Journal of Molecular Graphics and Modelling 39 (2013) 145–164

Contents lists available at SciVerse ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Pharmacophore modeling and virtual screening studies to design potential COMT inhibitors as new leads

Nidhi Jatana, Aditya Sharma, N. Latha*

Bioinformatics Infrastructure Facility, Sri Venkateswara College (University of Delhi), Benito Juarez Road, Dhaula Kuan, New Delhi 110 021, India

ARTICLE INFO

Article history:
Received 21 June 2012
Received in revised form
15 September 2012
Accepted 8 October 2012
Available online 28 November 2012

Keywords:
COMT inhibitors
Docking
Virtual screening
Pharmacophore
Binding free energy estimates

ABSTRACT

Catechol-O-methyltransferase (COMT) catalyzes the methylation of catecholamines, including neurotransmitters like dopamine, epinephrine and norepinephrine, leading to their degradation. COMT has been a subject of study for its implications in numerous neurological disorders like Parkinson's disease (PD), schizophrenia, and depression. The COMT gene is associated with many allelic variants, the Val108Met polymorphism being the most clinically significant.

Availability of crystal structure of both 108V and 108M forms of human soluble-COMT (S-COMT) facilitated us to use structure-based virtual screening approach to obtain new hits by screening a library of CNS permeable compounds from ZINC database. In this study, E-pharmacophore was also used to generate pharmacophore models based on a series of known COMT inhibitors. A five-point pharmacophore model consisting of one hydrogen-bond acceptor (A), two hydrogen bond donors (D), and two aromatic rings (R) was generated for both the polymorphic forms of COMT. These models were then used for filtering ZINC-CNS permeable library to obtain new hits. Physicochemical properties were also calculated for all the hits obtained from both the approaches for favorable ADME properties. These identified hits maybe of interest for further structural optimization and biological evaluation assays.

1. Introduction

Catechol-O-methyltransferase (COMT; EC 2.1.1.6) is a monomeric enzyme involved in the inactivation of catechols and catecholamines including neurotransmitters like dopamine, epinephrine and norepinephrine [1]. It transfers the methyl group from S-adenosylmethionine (SAM) in the presence of Mg^{2+} to various catechols and catecholamines, leading to their degradation (Fig. 1). COMT has been implicated in various human disorders like Parkinson's disease (PD) [2], depression [2,3], bipolar disorders [4], attention deficit hyperactivity disorder (ADHD) [5], schizophrenia [4,6], hypertension, heart failure [7] and many other diseases.

The enzyme is found to be present in almost all mammalian tissues investigated, the highest activity being in liver, followed by the kidneys and gastrointestinal tract [8,9]. COMT exists in two isoforms-soluble (S-COMT) and membrane-bound (MB-COMT) [10]. S-COMT is expressed at higher levels in most of the tissues and MB-COMT is more prevalent in the brain [11,12]. Human S-COMT contains 221 amino acids whereas the membrane-bound form consists of an additional 50 amino acids at the N-terminus that contain the hydrophobic membrane-anchor region.

COMT gene is located on the long (q) arm of chromosome 22 and is highly polymorphic. Polymorphisms in upstream, downstream regions and within the COMT gene (Table 1) are involved in a variety of neuropsychiatric phenotypes, most studied being the Val108Met polymorphism [13]. It has been shown that the allele frequency of this polymorphism significantly differs across the populations studied world-wide [14]. Both the allelic forms have been linked to various disorders in human; hence it becomes evident to target both the allelic forms for the design of novel therapeutics. The Val108Met genetic polymorphism has been associated with PD [15], Alzheimer's disease [16], breast cancer [17], obsessive–compulsive disorder [18], adult-onset alcoholism [19], aggressive and suicidal manifestations of schizophrenia [20], Irritable Bowel Syndrome [21] and many more.

Vidgren and coworkers [22] reported the first crystal structure of rat S-COMT (PDB# 1VID) in 1994, after which the efforts for the design of inhibitors significantly improved. The crystal structure of human S-COMT bound to SAM, 3,5-dinitrocatechol (DNC) and Mg^{2+} was resolved by X-ray crystallography for both the 108V and 108M forms (PDB# 3BWM (108V) and 3BWY (108M)) [23]. Subsequently, binding studies with substrates and various types of inhibitors have played an important role in the design of novel COMT inhibitors [2].

The design of inhibitors has been in practice since the discovery of COMT in 1958 [24]. The first generation of COMT inhibitors such as tropolone, N-butylgallate, catechol, 2-hydroxylated oestrogens and U-0521 showed low efficacy *in vivo*, were short acting,

* Corresponding author. Tel.: +91 11 24111742; fax: +91 11 24118535.
E-mail address: lata@bic-svc.ac.in (N. Latha).

1093-3263/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2012.10.010

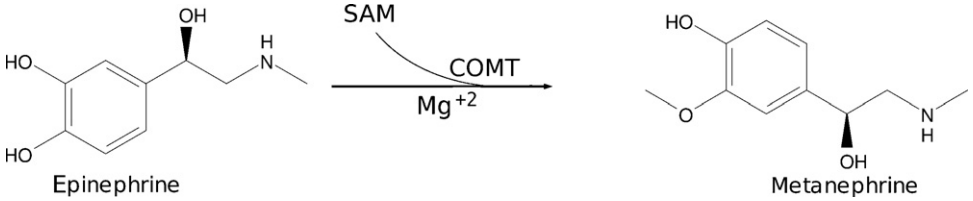


Fig. 1. Reaction catalyzed by catechol-O-methyltransferase (COMT).

lacked selectivity and were quite toxic [25]. The second-generation inhibitors included nitrocatechols like entacapone (OR-611), nitecapone (OR-462) and tolcapone (Ro 40-7592) [26,27]. Entacapone and tolcapone are currently being used in the treatment of Parkinson's disease where the COMT inhibitor has a beneficial effect of prolonging the half life of levodopa [28]. Many other approaches were used for the design of COMT inhibitors where two catecholic pharmacophores are incorporated in the same inhibitor molecule [29,30]. Attempts were also made for the design of bisubstrate inhibitors by incorporating both the catechol and the SAM substructures in the same molecule [31].

Currently used COMT inhibitors namely tolcapone and entacapone, in the treatment of Parkinson's disease, have various dopaminergic and gastrointestinal side-effects. Commonly observed dopaminergic events include worsening of levodopa-induced dyskinesia, nausea, anorexia, vomiting, orthostatic hypotension, sleep disorders and hallucinations. The most common gastrointestinal side-effects include diarrhea and this is a cause of treatment discontinuation [32]. Hence, there still exists a requirement for improved COMT inhibitors to address the unmet medical needs of many Parkinson's disease patients.

In this study, structures of both the 108V and 108M forms of S-COMT have been utilized for the design of novel inhibitors. A combination of structure-based and ligand-based approaches have been employed to obtain a novel set of potential inhibitors. Structure-based approach utilizes screening of hits with a pre-filtered ZINC dataset for central nervous system (CNS) activity. On the other hand, ligand-based approach generates a pharmacophore using an active dataset of known COMT inhibitors which is then screened with ZINC CNS permeable library to obtain hits. Top hits from these approaches are then evaluated for their ADME profile based on the calculation of physicochemical properties.

2. Materials and methods

All computations were performed on 64 Processor-8-node Sun Fire X2200 cluster running Cent OS 5 and on Sun Fire X4600 M2 server, dual-core AMD Opteron with eight processors running SUSE Linux Enterprise Server Edition 10.0.

2.1. Crystallographic structure and binding site analysis

The 221 amino acids long sequence of human S-COMT has been retrieved from NCBI (Genbank # ACJ38231). The crystal structures of both 108V and 108M forms of human S-COMT (PDB # 3BWM, 3BWY) bound with SAM, DNC and Mg^{2+} has been obtained from protein data bank (PDB) [33]. The binding site of the protein has been characterized using Q-SiteFinder [34] and CASTp [35]. The interactions of the ligands (substrates and inhibitors) with the protein residues in the active site were visualized using LIGPLOT [36].

Table 1
Table showing polymorphism in COMT.

S. No.	Polymorphism	NCBI SNP id	COMT activity[a]
1.	Val108Met (G/A)	rs4680	Yes
2.	Ala22Ser (G/T)	rs6267	No
3.	His12His (C/T)	rs4633	No
4.	Leu86Leu (C/G and C/T)	rs4818	No
5.	Ala52Thr (G/A)	rs5031015	No

[a] COMT activity refers to if function of COMT is affected by the mutation.

2.2. Compilation of known COMT inhibitors

The literature was extensively surveyed to collect diverse set of COMT inhibitors from experimental investigations and the three-dimensional structures of all the known COMT inhibitors (806 in number) were generated using Maestro module of Schrödinger [37]. Based on their structural frameworks, they were classified as catechol-based COMT inhibitors (Type I), SAM-based COMT inhibitors (Type II), natural-products as COMT inhibitors (Type III), bifunctional COMT inhibitors (Type IV), bisubstrate type of COMT inhibitors (Type V) and other types of COMT inhibitors (Type VI) (Fig. 2) for the purpose of convenience.

Type I class of inhibitors are designed based on the catechol moiety. Modifications of the catechol substructure resulted in a diverse set of compounds including nitrocatechols [26,27], pyrogallol [38], salsolinol [39], polychlorinated biphenyls [40], *etc.* (Fig. 3). Of these, tolcapone [41] and entacapone [42] have been used clinically for the treatment of Parkinson's disease.

Type II class of inhibitors are mostly analogs of S-adenosyl-L-homocysteine (SAH), demethylated product of SAM, which is a potent inhibitor of many methyltransferases [43]. A series of SAH derivatives were prepared with modification in the sugar portion [44], amino acid portion [45] and base portion [46] of which a few are represented in Fig. 4.

Type III class of inhibitors are natural product derivatives, of which flavonoids have been found to have great inhibitory action against COMT. Four major green tea catechins showed inhibitory activity in the following order of epigallocatechin gallate (EGCG) > (−)-epicatechin-3-gallate (ECG) > epigallocatechin (EGC) > (−)-epicatechin (EC) (Fig. 5) [47]. Constituents of coffee (chlorogenic acid, caffeic acid and caffeic acid phenethyl ester (CAPE)) [48] are also effective inhibitors of COMT. Studies carried out on natural compounds like flavonols, flavones, catechins and aromatics reported for their inhibitory action against COMT [49]. Also, natural products extracted from various plants like *Cistus parviflorus* and *Vitex agnus-cactus* and from the seeds of *Peganum harmala* have been reported for their inhibitory activity [50].

Type IV class of inhibitors have duplicated substructures which interact with binding domains of the enzyme. Some of them are D-catechin, desmethylpapaverine and nordihydroxyguaiaretic acid (Fig. 6) [51,52], which belong to first generation of inhibitors. Later, bifunctional inhibitors were designed with dual substituted catechols and were mostly catechol derivatives like 3,4-dihydroxybenzamide or 3,4,5-trihydroxybenzamide (Fig. 6), linked by a spacer section consisting of varying numbers of methylene units. Also, bifunctional inhibitors of varying length have been designed to investigate the effect of electron withdrawing groups

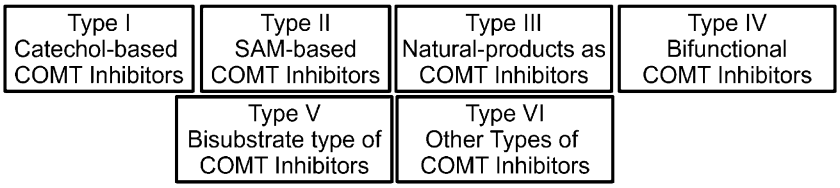


Fig. 2. Classification of COMT inhibitors. This classification is based on the structural framework (SAM = S-adenosylmethionine).

on the catechol rings that incorporated nitro group substitutions. On the whole, it was observed that some of them were significantly more potent than their monofunctional counterparts [29,30].

Type V class of inhibitors includes bisubstrate inhibitors which are designed by replacing both the cofactor SAM and the catechol substrate. The first potent bisubstrate inhibitor was reported in the year 2000 with an IC50 value of 2 μM (Fig. 7) [53]. Later, many of its structural analogs have been prepared by modifications of the ribose moiety that led to the observation that binding mode was extremely sensitive toward modifications of the ribose moiety [54,55]. However, their *in vivo* efficacy has not been established yet.

Type VI comprise of all those inhibitors that structurally cannot be included in any of the above classes. These constitute tropolones [56], benzotropolones [57], bicyclic compounds [58] and many others. Some of them are represented in Fig. 8.

The binding studies of these known inhibitors with COMT provided us with a clue to design novel, selective COMT leads using ligand-based approaches.

2.3. Structure-based and ligand-based drug design approaches

In the current study, structure-based approach was utilized due to the availability of the crystal structure of the protein which provides a scope for search of new scaffolds that could bind in the active site of protein. The ligand-based strategies propose and evaluate potential lead compounds based on pharmacophore generated so as to conserve the three-dimensional arrangement of functional groups on a scaffold believed to be most important in the activity of existing ligands. We then propose to integrate the ligand- and structure-based drug design methodologies. The ligand-based approach used here takes up the top docked compounds to generate a structure-based pharmacophore, therefore combining the strengths of both structure-based docking and ligand-based pharmacophore.

2.3.1. Structure-based approaches

Availability of crystal structure of human S-COMT facilitated the use of structure-based virtual screening for the search of novel hits. Structure-based virtual screening depends on fast and accurate computational methods for the estimation of receptor–ligand binding modes and binding affinities.

