In-silico Drug Repurposing Pipeline for Epilepsy: Integrating Deep Learning and Structure-based Approaches

Xiaoying Lv[1], Jia Wang[2], Ying Yuan[1], Lurong Pan[3], Jinjiang Guo[1*]

[1] Global Health Drug Discovery Institute, Beijing, China

[2] Cipher Gene Limited, Beijing, China

[3] Ainnocence Limited, HongKong, China

* corresponding author

Email: jinjiang.guo@ghddi.org

Abstract

Due to considerable global prevalence and high recurrence rate, the pursuit of effective new medication for epilepsy treatment remains an urgent and significant challenge. Drug repurposing emerges as a cost-effective and efficient strategy to combat this disorder. This study leverages the transformer-based deep learning methods coupled with molecular binding affinity calculation to develop a novel *in-silico* drug repurposing pipeline for epilepsy. The number of candidate inhibitors against 24 target proteins encoded by gain-of-function (GOF) genes implicated in epileptogenesis ranged from zero to several hundreds. Our pipeline has repurposed the medications with most anti-epileptic drugs (AEDs) and nearly half psychiatric medications, highlighting the effectiveness of our pipeline. Furthermore, Lomitapide, a cholesterol-lowering drug, first emerged as particularly noteworthy, exhibiting high binding affinity for 10 targets and verified by molecular dynamics (MD) simulation and mechanism analysis. These findings provided a novel perspective on therapeutic strategies for other central nervous system (CNS) disease.

Keywords: drug repurposing, epilepsy, gain-of-function gene, docking, deep learning, BBB (blood-brain barrier) permeability

Introduction

Epilepsy, a central neurological disease, can be caused by any direct or indirect insult to the CNS, and often accompanied by cognitive and psychological impact due to its recurrent nature[1]. A meta-analysis study revealed that worldwide lifetime prevalence of epilepsy was approximately 1,099 per 100,000 individuals worldwide[2]. Recent treatment strategies for epilepsy primarily focused on the relief of symptoms, and advancements in treatment have enabled most patients to achieve seizure control[3]. Beside AEDs, alternative treatments such as dietary therapies[4], surgery[5], neuromodulation therapies[6] can also aid in seizures control. Despite therapeutic advancements, reducing high recurrence rate for intractable epilepsy still poses a significant challenge for approximately 30% of patients[7], even with monotherapy or polytherapy AED regimens[8,9]. Thus, the discovery of new AEDs would serve for not only improving seizure management but also for diminishing the substantial societal and healthcare burdens associated with these neurological conditions[10].

Drug repurposing, the practice of finding new therapeutic usages for FDA-approved drugs, provides a cost-effective and time-efficient approach in drug discovery. Meanwhile, the integration of structure-based and deep learning methods has significantly enhanced drug repurposing, particularly in elucidating structure-activity relationships and accelerating drug discovery[11,12]. The Experimental Epilepsies (PDE3) database, developed by Nasir group through literature mining, has cataloged drugs that have demonstrated efficacy in experimental epilepsy models[13]. Reutens's laboratory had attempted to repurpose drugs related to complement system for post-traumatic epilepsy, leveraging shared mechanism of action between these drugs and neuroinflammatory responses associated with epileptogenesis[14]. Moreover, potential epilepsy treatments have been identified by contrasting gene expression profiles in epileptic patients against healthy controls,

using drug perturbation data from the Connectivity Map (CMAP) database[15]. A successful example of drug repurposing is Igalmi, initially approved by FDA in 1999 for sedation and analgesia in intensive care settings. BioXcel biopharmaceuticals repurposed it for epilepsy treatment, capitalizing on its properties as a selective α2-adrenergic receptor modulator known to alleviate agitation in adults[16].

In this study, we developed an innovative *in-silico* drug repurposing pipeline by integrating structure-based and deep learning approaches. The pipeline is specifically designed to identify new potential inhibitors for epilepsy (Fig. 1). We selected 24 targets encoded by GOF genes for drug repurposing, restricted by the unavailability of crystal structures for other targets[17,18,19]. The identified potential inhibitors displayed a high BBB permeable probability to effectively reach brain tissue. We prioritized candidates that exhibited low cytotoxicity, especially those intended for chronic management. Consequently, our pipeline selected a range of drugs which may demonstrated anti-seizure activities against specific targets. Moreover, for Lomitapide repurposed by our pipeline, 100-ns MD simulations against each selected target and mechanism analysis were conducted to further validate its potential inhibition for epilepsy.

Results

2.1 Candidates number reduction by using individualized reference compounds

In the preliminary stage of our pipeline development, Glide standard precision (SP) docking was utilized to screen a set of 11,008 compounds against 24 target proteins encoded by GOF genes. When applying a uniformly established docking score threshold of -7.0, a highly disproportionate distribution of compounds meeting the criterion was observed across different targets (Fig. S1). To rectify this discrepancy, individualized docking score cutoffs were proposed

for targets with a ligand-bound structure in the PDB database or known drug association. Notably, most reference compounds achieved docking scores below -7.0. Nevertheless, five reference compounds deviated from this trend, as detailed in Table 1. Remarkably, the *RHEB* encoding target, which has two distinct structures (PDB IDs: 6BT0, 6BSX), manifested docking scores for the non-GDP binding pocket of -6.08642 and -6.71108, respectively. When the individualized reference compounds were utilized as positive controls, a notable decrease in the pool of candidate compounds observed.

In the case of seven structures featuring two putative binding pockets, the quantity of compounds exhibiting strong binding affinity was generally higher in one pocket compared to the other. For example, when applying a cutoff of -7, the number of compounds with favorable binding affinity to the two pockets of protein (PDB ID: 4JTA) was 1637 for the first and 193 for the second (Fig. S1).

2.2 Compound BBB Permeability Evaluation

BBB permeability is paramount in determining the efficacy of drugs for neurological disorders by dictating the accessibility of compounds to CNS. To evaluate compound's BBB permeability, we adopted a self-attention-based Graph Neural Network (GNN) model, namely Ligandfomer[20], to predict its potential brain accessibility. For a given compound, the model generates a probability score within a range of 0 to 1, where a higher score indicates greater likelihood of brain accessibility. The Ligandformer model was rigorously trained and validated on the DB3B dataset[21]. The model's performance was assessed by various metrics, yielding a harmonic average score of 0.932 for the area under the receiver operating characteristic curve (AUROC) and a harmonic average score of 0.925 for the area under the precision-recall curve (AUPR).

In our analysis of the predicted probability scores for 11,008 small molecule drugs, the distribution across various BBB permeability intervals appeared relatively uniform, with each interval containing approximately 1,000 compounds, except in the score ranges below 0.1 and above 0.9 (Fig. S2). Notably, the BBB permeability predictions for AEDs, which are commonly available in the market, exceeded a probability of 0.8 (Table S1). The consistent pattern underscores the model's robust capability for predicting BBB permeability. To avoid omitting any potential drugs for epilepsy, we selected 3,200 candidate drugs with BBB permeable probability score above 0.75 for the subsequent stage.

2.3 The potential drugs screened against each target

In our strategic pipeline, there was a preferential selection for the orally administered and FDA-approved drugs for chronic diseases. We identified a multitude of prospective compounds for most targets. However, for eight target proteins encoded by—namely, *KCNQ2*, *ATP6V1A*, *FGF12*, *FGF13*, *HCN2*, *RHEB* and *SCN8A* genes—there were no suitable potential medications available, as detailed in Table 1. Despite the compatibility of numerous drugs with their target crystal structures in the docking analysis, none exhibited sufficient probabilities of BBB permeability, based on the Ligandformer predictions.

