# Molecular Mingling: Multimodal Predictions of Ligand Promiscuity in Pentameric Ligand-Gated Ion Channels

Filip Koniuszewski[1], Florian D. Vogel[1], Konstantina Bampali[1], Jure Fabjan[1], Thomas Seidel[2], Petra Scholze[1], Philip B. Schmiedhofer[1], Thierry Langer[2] and Margot Ernst[1]*

[1]Department of Pathobiology of the Nervous System, Center for Brain Research, Medical University Vienna, Vienna, Austria, [2]Department of Pharmaceutical Sciences, Division of Pharmaceutical Chemistry, University of Vienna, Vienna, Austria

**Background:** Human pentameric ligand-gated ion channels (pLGICs) comprise nicotinic acetylcholine receptors (nAChRs), 5-hydroxytryptamine type 3 receptors (5-HT$_3$Rs), zinc-activated channels (ZAC), γ-aminobutyric acid type A receptors (GABA$_A$Rs) and glycine receptors (GlyRs). They are recognized therapeutic targets of some of the most prescribed drugs like general anesthetics, anxiolytics, smoking cessation aids, antiemetics and many more. Currently, approximately 100 experimental structures of pLGICs with ligands bound exist in the protein data bank (PDB). These atomic-level 3D structures enable the generation of a comprehensive binding site inventory for the superfamily and the *in silico* prediction of binding site properties.

**Methods:** A panel of high throughput *in silico* methods including pharmacophore screening, conformation analysis and descriptor calculation was applied to a selection of allosteric binding sites for which *in vitro* screens are lacking. Variant abundance near binding site forming regions and computational docking complement the approach.

**Results:** The structural data reflects known and novel binding sites, some of which may be unique to individual receptors, while others are broadly conserved. The membrane spanning domain, comprising four highly conserved segments, contains ligand interaction sites for which *in vitro* assays suitable for high throughput screenings are critically lacking. This is also the case for structurally more variable novel sites in the extracellular domain. Our computational results suggest that the phytocannabinoid Δ$^9$-tetrahydrocannabinol (Δ$^9$-THC) can utilize multiple pockets which are likely to exist on most superfamily members.

**Conclusion:** With this study, we explore the potential for polypharmacology among pLGICs. Our data suggest that ligands can display two forms of promiscuity to an extent

**Abbreviations:** AEs, adverse events; Bz, benzodiazepine; CBD, cannabidiol; CNS, central nervous system; ACh, acetylcholine; nAChR, nicotinic acetylcholine receptor; ECD, extracellular domain; EMA, European medicines agency; FDA, US Food and drug administration; GABA, gamma- aminobutyric acid; GABA$_A$R, GABA$_A$ receptor; ICD, intracellular Domain; INDEL, insertion/deletion; PDB, protein database; THC, Δ$^9$-tetrahydrocannabinol (= dronabinol); TMD, transmembrane domain; pLGIC, pentameric ligand gated ion channel; ZAC, zinc-activated ligand gated ion channel; 5-HT$_3$, 5-hydroxytryptamine type 3 receptors.

greater than what has been realized: 1) Ligands can interact with homologous sites in many members of the superfamily, which bears toxicological relevance. 2) Multiple pockets in distinct localizations of individual receptor subtypes share common ligands, which counteracts efforts to develop selective agents. Moreover, conformational states need to be considered for *in silico* drug screening, as certain binding sites display considerable flexibility. In total, this work contributes to a better understanding of polypharmacology across pLGICs and provides a basis for improved structure guided *in silico* drug development and drug derisking.

## INTRODUCTION

In mammalian organisms, the superfamily of pentameric ligand gated ion channels (pLGICs, also often termed cys-loop receptors) comprises ligand gated cation and anion channels, many of which are highly expressed in the mammalian nervous system. Several members of the superfamily are well established targets of important neuropsychiatric medications (Young and Snyder, 1973; Walstab et al., 2010; Dineley et al., 2015; Sieghart and Savic, 2018; Tregellas and Wylie, 2019; Breitinger and Breitinger, 2020; Caton et al., 2020). Beyond clinically approved drugs, several pLGICs are also targets of recreational and illicit drugs, and mediate toxicological effects by many synthetic, plant- and animal derived toxins (Ho and Flood, 2004; Johnston et al., 2006). Thus, all of the pLGIC genes and their resulting protein products which are found in nervous system tissues comprise a large group of potential off-targets that might mediate neuropsychiatric toxicological liabilities.

The EU funded NeuroDeRisk project (https://neuroderisk.eu/) is concerned with improving the preclinical prediction of neuropsychiatric adverse drug effects, which cause considerable problems in drug development. Its scope covers three types of drug induced adverse events (AEs): 1) seizures and convulsions; 2) psychological and psychiatric adverse events; and 3) peripheral neuropathies. The overarching project goal is to improve *in silico* and *in vitro* alerts for these three groups of nervous system AEs. **Table 1** provides some examples of pLGIC mediated toxicities of interest for the NeuroDeRisk project.

The human pLGICs comprise the γ-aminobutyric acid- (GABA) and glycine- gated anion channels (GABA$_A$Rs, GlyRs), and the cation permeable acetylcholine-, 5-HT- type 3,

and zinc- activated channels (nAChRs, 5-HT$_3$Rs, ZAC). As far as it is known, all vertebrate pLGICs have agonist binding sites in the extracellular domain at specific subunit interfaces. The interface forming subunit sides are traditionally called "principal" or (+), and "complementary" or (−). All pLGICs also feature a site for channel blocking ligands in the channel pore, and multiple allosteric binding sites in the extracellular and the transmembrane domains (**Figure 1**) (Thompson et al., 2010). Thus, drug effects can arise from multiple ligand interaction sites.

The GABA$_A$ receptor (GABA$_A$Rs) assemblies are drawn from a panel of 19 mammalian subunits (six α, three β, three γ, one δ, three ρ, one ε, one π and one θ) and their respective variants (splice isoforms, RNA-editing variants). Their ubiquitous presence in neuronal synaptic and extra-synaptic compartments (Olsen and Sieghart, 2008; Brickley and Mody, 2012) and in many glial cell types (such as astrocytes and microglia), along with a large number of identified and potential subunit assemblies, results in great pharmaco-toxicological complexity. Many subunit assemblies are thought to consist of two α, two β and one γ subunit (Sieghart, 2015). The canonical GABA binding site is located at the extracellular domain (ECD) β+/α− subunit interface of these receptors, while the homologous α+/γ− subunit interfaces form the high affinity benzodiazepine (Bz) binding sites (Sieghart, 2015). Many allosteric binding sites have been described and mediate the complex mode of action of GABA$_A$R targeting compounds (Sieghart, 2015; Puthenkalam et al., 2016; Sieghart and Savic, 2018). Most drugs in current use, such as the well-known and widely used Bzs, are allosteric modulators of these receptors and as such alter GABA elicited channel activity (Rudolph and Knoflach, 2011;

**TABLE 1 |** Examples of known neuropsychiatric adverse events from the NeuroDeRisk scope elicited by compounds which bind at the different pLGIC family members.

| Protein family | Ligand | Adverse effects |
| --- | --- | --- |
| GABA$_A$Rs | Picrotoxin, bicuculline | Seizures |
| | Diazepam/benzodiazepines | Amnestic effects, unwanted over-sedation, addiction and many others |
| | Barbiturates, propofol | Addiction |
| GlyRs | Strychnine, tranexamic acid | Convulsions |
| nAChR | Nicotine | Addiction, insomnia, nausea/vomiting |
| | Varenicline | Psychiatric adverse events, seizures, headache, nausea |
| | Mecamylamine | Amnestic effect, negative influence on cognition and learning |
| 5-HT$_3$Rs | Antiemetics (setrons) | Headache, fatigue, malaise |

**FIGURE 1 |** Binding sites in pLGICs. For each pLGIC family, one representative subunit dimer is displayed as ribbon structure, in superposition with selected ligand bound structures. The superposition of multiple ligands (where available) serves to give an impression of the overall volume of the pockets. The displayed structures are as follows: GABA$_A$Rs: ribbon and picrotoxin (yellow, channel blocker site 6): 6X40; canonical ECD site 1 (red in all panels): superposition of diazepam/6HUP, bicuculline/6X3S, flumazenil/6X3U; upper TMD interface site 3 (green shades in all panels): etomidate/6X3V and propofol/6X3T; lower TMD interface site 4 (cyan): alphaxalone/6CDU; TM3/TM4- lipid associated site 5 (ocean blue): pregnenolone sulfate/5OSC. GlyR: ribbon and glycine in canonical ECD site: 5BKG; upper ECD interface site 2 (brown): AM-3607/5TIO; site 3: ivermectin/5VDI; channel blocker site 6: 6UD3. nAChR: ribbon with alpha-bungarotoxin at site 1 in red tube: 7KOO, site 1: varenicline/6UR8 and EVP-6124/7EKT, site 3: PNU-120596/7EKT. 5-HT$_3$R: ribbon and site 3 granisetron/6NP0. Next to the ribbon renderings, some representative pentameric arrangements are shown schematically. For each family, the binding sites for the orthosteric agonists, and the high affinity benzodiazepine site (Bz) of GABA$_A$ receptors are displayed in the selected pentameric arrangements.

Sieghart and Savic, 2018). Broadly speaking, enhancing GABA effects leads to CNS depressant effects such as sedation, and to an increase of seizure threshold in anti-epileptic treatment, while reduction of GABA effects leads to CNS stimulatory effects such as arousal, anxiogenesis, and seizurogenesis. For example, the channel blocker picrotoxin induces seizures. This broad stroke picture needs to be interpreted cautiously, as evidenced by low dose paradoxical effects which are a common unwanted effect of many GABA$_A$R targeting sedatives (Mancuso et al., 2004). Unwanted neuropsychiatric effects associated with GABA$_A$Rs cover a very broad range and include seizures and convulsions, sedation, amnesia, addiction, mood changes such as the low dose aggression induction observed for Bzs, and many more (Bond, 1998; Paton, 2002; Mancuso et al., 2004; Albrecht et al., 2014). Despite many years of intense preclinical research and a large body of clinical studies, the exact molecular substrates of many unwanted effects mediated by human GABA$_A$R populations remain unclear with the literature containing many controversies (Skolnick, 2012; Sieghart and Savic, 2018).

Glycine receptors (GlyRs) are a smaller family of pLG anion channels, featuring five human GlyR subunits (α1–4, β) (Lynch, 2004; Lynch et al., 2017). The α4 subunit encoding gene is a pseudogene in humans due to an early stop codon (Simon et al., 2004). Similar to GABA$_A$ receptors, they chiefly mediate inhibitory effects in the CNS. Their spatial expression is more restricted compared to the almost omnipresent GABA$_A$Rs. In the spinal cord and brain stem they are mainly located at the post-synapse, whereas in the brain pre-synaptic and extra-synaptic GlyRs are more abundant (Lynch, 2004). Despite the high number of GlyR modulators, so far no specific glycine receptor targeting therapeutic has emerged, in contrast to the widely targeted GABA$_A$Rs. A pharmacological distinction between GABA$_A$Rs and GlyRs can be made with the highly specific orthosteric antagonist strychnine, a plant derived toxin. Pathophysiologically, strychnine-induced convulsions occur at low doses due to the block of GlyR-mediated inhibition, whereas GABA$_A$Rs are blocked at much higher concentrations (Breitinger and Breitinger, 2020). Additionally, tranexamic acid and ε-aminocaproic acid are two clinically used

antifibrinolytics which can cause convulsions as a GlyR-mediated side effect (Lecker et al., 2012). The GlyRs high promiscuity within the pLGIC family is reflected by the number and diversity of compounds that bind to and modulate the receptors, such as taurine, beta-alanine, ivermectin, bicuculline, gabazine, as well as fluoxetine and clozapine (Breitinger and Breitinger, 2020). Many more clinically used therapeutics modulate GlyRs and might elicit some fraction of their adverse effects via unwanted modulation of these receptors, especially at higher dosages.

nAChRs are pLG cation channels that are found both in the central and peripheral nervous system including the neuromuscular junction. They are built out of a choice of 16 subunits (α1–7, α9–10, β1–4, γ, δ and ε) assembled around a central sodium- and calcium-selective pore. Muscle specific and neuron specific hetero-meric and homopentameric nAChRs have been described (Gotti et al., 2006; Gotti et al., 2007; Zoli et al., 2018). The endogenous ligand acetylcholine binds at the extracellular domain of the nAChR, at the interface between an α (+) and a β (−) subunit or two α7− subunits. Both α4β2− in the central nervous system, and α3β4− receptors in the peripheral nervous system can contain additional accessory subunits. These confer altered physiological, pharmacological and/or trafficking properties as has been shown for the α5-subunit (Scholze and Huck, 2020). As for GABA$_A$Rs and GlyRs, allosteric modulatory sites have been described (Chatzidaki and Millar, 2015). Interestingly, allosteric ligands of nAChRs can mediate additional metabotropic-like signal transduction behavior (Horenstein and Papke, 2017). All nAChR targeting therapeutics in current use bind at agonist sites with different subtype preferences, and are chiefly used as smoking cessation aids. Varenicline and nicotine (applied orally as chewing gum, as nasal spray, or as transdermal patch), are approved by the U.S. Food And Drug Administration (FDA) and the European Medicines Agency (EMA), and proved to be effective medications to reduce smoking (Rosen et al., 2018). Nicotine patches have been described to cause side-effects such as addiction, insomnia, nausea and vomiting (Tønnesen et al., 1999; Schnoll et al., 2010). Although no direct effects on peripheral neuropathy is known to date, nicotine also worsens nerve regeneration after injury due to cytokine interactions (Rodriguez-Fontan et al., 2020). Initially varenicline was suspected to trigger epileptic seizures in rats as well as in humans (Erken et al., 2014), but a recent case-control study showed no significant association (Chopra et al., 2019). Neuropsychiatric side effects (depression, anxiety, panic attacks, etc.) are however still discussed (Menkü et al., 2021). Mecamylamine is a nAChR competitive antagonist. It was initially developed as a ganglion-blocker that inhibits the action of the sympathetic nervous system, and has antihypertensive properties. It has been used for this purpose for several decades before finally being replaced (Young et al., 2001).

5-HT$_3$Rs, similar to nAChRs, are cation selective and the only ion channels among serotonin (5-HT) receptors. Five 5-HT$_3$R subunits are known (namely A, B, C, D and E) with various post-translational modifications and splice variants (Brüss et al., 2000; Tzvetkov et al., 2007; Walstab et al., 2010). Only the 5-HT$_3$A

subunit is able to form functional homopentamers. Furthermore, its presence in a heteropentamer seems to be necessary since it has been shown to play a key role in assembly and trafficking (Niesler et al., 2007; Holbrook et al., 2009) [heterogeneity of native receptors reviewed in (Jensen et al., 2008)]. The majority of 5-HT$_3$Rs in the brain are located presynaptically as evidenced in rodent studies, with the exception of the hippocampus where they are localized mainly postsynaptically in somatodendritic regions (Miquel et al., 2002). Most ligands interact with 5-HT$_3$Rs through the orthosteric site (ECD interface between two subunits) and the channel site (Thompson and Lummis, 2013) [molecular determinants of ligand binding reviewed also in (Thompson et al., 2010)]. An allosteric intersubunit site has also been described in the transmembrane domain (TMD) (Trattnig et al., 2012). Currently, the main therapeutic application of drugs, specifically antagonists, targeting 5-HT$_3$Rs is for the management of chemotherapy- or radiation-induced and post-operative nausea and vomiting. 5-HT$_3$R antagonists used as antiemetics, also called setrons, include ondansetron, granisetron, dolasetron and palonosetron. They target 5-HT$_3$Rs on the vagal afferent nerve in the gut and on the chemoreceptor trigger zone in the brainstem (Basak et al., 2020). Adverse effects are common among 5-HT$_3$R antagonists and include mild to moderate headache which is the most frequently observed side effect, appearing in approximately 20%–30% of patients (Hainsworth, 2014). Other side effects include fatigue and malaise occurring in up to 13% of patients (Theriot et al., 2021).

