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

Epilepsy is a severe neurological disorder characterized by its chronic predisposition to recurrent epileptic seizures. Due to its considerable global prevalence and healthcare burden, the pursuit of effective new medication for epilepsy treatment remains an urgent and significant challenge in medical practice. 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 approved for chronic conditions against 24 gain-of-function (GOF) encoding targets implicated in epileptogenesis ranged from zero to several hundred. Nearly half of candidate medications for each target were used for psychiatric diseases and most of anti-epileptic drugs (AEDs) in market were identified, highlighting the effectiveness of our pipeline. Furthermore, Lomitapide, a cholesterol-lowering drug, emerged as particularly noteworthy, exhibiting high binding affinity for ten targets and verified by molecular dynamics (MD) simulation. These findings demonstrated the practical potential of our *in-silico* drug repurposing pipeline in epilepsy treatment, providing 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 permeability

Introduction

Epilepsy, a central neurological disease, can be precipitated by any direct or indirect insult to the CNS, and often accompanied by cognitive, psychological, and social 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 symptomatic treatment with AEDs, 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 along with 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 hold potential 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[11]. Meanwhile, the integration of structure based and deep learning methods have significantly enhanced drug repurposing, particularly in elucidating structure-activity relationships and accelerating drug discovery[12,13]. The Experimental Epilepsies (PDE3) database, developed by Nasir group through literature mining, has cataloged drugs that have demonstrated efficacy in experimental epilepsy models[14]. Reutens's laboratory has 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[15]. Moreover, potential epilepsy treatments have been identified by contrasting gene expression profiles of epileptic patients against

healthy controls, using drug perturbation data from the Connectivity Map (CMAP) database[11]. A successful example of drug repurposing is Igalmi, initially approved by FDA in 1999 for sedation and analgesia in intensive care settings. BioXcel biopharmaceuticals has 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 proteins encoded by 24 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 demonstrate anti-seizure activities against specific targets. Moreover, for Lomitapide repurposed by our pipeline, 100-ns MD simulation against each selected target were conducted to further validate its potential inhibition for epilepsy.

Results

2.1 Reduce candidates number using individualized reference compounds

In the preliminary stage of our pipeline, Glide SP docking was utilized to screen a set of 11,008 compounds against 24 targets. 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 was 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 occurred.

In the case of seven structures featuring two putative binding pockets, the quantity of compounds exhibiting strong binding affinity was generally higher in the first pocket compared to the second. 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.

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 adapted 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 an harmonic average score of 0.925 for the area under the precision-recall curve (AUPR).

In our analysis of 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 S2). The consistent pattern underscores the model's robust capability for predicting BBB permeability. To avoid omitting any potential drugs for epilepsy, we selected candidate drugs with BBB permeability probability score above 0.75 for subsequent stage.

2.3 The potential drugs screened for 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 GOF targets. However, for eight targets encoded by—namely, *KCNQ2*, *ATP6V1A*, *FGF12*, *FGF13*, *HCN2*, *RHEB* and *SCN8A*—there were no suitable potential medications available, as detailed in Table 1. Despite the compatibility of numerous drugs with their crystal structures in the docking analysis, none exhibited sufficient probabilities of BBB permeability, based on Ligandformer predictions.

In the joint filter analysis, the number of potential candidate medications range from zero to hundreds, particularly exceeding one hundred for target (pocket 1) encoded by *EMP2A* and *GAGRB2* genes (Fig 2). The lack of a reference compound for protein encoded by *EMP2A* likely contributed to the identification of 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 screened twenty-seven antiepileptic drugs (AEDs), confirming its efficacy in identifying suitable treatments for

epilepsy. Additionally, a few antihypertensive and antidiabetic drugs were recognized for their potential efficacy in epilepsy management.

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. 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 repurposable 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 ten 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 against ten selected targets

To 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 ten systems, as evidenced by convergence of backbone RMSD within the 100 ns production phase (Fig. 4E), with the exception of protein encoded by *CHRNA2* and *MTOR* gene. 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 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 hydrogen bonds with Ile57 and Tyr83 in *FKBP1A* encoding protein (chain B of 8ERA) were not present. 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. S4B).

