

Building structure activity relationships to avoid toxicity due to unwanted central nervous system ion channel activity

Louisa A.K. Zolkiewski ¹, Kimberly L. Rockley¹, Ruth A. Roberts ^{1,2,*}, Hannah G. Jennings¹, Magali-Anne Maizières¹, Karen Jones¹, James L. McDonagh¹, Katie Newman¹, Ed J. Griffen³, Jessica E. Stacey³, Michael J. Morton¹

¹Apconix, Macclesfield SK10 4TG, United Kingdom

²Department of Biosciences, University of Birmingham, Edgbaston B15 1TT, United Kingdom

³MedChemica Consultancy Ltd, Motorworks, Macclesfield, Cheshire SK11 6DU, United Kingdom

*Corresponding author: Department of Safety Science, Apconix, Alderley Park, Cheshire SK10 4TG, United Kingdom. Email: ruth.roberts@apconix.com

Abstract

We previously described an integrated in vitro liability assay for seizure, a new approach methodology (NAM) to reduce toxicity due to central nervous system (CNS) liability in drug discovery and development. Here we report the development of structure activity relationships (SAR) to guide drug design away from this liability. SAR test compounds were selected from the Enamine REadily Accesible (REAL) database using pharmacophore features and similarity to previous test compounds (amoxapine, diphenhydramine, quetiapine, 4-AP, linopirdine) or to ion channel positive reference compounds (bepridil, NS1619, quinidine, verapamil, XE991). These 88 compounds (10 parent compounds and 78 structurally related derivatives) were screened by automated electrophysiology in cell lines expressing human $K_v2.1$, $Na_v1.2$, the $\alpha_1\beta_2\gamma_2$ GABA_A or $\alpha_4\beta_2$ nicotinic receptors to generate IC₅₀ values. Across all 4 ion channels, the derivative compounds exhibited increased, equivalent or reduced potency compared with the parent compounds, providing a complex and rich dataset for SAR. Regarding individual pharmacophoric features, statistical analysis identified 12 features significantly associated with activity at the $\alpha_4\beta_2$ nicotinic receptor; 4 of these were also significantly associated with activity at $K_v2.1$. Selected parent compounds (amoxapine, diphenhydramine, quetiapine) and structural analogues were screened for seizure-like activity in human induced pluripotent stem cell neurons using microelectrode array. The seizure-like phenotype was altered with the derivatives, as expected from the respective ion channel IC₅₀ values. These data provide insight into specific substructures associated with seizure-like drug toxicity, offering the opportunity to avoid CNS liability in the development of novel compounds, saving time, money, and resources.

Key words: SAR; CNS toxicity; drug development; drug discovery; seizure; ion channels; MEA; hiPSC neurons; iLA^{Seizure}.

Central nervous system (CNS) toxicities, specifically drug-induced seizures, frequently occur during drug discovery and development (Cook et al. 2014; Authier et al. 2016), leading to project delays and failures. Compounds associated with CNS liability span a variety of therapy areas including CNS, cardiovascular, gastrointestinal, respiratory, and inflammation (Easter et al. 2009; Authier et al. 2016; Walker et al. 2018). Seizures are usually detected during preclinical rodent and nonrodent studies or even later, once the drug has entered clinical development, leading to a significant waste of time and resources. Progress has been made in seizure detection using automated video systems that record and analyze animal movements (Yip et al. 2019), and substantial effort has been invested in developing screening methods that could be used earlier in drug discovery, such as hippocampal brain slice (Easter et al. 2007) or zebrafish larval locomotor assays (Winter et al. 2008). However, these methods have had limited success, with concerns regarding throughput and translation of rodent and zebrafish data to humans.

We have previously described an in vitro liability assay for seizure (iLA^{Seizure}) based on measuring electrical activity in induced human pluripotent stem cell (hiPSC) neurons using microelectrode array (MEA), coupled with a neuronal ion channel panel (Rockley et al. 2023). Activity of 15 known seizurogenic and 1 negative compounds at 15 seizure-related ion channels ($Na_v1.1$, $Na_v1.2$, $Na_v1.6$, $K_v7.2/7.3$, $K_v7.3/7.5$, $K_v1.1$, $K_v4.2$, $K_{Ca4.1}$, $K_v2.1$, $K_v3.1$, $K_{Ca1.1}$, $Ca_v2.1$, and the $\alpha_1\beta_2\gamma_2$ GABA_A, $\alpha_4\beta_2$ nicotinic and GluN1/2A NMDA receptors) was assessed using automated patch-clamp electrophysiology. The data demonstrated inhibition of function across ion channels with some notable patterns. Specifically, there was a similar inhibition profile across the 3 sodium (Na_v) channels, the $K_v2.1$ potassium channel was inhibited the most among the CNS active therapies and the nicotinic $\alpha_4\beta_2$ receptor was the most frequently hit channel (Rockley et al. 2023). A further 7 negative compounds have been assessed using iLA^{Seizure} since the publication of this work. Of these 23 screened compounds, there were 13 true positives, 7 true negatives, 1 false positive, and 2 false negatives.

This resulted in predictivity, sensitivity, specificity, and positive and negative predictive values for the assay of 87%, 87%, 88%, 93%, and 78%, respectively.

For compounds identified as a potential seizure risk, medicinal chemists face the challenge of redesigning the molecule to avoid the potential termination of a promising lead series. The ability to link chemical structures to molecular properties, such as potency or stability, is known as structure activity relationship (SAR); for example, “adding a chlorine at 2-position adds a 3-fold potency but reduces solubility 10-fold” (Griffen et al. 2011). Methods to establish these relationships and develop rules to guide design have been used for conventional small molecules, novel modalities, and even materials design (Allen et al. 2024; McDonagh et al. 2024). As seen with the human *Ether-à-go-go*-Related Gene (hERG), a potassium channel associated with cardiac liability, understanding, and modelling SAR can direct redesign to avoid substructures associated with unwanted hERG activity (Gund 1977; Garrido et al. 2020; Creanza et al. 2021). Alternatively, considering the impact of specific changes, such as “How much potency does adding a chlorine at the 2-position generally give?” can prove informative. Indeed, using quantitative SAR (QSAR) approaches, several well-established pharmacophore features have been associated with hERG inhibition, which can be used to guide strategies to design out hERG liabilities (Garrido et al. 2020). A pharmacophore is defined as “a set of structural features in a molecule that is recognized at a receptor site and is responsible for the molecule’s biological activity”; typical features include hydrophobic centroids, aromatic rings, hydrogen bond acceptors or donors, cations, and anions (Gund 1977; Wermuth et al. 1998).

Inspired by the hERG successes, we embarked on an SAR study to identify the pharmacophore features responsible for the inhibition of seizure-related ion channels: Na_v1.2, K_v2.1, and the $\alpha_1\beta_2\gamma_2$ GABA_A and $\alpha_4\beta_2$ nicotinic receptors. For this study, the key pharmacophore features of selected original test compounds (amoxapine, diphenhydramine, quetiapine, 4-AP [4-Aminopyridine], linopirdine) plus several ion channel positive reference compounds (bepridil, NS1619, quinidine, verapamil, XE991) were identified. Structural analogues were extracted from the Enamine REadily Accesible (REAL) database by scaffold searching and then selected based on the similarity of the pharmacophore features to the parent compounds. The Enamine REAL database is an extremely large enumerated database of synthetically feasible molecules (>1.2 billion) that can assist in SAR investigations (Grygorenko et al. 2020; “REAL Database—Enamine,” n.d.). We also investigated electrical activity by MEA in hiPSC neurons to confirm that significant differences in ion channel potencies translated to a phenotypic change between parent and derivative compounds.

The data demonstrate that certain pharmacophore features present a significant risk of CNS ion channel inhibition, which translates to perturbed electrical signaling in hiPSC neurons. Together, this highlights the significant value of developing a robust predictive model that can assess SAR of compounds with seizurogenic potential to enable structural chemists to design-out CNS liability in the development of novel compounds by ensuring test molecules do not have high-risk groups, thus eliminating costly project failures.

Materials and methods

Ion channels

Chinese hamster ovary (CHO-K1) or human embryonic kidney (HEK293) cells stably expressing the wild-type human ion

channels Na_v1.2, K_v2.1 and the $\alpha_1\beta_2\gamma_2$ GABA_A and $\alpha_4\beta_2$ nicotinic receptors were generated by standard recombinant techniques (Green and Sambrook 2012). The voltage-gated sodium and potassium channels were created at ApconIX, and the ligand-gated channels were purchased from B'SYS GmbH (Witterswil, Switzerland).

Cell culture

CHO-K1 cells were cultured in F-12 Ham medium (cat. N6658, Sigma-Aldrich, Merck, St. Louis, MO, United States) supplemented with 10% v/v fetal bovine serum (cat. SV30160.03, Cytiva, Marlborough, MA, United States), 1% v/v nonessential amino acids (cat. 11140-050, Gibco, ThermoFisher Scientific, Waltham, MA, United States), 1% v/v penicillin-streptomycin (cat. P4333, Sigma-Aldrich), and the appropriate selection antibiotics. HEK293 cells were cultured in a 50:50 mix of F-12 Ham and Dulbecco's Modified Eagles Medium (cat. SH30081.01, Cytiva) supplemented with 10% v/v fetal bovine serum, 1% v/v nonessential amino acids, 1% v/v penicillin-streptomycin, and the appropriate selection antibiotics.

Compounds

The compound set used to investigate SAR at CNS ion channels consisted of 10 parent compounds, either seizure-associated compounds from the initial in vitro assay validation study (Rockley et al. 2023) or ion channel positive reference compounds, purchased from either Sigma-Aldrich or TOCRIS (Bristol, United Kingdom). A further 78 compounds were then selected and purchased from the Enamine REAL database (Kyiv, Ukraine). Scaffold searches were performed for each of the reference compounds within the Enamine REAL database, the closest 140 compounds in each case by extended-connectivity fingerprints (ECFP) were then further filtered by counts of pharmacophore features.

Pharmacophore features were defined as a pair of substructures connected by their shortest topological path, similar to that described by Reutlinger et al. (2013). The shortest path distance between 2 features is within the same group or a direct link, and the longest link considered was 8 atoms separated. A diagrammatic example is provided in Fig. S1. The pharmacophore features were hydrophobic group; hydrogen bond donors; hydrogen bond acceptors; aromatic ring attached; aliphatic ring attached; halogen bonds; basic group; acidic group (Table 1). Compounds were then picked to iteratively remove or reduce the number of each pharmacophoric feature present in the reference compound and the candidate compound to represent a test of each feature. This resulted in parent and (derivative test compound numbers) as follows: Amoxapine (*n* = 10), diphenhydramine (*n* = 9), quetiapine (*n* = 10), 4-AP (*n* = 10), linopirdine (*n* = 8), bepridil (*n* = 9), NS1619 (*n* = 9), quinidine (*n* = 9), verapamil (*n* = 2), and XE991 (*n* = 2).

Stock solutions were made in dimethyl sulfoxide (DMSO; cat. D8148, Sigma-Aldrich) to 33 mM and stored at −20°C until use.

Screening of compounds

QPatch

Automated patch-clamp electrophysiology experiments (QPatch II, Sophion Biosciences, Ballerup, Denmark) (Bell and Dallas 2018) were performed as previously described (Rockley et al. 2023). Briefly, stock compounds in DMSO were diluted to their top test concentration in HEPES-buffered saline. Single-cell ionic currents were measured in whole-cell configuration at room temperature (25°C). Compounds were incubated for at least 120 s. Concentration-response curves were generated by cumulative addition of compound with concentrations low to high. In all cases, steady-state inhibition was achieved before the next concentration of

Table 1. Pharmacophore descriptions used to define a pharmacophore feature.