Another, sometimes termed atypical, member of the cys-loop receptor superfamily is the zinc-activated channel (ZAC) (Davies et al., 2003). ZAC is directly gated by the ions $Zn^{2+}$, $Cu^{2+}$ and $H^+$, a property that seems to be unique among mammalian cys-loop receptors (Trattnig et al., 2016; Madjroh et al., 2021). Additionally, ZAC has distinct gating characteristics as seen with homomeric ZACs expressed in heterologous systems, such as considerable spontaneous activity, and slow activation and desensitization kinetics (Madjroh et al., 2021). However, little is known about the physiological role of this receptor. It has low amino acid sequence homology with the other members of the superfamily, but contains most of the major structural elements of a cys-loop receptor subunit (Madjroh et al., 2021). ZAC is expressed in several organs, such as pancreas, placenta, prostate, thyroid, as well as the adult and fetal brain (Davies et al., 2003; Houtani et al., 2005). The exploration of pharmacological tools to specifically target ZAC has proven to be challenging due to off-target action of several allosteric modulators in other cys-loop receptors (Madjroh et al., 2021), thus it is likely also an unknown player in adverse events elicited by pLGIC targeting substances.

In total, 45 human cys-loop receptor subunits contribute to a wide range of toxicological liabilities, which are only in part understood as the physiological function of many family members is still unknown. In the spectrum of unwanted effects, seizures and convulsions, adverse psychological and psychiatric effects, and some indirect impact on processes involved in peripheral neuropathies, place them in the center of NeuroDeRisk's attention. In this study we thus examine all binding sites present in vertebrate pLGICs with structural

evidence for their existence. Moreover, we examine a panel of *in silico* tools which facilitate prediction of off-target effects coming from binding sites that are highly conserved across the different family members. The complete list of structures that was integrated into our study comprises 31 GABA$_A$Rs, 41 GlyRs, 26 nAChRs and 16 5-HT$_3$Rs structures (**Supplementary Item S1**). These structures cover a total of 20 unique subunits, or fragments of subunits.

All cys-loop receptor subunits are glycoproteins with three domains (**Figure 1**). The ECDs and TMDs are highly conserved across the entire superfamily, while the intracellular domains (ICDs) are much more variable. Atom-level structural knowledge provides most data for the TMDs and partial or complete ECDs of multiple superfamily members. So far, fragments of ICDs have been structurally resolved only for nAChRs and 5-HT$_3$Rs (Polovinkin et al., 2018; Gharpure et al., 2019).

The available structures comprise mainly homopentameric and a few heteropentameric GABA$_A$R and nAChR assemblies. For all families which contain heteropentameric assemblies, the existing subunit stoichiometries and arrangements are still under investigation and in part controversially debated in the literature (Thompson et al., 2010) (see schematic pentamers in **Figure 1**). For heteromeric GABA$_A$Rs, the subunit arrangement for α1βγ2 assemblies is known (**Figure 1**). Binary α1β3 assemblies feature 2α:3β stoichiometry (Kasaragod et al., 2022), containing two GABA sites and a histamine site. Information for arrangements with other subunits is lacking or controversially debated at the time of writing. Concerning GlyRs, the arrangement and stoichiometry of heteromeric receptors has been extensively debated, however recent studies show an arrangement of 4α:1β, as depicted in **Figure 1** (Yu et al., 2021; Zhu and Gouaux, 2021). In neurons, heteromeric nAChRs consisting of combinations of α2, 3, 4, or 6 with β2 or 4 and homomeric α7 receptors are most dominant (Gotti et al., 2006; Gotti et al., 2007; Zoli et al., 2018). The arrangement of the muscle type receptor is also known (Zuber and Unwin, 2013), and arrangements with and without the accessory α5 subunit have been described (Scholze and Huck, 2020). Conflicting evidence regarding the exact arrangement of subunits in 5-HT$_3$AB receptors exists. Barrera et al. showed using atomic force microscopy a subunit stoichiometry of 2A:3B and a proposed arrangement of B-A-B-B-A (Barrera et al., 2005). This is in contrast to other studies suggesting an A-A interface being present in 5-HT$_3$AB receptors (Brady et al., 2001; Lochner and Lummis, 2010). However, it is acknowledged that the incorporation of the 5-HT$_3$A subunit is necessary in order to form functional heteromeric receptors (**Figure 1**) (Jensen et al., 2008).

For homo- and heteropentameric mammalian pLGICs or closely homologous vertebrate family members, ligand bound pockets with a broad diversity of ligands are by now available in the PDB (**Supplementary Figure S1**). Multiple structures of ECD ligand bound family members are available with several ligands for all families with the exception of the ZAC (**Figure 1**; **Supplementary Item S1**). The known sites at the ECD subunit interface form either the orthosteric agonist/inverse agonist sites, or allosteric modulatory sites, depending on the subunits which contribute to the interface. At the ECD interface, a novel pocket has been observed in GlyRs (**Figure 1**) (Huang et al., 2017). Ligated pockets have been reported at the upper TMD interface for nAChRs, GABA$_A$Rs and GlyRs, while the lower TMD modulatory steroid site and a lipid associated steroid bound site (Laverty et al., 2017; Miller et al., 2017) have so far only been studied in GABA$_A$Rs (**Figure 1**).

Additional candidate binding sites have been observed in more remote members of the pLGIC family that belong to evolutionary distant organisms and share in part low sequence identity. Their 3D architecture reveals that the overall structural features are conserved from bacteria to complex eukaryotes (**Supplementary Figure S1**) (Salari et al., 2014). Moreover, cholesterol and derivatives thereof have been identified in a number of structures (Laverty et al., 2017; Walsh et al., 2018; Zhao et al., 2021; Zhu and Gouaux, 2021). Recently, mutational evidence for a novel, intrasubunit site in the ECD of GABA$_A$R α5 subunits was presented (Bampali et al., 2022).

While it is well established that the pLGICs can mediate a wide range of neuropsychiatric AEs including seizures, alterations in various cognitive and mood functions (e.g., sedation, amnesia, anxiogenesis, paradoxical excitation and aggression, addiction and many more), the currently available *in vitro* high throughput screens still miss interactions of small molecules with pLGICs. This is partly due to the fact that standard screening assays cover only a fraction of the existing binding sites (Chatzidaki and Millar, 2015; Puthenkalam et al., 2016). Moreover, in the case of many compounds, functional studies show they bind at, and interact with specific pLGICs, but fail to identify the binding sites through which the effects are elicited. The binding sites displayed in **Figure 1** form the basis for the multimodal *in silico* prediction of polypharmacology alerts we undertook with the aim to generate testable predictions for binding sites lacking radioligands.

# METHODS

## Protein Sequences and Structures
All human canonical sequences have been obtained from UniProt. The full list of UniProt IDs and the corresponding gene names are listed in the **Supplementary Table S1**.

## Variant Data Mining
Data for the genes was downloaded from GnomAD versions 3.1.1 and 2.1.1. We used the former for the analysis and the latter portion of the data for cases of variants for which we could not find an amino acid match in the reference sequences. The data was fetched manually and then analyzed using Python 3 (Van Rossum and Drake Jr, 2009) and Pandas (McKinney, 2010; team, T.p.d., 2020). As the variants have a wide spread of screened alleles we filtered the outliers (**Supplementary Figure S2**). Next, we filtered out the variants, which have the frequency lower than 0.001%. Furthermore, we retained only missense variants. Those

we transferred onto a reference numbering based on an alignment of all reference protein sequences.

## PDB Files

All experimental structures analyzed in this study have been taken from the PDB (**Supplementary Item S1**) (Unwin and Fujiyoshi, 2012; Miller et al., 2017; Basak et al., 2018a; Basak et al., 2018b; Chen et al., 2018; Gharpure et al., 2019; Noviello et al., 2021; Zhang et al., 2021; Zhu and Gouaux, 2021).

## Drug Dataset Preparation and Pharmacophore Screening Database Generation

The molecular structures of the majority of the 168 investigated drugs (**Supplementary Item S2**) were downloaded from DrugBank (https://go.drugbank.com/) (Wishart et al., 2018) (154 in total). 12 of the 168 drugs were not available on DrugBank and have been retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov/) (Kim et al., 2016) instead. The remaining two drugs—ivermectin and picrotoxin represent receptor-bound ligand structures and were extracted from the PDB complexes 5VDH and 6X40 (Berman et al., 2000), respectively. Structures of drugs taken from PubChem and DrugBank were all downloaded and stored as SD-files containing 2D atom coordinates. This format was chosen because several of the investigated drugs get administered as mixtures of stereoisomers and information about undefined stereocenters is not present in SD-files containing 3D atom coordinates. The individual drug SD-files were then concatenated to a single SD-file for further processing by the software Flipper (https://docs.eyesopen.com/applications/omega/flipper.html, version 3.1.1.2, with default settings for all non-mandatory options) in order to enumerate all possible stereoisomers of drugs having one or more undefined stereocenters. The resulting SD-file contained a total of 246 structures which were then subjected to molecule conformer ensemble generation using the program OMEGA (http://www.eyesopen.com/, version 3.1.1.2, with default settings for all non-mandatory options). For the generation of the pharmacophore screening database, a simple KNIME workflow (Berthold et al., 2008) using nodes from the LigandScout KNIME extension (Wolber and Langer, 2005) (version 1.8.0, http://www.inteligand.com/download/LigandScout_Knime_Nodes.pdf, http://www.inteligand.com/ligandscout3/) was created which consisted of a SDF-reader node connected to a LDB-writer node. In the SDF-reader node configuration dialog the multi-conformer SD-file generated by OMEGA was chosen as input file, molecule conformer detection got enabled and "Compare by Name" has been selected as associated conformer matching strategy. After execution of the workflow, a pharmacophore screening database with 167 records has been obtained. Ethanol was not included in the output database since its pharmacophores do not comprise at least three features with distinct spatial positions (a prerequisite for LigandScouts pharmacophore alignment algorithm). In the generated database, every entry

corresponds to a particular drug of the dataset and comprises all of its stereoisomers and associated conformers. This means, a drug will be considered as a hit in the screening process if at least one of its stereoisomers matches the query pharmacophore. This mimics the situation in physiological systems when a drug gets administered as a mixture of stereoisomers and not all isomers of the drug are active towards the target receptor.

## Structure-Based Pharmacophore Generation

Pharmacophores of ligands bound to the assessed target binding sites were generated using the software LigandScout (http://www.inteligand.com/ligandscout3/, version 4.4.8) (Wolber and Langer, 2005) by employing the following general procedure: First, the PDB file providing the structure(s) of the receptor bound ligand was loaded into the "structure-based" perspective of LigandScout. Afterwards, for every binding site of interest (**Supplementary Table S2**), a structure-based pharmacophore of the ligand was generated (using default settings) by zooming into the corresponding site and initiating the pharmacophore generation process. The obtained structure-based pharmacophores were continuously collected in LigandScouts "alignment" perspective and, after all PDB files have been processed, stored in a single file for further processing.

## Generation of "Merged" Pharmacophores

Pharmacophores representing combinations of multiple structure-based pharmacophores were generated in the "alignment" perspective of LigandScout using the "merged" pharmacophore generation functionality. The pharmacophores "[3/Ach (α7+/α7−)_18 Hits]" (merged pharmacophores of PNU-120596 in PDB entry 7EKT), "[3/Gly (α3+/α3−)_119 Hits]" (merged pharmacophores of ivermectin in PDB entry 5VDH), "[3/GABA (α1+/β2−); (γ2+/β2−)_36 Hits]" (merged pharmacophores of phenobarbital in PDB entry 6X3W) and "[4/(GABA (α1+/α1−)_74 Hits]" (merged pharmacophores of pregnanolone in PDB entry 5O8F, tetrahydrodeoxycorticosterone in PDB entry 5OSB and alphaxolone in PDB entry 6CDU) (**Supplementary Table S2**) were all obtained via a feature-based alignment strategy since the ligands in each of these complexes displayed identical binding modes. However, slight differences were observed for the structure-based pharmacophores of ivermectin in 5VDH. Two of the ivermectin features (14 in total) that were not present in all sites of 5VDH have therefore been marked as optional for matching in the performed pharmacophore screening runs.

For the generation of the pharmacophore models "[3/GABA (γ2+/β2−)_103 Hits]" (merged pharmacophores of diazepam in 6X3X [E] and phenobarbital in 6X3W [E]), "[3/GABA (β(2/3)+/α1−)_84 Hits_]" (merged pharmacophores of propofol in 6X3T [D], etomidate in 6X3V [A] and diazepam in 6HUP [E]) a reference point (Cαs of proximal binding site residues) alignment strategy has been chosen. This was appropriate

since the ligands do not share a common binding mode but nevertheless bind to overlapping regions of homologous receptor sites.

## Pharmacophore Screening

The parallel screening of the generated structure-based and merged pharmacophores against the prepared drug database was performed by means of KNIME workflows using the activity profiling node provided by the LigandScout KNIME extension (version 1.8.0). The corresponding inputs of the activity profiling node were connected to a PMZ-reader node providing the screened pharmacophore dataset and a LDB-reader node specifying the drug database LDB file. For screening with the structure-based pharmacophore dataset, the retrieval mode in the activity profiling node configuration has been set to "Get best matching conformation," exclusion volume checks were enabled and the omitted feature count was set to zero. For screening with the merged pharmacophore dataset, the retrieval mode was set set to "Stop after first matching conformation," exclusion volume checks were enabled and the omitted feature count has been set to 9 in order to require only 3 matching drug features for any of the query pharmacophores (even in the case of the largest pharmacophore "[3/Gly (α3+/α3−)_119 Hits]" with 12 mandatory features). After execution of the screening workflows, the obtained results were saved in CSV format via the heatmap view of the activity profiling node.

## Generation of Reference Alignment for Variant Analysis

Based on the data from GnomAD versions 3.1.1 and 2.1.1 and the canonical sequences, an alignment was generated: The alignment was produced using MOE (Inc, C.C.G, 2016) by first aligning AlphaFold-predicted structures (Tunyasuvunakool et al., 2021) for all the main isoforms. For the section of the sequences before strand 1 the published structures were used for reference. The intracellular part of the proteins between TM3 and TM4 was separately aligned for minimization of gaps. After finishing the alignment of the major isoforms, other isoforms were added to the alignment with the exception of the ones too short and diverging for a confident mapping to be possible. Alignment was subsequently used in the Python environment through Biopython (Cock et al., 2009). The alignment is available as **Supplementary Item S3**.

## Sequence and Sequence-To-Structure Alignments and Similarity Calculation

Sequence alignments have been produced with Clustal X (Larkin et al., 2007) and Promals3D (Pei et al., 2008). Sequence-to-structure alignments were produced with Promals3D, PDBeFold (Krissinel and Henrick, 2004) and MOE. Specific alignments and supporting information are provided as individual FASTA files (**Supplementary Item S3**) and **Supplementary Figures S3, S10**. Pairwise sequence similarity was calculated by using the sequence similarity monitor script from MOE (ULC, 2022) with PAM180 substitution matrix.

## Analysis of Binding Sites

Binding sites have been identified through the 3D structures of the pentameric ligand-gated ion channels that are found in the PDB. These structures with ligands in the binding sites of interest were further analyzed in MOE. The ligplot function and visual inspection of the binding sites were used to identify the pocket-forming amino acids for subsequent local sequence similarity calculations and the descriptor analysis. For the TMD, the chosen amino acids (colored in the sequence alignment in the **Supplementary Item S4**) were used for all subsequent procedures (similarity and descriptor calculation, mapping of amino acids and variants to a representative ribbon structure, computation of Cα distances for the conformational analysis).

## Pocket Properties, Descriptor Set Calculations

For the proteochemometric (PCM) modeling to model the ligand-target interaction space the z-Scales(3) (Sandberg et al., 1998) and the MSWIM (Zaliani and Gancia, 1999) descriptors have been calculated using the R package "peptides" (Osorio et al., 2015) for each binding site in RStudio 1.4.1717 (http://www.rstudio.com).

## Analysis and Visualization of Descriptors

The z-Scale(3) and MSWHIM have been plotted in a 3D scatter plot and all pairwise euclidean distances have been calculated by using the Plotly (Plotly Technologies Inc., 2015) and Numpy (Harris et al., 2020) libraries in Python 3.8. Through these distances, a dendrogram and a heatmap have been created with hierarchical clustering by using Seaborn (Waskom et al., 2021). The clusters have been visualized by using Plotly.