In the joint filtering analysis, the number of potential candidate medications range from zero to several hundreds, particularly exceeding one hundred for the target (pocket 1) encoded by *EMP2A* and *GAGRB2* genes (Fig 2). The lack of a reference compound for target encoded by *EMP2A* likely contributed to the substantial number of potential drugs. Nearly half of candidate drugs are indicated for psychiatric conditions, such as depression, schizophrenia, insomnia, bipolar, Parkinson's disease, dementia, anxiety and migraine, all of which display symptoms and mechanisms closely associated with the onset of seizures. Consequently, these drugs are deemed

promising for the prophylaxis of seizure incidents. Furthermore, our pipeline identified 27 antiepileptic drugs (AEDs), confirming its effectiveness in identifying suitable treatments for epilepsy. Additionally, a few antihypertensive and antidiabetic drugs were recognized for their potential efficacy in epilepsy management (Table S2).

In our investigation of voltage-gated sodium channels, we analyzed five structures (encoded by *SCN1-3A, SCN8-9A*). Despite the structures' high degree of similarity, indicated by a root mean square deviation (RMSD) of less than 2 Å (Fig. S3a), distinct binding sites were identified (Fig. S3b). This dissimilarity in binding sites could explain the observed inconsistency in the efficacy of candidate drugs targeting the SCN family (Table S2). For the *SCN1A* encoding target, Midazolam was recognized for its ability to generate π-π interactions with Phe1789 (Fig. 3a), exhibiting a notable docking score of -8.21737. Fluphenazine (Fig. 3b) and Lomitapide (Fig. 3c) emerged as the most promising drugs for target encoded by *SCN1A* gene, offering multiple protein-ligand interactions that enhance pose stability.

2.4 Potential medications targeting multiple proteins

In this study, several drugs were identified as candidates for repurposing with actions on multiple targets implicated in seizure control. Among the identified drugs, Nebivolol, Lomitapide, and Pimozide exhibited the most promise, engaging with 11, 10 and 9 targets respectively (Table 2).

Nebivolol was a selective β1-blocker with vasodilatory properties commonly prescribed. Extensive research endeavors have sought to unravel the intricacies of drug-drug interactions between nebivolol and prevalent adverse drug events using a murine model[22]. Additionally, a synergistic effect is observed when nebivolol is combined with a low dose of phenytoin, effectively mitigating seizures precipitated by maximal electroshock seizures (MES)[23]. Conversely, nebivolol

has been demonstrated to diminish the anticonvulsant efficacy of both carbamazepine and phenobarbital in MES models[24].

Lomitapide, utilized for cholesterol reduction in patients with homozygous familial hypercholesterolemia, has emerged as a candidate with potential efficacy for 10 distinct targets (Fig. 4a). Despite its demonstrated affinity for multiple GOF targets, Lomitapide lacks documented evidence in treating neurological disorders. The stability of the protein-ligand interaction in *CACNA1G* encoding target (PDB ID: 6KZP), is largely due to a π-π interaction with Phe917 and a hydrophobic interaction, as depicted in Figure 4b. Comparable interactions are seen with the target encoded by *MTOR* gene, where Lomitapide binds similarly through π-π and hydrophobic interactions (Fig. 4c). Moreover, the significant docking score of -10.6177 for the target encoded by *SCN3A* gene is primarily the result of hydrophobic interactions, highlighted in Figure 4d.

Pimozide, associated with nine targets as indicated in Table 2, is a potent dopamine antagonist. Studies have shown this medication to be correlated with a reduced risk of inducing seizures[25,26].

2.5 MD simulation for Lomitapide stability validation against 10 selected targets

To further substantiate the potential of Lomitapide in mitigating epilepsy progression, MD simulations were conducted for each target. Stability of the ligand-binding domains was observed in all 10 systems, as evidenced by convergence of backbone RMSD within the 100 ns production phase (Fig. 4e), with the exception of the proteins encoded by *CHRNA2* and *MTOR* genes. For the *CHRNA2* encoding target (PDB ID: 5FJV), a PDB resolution of 3.2 Å suggests that suboptimal side-chain packing may account for the observed instability during MD simulation. In the case of target encoded by *MTOR* gene (PDB ID: 8ERA), the elevated RMSD was primarily attributed to the relative movement of chains A and B, yet the RMSDs of these individual chains remained approximately 2 Å (Fig. S4a).

In most systems, Lomitapide maintained a pose similar to that of the initial docking conformation, with only minor deviations observed within the binding pocket during the 100 ns MD simulations, as depicted in Movies S1-S10. Stability of the ligand within the binding site was observed, with the notable exception of the ligand segments located in the solvent or cavity, such as those in 4RKK and 5FJV, which exhibited flexibility due to the lack of direct interactions with the receptor. In the sirolimus binding pocket, situated at the protein-protein interface of the *TSC* complex (PDB ID: 8ERA), the two chains engaged by sirolimus binding showed significant separation, potentially resulting from the weaker interactions of Lomitapide compared to the macrocyclic lactones sirolimus. Despite Lomitapide's high docking score, the 5 hydrogen bonds with Asp38, Gln54, Glu55, Ile57 and Tyr83 of protein domain (chain B of 8ERA) encoded by *FKBP1A* gene were not present (Fig. S4b). Additionally, the new pi-pi interaction with His88 in *FKBP1A* encoding protein proved insufficient to sustain the interaction between chain A and chain B (Fig. S4c).

2.6 Mechanism of action exploration of Lomitapide

To further explore Lomitapide's mechanism of action, we investigated its target's interacting genes within a protein-protein interaction (PPI) database and searched for shared pathways within AEDs in Gene Ontology (GO) database. The reported target of Lomitapide in FDA is microsomal triglyceride transfer protein, encoded by *MTTP* gene. In our curated PPI database[27], six genes (*APOA1*, *APOA2*, *APOA4*, *APOB*, *P4HB*, *HSP90B1*) demonstrated interactions with the *MTTP* gene. The majority are apolipoprotein-related gene, with the exception of the *HSP90B1* gene (Fig 5a), which has been previously associated with diseases such as Pelizaeus-Merzbacher-Like disease and bipolor disorder[28]. Notably, the former, Pelizaeus-Merzbacher-Like disease type 1, is

a hereditary central nervous system disorder that exhibits several symptoms reminiscent of epilepsy.

We conducted a search for pathways related to Lomitapide in the GO database, identifying 41 pertinent pathways (Table S3). Subsequently, we enriched the target genes of AEDs to analyze pathways using the DAVID website[29]. Our analysis identified four shared pathways: response to calcium ions, macromolecular complex binding, circadian rhythm regulation, and endoplasmic reticulum (Fig 5b). These findings suggest the potential existence of a common mechanism of action for Lomitapide and AEDs.

Discussion

Several prior studies on drug repurposing for epilepsy have primarily utilized approaches such as literature search[13], drug-target network[7] and gene expression profiling[15], as previous mentioned. In contrast, our proposed pipeline aims to repurpose drugs for seizure management by integrating structure-based and deep learning methods. Since FDA-approved drugs already conform to established clinical standards of safety and bioavailability, our focus is primarily on evaluating their on-target binding affinity and CNS bioavailability.