Pharmacophore	Abbreviation	Description
Basic Group	BG	Atom or group most likely protonated at pH 7.4 (definitions derived from MedChemica proprietary SMARTS library)
Acidic Group	AG	Atom or group most likely deprotonated at pH 7.4, includes N and C acids (definitions derived from MedChemica proprietary SMARTS library)
Hydrogen Bond Acceptor	HA	Definitions derived from Taylor et al. (2012)
Hydrogen Bond Donor	HD	Definitions derived from Taylor et al. (2012)
Hydrophobic	LH	C4 or greater, cyclic or acyclic alkyl group
Aromatic Attachment	AR	Connection of any group to an aromatic atom excluding connections within rings
Aliphatic Attachment	RR	Connection of any atom to an aliphatic group not in a ring
Halogen bonds	Hal	F, Cl, Br, I, CF ₃ , OCF ₃

A pharmacophore features is a pair of pharmacophores connected by their shortest topological path. The shortest path distance between 2 distances are: Same group, direct link and 1 to 8 atoms.

Table 2. Cell lots used in hiPSC neuronal MEA assay.

MEA replicate	GlutaNeuron lot number	Astrocyte lot number	Final cell density per well
1	107487 (81% glutamatergic neurons/ 19% GABAergic neurons)	107580	97,200 glutamatergic neurons, 22,800 GABAergic neurons and 20,000 astrocytes
2	108288 (84% glutamatergic neurons/ 16% GABAergic neurons)	107638	100,800 glutamatergic neurons, 19,200 GABAergic neurons and 20,000 astrocytes
3	107729 (83% glutamatergic neurons/ 17% GABAergic neurons)	107580	99,600 glutamatergic neurons, 20,400 GABAergic neurons and 20,000 astrocytes

iCell GlutaNeuron and iCell Astrocyte lot numbers used for 3 independent biological replicates of the MEA analysis experiment. Final seeding densities of each cell type, per well, is also provided.

compound was added. The difference in current between the pre- and post-compound measurements was exported and analyzed using GraphPad Prism software (version 10.6.1; GraphPad Software, Boston, MA, United States). A minimum of 6- ($K_v2.1$, $Na_v1.2$), 5- ($\alpha_1\beta_2\gamma_2$ GABA_A receptor), or 4-point ($\alpha_4\beta_2$ nicotinic receptor) concentration-response curves were generated for each compound to generate IC₅₀ (half-maximal inhibitory concentration) values. IC₅₀ values were obtained from a 4-parameter logistic fit of normalized concentration-response data. Data were transformed to the pIC₅₀ (the negative logarithmic measure of the IC₅₀), a value often used in pharmacology to easily compare compound potency, where a higher value indicates greater potency.

Microelectrode array

iCell GlutaNeurons (cat. R1034, FUJIFILM Cellular Dynamics Inc. [CDI], Maddison, WI, United States) and iCell astrocytes (cat. R1092, FUJIFILM CDI) were mixed and seeded into 48-well MEA plates (16 electrodes/well; Axion Biosystems, Atlanta, GA, United States) precoated with 0.07% v/v Polyethylenimine (cat. 181978, Sigma-Aldrich) at a final ratio of 8:3 (GlutaNeurons: Astrocytes), following manufacturer's protocols. The lot numbers and final cell densities per well for each biological replicate are provided in Table 2. Cells were cultured in complete BrainPhys™ media: BrainPhys™ media (cat. 05790, Stem Cell Technologies, Vancouver, Canada) supplemented with 1% v/v iCell Nervous System Supplement (cat. M1031, Fujifilm CDI), 2% v/v iCell Neural Supplement B (cat. M1029, Fujifilm CDI), 1% v/v N2 supplement (cat. 17502048, ThermoFisher Scientific), 1% v/v penicillin-streptomycin, and 1 µg/mL laminin (cat. 23017-015, Gibco) with 50% media changes performed every 24 to 72 h. Spontaneous electrical activity was assessed weekly during cell maturation on day in vitro (DIV) 5, 12, and 19. Compounds were added on—DIV17-30.

Stock compounds were solubilized at 1000× their top concentration in DMSO. Serial dilutions in complete BrainPhys™ media were prepared at 10x final concentration which were added

directly to media in the wells. A 0.1% DMSO vehicle control was included on every plate.

Before compound addition, the MEA plate was equilibrated in the MEA Pro instrument (Axion Biosystems) for at least 15 min. Then, 15-min baseline recordings were taken, and 15-min treatment recordings were taken after 1 h incubation with test compounds. The sampling frequency was 12.5 kHz. The digital low-pass filter was 3 kHz and the high-pass filter was 200 Hz.

Spike files were created from the original 15-min recordings for analysis. Raster plots of the first 150s of the recordings were created using the Axion Neural Metric Tool (v4.1.1), and the effects of compound treatment on selected MEA parameters were determined using the Axion Axis Metric Plotting Tool (v2.5.1).

The effect of each drug was measured relative to its own pre-drug baseline (from the same well). Compounds ($n=4$) or vehicle ($n=6$ to 12) were applied directly to wells, and the average change from baseline was calculated. The DMSO vehicle control is presented for each parameter and usually shows minimal changes ($\pm 20\%$).

A 1-way ANOVA with Šidák's multiple comparisons test was performed on parent and derivative compound data (Graphpad Prism software v10.6.1); $P < 0.05$ was considered significant. Many different MEA parameters can be used to describe electrical activity, such as weighted mean firing rate (WMFR), interspike interval (ISI) coefficient of variation, area under normalized cross correlation (AUNCC), network burst frequency, network burst duration, and number of spikes per network burst (see Table 3 for parameter definitions). Similar parameters have been selected in analyses by other groups (Kreir et al. 2018; Tukker et al. 2020a, 2020b).

Pharmacophore analysis

Using the 2-sided Mann-Whitney U-test as implemented in Scipy (version 1.16.3), P-values were calculated using the Bonferroni Correction (BC) to account for multiple comparisons. Strict

Table 3. Definitions of selected MEA parameters, as defined by Axion Biosystems.

Firing rate parameters	Definition
Weighted mean firing rate (Hz)	The mean firing rate based on only electrodes with activity greater than minimum spike rate.
ISI coefficient of variation	The coefficient of variation (SD/mean) of the interspike interval, the time between spikes. This is a measure of spike regularity.
Network burst parameters	
Network burst frequency (Hz)	Total number of network bursts divided by the recording time.
Network burst duration (s)	Average time from the first spike to the last spike in a network burst.
Number of spikes per network burst	Average number of spikes in a network burst.
Synchrony parameter	
Area under normalized cross-correlation	Area under well-wide pooled interelectrode cross-correlation normalized to the auto-correlations. High values indicate greater synchronicity of the network.

criteria were applied to ensure that pharmacophore features identified as significant were those of the greatest risk and not statistical artifacts. First, a feature was only considered significant if the Mann–Whitney *U*-test *P*-value was less than $0.05/207 = 0.00024$ (BC corrected 95% threshold) when comparing the shift in average pIC_{50} value between the group of compounds with or without the pharmacophore feature. Theoretically, there are $(N[N-1]) \times 10/2 = 360$ pharmacophore features within the compound set tested, where $N = 8$ (i.e. the number of pharmacophore features), 10 is the number of distances considered between features and the division by 2 is due to the symmetry of the dyad pairs. Here, only 207 pharmacophore features were actually observed in the molecules screened, hence why this was used for the BC. Secondly, there needed to be a minimum difference of 0.3 in the pIC_{50} value between the mean of compounds with and without a specific pharmacophore feature. This delta pIC_{50} (ΔpIC_{50}) threshold of 0.3 corresponds to a 2-fold change in IC_{50} values and is commonly used within QSAR modelling as a threshold to define activity cliffs and account for assay variability (Cramer 2011; Shen et al. 2024). Finally, 8 or more examples of compounds with or without a given feature were required in the dataset tested. Given the small size of the dataset, this value was chosen to increase the likelihood that any statistically significant data were not random. For example, when conducting a binomial test, the probability of observing 8/8 positive measurements at random (based on a ΔpIC_{50} of 0.3) = 0.007812, which is below the $P < 0.05$ threshold required to reject a null hypothesis.

Graphical representation of structural and spatial distribution of compounds

To visualize the distribution of IC_{50} values across the compound set, each parent molecule and the associated structural relatives from the Enamine REAL set were extracted and grouped by IC_{50} values at each of the 4 ion channels screened. A boxplot was constructed to represent the distribution of IC_{50} values for each group. The parent molecule's IC_{50} value was highlighted in blue and each associated molecule's IC_{50} was overlayed as a black dot. Visualizations of the structures of the Enamine REAL molecules relative to their respective parent molecules were generated using Networkx (v3.5). These images show the parent molecule in the center and each associated molecule around the outside connected to the parent with a molecular similarity (ECFP4 [2048 bits] and Tanimoto similarity) to summarize the degree of similarity between the two. The IC_{50} values and the difference (ΔIC_{50}) compared with the parent are quoted for each associated molecule. Molecule images were all rendered using the RDKit (v2023.9.6) ("RDKit: Open-Source Cheminformatics Software," n.d.).

Results

Ion channel potency

The potencies of the Enamine REAL compounds compared with the parent compounds are presented as representative graphs with IC_{50} values in Fig. 1 and summarized in Table 4. All raw IC_{50} values and pharmacophore feature labels for each compound are provided in Supplementary Data. At the $K_V2.1$ ion channel, the potency of the parent compounds ranged from <2 to $>100 \mu M$. Eight of the 10 parent compounds were highly potent ($IC_{50} < 30 \mu M$), with the exception of quetiapine and 4-AP. Against this background, 36 compounds exhibited similar potencies compared with the respective parent molecules, and 32 and 10 compounds had reduced and increased potencies, respectively. Notably, all the 4-AP derivative compounds remained inactive, in line with the parent molecule. Further, all the derivative compounds for NS1619, quinidine and verapamil had reduced potencies compared with their parent molecules, whereas all the XE991 derivative compounds were more potent than the parent (Table 4).

For the $Na_V1.2$ ion channel, the potency of the parent compounds ranged from <10 to $>100 \mu M$. Amoxapine and bepridil parent compounds were the most potent, with 6 of the remaining parent compounds inactive ($IC_{50} > 100 \mu M$). Against this background, most of the compounds (45) had equivalent potency to the parent molecules, with 25 and 8 compounds with reduced or increased potency, respectively. Notably, all the quinidine derivative compounds remained inactive, in line with the parent molecule. In the case of bepridil and verapamil, all derivative compounds were less potent than their respective parent molecules (Table 4).

At the $\alpha_1\beta_2\gamma_2$ GABA_A receptor, the potency of the parent compounds ranged from <3 to $>100 \mu M$. The amoxapine parent compound was highly potent ($IC_{50} = 2.6 \mu M$), with all remaining parent compounds inactive ($IC_{50} > 100 \mu M$). Against this background, most of the derivative compounds (63) had equivalent potency to the parent molecules, with 8 and 7 compounds with reduced or increased potency, respectively (Table 4).

For the $\alpha_4\beta_2$ nicotinic receptor, the potency of the parent compounds ranged from <3 to $>40 \mu M$. All 10 parent compounds were active, with at least moderate potency detected ($IC_{50} < 45 \mu M$). Against this background, 37 compounds had equivalent potency to the parent molecules, and 36 and 5 compounds had reduced or increased potencies, respectively (Table 4).