Discussion

Several prior studies on drug repurposing for epilepsy have primarily utilized approaches such as literature search[14], drug-target network[7] and gene expression profiling[11], 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. In the Schrödinger suite, docking scores are calculated through mathematical equations that account for a range of interactions such as van der Waals forces, hydrogen bonds, and ionic interactions. This simulation-based methodology inherently causes deviations from the true binding affinities, which is particularly evident in the inconsistent conformations of esketamine seen in docking simulations when compared to its resolved crystal structure (Fig. S5). Therefore, it is essential to validate these computational predictions with experimental binding assays.

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 hundred, contingent on the reference compounds used for each target. Notably, our analysis identified common AEDs as possible candidates, underscoring the efficacy and robustness of our strategy in drug repurposing efforts. For the first time, Lomitapide emerged as a promising candidate medication, demonstrating stability across ten targets in MD simulations. 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.

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[27], Online Mendelian Inheritance in Man (OMIM)[28] and Development Disorder Genotype - Phenotype Database (DDG2P) using "epilepsy" and "gain of function" as keywords. Finally, seventy-two genes were selected as target genes for drug repurposing. The reported evidence for GOF spans three levels-variant, gene, disease- detailed in Supplemental Table S1.

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 targets. Only 25 available crystal structures with appropriate binding pockets were downloaded from the PDB database[29] 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. S6). From the DrugBank database (version *2022.1.17*)[30], 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[31] 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[32]. 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 11008 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 standard precision (SP) docking protocol, which utilized the OPLS_2005 forcefield and default parameters[33,34].

4.4 BBB prediction Evaluation

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 S7.

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 Potential candidate drugs approved by FDA-approved 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 ten potential Lomitapide targets

We conducted MD simulations on the docking conformations produced by Lomitapide for ten 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. S8). We utilized the Tleap module in AMBER 22[35] to systematically delete and subsequently reintroduce hydrogen atoms. For the three solvent proteins (PDB ID: 4RKK, 5FJV, 8ERA) systems, we built a TIP3P water box along XYZ axis, extending 12 Å and neutralized the entire systems with a NaCl concentration of 0.15 M. For the rest seven membrane protein systems, we additionally oriented the direction of transmembrane region along Z-axis using PPM 2.0[36] and embedded the proteins into 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) bilayer using CHARMM-GUI membrane builder[37].

All-atom MD simulations were conducted employing PMEMD engine in AMBER 22[35]. The AMBER FF19SB force field[38] and AMBER lipid force field LIPID21[39] 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[40]. The particle mesh Ewald (PME) algorithm[41] 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[42]. 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.

Author contributions

J.J.G carried out study design and performed the BBB permeability prediction. X.Y.L. handled the docking part and write the manuscript. J.W. conducted GOF genes search. Y.Y. and L.R.P 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.

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).

[Type here]

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. Brueggeman, L. *et al.* Drug repositioning in epilepsy reveals novel antiseizure candidates. *Ann. Clin. Transl. Neurol.* **6**, 295–309 (2019).
12. Pinzi, L. & Rastelli, G. Molecular Docking: Shifting Paradigms in Drug Discovery. *Int. J. Mol. Sci.* **20**, 4331 (2019).
13. 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).
14. 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).
15. 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).
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. White, J. PubMed 2.0. *Med. Ref. Serv. Q.* **39**, 382–387 (2020).
28. 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).

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

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

31. 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).

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

33. 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).

34. 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).

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

36. 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).

37. 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).

38. 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).

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

40. 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).

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

42. 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).

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

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”.