## Analysis and Visualization of the Variants

All the following visualizations were produced using Matplotlib (Hunter, 2007).

## Ribbon Heatmaps of Conserved Regions

For each position in the reference sequence the sum of minor allele count per 100,000 was calculated. The values were transferred to 7EKT [A] (Zhao et al., 2021). The conserved regions of the proteins were also mapped onto 7EKT [A] and for those positions the values were transformed into RGB values. The values for the rest of the sequence were set to yellow (0xb3b300). The colors were then transferred to MOE and applied to the 7EKT [A] structure.

## Heatmaps of Variable Regions

Variable regions were mapped on the reference alignment. For each gene and each region, we calculated the sum of minor allele count per 100,000, divided by the length of the region in the gene's major isoform.

## Binding Site Heatmaps

The amino acids involved in the formation of binding sites were first assigned to the GABRA1 sequence used in AlphaFold

predictions. Then they were transferred to the reference numbering. For each of the binding sites of interest (and the pore facing amino acids on TM2) the per-amino acid number of variants and the sum of minor allele count per 100,000 were calculated and separately visualized.

## Conformational Analysis

The lists of pocket forming amino acids for each binding site were used to extract the Cα coordinates for each individual amino acid in the original PDB file using a Python script. To identify the different local conformations of the protein binding sites, distances and angles formed by the selected Cα were calculated using the Python library Biopython (https://biopython.org/). For the site present on a single subunit (5OSC), single subunits were used. For the sites at subunit interfaces, all dimers were used. All datasets of distances and angles (per subunit, or per dimer) have been processed using the principal component analysis (PCA) to reduce the dimensionality to three dimensions. The PCA was performed by using the Python library scikit-learn (Pedregosa et al., 2012).

All structures that are listed in **Supplementary Item S1** were used in the analyses, except those that have missing segments or amino acids (4X5T, 6QFA, 4BOT, 4BOR, 4BOI, 4BON, 4BOO, 4AQ5) or are partly unfolded or misfolded (6D6T, 6D6U).

## Homology Modeling and Computational Docking

Homology modeling of GABA$_A$R α6 was performed by the use of MODELLER 9.21 (Webb and Sali, 2016). The alignment (which is free of INDELS) was done with Promals3D with 6HUO as the template. All the dockings were performed using GOLD (Jones et al., 1997) from Hermes 2020.2.0. The ligand was set flexible by allowing to flip ring corners and detecting internal H bonds. 300 poses were generated per docking run with the option to generate diverse solutions turned on. The CHEMPLP was set as the main scoring function and chemscore for the rescoring. The best 20 docking poses from both scoring functions were used for further analysis.

## Flexible Side-Chains

site 2 (GlyR α3+/α3−): 5TIO: F13, R27, I28, R29, F32, Y78, L83, D84, L85, D86

site 3 (nAChR α7+/α7−): 7EKT: L235, N236, I244, T273, F275, M276, L277, L278, Q295, M301

site 3 (GABA$_A$ β2+/α1−): 6X3T: I228, Q229, L232, M236, T262, N265, D282, L285, M286, F289

site 4 (GABA$_A$ α1+/α1−): 5OSB: I238, Q241, V242, W245, F297, I301, T305, Y308, F309, R396

site 5 (GABA$_A$ α1): 6HIO: I296, E300, F303, L307, V432, L433, L436, L437, I440, Y441

site 5 (GABA$_A$ α6): α6 homology model: F233, F284, S287, E291, K317, I318, Y321, S322, L325, F326 (numbering as in 6HUO, with ICD numbers missing and TM4 numbering offset).

# RESULTS

The binding sites for which structural data exists were analyzed from the viewpoint of similarities across the sub-families. The study's focus is on sites for which no radioligands are available, but structural data allows a variety of *in silico* predictions. Specifically, sites 2, 3, 4 and 5 (**Figure 1**) were analyzed in more detail with the methods shown in **Figure 2** as applicable. All chosen methods except for the computational docking are suitable for simple high throughput *in silico* screenings.

In a first step, the available bound state structures and homologous apo- states were analyzed to generate an exact inventory of binding site segments, and side-chain contributing amino acids. Where structural equivalence allows this, they were extrapolated to subunits with unknown structures. The high structural variability in the ECD (**Supplementary Figures S3, S4**) leads to ambivalent sequence to structure alignments in several binding site segments for site 1 and 2, but can be performed for smaller subsets of subunits. Pocket descriptor calculation was thus applied to all TMD sites, and to a smaller set of subunits for site 1. Pharmacophore screening of a small library of known binders as reflected by Drug Central was performed for representative models of sites 1, 3, 4 and 5. The predicted pockets for sites 3, 4 and 5 which largely lack INDELS were examined with the two descriptor methods to derive predictions for highly similar pockets. Consensus predictions derived from pharmacophore screening and pocket analysis were examined and followed up with homology modeling and computational docking for the case of the phytocannabinoid Δ$^9$-THC. Prior to the selection of structures for computational docking, conformational analysis was performed as ligand bound states often display induced fit conformations, and apo-states can feature collapsed pockets.

Drug effects display large individual variability, ranging from non-responding to rare side effects observed only in a small fraction of treated individuals. As an example, midazolam, which is popular as a sedative for unpleasant medical procedures such as dental work, can induce severe paradoxical effects (Mancuso et al., 2004). Among the possible causes of such effects, genetic variants that directly impact the effect of a drug on its targets can play a role. These variants can be positioned in the binding site and thus affect the interaction of the ligand with the protein. On the other hand, variants positioned near the binding site can affect the transfer of the ligand's conformational effect. Thus, we also investigated the occurrence and frequency of missense variants in the greater region of drug binding sites.

## Binding Sites in the ECD

The ECD features a modified immunoglobulin fold (Brejc et al., 2001), consisting of several highly conserved segments with rather large variable regions at the N-termini, and interspersed into the domain in multiple places (**Supplementary Figures S3, S4**). The canonical agonist site is localized at ECD interfaces of specific subunits in all families. Allosteric sites are formed by other

**FIGURE 2 |** Graphical methods: The binding sites which were selected for this study were subjected to an array of computational tools: Amino acids contributing to the pockets were extracted. Subsequently, local structure based alignments and sequence- to- structure alignments were employed to identify or predict their topological equivalents in the other subunits where possible. These pocket forming segments were then subjected to local sequence similarity and pocket descriptor computation. Hierarchical clustering of pairwise euclidean distances between pocket descriptor coordinates was performed to predict similarities (see *Methods*). Pocket conformations were analyzed in order to guide the selection of structures for computational docking. The relative abundance of variants at and near ligand binding sites was extracted from gnomAD. Structure based pharmacophore screening was employed to derive selected bound state hypotheses to complement the pocket analysis. For dronabinol (the phytocannabinoid $\Delta^9$-THC), multimodal predictions from literature search, mining of Drug Central, pharmacophore screening and pocket similarity predictions were ranked and strong candidates were followed op with homology modeling and computational docking.

specific subunit interfaces, such as the high affinity Bz-binding site of GABA$_A$Rs (**Figure 1**). For the agonist sites of all families, as well as for the Bz-site, high affinity radioligands exist for efficient *in vitro* screenings. The binding site forming segments comprise conserved and variable parts (**Supplementary Figures S3, S4**), and the variable parts contain fairly large INDELS. Existing structures provide a glimpse into the structural and conformational variability of the canonical interface pockets, which renders them very challenging targets for rapid *in silico* methods due to the ambiguities of the INDEL positioning and the appropriate choice of ligand bound conformations (Puthenkalam et al., 2016). While there is considerable interest in allosteric sites for which radioligands are lacking (Ramerstorfer et al., 2011; Puthenkalam et al., 2016), the need for more elaborate *in silico* methods for suitable INDEL placement go beyond the scope of this study.

For a subset of GABA$_A$R receptor subunits which align well, an explorative computation of pocket descriptors was performed. For this, we used existing Bz bound structures and applied the pocket descriptor calculation in order to test and compare [z-Scales(3) and MSWHIM]. As the arrangements of α-β-γ containing receptors are generally assumed to be

identical with the known situation for α1-β-γ2 as depicted in **Figure 1**, and α-β receptors are assumed to have a β subunit in the place of the γ subunit, we constructed all hypothetical α+/β−, β+/α−, α+/γ−, γ+/β− and β+/β− interfaces. Based on the available bound states we extracted the binding site forming sidechains and examined the resulting z-Scales(3) and MSWHIM descriptors (**Supplementary Figures S5–S7**). Both descriptors separate the GABA binding β+/α− interfaces from the Bz-binding α+/γ− interfaces, thus recapitulating the known pharmacology (**Supplementary Figure S5**). In line with experimental findings, the modulatory α+/β− interfaces share more properties with the Bz-sites than with the GABA sites (**Supplementary Figure S6**).

In addition to the canonical interface pocket, a novel pocket has been observed in a crystal structure of a GlyR α3 homopentamer (**Figures 1**, **3**) (Huang et al., 2017). Structural superposition with other superfamily members and sequence to structure alignments indicate that this site comprises conserved and variable segments and considerable structural diversity (**Figures 3B,C**). Sequence similarity analysis for the conserved parts of this site suggest that homologous sites likely exist at interfaces of other subunits across the entire superfamily. The principal component of GABA$_A$R β, δ and θ subunits is most

**FIGURE 3 |** Novel binding site at the upper ECD interface of GlyR α3. **(A)** the binding site from 5TIO with pocket forming segments labeled (see **Supplementary Figure S4** for the topology) and amino acids displayed in stick rendering. **(B)** 5TIO (in grey) with the ligand AM-3607 bound at the novel pocket in superposition with closely homologous proteins glycine α2 (5BKF chain D in gold), glycine β (5BKF chain E in blue), GABA$_A$ α1 (6X3V chain B in pink) and GABA$_A$ β2 (6X3V chain C in red) at the complementary subunit to show the structural variability in VR3. **(C)** Amount of variants across all genes of interest in the region of this binding site mapped onto 7EKT. Ribbon heatmap represents the sum of minor allele counts per 100,000 at each aligned position. Variable regions, where the 44 subunits cannot be aligned to a single reference structure, are shown in gray.

similar, the much more variable complementary component of GABA$_A$R β, δ, θ and γ subunits are the strongest candidates for sites that might be used by ligands common to multiple family members. However, the overall size and shape is strongly affected by the variable region 3. A pharmacophore screen into this site provides first insights into the chemical space of candidate ligands. Of the small library we used here (**Supplementary Item S2**; **Supplementary Figure S11**), 46 drugs with a wide range of chemical scaffolds generated hits in this site (**Supplementary Figures S12, S13**). This result is suggestive of a possible contribution from this novel binding site to the polypharmacology within the pLGICs.

Generally, the large number of variable segments present at and near ECD binding site forming regions (**Supplementary Figures S3, S4**) indicates that these binding sites are of concern only for limited polypharmacology (e.g., the interactions of Bz-site binders with the homologous GABA$_A$ receptor α+/β− interfaces (Iorio et al., 2020), or for the interactions of e.g., taurine with glycine- and GABA$_A$ receptors. The mapping of variants to a representative structure is also not possible for large parts of the ECD and needs to be considered per- subunit. An overview of the variable segments is provided in **Supplementary Figure S8**.

## Binding Sites in the TMD

The most conserved domain is the TMD, where all anion channels share >75% sequence similarity. Among nAChR subunits, conservation is as high, but 5-HT$_3$Rs and the ZAC- protein feature remarkable variability within the 5-HT$_3$ family and also relative to nAChR subunits (**Supplementary Figure S9**). The similarity index for the whole domain is not informative about the individual binding sites, it is only suggesting that the TMD is likely to contain binding sites conserved across many superfamily

members. This is also evidenced by the observation that the channel blocker picrotoxin binds to most superfamily members with comparable affinity (Erkkila et al., 2004; Das and Dillon, 2003; Das and Dillon, 2005; Yang et al., 2007; Thompson et al., 2010). The channel blocker site is the only site in the TMD for which radioligands exist, thus, there is a big need to better characterize the other allosteric sites present in this domain and to develop reliable *in silico* screening methods for these.

The best characterized site in the TMD is localized at the upper portion of the subunit interfaces, where 9 ligand bound structures exist which cover GABA$_A$Rs, nAChRs and GlyRs (**Figure 4**; **Supplementary Item S1**). The TMD is nearly free of INDELS, and structurally very conserved. Thus, it is ideally suited to compare pockets on the basis of local sequence similarity, and with the descriptor methods that have been demonstrated as computationally very inexpensive tools to estimate overlapping chemical spaces across pockets. The two descriptors used here (z- Scale(3) and MSWHIM) broadly agree on the similarities within and across family members for both the principal and the complementary components of this site (**Figure 4B**). The principal component contributing subunits fall into two clusters, one containing all GABA$_A$R subunits and individual members from all other families. The smaller cluster comprises the β subunits and the majority of the α subunits from the nAChR family, alongside with the GlyR β subunit. The complementary component segments cluster differently, with almost all subunits sharing high similarity for the MSWHIM descriptor with only the ZAC protein and the GLRB protein forming a small separate cluster (**Figure 4B**; **Supplementary Figure S14B**). This suggests that the upper TMD interface is likely to have very broadly overlapping ligands for many superfamily members. This analysis was limited to the amino acids

**FIGURE 4 |** Representative structures of the upper TMD interface pocket, and descriptor analysis: **(A)** Overview of the TMD interface with ivermectin (cyan sticks) in 5VDH. The ribbon segments that provide ligand contacts with ivermectin and with small molecules are orange, those with small molecule contacts only are yellow, and those with ivermectin only contacts are green. Amino acid numbering as in 5VDH. **(B)** Comparison diagrams for the subsite used by small molecules derived from hierarchical clustering of z-Scale(3) and MSWHIM descriptors as estimates for pocket similarities. **Supplementary Figure S14A** depicts the full dendrograms, **Supplementary Figures S14B,C** depict the analogous comparison for the ivermectin site. **(C)** Ribbon of the GABA_A R β+/α− structure with etomidate (6X3V in red), phenobarbital (6X3W in light- and dark-green). **(D)** 7EKT with PNU-120596. The yellow positions on the ribbon indicate the binding site forming segments which were used for panel **(B)**. **(E)** Amount of variants across all genes of interest in the region of this binding site mapped onto 7EKT. Ribbon heatmap represents the sum of minor allele counts per 100,000 at each aligned position. See **Supplementary Figures S18,S19** for detailed heatmaps.

that contribute to the binding sites for small molecules as determined from 7EKT, 6HUO, 6X3V, 6X3X, 6X3T and 6X3W. In order to gauge the performance for more detailed predictions, we re-analyzed the amino acid lists for the upper TMD pockets of specific GABA_ARs, as was done for the ECD interface pockets. The resulting clusters follow closely the known distinctions between barbiturate-preferring and etomidate- selective pockets, see