ZINC [59], a free database for virtual screening, containing over 21 million compounds has been utilized. Since the current study requires the generation of novel inhibitors that can cross blood–brain barrier, therefore ZINC was pre-filtered for CNS permeability [60]. The compounds with molecular weight ranging from 150 to 400 Da, polar surface area from 0 to 60 Å and log *P* values from 1.5 to 2.7 were part of the CNS permeable dataset. The ligands were prepared for docking using Ligprep module (Schrödinger) [61]. All the possible stereoisomers were generated at a selected pH range (7 ± 2, by default). Also, Ligprep was used to add hydrogens and metal binding states. Low energy 3D structures for ligands

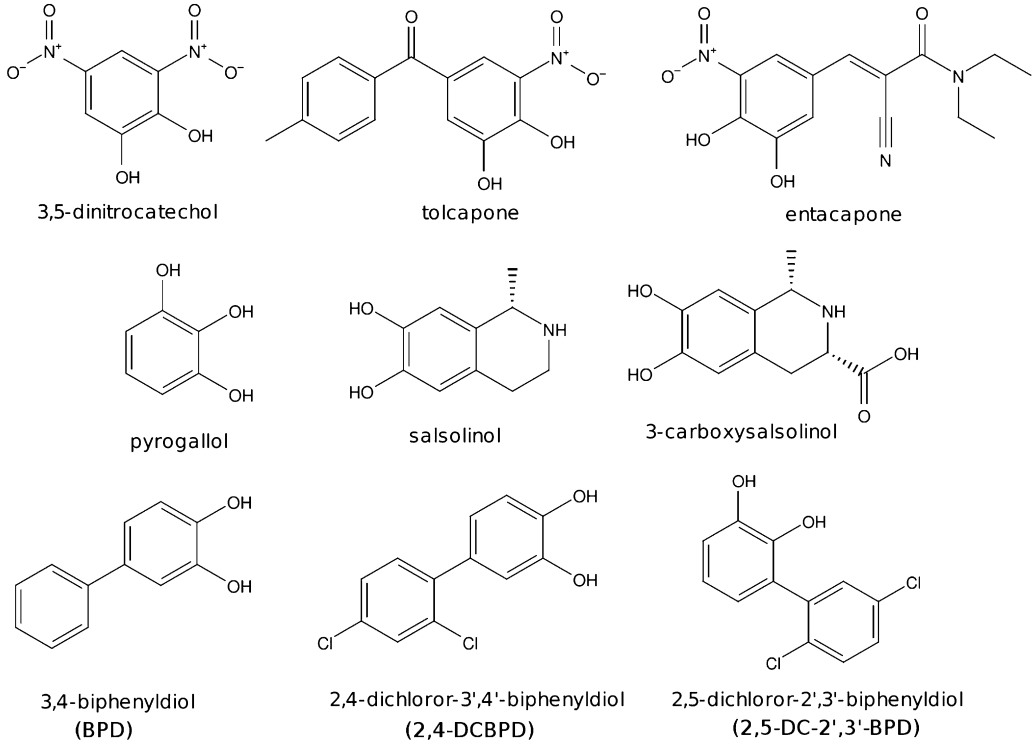
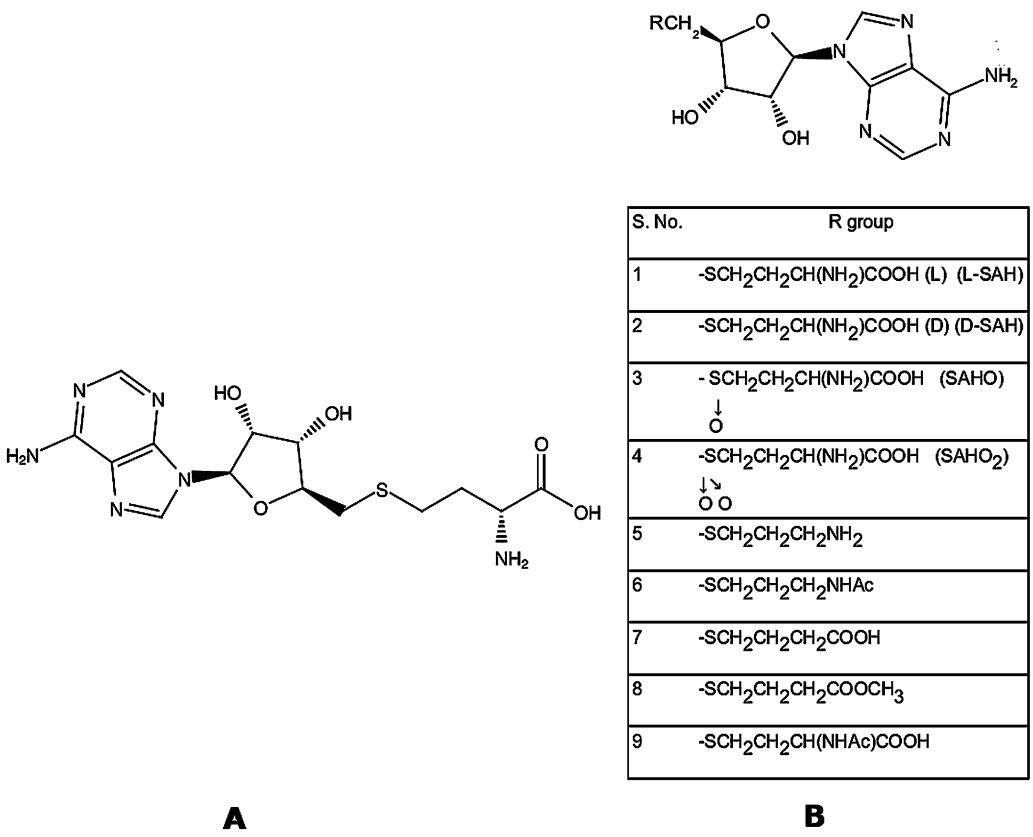


Fig. 3. Some of the Type I inhibitors (catechol-based inhibitors).



S. No.	R group
1	-$SCH_2CH_2CH(NH_2)COOH$ (L) (L-SAH)
2	-$SCH_2CH_2CH(NH_2)COOH$ (D) (D-SAH)
3	- $SCH_2CH_2CH(NH_2)COOH$ (SAHO) ↓ O
4	-$SCH_2CH_2CH(NH_2)COOH$ ($SAHO_2$) ↓↘ O O
5	-$SCH_2CH_2CH_2NH_2$
6	-$SCH_2CH_2CH_2NHAc$
7	-$SCH_2CH_2CH_2COOH$
8	-$SCH_2CH_2CH_2COOCH_3$
9	-$SCH_2CH_2CH(NHAc)COOH$

Fig. 4. Type II class of Inhibitors (SAM-based inhibitors). Structure of (A) S-adenosyl-L-homocysteine (SAH), demethylated product of SAM and (B) some of the derivatives of SAH.

were obtained using Optimized Potentials for Liquid Simulations (OPLS_2005) force field [62] by optimizing the geometry and minimizing the energy. The ligand atoms were assigned with force field parameters using default treatment for possible tautomers and up to 4 stereoisomers and 1 low ring conformation were retained per ligand. Also, ligands with reactive functional groups like acyl halides and alkali metals were filtered out.

The protein structures were prepared for docking by adding hydrogen atoms, assigning bond orders, removal of K^+ ion and other ligands from the active site. Only Mg^{2+} ion was retained and all the water molecules were removed except the one that forms coordination bond with Mg^{2+} ion. Also, metal binding states were generated for Mg^{2+} ion at pH 7.0 ± 4.0 and the protein molecule was optimized at neutral pH. Metal binding states were also included for ligands during ligand preparation. Since, metal-binding states are additional ionization states of ligand-like molecules; the adjusted metal state penalty is used in scoring instead of the normal penalty. Further, minimization was carried out using OPLS_2005 as the force field, by converging heavy atoms to RMSD of 0.3 Å.

Receptor grid was generated for both human S-COMT 108V and 108M around the SAM and DNC binding site. Virtual screening workflow [63] was implemented to screen the prepared CNS permeable small molecule database. The virtual screening workflow of Glide [64,65] consists of four stages – a three-tier docking pipeline to selectively filter ligands at every stage with increased stringency that also allows restricting the number of ligands that were passed onto the next stage. The final step ends with calculation of binding free energy estimates using MM-GBSA (Molecular Mechanics Generalized Born Surface Area) [66].

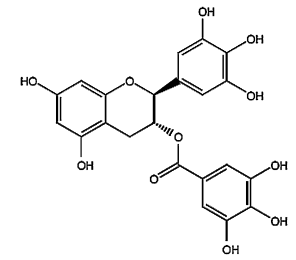
All the docking stages were run using OPLS_2005 force field. The first stage of docking performs high throughput virtual screening (HTVS) where conformational sampling was significantly reduced to increase the computational speed and to screen large database. The docking criteria in HTVS stage was set in such a way so that only the top 10% of all the docked states were passed on to the next stage. The next stage is standard precision (SP) docking, where a more extensive sampling using Monte Carlo procedure was carried out. The compounds were ranked using GlideScore which is as follows:

$$\text{GlideScore} = 0.065 \times \text{vdW} + 0.130 \times \text{Coul} + \text{Lipo} + \text{Hbond} + \text{Metal} + \text{BuryP} + \text{RotB} + \text{Site}$$

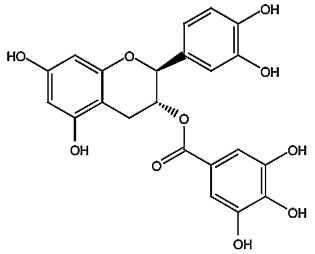
where vdW is van der Waals energy, Coul is coulomb energy, Lipo is lipophilic term derived from hydrophobic grid potential and rewards favorable hydrophobic interactions. Hbond is hydrogen-bonding term, Metal is metal-binding term which only includes the interactions with anionic or highly polar acceptor atoms, BuryP is penalty for buried polar groups while RotB signifies penalty for freezing rotatable bonds and Site indicates polar interactions in the active site.

From here, only top 10% of all the good scoring states were retained and passed on to the third stage, which performs extra precision (XP) docking [67]. XP docking does a more extensive sampling and uses anchor-and-grow strategy to weed out false positives. XP docking employs a more stringent scoring function than the SP GlideScore, which includes terms for hydrophobic enclosure and desolvation penalties.

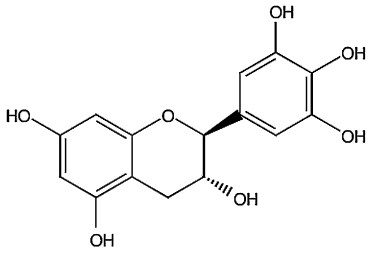
All the docked compounds were then evaluated for binding free energy estimates using MM-GBSA which employs VSGB 2.0 as the solvent model [68]. Protein–ligand contacts of the top hit with the interacting residues were visualized using ligand interaction diagram tool of Schrödinger.



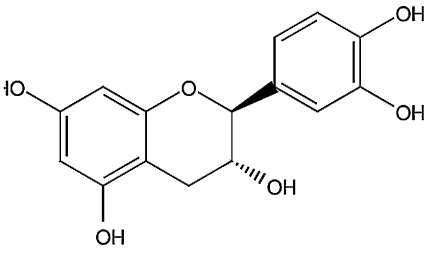
epigallocatechin-3-gallate (EGCG)



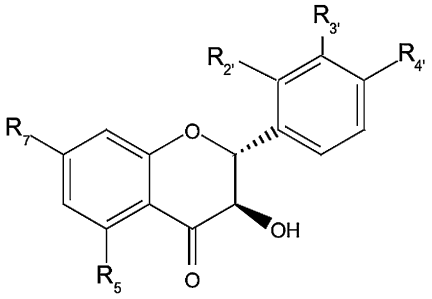
epicatechin-3-gallate (ECG)



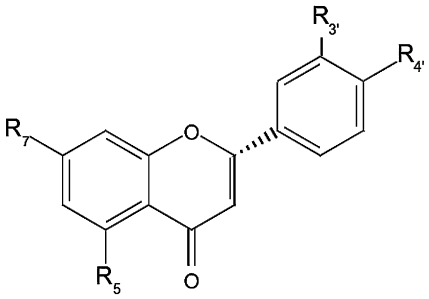
epigallocatechin (EGC)



epicatechin (EC)



Flavonol	R_5	R_7	$R_{2'}$	$R_{3'}$	$R_{4'}$
Quercetin	OH	OH	-	OH	OH
Morin	OH	OH	OH	-	OH
Kaempferol	OH	OH	-	-	OH
Galangin	OH	OH	-	-	-



Flavone	R_5	R_7	$R_{3'}$	$R_{4'}$
Luteolin	OH	OH	OH	OH
Apigenin	OH	OH	-	OH
Chrysin	OH	OH	-	
Diosmetin	OH	OH	OH	OCH_3

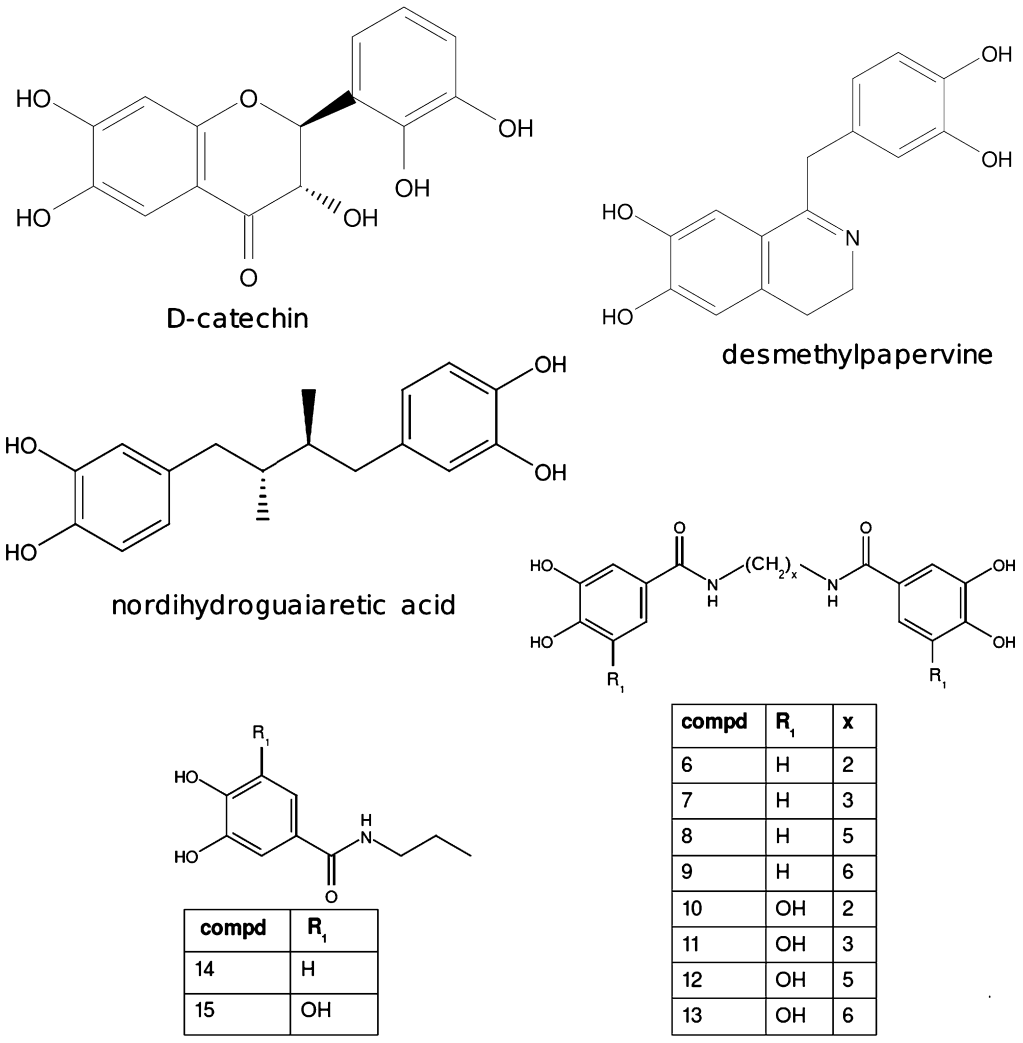
Fig. 5. Structure of some Type III COMT Inhibitors (natural product-based inhibitors). Structure of tea catechins (epigallocatechin gallate (EGCG), (−)-epicatechin-3-gallate (ECG), epigallocatechin (EGC) and (−)-epicatechin (EC)) and flavonoids (flavonols and flavones).

2.3.2. Ligand-based approaches

Ligand-based drug design approaches require molecular descriptor information of active molecules and mainly constitute methods like pharmacophore modeling and QSAR studies for the design of novel inhibitors. The workflow used for ligand-based approach is presented in Fig. 9. Here, COMT inhibitors (compiled from experimental investigations) were docked with crystal structure of human S-COMT to determine active ligands. For this, the crystal structure of protein was prepared using protein preparation wizard while the tautomers were generated for the inhibitors using Ligprep (Schrödinger) [61]. These were then docked to the protein structure using Glide XP docking (Schrödinger) [67] and binding free energy estimates were calculated for the docked complexes using MM-GBSA [66]. All the docked compounds were ranked in order of their binding energies. Top 200 compounds were selected for energy-optimized pharmacophore (E-pharmacophore) [69] modeling.