Our *in-silico* pipeline has effectively identified numerous potential therapeutics; nonetheless, it is subject to inherent constraints of the approaches employed in our study. The accuracy of docking scores depends on various factors, including, but not limited to, the resolution of the receptor, the conformation of the ligand, and the selected scoring function. On the other hand, the deep learning model for predicting BBB permeability demonstrated promising performance, however its predictive accuracy is still limited by the quality and diversity of training data and inherent

algorithmic constraints. Therefore, it is essential to validate these computational predictions with experimental edassays.

In conclusion, we have established an *in-silico* drug repurposing pipeline, leveraging structural compatibility and deep learning method to screen FDA-approved drugs for potential inhibitory effects on targets encoded by 24 GOF genes implicated in epilepsy. We observed a considerable variation in the number of potential repurposable drugs, which ranged from zero to several hundreds, contingent on the reference compounds used for each target. Notably, our analysis identified common AEDs as possible candidates, underscoring the effectiveness and efficiency of our pipeline in drug repurposing efforts. For the first time, Lomitapide emerged as a promising candidate medication, demonstrating stability across 10 targets in MD simulations and shared similar mechanism with AEDs in PPI interaction and GO enrichment analysis. This finding warrants further investigation into its therapeutic potential using zebrafish models and subsequent validation in human patients. While our approach offers insightful directions for epilepsy therapy, extensive research remains imperative to realize the ultimate objective of curing this complex disease.

Materials and Methods

4.1 GOF genes searched from three databases

Potential candidate genes implicated in epileptogenesis via GOF mechanisms were identified through database search in PubMed Central[30], Online Mendelian Inheritance in Man (OMIM)[31] and Development Disorder Genotype - Phenotype Database (DDG2P) using "epilepsy" and "gain of function" as keywords. Finally, 72 genes were selected as target genes for drug repurposing.

The reported evidence for GOF spans three levels-variant, gene, disease- detailed in Supplemental Table S4.

4.2 Crystal structures on target proteins and small molecules

Given that certain proteins function as complexes, a single crystal structure can suffice for multiple genes. Only 25 available crystal structures encoded by GOF genes with appropriate binding pockets were downloaded from the PDB database[32] for the repurposing of small molecules to enhance docking precision (Table 1). Specially for the *RHEB* encoding target, two distinct crystal structures were selected due to a significant conformational alteration at the Tyr74 amino acid, as illustrated in structures 6BT0 and 6BSX (see Fig. S5). From the DrugBank database (version *2022.1.17*)[33], we imported 11,008 small molecules for the purpose of docking with the targets.

4.3 Structure-based molecule docking

The docking workflow encompassed three steps: protein preparation, ligand preparation and the docking operation. Protein preparation utilized schrodinger Protein Preparation Workflow panel[34] with default parameters via Maestro interface (version *12.9.123*, released at 2021.3). This process entailed the designation of bond orders, addition of hydrogen atoms, retention of waters beyond 5Å distance with ligands, and the construction of missing side chains and loops. Subsequently, H-bond assignments were optimized based on PROPKA-predicted pKa values, followed by a restrained energy minimization using the default force field.

The putative binding pocket was defined either by referring the position of co-crystallized ligands or through prediction utilizing the SiteMap suite[35]. While most proteins possessed a single binding pocket, six proteins presented two putative binding pockets (Table 1). Protein grid files were generated by Receptor Grid Generation panel. After ligand preparation using command:

ligprep -ismi *.smi -omae *maegz, a total of 11,008 small molecules were successfully converted from 2D SMILES representations to 3D structures. These compounds were then docked into respective putative binding pockets employing the Glide SP docking protocol, which utilized the OPLS_2005 forcefield and default parameters[36,37].

4.4 BBB Evaluation and Filtering

We utilized Ligandformer, a deep learning model based on transformer architecture to estimate BBB permeability. Ligandformer, as a graph neural network that employs a multi-layer, single-head self-attention mechanism to predict ligand properties. This model leverages the self-attention mechanism inherent in deep learning architecture, allowing a detailed representation of the relationship between chemical structures and their attributes. This capability facilitates more informed interpretations, as demonstrated in Figure S6.

The predictive model was trained and validated with a dataset comprising experimentally determined BBB properties, as reported by Meng et al[21]. This dataset contained 7,806 data points, which included 4,955 positive samples (permeable) and 2,851 negative samples (non-permeable). To enhance the model's generalization capabilities, we employed a 5-fold cross-validation method. During each iteration of this process, the dataset was divided into an 80% training set and a 20% testing set, guaranteeing exposure to diverse data distributions.

4.5 Candidate drugs approved by FDA for chronic conditions

Any medications related to cancer treatment were excluded using keywords such as "malignant tumor", "cancer", "carcinoma" in their indication and their approved status was "true" in the DrugBank database.

4.6 System setup and MD simulations for 10 potential Lomitapide targets

We conducted MD simulations on the docking conformations produced by Lomitapide for 10 distinct targets. The initial structures were derived from the optimal docking conformations and maintained the same protonation state for the titratable residues. Given the substantial size of *GRIN1* encoding protein (PDB ID: 7EU7), MD simulations were confined to the transmembrane region alone (Fig. S7). We utilized the Tleap module in AMBER 22[38] to systematically delete and subsequently reintroduce hydrogen atoms. For the three solvent proteins (PDB IDs: 4RKK, 5FJV, 8ERA) systems, we built a TIP3P water box along XYZ axis, extending 12 Å and neutralized the entire systems with 0.15 M NaCl concentration. For the rest seven membrane protein systems, we additionally oriented the direction of transmembrane region along Z-axis using PPM 2.0[39] and embedded the proteins into 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) bilayer using CHARMM-GUI membrane builder[40].

All-atom MD simulations were conducted employing PMEMD engine in AMBER 22[38]. The AMBER FF19SB force field[41] and AMBER lipid force field LIPID21[42] was adopted for protein and POPC molecules, respectively. Lomitapide was parameterized using general AMBER force field (GAFF). A cutoff of 10 Å was implemented for nonbonded interactions and bond-length constrain involving hydrogen atoms were managed by the SHAKE algorithm[43]. The particle mesh Ewald (PME) algorithm[44] was used to address long-range electrostatic interactions. Simulations were conducted under periodic boundary conditions in all three dimensions. Initially, energy minimization was performed for 10,000 steps, followed by a gradual heating of the system from 0 K to 310 K across a span of 500 ps with the aid of the Langevin thermostat[45]. During heating, positional restraints were applied to the heavy atoms of the proteins and the ligand with a force constant of 50 kcal/mol/Å^2. This was succeeded by 2 ns of pre-equilibration in the isochoric-isothermal (NVT) ensemble with the imposition of decreasing positional restraint force from 50,

5 to 0.5 kcal/mol/Å^2. For seven membrane protein systems, the lipid heads were also subjected to positional restrained and a further 30 ns equilibration phase was administered with the proteins and ligand still under 0.5 kcal/mol/Å^2 constraint. The production phase spanned 100 ns for each system in a constant pressure (NPT) ensemble. The time step was set to 2 fs. The frames were saved every 5,000 steps for trajectory analysis.

4.7 PPI interaction and Gene Enrichment analysis

The target gene *MTTP*, associated with Lomitapide, was searched for interacting genes in the PPI database established in our previous study[27]. We investigated pathways related to Lomitapide within the GO database[46,47](Table S3). Furthermore, target genes of AEDs were retrieved from the DrugBank database and analyzed for GO pathway enrichment using the DAVID platform [48,29], examining three GO categories (biological process, cellular component and molecular function) using default settings (Table S3).