Data suitability for SAR

The distribution of the 78 derivative compound IC_{50} values compared with the 10 parent molecules for each ion channel are shown in Fig. 2. These box plots illustrated variability in the data

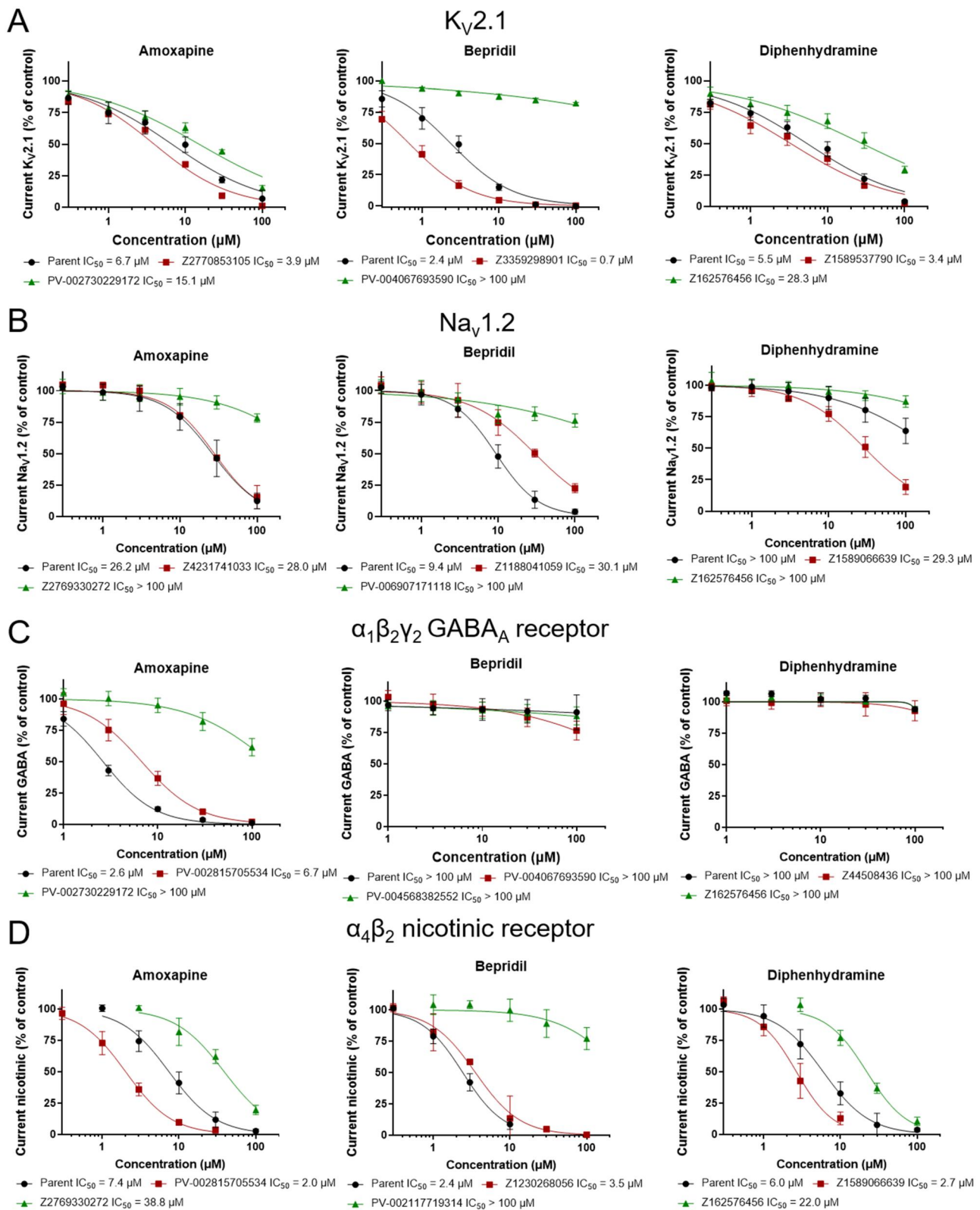


Fig. 1. Distribution of potencies of selected parent and derivative compounds across 4 seizure-related ion channels. Representative IC_{50} plots for amoxapine, bepridil, and diphenhydramine parent compounds (black circles) compared with derivative Enamine REAL compounds with the highest (red squares) and lowest (green triangles) potencies screened against seizure-related A. $K_{V2.1}$, B. $\text{Na}_{V1.2}$, C. $\alpha_1\beta_2\gamma_2$ GABA_A and D. $\alpha_4\beta_2$ nicotinic ion channels. Where Enamine REAL compounds with higher or lower potency were not identified during screening, alternative representative compounds have been shown. Data shown as mean $\pm$ SD from a minimum $n = 3$ cells.

Table 4. Potency of Enamine REAL compounds compared with 10 parent compounds screened against CNS-related ion channels K_v2.1, Na_v1.2 and the $\alpha_1\beta_2\gamma_2$ GABA_A and $\alpha_4\beta_2$ nicotinic receptors.

Ion channel	Parent compound	4-AP	Amoxapine	Bepridil	Diphenhydramine	Linopirdine	NS1619	Quetiapine	Quinidine	Verapamil	XE991
K _v 2.1	Parent compound IC ₅₀	>100 μ M	6.7 μ M	2.4 μ M	5.5 μ M	8.8 μ M	7.6 μ M	31.4 μ M	1.2 μ M	1.5 μ M	19.2 μ M
	IC ₅₀ <10 μ M	0	7	7	7	3	0	1	0	0	2
	IC ₅₀ 10–30 μ M	0	3	0	2	2	1	7	0	0	0
	IC ₅₀ 30–100 μ M	0	0	1	0	1	1	2	2	1	0
	IC ₅₀ >100 μ M	10	0	1	0	2	7	0	7	1	0
Na _v 1.2	Parent compound IC ₅₀	>100 μ M	26.2 μ M	9.4 μ M	>100 μ M	>100 μ M	56.9 μ M	>100 μ M	>100 μ M	35.6 μ M	>100 μ M
	IC ₅₀ <10 μ M	0	0	0	0	0	0	0	0	0	0
	IC ₅₀ 10–30 μ M	0	1	0	1	0	0	0	0	0	0
	IC ₅₀ 30–100 μ M	1	4	3	2	1	4	2	0	0	1
	IC ₅₀ >100 μ M	9	5	6	6	7	5	8	9	2	1
$\alpha_1\beta_2\gamma_2$ GABA _A	Parent compound IC ₅₀	>100 μ M	2.6 μ M	>100 μ M	>100 μ M	>100 μ M	>100 μ M	>100 μ M	>100 μ M	>100 μ M	>100 μ M
	IC ₅₀ <10 μ M	0	2	0	0	0	0	0	0	0	0
	IC ₅₀ 10–30 μ M	0	2	0	0	2	0	0	1	0	0
	IC ₅₀ 30–100 μ M	0	1	0	0	0	0	3	1	0	0
	IC ₅₀ >100 μ M	10	5	9	9	6	9	7	7	2	2
$\alpha_4\beta_2$ nicotinic	Parent compound IC ₅₀	21.0 μ M	7.4 μ M	2.4 μ M	6.0 μ M	43.6 μ M	25.9 μ M	26.9 μ M	12.4 μ M	13.4 μ M	19.3 μ M
	IC ₅₀ <10 μ M	0	6	3	6	1	0	0	1	0	0
	IC ₅₀ 10–30 μ M	1	3	3	3	3	1	9	4	1	2
	IC ₅₀ 30–100 μ M	2	1	1	0	4	3	1	3	1	0
	IC ₅₀ >100 μ M	7	0	2	0	0	5	0	1	0	0

The number of Enamine REAL compounds with IC₅₀ values of <10 μ M, 10–30 μ M, 30–100 μ M and >100 μ M relative to the respective parent compounds are included for each channel. Graphs showing IC₅₀ values for the parent, and representative Enamine REAL derivative, compounds amoxapine, bepridil and diphenhydramine can be seen in Fig. 1 4-AP, 4-aminopyridine.

between ion channels, with K_v2.1 and the $\alpha_4\beta_2$ nicotinic receptor offering a complex and intriguing dataset. In comparison, at Na_v1.2 and the $\alpha_1\beta_2\gamma_2$ GABA_A receptor, the parent compounds themselves were often inactive and structural modifications usually resulted in no change, or loss of any potency.

Figures S2–S5 show the structure and spatial distribution of IC₅₀ values for the Enamine REAL database compounds relative to parent molecules at K_v2.1, Na_v1.2, and the $\alpha_1\beta_2\gamma_2$ GABA_A and nicotinic- $\alpha_4\beta_2$ receptors, respectively.

Figure S6 shows the relative distribution of all 88 molecules in 2D using the t-SNE method. For K_v2.1 and the $\alpha_4\beta_2$ nicotinic receptor, there were some initial clusters of molecules which were consistent with IC₅₀ classes (Fig. S6A and D). In contrast, for Na_v1.2 and the $\alpha_1\beta_2\gamma_2$ GABA_A receptor, most molecules fell into a single class hence little could be gained from these plots (Fig. S6B and C). A larger data set is required to explore this behaviour in more detail and apply similar methods at the other ion channels considered. Figure S7 shows boxplots of Tanimoto similarity between all molecules in the dataset. This provides a standard chemical similarity metric upon which to compare the molecules and can help to interpret the chemical space covered by the data. The mean and median similarities are typically low for all molecules as would be expected. Further, there are typically several similar molecules, this is not unexpected based on the parent molecule selection process. Given the dataset was relatively small, care was taken not to overinterpret these plots. However, small initial clusters emerging for 2 of the 4 ion channels is suggestive of regions of chemical spaces which are likely to exhibit higher or lower IC₅₀ values at those targets. The range in IC₅₀ values and spatial distribution across compounds at K_v2.1 and the $\alpha_4\beta_2$ nicotinic receptor presented a valuable dataset for SAR, featuring increasing, equivalent or decreasing potency with the structural modifications conducted.

Significant pharmacophore features

Tables 5 and 6 show significant pharmacophore features, each of which is defined as a 2D dyad pattern using proprietary SMARTS notations (“Daylight Theory: SMARTS—A Language for Describing Molecular Patterns,” n.d.), at K_v2.1 and the $\alpha_4\beta_2$ nicotinic receptor, respectively. There were 4 significant features at K_v2.1: A hydrogen bond donor directly connected to an aromatic ring (HD_directlink_AR, Δ pIC₅₀ = −0.77 [decreased potency]); a hydrogen bond donor separated from a hydrogen bond acceptor by a single atom (HD_1atom_HA, Δ pIC₅₀ = −0.64 [decreased potency]); a hydrogen bond acceptor separated from a basic group by 2 atoms (HA_2atom_BG, Δ pIC₅₀ = 0.52 [increased potency]) and an atom which is part of an aromatic system separated from a basic group by 4 atoms (AR_4atom_BG, Δ pIC₅₀ = 0.51 [increased potency]) (Table 5). These 4 features were also significant at the $\alpha_4\beta_2$ nicotinic receptor (HD_directlink_AR, Δ pIC₅₀ = −0.53 [decreased potency]; HD_1atom_HA, Δ pIC₅₀ = −0.54 [decreased potency]; HA_2atom_BG, Δ pIC₅₀ = 0.56 [increased potency]; AR_4atom_BG, Δ pIC₅₀ = 0.57 [increased potency]) and conferred the same direction of potency (Table 6). Eight additional pharmacophore features achieved statistical significance, with 7 of these associated with an increased potency and 1 associated with a decreased potency: A hydrogen bond donor separated from a halogen by 3 atoms (HD_3atom_Hal, Δ pIC₅₀ = −0.56 [decreased potency]) (Table 6). In contrast, although there were some nominally differential features identified at the $\alpha_1\beta_2\gamma_2$ GABA_A receptor and Na_v1.2, none achieved statistical significance (Tables S3 and S4). Interestingly, an atom which is part of an aromatic system separated from a basic group by 4 atoms was also a differential feature at Na_v1.2, showing increased potency as seen for K_v2.1 and the $\alpha_4\beta_2$ nicotinic receptor (Table S4). The full lists of the impact of differentially present pharmacophore features for all ion channels are shown in Tables S1–S4.