E-pharmacophore (Schrödinger) is generated by energetically optimized, structure-based pharmacophore. It combines the two strategies of ligand-based pharmacophore screening and structure-based docking. It begins with the refinement of ligand poses of ligand–receptor complex and then computes the Glide XP scoring terms and maps the energies onto atoms. The pharmacophore sites were then generated and the Glide XP energies were summed for the atoms that comprise each pharmacophore site. In the next stage, the pharmacophore sites are scored on the basis of energies and the top scoring sites were selected based on energetic ranking to generate a pharmacophore hypothesis that was then used for database screening. The pharmacophore generated was used to screen the pre-filtered CNS permeable ZINC database. Conformers were generated for all the tautomers in the database and matched for four out of five pharmacophoric features with a distance matching tolerance of 3.0. The fitness score was taken as:

$$\text{Fitness} = 1.0 \times \left(\frac{1.0 - \text{align score}}{1.2}\right) + 1.0 \times \text{vector score} + 1.0 \times \text{volume score}$$



compd	R_1	x
6	H	2
7	H	3
8	H	5
9	H	6
10	OH	2
11	OH	3
12	OH	5
13	OH	6

compd	R_1
14	H
15	OH

Fig. 6. Structure of some Type IV COMT Inhibitors (bifunctional inhibitors). Structure of first generation bifunctional COMT inhibitors (D-catechin, desmethylpapaverine and nordihydroxyguaiaretic acid) and derivatives of 3,4-dihydroxybenzamide and 3,4,5-trihydroxybenzamide (6–15).

where fitness is a score that measures how well is the alignment of the matching pharmacophore site points to those of the hypothesis. It is a measure of how well the matching conformation in the database superimposes with that of the reference ligand conformation. The hits were rejected with align score greater than 1.2, vector score less than −1.0 and a volume score less than 0.0. Also, receptor-based excluded volumes were integrated to refine screening process, which helps in reducing false positives by rejecting inactive compounds.

The hits generated by database screening were then docked to crystal structure of human S-COMT using Glide XP docking protocol [67]. Binding free energies were calculated for all the docked compounds using MM-GBSA [66]. The hits were then ranked according to binding free energy estimates. The ligand interaction diagram tool of Schrödinger was used to visualize the protein–ligand contacts of the top hit with the interacting residues.

2.4. Screening for ADME properties

In addition to the binding affinity studies, it is increasingly becoming apparent that ADME (Absorption, Distribution, Metabolism and Excretion) properties are crucial determinants of the ultimate clinical success of a drug [70,71]. Unfavorable ADME properties can lead to rejection of a drug in the later stages of drug discovery process [72]. Physiochemical properties important for ADME considerations were predicted using Qikprop (Schrödinger) [73] that calculates properties like molecular weight, molecular volume, dipole moment, no. of H-bond donors, no. of H-bond acceptors, no. of rotatable bonds, polar surface area, CNS activity, QPlogBB to filter out compounds with clear-cut undesirable properties.

The prerequisite was to neutralize the compounds before being used by QikProp. The neutralization step was carried out using Ligprep after which all the hits from both the approaches were processed for calculation of ADME properties.

3. Results and discussion

3.1. Structure and binding site analysis

COMT belongs to the class of SAM-dependent methyltransferase fold family [74,75]. The crystal structure of human COMT reported seven-stranded β-sheet core sandwiched between two sets of helices. The structure is characteristics of this family except that it contains a loop in between β-strands 6 and 7 which is part of catechol-binding site. The structures of both the 108V and 108M

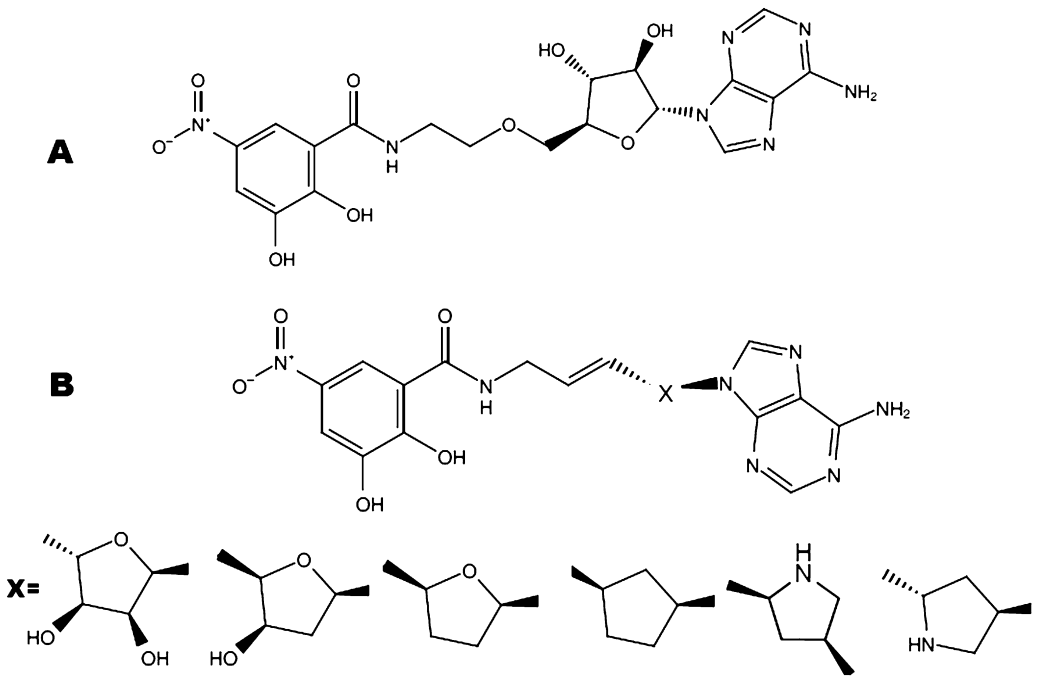


Fig. 7. Some Type V COMT Inhibitors (bisubstrate inhibitors). (A) Structure of first potent bisubstrate inhibitor and (B) its modifications in the ribose moiety.

forms of human S-COMT (Fig. 10) were found to be very similar (C^{α} RMSD = 0.2 Å) [23].

The binding site of COMT that constitutes a SAM binding site and a catechol binding pocket, was predicted by Q-SiteFinder [34] and CASTp [35] (Fig. 11). The catechol-binding site constitutes a shallow pocket while the SAM-binding pocket is deep seated. The adenine ring of SAM forms hydrogen bond with Ser119 and Gln120 and van der Waals interactions with residues Ile91, Gly117, Ala118, His142 and Trp143. The methionine portion of SAM shows hydrogen bonding with residues Asn41, Val42, Ser72 and Asp141 and hydrophobic interactions with Met40, Val42, Gly60, Tyr68 and Ile89. On the other hand, the catechol-binding site is formed by residues Met40, Leu198, Tyr200 and also the "gatekeeper" residues Trp38 and Pro174. The "gatekeeper" residues are involved in van der Waals interaction with DNC and thus hold the substrate in the correct position for methylation. Mg^{2+} is also a part of the active site where it is octahedrally coordinated to the side-chains of Asp141, Asp169 and Asn170 and also with one water molecule and two hydroxyl groups of the catechol.

Fig. 12A and B displays a schematic representation of the protein–ligand interactions of human S-COMT 108V and S-COMT 108M forming interactions with SAM, DNC and Mg^{2+}.

The polymorphic residue (108M/108V) is buried in a hydrophobic residue, around 16 Å away from the SAM-binding site. The crystal structure reports that the interactions in both the SAM- and catechol-binding sites are virtually identical in both the polymorphic forms but there still exist intramolecular packing differences in the proximity of the polymorphic residue. 108M interacts more closely with residues Ala22, Arg78, Val103 and Ala106 than does 108V, resulting in a 0.7 Å displacement between Cα in the backbone structure, which propagates toward the SAM-binding site. Structurally, the 108M form is thermally unstable, more prone to inactivation, has decreased enzyme activity and protein levels *in vivo*. The co-substrate SAM stabilizes the 108M form against both thermal inactivation and denaturation. Molecular dynamics (MD) simulations of 108M in the absence of SAM reported that structural differences occur at the polymorphic site which extends throughout the protein becoming more prominent around the catalytic site [76,77].

3.2. Structure-based and ligand-based approaches

A combination of both structure- and ligand-based approaches were employed to screen novel inhibitors against both 108V and 108M forms of COMT.

3.2.1. Structure-based approaches

Virtual screening was employed to screen the ZINC CNS permeable database (435 388 compounds) for the generation of novel hits against human S-COMT. Fig. 13 shows the workflow of the methodology that we used for structure-based virtual screening.

A total of 718 180 tautomers were generated which were docked to both human COMT-108V and 108M. The ligands were first subjected to HTVS from where a total of 96 188 structures were passed onto the next stage in case of 108V and a total of 99 205 in case of 108M. The next stage used SP docking protocol to screen the database from where a total of 8785 compounds for 108V and 10 074 for 108M were passed on to the XP docking stage. The last stage retained only 661 compounds for 108V and 877 compounds in case of 108M. All of them were processed for binding energy estimates using MM-GBSA. The ligands were ranked according to binding free energy values. The docking scores (XP Gscore) and binding free energy estimates of the top five compounds docked to both 108V and 108M are listed in Table 2. The compounds were ranked according to binding free energy values and all of them displayed a docking score in the range of −12 to −8 kcal/mol.

The top five compounds docked to COMT-108V and 108M in the active site are depicted in Fig. 14. These top docked compounds were found to superimpose at the same site as the co-crystallized ligand in the crystal structure of both 108V and 108M. The two-dimensional structures of these top hits are represented in Fig. 15. Structural analyses of the top hits clearly showed the presence of two to three-ring systems connected by a spacer region consisting

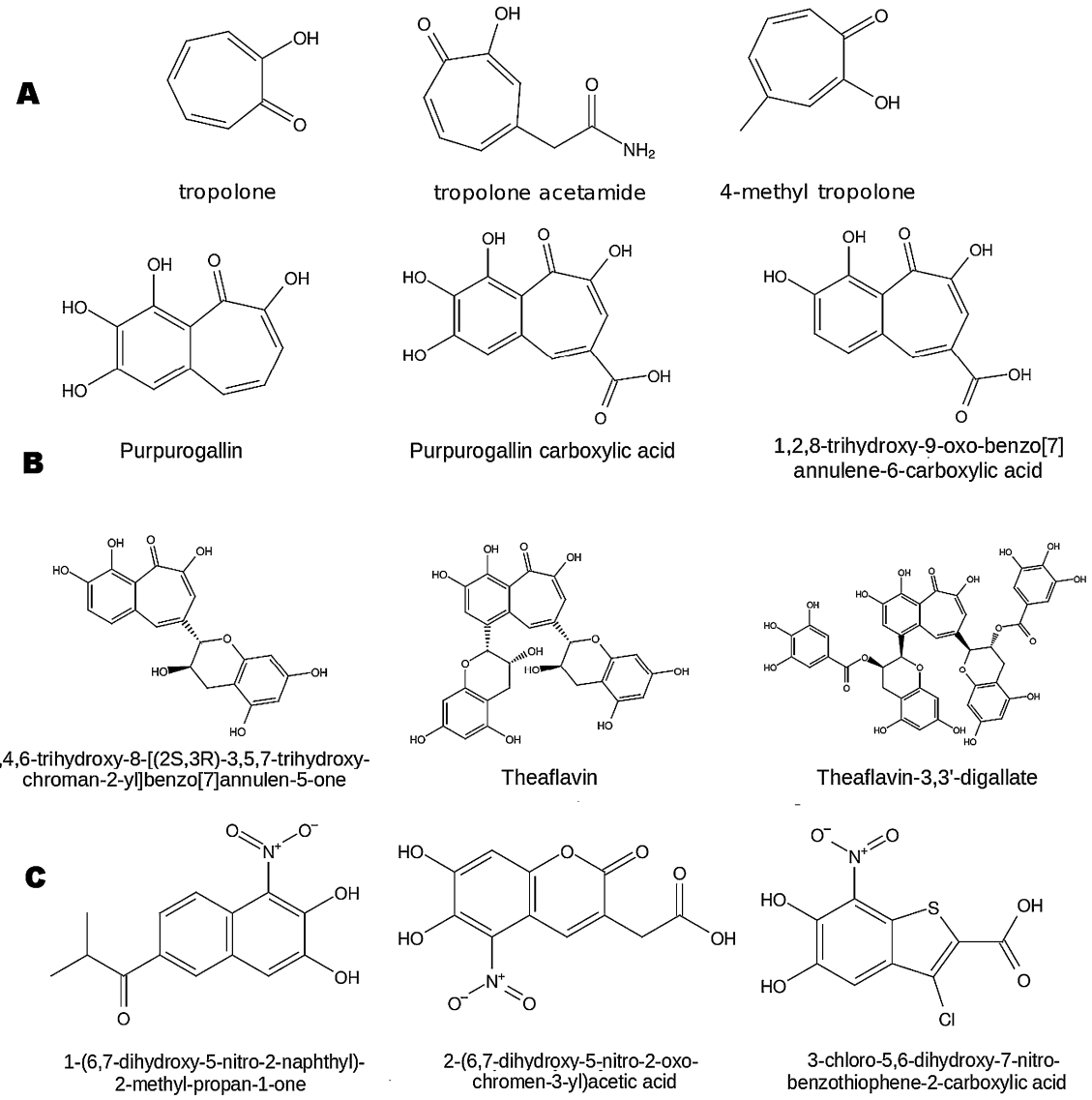


Fig. 8. Chemical structure of some Type VI COMT Inhibitors. (A) Tropolones, (B) benzotropolones, and (C) bicyclic compounds.

of an amide group in both the polymorphic forms of the target protein.

The non-covalent interactions of the top hit with the amino acid residues in the active site of 108V and 108M are displayed in Fig. 16. Analyses of the key amino acid residues interacting with the top hits in both 108V and 108M forms showed Ser119 forming hydrogen bond with nitrogen of the pyridine ring in the top hit (ZINC25947810) in case of 108V and with oxygen of benzofuran ring in the top hit (ZINC32457316) in 108M. While Trp38, Met40, Ala67, Tyr68, Ile89, Ile91, Cys95, Ala118, Trp143, Pro174 and Leu198 were found to be involved in hydrophobic interactions; His142 and Trp143 were shown to form π-stacking interactions with the ligand molecule. The top hits in both the polymorphic forms also show interactions with the Mg^{2+} ion in the binding site of the protein. Visual inspection of the protein-ligand complexes revealed that the top hits seem to bind at the same site in both the polymorphic forms of the protein.

Pre-filtered ZINC database utilizes size, polar surface area and lipophilicity to select CNS-active agents. It is believed that CNS activity is a complex function of physiochemical parameters of

Table 2
Table displaying the XP Gscore and binding free energy values of top 5 compounds docked to COMT-108V and 108M.

108V			108M		
Title (108V)	XP GScore[a]	Binding free energy estimates[a]	Title	XP GScore[a]	Binding free energy estimates[a]
ZINC25947810	−9.56	−94.03	ZINC32457316	−8.44	−101.18
ZINC27680940	−8.79	−87.42	ZINC32618117	−8.27	−98.30
ZINC20536627	−8.39	−83.93	ZINC27474029	−8.08	−98.15
ZINC32851011	−11.80	−82.88	ZINC05267754	−8.80	−94.89
ZINC05265226	−8.28	−81.24	ZINC27680940	−8.71	−94.57

[a] Expressed as kcal/mol.