Reference

1. Beghi, E. The Epidemiology of Epilepsy. *Neuroepidemiology* **54**, 185–191 (2020).
2. Fiest, K. M. *et al.* Prevalence and incidence of epilepsy: A systematic review and meta-analysis of international studies. *Neurology* **88**, 296–303 (2017).
3. Goldenberg, M. M. Overview of drugs used for epilepsy and seizures: etiology, diagnosis, and treatment. *P T Peer-Rev. J. Formul. Manag.* **35**, 392–415 (2010).
4. Dyńka, D., Kowalcze, K. & Paziewska, A. The Role of Ketogenic Diet in the Treatment of Neurological Diseases. *Nutrients* **14**, 5003 (2022).
5. Rugg-Gunn, F., Miserocchi, A. & McEvoy, A. Epilepsy surgery. *Pract. Neurol.* **20**, 4–14 (2020).
6. Ryvlin, P., Rheims, S., Hirsch, L. J., Sokolov, A. & Jehi, L. Neuromodulation in epilepsy: state-of-the-art approved therapies. *Lancet Neurol.* **20**, 1038–1047 (2021).
7. Ko, Y., Lee, C., Lee, Y. & Lee, J.-S. Systematic Approach for Drug Repositioning of Anti-Epileptic Drugs. *Diagn. Basel Switz.* **9**, 208 (2019).
8. Nevitt, S. J., Sudell, M., Cividini, S., Marson, A. G. & Tudur Smith, C. Antiepileptic drug monotherapy for epilepsy: a network meta-analysis of individual participant data. *Cochrane Database Syst. Rev.* **4**, CD011412 (2022).
9. Kaplan, Y. C. & Demir, O. Use of Phenytoin, Phenobarbital Carbamazepine, Levetiracetam Lamotrigine and Valproate in Pregnancy and Breastfeeding: Risk of Major Malformations, Dose-dependency, Monotherapy vs Polytherapy, Pharmacokinetics and Clinical Implications. *Curr. Neuropharmacol.* **19**, 1805–1824 (2021).

10. Chen, Z., Brodie, M. J. & Kwan, P. What has been the impact of new drug treatments on epilepsy? *Curr. Opin. Neurol.* **33**, 185–190 (2020).
11. Pinzi, L. & Rastelli, G. Molecular Docking: Shifting Paradigms in Drug Discovery. *Int. J. Mol. Sci.* **20**, 4331 (2019).
12. Vemula, D., Jayasurya, P., Sushmitha, V., Kumar, Y. N. & Bhandari, V. CADD, AI and ML in drug discovery: A comprehensive review. *Eur. J. Pharm. Sci. Off. J. Eur. Fed. Pharm. Sci.* **181**, 106324 (2023).
13. Sivapalarajah, S. *et al.* The prescribable drugs with efficacy in experimental epilepsies (PDE3) database for drug repurposing research in epilepsy. *Epilepsia* **59**, 492–501 (2018).
14. Chen, M., Edwards, S. R. & Reutens, D. C. Complement in the Development of Post-Traumatic Epilepsy: Prospects for Drug Repurposing. *J. Neurotrauma* **37**, 692–705 (2020).
15. Brueggeman, L. *et al.* Drug repositioning in epilepsy reveals novel antiseizure candidates. *Ann. Clin. Transl. Neurol.* **6**, 295–309 (2019).
16. Smith, C. M., Santalucia, M., Bunn, H. & Muzyk, A. Sublingual Dexmedetomidine for the Treatment of Agitation in Patients with Schizophrenia and Bipolar Disorder. *Clin. Psychopharmacol. Neurosci. Off. Sci. J. Korean Coll. Neuropsychopharmacol.* **21**, 215–221 (2023).
17. McGinn, R. J., Von Stein, E. L., Summers Stromberg, J. E. & Li, Y. Precision medicine in epilepsy. *Prog. Mol. Biol. Transl. Sci.* **190**, 147–188 (2022).
18. Symonds, J. D., Zuberi, S. M. & Johnson, M. R. Advances in epilepsy gene discovery and implications for epilepsy diagnosis and treatment. *Curr. Opin. Neurol.* **30**, 193–199 (2017).
19. Meldrum, B. S. & Rogawski, M. A. Molecular targets for antiepileptic drug development. *Neurother. J. Am. Soc. Exp. Neurother.* **4**, 18–61 (2007).

20. Guo, J., Liu, Q., Guo, H. & Lu, X. Ligandformer: A Graph Neural Network for Predicting Compound Property with Robust Interpretation. Preprint at http://arxiv.org/abs/2202.10873 (2022).
21. Meng, F., Xi, Y., Huang, J. & Ayers, P. W. A curated diverse molecular database of blood-brain barrier permeability with chemical descriptors. *Sci. Data* **8**, 289 (2021).
22. Goel, R. & Goel, A. Nebivolol enhances the anticonvulsant effect of gabapentin against various animal models of epilepsy in mice. *J. Pharm. Res.* **6**, 364–367 (2013).
23. Dhamodharan, A., Selvaraj, N. & Meenakshi, R. Anticonvulsant Effect of Nebivolol alone and in Combination with Phenytoin against Maximal Electroshock-Induced Seizures in Mice. *J. Clin. Diagn. Res.* (2021) doi:10.7860/JCDR/2021/47167.14669.
24. Borowicz-Reutt, K. K., Banach, M. & Rudkowska, M. Nebivolol attenuates the anticonvulsant action of carbamazepine and phenobarbital against the maximal electroshock-induced seizures in mice. *Pharmacol. Rep. PR* **72**, 80–86 (2020).
25. Hedges, D., Jeppson, K. & Whitehead, P. Antipsychotic medication and seizures: a review. *Drugs Today Barc. Spain 1998* **39**, 551–557 (2003).
26. Larkin, C. Epileptogenic effect of pimozide. *Am. J. Psychiatry* **140**, 372–373 (1983).
27. He, S. *et al.* Drug Repositioning for Amyloid Transthyretin Amyloidosis by Interactome Network Corrected by Graph Neural Networks and Transcriptome Analysis. *Hum. Gene Ther.* **35**, 70–79 (2024).
28. Kakiuchi, C. *et al.* Association analysis of HSP90B1 with bipolar disorder. *J. Hum. Genet.* **52**, 794–803 (2007).
29. Sherman, B. T. *et al.* DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* **50**, W216–W221 (2022).

30. White, J. PubMed 2.0. *Med. Ref. Serv. Q.* **39**, 382–387 (2020).

31. Amberger, J. S. & Hamosh, A. Searching Online Mendelian Inheritance in Man (OMIM): A Knowledgebase of Human Genes and Genetic Phenotypes. *Curr. Protoc. Bioinforma.* **58**, 1.2.1-1.2.12 (2017).

32. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

33. Wishart, D. S. *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **46**, D1074–D1082 (2018).

34. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27**, 221–234 (2013).

35. Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **49**, 377–389 (2009).

36. Friesner, R. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).

37. Halgren, T. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **47**, 1750–1759 (2004).

38. Case, D. A. *et al.* AmberTools. *J. Chem. Inf. Model.* **63**, 6183–6191 (2023).

39. Lomize, A. L., Pogozheva, I. D., Lomize, M. A. & Mosberg, H. I. Positioning of proteins in membranes: A computational approach. *Protein Sci.* **15**, 1318–1333 (2006).

40. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).