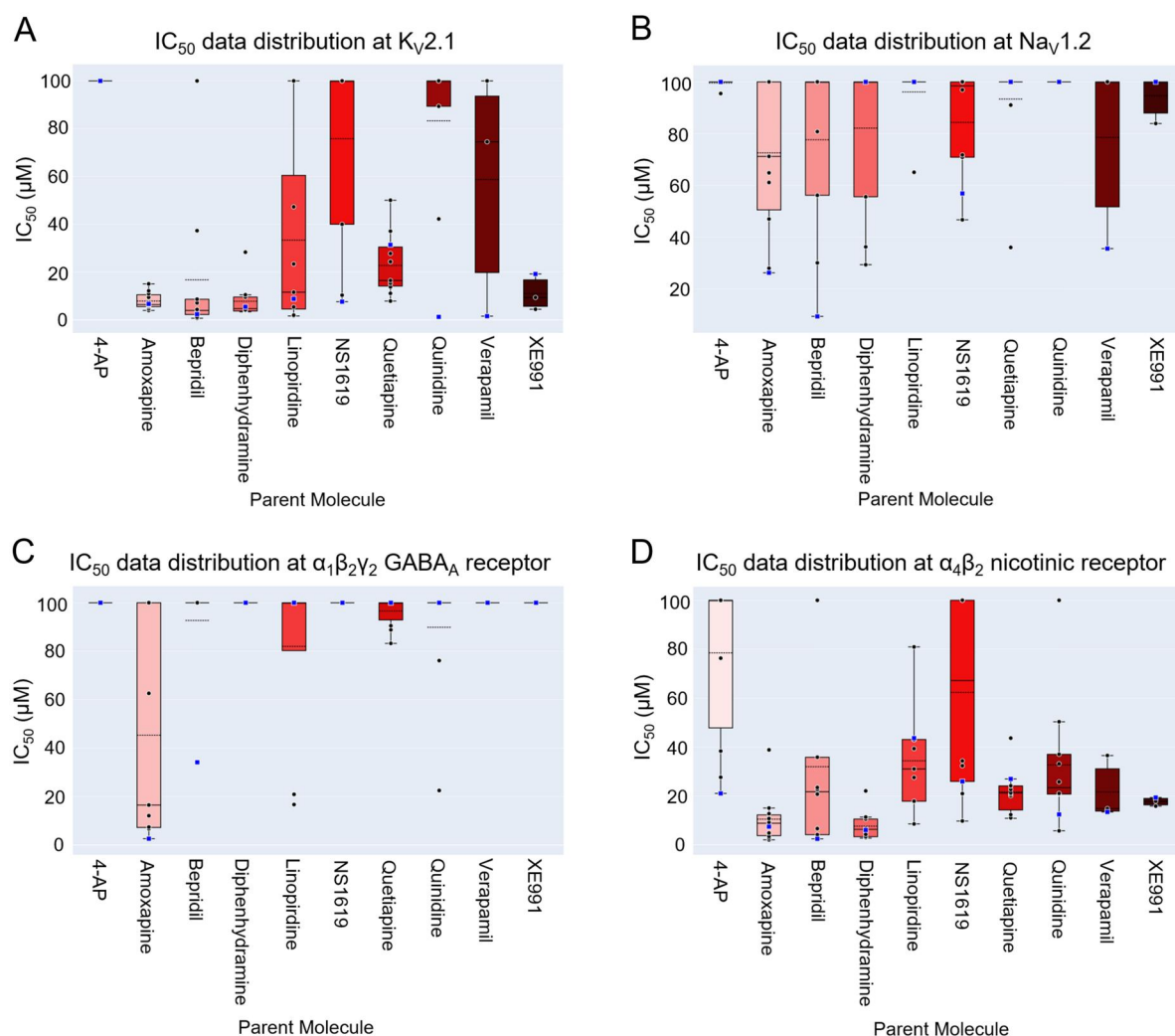


Fig. 2. Distribution of potencies of parent and derivative compounds across 4 seizure-related ion channels. Distribution of IC_{50} values for Enamine REAL database compounds (black circles) relative to parent compounds (blue squares), separated by parent compound at seizure-related ion channels A. $K_v2.1$, B. $Na_v1.2$, C. $\alpha_1\beta_2\gamma_2$ $GABA_A$ receptor, and D. $\alpha_4\beta_2$ nicotinic receptor. Mean values are represented by a dashed line, and median values are represented by a solid line. 4-AP, 4-aminopyridine.

Table 5. Statistically significant pharmacophore features at $K_v2.1$.

Pharmacophore feature	Mean pIC_{50} with feature	With feature (N =)	Mean pIC_{50} without feature	Without feature (N =)	ΔpIC_{50}	P-value
HD_directlink_AR	4.127602	21	4.899596	67	-0.77199	4.77E-07
HD_1atom_HA	4.222548	20	4.860318	68	-0.63777	6.86E-05
HA_2atom_BG	5.020858	36	4.503878	52	0.51698	8.49E-05
AR_4atom_BG	5.019541	35	4.514502	53	0.505039	0.000125699

Data analyzed by 2-sided Mann-Whitney U-test with Bonferroni Correction; $P < 0.0002415$ considered significant ($P = 0.05/207$ [number of tests performed]). Positive (red text) and negative (green text) ΔpIC_{50} (ΔpIC_{50}) values represent increased and decreased potency, respectively, for compounds with the given pharmacophore feature. **Bold shows pharmacophore features common across both $K_v2.1$ and the $\alpha_4\beta_2$ nicotinic receptor.** Definitions and abbreviations for pharmacophores are listed in Table 1. The full dataset can be seen in Table S1.

Figure 3 shows box plots of the impact of the 4 common pharmacophore features on pIC_{50} values at $K_v2.1$ (Fig. 3A) and the $\alpha_4\beta_2$ nicotinic receptor (Fig. 3B). For both $K_v2.1$ and the $\alpha_4\beta_2$ nicotinic receptor, the presence of either a hydrogen bond donor directly connected to an aromatic ring (panel 1) or hydrogen bond donor separated from a hydrogen bond acceptor by a single atom (panel 2) resulted in lower pIC_{50} values, equating to lower potency. In contrast, the presence of a hydrogen bond acceptor separated from a basic group by 2 atoms (panel 3) or an atom which is part of an aromatic system separated from a basic group

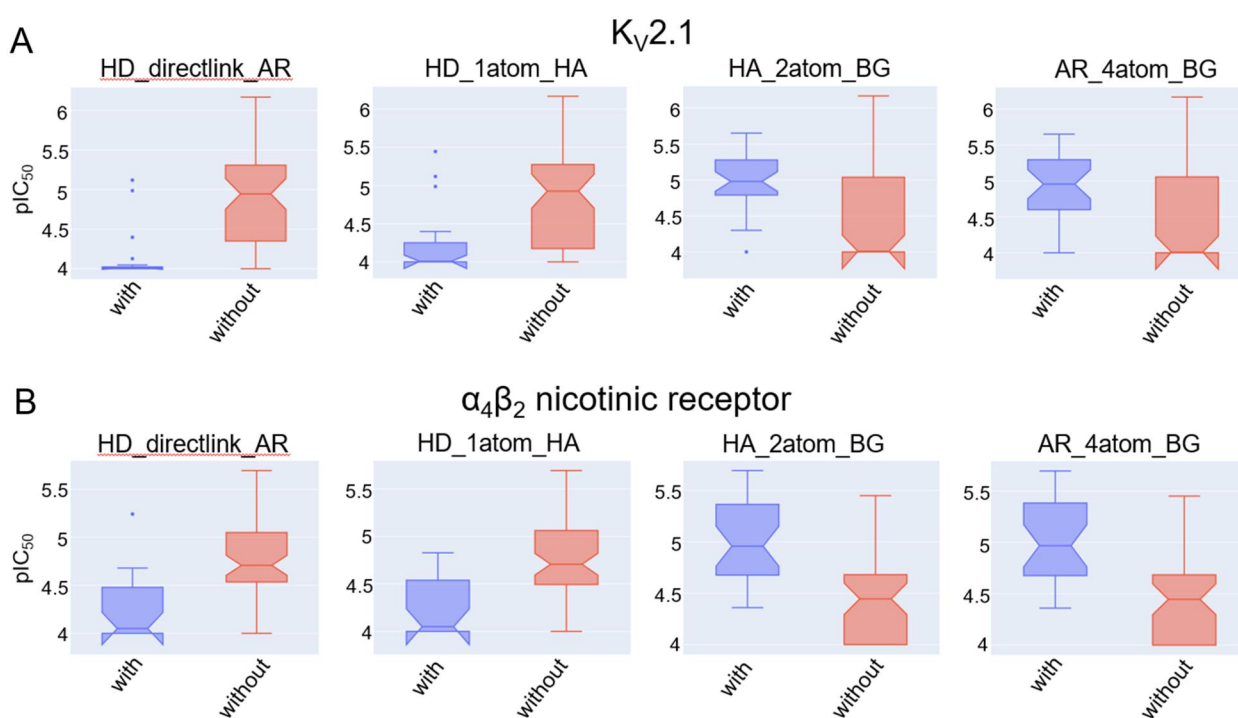
by 4 atoms (panel 4) resulted in higher pIC_{50} values, equating to higher potency. Generally, the pIC_{50} values were more variable at $K_v2.1$ with these 4 features compared with without (Fig. 3A). In contrast, the pIC_{50} values were equally variable at the $\alpha_4\beta_2$ nicotinic receptor with or without these 4 features (Fig. 3B).

An interesting question followed from this analysis: “do any of the significant pharmacophore features also impact other safety targets?”. We therefore additionally screened the 2 hERG datasets retrieved from ChEMBL (version 36). The molecules identified had measured K_i and/or EC_{50} . The structures were cleaned by

Table 6. Statistically significant pharmacophore features at the $\alpha_4\beta_2$ nicotinic receptor.

Pharmacophore feature	Mean pIC ₅₀ with feature	With feature (N =)	Mean pIC ₅₀ without feature	Without feature (N =)	Δ pIC ₅₀	P-value
HA_samegroup_BG	4.89039	56	4.304497	32	0.585892	7.54E-10
AR_4atom_BG	5.022346	35	4.449502	53	0.572843	3.16E-09
HA_2atom_BG	5.010487	36	4.446696	52	0.56379	5.98E-09
HD_directlink_AR	4.274813	21	4.803502	67	-0.52869	1.15E-06
AR_5atom_BG	4.947246	36	4.490479	52	0.456767	2.02E-06
HA_2atom_HA	4.909389	44	4.445286	44	0.464103	2.23E-06
HD_1atom_HA	4.262416	20	4.799374	68	-0.53696	2.31E-06
AR_1atom_BG	4.932171	37	4.492459	51	0.439712	6.96E-06
HD_samegroup_BG	4.964284	31	4.521279	57	0.443005	2.43E-05
BG_2atom_BG	5.193379	11	4.603618	77	0.589761	3.67E-05
HD_3atom_Hal	4.193404	12	4.753749	76	-0.56035	3.68E-05
AR_2atom_BG	4.882711	37	4.528342	51	0.354369	0.000202

Data analyzed by 2-sided Mann-Whitney U-test with Bonferroni Correction; $P < 0.0002415$ considered significant ($P = 0.05/207$ [number of tests performed]). Positive (red text) and negative (green text) Δ pIC₅₀ (Δ pIC₅₀) values represent increased and decreased potency, respectively, for compounds with the given pharmacophore feature. **Bold shows pharmacophore features common across both K_v2.1 and the $\alpha_4\beta_2$ nicotinic receptor.** Definitions and abbreviations for pharmacophores are listed in Table 1. The full dataset can be seen in Table 2.

**Fig. 3.** Impact of significant pharmacophore features at K_v2.1 and the $\alpha_4\beta_2$ nicotinic receptor. pIC₅₀ values are shown for 88 compounds screened against A. K_v2.1 or B. the $\alpha_4\beta_2$ nicotinic receptor with or without the specified pharmacophore feature (see Tables 5 and 6 for details).

desalting, normalizing protonation states and standardizing tautomers. The K_i and EC_{50} values were converted to pK_i and pEC_{50} as $-\log_{10}(x)$ where x is the K_i or EC_{50} value. All statistical tests were applied to the transformed pK_i and pEC_{50} values. These datasets were screened using the same featurization where pK_i encountered 306 pharmacophore features and the pEC_{50} encountered 313 pharmacophore features out of the 360 possible features. The BC correction was applied based upon the number of encountered pharmacophore features. As with seizure, we split the datasets based on the molecules containing, or not, each pharmacophore feature independently and tested for a statistically significant shift in the average pK_i or pEC_{50} . We found that none of the 4 statistically significant pharmacophore features identified for seizure were identified for hERG. This suggested these identified features for seizure showed at least some specificity to those targets. The hERG datasets are provided in

Supplementary Data. Figure S8 shows the pharmacophore features which were statistically significant at hERG for both pK_i and pEC_{50} .