**Supplementary Figure S14A**: Both descriptors generate a homogenous α+/β− cluster, reflecting one of the barbiturate preferring groups. The other barbiturate preferring interface, localized at γ+/β− interfaces, forms clusters containing several β+/α− interfaces for the z-Scale(3) and MSWHIM, suggesting that barbiturate- and etomidate- preferring pockets might have common ligands. This is in fact observed for the low affinity diazepam sites (Kim et al.,
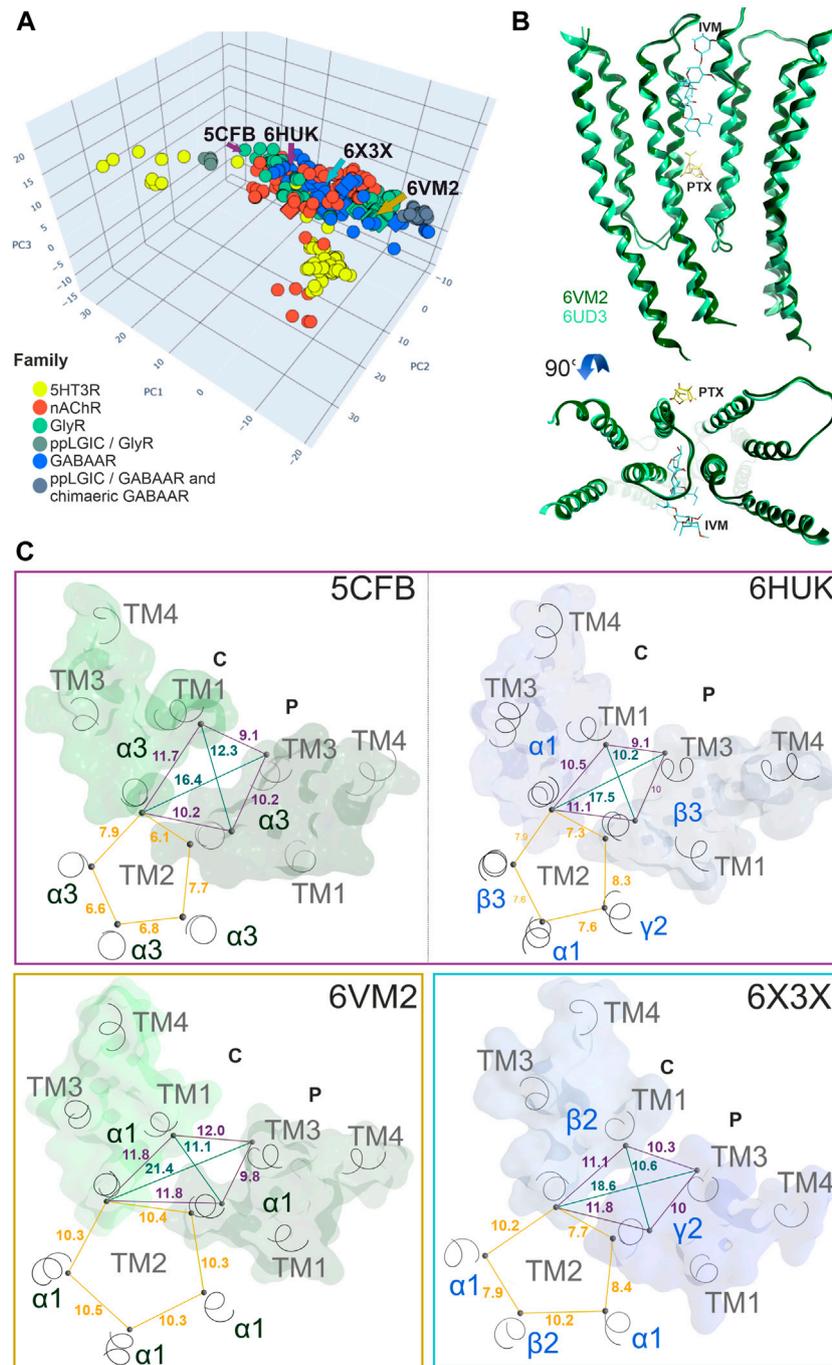
**FIGURE 5 |** Conformation analysis of the upper TMD interface pocket: **(A)** Scatter plot of the pocket conformations. Each dot represents the results for an individual upper TMD interface, diamond shaped symbols indicate the presence of a ligand in the analysed site. The colour of the dots indicates the protein family as defined in the legend, gray is used for diverse chimeras. ppLGIC: procaryotic pentameric ligand gated ion channel. Arrows and PDB IDs indicate the localization of structures which are rendered in panels **(B)** and **(C)**, arrow and box colours match across the panels. **(B)** Superposition and comparison of two interface forming subunits of the ivermectin bound Gly R structure 6VM2 and the picrotoxin bound 6UD3, featuring nearly identical conformations of the entire TMD. **(C)** Different conformations result in different distances between alpha carbon atoms: A plane through the upper TMD is depicted with all five TM2 segments and two subunits. Distances between alpha carbons are shown to illustrate backbone conformation, where yellow lines are distances between homologous pore forming TM2 residues as an indication for asymmetry, purple lines are distances between pocket forming residues defining the circumference, and green lines the pocket "diagonals".

2020). The etomidate insensitive β1+/α− form a separate cluster, again recapitulating the known pharmacology.

The ivermectin binding site, also localized at the upper TMD interface, is larger and overlaps partly with the sites for the small molecule ligands discussed above, see **Figure 4**. In our pharmacophore screen, we observed that a stringent screen into ivermectin bound state, as expected, cannot retrieve small molecule binders, but a less stringent screen will produce many weak hits and is prone to false negatives. Applying the descriptor analysis to the amino acids that form ivermectin contacts results in partly controversial findings between the two used methods, see **Supplementary Figure S14C**: For example, the principal component of ZACN clusters with GABA$_A$R subunits in the z-Scale(3) and MSWHIM dataset. It would be interesting to have experimental data with which the results could be validated, requiring more subtypes of pLGICs to be tested. At the present time, several high and low affinity interactions are known suggesting a smaller group of high affinity sites, and a very broadly promiscuious profile in the micromolar range (Lynagh and Lynch, 2010).

The high similarity for the pockets not only within, but also across families suggests common ligands. Most of the experimental structures that can be employed for structure-based pharmacophore screens are ligand-bound GABA$_A$Rs. Since homologous upper TMD binding pockets exist within receptor families (low affinity diazepam/etomidate site and both barbiturate sites) as well as in different pLGIC family members (7EKT for nAChR and 5VDH for glycine receptors), multiple pharmacophore screens to this binding site were performed (**Supplementary Figures S11, S12**). Most pharmacophores derived from ligand bound GABA$_A$Rs predict a wide range of drugs to be potentially accommodated by upper TMD interface sites (**Supplementary Figure S12**), in agreement with the promiscuity of this pocket according to literature (Iorio et al., 2020). All of the 18 substances that produced drug-target hits from pharmacophore screens derived from the homologous site in the nAChR α7 receptor (PDB ID: 7EKT) also generated a hit in at least one homologous site of GABA$_A$Rs or GlyRs. As an example, the plant compound dronabinol (Δ$^9$-THC), was predicted to potentially bind at the upper TMD sites of nAChRs, GABA$_A$Rs and GlyRs (**Supplementary Figures S11, S12**).

Conformational analysis indicates, in line with previous work (Puthenkalam et al., 2016), substantial flexibility of this region. Structures of all family members share overlapping conformations, irrespective of the occupancy of the upper TMD binding site (**Figure 5A**), but for the cation channel branch of the superfamily, additional conformations have been observed. In these, the RMSD of the pocket backbone differs by up to 2.8 Å from the more frequently observed conformations (**Supplementary Figure S17A**). The majority of structures populate a conformation space shared by receptors with ligands present at one or several upper TMD interfaces and receptors without ligands in this pocket (apo- site 3, **Figures 5B,C**). This indicates that ligands of site 3 do not lead to

induced fit conformations. Intriguingly, the picrotoxin bound 6UD3 and an ivermectin bound 6VM2 GlyRα1 structure display nearly identical conformation not only of site 3 in the apo- state versus ivermectin- bound state, but of the entire TMD. In contrast, a very different conformation of the upper TMD interface is featured in the GlyR α3 5CFB in complex with the orthosteric antagonist strychnine, see **Figures 5A,C**. Following on this observation we selected further structures from the region around 5CFB and noted many instances of orthosteric antagonist bound structures in this region of the conformation space, including the bicuculline bound 6HUK β3+/α1− interface, **Figures 5A,C** and **Supplementary Figure S17A**. For nAChRs, we observe the same dominant effect of orthosteric antagonists on the TMD pocket conformation, see **Supplementary Figure S17A**. The strongest effect appears to be a contraction of the upper TMD pocket, mainly on the "long" diagonal between complementary TM2 and principal TM3, see **Figure 5C**. These observations suggest that orthosteric ligands are dominant drivers of the conformation in the upper TMD, which is further weakly influenced by other factors. Of further interest is the high degree of asymmetry relative to a pseudo symmetrical pentagon that occurs in many structures, such as the 6X3X GABA$_A$R with diazepam present in the upper TMD, **Figure 5C**. Interestingly, we find the conformations of chimeric constructs with ECDs from bacterial homologues to occupy distinct regions of the conformation space (**Figure 5A**; **Supplementary Figure S17A**).

The lower TMD interface also contains a binding site, which has been characterized as a modulatory site for neuroactive steroids (**Figure 6**) (Laverty et al., 2017; Miller et al., 2017; Chen et al., 2018). There is big interest in the development of neurosteroid-mimetics as putative drug candidates (Blanco et al., 2018), which also raises the question how prone to cross-reactivity such ligands might be. The analysis we performed based on local sequence similarity and pocket- descriptors indicates that the 45 subunits fall into two clusters, which are similar for the principal and the complementary component (**Figure 5C**; **Supplementary Figure S15**). In GABA$_A$Rs receptors, α subunits contribute a glutamine sidechain from TM1, which drives the high efficacy modulation and is absent from all other subunits. The remainder of the site is quite conserved for GABA$_A$Rs and GlyR α subunits, and a few members from the cation permeable family members fall into this cluster. Thus, the unique high efficacy determinant present exclusively on GABA$_A$R α subunits might render this site sufficiently unique to be of less concern for toxicological alerts.

An additional steroid binding site has been described, which is localized in a position between the lower end of TM3 and TM4 of a GABA$_A$R α1− subunit and the lipid collar, the TM3/TM4 lipid associated site (Miller et al., 2017, see **Figure 7**). Interestingly, in another structure, a PIP2 molecule was found in an overlapping localization (Laverty et al., 2019). This region has been subjected to mutational studies aimed at the identification of binding sites for endo- and phytocannabinoids (Sigel et al., 2011; Xiong et al., 2011; Bakas et al., 2017), and is thus considered to be an

**FIGURE 6 |** Representative structures of the lower TMD interface steroid pocket, and descriptor analysis: **(A)** Ribbon of the structures with modulatory steroids at the lower TMD interface: 6CDU—alphaxalone (dark green), 5O8F—pregnanolone (cyan), 5OSB—THDOC (light green). The yellow places on the ribbon indicate the binding site forming positions which were used for the pocket analysis computations. **(B)** Amount of variants across all genes of interest in the region of this binding site mapped onto 7EKT. Ribbon heatmap represents the sum of minor allele counts per 100,000 at each aligned position. See **Supplementary Figures S18,S19** for detailed heatmaps. **(C)** Comparison diagrams derived from hierarchical clustering of z- Scale(3) and MSWHIM descriptors as estimates for similarities. **Supplementary Figures S15,S16** depict the full dendrograms.

**FIGURE 7 |** Representative structures of the lipid/TM3/TM4 associated site, and descriptor analysis: **(A)** Ribbon of the 5OSC structure with pregnanolone sulfate bound. The yellow places on the ribbon indicate the binding site forming positions which were used for the pocket analysis computations. **(B)** Amount of variants across all genes of interest in the region of this binding site mapped onto 7EKT. Ribbon heatmap represents the sum of minor allele counts per 100,000 at each aligned position. See **Supplementary Figures S18,S19** for detailed heatmaps. **(D)** Comparison diagrams derived from hierarchical clustering of z-Scale(3) and MSWHIM descriptors as estimates for similarities. **Supplementary Figure 16** depicts the full dendrograms.

interaction site for both endogenous and exogenous lipophilic ligands.

The descriptor analysis does not give a consensus picture, as the z-Scales(3) predicts two clusters (one comprising all nAChR subunits, the A and C subunits of the 5-HT$_3$R and the ZAC protein), while MSWHIM predicts high similarity across the entire superfamily except for ACHD, ACHE and ACHG (**Figure 7C**; **Supplementary Figure S16**).

Conformation analysis of the lower TMD reveals that the majority of structures of all superfamily members cluster very tightly for both sites, indicative of very similar conformations irrespective of the presence or absence of ligands in this site, see

**Figure 8**. Additionally, a few structures of nAChRs and 5-HT$_3$ARs adopt a very different conformation with a local RMSD up to 2.6 Å (**Supplementary Figure S17C**). In the case of the lipid associated site, this reflects very different conformations of the TM4 helices in the different experimental structures, resulting in local RMSD differences > 3 Å (**Supplementary Figure S17D**). For this site, the 5-HT$_3$R structures are uniquely different from all others, see **Figure 8B** and **Supplementary Figure S17D**. Surprisingly, the lower TMD ligands seem to not induce any induced fit at all, as the structure with the inhibitory steroid preganolone sulfate in site 5 (5OSC) and all structures with modulatory steroids in sites 4 (5OSB, 6CDU, and 5O8F) superpose very tightly, see **Figure 8C**.
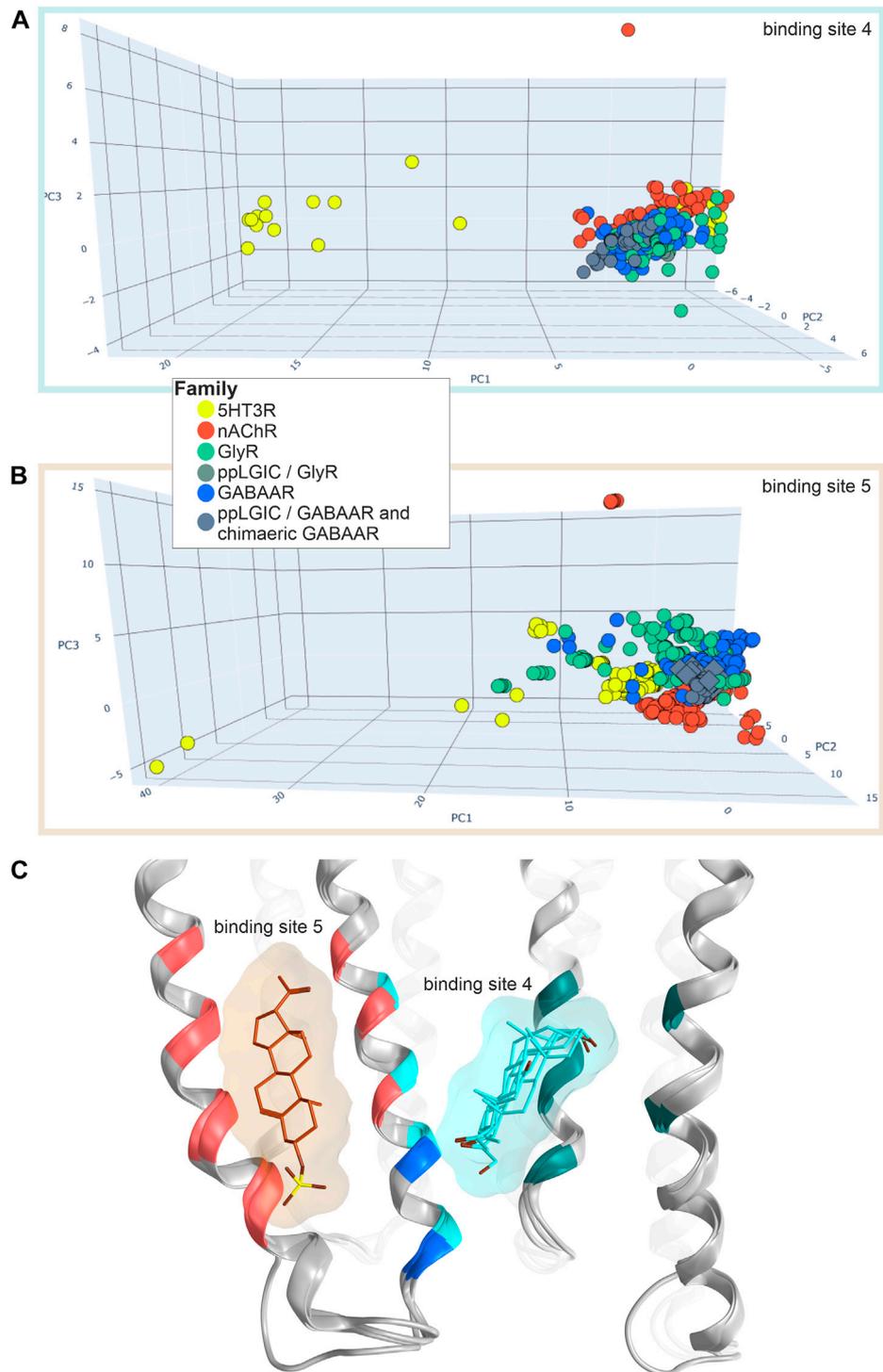
**FIGURE 8 |** Conformation analysis of the lower TMD with sites 4 and 5: **(A,B)** Scatter plots of the pocket conformations. Each dot represents the results for an individual site, diamond shaped symbols indicate the presence of a ligand in the analysed site. The colour of the dots indicates the protein family as defined in the legend, gray is used for diverse chimeras. ppLGIC: procaryotic pentameric ligand gated ion channel. **(C)** Superposition and comparison of 5OSB, 5OSC, 5O8F and 6CDU. Pregnenolone sulfate bound to site 5 is rendered in brown sticks, THDOC, alphaxalone and pregnenolone bound at site 4 in cyan sticks. Ribbon colouring indicates contributions to site 5 (red), 4 (cyan) or both (blue).