41. Tian, C. *et al.* ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **16**, 528–552 (2020).

42. Dickson, C. J., Walker, R. C. & Gould, I. R. Lipid21: Complex Lipid Membrane Simulations with AMBER. *J. Chem. Theory Comput.* **18**, 1726–1736 (2022).

43. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).

44. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).

45. Pastor, R. W., Brooks, B. R. & Szabo, A. An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol. Phys.* **65**, 1409–1419 (1988).

46. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

47. Gene Ontology Consortium *et al.* The Gene Ontology knowledgebase in 2023. *Genetics* **224**, iyad031 (2023).

48. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).

Author contributions

J.J.G carried out study design and performed the BBB permeability prediction. J.W. conducted GOF genes search and help to write the partial Introduction part. X.Y.L. handled the remaining part and write the manuscript. Y.Y. and L.R.P provided advice on study design and help to revise the manuscript. All authors reviewed the results and approved the final version of the manuscript.

Competing interests

The author(s) declare no competing interests.

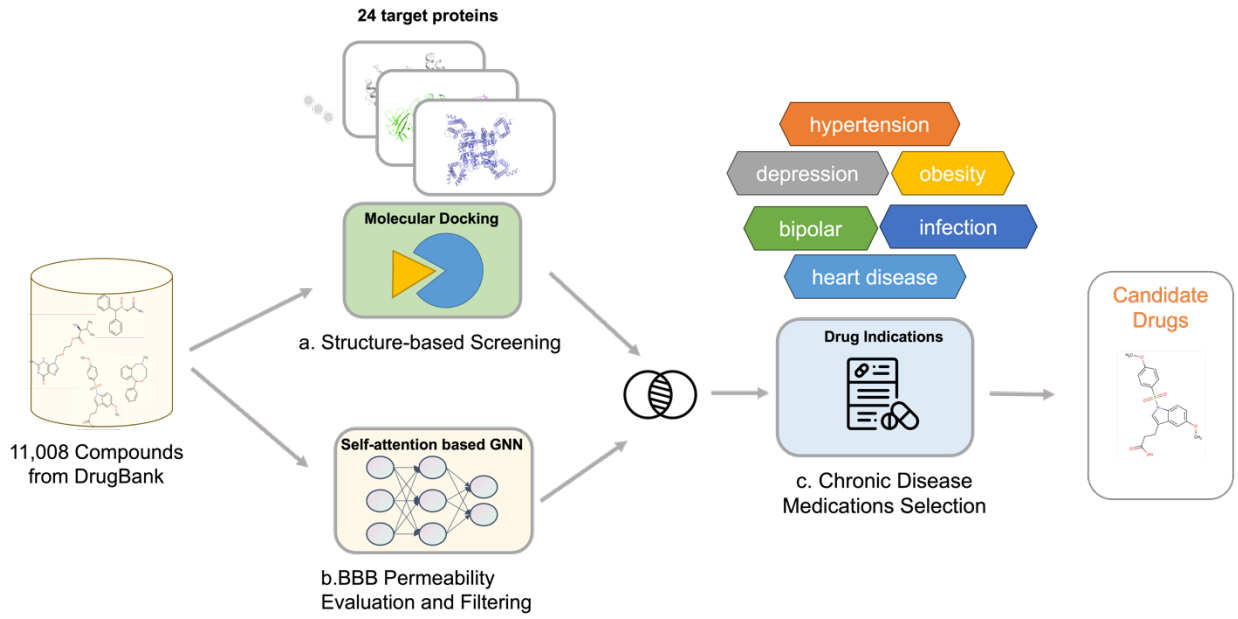

Fig 1 The *in-silico* drug repurposing pipeline for epilepsy. In our pipeline, we integrate two stages: structure-based screening (a) and BBB permeability evaluation and filtering (b) to screen potential medications from 11,008 compounds sourced from DrugBank database. (c) Medications prescribed for chronic diseases, including hypertensin, depression and infection disease, are preferred for seizure control in our *in-silico* pipeline.

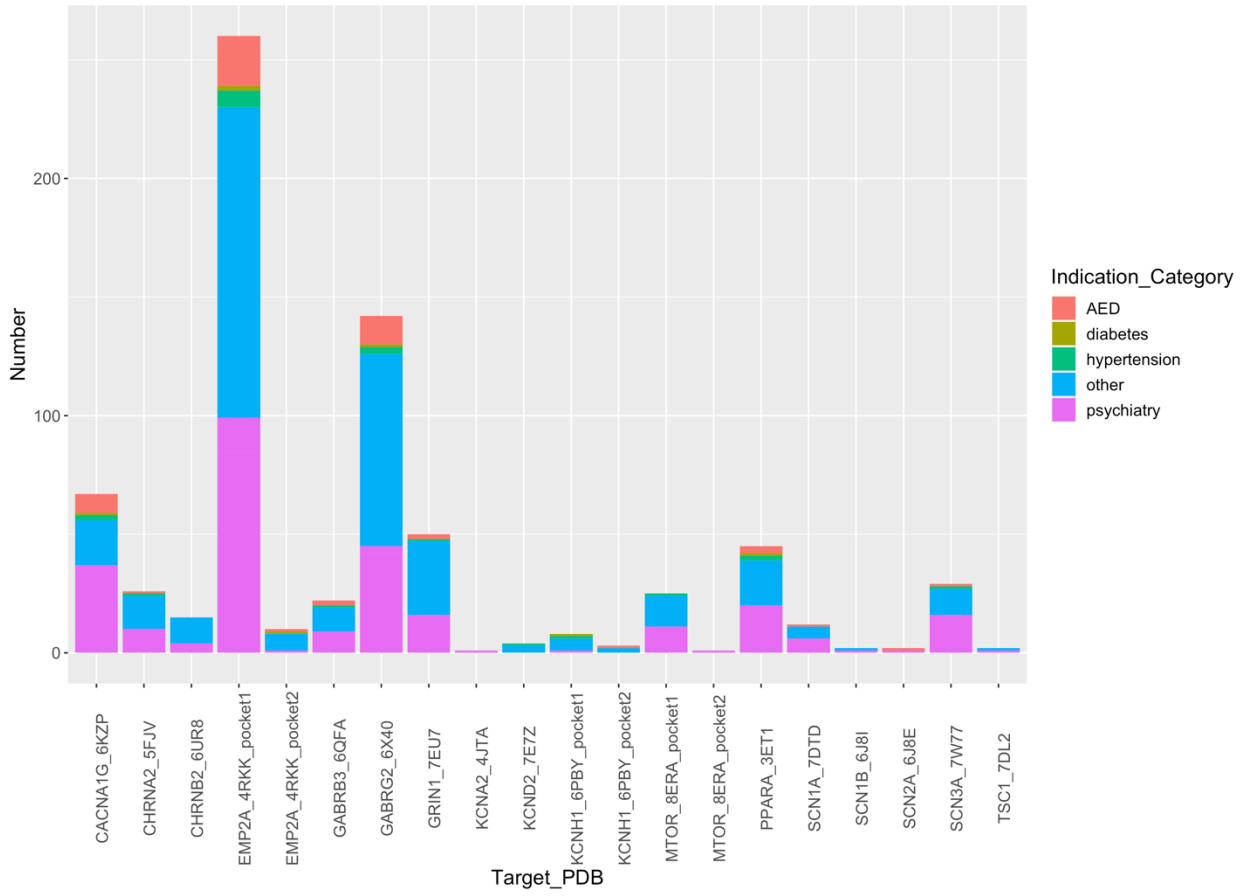

Fig 2 The distribution of candidate drugs selected by our *in-silico* pipeline. AED means anti-epileptic drug.