MEA analysis

MEA analysis of amoxapine, diphenhydramine, quetiapine, and selected Enamine REAL derivative compounds was conducted to investigate if the modified potency at ion channels translated into changes in seizurogenic phenotypes at a whole-cell level. Amoxapine showed a clear seizurogenic phenotype compared with baseline activity at 3 μ M (Fig. 4A, panel 1 raster plots), with a high potency at K_v2.1, the $\alpha_4\beta_2$ nicotinic and $\alpha_1\beta_2\gamma_2$ GABA_A receptors, and moderate potency at Na_v1.2 (Fig. 4A). The derivative molecule Z4456062749 showed a less potent seizurogenic phenotype (Fig. 4A, panel 2 raster plots) and was associated with lower potency at Na_v1.2 and the $\alpha_1\beta_2\gamma_2$ GABA_A receptor (Fig. 4A).

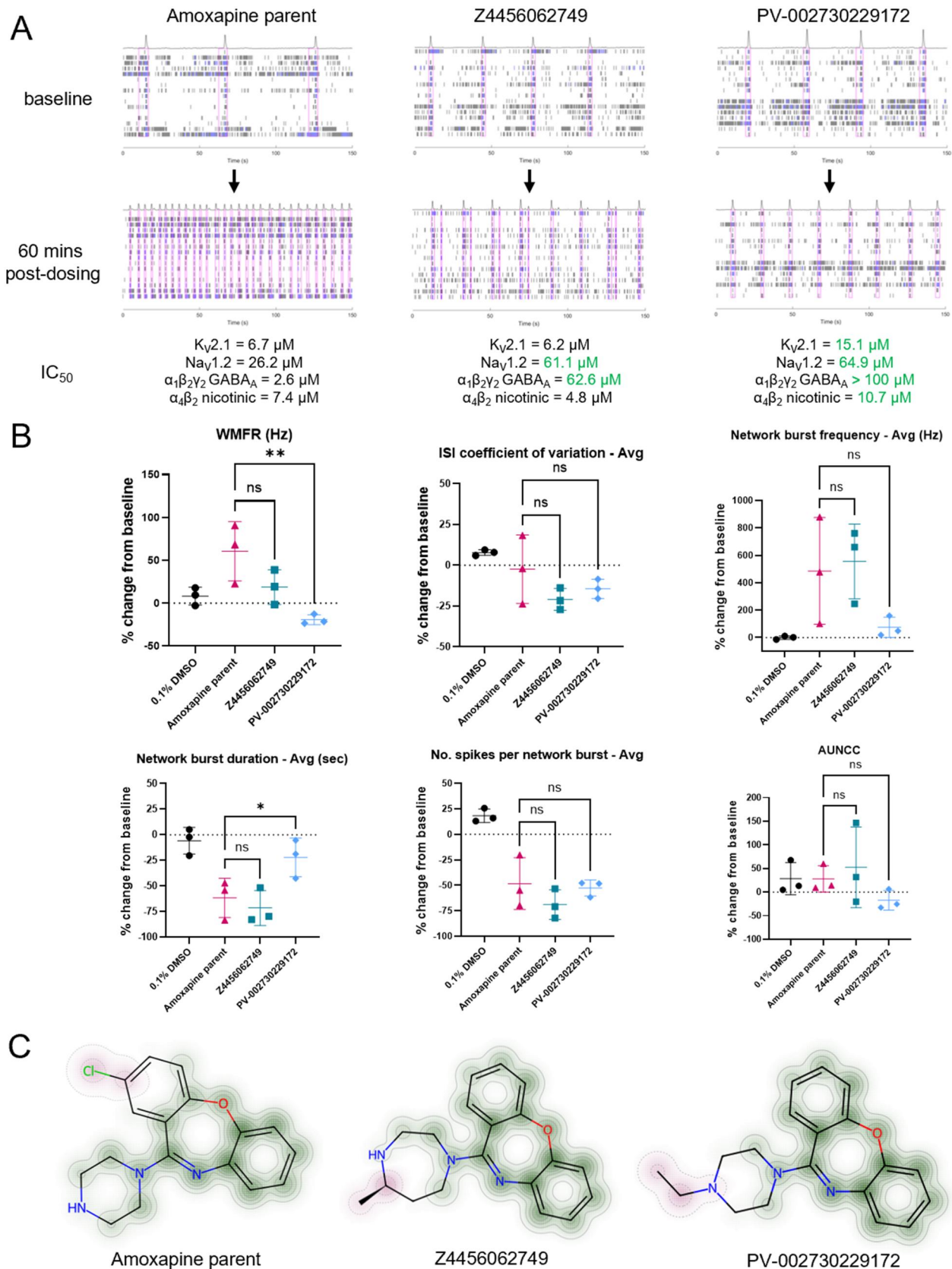


Fig. 4. MEA analysis of amoxapine and derivatives at $3 \mu\text{M}$. A. Representative raster plots depicting 0 to 150 s activity at baseline (top row) and 60 min postdose (bottom row) with parent or derivative Enamine REAL compounds, Z4456062749 and PV-002730229172. IC₅₀ values at 4 ion channels are shown below raster plots (black text = similar potency and green text = lower potency of Enamine REAL compound compared with parent at the given ion channel). B. Key MEA parameters highlighting significant differences. Data shown as mean $\pm$ SD from $n = 3$ biological replicates. Data analyzed by 1-way ANOVA with Šidák's multiple comparisons test; * $P < 0.05$, ** $P < 0.01$. C. Chemical structures of compounds show areas of structural similarity (green) and differences (purple). AUNCC, area under normalized cross-correlation; ISI, interspike interval; WMFR, weighted mean firing rate.

However, this was not associated with statistically significant differences in MEA parameters compared with the parent compound (Fig. 4B). The derivative PV-002730229172 showed an even less potent phenotype (Fig. 4A, panel 3 raster plots) quantified by statistically significant differences in several MEA parameters (Fig. 4B). For example, PV-002730229172 exhibited significantly reduced WMFR ($P=0.0032$) with an increased network burst duration ($P=0.0457$). This was associated with considerably lower potencies at all 4 ion channels (Fig. 4A). The 3 compounds (Fig. 4C) share many structural similarities, with the loss of the halogen group (chlorine; Cl) from the amoxapine parent an obvious change as seizure risk is reduced from left to right. Noticeably, shifting the piperazine NH group position, through ring substitution to a diazepane (in compound Z4456062749), or replacing a hydrogen with a bulkier and flexible ethyl chain (in compound PV-002730229172) could alter the accessibility of an otherwise relatively rigid group, which may impact binding at the ion channels screened.

Diphenhydramine showed a seizurogenic phenotype compared with baseline activity at $10\mu\text{M}$, and was associated with potent inhibition of $K_v2.1$ and the $\alpha_4\beta_2$ nicotinic receptor (Fig. 5A, panel 1). The derivative Z162576456 showed a reduced seizure phenotype in the MEA raster plots and had a similar ion channel profile with slightly reduced potency at $K_v2.1$ and the $\alpha_4\beta_2$ nicotinic receptor (Fig. 5A, panel 2). This difference was quantified by a statistically significant increase in ISI coefficient of variation ($P=0.0168$) and number of spikes per network burst ($P=0.0011$, Fig. 5B), in line with vehicle control values. In contrast, the derivative molecule Z1589066639 showed a clear seizure phenotype, resulting in a complete dysregulation of neuronal activity, associated with higher potency at $K_v2.1$, $\text{Na}_v1.2$ and the $\alpha_4\beta_2$ nicotinic receptor (Fig. 5A, panel 3). This was reflected in significant decreases in several MEA parameters, including weighted mean firing rate ($P=0.0093$), ISI coefficient of variation ($P=0.0199$), network burst duration ($P=0.0343$), number of spikes per network burst ($P=0.0159$), and area under normalized cross correlation ($P=0.0115$) compared with the parent compound (Fig. 5B). The chemical structures of the 3 compounds (Fig. 5C) show several areas of similarity and differences between Z162576456 and Z1589066639 compared with diphenhydramine. Replacing the benzene ring with a pyridine ring and replacing the ether (O) group with a secondary amine (NH) functionality from the parent molecule to Z162576456 may offer some reduction in seizure risk. Noticeably, addition of a halobenzyl ring and replacing the ether (O) group with a secondary amine in Z1589066639 could confer increased seizure risk. Clearly, the ether to amine occurs in both with opposing seizure risk outcomes. As some of these structural modifications are not independent, i.e. not the sole modification between 2 of the structures, we cannot assign the seizurogenic change to 1 of these modifications, as there are potentially joint effects between the modifications. Further work is needed to establish these links.

Quetiapine showed a mild seizurogenic phenotype compared with baseline activity at $10\mu\text{M}$, and showed moderate inhibition of $K_v2.1$ and the $\alpha_4\beta_2$ nicotinic receptor (Fig. 6A, panel 1). The derivative Z3649159890 showed a slightly less seizurogenic phenotype and slightly reduced potency at the $\alpha_4\beta_2$ nicotinic receptor (Fig. 6A, panel 2). However, this was not associated with statistically significant differences in MEA parameters compared with the parent compound (Fig. 6B). In contrast, the derivative molecule PV-004814057247 showed a potent seizurogenic phenotype, associated with a higher potency at all 4 ion channels (Fig. 6A, panel 3). There were significant changes in several MEA

parameters including an increased network burst frequency ($P=0.0014$), and decreased ISI coefficient of variation ($P=0.0398$), network burst duration ($P=0.0011$), and number of spikes per burst ($P=0.0207$) compared with the parent compound (Fig. 6B). The chemical structures of the 3 compounds (Fig. 6C) revealed significant areas of difference; the major structural differences occurred on the substituents of the piperazine ring. The parent molecule possessed an extended bifunctional chain containing an ether and primary alcohol. The 2 sampled molecules, Z3649159890 and PV-004814057247, respectively possessed a cyclized version of the parents chain based on a tetrahydrofuran ring and a bifurcated monofunctional chain which retains the ether functionality of the parent molecule. Additionally, both of the sampled molecules substituted the main Dibenzo[B, F][1,4]thiazepine framework for the Dibenzo[B, F](1,4)Oxazepine framework removing the sulfur and replacing with an oxygen. The 2 sampled molecules showed divergent seizurogenic potential suggesting the framework modification leads to minimal impact. The pyrazine substituent modifications which differ between the 2 sampled molecules are more likely to be the source of this divergent behaviour, with the bulkier ring system (Z3649159890) modification raising the IC_{50} at 2 of the ion channels and the bifurcated chain (PV-004814057247) reducing the IC_{50} at all of the ion channels screened.

Discussion

CNS liability, specifically drug-induced seizures, is a leading cause of project delays and failures during drug discovery and development. CNS toxicity is associated with compounds spanning therapy areas including CNS, cardiovascular, gastrointestinal, respiratory, and inflammation (Easter et al. 2009; Cook et al. 2014; Walker et al. 2018). Typically, CNS liabilities are not assessed until late-stage preclinical in vivo studies, often going undetected until clinical phase studies, after significant investment of time and money. We have previously shown that a panel of CNS ion channels can be used to screen for toxicity due to seizure liability in lead generation/lead optimization and can provide mechanistic insight where seizures have been observed in GLP toxicology studies or in clinical development (Rockley et al. 2023).

Here, to assist in the development of therapies without liability, we aimed to develop SAR by identifying the specific structural characteristics of compounds associated with ion channel-related seizurogenic activity. As a proof-of-concept study, we initially selected 4 ion channels (from the 15 ion channels in the $i\text{LA}^{\text{Seizure}}$ panel) that would be representative of the key ion channel families associated with seizure risk. Compounds were selected from the Enamine REAL database using pharmacophore features to assess similarity to reference compounds. The final set consisted of 10 parent and 78 derivative compounds.

We initially assessed ion channel activity by automated patch-clamp as this is a relatively high-throughput assay and thus more practical than MEA for the >350 datapoints that were to be evaluated. The pharmacophore analysis of the $\alpha_4\beta_2$ nicotinic receptor provided the richest dataset, with 63 differential features detected (where pIC_{50} values differ with the pharmacophore feature vs without), 12 of which reached statistical significance. Within this dataset, several features conferred decreased potency but most of the significant features increased potency, which was unexpected.