## Cross-Reactivity Within Non-Homologous Sites in Individual Receptor Pentamers

An unexpected finding in our pharmacophore screens is the extensive overlap of hits for the novel ECD site 2 and the upper TMD interface sites. This raises the question whether it reflects spurious hits, or hints at a genuine overlap of ligand space shared by non-homologous pockets. A similar situation is well established for GABA$_A$ receptors: Diazepam binds with high affinity at the canonical ECD α+/γ− interfaces of certain receptor subtypes, and additionally with lower affinity at two of the upper TMD sites (Masiulis et al., 2019; Kim et al., 2020).

In this study, we did not include the canonical ECD interface sites, as the large number of distinct interfaces and the multiple segments with INDELS would have exceeded the scope of this paper. Our results, however, suggest that the multi-site usage of ligands as displayed by diazepam is not limited to the sites 1 and 3, but may also affect other combinations of binding sites. For the sites localized in the TMD, structural data is yet not very suggestive of ligand promiscuity between the different sites (3, 4, and 5). In contrast, photoaffinity studies strongly suggest that steroid derivatives have multiple binding sites not limited to the lower TMD, but also in the upper TMD (Chen et al., 2019; Sugasawa et al., 2019). It would be very interesting to find out whether endogenous ligands also interact with multiple sites in the ECD and TMD.

While our limited pharmacophore screen does not suggest steroids as hits for the upper TMD sites, we note that 17 compounds, representing considerable chemical diversity, are predicted with moderate to strong scores to match with the pharmacophore models from all three types of sites. This is the case for dronabinol (Δ$^9$-THC), another example is the 5-HT$_3$R interacting antiemetic metoclopramide (**Supplementary Figures S11, S12**). The latter thus is potentially another example for a compound which interacts with the canonical ECD interface site and additional sites in the TMD, akin to the Bzs. In the screens performed here, we also note that a fraction of the screened Bzs, as well as the Bz-site ligand alpidem, score with 0.7 or higher for all three TMD sites in at least one family member (**Supplementary Figures S11, S12**).

## Assessing Candidate Binding Sites for the Phytocannabinoid Δ$^9$-THC

For phytocannabinoids, the literature demonstrates interactions with multiple family members (Barann et al., 2002; Oz et al., 2004; Yang et al., 2008; Yang et al., 2010; Xiong et al., 2011; Bakas et al., 2017; Schmiedhofer, 2017). Effects differ depending on the protein subtype, as well as the cannabinoid molecule. For example, Δ$^9$-THC is relatively ineffective at inhibiting α7 nAChRs, while inhibiting 5-HT$_3$Rs with similar potency to CB$_1$ receptors (Barann et al., 2002; Oz et al., 2004). On the other hand, CBD and Δ$^9$-THC are potentiating GlyRs and most GABA$_A$R subtypes (Xiong et al., 2011; Schmiedhofer, 2017). One candidate binding site has been delineated by mutational studies: In GlyR α1 and α3 subunits, a site which was proposed to correspond with the later described pregnanolone sulfate site on GABA$_A$ R α1 subunits in the 5OSC structure, was shown to mediate a large fraction of a strong PAM effect elicited by CBD (Xiong et al., 2011). For 5-HT$_3$Rs an allosteric modulatory binding site has been suggested (Barann et al., 2002).

Our PH4 screen indicates matches for Δ$^9$-THC with the TM3/TM4 lipid associated site 5, as well as with most upper TMD interface sites (3), and with the steroid bound lower TMD-interface (site 4) (**Supplementary Figure S12**). In addition, another match occurred for the novel upper ECD interface site. We thus performed an exploratory computational docking screen into all matched sites. In turn, we ruled out the upper ECD interface as it displays multiple clashes, and the docking results cannot recapitulate the overlap of the pharmacophore features at all. This leaves the binding sites in the TMD as candidates. We again turned to the literature to search for evidence on binding sites, and to prioritize docking for family members with known effects. For GABA$_A$Rs, both CBD and Δ$^9$-THC have been investigated in multiple subunit combinations (Bakas et al., 2017; Schmiedhofer, 2017). Both phytocannabinoids elicit similar effects in most GABA$_A$R subunit combinations that were tested, enhancing GABA currents in most of them and reducing currents in β1-containing assemblies (Schmiedhofer, 2017). A mutation in TM3 of the GlyR α3 subunit (Xiong et al., 2011) is localized in a position shared by sites 4 and 5, and in close proximity to site 3 (**Figure 9**).

In order to select subunits for further computational docking, we examined the predicted similarity between GlyR α3 and other candidate subunits. The descriptor analysis of the candidate sites is ambiguous for the upper TMD interface, where the two descriptors place GlyR α3 into different clusters. The placement by MSWHIM with nAChR subunits (**Supplementary Figure S14b**) would be consistent with the strong hit in the 7EKT structure's site 3, and was thus followed up in more detail. For the lower TMD interface site, both descriptors cluster the GlyR α3 subunit with GABA$_A$R subunits (**Supplementary Figure S15**), in line with the hit for this steroid binding site. Interestingly, for the TM3/TM4 lipid associated site, where MSWHIM suggests broad conservation, the z-Scales(3) places GlyR α3 into a cluster with the GABA$_A$R α6 subunit (**Supplementary Figure S16**). This is interesting because the current enhancement by Δ$^9$-THC that was observed in a pilot experimental study is stronger for α6-containing assemblies compared to other α- isoforms (Schmiedhofer, 2017).

In order to derive testable structural hypotheses for pockets with combined experimental and computational evidence, we generated docking poses for site 3 in the nAChR α7 homomer, in the steroid bound site 4 of the GABA$_A$R lower TMD α1+/α1− (original PDBID: 5OSB), and in site 5 in GlyR α3 [to compare with the CBD docking from (Xiong et al., 2011)] as well as in a homology model of GABA$_A$R α6. The docking results for these four cases were then subjected to two scoring functions to obtain poses with highly ranked consensus scores (**Figure 9**).

For all four dockings, multiple binding modes with high scores were obtained with both scoring functions. With additional restraints coming from indirect experimental evidence, poses were identified for site 5 which feature GlyR α3 Ser307 in close proximity to the benzo[C]-chromen-1-ol hydroxy group on the ligand, which is in agreement with the pharmacophore screen as well. In addition, this pose has a high overlap with a pose in the GABA$_A$R α6 model, where both are among the top ten of both scoring functions, see **Figure 9**. For the docking into the site 3 of the nAChR α7 homomer, the top pose of both scoring functions features high overlap with the PAM ligand of the experimental structure, and

| chemPLP score [rank] | chemscore [rank] |
|---|---|
| 63.617 [4] | 27.9862 [8] |

| chemPLP score [rank] | chemscore [rank] |
|---|---|
| 87.2618 [1] | 42.7347 [1] |

| chemPLP score [rank] | chemscore [rank] |
|---|---|
| 61.3012 [7] | 28.7800 [5] |

| chemPLP score [rank] | chemscore [rank] |
|---|---|
| 78.3396 [2] | 33.7103 [1] |

**FIGURE 9 |** Computational docking into candidate binding sites for $\Delta^9$-THC. **(A)** Consensus binding mode for site 5 in the GlyR α3 subunit from 5VDI which is also present in the GABA$_A$R α6 docking in panel **(D)**. **(B)** The highest ranked consensus scored representative binding mode from the docking into site 3 from 7EKT with $\Delta^9$-THC in blue superposed with PNU-120596 in pink. **(C)** All three sites of interest displayed on a representative ribbon structure with GlyR α3 Ser307 indicated by a yellow segment on TM3. **(D)** Docking result of $\Delta^9$-THC into site 5 of a GABA$_A$R α6 homology model. **(E)** Representative binding mode from the docking into 5OSB superposed with THDOC in pink.

thus is in agreement with all available evidence. Lastly, for the results obtained for site 4, the best ranked binding mode overlaps very closely with the ligand from the experimental structure (**Figure 9E**). An alternative binding mode brings the ligand closer to the TM3 of the principal subunit and the site of mutagenesis (**Supplementary Figure S20**) (Xiong et al., 2011). These structural hypotheses can be used to guide further mutational studies, see **Supplementary Figure S21**.

# DISCUSSION

pLGICs mediate drug induced seizures and convulsions as well as a broad range of neuropsychiatric adverse effects impacting on sleep, vigilance, mood, memory, autonomic NS functions, and many more, including indirect effects on peripheral nerve regeneration. The NeuroDeRisk project (https://neuroderisk. eu/) aims to improve *in silico* methods for the efficient prediction of such liabilities. The need for such methods is underscored by the large number of allosteric binding sites present on pLGICs, where the development of *in vitro* screening tools such as specific high affinity radioligands is lagging behind. We thus took advantage of the recent surge in structural data and compiled an inventory of allosteric sites for further study.

A multitude of molecules and drugs exist that target several family members. Histamine is a positive modulator of some GABA$_A$R subtypes, whereas it was shown to inhibit GlyRs currents with low potency (Breitinger and Breitinger, 2020). Similarly, neuroactive steroids are potent positive or negative allosteric modulators of GABA$_A$Rs, while in contrast inhibiting glycine elicited chloride current (Breitinger and Breitinger, 2020). The non-steroidal anti-inflammatory drug niflumic acid, which also binds to GABA$_A$Rs where it acts as a positive or negative modulator depending on the subtype (Rossokhin et al., 2019), inhibits GlyRs (Breitinger and Breitinger, 2020).

Apart from their well-recognized use as antiemetics, preclinical studies have suggested a possible therapeutic use of 5-HT$_3$ receptor antagonists in depression, substance abuse, cognitive and psychiatric disorders, and pain (Costall and Naylor, 2004; Thompson and Lummis, 2007; Rajkumar and Mahesh, 2010; Walstab et al., 2010). However, clinical studies have so far failed to substantiate the use of 5-HT$_3$ receptor antagonists in the treatment of CNS disorders (Greenshaw and Silverstone, 1997). Interestingly, the 5-HT$_3$ inhibiting antiemetic therapeutic tropisetron potentiates glycine elicited effects of some GlyR subtypes (Breitinger and Breitinger, 2020). The literature thus provides many examples of largely unwanted promiscuity of ligands across the pLGIC families, resulting in a need to study it systematically and to develop *in vitro* tools for ligand interaction sites which are highly conserved among the pentameric ligand gated ion channel superfamily.

Several methods are available to perform computationally inexpensive, and thus high throughput, *in silico* predictions for multiple aspects of ligand-protein interactions and the properties and target space of a binding site with known structure. Pharmacophore screening has a long and successful history and serves well to rank libraries of test compounds as potential hits for a site which can be characterized by mandatory and optional interaction features. By itself, the method has known limitations (Kaserer et al., 2015): The stringency of the screen needs to be carefully optimized for each application in order to optimize enrichment. Here, we reasoned that rapid, simple screens with moderate optimization can generate useful hypotheses, and that the combination of multiple *in silico* methods that cover different aspects of the protein-ligand interaction theme can be combined to produce ranked toxicological alerts for more elaborate follow-up investigations. In this vein, less strict screenings and accepting false positive hits is a compromise that shifts a more costly problem of false negatives into a less costly acceptance of true hits mixed with false positives in the "early alert" stage.

If structural information is available in sufficient amount and quality, *in silico* methods that result in simple models of the binding sites can be added to the workflow and further prioritize the experimental follow-up work. A large protein superfamily, such as the pLGICs with 45 human subunits that comprise four families of transmitter receptors, is inherently prone to issues with polypharmacology. The efforts that would be needed for systematic *in vitro* testing of drug candidates on all proteins in the superfamily are not feasible. Given enough experimental structures, ligand binding pockets can be characterized in terms of their structural variability. Currently, available data shows that binding sites present in the ECD of pLGICs are composed of a mixture of conserved and variable protein segments, and thus ligand promiscuity is theoretically limited to well-defined groups of family members. In contrast, binding sites in the TMD share high structural conservation and differ only in sidechains and a few very minor INDEL bearing segments. This is an ideal scenario for the use of methods which rest on the properties which amino acid sidechains contribute to ligand binding pockets independent of their 3D arrangement. Various descriptors have been developed and are widely used to predict properties of peptides, but the concept is increasingly also applied to binding sites (van Westen et al., 2013). The application of descriptors to binding site forming amino acids at the canonical ECD interface site and the upper TMD "anaesthetics" sites of GABA$_A$Rs demonstrates that the gross pharmacology is recapitulated, but also that the method is very sensitive to the selected amino acids. If used to predict shared ligand space of a pocket, it is essential to use strictly those amino acids that contribute with sidechains to the pocket in question, which is a considerable limitation for practical applications. Thus, for its routine use, automated extraction of binding site forming sidechains and separate predictions for smaller and larger ligands would still need to be developed and benchmarked. Here, we used the results to guide the selection of candidate pockets for the exploratory computational docking that was performed for $\Delta^9$-THC.

Prior to embarking on computationally more expensive methods such as more elaborate computational docking and different types of molecular dynamics simulations, it pays off

to select among the experimental structures those that are suited to address a given scientific question. One of the inherent limitations is the static nature of methods such as cryo-EM and X-ray crystallography, which captures proteins in long-lived conformational states. The automated analysis of protein conformations which was employed here provides a bias-free approach to compare the pocket of interest across all proteins in which the pocket is found, independent of absence or presence of ligands, and independent of global protein conformations (such as closed- desensitized). An interesting finding along these lines was the dominant influence of ECD ligands on the conformations of TMD binding sites, and even more strikingly, the apparent absence of steroid impact on the TMD conformations (see **Figures 5**, **8**). This underscores the importance to consider all available structures as templates for high throughput methods such as structure based pharmacophore screening, and the inherent limitations of static structural snapshots. The increase in computational performance we currently witness may soon allow certain MD methods to be integrated into high throughput workflows, but at this time, MD still cannot be termed a fast computational technique.

Evidence accumulated that ligands display two types of distinct cross-reactivity: Interactions with homologous pockets of other family members, and interactions with non-homologous pockets in the same receptor. Benzodiazepines, and likely many other compounds interact with multiple sites in the same receptor. In the case of Bz-site ligands, the differences in affinity seem to be high, but data still is scarce (Iorio et al., 2020). While the small scale pharmacophore screen we performed here has clear limitations due to high heterogeneity in the resolution and number of available structures per binding site and the small size of the library, as well as lacking decoys, it does recapitulate known findings well, and it does suggest that the ligand promiscuity as depicted by the data at DrugCentral is severely underestimated (**Supplementary Figure S13**).

Many of the screened compounds are seen to be moderately or highly scored hits at two and more binding sites, including dronabinol (the trade name of a specific $\Delta^9$-THC formulation). Since the literature confirms that it has functional effects on all pLGIC families (Barann et al., 2002; Oz et al., 2004; Yang et al., 2008; Yang et al., 2010; Xiong et al., 2011; Bakas et al., 2017; Schmiedhofer, 2017), we explored these hits in more detail. The computational docking results are consistent with most of the high scoring pharmacophore hits, leaving us with three candidate interaction sites for this phytocannabinoid in the pLGICs—the upper and lower TMD interface sites, and the TM3/TM4 lipid associated site. The mutational work which was performed for CBD in previous studies (Xiong et al., 2011; Bakas et al., 2017) probed an amino acid which is localized in a position in close proximity to all three of these sites, thus, further mutational work should be structure guided and probe amino acids that are unique to the individual pockets (see **Supplementary Figure S21**). The published functional studies show a mix of enhancement or reduction of agonist mediated effects (Barann et al., 2002; Oz et al., 2004; Yang et al., 2008; Yang et al., 2010; Xiong et al., 2011; Bakas et al., 2017; Schmiedhofer, 2017). Of note, in GABA$_A$R a limited pilot study observed the net effect to depend on the isoform of α and β subunits (Schmiedhofer, 2017). This is more compatible with a site of action at a subunit interface, or with the use of multiple binding sites in the same receptor. The GABA$_A$R β isoforms are known for a single amino acid difference in the upper TMD interface complementary component, but also feature different amino acids in the TM3/TM4 segments that contribute to an endocannabinoid site in the β2 subunit (Sigel et al., 2011). Thus, further experimental studies are needed to clarify the mechanism of action by which $\Delta^9$-THC modulates 5-HT$_3$R, nAChR, GlyR and GABA$_A$Rs. Future studies can be guided by the bound state models presented here.