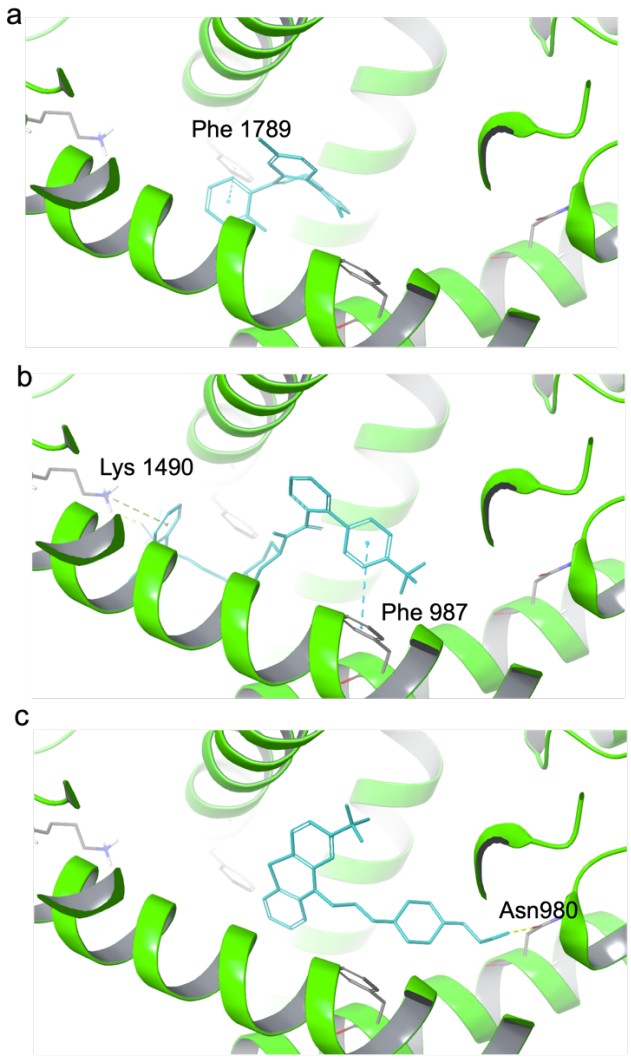

Fig 3 The binding pose of Midazolam, Lomitapide and Fluphenazine in SCN1A target. The 7DTD protein (cartoon) is colored by green and the ligands (stick) (a: Midazolam, b: Lomitapide, c: Fluphenazine) by cyan. The key residues (stick) in protein are colored by gray. The hydrogen-bond, cation-pi interaction and pi-pi interaction are colored by yellow, green and blue dotted line, respectively.

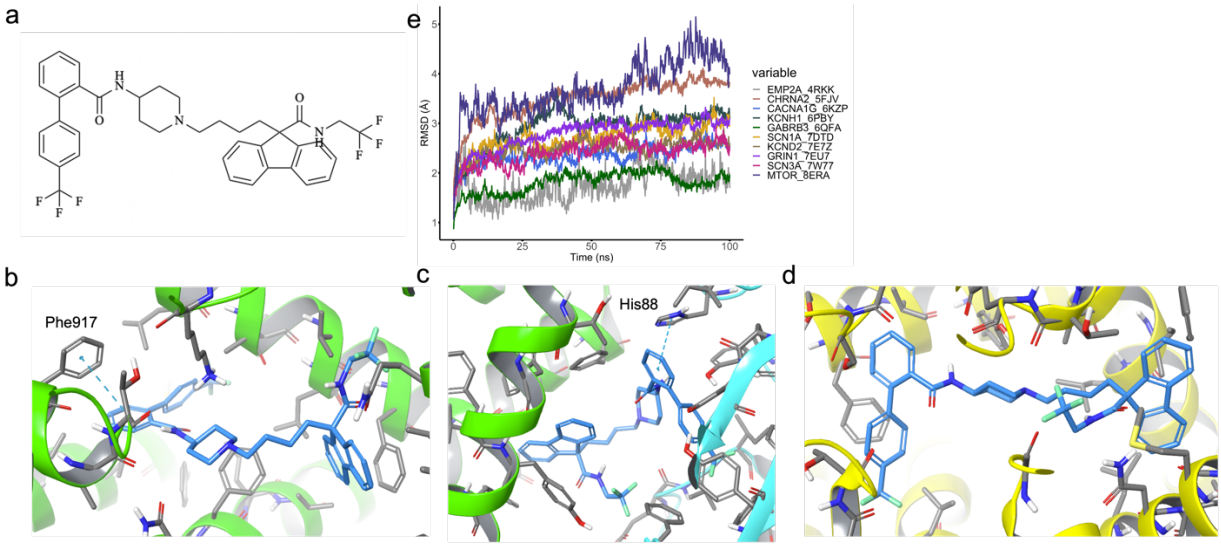

Fig 4 Lomitapide structure and binding pose in selected multiple targets with protein backbone RMSD in 100-ns MD simulations. The 2D structure of Lomitapide is in panel a. The binding poses of Lomitapide in CACNA1G (panel b), MTOR (panel c) and SCN3A (panel d) targets. These targets are colored by chain name (chain A: green, chain B: blue, chain D: yellow). The key residues and ligands (stick) in protein are colored by gray and blue, respectively. The pi-pi interaction is colored by blue dotted line. The RMSD of protein backbone (referred by initial docking pose) for 100-ns MD simulation in each target are showed in panel e.

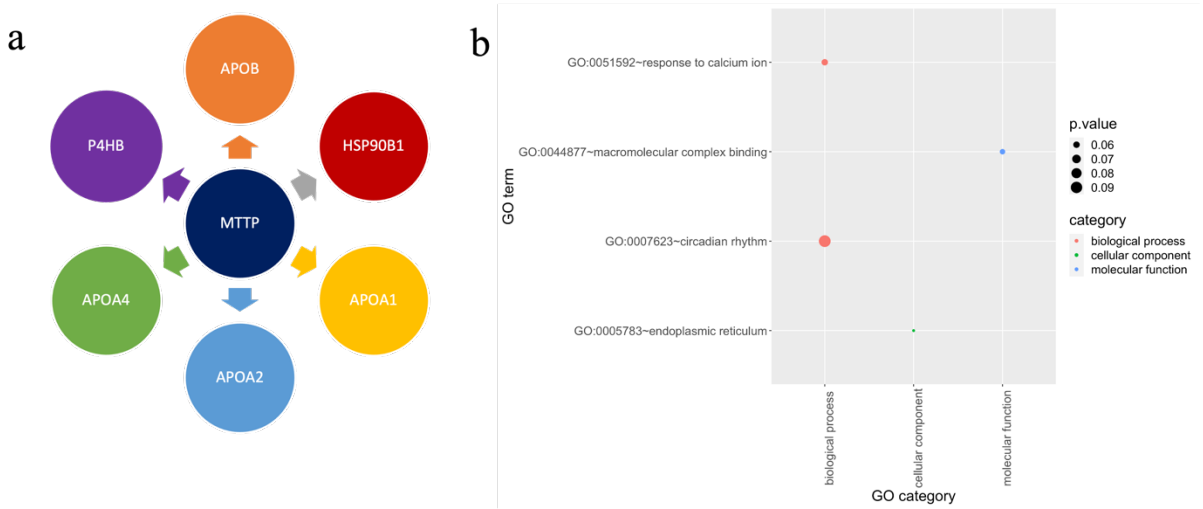

Fig 5. The mechanism of action exploration of Lomitapide. The genes that interact with MTTP gene in PPI database are shown in panel a. Panel b is the shared pathway of targets between Lomitapide and AEDs. The size of cycle represents p-values and the GO categories being distinguished by color: red for biological process, green for cellular component and blue for molecular function.