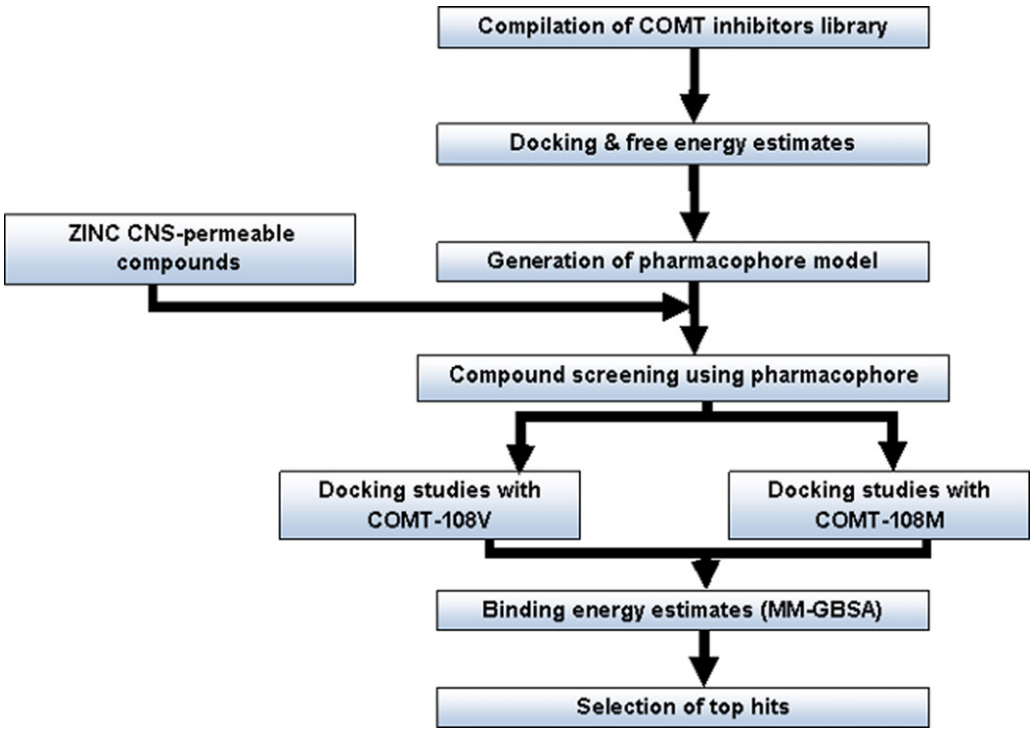


Fig. 9. Flowchart depicting the methodology for ligand-based drug design.

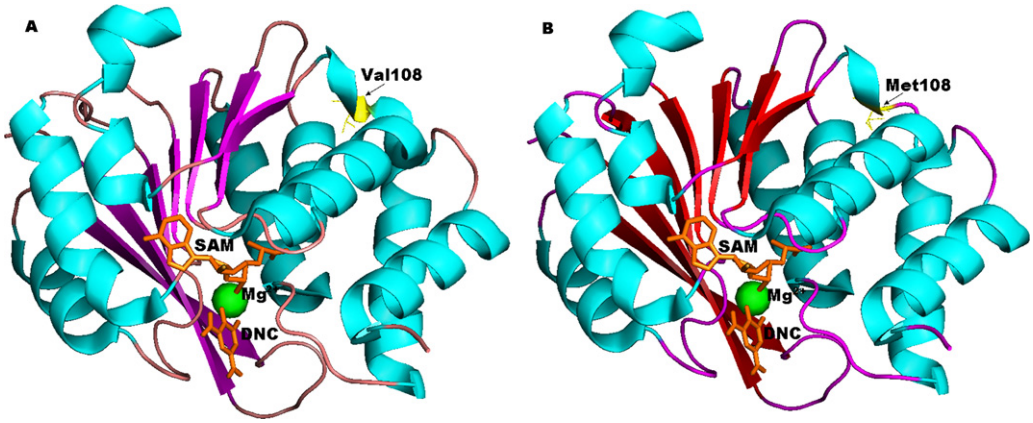


Fig. 10. Crystal structure of human S-COMT. Structure of human S-COMT (A) 108V and (B) 108M bound to SAM, dinitrocatechol (DNC) and Mg^{2+}.

Table 3
Physiochemical descriptors of the top 5 hits docked to COMT-108V obtained using Virtual screening.

Title	No. of rotatable bonds[a]	CNS[b]	Molecular weight	Dipole moment	Volume[c]	HB donors[d]	HB acceptors[e]	QPlogBB[f]	PSA[g]
ZINC25947810	9	1	375.47	7.32	1267.54	1	6.5	−0.37	62.52
ZINC27680940	3	0	326.78	7.36	1042.13	1	5	−0.40	64.03
ZINC20536627	4	1	394.47	5.04	1283.46	1	7	−0.16	63.66
ZINC32851011	9	0	282.38	3.00	1032.97	3	2.5	−0.62	54.79
ZINC05265226	3	0	321.40	7.81	984.58	1	4	−0.32	54.28

[a] Number of non-trivial (not CX3), non-hindered (not alkene, amide, small ring) rotatable bonds.
[b] Predicted central nervous system activity on a −2 (inactive) to +2 (active) scale.
[c] Total solvent-accessible volume in cubic angstroms using a probe with a 1.4 Å radius.
[d] Estimated number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer.
[e] Estimated number of hydrogen bonds that would be accepted by the solute from water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer.
[f] Predicted brain/blood partition coefficient.
[g] Van der Waals surface area of polar nitrogen and oxygen atoms in square angstroms.

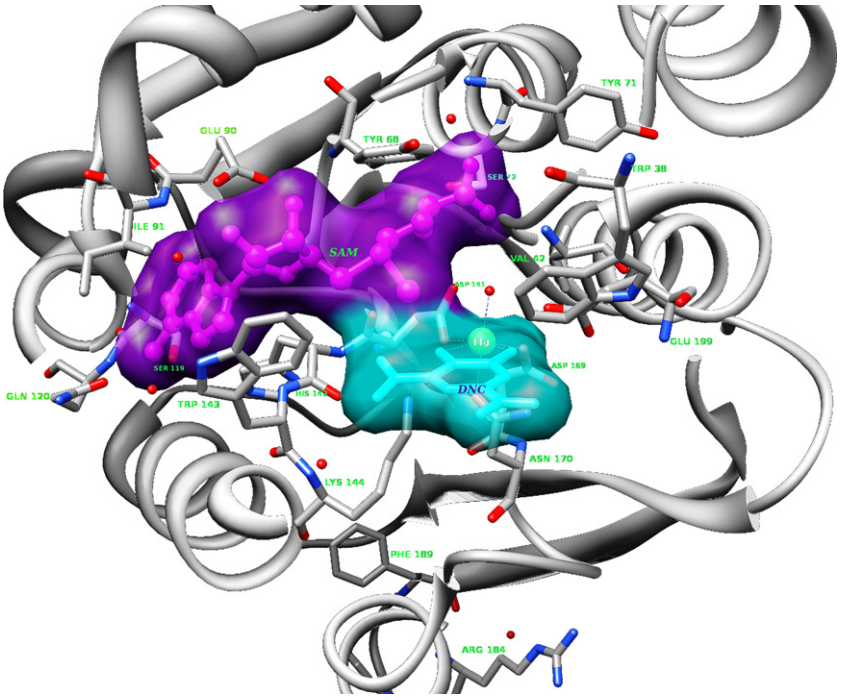


Fig. 11. Protein–ligand interactions highlighting amino acid residues in the binding site of DNC (shown in blue) and SAM (shown in pink) in human S-COMT. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

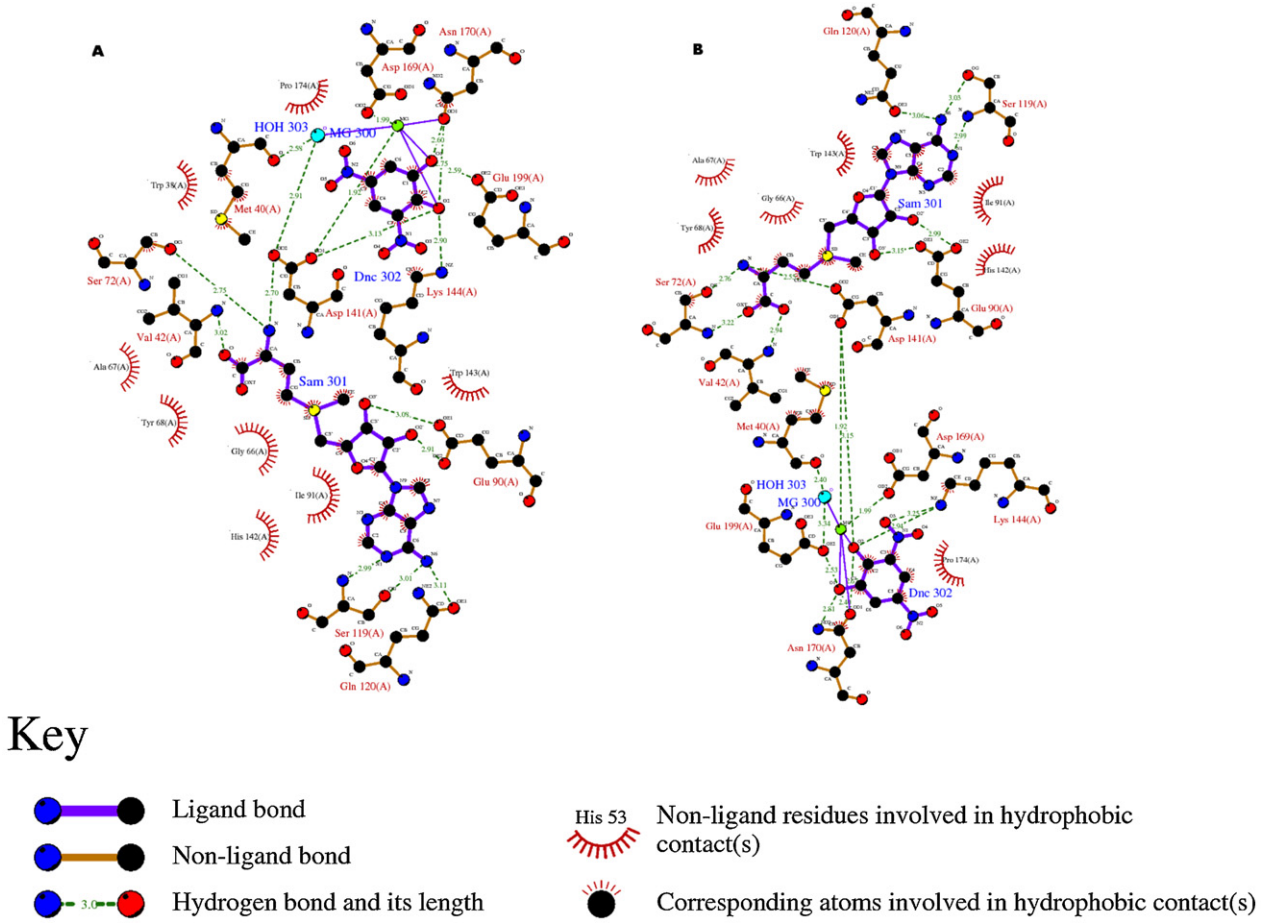


Fig. 12. Interaction Map depicting protein–ligand contacts in human S-COMT. Ligplot showing the protein–ligand contacts in (A) 108V and (B) 108M complexed with SAM, DNC and Mg^{2+}. SAM is being represented by Sam301, DNC by Dnc302 and Mg^{2+} by MG 300. The figure also contains a water molecule which is involved in coordination with Mg^{2+} and is being represented by HOH 303.

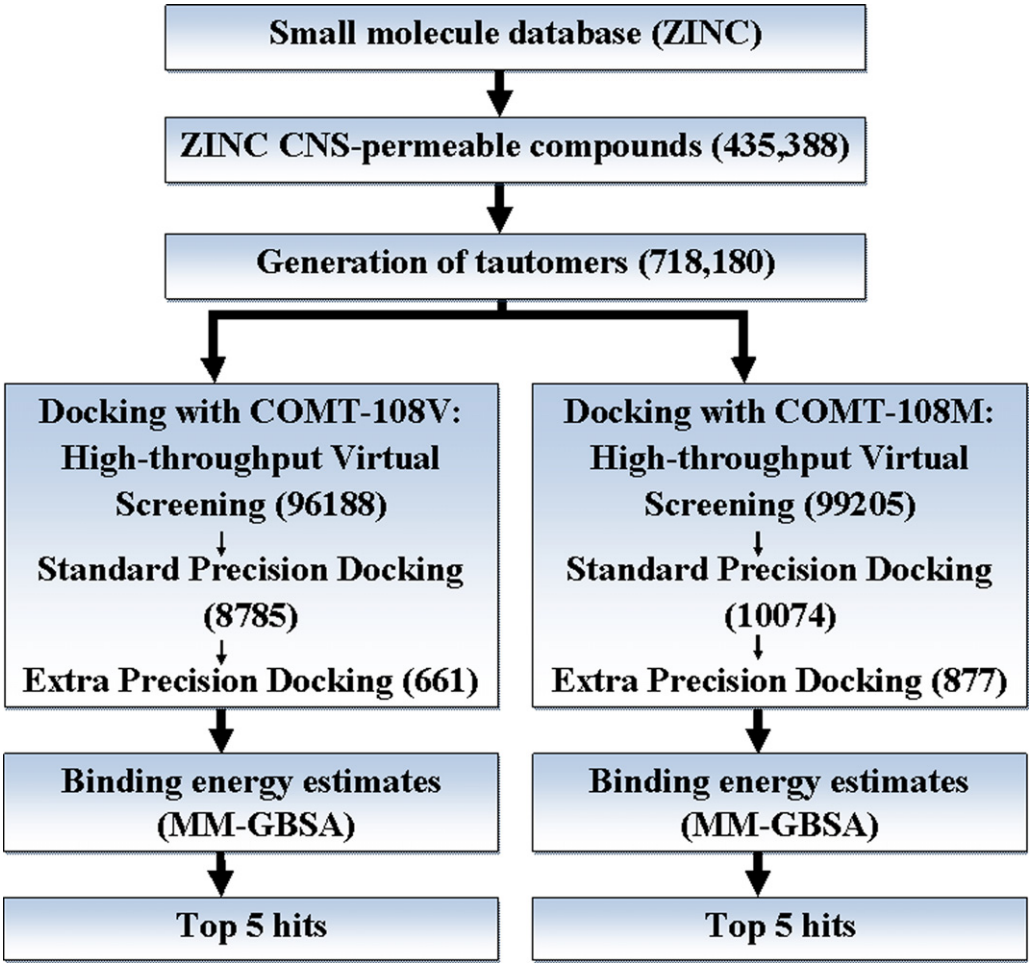


Fig. 13. Schematic representation of structure-based virtual screening for COMT-108V and COMT-108M. The number in the bracket indicates the number of compounds retained at that stage.