The panel of methods employed in this study, which we benchmarked with a small library of drugs and selected showcases, can deliver useful toxicological alerts with inexpensive *in silico* toolchains that are easy to implement with many available software packages. For proteins with many allosteric sites, multimodal computational screening can help to prioritize the slow and expensive functional studies that need to be performed when radioligands are lacking.

## DATA AVAILABILITY STATEMENT

The conformational analysis scripts can be found under https://github.com/FilipKon/CoDiP. To see only the results (3D scatter plots) and distances with the corresponding amino acids, visit https://fkoniuszewski.pythonanywhere.com/. The scripts used to compare the protein sequences and the variant analysis are available at https://github.com/JureFabjan/GABR_genetics. Other data is available upon request to the corresponding author.

## AUTHOR CONTRIBUTIONS

FK, ME, KB, FV, TS and PBS contributed to writing. FK, JF, TS and PBS performed computational work. FK, JF, TS and FV analyzed data. FK and JF prepared figures. ME and TL supervised work.

## FUNDING

# ACKNOWLEDGMENTS

# SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.860246/full#supplementary-material

**Supplementary Figure S1 |** Superposed structures and their pockets derived from remote homologues. Chlorpromazine from PDB ID 5LG3 (violet), isoflurane/propofol from PDB ID 4Z90/5MUO (green), ketamine from PDB ID 4F8H (red), bromoethanol from PDB ID 5SXV (blue), memantine from PDB ID 4TWD (pink).

**Supplementary Figure S2 |** Total number of alleles per variant in each gene of interest. On the x-axis the respective genes of interest are listed; the y-axis reflects the total number of alleles per variant. Each point represents a variant. Red colored variants were excluded from the analysis since they did not pass the outlier test (more than 3 standard deviations away from the mean of total number of alleles). The retained variants (less than 3 standard deviations away from the mean) are colored in gray.

**Supplementary Figure S3 |** Structure-based alignment of the N-terminally truncated ECDs of a GABAA β3 subunit, a GlyR α3 subunit, a nAChR α7 subunit, a 5-HT3R A subunit, a nAChR α3 subunit and a nAChR β4 subunit in this order to visualize the variable regions which are schematically depicted on the topology diagram in **Supplementary Figure S4**.

**Supplementary Figure S4 |** Topology diagram of cys loop receptor extracellular domain (ECD): The ECD features a modified immunoglobulin-like fold with a twisted beta-sandwich. The topology is as indicated in the image and comprises ten strands (some are broken in many structures) connected as greek key motif. Between conserved blocks, that mainly form the packing core of the domain, multiple variable regions are located (VR1-7) with and without INDELS. Notably, not all variable regions are localized in loop regions, but are in part also interspersed in strands. Strand 9 has a short INDEL. The image depicts the variable regions resulting from a structural comparison of all families, i.e., GABA-A, GlyR, nAChR, and 5HT3-R subunits. Within families, fewer variable regions occur. A representative alignment derived from a structural superposition generated with PDBe Fold is shown in **Supplementary Figure S3**.

**Supplementary Figure S5 |** 3D scatter plots of z-Scale(3) and MSWHIM results for the selected subunit ECD interface binding sites in which the principal and complementary sites were merged to a full pocket amino acid list without the loop F. Both descriptors predict a clear separation between the GABA binding sites and the benzodiazepine binding sites. The diazepam sensitive (DS) sites and the diazepam insensitive (DI) sites, which are formed by different principal subunits, also separate, as indicated on the images. In a next step, pairwise euclidean distances can be used for hierarchical clustering to extract from the descriptor results to clusters (**Supplementary Figures S6,S7**).

**Supplementary Figure S6 |** z-Scale(3) **(A)** and MSWHIM **(B)** results of the hierarchical clustering for all analyzed ECD interfaces in which the principal and complementary sites were merged to a full pocket amino acid list without loop F.

**Supplementary Figure S7 |** z-Scale(3) **(A)** and MSWHIM **(B)** results of the hierarchical clustering for all analyzed ECD interfaces in which the principal and complementary sites were merged to a full pocket amino acid list, with amino acids on loop F. Note the different results obtained with loop F amino acids, suggesting that smaller and larger pockets need to be analyzed separately.

**Supplementary Figure S8 | (A)** Amount of variants across all genes of interest in the dimer onto 7EKT. Ribbon heatmap represents the sum of minor allele counts per 100,000 at each aligned position. **(B)** Variability of variable regions in individual genes. (a) Sum of minor allele numbers per 100,000 in individual variable regions on a logarithmic scale. The ECD variable regions are displayed schematically in **Supplementary Figure S4**. If not stated otherwise, a variable region is defined to contain insertions or deletions of two or more amino acids.

(b) Number of variants divided by the length of the variable region in major isoforms of individual genes.

**Supplementary Figure S9 |** % identity and % similarity for the whole TMD for all pentameric ligand-gated ion channels. The alignment which was used for the calculation is shown in **Supplementary Item S4**.

**Supplementary Figure S10 |** Summary of molecules investigated in pharmacophore screenings. As described in the methods, all drugs for which positive bioactivity at any of the pLGIC families is reported at DrugCentral were used as a small library which represents an incomplete, but representative snapshot of the chemical space comprising approved drugs with desired or off-target interactions with these receptors, irrespective of subtypes or isoforms. Blue tiles = GABAA receptor targeting; green = 5HT3 targeting; red = nACh receptor targeting; purple = glycine receptor targeting; targeting; gray = multi targeting substances. Note that this gallery reflects strictly the data at Drug Central, while for many of these compounds more promiscuity within the superfamily is known already.

**Supplementary Figure S11 |** Heat map of pharmacophore predicted drug-target interaction for the novel glycine receptor site (2), the five screens into the upper TMD pockets (site 3), the lower TMD interface site (4) and the lipid TM3/TM4 associated site bound with pregnanolone sulfate (5). For a detailed description of the generation of pharmacophores, see the *Methods* section of the manuscript. Hits = number of predicted drug-target interactions for the respective site according to the matched pharmacophore model. The colors represent score values ranging from 0 (light blue) to 1 (dark blue), reflecting the number and quality of overlap of matching features between the respective ligand and the pharmacophore model.

**Supplementary Figure S12 |** Summary of the five screens into upper TMD pockets. Chord diagram of pharmacophore predicted drug-target interactions within the upper TMD binding pocket. For this site most structural templates exist in the pLGIC family (PDB IDs: 5VDH (GlyR- α3 homopentamer), 6HUP (GABAA receptor - α1β3γ2L), 6X3T (GABAA receptor - α1β2γ2) , 6X3V (GABAA receptor - α1β2γ2), 6X3W (GABAA receptor - α1β2γ2), 6X3X (GABAA receptor - α1β2γ2), 7EKT (nAChR - α7 homopentamer). The whole substance library was screened into five merged pharmacophores, indicated by the colored bars. For details regarding the generation of pharmacophores, see *Methods* section. Every substance is connected to the respective pharmacophore models for which a hit is predicted. Note, the number of hits ranges from 18 to 119, showcasing that pharmacophore models based on homologous, but not identical pockets between of the pLGIC family members are sensitive enough to differ between them.

**Supplementary Figure S13 |** Total counts of drugs (y-axis) associated with the target families as reflected by DrugCentral only (blue), and with the addition of our pharmacophore screening results (red). The pharmacophore screening reflects only those binding sites for which structural data was available and for which screens were performed with the exception of the screen into the ivermectin site (**Supplementary Figure S11**).

**Supplementary Figure S14 | (A)** Descriptor set results of the hierarchical clustering for the z-Scale(3) and MSWHIM for the GABAARs full pockets of site 3. **(B)** Descriptor set results of the hierarchical clustering for the z-Scale(3) and MSWHIM for the site 3. The numbers on the branching points refer to the comparison diagrams in **Figure 4**. **(C)** Descriptor set results of the hierarchical clustering for the z-Scale(3) and MSWHIM for the ivermectin binding site. The numbers on the branching points refer to the comparison diagrams in **Supplementary Figure S14D**. **(D)** Comparison diagram for the ivermectin site in the upper TMD. The left half of the diagram reflects the z- Scale (3) results, and the right half the MSWHIM results.

**Supplementary Figure S15 |** Descriptor set results of the hierarchical clustering for the z-Scale(3) and MSWHIM for the site 4. The numbers on the branching points refer to the comparison diagrams in **Figure 6**.

**Supplementary Figure S16 |** Descriptor set results of the hierarchical clustering for the z-Scale(3) and MSWHIM for the site 5. The numbers on the branching points refer to the comparison diagrams in **Figure 7**.

**Supplementary Figure S17 | (A)** Multiple views of the 3d scatter plot from **Figure 5** with some conformational differences of three different receptor subtypes from the nicotinic acetylcholine receptors (red structures) and a phenobarbital bound site 3 of a GABAA receptor (blue). The numbers in the scatter plots identify the structures used for the RMSD matrix, reflecting representative conformations. Specific examples are displayed in the same style as in **Figure 5**: Different conformations

result in different distances between alpha carbon atoms: A plane through the upper TMD is depicted with all five TM2 segments and two subunits. Distances between alpha carbons are shown to illustrate backbone conformation, where yellow lines are distances between homologous pore forming TM2 residues as an indication for asymmetry, purple lines are distances between pocket forming residues defining the circumference, and green lines the pocket "diagonals." **(B)** Example of two different conformations of 5HT3A. **(C)** Scatter plots of the site 4 conformation analysis matching **Figure 8**, with representative RMSD values from the labeled structures. **(D)** Scatter plots of the site 5 conformation analysis matching **Figure 8**, with representative RMSD values from the labeled structures.

**Supplementary Figure S18 |** Minor allele frequency of variants in TMD subregions. **(A)** Heatmap representing the upper TMD binding site. **(B)** Heatmap representing the extension of the upper TMD binding site with amino acids involved in binding of ivermectin. **(C)** Heatmap representing the steroid binding site. **(D)** Heatmap representing amino acids on TM2 helix oriented with side chains towards the pore. Heatmaps show the sum of minor allele counts per 100,000 in gene and amino acid combinations. Amino acids are numbered according to the reference alignment (**Supplementary Item S3**). The amino acids correspond to those which are color coded in **Supplementary Figure S18** for the upper and lower TMD binding sites, the PS-site.

**Supplementary Figure S19 |** Number of different variants in TMD subregions. **(A)** Heatmap representing the upper TMD binding site. **(B)** Heatmap representing the extension of the upper TMD binding site with amino acids involved in binding of ivermectin. **(C)** Heatmap representing the steroid binding site. **(D)** Heatmap representing amino acids on TM2 helix oriented with side chains towards the pore. Amino acids are numbered according to the reference alignment.

**Supplementary Figure S20 |** An alternative binding mode for the docking of THC into 5OSC showing close proximity to the glycine α3.

**Supplementary Figure S21 |** Tentative THC binding sites and candidate positions for mutational analysis. Ribbon rendering of two subunits' TMD domains, helices that contribute to sites 3, 4, or 5 are labeled. Colour codes match **Figure 9**. Site 3 in the upper TMD: Representative docking pose of THC (green) in superposition with PNU-120596 (light pink, 7EKT). Green ribbon segments on principal TM2, TM3 and on complementary TM1 form ligand contacts. The segment numbering matches the table below the image, amino acid numbering as in (PDBID - 5OSB) for all sites. Site 4: Representative docking pose of THC (cyan) in superposition with THDOC (light pink, 5OSB). Cyan ribbon segments on principal TM3 and complementary TM4 form ligand contacts. Site 5: Representative docking pose of THC (gold) in (partial) superposition with preganolone sulfate (PS, light pink, 5OSC). The best scoring docking poses overlap only to a small extent with PS. Golden ribbon segments on TM3 and TM4 form ligand contacts. The bright yellow segment with the * marks the position which has been mutated in the GlyR α3 subunit, see **Figure 9**.

**Supplementary Item S1 |** This is an Excel file with a comprehensive list of PDB files of pLGICs, those used for this study and many more.

**Supplementary Item S2 |** List of drugs which were used for the structure-based pharmacophore screening.

**Supplementary Item S3 |** FASTA formatted file containing the aligned sequences used in variant analysis.

**Supplementary Item S4 |** This file contains a color coded alignment reflecting the amino acids used for the TMD binding sites.

**Supplementary Table S1 |** Gene names and their corresponding uniprot IDs.

**Supplementary Table S2 |** Structures used for the pharmacophore screens.

# REFERENCES

Albrecht, B., Staiger, P. K., Hall, K., Miller, P., Best, D., and Lubman, D. I. (2014). Benzodiazepine Use and Aggressive Behaviour: a Systematic Review. *Aust. N. Z. J. Psychiatry* 48 (12), 1096–1114. doi:10.1177/0004867414548902

Bakas, T., van Nieuwenhuijzen, P. S., Devenish, S. O., McGregor, I. S., Arnold, J. C., and Chebib, M. (2017). The Direct Actions of Cannabidiol and 2-arachidonoyl Glycerol at GABA A Receptors. *Pharmacol. Res.* 119, 358–370. doi:10.1016/j.phrs.2017.02.022

Bampali, K., Koniuszewski, F., Silva, L., Rehman, S., Vogel, F., Seidel, T., et al. (2022). Tricyclic Antipsychotics and Antidepressants Can Inhibit α5-containing GABAA Receptors by Two Distinct Mechanisms. *Authorea.* doi:10.22541/au.161797799.93015888/v2

Barann, M., Molderings, G., Brüss, M., Bönisch, H., Urban, B. W., and Göthert, M. (2002). Direct Inhibition by Cannabinoids of Human 5-HT3A Receptors: Probable Involvement of an Allosteric Modulatory Site. *Br. J. Pharmacol.* 137 (5), 589–596. doi:10.1038/sj.bjp.0704829

Barrera, N. P., Herbert, P., Henderson, R. M., Martin, I. L., and Edwardson, J. M. (2005). Atomic Force Microscopy Reveals the Stoichiometry and Subunit Arrangement of 5-HT 3 Receptors. *Proc. Natl. Acad. Sci. U.S.A.* 102 (35), 12595–12600. doi:10.1073/pnas.0503253102

Basak, S., Gicheru, Y., Rao, S., Sansom, M. S. P., and Chakrapani, S. (2018a). Cryo-EM Reveals Two Distinct Serotonin-Bound Conformations of Full-Length 5-HT3A Receptor. *Nature* 563 (7730), 270–274. doi:10.1038/s41586-018-0660-7

Basak, S., Gicheru, Y., Samanta, A., Molugu, S. K., Huang, W., Fuente, M. l. d., et al. (2018b). Cryo-EM Structure of 5-HT3A Receptor in its Resting Conformation. *Nat. Commun.* 9 (1), 514. doi:10.1038/s41467-018-02997-4

Basak, S., Kumar, A., Ramsey, S., Gibbs, E., Kapoor, A., Filizola, M., et al. (2020). High-resolution Structures of Multiple 5-HT3AR-setron Complexes Reveal a Novel Mechanism of Competitive Inhibition. *eLife* 9, e57870. doi:10.7554/eLife.57870

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., et al. (2000). The Protein Data Bank. *Nucleic Acids Res.* 28 (1), 235–242. doi:10.1093/nar/28.1.235

Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinl, T., et al. (2008). "KNIME: The Konstanz Information Miner,". Editors C. Preisach, H. Burkhardt, L. Schmidt-Thieme, and R. Decker (Springer Berlin Heidelberg), 319–326.()*Data Anal. Machine Learn. Appl.*

Blanco, M. J., La, D., Coughlin, Q., Newman, C. A., Griffin, A. M., Harrison, B. L., et al. (2018). Breakthroughs in Neuroactive Steroid Drug Discovery. *Bioorg. Med. Chem. Lett.* 28 (2), 61–70. doi:10.1016/j.bmcl.2017.11.043