Table 1 24 GOF targets and the number of repurposable drugs in our pipeline

GOF gene	PDB ID	Reference compound	Docking score	Number of candidates
CACNA1G	6KZP	DZR (co-cry)	-6.94518	73
KCNA2-KCNB1-KCNB2-KCNB3	4JTA	DB03461 (co-cry)/-	-9.88066/-7	0/1
KCNH1	6PBY	-/-	-7/-7	10/5
KCNQ2	7CR2	DB04593	-10.7923	0
SCN1A	7DTD	DB00564	-7.91946	14
SCN1B-SCN9A	6J8I	DB05232 (co-cry)	-7.07374	2
ATP6V1A	6WM3	ADP (co-cry)	-9.73841	0
CHRNA2	5FJV	DB07720	-7.94885	28
CHRNB2	6UR8	DB01273	-8.32579	17
EMP2A	4RKK	-/-	-7/-7	282/12
FGF12(FHF1)-SCN5A	4JQ0	-	-7	0/0
FGF13(FHF2)-SCN2A	4JPZ	-	-7	0
GABRB3	6QFA	DB01174	-6.68507	25
GABRG2	6X40	DB01068	-6.51669	149
GRIN1	7EU7	-	-7	53
HCN2	3U10	CMP (co-cry)	-12.2342	0

KCND2	7E7Z	-	-7	4
MTOR	8ERA	DB00877 / XZ9	-7/-9.12847	30/1
PPARA	3ET1	DB13115 (co-cry)	-8.37817	49
RHEB	6BT0	E7V(co-cry)/ GDP (co-cry)	-6.08642/-10.8434	0/0
RHEB	6BSX	E7S(co-cry)/ GDP (co-cry)	-6.71108/-10.5601	0/0
SCN2A	6J8E	DB00564	-9.20486	2
SCN3A-SCN1B-SCN2B	7W77	DB00564	-8.34858	33
SCN8A	8GZ2	DB05232	-7.34028	0
TSC1-TSC2	7DL2	-/-	-7	2

The "-" character in "GOF gene" column means that multi-genes consist of a complex structure. The second column is the protein PDB ID encoded by corresponding GOF targets. In third column, the "/" represents there are two binding pockets in structure and the "co-cry" means the reference compound from crystal structure. The fourth column is the binding score of reference compounds and dash means the docking score threshold is set to -7 due to no available reference compound. The fifth column is the number of repurposable drugs selected using our pipeline.

Table 2 The docking score and rank of Nebivolol, Lomitapide and Pimozide in drug list

Target (nebi)	**Docking score**	**Docking rank**	**Target (lomi)**	**Docking score**	**Docking rank**	**Target (pimo)**	**Docking score**	**Docking rank**
3ET1	-10.3592	5	4RKK-1	-9.44587	7	3ET1	-11.3451	1
4RKK-1	-9.08446	13	5FJV	-8.27674	12	4RKK-1	-9.91056	2
5FJV	-8.66181	6	6KZP	-9.59345	1	6J8E	-9.30785	1
6KZP	-8.0873	5	6PBY-1	-8.75374	1	6KZP	-7.42988	31
6PBY-1	-7.28706	7	6QFA	-7.77517	2	6QFA	-7.53916	3
6QFA	-7.23133	8	7DTD	-8.63426	2	6X40	-8.20083	4
6X40	-8.05761	7	7E7Z	-7.35713	2	7DTD	-8.55467	5
7E7Z	-7.27256	3	7EU7	-7.71214	21	7EU7	-7.15027	45
7EU7	-7.44296	28	7W77	-10.6177	1	7W77	-9.93179	5
7W77	-9.31145	12	8ERA-1	-8.86007	1			
8ERA-1	-7.61246	9						

The “-1” in “Target” column represents “pocket 1”.

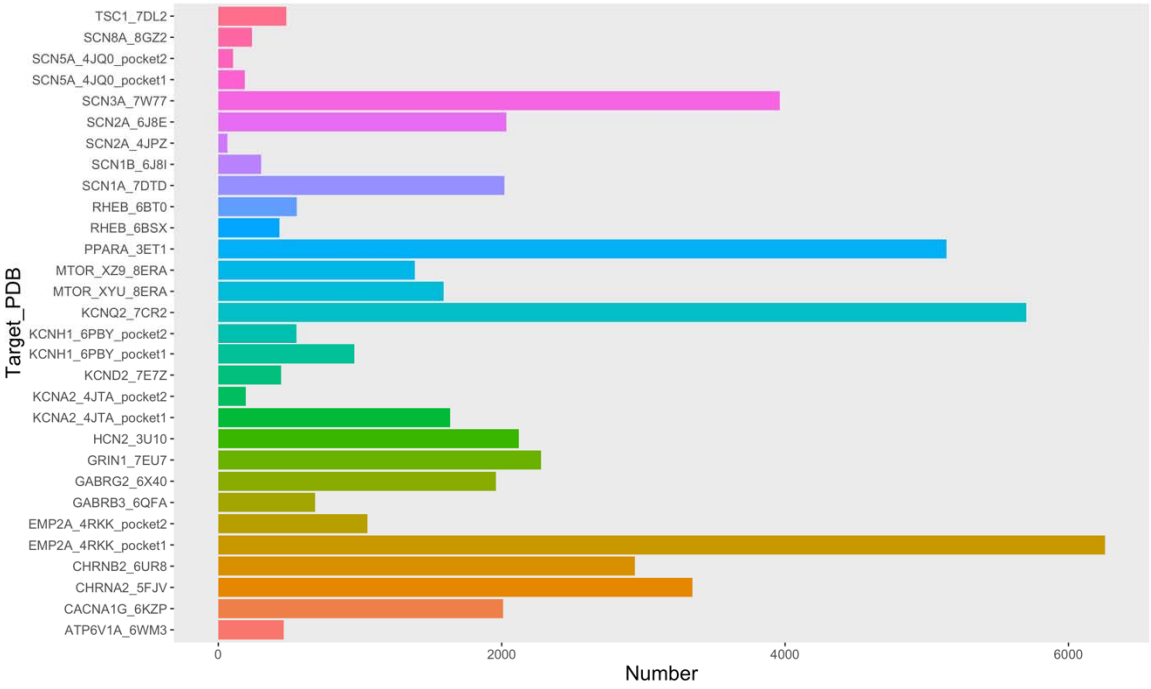

Fig S1 The number of candidates after docking filter using -7.0 as threshold criterion.