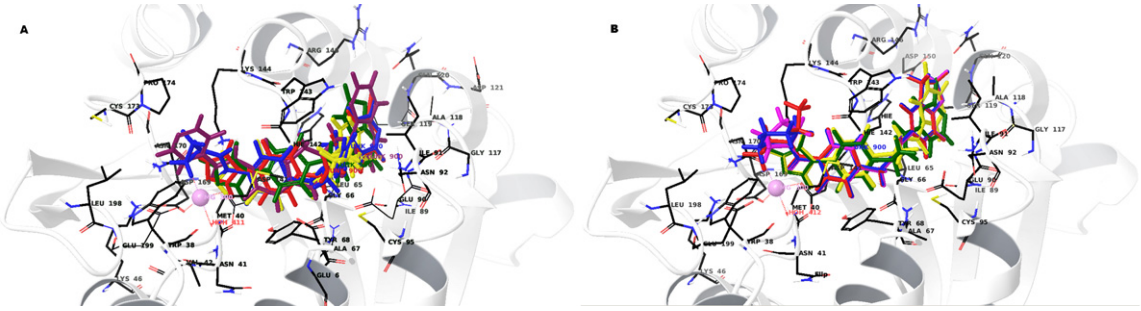


Fig. 14. Top hits from structure-based virtual screening approach. Docked conformation of the top 5 hits which were overlapped in the active site of (A) COMT-108V and (B) COMT-108M. The figure also displays interactions with Mg^{2+} (shown in pink sphere) in the active site. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

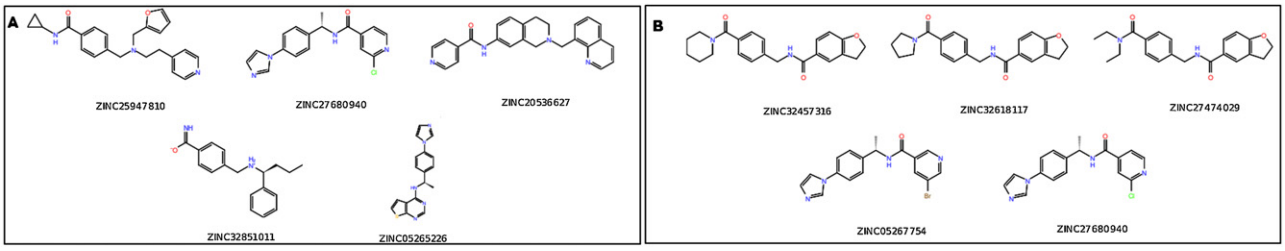


Fig. 15. The 2D structures of the top 5 hits along with their ZINC ID's obtained from structure-based approach for (A) COMT-108V and (B) COMT-108M.

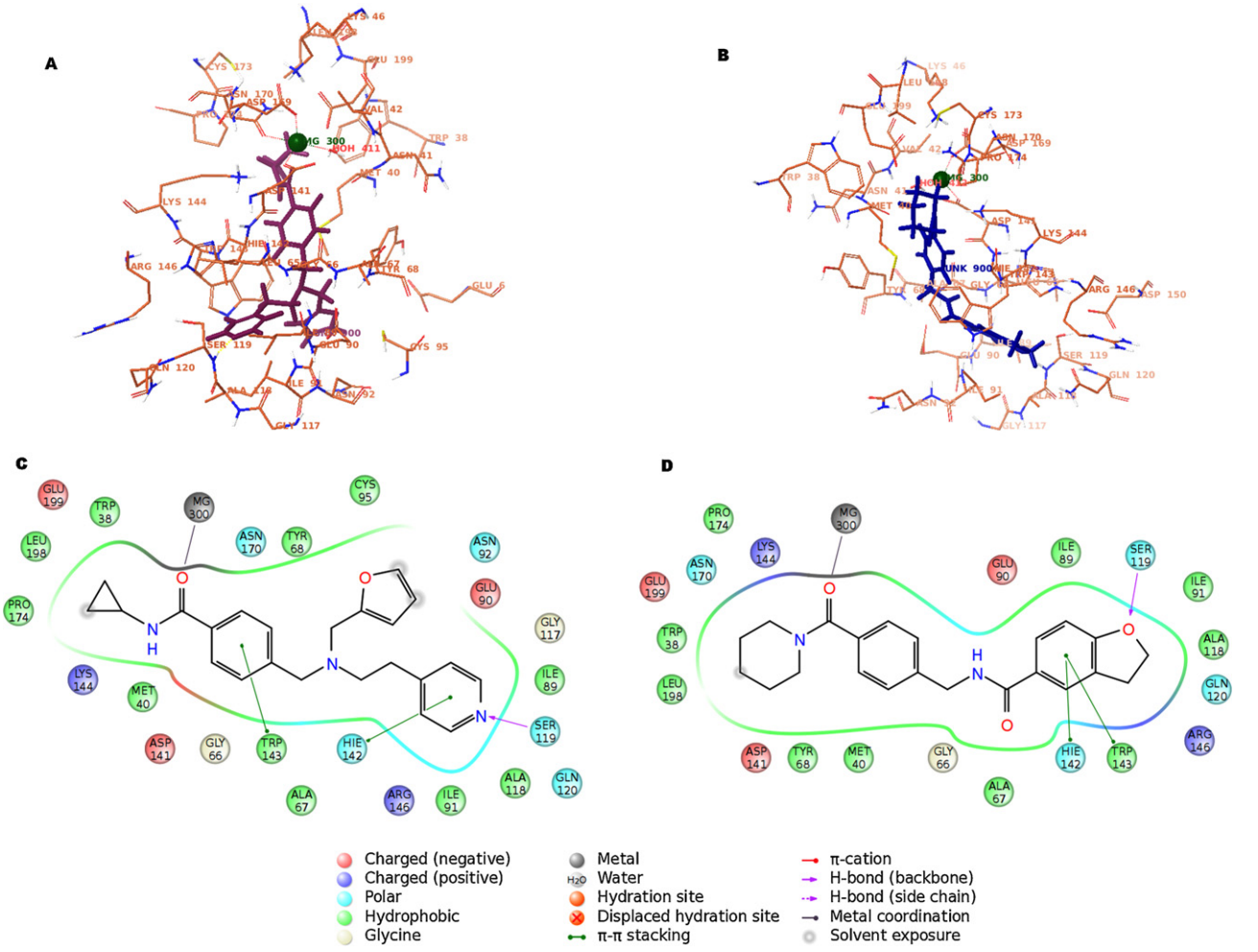


Fig. 16. Protein–ligand contacts of human S-COMT with the top hit obtained using virtual screening for (A) COMT-108V and (B) COMT-108M. Ligand-interaction diagram representing protein–ligand of the top hit in two-dimension with (C) COMT-108V and (D) COMT-108M.

molecules such as size, lipophilicity, H-bond donors and acceptors, rotatable bonds, polar surface area, aqueous solubility and many more, hence, detailed analyses of ADME properties of top hits were carried out. Tables 3 and 4 shows the ADME properties for the top 5 hits docked to COMT-108V and 108M. The compounds were found to lie well within the allowed range of CNS drugs [60].

3.2.2. Ligand-based approaches

Ligand-based approach was also utilized for the generation of novel inhibitors against COMT. The workflow of the methodology used is shown in Fig. 17. Energy-optimized pharmacophore was used which combines attributes of both structure-based and ligand-based techniques. Structural and energetic information computed from the Glide XP scoring function was used to identify the important pharmacophoric features. The energetic terms were derived from protein–ligand contacts of active ligands which were identified by docking of all known COMT inhibitors with the crystal structure of both COMT-108V and 108M. A total of 9877 tautomers were generated for 806 COMT inhibitors compiled from the literature. Out of a total of 9877 tautomers, 9876 docked to

Table 4
Physiochemical descriptors of the top 5 hits docked to COMT-108M obtained using Virtual screening.

Title	No. of rotatable bonds[a]	CNS[b]	Molecular weight	Dipole moment	Volume[c]	HB donors[d]	HB acceptors[e]	QPlogBB[f]	PSA[g]
ZINC32457316	4	0	364.44	7.22	1206.39	1	6.25	−0.53	74.21
ZINC32618117	4	0	350.42	8.98	1154.32	1	6.25	−0.58	74.95
ZINC27474029	6	0	352.43	7.67	1178.51	1	6.25	−0.60	71.55
ZINC05267754	3	0	371.24	7.40	1050.35	1	5.5	−0.44	64.28
ZINC27680940	3	0	326.79	9.12	1042.10	1	5	−0.40	64.04

[a] Number of non-trivial (not CX3), non-hindered (not alkene, amide, small ring) rotatable bonds.
[b] Predicted central nervous system activity on a −2 (inactive) to +2 (active) scale.
[c] Total solvent-accessible volume in cubic angstroms using a probe with a 1.4 Å radius.
[d] Estimated number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer.
[e] Estimated number of hydrogen bonds that would be accepted by the solute from water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer.
[f] Predicted brain/blood partition coefficient.
[g] Van der Waals surface area of polar nitrogen and oxygen atoms in square angstroms.

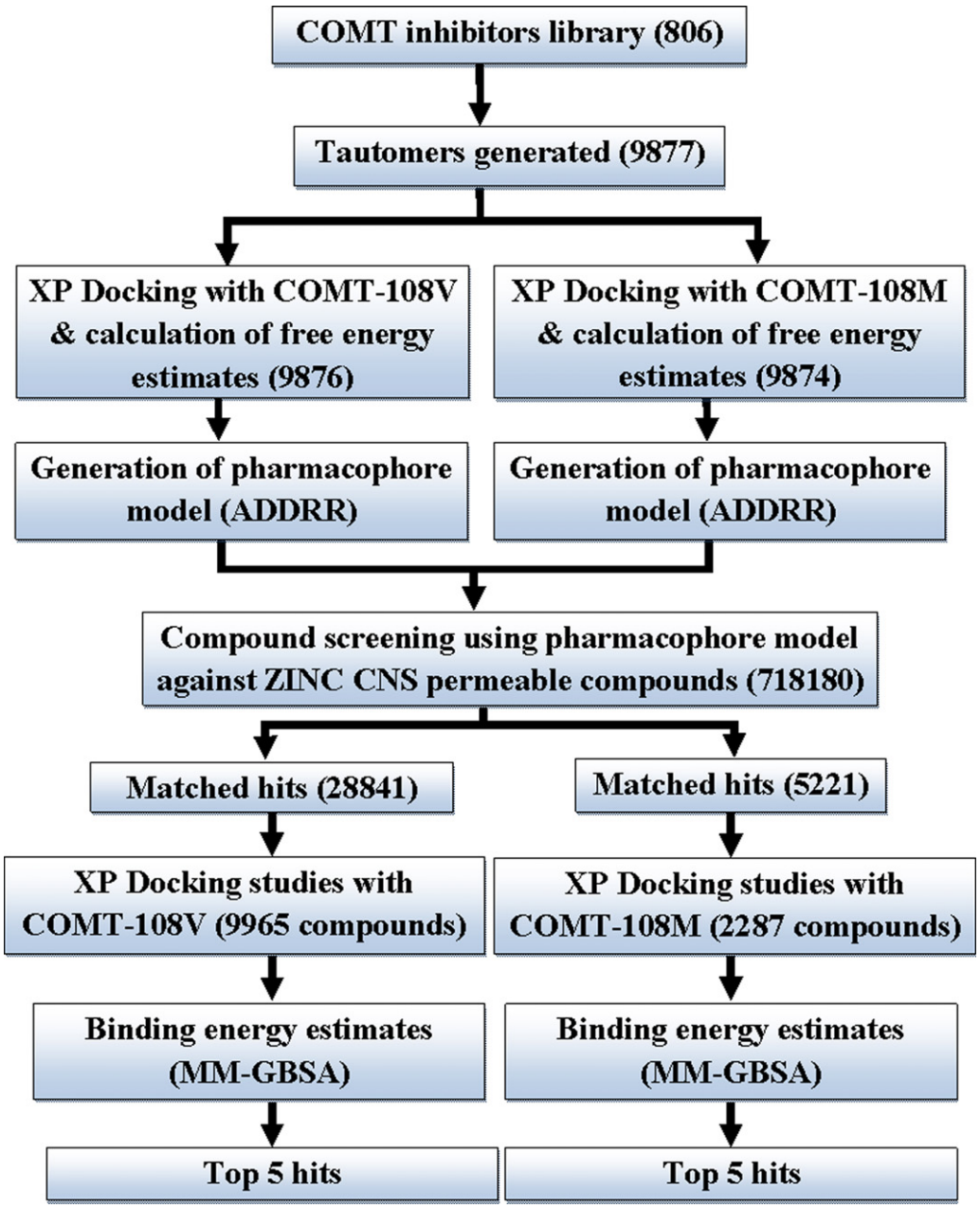


Fig. 17. Schematic representation of ligand-based drug design approach for COMT-108V and COMT-108M. The numbers in the bracket indicates the number of compounds retained at that stage.

Table 5
Table displaying the XP Gscore, binding free energy and pIC50 values of top 5 bisubstrate inhibitors docked to 108V.

Title	XP GScore[a]	Binding free energy estimates[a]	pIC50 values
3-[[(E)-3-[(2S,3R,4R,5S)-5-(6-aminopurin-9-yl)-3,4-dihydroxy-tetrahydrofuran-2-yl]prop-2-enyl]carbamoyl]-2-hydroxy-5-(p-tolylsulfonyl)phenolate	−10.70	−95.25	6.67
N-[3-[(1S,2S,3S,4S)-4-(6-aminopurin-9-yl)-2,3-dihydroxy-cyclopentyl]propyl]-2,3-dihydroxy-5-(p-tolyl)benzamide	−9.00	−94.90	>3
N-[(E)-3-[(1R,2S,3S,4S)-4-(6-aminopurin-9-yl)-2,3-dihydroxy-cyclopentyl]prop-2-enyl]-2,3-dihydroxy-5-(p-tolyl)benzamide	−8.22	−93.09	>3
3-[[(E)-3-[(2R,3R,4S,5S)-5-(6-aminopurin-9-yl)-3,4-dihydroxy-tetrahydrofuran-2-yl]prop-2-enyl]carbamoyl]-5-(cyclohexanecarbonyl)-2-hydroxy-phenolate	−8.15	−91.68	7.08
3-[3-[(1S,2S,3S,4S)-4-(6-aminopurin-9-yl)-2,3-dihydroxy-cyclopentyl]propylcarbamoyl]-2-hydroxy-5-nitro-phenolate	−9.03	−91.65	>3

[a] Expressed as kcal/mol.

Table 6
Table displaying the XP Gscore binding free energy and pIC50 values of top 5 bisubstrate inhibitors docked to 108M.

Title	XP GScore[a]	Binding free energy estimates[a]	pIC50 values
N-[(E)-3-[(2R,4R,5R)-5-(6-aminopurin-9-yl)-4-hydroxy-tetrahydrofuran-2-yl]prop-2-enyl]-2,3-dihydroxy-5-(p-tolyl)benzamide	−10.81	−111.01	6.74
N-[(E)-3-[(2R,3R,4R,5R)-5-(6-aminopurin-9-yl)-3,4-dihydroxy-tetrahydrofuran-2-yl]prop-2-enyl]-5-(cyclohexanecarbonyl)-2,3-dihydroxy-benzamide	−10.62	−110.02	7.08
N-[(E)-3-[(2R,3S,4S,5R)-5-(6-aminopurin-9-yl)-3,4-dihydroxy-tetrahydrofuran-2-yl]prop-2-enyl]-5-(1,3-benzothiazol-2-yl)-2,3-dihydroxy-benzamide	−11.17	−107.78	7.54
N-[(E)-3-[(1S,2R,3R,4R)-4-(6-aminopurin-9-yl)-2,3-dihydroxy-cyclopentyl]prop-2-enyl]-2,3-dihydroxy-5-(p-tolyl)benzamide	−11.06	−107.12	>3
N-[(E)-3-[(2R,3S,4R,5R)-5-(6-aminopurin-9-yl)-3,4-dihydroxy-tetrahydrofuran-2-yl]prop-2-enyl]-2,3-dihydroxy-5-(4-methylbenzoyl)benzamide	−11.70	−106.75	7.47

[a] Expressed as kcal/mol.