Bond, A. J. (1998). Drug-Induced Behavioural Disinhibition. *CNS Drugs* 9 (1), 41–57. doi:10.2165/00023210-199809010-00005

Brady, C. A., Stanford, I. M., Ali, I., Lin, L., Williams, J. M., E. Dubin, A., et al. (2001). Pharmacological Comparison of Human Homomeric 5-HT3A Receptors versus Heteromeric 5-HT3A/3B Receptors. *Neuropharmacology* 41 (2), 282–284. doi:10.1016/s0028-3908(01)00074-0

Breitinger, U., and Breitinger, H.-G. (2020). Modulators of the Inhibitory Glycine Receptor. *ACS Chem. Neurosci.* 11 (12), 1706–1725. doi:10.1021/acschemneuro.0c00054

Brejc, K., van Dijk, W. J., Klaassen, R. V., Schuurmans, M., van Der Oost, J., Smit, A. B., et al. (2001). Crystal Structure of an ACh-Binding Protein Reveals the Ligand-Binding Domain of Nicotinic Receptors. *Nature* 411 (6835), 269–276. doi:10.1038/35077011

Brickley, S. G., and Mody, I. (2012). Extrasynaptic GABAA Receptors: Their Function in the CNS and Implications for Disease. *Neuron* 73 (1), 23–34. doi:10.1016/j.neuron.2011.12.012

Brüss, M., Barann, M., Hayer-Zillgen, M., Eucker, T., Göthert, M., and Bönisch, H. (2000). Modified 5-HT3A Receptor Function by Co-expression of Alternatively Spliced Human 5-HT3A Receptor Isoforms. *Naunyn-schmied Arch. Pharmacol.* 362 (4-5), 392–401. doi:10.1007/s002100000342

Caton, M., Ochoa, E. L. M., and Barrantes, F. J. (2020). The Role of Nicotinic Cholinergic Neurotransmission in Delusional Thinking. *npj Schizophr* 6 (1), 16. doi:10.1038/s41537-020-0105-9

Chatzidaki, A., and Millar, N. S. (2015). Allosteric Modulation of Nicotinic Acetylcholine Receptors. *Biochem. Pharmacol.* 97 (4), 408–417. doi:10.1016/j.bcp.2015.07.028

Chen, Q., Wells, M. M., Arjunan, P., Tillman, T. S., Cohen, A. E., Xu, Y., et al. (2018). Structural Basis of Neurosteroid Anesthetic Action on GABAA Receptors. *Nat. Commun.* 9 (1), 3972. doi:10.1038/s41467-018-06361-4

Chen, Z. W., Bracamontes, J. R., Budelier, M. M., Germann, A. L., Shin, D. J., Kathiresan, K., et al. (2019). Multiple Functional Neurosteroid Binding Sites on Gabaa Receptors. *Plos Biol.* 17 (3), e3000157. doi:10.1371/journal.pbio.3000157

Chopra, D. A., Shah, A. B., Vadhariya, A. H., and Painter, J. T. (2019). The Risk of Varenicline-Induced Seizure Among Those Who Have Attempted to Quit Smoking Using Pharmacotherapy. *Epilepsy Behav.* 97, 169–173. doi:10.1016/j.yebeh.2019.05.031

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., et al. (2009). Biopython: Freely Available Python Tools for Computational Molecular Biology and Bioinformatics. *Bioinformatics* 25 (11), 1422–1423. doi:10.1093/bioinformatics/btp163

Costall, B., and Naylor, R. (2004). 5-HT3 Receptors. *Cdtcnsnd* 3 (1), 27–37. doi:10.2174/1568007043482624

Das, P., and Dillon, G. H. (2005). Molecular Determinants of Picrotoxin Inhibition of 5-hydroxytryptamine Type 3 Receptors. *J. Pharmacol. Exp. Ther.* 314 (1), 320–328. doi:10.1124/jpet.104.080325

Das, P., and Dillon, G. H. (2003). The 5-HT3B Subunit Confers Reduced Sensitivity to Picrotoxin when Co-expressed with the 5-HT3A Receptor. *Mol. Brain Res.* 119 (2), 207–212. doi:10.1016/j.molbrainres.2003.09.003

Davies, P. A., Wang, W., Hales, T. G., and Kirkness, E. F. (2003). A Novel Class of Ligand-Gated Ion Channel Is Activated by Zn2+. *J. Biol. Chem.* 278 (2), 712–717. doi:10.1074/jbc.m208814200

Dineley, K. T., Pandya, A. A., and Yakel, J. L. (2015). Nicotinic ACh Receptors as Therapeutic Targets in CNS Disorders. *Trends Pharmacological Sciences* 36 (2), 96–108. doi:10.1016/j.tips.2014.12.002

Erken, H. A., Erken, G., Şimşek, H., Korkut, O., Koç, E. R., Yavuz, Ö., et al. (2014). Single Dose Varenicline May Trigger Epileptic Activity. *Neurol. Sci.* 35 (11), 1807–1812. doi:10.1007/s10072-014-1845-y

Erkkila, B. E., Weiss, D. S., and Wotring, V. E. (2004). Picrotoxin-mediated Antagonism of α3β4 and α7 Acetylcholine Receptors. *Neuroreport* 15 (12), 1969–1973. doi:10.1097/00001756-200408260-00027

Gharpure, A., Teng, J., Zhuang, Y., Noviello, C. M., Walsh, R. M., Jr., Cabuco, R., et al. (2019). Agonist Selectivity and Ion Permeation in the α3β4 Ganglionic Nicotinic Receptor. *Neuron* 104 (3), 501–511. e506. doi:10.1016/j.neuron.2019.07.030

Gotti, C., Moretti, M., Gaimarri, A., Zanardi, A., Clementi, F., and Zoli, M. (2007). Heterogeneity and Complexity of Native Brain Nicotinic Receptors. *Biochem. Pharmacol.* 74 (8), 1102–1111. doi:10.1016/j.bcp.2007.05.023

Gotti, C., Zoli, M., and Clementi, F. (2006). Brain Nicotinic Acetylcholine Receptors: Native Subtypes and Their Relevance. *Trends Pharmacol. Sci.* 27 (9), 482–491. doi:10.1016/j.tips.2006.07.004

Greenshaw, A. J., and Silverstone, P. H. (1997). The Non-antiemetic Uses of Serotonin 5-HT3 Receptor Antagonists. *Drugs* 53 (1), 20–39. doi:10.2165/00003495-199753010-00003

Hainsworth, J. D. (2014). "Nausea and Vomiting," in *Abeloff's Clinical Oncology*. Editors J. E. Niederhuber, J. O. Armitage, J. H. Doroshow, M. B. Kastan, and J. E. Tepper. Fifth Edition (Philadelphia: Churchill Livingstone), 626–634. e622. doi:10.1016/b978-1-4557-2865-7.00042-4

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., et al. (2020). Array Programming with NumPy. *Nature* 585 (7825), 357–362. doi:10.1038/s41586-020-2649-2

Ho, K. K., and Flood, P. (2004). Single Amino Acid Residue in the Extracellular Portion of Transmembrane Segment 2 in the Nicotinic α7 Acetylcholine Receptor Modulates Sensitivity to Ketamine. *Anesthesiology* 100 (3), 657–662. doi:10.1097/00000542-200403000-00028

Holbrook, J. D., Gill, C. H., Zebda, N., Spencer, J. P., Leyland, R., Rance, K. H., et al. (2009). Characterisation of 5-HT3C, 5-HT3D and 5-HT3E receptor Subunits: Evolution, Distribution and Function. *J. Neurochem.* 108 (2), 384–396. doi:10.1111/j.1471-4159.2008.05775.x

Horenstein, N. A., and Papke, R. L. (2017). Anti-inflammatory Silent Agonists. *ACS Med. Chem. Lett.* 8 (10), 989–991. doi:10.1021/acsmedchemlett.7b00368

Houtani, T., Munemoto, Y., Kase, M., Sakuma, S., Tsutsumi, T., and Sugimoto, T. (2005). Cloning and Expression of Ligand-Gated Ion-Channel Receptor L2 in central Nervous System. *Biochem. Biophysical Res. Commun.* 335 (2), 277–285. doi:10.1016/j.bbrc.2005.07.079

Huang, X., Shaffer, P. L., Ayube, S., Bregman, H., Chen, H., Lehto, S. G., et al. (2017). Crystal Structures of Human glycine Receptor α3 Bound to a Novel Class of Analgesic Potentiators. *Nat. Struct. Mol. Biol.* 24 (2), 108–113. doi:10.1038/nsmb.3329

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* 9 (3), 90–95. doi:10.1109/MCSE.2007.55

Inc, C.C.G. (2016). *Molecular Operating Environment (MOE)*. Montréal, QC, Canada: Chemical Computing Group Inc.

Plotly Technologies Inc. (2015). *Collaborative Data Science*. Montréal, QC, Canada: Plotly Technologies Inc.

Iorio, M. T., Vogel, F. D., Koniuszewski, F., Scholze, P., Rehman, S., Simeone, X., et al. (2020). GABAA Receptor Ligands Often Interact with Binding Sites in the Transmembrane Domain and in the Extracellular Domain-Can the Promiscuity Code Be Cracked? *Ijms* 21 (1), 334. doi:10.3390/ijms21010334

Jensen, A., Davies, P., Braunerosborne, H., and Krzywkowski, K. (2008). 3B but Which 3B? and That's Just One of the Questions: the Heterogeneity of Human 5-HT3 Receptors. *Trends Pharmacol. Sci.* 29 (9), 437–444. doi:10.1016/j.tips.2008.06.001

Johnston, G. A. R., Hanrahan, J. R., Chebib, M., Duke, R. K., and Mewett, K. N. (2006). Modulation of Ionotropic GABA Receptors by Natural Products of Plant Origin. *Adv. Pharmacol.* 54, 285–316. doi:10.1016/s1054-3589(06)54012-8

Jones, G., Willett, P., Glen, R. C., Leach, A. R., and Taylor, R. (1997). Development and Validation of a Genetic Algorithm for Flexible Docking 1 1Edited by F. E. Cohen. *J. Mol. Biol.* 267 (3), 727–748. doi:10.1006/jmbi.1996.0897

Kasaragod, V. B., Mortensen, M., Hardwick, S. W., Wahid, A. A., Dorovykh, V., Chirgadze, D. Y., et al. (2022). Mechanisms of Inhibition and Activation of Extrasynaptic αβ GABAA Receptors. *Nature* 602 (7897), 529–533. doi:10.1038/s41586-022-04402-z

Kaserer, T., Beck, K., Akram, M., Odermatt, A., and Schuster, D. (2015). Pharmacophore Models and Pharmacophore-Based Virtual Screening: Concepts and Applications Exemplified on Hydroxysteroid Dehydrogenases. *Molecules* 20 (12), 22799–22832. doi:10.3390/molecules201219880

Kim, J. J., Gharpure, A., Teng, J., Zhuang, Y., Howard, R. J., Zhu, S., et al. (2020). Shared Structural Mechanisms of General Anaesthetics and Benzodiazepines. *Nature* 585 (7824), 303–308. doi:10.1038/s41586-020-2654-5

Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., et al. (2016). PubChem Substance and Compound Databases. *Nucleic Acids Res.* 44 (D1), D1202–D1213. doi:10.1093/nar/gkv951

Krissinel, E., and Henrick, K. (2004). Secondary-structure Matching (SSM), a New Tool for Fast Protein Structure Alignment in Three Dimensions. *Acta Crystallogr. D Biol. Cryst.* 60 (Pt 12 Pt 1), 2256–2268. doi:10.1107/s0907444904026460

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X Version 2.0. *Bioinformatics* 23 (21), 2947–2948. doi:10.1093/bioinformatics/btm404

Laverty, D., Desai, R., Uchański, T., Masiulis, S., Stec, W. J., Malinauskas, T., et al. (2019). Cryo-EM Structure of the Human α1β3γ2 GABAA Receptor in a Lipid Bilayer. *Nature* 565, 516–520. doi:10.1038/s41586-018-0833-4

Laverty, D., Thomas, P., Field, M., Andersen, O. J., Gold, M. G., Biggin, P. C., et al. (2017). Crystal Structures of a GABAA-Receptor Chimera Reveal New Endogenous Neurosteroid-Binding Sites. *Nat. Struct. Mol. Biol.* 24, 977–985. doi:10.1038/nsmb.3477

Lecker, I., Wang, D.-S., Romaschin, A. D., Peterson, M., Mazer, C. D., and Orser, B. A. (2012). Tranexamic Acid Concentrations Associated with Human Seizures Inhibit glycine Receptors. *J. Clin. Invest.* 122 (12), 4654–4666. doi:10.1172/JCI63375

Lochner, M., and Lummis, S. C. R. (2010). Agonists and Antagonists Bind to an A-A Interface in the Heteromeric 5-HT3AB Receptor. *Biophysical J.* 98 (8), 1494–1502. doi:10.1016/j.bpj.2009.12.4313

Lynagh, T., and Lynch, J. W. (2010). An Improved Ivermectin-Activated Chloride Channel Receptor for Inhibiting Electrical Activity in Defined Neuronal Populations. *J. Biol. Chem.* 285 (20), 14890–14897. doi:10.1074/jbc.M110.107789

Lynch, J. W. (2004). Molecular Structure and Function of the glycine Receptor Chloride Channel. *Physiol. Rev.* 84 (4), 1051–1095. doi:10.1152/physrev.00042.2003

Lynch, J. W., Zhang, Y., Talwar, S., and Estrada-Mondragon, A. (2017). Glycine Receptor Drug Discovery. *Adv. Pharmacol.* 79, 225–253. doi:10.1016/bs.apha.2017.01.003

Madjroh, N., Mellou, E., Æbelø, L., Davies, P. A., Söderhielm, P. C., and Jensen, A. A. (2021). Probing the Molecular Basis for Signal Transduction through the Zinc-Activated Channel (ZAC). *Biochem. Pharmacol.* 193, 114781. doi:10.1016/j.bcp.2021.114781

Mancuso, C. E., Tanzi, M. G., and Gabay, M. (2004). Paradoxical Reactions to Benzodiazepines: Literature Review and Treatment Options. *Pharmacotherapy* 24 (9), 1177–1185. doi:10.1592/phco.24.13.1177.38089

Masiulis, S., Desai, R., Uchański, T., Serna Martin, I., Laverty, D., Karia, D., et al. (2019). GABAA Receptor Signalling Mechanisms Revealed by Structural Pharmacology. *Nature* 565 (7740), 454–459. doi:10.1038/s41586-018-0832-5

McKinney, W. (2010). Data Structures for Statistical Computing in Python. *Proc. Python Sci. Conference,Proceedings 9th Python Sci. Conf.* doi:10.25080/majora-92bf1922-00a

Menku, B., Altiparmak, T., Genis, B., and Cosar, B. (2021). Adverse Neuropsychiatric Events Associated with Varenicline: A Case Series. *Pbs* 11, 80–83. doi:10.5455/PBS.20201231094830

Miller, P. S., Scott, S., Masiulis, S., De Colibus, L., Pardon, E., Steyaert, J., et al. (2017). Structural Basis for GABAA Receptor Potentiation by Neurosteroids. *Nat. Struct. Mol. Biol.* 24 (11), 986–992. doi:10.1038/nsmb.3484

Miquel, M.-C., Emerit, M. B., Nosjean, A., Simon, A., Rumajogee, P., Brisorgueil, M.-J., et al. (2002). Differential Subcellular Localization of the 5-HT3-Asreceptor Subunit in the Rat central Nervous System. *Eur. J. Neurosci.* 15 (3), 449–457. doi:10.1046/j.0953-816x.2001.01872.x

Niesler, B., Walstab, J., Combrink, S., Möller, D., Kapeller, J., Rietdorf, J., et al. (2007). Characterization of the Novel Human Serotonin Receptor Subunits 5-Ht3c,5-Ht3d, and 5-HT3E. *Mol. Pharmacol.* 72 (1), 8–17. doi:10.1124/mol.106.032144

Noviello, C. M., Gharpure, A., Mukhtasimova, N., Cabuco, R., Baxter, L., Borek, D., et al. (2021). Structure and Gating Mechanism of the α7 Nicotinic Acetylcholine Receptor. *Cell* 184 (8), 2121–2134. e2113. doi:10.1016/j.cell.2021.02.049

Olsen, R. W., and Sieghart, W. (2008). International Union of Pharmacology. LXX. Subtypes of γ-Aminobutyric AcidAReceptors: Classification on the Basis of Subunit Composition, Pharmacology, and Function. Update. *Pharmacol. Rev.* 60 (3), 243–260. doi:10.1124/pr.108.00505

Osorio, D., Rondón-Villarreal, P., and Torres, R. (2015). Peptides: A Package for Data Mining of Antimicrobial Peptides. *R. J.* 7 (1), 4–14. doi:10.32614/rj-2015-001

Oz, M., Zhang, L., Ravindran, A., Morales, M., and Lupica, C. R. (2004). Differential Effects of Endogenous and Synthetic Cannabinoids on α7-Nicotinic Acetylcholine Receptor-Mediated Responses in Xenopus Oocytes. *J. Pharmacol. Exp. Ther.* 310 (3), 1152–1160. doi:10.1124/jpet.104.067751

Paton, C. (2002). Benzodiazepines and Disinhibition: a Review. *Psychiatr. Bull.* 26 (12), 460–462. doi:10.1192/pb.26.12.460

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2012). Scikit-learn: Machine Learning in Python. *J. Machine Learn. Res.* 12.