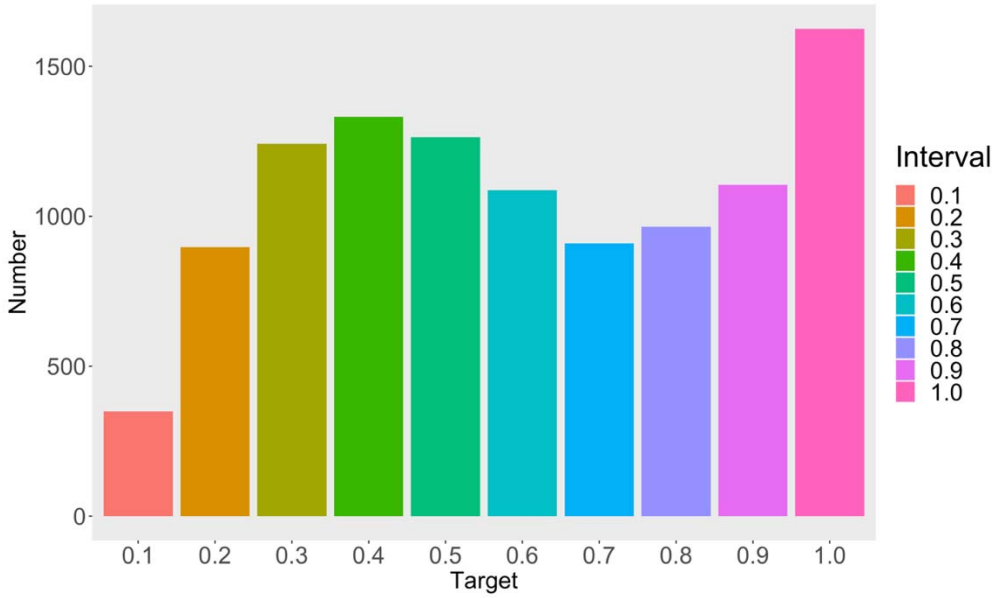

Fig S2 The distribution of predicted BBB permeability probability for compounds from DrugBank database.

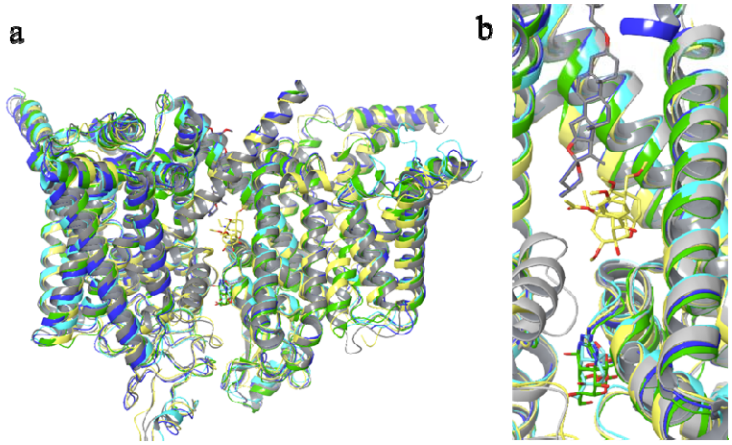

Fig S3 The five voltage-gated sodium channel structures. a. Using SCN1A (7DTD:gray) as a reference structure, the other four structures were aligned using protein (cartoon). The RMSD for SCN2A (6J8E:blue), SCN3A (7W77:yellow), SCN8A (8GZ2:cyan) and SCN9A (6J8I:green) are 1.664 , 1.373 , 1.779 and 1.664 , respectively. b. The close-up view of panel a. These ligands are represented by stick with the same color of corresponding protein in panel a.

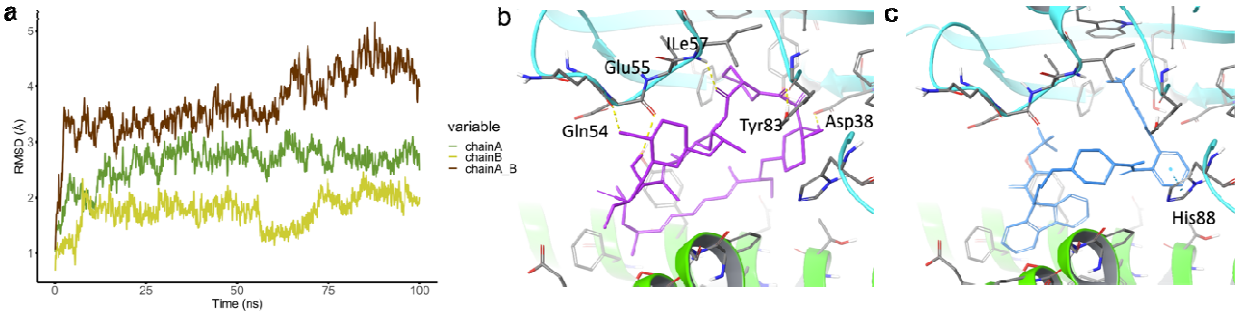

Fig. S4 The RMSD plot and ligand-receptor interaction for the target (PDB ID: 8ERA). Panel a, RMSD plot for the protein backbone, calculated across 8ERA complex and its chain A and B during the 100ns production phase. In panel b and c, chain A encoded by *MTOR* gene and chain B encoded by *FKBP1A* gene was highlighted in green and cyan, respectively. The co-crystal XYU (panel b) and Lomitapide (panel c) were colored by purple and blue, respectively.

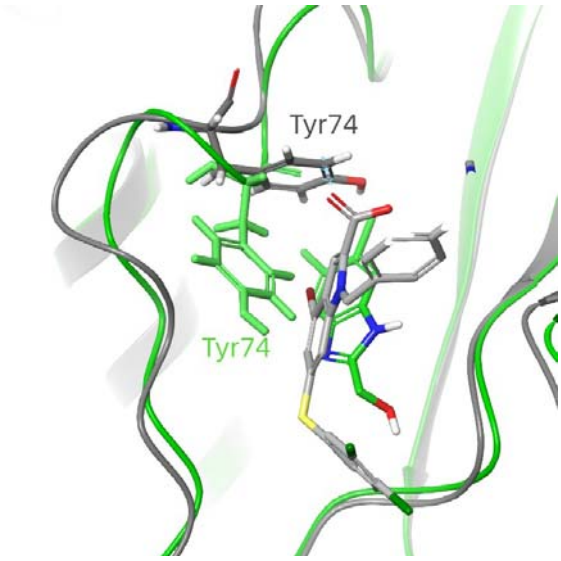

Fig S5 The two conformations of Tyr74 in 6BT0 (gray) and 6BSX (green) crystal structures. For RHEB target, the binding poses of ligand E7V (gray) and E7S (green) are shown.

Fig S6 Ligandformer architecture. This is a multi-layer single-head self-attention-based Graph Neural Network framework, for predicting compound chemical property with robust interpretation. Facilitated with visualization technique, the map shows insights on AI model's rationales on judging which parts of an input molecule impact certain property predictions.

Fig. S7 The structure of 7EU7. Only transmembrane region was retained in 100-ns MD simulation (below dotted line.)

Table S1 The BBB permeable probability of anti-epileptic drugs in market predicted by Ligandformer model

Drug	**ID**	**Probability**
Carbamazepine	DB00564	0.9117
Valproic acid	DB00313	0.9282
Phenytoin sodium	DB00252	0.8325
Oxcarbazepine	DB00776	0.94
Diazepam	DB00829	0.97
Levetiracetam	DB01202	0.95
Fosphenytoin	DB01320	0.8445
Flunarizine	DB04841	0.9560

Movie S1-S10 The 100-ns MD simulation for lomitapide in EMP2A (PDB ID: 4RKK), CHRNA2 (PDB ID: 5FJV), CACNA1G (PDB ID: 6KZP), KCNH1 (PDB ID: 6PBY), GABRB3 (PDB ID: 6QFA), SCN1A (PDB ID: 7DTD), KCND2 (PDB ID: 7E7Z), GRIN1 (PDB ID: 7EU7), SCN3A (PDB ID: 7W77), MTOR (PDB ID: 8ERA) targets.