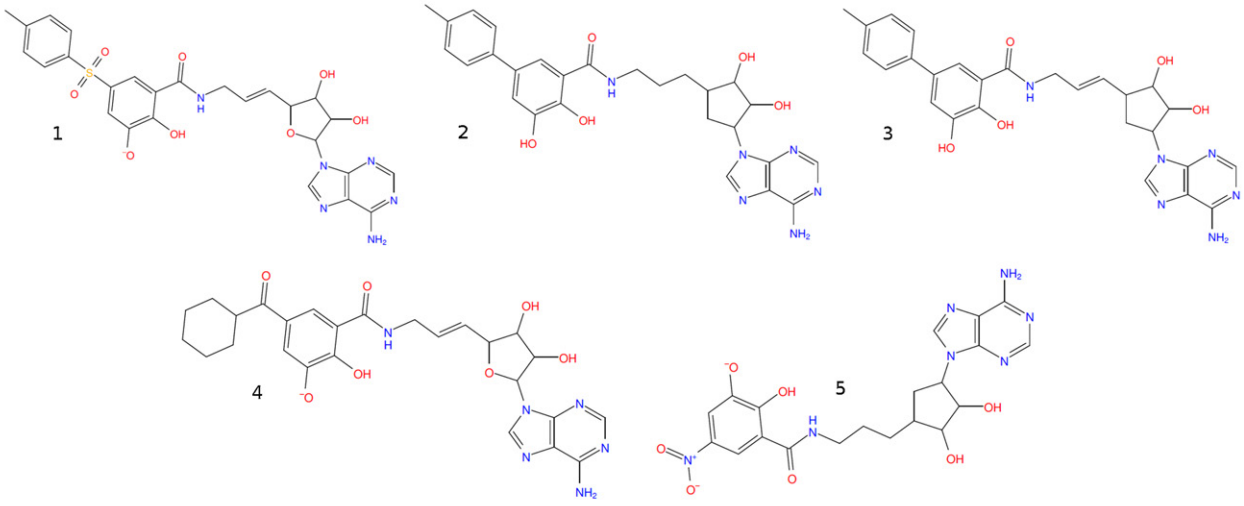


Fig. 18. Chemical structure of top docked compounds to COMT-108V.

COMT-108V and 9874 docked to COMT-108M. The ligands were ranked according to binding energy estimates computed using MM-GBSA. It was observed that the top docked compounds mostly belong to bisubstrate (Type V) class of inhibitors. The docking scores, binding free energy and pIC50 (−log IC50) values [55,78] of the top bisubstrate inhibitors docked are presented in Tables 5 and 6 for 108V-COMT and 108M-COMT, respectively. Figs. 18 and 19 shows the two-dimensional structure of top docked bisubstrate inhibitors to COMT-108V and 108M, respectively. Binding energy estimates of the docked complexes show that bisubstrate inhibitors rank better than the clinically used nitrocatechols. This is also desirable because the nitrocatechols used in treatment possess many gastro-intestinal and dopaminergic side-effects and is a cause of treatment discontinuation. Furthermore, it has been reported that bisubstrate inhibitors show effective inhibition in spite of eliminating the nitro group of the catechol structural

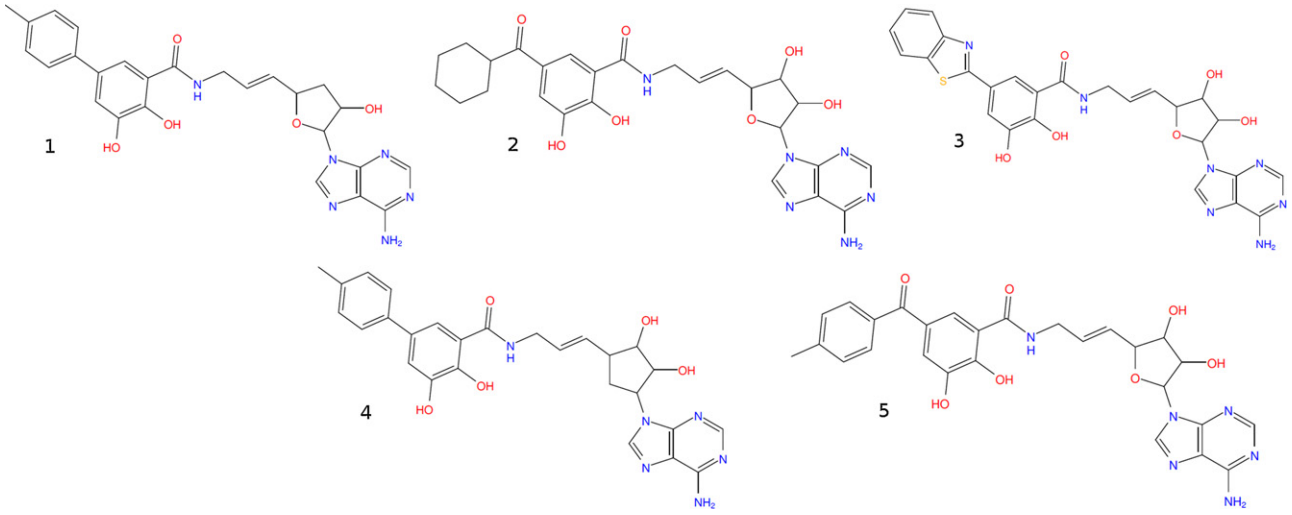


Fig. 19. Chemical structure of top docked compounds to COMT-108M.

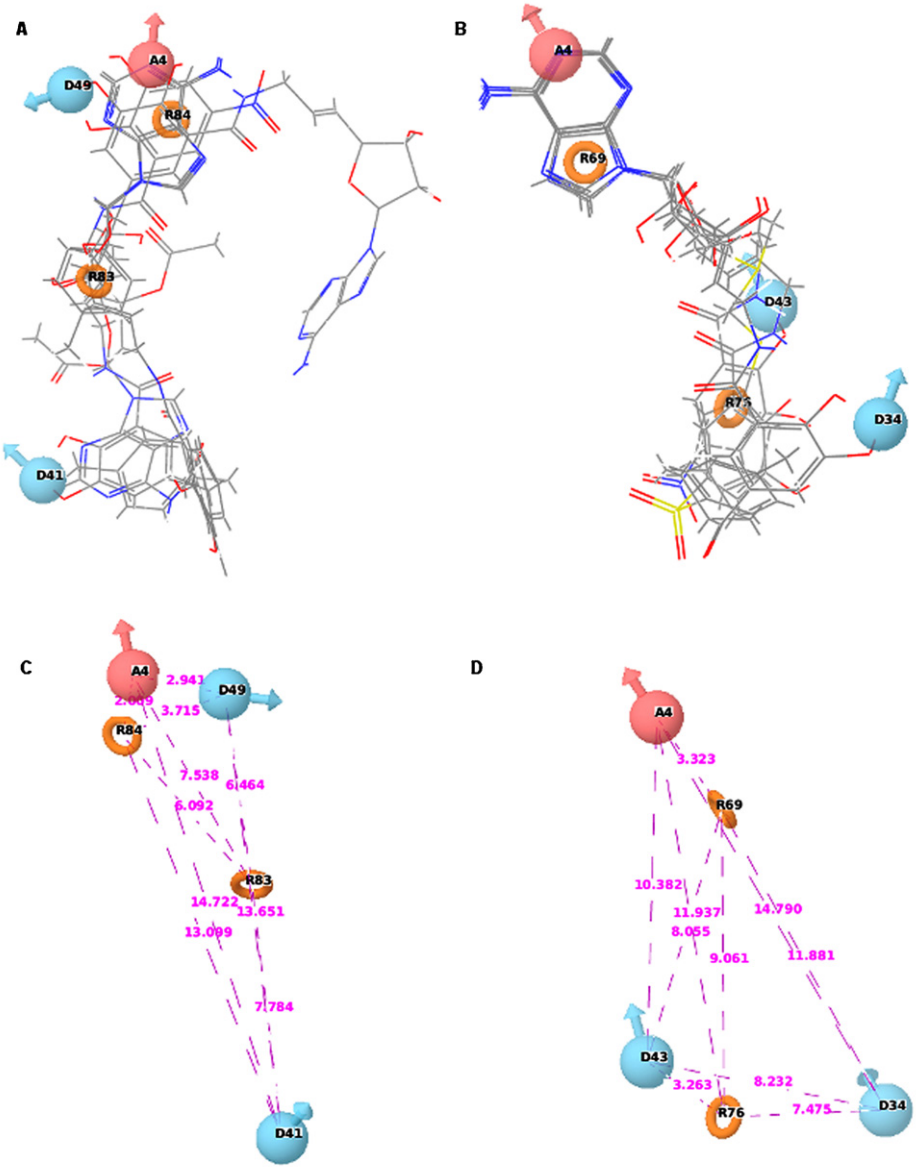


Fig. 20. Pharmacophore hypothesis (ADDRR). Five-point pharmacophore predicted for (A) COMT-108V and (B) COMT-108M. The distance between the pharmacophoric features (in Å) is shown in (C) COMT-108V and (D) COMT-108M. Orange torus indicate aromatic ring feature, pink sphere indicates acceptor feature, light blue sphere indicate donor feature. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

unit [78]. On the other hand, bisubstrate inhibitors that occupy both the SAM and catechol-binding sites, generate more interactions with the target resulting in improved affinity and specificity. The interactions are also stabilized by coordination with Mg^{2+} ion. Therefore, top scored compounds including mainly bisubstrate inhibitors were taken up for the generation of E-pharmacophore. XP descriptor information was generated for the top docked compounds which were used for prediction of pharmacophoric features. Seven pharmacophoric sites were predicted out of which only five were retained based on the score. The final hypothesis contains one hydrogen-bond acceptor (A), two hydrogen-bond donor (D) and two aromatic rings (R) for both 108V and 108M (Fig. 20). Although, the pharmacophore predicted for both forms were same, however, it displayed difference in pharmacophoric features in terms of bond angles and bond distances.

Taking these pharmacophore hypotheses, database screening was performed against prepared ZINC CNS-permeable dataset. Conformers were generated for all the tautomers in the dataset and matched for four out of five pharmacophoric sites. Compounds with fitness score of greater than 1.0 were retained which resulted in around 28 841 matches for COMT-108V and 5221 for COMT-108M. These were then subjected to Glide XP docking protocol with human COMT-108V and 108M.

XP docking resulted in around 9965 compounds for 108V and 2287 for 108M form of protein. The docked compounds were then ranked according to binding free energy estimates computed using MM-GBSA. The XP Gscore and binding energy of the top 5 compounds docked to COMT-108V and COMT-108M are reported in Tables 7 and 8, respectively.

The binding poses of the top five hits within the active site of both COMT-108V and 108M are illustrated in Fig. 21 and their chemical structures are represented in Fig. 22. The protein–ligand interactions of the top hit with the crystal structure of COMT-108V and 108M are shown in Fig. 23. Structural analyses of the top hits clearly show the presence of two-ring systems linked by a spacer for both forms of COMT.

Table 7
Table displaying the XP Gscore and binding free energy values of top 5 hits docked to 108V obtained after pharmacophore analysis.

Title	Align score	Vector score	Volume score	Fitness score	XP GScore[a]	Binding free energy estimates[a]
ZINC28624908	0.82	0.77	0.11	1.19	−7.56	−99.88
ZINC07129527	0.70	0.67	0.11	1.20	−7.36	−99.77
ZINC13679633	0.68	0.72	0.12	1.27	−5.46	−95.86
ZINC23115185	0.97	0.90	0.13	1.21	−9.40	−93.86
ZINC14399420	0.79	0.57	0.11	1.02	−6.34	−93.84

[a] Expressed as kcal/mol.

Table 8
Table displaying the XP Gscore and binding free energy values of top 5 hits docked to 108M obtained after pharmacophore analysis.

Title	Align score	Vector score	Volume score	Fitness score	XP GScore[a]	Binding free energy estimates[a]
ZINC32235251	0.62	0.64	0.22	1.34	−7.59	−97.50
ZINC23740255	0.93	0.66	0.12	1.01	−7.51	−96.40
ZINC26926448	0.86	0.60	0.15	1.03	−7.10	−95.06
ZINC24804757	1.09	0.80	0.15	1.04	−6.34	−94.74
ZINC24148834	0.99	0.67	0.18	1.03	−6.51	−94.59

[a] Expressed as kcal/mol.

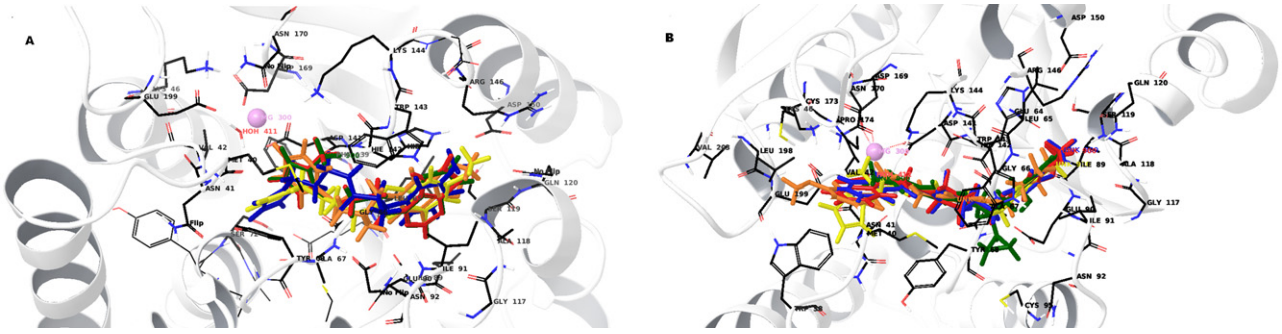


Fig. 21. Top hits from ligand–based pharmacophore search. Docked conformation of the top 5 hits overlapping in the active site of (A) COMT-108V and (B) COMT-108M, displaying protein–ligand contacts and interaction with the Mg^{2+} ion.

The protein–ligand interactions clearly depicts Ser119 and Ile91 forming H-bonds with nitrogen atoms of the thieno[2,3-d]pyrimidine ring and Glu90 forming H-bond with the nitrogen of amine group attached to the thieno[2,3-d]pyrimidine ring in the top hit (ZINC28624908) in case of 108V whereas in case of 108M, only Ile91 forms H-bond with one of the oxygen of the methoxy group attached to the phenyl ring in the top hit (ZINC32235251). The residues involved in hydrophobic interactions include Trp38, Met40, Leu65, Ala67, Tyr68, Ile89, Ile91, Ala118, Trp143, Cys173, Pro174, and Leu198 in both the polymorphic forms. Analyses of protein–ligand complex also show π-stacking interactions mediated by His142 and Trp143 in 108V and by Trp143 in 108M. Additionally, Lys144 forms π-cation interactions with the indole ring in the top hit (ZINC32235251) in 108M. The top hits also seem to interact with the Mg^{2+} ion in the binding site of both 108V and 108M.