Pei, J., Kim, B.-H., and Grishin, N. V. (2008). PROMALS3D: a Tool for Multiple Protein Sequence and Structure Alignments. *Nucleic Acids Res.* 36 (7), 2295–2300. doi:10.1093/nar/gkn072

Polovinkin, L., Hassaine, G., Perot, J., Neumann, E., Jensen, A. A., Lefebvre, S. N., et al. (2018). Conformational Transitions of the Serotonin 5-HT3 Receptor. *Nature* 563 (7730), 275–279. doi:10.1038/s41586-018-0672-3

Puthenkalam, R., Hieckel, M., Simeone, X., Suwattanasophon, C., Feldbauer, R. V., Ecker, G. F., et al. (2016). Structural Studies of GABAA Receptor Binding Sites: Which Experimental Structure Tells Us what? *Front. Mol. Neurosci.* 9. doi:10.3389/fnmol.2016.00044

Rajkumar, R., and Mahesh, R. (2010). Review: The Auspicious Role of the 5-HT3 Receptor in Depression: a Probable Neuronal Target? *J. Psychopharmacol.* 24 (4), 455–469. doi:10.1177/0269881109348161

Ramerstorfer, J., Furtmuller, R., Sarto-Jackson, I., Varagic, Z., Sieghart, W., and Ernst, M. (2011). The GABAA Receptor + - Interface: A Novel Target for Subtype Selective Drugs. *J. Neurosci.* 31 (3), 870–877. doi:10.1523/JNEUROSCI.5012-10.2011

Rodriguez-Fontan, F., Reeves, B., Tuaño, K., Colakoglu, S., Laura, D. A., and D'Agostino, R. (2020). Tobacco Use and Neurogenesis: A Theoretical Review of Pathophysiological Mechanism Affecting the Outcome of Peripheral Nerve Regeneration. *J. Orthopaedics* 22, 59–63. doi:10.1016/j.jor.2020.03.026

Rosen, L. J., Galili, T., Kott, J., Goodman, M., and Freedman, L. S. (2018). Diminishing Benefit of Smoking Cessation Medications during the First Year: a Meta-Analysis of Randomized Controlled Trials. *Addiction* 113 (5), 805–816. doi:10.1111/add.14134

Rossokhin, A. V., Sharonova, I. N., Dvorzhak, A., Bukanova, J. V., and Skrebitsky, V. G. (2019). The Mechanisms of Potentiation and Inhibition of GABAA

Receptors by Non-steroidal Anti-inflammatory Drugs, Mefenamic and Niflumic Acids. *Neuropharmacology* 160, 107795. doi:10.1016/j.neuropharm.2019.107795

Rudolph, U., and Knoflach, F. (2011). Beyond Classical Benzodiazepines: Novel Therapeutic Potential of GABAA Receptor Subtypes. *Nat. Rev. Drug Discov.* 10 (9), 685–697. doi:10.1038/nrd3502

Salari, R., Murlidaran, S., and Brannigan, G. (2014). Pentameric Ligand-Gated Ion Channels: Insights from Computation. *Mol. simulation* 40 (10-11), 821–829. doi:10.1080/08927022.2014.896462

Sandberg, M., Eriksson, L., Jonsson, J., Sjöström, M., and Wold, S. (1998). New Chemical Descriptors Relevant for the Design of Biologically Active Peptides. A Multivariate Characterization of 87 Amino Acids. *J. Med. Chem.* 41 (14), 2481–2491. doi:10.1021/jm9700575

Schmiedhofer, P. B. (2017). Pharmacology of GABAA Receptor Subtypes : Phytocannabinoids Act on GABAA Receptors.

Schnoll, R. A., Patterson, F., Wileyto, E. P., Heitjan, D. F., Shields, A. E., Asch, D. A., et al. (2010). Effectiveness of Extended-Duration Transdermal Nicotine Therapy. *Ann. Intern. Med.* 152 (3), 144–151. doi:10.7326/0003-4819-152-3-201002020-00005

Scholze, P., and Huck, S. (2020). The α5 Nicotinic Acetylcholine Receptor Subunit Differentially Modulates α4β2* and α3β4* Receptors. *Front. Synaptic Neurosci.* 12 (54). doi:10.3389/fnsyn.2020.607959

Sieghart, W. (2015). Allosteric Modulation of GABAA Receptors via Multiple Drug-Binding Sites. *Adv. Pharmacol.* 72, 53–96. doi:10.1016/bs.apha.2014.10.002

Sieghart, W., and Savić, M. M. (2018). International Union of Basic and Clinical Pharmacology. CVI: GABAA Receptor Subtype- and Function-Selective Ligands: Key Issues in Translation to Humans. *Pharmacol. Rev.* 70 (4), 836–878. doi:10.1124/pr.117.014449

Sigel, E., Baur, R., Rácz, I., Marazzi, J., Smart, T. G., Zimmer, A., et al. (2011). The Major central Endocannabinoid Directly Acts at GABA A Receptors. *Proc. Natl. Acad. Sci. U.S.A.* 108 (44), 18150–18155. doi:10.1073/pnas.1113444108

Simon, J., Wakimoto, H., Fujita, N., Lalande, M., and Barnard, E. A. (2004). Analysis of the Set of GABAA Receptor Genes in the Human Genome. *J. Biol. Chem.* 279 (40), 41422–41435. doi:10.1074/jbc.M401354200

Skolnick, P. (2012). Anxioselective Anxiolytics: On a Quest for the Holy Grail. *Trends Pharmacol. Sci.* 33 (11), 611–620. doi:10.1016/j.tips.2012.08.003

Sugasawa, Y., Bracamontes, J. R., Krishnan, K., Covey, D. F., Reichert, D. E., Akk, G., et al. (2019). The Molecular Determinants of Neurosteroid Binding in the GABA(A) Receptor. *J. Steroid Biochem. Mol. Biol.* 192, 105383. doi:10.1016/j.jsbmb.2019.105383

team, T.p.d. (2020). *Pandas-Dev/Pandas: Pandas.* doi: doi:10.5281/zenodo.3509134

Theriot, J., Wermuth, H. R., and Ashurst, J. V. (2021). "Antiemetic Serotonin-5-HT3 Receptor Blockers," in *StatPearls* (Treasure Island, FL: StatPearls Publishing Copyright © 2021StatPearls Publishing LLC.).

Thompson, A. J., Lester, H. A., and Lummis, S. C. R. (2010). The Structural Basis of Function in Cys-Loop Receptors. *Quart. Rev. Biophys.* 43 (4), 449–499. doi:10.1017/S0033583510000168

Thompson, A. J., and Lummis, S. C. (2007). The 5-HT3receptor as a Therapeutic Target. *Expert Opin. Ther. Targets* 11 (4), 527–540. doi:10.1517/14728222.11.4.527

Thompson, A., and Lummis, S. (2013). Discriminating between 5-HT 3 A and 5-HT 3 AB Receptors. *Br. J. Pharmacol.* 169 (4), 736–747. doi:10.1111/bph.12166

Tønnesen, P., Paoletti, P., Gustavsson, G., Russell, M. a., Saracci, R., Gulsvik, A., et al.members of the Steering Committee of Cease on behalf of the European Respiratory Society (1999). Higher Dosage Nicotine Patches Increase One-year Smoking Cessation Rates: Results from the European CEASE trialEuropean Respiratory Society. *Eur. Respir. J.* 13 (2), 238–246. doi:10.1034/j.1399-3003.1999.13b04.x

Trattnig, S. M., Gasiorek, A., Deeb, T. Z., Ortiz, E. J. C., Moss, S. J., Jensen, A. A., et al. (2016). Copper and Protons Directly Activate the Zinc-Activated Channel. *Biochem. Pharmacol.* 103, 109–117. doi:10.1016/j.bcp.2016.02.004

Trattnig, S. M., Harpsøe, K., Thygesen, S. B., Rahr, L. M., Ahring, P. K., Balle, T., et al. (2012). Discovery of a Novel Allosteric Modulator of 5-HT3 Receptors. *J. Biol. Chem.* 287 (30), 25241–25254. doi:10.1074/jbc.m112.360370

Tregellas, J. R., and Wylie, K. P. (2019). Alpha7 Nicotinic Receptors as Therapeutic Targets in Schizophrenia. *Nicotine Tob. Res.* 21 (3), 349–356. doi:10.1093/ntr/nty034

Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., et al. (2021). Highly Accurate Protein Structure Prediction for the Human Proteome. *Nature* 596 (7873), 590–596. doi:10.1038/s41586-021-03828-1

Tzvetkov, M. V., Meineke, C., Oetjen, E., Hirsch-Ernst, K., and Brockmöller, J. (2007). Tissue-specific Alternative Promoters of the Serotonin Receptor Gene HTR3B in Human Brain and Intestine. *Gene* 386 (1-2), 52–62. doi:10.1016/j.gene.2006.08.002

Chemical Computing Group ULC (2022). *Scientific Vector Language (SVL) Source Code provided by Chemical Computing Group ULC. Chemical Computing Group ULC, 1010.* Montreal, QC, Canada, H3A 2R7. Sherbooke St. West, Suite #910.

Unwin, N., and Fujiyoshi, Y. (2012). Gating Movement of Acetylcholine Receptor Caught by Plunge-Freezing. *J. Mol. Biol.* 422 (5), 617–634. doi:10.1016/j.jmb.2012.07.010

Van Rossum, G., and Drake, F. L., Jr (2009). *Python 3 Reference Manual*. Scotts Valley, CA, United States: CreateSpace.

van Westen, G. J., Swier, R. F., Cortes-Ciriano, I., Wegner, J. K., Overington, J. P., Ijzerman, A. P., et al. (2013). Benchmarking of Protein Descriptor Sets in Proteochemometric Modeling (Part 2): Modeling Performance of 13 Amino Acid Descriptor Sets. *J. Cheminform* 5 (1), 42. doi:10.1186/1758-2946-5-42

Walsh, R. M., Jr., Roh, S.-H., Gharpure, A., Morales-Perez, C. L., Teng, J., and Hibbs, R. E. (2018). Structural Principles of Distinct Assemblies of the Human α4β2 Nicotinic Receptor. *Nature* 557 (7704), 261–265. doi:10.1038/s41586-018-0081-7

Walstab, J., Rappold, G., and Niesler, B. (2010). 5-HT3 Receptors: Role in Disease and Target of Drugs. *Pharmacol. Ther.* 128 (1), 146–169. doi:10.1016/j.pharmthera.2010.07.001

Waskom, M., Gelbart, M., Botvinnik, O., Ostblom, J., Hobson, P., Lukauskas, S., et al. (20212021). Mwaskom/Seaborn: V0. *Zenodo* 11, 2. doi:10.5281/zenodo.5205191

Webb, B., and Sali, A. (2016). Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinformatics* 54 (1), 1. doi:10.1002/cpbi.3

Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., et al. (2018). DrugBank 5.0: a Major Update to the DrugBank Database for 2018. *Nucleic Acids Res.* 46 (D1), D1074–d1082. doi:10.1093/nar/gkx1037

Wolber, G., and Langer, T. (2005). LigandScout: 3-D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters. *J. Chem. Inf. Model.* 45 (1), 160–169. doi:10.1021/ci049885e

Xiong, W., Cheng, K., Cui, T., Godlewski, G., Rice, K. C., Xu, Y., et al. (2011). Cannabinoid Potentiation of glycine Receptors Contributes to Cannabis-Induced Analgesia. *Nat. Chem. Biol.* 7 (5), 296–303. doi:10.1038/nchembio.552

Yang, K.-H., Galadari, S., Isaev, D., Petroianu, G., Shippenberg, T. S., and Oz, M. (2010). The Nonpsychoactive Cannabinoid Cannabidiol Inhibits 5-hydroxytryptamine3A Receptor-Mediated Currents in *Xenopus laevis* Oocytes. *J. Pharmacol. Exp. Ther.* 333 (2), 547–554. doi:10.1124/jpet.109.162594

Yang, Z., Aubrey, K. R., Alroy, I., Harvey, R. J., Vandenberg, R. J., and Lynch, J. W. (2008). Subunit-specific Modulation of glycine Receptors by Cannabinoids and N-Arachidonyl-glycine. *Biochem. Pharmacol.* 76 (8), 1014–1023. doi:10.1016/j.bcp.2008.07.037

Yang, Z., Cromer, B. A., Harvey, R. J., Parker, M. W., and Lynch, J. W. (2007). A Proposed Structural Basis for Picrotoxinin and Picrotin Binding in the glycine Receptor Pore. *J. Neurochem.* 103 (2), 580–589. doi:10.1111/j.1471-4159.2007.04850.x

Young, A. B., and Snyder, S. H. (1973). Strychnine Binding Associated with glycine Receptors of the central Nervous System. *Proc. Natl. Acad. Sci. U.S.A.* 70 (10), 2832–2836. doi:10.1073/pnas.70.10.2832

Young, J., Shytle, R. D., Sanberg, P. R., and George, T. P. (2001). Mecamylamine: New Therapeutic Uses and Toxicity/risk Profile. *Clin. Ther.* 23 (4), 532–565. doi:10.1016/s0149-2918(01)80059-x

Yu, H., Bai, X.-C., and Wang, W. (2021). Characterization of the Subunit Composition and Structure of Adult Human glycine Receptors. *Neuron* 109 (17), 2707–2716. e2706. doi:10.1016/j.neuron.2021.08.019

Zaliani, A., and Gancia, E. (1999). MS-WHIM Scores for Amino Acids: A New 3D-Description for Peptide QSAR and QSPR Studies. *J. Chem. Inf. Comput. Sci.* 39 (3), 525–533. doi:10.1021/ci980211b

Zhang, Y., Dijkman, P. M., Zou, R., Zandl-Lang, M., Sanchez, R. M., Eckhardt-Strelau, L., et al. (2021). Asymmetric Opening of the Homopentameric 5-HT3A Serotonin Receptor in Lipid Bilayers. *Nat. Commun.* 12 (1), 1074. doi:10.1038/s41467-021-21016-7

Zhao, Y., Liu, S., Zhou, Y., Zhang, M., Chen, H., Eric Xu, H., et al. (2021). Structural Basis of Human α7 Nicotinic Acetylcholine Receptor Activation. *Cell Res* 31 (6), 713–716. doi:10.1038/s41422-021-00509-6

Zhu, H., and Gouaux, E. (2021). Architecture and Assembly Mechanism of Native glycine Receptors. *Nature* 599, 513–517. doi:10.1038/s41586-021-04022-z

Zoli, M., Pucci, S., Vilella, A., and Gotti, C. (2018). Neuronal and Extraneuronal Nicotinic Acetylcholine Receptors. *Cn* 16 (4), 338–349. doi:10.2174/1570159x15666170912110450

Zuber, B., and Unwin, N. (2013). Structure and Superorganization of Acetylcholine Receptor-Rapsyn Complexes. *Proc. Natl. Acad. Sci. U.S.A.* 110 (26), 10622–10627. doi:10.1073/pnas.1301277110