at cost of higher toxicity profile [71].

**Table 2: ADME Analysis of The Ligand Library**

Compound CID	BBB Permeable	Lipinski Violation	TPSA Value (in Å <sup>2</sup> )	Consensus logP	GI absorption	log Kp skin permeation (in cm/s)
163809280	Yes	No	60.03	3.95	High	-5.57
156193539	Yes	No	60.03	4.02	High	-5.32
57730406	Yes	No	48.00	4.33	High	-5.39
140537505	Yes	No	48.00	4.33	High	-5.39
140537506	Yes	No	48.00	4.33	High	-5.39
148738001	Yes	No	60.03	3.85	High	-5.55
156174442	Yes	No	60.03	3.96	High	-5.33
156174445	Yes	No	60.03	3.92	High	-5.33
57730405	Yes	No	48.00	4.33	High	-5.39

20

## 5. CONCLUSION

At present, the realm of treatment for Huntington's disease places its focus on alleviating symptoms through potential therapies and anti-dopaminergic medication targeting inhibition mechanisms. Among all, VMAT2 inhibition by Valbenazine emerges as a promising strategy to reduce the symptoms of chorea in HD. The objective of this study was to identify potential chemical structures structurally resembling Valbenazine with acceptable drug-like properties which can be utilized for symptomatic cure for chorea. We have deduced nine compounds which exhibited predicted binding affinities comparable to and better than Valbenazine with binding modes that preserved essential interactions with critical residues within the VMAT2 central cavity. The compound CID 163809280 turned out to have highest binding affinity and potentially serves as an alternative.

The structure-based docking results were integrated with early ADME profiling of shortlisted drug candidates and toxicity profiling of lead compound CID 163809280. The workflow adopted in this study provides rational and inexpensive strategy to focus only on limited promising candidates instead of large libraries of drug compounds. Currently, no prodrug has been developed that can cater to both motor and psychological symptoms of HD. The difference lies in underlying mechanism, that is degeneration of striata in HD and dysregulation of cortex and sub-cortex of brain which yet cannot be targeted together at same time[72]Future attention is required on developing therapeutic drugs that can target both motor and psychiatric manifestations of HD simultaneously.

All the results in this study are calculated from computational models, docking scores and ADMET profile predictions which is merely an estimate of real binding affinity values and in vivo mechanism. Effects of solvents, flexibility of protein and complex aspects of toxicity and metabolism are only partially evaluated by the methods utilized. Thus, there is necessity of experimentally validating VMAT2 inhibition efficiency by the prioritized compound CID 163809280 via vesicular uptake assays. Pharmacokinetic evaluation, safety and efficacy of CID 163809280 compound shall be performed in appropriate cellular and animal models of Huntington's disease. Additionally, future work can also extend the similar approach by identifying plausible chemical modifications to Valbenazine, employing advanced simulation technologies such as molecular dynamics and calculation of free-energy, further comparing candidates derived from chemically modified Valbenazine with structurally distinct VMAT2 inhibitor chemical structures.