The pharmacophore analysis of $K_v2.1$ yielded 37 differential features, 4 of which achieved statistical significance. Notably, these 4 features were also significantly different at the $\alpha_4\beta_2$

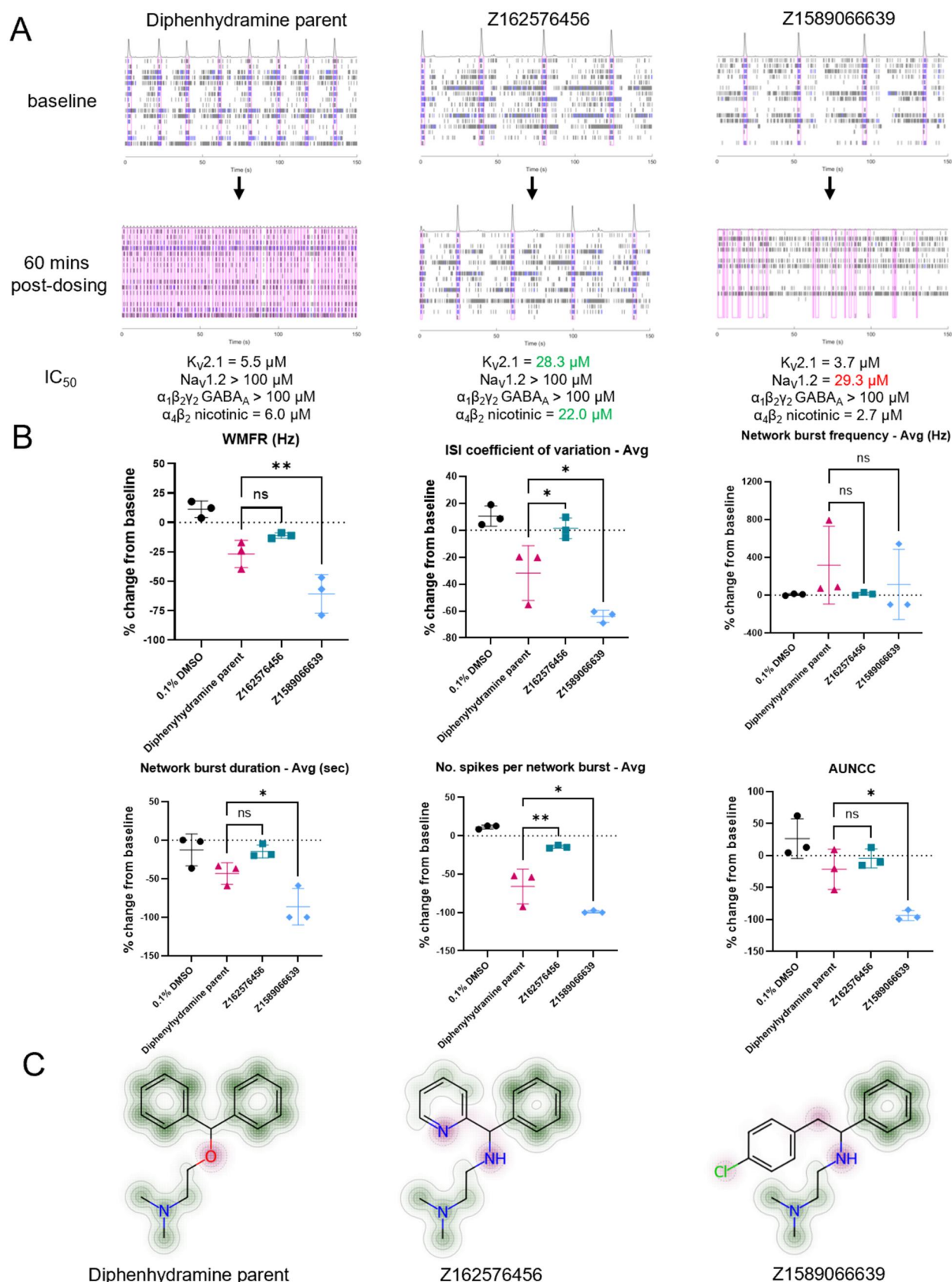


Fig. 5. MEA analysis of diphenhydramine and derivatives at 10 μM . A. Representative raster plots depicting 0 to 150 s activity at baseline (top row) and 60 min postdose (bottom row) with parent or derivative Enamine REAL compounds, Z162576456 and Z1589066639. IC₅₀ values at 4 ion channels are shown below raster plots (black text = similar potency, green text = lower potency, and red text = higher potency of Enamine REAL compound compared with parent at the given ion channel). B. Key MEA parameters highlighting significant differences. Data shown as mean $\pm$ SD from $n = 3$ biological replicates. Data analyzed by 1-way ANOVA with Šídák's multiple comparisons test; * $P < 0.05$, ** $P < 0.01$. C. Chemical structures of compounds show areas of structural similarity (green) and differences (purple). AUNCC, area under normalized cross-correlation; ISI, interspike interval; WMFR, weighted mean firing rate.

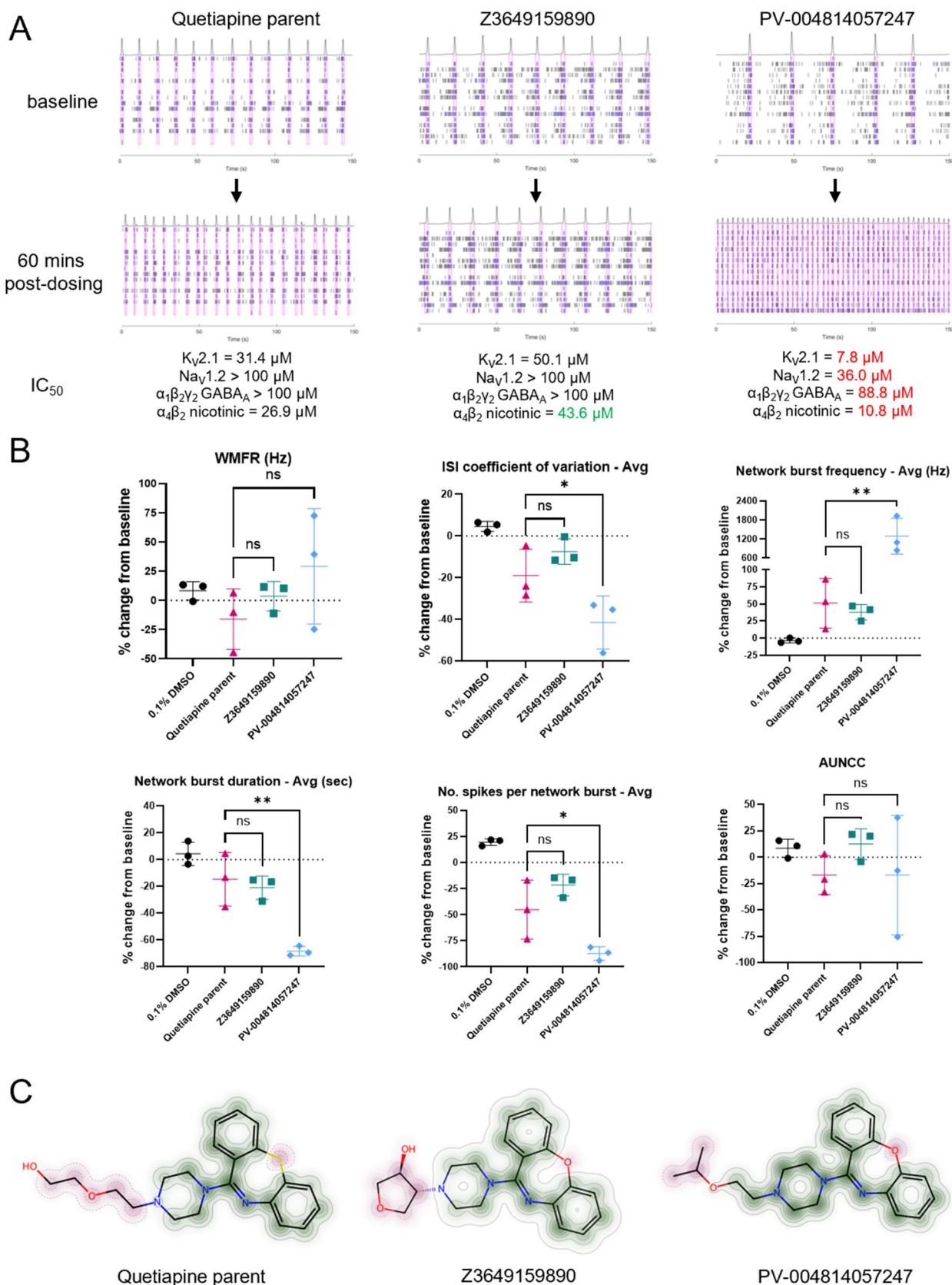


Fig. 6. MEA analysis of quetiapine and derivatives at 10 μ M. **A.** Representative raster plots depicting 0 to 150 s activity at baseline (top row) and 60 min postdose (bottom row) with parent or derivative Enamine REAL compounds, Z3649159890 and PV-004814057247. IC₅₀ values at 4 ion channels are shown below raster plots (black text = similar potency, green text = lower potency, and red text = higher potency of Enamine REAL compound compared with parent at the given ion channel). **B.** Key MEA parameters highlighting significant differences. Data shown as mean $\pm$ SD from $n = 3$ biological replicates. Data analyzed by 1-way ANOVA with Šidák's multiple comparisons test; * $P < 0.05$, ** $P < 0.01$. **C.** Chemical structures of compounds show areas of structural similarity (green) and differences (purple). AUNCC, area under normalized cross-correlation; ISI, interspike interval; WMFR, weighted mean firing rate.

nicotinic receptor. In both ion channels, 2 pharmacophore features increased potency whereas 2 features decreased potency. The SAR analysis for Nav1.2 and the $\alpha_1\beta_2\gamma_2$ GABA_A receptor revealed many differential pharmacophore features, but none achieved statistical significance. Interestingly, 1 differential pharmacophore feature at Nav1.2, showed increased potency in line with Kv2.1 and the $\alpha_4\beta_2$ nicotinic receptor. The commonality of these preliminary signals between ion channels highlights the value in developing SAR for CNS-related toxicities. This warrants further investigation to build a larger dataset that can be utilized to build a robust predictive CNS SAR model.

Typically, CNS-penetrating molecules are of modest molecular weight, low polar surface area, have $pK_a \sim 7.5$ to 10.5 and possess few hydrogen bond donors (Pajouhesh and Lenz 2005). The identified pharmacophore features provide defined chemical substructures: Although considering physicochemical properties of these substructures in isolation is unlikely to be beneficial (as the molecular scale determines CNS penetration not isolated substructural elements), we note some links. For example, the pharmacophores leading to increased potency both include bases. Although none of the identified significant pharmacophores contain acidic groups which would be excluded by the typical CNS-penetrating molecular pK_a range. Further, consideration should be given to the effect of metabolites. Changes in molecular structure due to metabolism inevitably modify a molecule's properties and behaviour, but these changes remain challenging to predict. For example, if a primary metabolite has increased seizure potential and it crosses the blood–brain barrier, then this may induce seizures. However, usually, metabolism leads to more polar compounds that are easier to excrete and are less CNS-permeable, therefore are less likely to be seizurogenic in vivo. It is likely that additional SAR will be introduced by primary metabolites, warranting consideration in future work.

For machine learning projects, it is common to develop applicability domains to approximately define where a model can reasonably be applied. A question that emerges for this work is “can we define a suitable applicability domain?”. Given the size of the dataset for this work the standard feature structure-based spaces are quite sparse and descriptor/property spaces show very limited ranges in some dimensions which can make an applicability domain definition less reliable. This work provides a preliminary analysis and small dataset in the supporting information. This enables others to build from this work and combine the data with new measurements to circumvent some of these issues.

As the ion channel data indicated SAR, we selected representative molecules for MEA (a lower throughput technology) to confirm that the altered ion channel potency for derivative molecules translated into phenotypic differences in seizure in hiPSC neurons. The data showed that altered IC_{50} values translated to differential changes in MEA for the 3 parent molecules tested compared with their derivatives. A chemical structure analysis revealed areas of similarity and difference associated with these potency changes, guiding structural chemistry away from unwanted structural features. There were some notable observations for amoxapine in that shifting the piperazine NH group position, through ring substitution to a diazepane, or replacing a hydrogen with a bulkier and flexible ethyl chain altered potency; this might be due to altering the accessibility of an otherwise relatively rigid group, which could affect binding at the ion channels screened. Similarly for diphenhydramine, replacing the benzene ring with a pyridine ring and replacing the ether group with a secondary amine functionality appeared to confer reduced seizure risk, whereas addition of a halobenzyl

ring, along with the secondary amine substitution, increased seizure risk. Unfortunately, in the current dataset, from the diphenhydramine parent, the secondary amine replacement occurs jointly with other replacements and with divergent IC_{50} impact, so we cannot decompose the individual group contributions to seizurogenic risk mitigation any further at this stage.