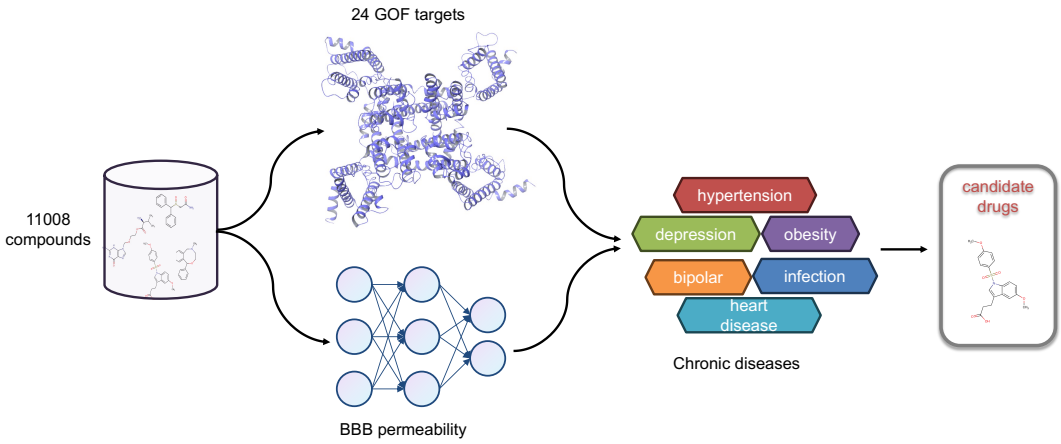

Fig 1 The *in-silico* drug repurposing pipeline for epilepsy.

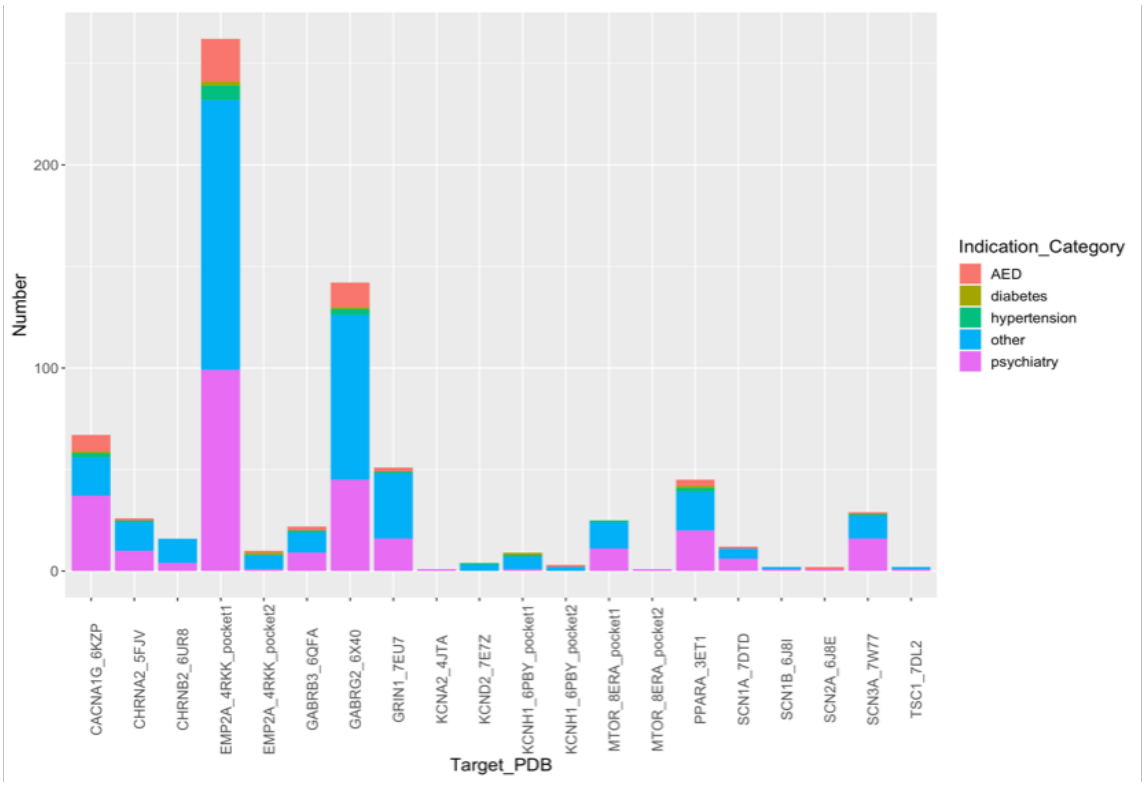

Fig 2 The distribution of candidate drugs selected by our *in-silico* pipeline. AED means anti-epileptic drugs.