Physicochemical properties were also calculated for all the compounds resulting from docking studies of screened compounds with COMT-108V and 108M. The ADME properties of the top 5 hits for both the polymorphic forms are presented in Tables 9 and 10, respectively. The values of all the properties were found to lie in the permissible range of CNS drugs [60].

The top hits generated by both the approaches show a better binding affinity than the clinically used nitrocatechols, as observed from their binding energy values. A comparison of top 50 compounds obtained by structure-based approach in both allelic forms (COMT-108M and COMT-108V) showed 21 common hits. On the other hand, only 3 common hits resulted when top 50 compounds obtained by pharmacophore-based approach, were compared in COMT-108V and COMT-108M. These common hits can be taken up for further studies to investigate their biological relevance as potent inhibitors.

In summary, this study reports the top hits from both the structure-based and ligand-based approaches showing significant binding affinity to the target protein. The top hits also possess physicochemical properties that lie well within the desired range

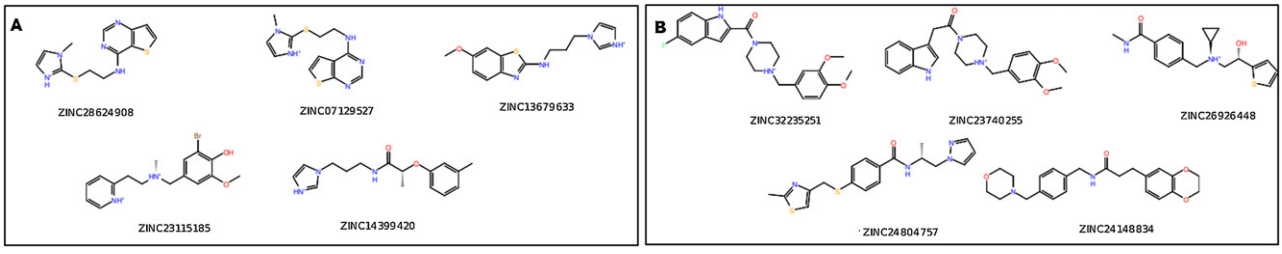


Fig. 22. The 2D structures of the top 5 hits along with their ZINC ID's investigated by pharmacophore search and docked with (A) COMT-108V and (B) COMT-108M.

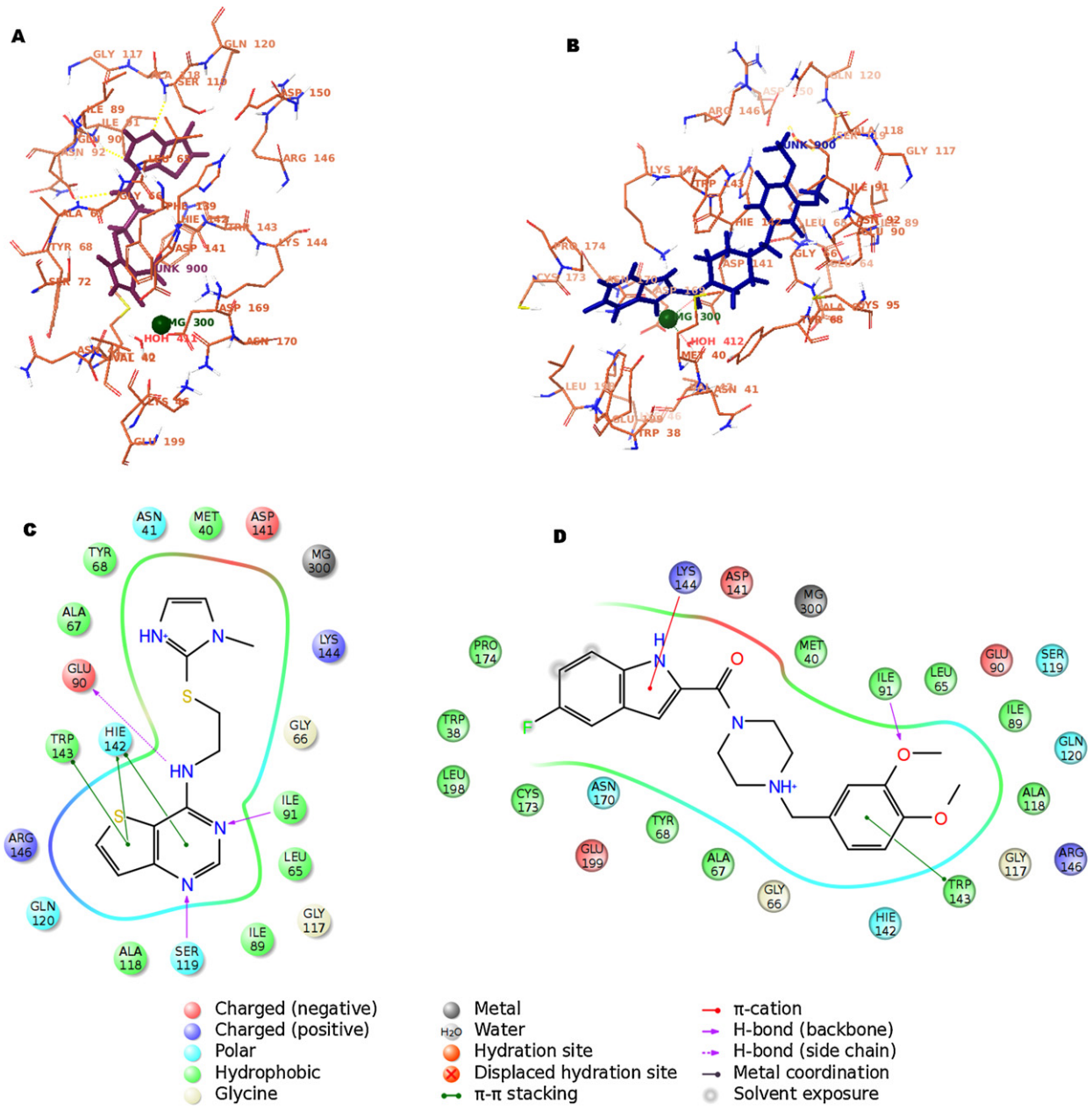


Fig. 23. Protein–ligand contacts of protein with the top hit obtained using ligand-based drug design for (A) COMT-108V and (B) COMT-108M. Ligand-interaction diagram depicting the protein–ligand interactions of the top hit in two-dimension with (C) COMT-108V and (D) COMT-108M.

Table 9
Physiochemical descriptors of top 5 hits docked to COMT-108V obtained from pharmacophore analysis.

Title	No. of rotatable bonds[a]	CNS[b]	Molecular weight	Dipole moment	Volume[c]	HB donors[d]	HB acceptors[e]	QPlogBB[f]	PSA[g]
ZINC28624908	5	0	291.388	5.52	917.34	1	4	−0.37	52.43
ZINC07129527	5	0	291.388	8.04	903.43	1	4	−0.33	52.17
ZINC13679633	6	0	288.367	4.89	955.04	1	4.75	−0.46	50.03
ZINC23115185	7	1	351.242	4.66	1000.25	1	4.5	0.24	46.15
ZINC14399420	7	0	287.361	6.75	1050.44	1	5.25	−0.65	59.96

[a] Number of non-trivial (not CX3), non-hindered (not alkene, amide, small ring) rotatable bonds.
[b] Predicted central nervous system activity on a −2 (inactive) to +2 (active) scale.
[c] Total solvent-accessible volume in cubic angstroms using a probe with a 1.4 Å radius.
[d] Estimated number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer.
[e] Estimated number of hydrogen bonds that would be accepted by the solute from water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer.
[f] Predicted brain/blood partition coefficient.
[g] Van der Waals surface area of polar nitrogen and oxygen atoms in square angstroms.

Table 10
Physiochemical descriptors of top 5 hits docked to COMT-108M obtained from pharmacophore analysis.

Title	No. of rotatable bonds[a]	CNS[b]	Molecular weight	Dipole moment	Volume[c]	HB donors[d]	HB acceptors[e]	QPlogBB[f]	PSA[g]
ZINC32235251	5	1	397.45	7.48	1254.17	1	6.5	0.23	59.12
ZINC23740255	6	1	393.48	2.62	1292.27	1	6.5	−0.09	61.80
ZINC26926448	8	1	330.44	6.06	1119.87	2	6.2	−0.32	61.55
ZINC24804757	7	0	372.50	7.94	1222.95	1	6	−0.48	61.88
ZINC24148834	7	1	396.48	4.45	1300.32	1	7.7	−0.02	68.71

[a] Number of non-trivial (not CX3), non-hindered (not alkene, amide, small ring) rotatable bonds.
[b] Predicted central nervous system activity on a −2 (inactive) to +2 (active) scale.
[c] Total solvent-accessible volume in cubic angstroms using a probe with a 1.4 Å radius.
[d] Estimated number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer.
[e] Estimated number of hydrogen bonds that would be accepted by the solute from water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer.
[f] Predicted brain/blood partition coefficient.
[g] Van der Waals surface area of polar nitrogen and oxygen atoms in square angstroms.

for a potential CNS-permeable lead molecule. Thus, a set of rationally designed compounds can be provided as potential inhibitors for assaying against the target protein.

4. Conclusion

Several lead molecules for COMT have been developed in recent years using both experimental and the classical evaluation of structure–activity relationships. An emerging concern, however, is side effects of the current drugs, namely tolcapone and entacapone that causes severe dopaminergic, gastro-intestinal and other adverse reactions. Experimental studies on COMT inhibition assays reported the design of scaffolds based on either catechol/SAM or both. In this study, we aim to generate new core structures which can be taken up as leads for the design of COMT inhibitors.

Structure-based virtual screening and pharmacophore modeling has emerged as complementary methods to high throughput screening of large chemical databases. However, these methods are individually far from perfect in many aspects, including a low hit rate and a low enrichment factor, as well as a high false positive rate. This work has taken a significant step toward the integration of ligand-based and structure-based methodologies for the design of novel COMT inhibitors to compensate for these limitations.

In the current study, both structure-based and ligand-based approaches were utilized to screen a CNS permeable library from ZINC database. The search for novel hits was carried out against both the polymorphic forms (108V and 108M) of the target protein.

A structure-based virtual screening approach has facilitated the identification of compounds that effectively inhibit COMT activity. ZINC database consisting of 21 million molecules was filtered for CNS permeability and screened for top hits using three docking stages with increasing stringency.

The ligand-based approach involved generation of five-point energy-based pharmacophore using energetic terms from docking of active ligands with human S-COMT. Further, this pharmacophore was used to screen the CNS permeable library. The screened molecules from the ZINC CNS-permeable library were then docked to crystal structure of both COMT-108V and 108M. The active compounds used for the generation of pharmacophore were mainly bisubstrate inhibitors with the goal of minimizing the side-effects of the clinically used mono-substrate inhibitors. The bisubstrate type of inhibitors will bind more specifically, employing both the SAM- and catechol-binding sites, thus prove to be better binders than single-site inhibitors.

The hits from both the approaches were ranked according to binding energy estimates. Also, these small-molecule inhibitors showed favorable ADME properties for CNS permeability and activity. The binding studies of the ligand–receptor complexes also indicated the key amino acids present in the active site of COMT that are crucial for ligand binding in both the forms. The two different approaches used here generated different set of hits for both the polymorphic forms of COMT. These diverse backbones will provide frameworks for the development of future inhibitors through exploitation of sub-sites both within and outside the main binding cleft.

The top hits that are synthesizable and selective for individual COMTs, in both their polymorphic forms, will be important as therapeutic agents and have good potential for a future hit-to-lead optimization. It has been reported that the Val(108/158)Met polymorphism affects enzymatic activity and protein levels in a similar way in both S-COMT and MB-COMT [79,80], so our conclusions concerning S-COMT apply to MB-COMT that is responsible for the metabolism of neurotransmitters and risk for several neurological dysfunctions.

Acknowledgements

This work is supported by funding from Department of Biotechnology, Ministry of Science and Technology, Government of India. NJ is thankful to University Grants Commission for providing fellowship.

References

[1] J. Axelrod, Methylation reactions in the formation and metabolism of catecholamines and other biogenic amines, Pharmacological Reviews 18 (1966) 95–113.
[2] M.J. Bonifácio, P.N. Palma, L. Almeida, P. Soares-da-Silva, Catechol-O-methyltransferase and its inhibitors in Parkinson's disease, CNS Drug Reviews 13 (2007) 352–379.
[3] T. Tom, J.L. Cummings, Depression in Parkinson's disease. Pharmacological characteristics and treatment, Drugs and Aging 12 (1998) 55–74.
[4] S. Shifman, M. Bronstein, M. Sternfeld, A. Pisanté, A. Weizman, I. Reznik, B. Spivak, N. Grisaru, L. Karp, R. Schiffer, M. Kotler, R.D. Strous, M. Swartz-Vanetik, H.Y. Knobler, E. Shinar, B. Yakir, N.B. Zak, A. Darvasi, COMT: a common susceptibility gene in bipolar disorder and schizophrenia, American Journal of Medical Genetics Part B: Neuropsychiatric Genetics 128B (2004) 61–64.
[5] D.K. Cheuk, V. Wong, Meta-analysis of association between a catechol-O-methyltransferase gene polymorphism and attention deficit hyperactivity disorder, Behavior Genetics 36 (2006) 651–659.
[6] K.E. Lewandowski, Relationship of catechol-O-methyltransferase to schizophrenia and its correlates: evidence for associations and complex interactions, Harvard Review of Psychiatry 15 (2007) 233–244.
[7] M.C. Houston, The role of mercury and cadmium heavy metals in vascular disease, hypertension, coronary heart disease, and myocardial infarction, Alternative Therapies in Health and Medicine 13 (2007) S128–S133.
[8] P.T. Männistö, I. Ulmanen, K. Lundström, J. Taskinen, J. Tenhunen, C. Tilgmann, S. Kaakkola, Characteristics of catechol-O-methyltransferase (COMT) and properties of selective COMT inhibitors, Progress in Drug Research 39 (1992) 291–350.
[9] E. Nissinen, R.K. Tuominen, V. Perhoniemi, S. Kaakkola, Catechol-O-methyltransferase activity in human and rat small intestine, Life Sciences 42 (1988) 2609–2614.
[10] M. Assicot, C. Bohuon, Presence of two distinct catechol-O-methyltransferase activities in red blood cells, Biochimie 53 (1971) 871–874.