Secondary safety pharmacology data are included in Investigational New Drug applications, yet only hERG screening requirements are explicitly described in ICH S7B guidelines (ICH 2005). Recently, there have been calls to revise and update the ICH S7A guidelines following their original publication in 2000 (Bowes et al. 2012; Lynch et al. 2017; Valentin and Leishman 2023, 2025; Brennan et al. 2024). A key proposal for the revisions focuses on human-based in vitro and in silico secondary pharmacology models. Indeed, Brennan et al. noted that the field could be revolutionized with the combination of high-quality functional data and machine learning models (Brennan et al. 2024). Recently, there has been a significant global commitment to reduce reliance on animal studies across the life sciences, particularly in preclinical safety testing (FDA 2022, 2025; UK Government 2025). Focus has been on developing and implementing NAMs (in vitro systems, in silico models, ex vivo tissues) for regulatory decision-making to increase accuracy and ethical acceptance (FDA 2025). Reflecting this commitment, the iLA^{Seizure} work was cited by The US Food and Drug Administration (US FDA) in their publication on promising new approach methodologies (NAMs) (Avila et al. 2023). Correspondingly, between 2015 and 2024 there were considerable increases of 35% and 27% in the use of stem cell-derived neurons and neuronal/glial co-cultures, respectively, for seizure liability testing within the safety pharmacology industry (Authier et al. 2016, 2025). This centrally positions iLA^{Seizure} and the proposed CNS SAR safety model as a combined human-based approach that could be implemented earlier in the drug discovery process to assess safety in line with these proposals.

Since the 1960s, QSAR approaches have been widely used in drug discovery to predict blood–brain barrier penetration and pharmacological activity based on chemical substructures (Gupta 1989). Today, several SAR platforms are commercially available from Certara (D360TM (Certara, n.d.) and LibraTM (“Libra,” n.d.)), BioAscent (BioAscent, n.d.), and Solvonis (“Initiation of artificial intelligence-supported CNS drug discovery programme—Solvonis Therapeutics,” n.d.). However, the software and data used to build these platforms are all proprietary. Further, while some have a CNS focus, their primary function is to predict brain penetration and primary pharmacology for drug candidate prioritization. Prediction models for drug-induced seizure liability have been published; however, these are not without caveats.

For example, while a support vector machine classification prediction model with an “excellent” prediction accuracy of 86% has been published (Zhang et al. 2011), the model lacked mechanistic or target prediction insights and was limited to general predictivity. More recently, publicly-available ChEMBL data from 230 (196 seizure-positive and 34 seizure-negative) compounds was used to build 3 machine learning models to predict seizure-liability across 67 CNS targets (Lane et al. 2025). Initially, these models appeared to correctly predict the activity and seizure-liability of the molecules assessed; however, the prediction accuracy was not quantified. Further, there were concerns over the quality and availability of the data used to train the models, with the authors noting that the models predicted moderate activity of eglumetad at the NMDA receptor, whereas experimental testing revealed no activity (no binding at $>100\mu M$) (Lane et al. 2025).

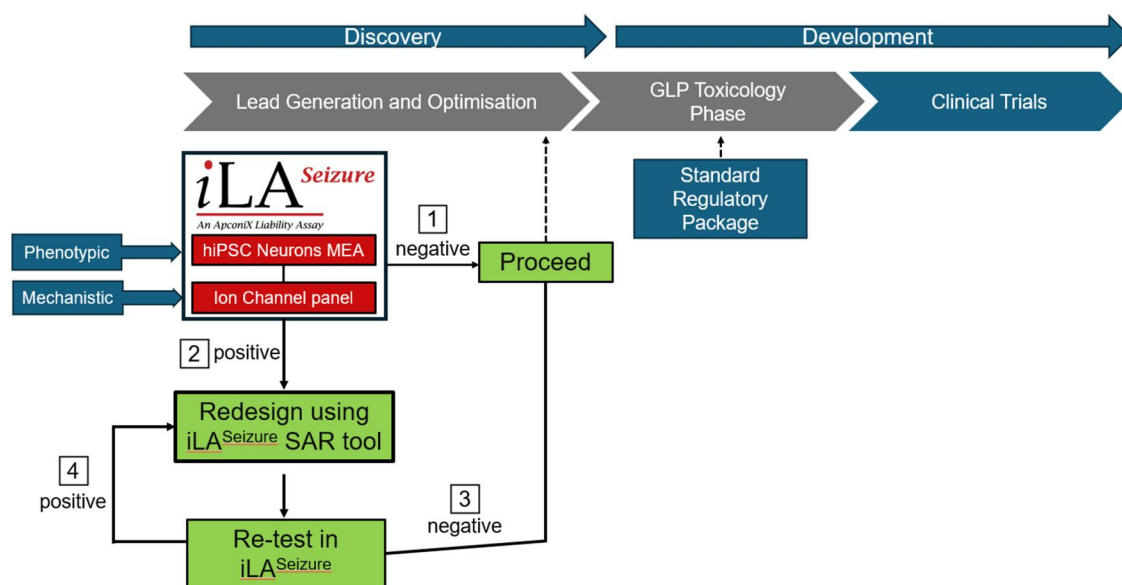


Fig. 7. Positioning iLA^{Seizure} SAR within the early stages of drug discovery and development. iLA^{Seizure} is currently used particularly in the lead generation and optimization phases of preclinical safety testing. An SAR tool could help evaluate and redesign compounds with a positive result in iLA^{Seizure}. 1. Compounds that have a negative result in the initial iLA^{Seizure} screen will proceed to the next stage of the discovery and development process. 2. For compounds with a positive result in iLA^{Seizure}, we propose these candidates will be applied to the iLA^{Seizure} SAR model and redesigned to remove any chemical structure flags. 3. The newly designed compound would then be screened in the iLA^{Seizure} panel, with a negative result allowing the compound to proceed to the next steps of development. 4. If the redesigned compound produced another positive result, it could be fed back into the iLA^{Seizure} SAR tool for further redesign.

Finally, while these models have been shown to predict drug-induced seizure liability across several CNS targets, there is no report of which substructures are associated with an increased seizure risk. In contrast, the data we present here, and the proposed CNS SAR model, specifically address SAR for secondary pharmacology with a focus on drug-induced CNS liability, presenting specific substructures associated with altered risk at the ion channels screened, using gold-standard electrophysiology data.

One concern we had for this ambitious project was that there would be no detectable SAR for seizure-related ion channels. Specifically, we anticipated that the dataset could be “flat,” in that any variation of pharmacophore features away from the parent compounds would reduce potency to $>100\mu\text{M}$, yielding little useful information. However, this was clearly not the case; the data provided a complex and detailed picture of the link between structure and potency. Indeed, many of the significant pharmacophore features were associated with increased potency; this was perhaps unexpected but advantageous in generating SAR.

One caveat of this work was the relatively small compound set size which prevented application of robust machine learning approaches, such as neural networks, Random Forest models, or support vector machines. However, as an initial proof-of-concept study, this compound set was sufficient to begin developing SAR for CNS-related ion channels. To train a reliable and robust predictive SAR model would likely require ~ 800 to 1000 small molecules with different chemical substructures spanning a range of IC_{50} values at the chosen ion channels. This is commensurate with previously published research showing datasets of ~ 1000 compounds are required to train and test reliable predictive QSAR models (Roy et al. 2008; Rácz et al. 2021; National Center for Toxicological Research (NCTR) 2023; Seal et al. 2024). Further, additional ion channels could be screened to build a more complete picture across the iLA^{Seizure} panel, with a focus on the most structurally diverse channels. It would also be important to

investigate which ion channels are most associated with seizure risk in the whole-cell hiPSC-neuronal model by testing a subset ($\sim 10\%$ of compounds) in MEA.

In conclusion, the data suggest that there are key pharmacophore features that present a significant risk for potency at ion channels associated with drug-induced toxicities to the CNS. Given the relatively small dataset, care has been taken not to overinterpret the data. However, small initial clusters based on chemical structure show some commonality in terms of activity (IC_{50}) for 2 of the 4 ion channels tested. This suggests there may be regions of chemical spaces which are privileged in terms of reducing seizure potential. Taken together, the data provide insight and a starting point for an SAR predictive tool (Fig. 7). These findings are the first report of seizure-related ion channel SAR and are the first step in helping medicinal chemists redesign their molecules to avoid CNS liability supporting optimal drug design. SAR modelling has proven successful in the cardiac liability field, with significant reductions in hERG liability and no market withdrawals due to TdP under ICH guidance (Gund 1977; Gintant et al. 2016; Garrido et al. 2020; Creanza et al. 2021). By analogy, early stage safety screening combined with an understanding of the structural features associated with ion channel potency provides an exciting opportunity to reduce CNS liability.

Supplementary material

Supplementary material is available at *Toxicological Sciences* online.

Funding

All work described was funded by Apconix either directly in our own research laboratory or by paid fee for service to MedChemica.

Conflicts of interest. RAR, KLR, HGJ, KJ, M-AM, LAKZ, JLMcD, KN, and M-AM are all employees of Apconix, an integrated ion channel research and toxicology consultancy that provides expertise to academia, industry, NGOs, and government on all aspects of nonclinical program design and delivery. EJG and JES are employees of MedChemica Consultancy Ltd.

Data availability

All data generated or analyzed in this study are included in this published article. IC₅₀ data for the seizure analysis (<https://doi.org/10.5281/zenodo.20658862>) and the ChEMBL hERG datasets (<https://doi.org/10.5281/zenodo.20660831>) have been deposited in Zenodo.