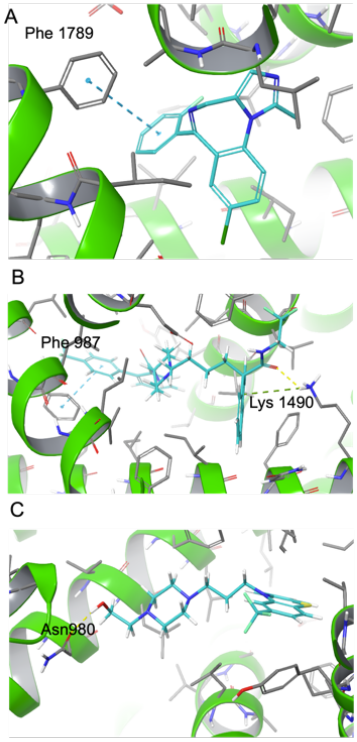

Fig 3 The binding pose of midazolam, Lomitapide and fluphenazine in SCN1A target. The 7DTD protein (cartoon) is colored by green and the ligands (stick) (A: midazolam, B: Lomitapide, C: fluphenazine) by cyan. The key residues (stick) in protein are colored by gray. The hydrogen-bond, cation-pi interaction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

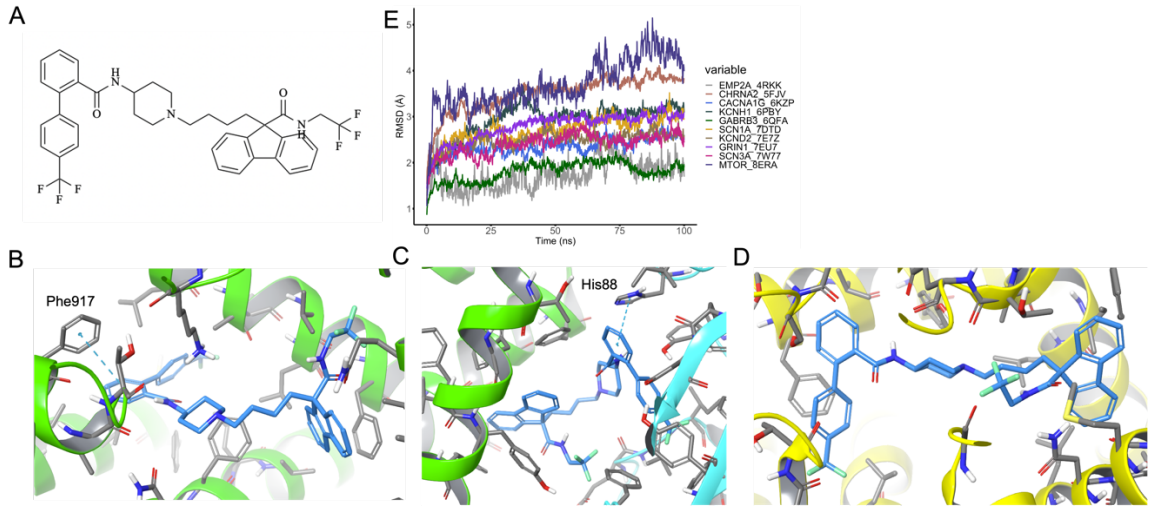

Fig 4 Lomitapide structure and binding pose in selected multiple targets with protein backbone RMSD in 100-ns MD simulations. The 2D structure of lomitapide is in panel A. The binding pose of Lomitapide in CACNA1G (panel B), MTOR (panel C) and SCN3A (panel D) targets. These targets are colored by chain 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.