[11] J. Tenhunen, M. Salminen, K. Lundström, T. Kiviluoto, R. Savolainen, I. Ulmanen, Genomic organization of the human catechol-O-methyltransferase gene and its expression from two distinct promoters, European Journal of Biochemistry 223 (1994) 1049–1059.
[12] M. Salminen, K. Lundström, C. Tilgmann, R. Savolainen, N. Kalkkinen, I. Ulmanen, Molecular cloning and characterization of rat liver catechol-O-methyltransferase, Gene 93 (1990) 241–247.
[13] T. Lotta, J. Vidgren, C. Tilgmann, I. Ulmanen, K. Melén, I. Julkunen, J. Taskinen, Kinetics of human soluble and membrane-bound catechol O-methyltransferase: a revised mechanism and description of the thermolabile variant of the enzyme, Biochemistry 34 (1995) 4202–4210.
[14] M.A. Palmatier, A.M. Kang, K.K. Kidd, Global variation in the frequencies of functionally different catechol-O-methyltransferase alleles, Biological Psychiatry 46 (1999) 557–567.
[15] H. Kunugi, S. Nanko, A. Ueki, E. Otsuka, M. Hattori, F. Hoda, H.P. Vallada, M.J. Arranz, D.A. Collier, High and low activity alleles of catechol O-methyltransferase gene: ethnic difference and possible association with Parkinson's disease, Neuroscience Letters 221 (1997) 202–204.
[16] M.F. Martínez, X.E. Martín, L.G. Alcelay, J.C. Flores, J.M. Valiente, B.I. Juanbeltz, M.A. Beldarraín, J.M. López, M.C. Gonzalez-Fernández, A.M. Salazar, R.B. Gandarias, S.I. Borda, N.O. Marqués, M.B. Amillano, M.C. Zabaleta, M.M. de Pancorbo, The COMT Val158 Met polymorphism as an associated risk factor for Alzheimer disease and mild cognitive impairment in APOE 4 carriers, BMC Neuroscience 10 (2009) 125.
[17] S. Wedrén, T.R. Rudqvist, F. Granath, E. Weiderpass, M. Ingelman-Sundberg, Catechol-O-methyltransferase gene polymorphism and post-menopausal breast cancer risk, Carcinogenesis 24 (2003) 681–687.
[18] E.C. Pooley, N. Fineberg, P.J. Harrison, The met (158) allele of catechol-O-methyltransferase (COMT) is associated with obsessive–compulsive disorder in men: case–control study and metaanalysis, Molecular Psychiatry 12 (2007) 556–561.
[19] T. Wang, P. Franke, H. Neidt, S. Cichon, M. Knapp, D. Lichtermann, W. Maier, P. Propping, M.M. Nöthen, Association study of the low-activity allele of catechol-O-methyltransferase and alcoholism using a family-based approach, Molecular Psychiatry 6 (2001) 109–111.
[20] R.D. Strous, K.A. Nolan, R. Lapidus, L. Diaz, T. Saito, H.M. Lachman, Aggressive behavior in schizophrenia is associated with the low enzyme activity COMT polymorphism: a replication study, American Journal of Medical Genetics Part B: Neuropsychiatric Genetics 120B (2003) 29–34.
[21] P. Karling, Å. Danielsson, M. Wikgren, I. Söderström, J. Del-Favero, R. Adolfsson, K.F. Norrback, The relationship between the Val158Met catechol-o-methyltransferase (COMT) polymorphism and irritable bowel syndrome, PLoS One 6 (2011) e18035.
[22] J. Vidgren, L.A. Svensson, A. Liljas, Crystal structure of catechol O-methyltransferase, Nature 368 (1994) 354–358.
[23] K. Rutherford, I.Le. Trong, R.E. Stenkamp, W.W. Parson, Crystal structures of human 108V and 108M catechol O-methyltransferase, Journal of Molecular Biology 380 (2008) 120–130.
[24] J. Axelrod, R. Tomchick, Enzymatic O-methylation of epinephrine and other catechols, Journal of Molecular Biology 233 (1958) 702–705.
[25] H.C. Guldberg, C.A. Marsden, Catechol-O-methyl transferase: pharmacological aspects and physiological role, Pharmacological Reviews 27 (1975) 135–206.
[26] R. Bäckström, E. Honkanen, A. Pippuri, P. Kairisalo, J. Pystynen, K. Heinola, E. Nissinen, I.B. Linden, P.T. Männistö, S. Kaakkola, P. Pohto, Synthesis of some novel potent and selective catechol-O-methyltransferase inhibitors, Journal of Medicinal Chemistry 32 (1989) 841–846.
[27] J. Borgulya, H. Bruderer, K. Bernauer, G. Zürcher, M. Da Prada, Catechol-O-methyltransferase-inhibiting pyrocatechol derivatives: synthesis and structure–activity studies, Helvetica Chimica Acta 72 (1989) 952–968.
[28] C. Marin, J.A. Obeso, Catechol-O-methyltransferase inhibitors in preclinical models as adjuncts of L-dopa treatment, International Review of Neurobiology 95 (2010) 191–205.
[29] S.E. Brevitt, E.W. Tan, Synthesis and *in vitro* evaluation of two progressive series of bifunctional polyhydroxybenzamide catechol-O-methyltransferase inhibitors, Journal of Medicinal Chemistry 40 (1997) 2035–2039.
[30] K. Bailey, E.W. Tan, Synthesis and evaluation of bifunctional nitrocatechol inhibitors of pig liver catechol-O-methyltransferase, Bioorganic and Medicinal Chemistry 13 (2005) 5740–5749.
[31] G.L. Anderson, D.L. Bussolotti, J.K. Coward, Synthesis and evaluation of some stable multisubstrate adducts as inhibitors of catechol-O-methyltransferase, Journal of Medicinal Chemistry 24 (1981) 1271–1277.
[32] K. Haasio, Toxicology and safety of COMT inhibitors, International Review of Neurobiology 95 (2010) 163–189.
[33] F.C. Bernstein, T.F. Koetzle, G.J. Williams, E.F. Meyer, M.D. Brice Jr., J.R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, The protein data bank: a computer-based archival file for macromolecular structures, Journal of Molecular Biology 112 (1977) 535–542.
[34] A.T.R. Laurie, R.M. Jackson, Q-SiteFinder: an energy-based method for the prediction of protein–ligand binding sites, Bioinformatics 21 (2005) 1908–1916.
[35] J. Dundas, Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz, J. Liang, CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues, Nucleic Acids Research 34 (2006) W116–W118.
[36] A.C. Wallace, R.A. Laskowski, J.M. Thornton, LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions, Protein Engineering 8 (1995) 127–134.
[37] Maestro, Version 9.2, Schrödinger, LLC, New York, NY, 2011.
[38] J.R. Crout, Inhibition of catechol-O-methyl transferase by pyrogallol in the rat, Biochemical Pharmacology 6 (1961) 47–50.
[39] G. Tunnicliff, T.T. Ngo, Kinetics of rat brain soluble catechol-O-methyltransferase and its inhibition by substrate analogues, International Journal of Biochemistry 15 (1983) 733–738.
[40] C.E. Garner, L.T. Burka, A.E. Etheridge, H.B. Matthews, Catechol metabolites of polychlorinated biphenyls inhibit the catechol-O-methyltransferase-mediated metabolism of catechol estrogens, Toxicology and Applied Pharmacology 162 (2000) 115–123.
[41] G. Zürcher, H.H. Keller, R. Kettler, J. Borgulya, E.P. Bonetti, R. Eigenmann, M. Da Prada, Ro 40-7592, a novel, very potent, and orally active inhibitor of catechol-O-methyltransferase: a pharmacological study in rats, Advances in Neurology 53 (1990) 497–503.
[42] E. Nissinen, I.B. Lindén, E. Schultz, P. Pohto, Biochemical and pharmacological properties of a peripherally acting catechol-O-methyltransferase inhibitor entacapone, Naunyn-Schmiedebergs Archives of Pharmacology 346 (1992) 262–266.
[43] T. Deguchi, J. Barchas, Inhibition of transmethylations of biogenic amines by S-adenosylhomocysteine. Enhancement of transmethylation by adenosylhomocysteinase, Journal of Biological Chemistry 246 (1971) 3175–3181.
[44] R.T. Borchardt, Y.S. Wu, Potential inhibitors of S-adenosylmethionine-dependent methyltransferases. 3. Modifications of the sugar portion of S-adenosylhomocysteine, Journal of Medicinal Chemistry 18 (1975) 300–304.
[45] R.T. Borchardt, Y.S. Wu, Potential inhibitors of S-adenosylmethionine-dependent methyltransferases. 1. Modification of the amino acid portion of S-adenosylhomocysteine, Journal of Medicinal Chemistry 17 (1974) 862–868.
[46] R.T. Borchardt, J.A. Huber, Y.S. Wu, Potential inhibitor of S-adenosylmethionine-dependent methyltransferases. 2. Modification of the base portion of S-adenosylhomocysteine, Journal of Medicinal Chemistry 17 (1974) 868–873.
[47] H. Lu, X. Meng, C.S. Yang, Enzymology of methylation of tea catechins and inhibition of catechol-O-methyltransferase by (-)-epigallocatechin gallate, Drug Metabolism and Disposition: The Biological Fate of Chemicals 31 (2003) 572–579.
[48] B. Zhu, P. Wang, M. Nagai, Y. Wen, H.W. Bai, Inhibition of human catechol-O-methyltransferase (COMT)-mediated O-methylation of catechol estrogens by major polyphenolic components present in coffee, Journal of Steroid Biochemistry and Molecular Biology 113 (2008) 65–74.
[49] M. Kadowaki, E. Ootani, N. Sugihara, K. Furuno, Inhibitory effects of catechin gallates on o-methyltranslation of protocatechuic acid in rat liver cytosolic preparations and cultured hepatocytes, Biological and Pharmaceutical Bulletin 28 (2005) 1509–1513.
[50] D. Yalcin, O. Bayraktar, Inhibition of catechol-O-methyltransferase (COMT) by some plant-derived alkaloids and phenolics, Journal of Molecular Catalysis B: Enzymatic 64 (2010) 162–166.
[51] J.V. Burba, M.F. Murnaghan, Catechol-O-methyl transferase inhibition and potentiation of epinephrine responses by desmethylpapavernine, Biochemical Pharmacology 14 (1965) 823–829.
[52] J.V. Burba, G.C. Becking, Effect of the antioxidant nordihydroguaiaretic acid on the *in vitro* activity of catechol-O-methyl transferase, Archives Internationales de Pharmacodynamie et de Therapie 180 (1969) 323–330.
[53] B. Masjost, P. Ballmer, E. Borroni, G. Zürcher, F.K. Winkler, R. Jakob-Roetne, F. Diederich, Structure-based design, synthesis, and *in vitro* evaluation of bisubstrate inhibitors for catechol-O-methyltransferase (COMT), Chemistry 6 (2000) 971–982.
[54] C. Lerner, R. Siegrist, E. Schweizer, F. Diederich, V. Gramlich, R. Jakob-Roetne, G. Zürcher, E. Borroni, Bisubstrate inhibitors for the enzyme catechol O-methyltransferase (COMT): dramatic effects of ribose modifications on binding affinity and binding mode, Helvetica Chimica Acta 86 (2003) 1045–1062.
[55] R. Paulini, C. Trindler, C. Lerner, L. Brändli, W.B. Schweizer, R. Jakob-Roetne, G. Zürcher, E. Borroni, F. Diederich, Bisubstrate inhibitors of catechol O-methyltransferase (COMT): the crucial role of the ribose structural unit for inhibitor binding affinity, ChemMedChem 1 (2006) 340–357.
[56] B. Belleau, J. Burba, Tropolones: a unique class of potent non-competitive inhibitors of S-adenosylmethionine-catechol methyltransferase by tropolones, Biophysica Acta 54 (1961) 195–196.
[57] J.D. Lambert, D. Chen, C.Y. Wang, N. Ai, S. Sang, C.T. Ho, W.J. Welsh, C.S. Yang, Benzotropolone inhibitors of estradiol methylation: kinetics and *in silico* modeling studies, Bioorganic and Medicinal Chemistry 13 (2005) 2501–2507.
[58] R. Bäckström, J. Pystynen, T. Lotta, M. Ovaska, J. Taskinen, Derivatives of naphthalene with COMT inhibiting activity, WO2002/22551 A1 (2002).
[59] J.J. Irwin, B.K. Shoichet, ZINC – a free database of commercially available compounds for virtual screening, Journal of Chemical Information and Modeling 45 (2005) 177–182.
[60] H. Pajouhesh, G.R. Lenz, Medicinal chemical properties of successful central nervous system drugs, NeuroRx 2 (2005) 541–553.
[61] LigPrep, version 2.4, Schrödinger, LLC, New York, NY, 2010.
[62] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, Journal of the American Chemical Society 118 (1996) 11225–11236.
[63] Schrödinger Suite 2011 Virtual Screening Workflow; Glide version 5.7, Schrödinger, LLC, New York, NY, 2011; LigPrep version 2.5, Schrödinger, LLC, New York, NY, 2011; QikProp version 3.4, Schrödinger, LLC, New York, NY, 2011.
[64] R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin,

Glide A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, Journal of Medicinal Chemistry 47 (2004) 1739–1749.
[65] T.A. Halgren, R.B. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard, J.L. Banks, Glide A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening, Journal of Medicinal Chemistry 47 (2004) 1750–1759.
[66] G. Rastelli, A. Del Rio, G. Degliesposti, M. Sgobba, Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA, Journal of Computational Chemistry 31 (2010) 797–810.
[67] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes, Journal of Medicinal Chemistry 49 (2006) 6177–6196.
[68] J. Li, R. Abel, K. Zhu, Y. Cao, S. Zhao, R.A. Friesner, The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling, Proteins 79 (2011) 2794–2812.
[69] N.K. Salam, R. Nuti, W. Sherman, Novel method for generating structure-based pharmacophores using energetic analysis, Journal of Chemical Information and Modeling 49 (2009) 2356–2368.
[70] H. van de Waterbeemd, E. Gifford, ADMET in silico modelling: towards prediction paradise? Nature Reviews Drug Discovery 2 (2003) 192–204.
[71] G. Cruciani, F. Milletti, L. Storchi, G. Sforna, L. Goracci, In silico p*K*a prediction and ADME profiling, Chemistry & Biodiversity 6 (2009) 1812–1821.
[72] F. Darvas, G. Keseru, A. Papp, G. Dormán, L. Urge, P. Krajcsi, In Silico, Ex silico ADME approaches for drug discovery, Current Topics in Medicinal Chemistry 2 (2002) 1287–1304.
[73] E.M. Duffy, W.L. Jorgensen, Prediction of properties from simulations: free energies of solvation in hexadecane, octanol, and water, Journal of the American Chemical Society 122 (2000) 2878–2888.
[74] X. Cheng, R.J. Roberts, AdoMet-dependent methylation, DNA methyltransferases and base flipping, Nucleic Acids Research 29 (2001) 3784–3795.
[75] J.L. Martin, F.M. McMillan, SAM. (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold, Current Opinion in Structural Biology 12 (2002) 783–793.
[76] K. Rutherford, B.J. Bennion, W.W. Parson, V. Daggett, The 108M polymorph of human catechol O-methyltransferase is prone to deformation at physiological temperatures, Biochemistry 45 (2006) 2178–2188.
[77] K. Rutherford, E. Alphandéry, A. McMillan, V. Daggett, W.W. Parson, The V108M mutation decreases the structural stability of catechol O-methyltransferase, Biochimica et Biophysica Acta 1784 (2008) 1098–1105.
[78] R. Paulini, C. Lerner, R. Jakob-Roetne, G. Zürcher, E. Borroni, F. Diederich, Bisubstrate inhibitors of the enzyme catechol O-methyltransferase (COMT): efficient inhibition despite the lack of a nitro group, ChemBioChem 5 (2004) 1270–1274.
[79] J. Chen, B.K. Lipska, N. Halim, Q.D. Ma, M. Matsumoto, S. Melhem, B.S. Kolachana, T.M. Hyde, M.M. Herman, J.A. Apud, M.F. Egan, J.E. Kleinman, D.R. Weinberger, Functional analysis of genetic variation in catechol-O-methyltransferase (COMT): effects on mRNA, protein, and enzyme activity in postmortem human brain, American Journal of Human Genetics 75 (2004) 807–821.
[80] A.E. Doyle, J.D. Yager, Catechol-O-methyltransferase: effects of the val108met polymorphism on protein turnover in human cells, Biochimica et Biophysica Acta 1780 (2008) 27–33.