References

- Allen TEH, McDonagh JL, Broncel M, Bryant CJ, Incarnato D, Vasudevan A, Khan RT. 2024. Physicochemical principles driving small molecule binding to RNA. *bioRxiv*, preprint. <https://doi.org/10.1101/2024.01.31.578268>
- Authier S, Arezzo J, Delatte MS, Kallman MJ, Markgraf C, Paquette D, Pugsley MK, Ratcliffe S, Redfern WS, Stevens J, et al. 2016. Safety pharmacology investigations on the nervous system: an industry survey. *J Pharmacol Toxicol Methods*. 81:37–46. <https://doi.org/10.1016/j.vascn.2016.06.001>
- Authier S, Delatte MS, Wakeford A, Halpern W, McMaster O, Mellon D, Rao DB, Traevert M, Valentin J-P, Sokolowski K, et al. 2025. Toxicology and safety pharmacology investigations on the nervous system: 2024 industry survey. *J Pharmacol Toxicol Methods*. 135:108378. <https://doi.org/10.1016/j.vascn.2025.108378>
- Avila AM, Bebenek I, Mendrick DL, Peretz J, Yao J, Brown PC. 2023. Gaps and challenges in nonclinical assessments of pharmaceuticals: an FDA/CDER perspective on considerations for development of new approach methodologies. *Regul Toxicol Pharmacol*. 139:105345. <https://doi.org/10.1016/j.yrtph.2023.105345>
- Bell DC, Dallas ML. 2018. Using automated patch clamp electrophysiology platforms in pain-related ion channel research: insights from industry and academia. *Br J Pharmacol*. 175:2312–2321. <https://doi.org/10.1111/bph.13916>
- BioAscent, n.d. Accelerating CNS drug discovery. *BioAscent Integrated Drug Discovery Services*. [Webpage accessed 2026 May 26]. <https://www.bioascent.com/integrated-drug-discovery/in-silico-discovery-and-data-analysis/accelerating-cns-drug-discovery>
- Bowes J, Brown AJ, Hamon J, Jarolimek W, Sridhar A, Waldron G, Whitebread S. 2012. Reducing safety-related drug attrition: the use of in vitro pharmacological profiling. *Nat Rev Drug Discov*. 11:909–922. <https://doi.org/10.1038/nrd3845>
- Brennan RJ, Jenkinson S, Brown A, Delaunoy A, Dumotier B, Pannirselvam M, Rao M, Ribeiro LR, Schmidt F, Sibony A, et al. 2024. The state of the art in secondary pharmacology and its impact on the safety of new medicines. *Nat Rev Drug Discov*. 23:525–545. <https://doi.org/10.1038/s41573-024-00942-3>
- Certara, n.d. D360 scientific informatics platform. D360 Scientific Data Informatics Platform. [Webpage accessed 2026 May 20]. <https://www.certara.com/d360-scientific-informatics-platform/>
- Cook D, Brown D, Alexander R, March R, Morgan P, Satterthwaite G, Pangalos MN. 2014. Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nat Rev Drug Discov*. 13:419–431. <https://doi.org/10.1038/nrd4309>
- Cramer RD. 2011. Rethinking 3D-QSAR. *J Comput Aided Mol Des*. 25:197–201. <https://doi.org/10.1007/s10822-010-9403-z>
- Creanza TM, Delre P, Ancona N, Lentini G, Saviano M, Mangiatordi GF. 2021. Structure-based prediction of hERG-related cardiotoxicity: a benchmark study. *J Chem Inf Model*. 61:4758–4770. <https://doi.org/10.1021/acs.jcim.1c00744>
- Daylight Theory: SMARTS – a language for describing molecular patterns, n.d. [Webpage accessed 2026 Mar 18]. <https://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>
- Easter A, Bell ME, Damewood JR, Redfern WS, Valentin J-P, Winter MJ, Fonck C, Bialecki RA. 2009. Approaches to seizure risk assessment in preclinical drug discovery. *Drug Discov Today*. 14:876–884. <https://doi.org/10.1016/j.drudis.2009.06.003>
- Easter A, Sharp TH, Valentin JP, Pollard CE. 2007. Pharmacological validation of a semi-automated in vitro hippocampal brain slice assay for assessment of seizure liability. *J Pharmacol Toxicol Methods*. 56:223–233. <https://doi.org/10.1016/j.vascn.2007.04.008>
- FDA. 2022. FDA Modernization Act 2.0. Pub. L. No. 117-328, § 3209, 136 Stat. 4459, 5841 (codified at 21 U.S.C. § 355(z)).
- FDA. 2025. Roadmap to reducing animal testing in preclinical safety testing. Silver Spring (MD): FDA; [released 2025 Apr 10; cited 2026 Mar 18]. https://www.fda.gov/files/newsroom/published/roadmap_to_reducing_animal_testing_in_preclinical_safety_studies.pdf
- Garrido A, Lepailleur A, Mignani SM, Dallemagne P, Rochais C. 2020. hERG toxicity assessment: useful guidelines for drug design. *Eur J Med Chem*. 195:112290. <https://doi.org/10.1016/j.ejmech.2020.112290>
- Gintant G, Sager PT, Stockbridge N. 2016. Evolution of strategies to improve preclinical cardiac safety testing. *Nat Rev Drug Discov*. 15:457–471. <https://doi.org/10.1038/nrd.2015.34>
- Green MR, Sambrook J. 2012. Molecular cloning: a laboratory manual. 4th ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.
- Griffen E, Leach AG, Robb GR, Warner DJ. 2011. Matched molecular pairs as a medicinal chemistry tool. *J Med Chem*. 54:7739–7750. <https://doi.org/10.1021/jm200452d>
- Grygorenko OO, Radchenko DS, Dziuba I, Chuprina A, Gubina KE, Moroz YS. 2020. Generating multibillion chemical space of readily accessible screening compounds. *iScience*. 23:101873. <https://doi.org/10.1016/j.isci.2020.101681>
- Gund P. 1977. Three-dimensional pharmacophoric pattern searching. In: Hahn FE, Kersten H, Kersten W, Szybalski W, editors. *Progress in molecular and subcellular biology*. Heidelberg, Berlin: Springer. p. 117–143. https://doi.org/10.1007/978-3-642-66626-1_4
- Gupta SP. 1989. QSAR studies on drugs acting at the central nervous system. *Chem Rev*. 89:1765–1800. <https://doi.org/10.1021/cr00098a007>
- ICH. 2005. The non-clinical evaluation of the potential for delayed ventricular repolarisation (QT interval prolongation) by human pharmaceuticals. S7B. London: ICH Harmonised Tripartite Guideline.
- Initiation of artificial intelligence-supported CNS drug discovery programme – Solvonis therapeutics, n.d. [Webpage accessed 2026 Jun 2]. <https://solvonis.com/news/initiation-of-artificial-intelligence-supported-cns-drug-discovery-programme>
- Kreir M, Van Deuren B, Versweyveld S, De Bondt A, Van den Wyngaert I, Van der Linde H, Lu HR, Teuns G, Gallacher DJ. 2018. Do in vitro assays in rat primary neurons predict drug-induced seizure liability in humans? *Toxicol Appl Pharmacol*. 346:45–57. <https://doi.org/10.1016/j.taap.2018.03.028>
- Lane TR, Snyder SH, Harris JS, Urbina F, Ekins S. 2025. Adverse outcome pathway and machine learning to predict drug induced seizure liability. *ACS Chem Neurosci*. 16:2085–2099. <https://doi.org/10.1021/acscchemneuro.5c00177>

- Libra, n.d. Certara. [Webpage accessed 2026 Jun 2]. <https://www.certara.com/toxstudio/libra-dili-prediction/>
- Lynch JJ, Van Vleet TR, Mittelstadt SW, Blomme EAG. 2017. Potential functional and pathological side effects related to off-target pharmacological activity. *J Pharmacol Toxicol Methods*. 87:108–126. <https://doi.org/10.1016/j.vascn.2017.02.020>
- McDonagh JL, Zavitsanou S, Harrison A, Zubarev D, Kessel T, van Wunsch BH, Cipcigan F. 2024. Chemical space analysis and property prediction for carbon capture solvent molecules. *Digit Discov*. 3:528–543. <https://doi.org/10.1039/D3DD00073G>
- National Center for Toxicological Research (NCTR). 2023. Drug induced liver injury rank (DILIrank) dataset. Silver Spring (MD): FDA (US); [cited 2026 Mar 18]. <https://www.fda.gov/science-research/liver-toxicity-knowledge-base-ltkb/drug-induced-liver-injury-rank-dilirank-20-dataset>
- Pajouhesh H, Lenz GR. 2005. Medicinal chemical properties of successful central nervous system drugs. *NeuroRx*. 2:541–553. <https://doi.org/10.1602/neurorx.2.4.541>
- Rácz A, Bajusz D, Héberger K. 2021. Effect of dataset size and train/test split ratios in QSAR/QSPR multiclass classification. *Molecules*. 26:1111. <https://doi.org/10.3390/molecules26041111>
- RDKit: Open-Source Cheminformatics Software, n.d. [Webpage accessed 2026 Mar 16]. <https://www.rdkit.org/>
- REAL Database – Enamine, n.d. [Webpage accessed 2025 Apr 22]. <https://enamine.net/compound-collections/real-compounds/real-database>
- Reutlinger M, Koch CP, Reker D, Todoroff N, Schneider P, Rodrigues T, Schneider G. 2013. Chemically advanced template search (CATS) for Scaffold-Hopping and prospective target prediction for ‘orphan’ molecules. *Mol Inform*. 32:133–138. <https://doi.org/10.1002/minf.201200141>
- Rockley K, Roberts R, Jennings H, Jones K, Davis M, Levesque P, Morton M. 2023. An integrated approach for early in vitro seizure prediction utilising hiPSC neurons and human ion channel assays. *Toxicol Sci*. 196:126–140. <https://doi.org/10.1093/toxsci/kfad087>
- Roy PP, Leonard JT, Roy K. 2008. Exploring the impact of size of training sets for the development of predictive QSAR models. *Chemom Intell Lab Syst*. 90:31–42. <https://doi.org/10.1016/j.chemolab.2007.07.004>
- Seal S, Williams D, Hosseini-Gerami L, Mahale M, Carpenter AE, Spjuth O, Bender A. 2024. Improved detection of drug-induced liver injury by integrating predicted in vivo and in vitro data. *Chem Res Toxicol*. 37:1290–1305. <https://doi.org/10.1021/acs.chemrestox.4c00015>
- Shen W, Cui C, Su X, Zhang Z, Velez-Arce A, Wang J, Shi XC, Zhang YB, Wu J, Chen YZ, et al. 2024. Activity cliff-informed contrastive learning for molecular property prediction. *ChemRxiv preprint*. <https://doi.org/10.26434/chemrxiv-2023-5cz7s-v2>
- Taylor R, Cole JC, Cosgrove DA, Gardiner EJ, Gillet VJ, Korb O. 2012. Development and validation of an improved algorithm for overlaying flexible molecules. *J Comput Aided Mol Des*. 26:451–472. <https://doi.org/10.1007/s10822-012-9573-y>
- Tukker AM, van Kleef RGDM, Wijnolts FMJ, de Groot A, Westerink RHS. 2020a. Towards animal-free neurotoxicity screening: applicability of hiPSC-derived neuronal models for in vitro seizure liability assessment. *ALTEX*. 37:121–135. <https://doi.org/10.14573/altex.1907121>
- Tukker AM, Wijnolts FMJ, De Groot A, Westerink RHS, Westerink RHS. 2020b. Applicability of hiPSC-derived neuronal co-cultures and rodent primary cortical cultures for in vitro seizure liability assessment. *Toxicol Sci*. 178:71–87. <https://doi.org/10.1093/toxsci/kfaa136/5899742>
- UK Government. 2025. DSIT (Department for Science, Innovation and Technology), Home Office, Defra (Department for Environment, Food & Rural Affairs). Replacing animals in science: a strategy to support the development, validation and uptake of alternative methods [Internet]. London (UK): DSIT; [published 2025 Nov 11; cited 2026 Aug 7]. <https://www.gov.uk/government/publications/replacing-animals-in-science-strategy/replacing-animals-in-science-a-strategy-to-support-the-development-validation-and-uptake-of-alternative-methods>
- Valentin J-P, Leishman D. 2023. 2000–2023 Over two decades of ICH S7A: has the time come for a revamp? *Regul Toxicol Pharmacol*. 139:105368. <https://doi.org/10.1016/j.yrtph.2023.105368>
- Valentin J-P, Leishman D. 2025. Points to consider for revising the ICH S7A guideline on safety and secondary pharmacology. *Regul Toxicol Pharmacol*. 159:105795. <https://doi.org/10.1016/j.yrtph.2025.105795>
- Walker AL, Imam SZ, Roberts RA. 2018. Minireview drug discovery and development: biomarkers of neurotoxicity and neurodegeneration. *Exp Biol Med*. 243:1037–1045. <https://doi.org/10.1177/1535370218801309>
- Wermuth CG, Ganellin CR, Lindberg P, Mitscher LA. 1998. Glossary of terms used in medicinal chemistry (IUPAC recommendations 1998). *Pure Appl Chem*. 70:1129–1143. <https://doi.org/10.1351/pac199870051129>
- Winter MJ, Redfern WS, Hayfield AJ, Owen SF, Valentin JP, Hutchinson TH. 2008. Validation of a larval zebrafish locomotor assay for assessing the seizure liability of early-stage development drugs. *J Pharmacol Toxicol Methods*. 57:176–187. <https://doi.org/10.1016/j.vascn.2008.01.004>
- Yip PK, Chapman GE, Sillito RR, Ip THR, Akhigbe G, Becker SC, Price AW, Michael-Titus AT, Armstrong JD, Tremoleda JL. 2019. Studies on long term behavioural changes in group-housed rat models of brain and spinal cord injury using an automated home cage recording system. *J Neurosci Methods*. 321:49–63. <https://doi.org/10.1016/j.jneumeth.2019.04.005>
- Zhang H, Li W, Xie Y, Wang W-J, Li L-L, Yang S-Y. 2011. Rapid and accurate assessment of seizure liability of drugs by using an optimal support vector machine method. *Toxicol In Vitro*. 25:1848–1854. <https://doi.org/10.1016/j.tiv.2011.05.015>